National Technical University Of Athens
School Of Electrical And Computer Engineering
Division Of Electromagnetics, Electrooptics and
Electronic Materials

Monte Carlo simulations to estimate the damage potential of electron beam and tests of beam loss detector based on quartz Cherenkov radiator read out by a silicon photomultiplier on CLIC Test Facility 3(CTF3).

A Diploma Thesis

By

Orfanelli S. Styliani

CERN Supervisor :
Mariusz Gracjan Sapinski

NTUA Supervisor:
Evangelos N.Gazis

Athens, 25th of November 2011

NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING
DIVISION OF ELECTROMAGNETICS, ELECTROOPTICS AND ELECTRONIC MATERIALS

Monte Carlo simulations to estimate the damage potential of electron beam and tests of beam loss detector based on quartz Cherenkov radiator read out by a silicon photomultiplier on CLIC Test Facility 3(CTF3).

A Diploma Thesis

By

Orfanelli S. Styliani

The undersigned have examined the thesis presented the 25th of November 2011 and hereby certify that it is worthy of acceptance :

....
Evangelos Gazis
Professor

....
Dimitrios Tsamakis
Professor

....
Elias Glytsis
Professor

....
Orfanelli S. Styliani
Diploma Graduate in Electrical and Computer Engineering

In presenting this thesis in partial fulfilment of the requirements for a Diploma degree at the NTUA, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Styliani S. Orfanelli
November 25, 2011

I. ΕΠΙΤΟΜΗ

Ο Συμπαγής Γραμμικός Επιταχυντής (CLIC) είναι μία μελέτη σκοπιμότητας που στοχεύει στην ανάπτυξη ενός γραμμικού επιταχυντή ηλεκτρονίου/ποζιτρονίου με ενέργεια κέντρου μάζας σε ενέργειας της τάξης των TeV. Κάθε Γραμμικό Τμήμα θα έχει μήκος 21 km, το οποίο σημαίνει ότι απαιτούνται πολύ υψηλές βαθμίδες επιτάχυνσης (>100 MV/m). Για να επιτευχθούν οι υψηλές επιταχυντικές βαθμίδες, ένα καινοτόμο σχήμα επιτάχυνσης δυο δεσμών, στο οποίο RF ισχύς θα μεταφέρεται από μια δέσμη οδηγό υψηλού ρεύματος και χαμηλής ενέργειας προς μια κύρια δέσμη χαμηλού ρεύματος και υψηλής ενέργειας, σχεδιάζεται. Ένα σύστημα ανίχνευσης απωλειών δέσμης (BLM system) θα σχεδιαστεί για τον CLIC ώστε να ανταποκριθεί στις απαιτήσεις του επιταχυντικού συστήματος. Ο κύριος ρόλος του, ως μέρος του σχήματος για την προστασία του μηχανήματος, θα είναι να ανιχνεύει πιθανές επικίνδυνες αστάθειες της δέσμης και να προλαβαίνει επακόλουθη έγχυση στην κύρια δέσμη ή στους επιβραδυντές της δέσμης οδηγού. Το πρώτο μέρος της παρούσας εργασίας περιγράφει τις GEANT 4 Μόντε Κάρλο προσομοιώσεις που έγιναν για να υπολογιστεί το ενδεχόμενο καταστροφής από δέσμες ηλεκτρονίων υψηλής ενέργειας επιδρώντας σε ένα στόχο χαλκού. Το δεύτερο μέρος αφιερώθηκε στη μελέτη χαρακτηριστικών ενός ανιχνευτή απωλειών δέσμης που βασιζόταν σε ένα Cherenkov ακτινοβολητή από πυρητία ο οποίος διαβαζόταν από έναν Φωτοπολλαπλασιαστή πυριτίου (SiPM) ώστε να αποφασιστεί η καταλληλότητα του ως μια πιθανή τεχνολογική επιλογή για τον CLIC. Προκαταρκτικές μετρήσεις από την εγκατάσταση στον CLIC Test Facility (CTF3) περιγράφονται. Και τα δυο μέρη της εργασίας, προσομοιώσεις και πειράματα, πραγματοποιήθηκαν στο CERN ως μέρος της μελέτης CLIC BLM. Και τα δυο μέρη της εργασίας, προσομοιώσεις και πειράματα, πραγματοποιήθηκαν στο CERN ως μέρος της μελέτης του BLM συστήματος για τον CLIC.

I. ABSTRACT

The Compact Linear Collider (CLIC) study is a feasibility study aiming at the development of an electron/positron linear collider with a centre of mass energy in the multi-TeV energy range. Each Linac will have a length of 21 km, which means that very high accelerating gradients (>100 MV/m) are required. To achieve the high accelerating gradients, a novel two-beam acceleration scheme, in which RF power is transferred from a high-current, low-energy drive beam to the low-current, high energy main accelerating beam is designed. A Beam Loss Monitoring (BLM) system will be designed for CLIC to meet the requirements of the accelerator complex. Its main role as part of the machine protection scheme will be to detect potentially dangerous beam instabilities and prevent subsequent injection into the main beam or drive beam decelerators. The first part of this work describes the GEANT4 Monte Carlo simulations performed to estimate the damage potential of high energy electron beams impacting a copper target. The second part was dedicated to the characterization of a beam loss detector based on quartz Cerenkov radiator read out by Silicon Photomultipliers (SiPMs) to determine its suitability as a possible technology choice for CLIC. Preliminary measurements from the installation at the CLIC Test Facility (CTF3) are described. Both the simulations and the experiments were performed at CERN as part of the CLIC BLM study.

II. ACKNOWLEDGMENTS

This dissertation would not have been possible to write without the guidance and the help of several individuals who in one way or the other contributed in the preparation and completion of this study.

Firstly and foremost, it is difficult to overstate my gratitude to my Professor Evangelos Gazis, who since 2009 he has been continuously supporting me and encouraging me to take the leap and come at CERN. I would also like to thank my CERN supervisor Mariusz Sapinski for his trust and his unselfish and stead fast encouragement to complete this study. I am indebted to many of my colleagues at CERN but most of all Sophie Mallows for her untiring help; without her that thesis would never have been completed. More than half of the thesis and the experiments performed would not have taken place if Sophie wasn't there to support me not only scientifically but also psychologically.

Another colleague that I would like to thank is William Vigano who offered me the chance to be his apprentice and has shared with me the approach of a professional engineer when working in a lab. A big thank you for the rest of the staff of the BE/BI/BL group for welcoming me in their team and being so helpful despite their heavy load of work. Special thanks to my officemate Christoph Kurfuerst who has offered me his smile every morning and his support during my stay at CERN. I could not also forget to express my gratitude to Thibaut Lefevre for his support from the beginning of my work at CERN. I would also like to thank Mike Albrow who lent the Cherenkov detector to BL group and I was able to work with it.

Last but not least, I would like to thank my friends and my family for being always there for me and supporting my choices.

III. TABLE OF CONTENTS

IV. LIST OF FIGURES

V. LIST OF TABLES

V. MAIN BODY

CHAPTER 1. SYNOPSIS

The Compact Linear Collider (CLIC) study is a feasibility study aiming at the development of a realistic technology to extend electron/positron linear colliders into the multi-TeV energy range. CLIC will be a large facility with many subsystems. Each Linac will have a length of 21 km, which means that very high accelerating gradients (>100 MV/m) are required. To achieve the high accelerating or gradients, a novel two-beam acceleration scheme, in which energy is transferred from a high-current, low-energy drive beam to the low-current, high energy main accelerating beam is designed. A Beam Loss Monitoring system will be designed for CLIC to meet the new requirements of the accelerator. Its task will be to prevent damage to accelerator components from beam losses from either of the beams.

Radiation damage and other radiation induced effects have become a major concern for the operation of accelerators and detectors fro High Energy Physics experiments. The need to understand and master radiation induced effects will be of great importance for CLIC as they very well might represent limiting factors for the feasibility of an experiment. The design of protection elements for HEP experiments is based on assumption on dmage levels, which are derived from simulations. For the first part of the thesis, simulations were performed with Geant4 toolkit in order to estimate potential radiation effect problem to a copper cylindrical solid target irradiated by high energy electron beam. In combination with dedicated experiments at adequate radiation eptest facilities, the simulation results are useful to obtain optimum design parameters for accelerating structures such as various magnetic elements.

The second part of the thesis is dedicated to the characterization of a silicon photomultiplier and the testing of a Cherenkov radiator read out by the SiPM as a possible technology choice as a Beam Loss Monitor for CLIC. The characterization experiments took place at BE-BI-BL group laboratory and the BLM testing in the CLEX area of CLIC test facilty 3 (CTF3). Due to limited dynamic range of the SiPM and the direct irradiation of the BLM , the signal was saturated .

The present introductory chapter is followed by another six chapters:

- Chapter 2 is an introduction to CERN, LHC and Compact Linear Collider (CLIC) study, as well as explaining the need of a high energy future linear collider like CLIC and its novel two beam accelerating scheme. CTF3, where the final test of the performance of the Cherenkov detector took place, is also described.

- Chapter 3 gives an overview over beam loss monitoring systems and explains their necessity as part of an accelerating structure machine protection. Based on different types of the radiation detection methods short description of the most frequently used beam loss monitors is given, including Cherenkov radiators.

- Chapter 4 is dedicated to the simulations of beam losses performed with Geant4. Firstly, an introduction in Monte Carlo method and the modeling of particle-matter interactions is given. Afterwards, the basic concepts of Geant4 toolkit are described, followed by a detailed description of the method that was followed for the simulations. At the end of the chapter the obtained results of the simulations performed are presented.

- Chapter 5 is an introductory chapter in photodetection theory. The history of photodetectors and their evolution until the Silicon photomultipliers (SiPMs) and their characteristics are summarized.

- In Chapter 6 the setups, the steps and the results from the characterization lab tests of the SiPM that was used to read out the light produced by the Cherenkov radiator are presented.

- In Chapter 7 the conclusions and future thoughts about this study are stated.

CHAPTER 2. CERN, CLIC & CTF3

2.0 Introduction

The European Organization for Nuclear Research, known as CERN is one of the world's largest centers for scientific research. Founded in 1954, the CERN Laboratory sits astride the Franco–Swiss border near Geneva and now has 20 Member States.

Its main function is fundamental physics by finding out what the Universe is made of. The instruments used are particle accelerators and detectors. Accelerators boost beams of particles to high energies before they are made to collide with each other or with stationary targets. Detectors observe and record the results of these collisions. Numerous experiments have been constructed at CERN by international collaborations which aim to study and reveal phenomena in modern particle physics.

2.1 CERN's experimental program. The LHC.

The latest and foremost accelerator at CERN is the Large Hadron Collider [1]- the world's largest particle accelerator- which lies in a tunnel as deep as 175 meters. The precise circumference of the LHC accelerator is 26 659 m, with a total of 9300 magnets inside. This synchrotron is designed to collide opposing particle beams of either protons at an energy of 7 TeV per particle, or lead nuclei at an energy of 574 TeV per nucleus. The LHC will continue to operate at half energy until the end of 2012; it will not run at full energy (7 TeV per beam) until 2014.

LHC is set to provide a rich programme of physics at a new high-energy frontier over the coming years. LHC is expected to run for more than 10 years. Six detectors have been constructed at the LHC, located underground in large caverns excavated at the LHC's intersection points. Two of them, the ATLAS experiment and the Compact Muon Solenoid (CMS), are large, general purpose particle detectors. A Large Ion Collider Experiment (ALICE) and LHCb, have more specific roles and the last two, TOTEM and LHCf, are very much smaller and are for very specialized research.

Scientists anticipate for the LHC results as it should above all confirm or refute the existence of the Higgs boson to complete the Standard Model[2] (see figure 1), explaining how particles get mass through the so-called Higgs field. The LHC experiments will also explore the possibilities beyond the Standard Model such as supersymmetry, extra dimensions and new gauge bosons. Nevertheless, particle

---

[1] *The term hadron refers to particles composed of quarks.*

[2] *Because of its success in explaining a wide variety of experimental results, the Standard Model is sometimes regarded as a theory of almost everything.*

physics community worldwide has reached a consensus that the results from the LHC will need to be complemented by experiments at an electron-positron collider operating in the tera-electron-volt (TeV) energy range.

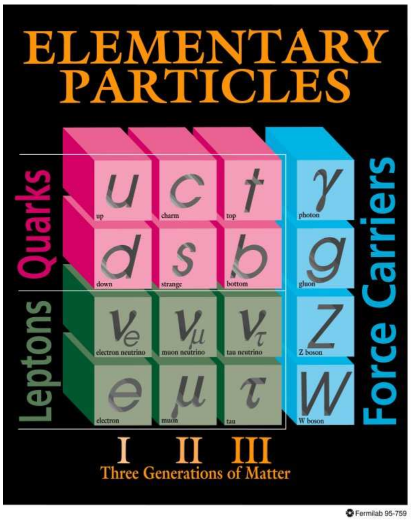


Figure 1: The Standard Model of elementary particles, with the gauge bosons in the rightmost column. The Standard model describes all the fundamental forces of the nature: electromagnetic, nuclear and weak forces; except for the gravitation force and its predicted carrier, graviton.

2.2 Future Linear Colliders

Typically hadron colliders are used as ‘discovery facilities’ and lepton colliders for precision physics. Discoveries at the LHC will set the direction for possible future high-energy colliders. The general consensus among particle physicists is that results from the LHC will need to be complemented by experiments at an electron–positron collider operating in the tera-electron-volt energy range. Until present, the highest centre-of-mass energy achieved at an electron-positron collider was 209 GeV at the LEP (Large Electron-Positron collider).. Due to the higher energies required by the next generation of lepton colliders, circular storage rings such as the LEP, are no longer an option due to the energy of the beam that is lost in the form of synchrotron radiation. (Particles travelling in a circular orbit emit synchrotron radiation proportional to the fourth power their mass). Linear colliders are therefore the only option for realizing electron–positron collisions at tera-electron-volt energies.

The concept of a lepton collider is simple: It consists of two linacs, one accelerating electrons ($e-$), the other positrons ($e+$) in opposite directions so that the particles collide head on. The scheme has inherent features that strongly influence the design: A linac must accelerate the particles in one single pass, therefore high electric fields for acceleration gradients are required to keep the length of the collider within reasonable limits. Furthermore, as the two beams collide only once, a linear collider in contrast to a circular collider would have a lower repetition frequency, typically 5–100 Hz. Therefore to maintain sufficient luminosity, very small beam dimensions at the interaction point and high bunch charges are required.

2.3 The Compact Linear Collider (CLIC) study

The Compact Linear Collider (CLIC) study is a feasibility study aiming at the development of a realistic technology to extend electron/positron linear colliders into the multi-TeV energy range. It would provide significant fundamental physics in post LHC-era. Currently 32 institutes participate from 17 countries (Europe, Australia, USA). The CLIC study aims at a center-of-mass energy range for electron/positron collision of 0.5 TeV-5 TeV, optimised for a nominal center-of-mass energy of 3TeV. The luminosity is expected to be approximately $6 \cdot 10^{34} cm^{-2} s^{-1}$. However, both of the E_{cm} and L values will be revised, when LHC physics results avail.

Each Linac will have a length of 21 km, which means that very high accelerating gradients (>100 MV/m) are required. To achieve the high accelerating or gradients, a novel two-beam acceleration scheme, in which energy is transferred from a high-current, low-energy drive beam to the low-current, high energy main accelerating beam is designed. The CLIC scheme is based on normal conducting travelling-wave accelerating structures, operating at a frequency of 12GHz and high electric fields.

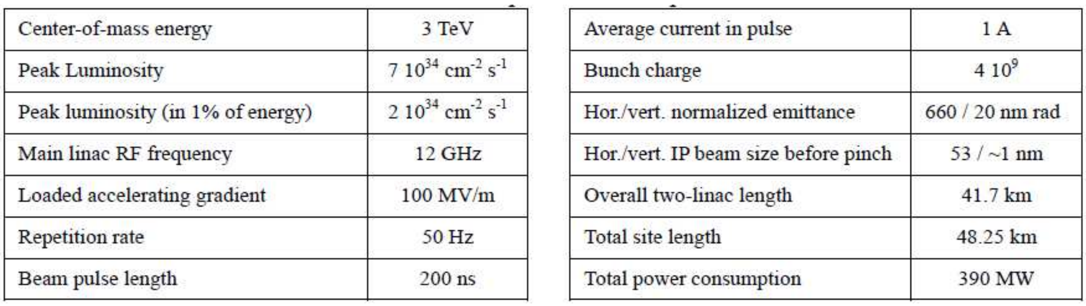
To transfer the energy to the main beam, the drive beam passes through novel Power Extraction and Transfer Structures (PETS), where it excites strong electromagnetic oscillations. Thus, as the beam is decelerated, the RF energy is extracted from the PETS and sent via waveguides to the accelerating structures in the parallel main beam. The PETS are travelling wave structures like the accelerating structures for the main beam, but with different parameters.

The two-beam scheme was the only solution in producing and providing the high peak RF power requested for the electric fields of 100MV/m -that is 275MW per active meter of accelerating structure. This value would be impossible to be achieved with the use of individual RF sources such as klystrons for a length of almost 50km. The absence of klystrons gives at the CLIC study the advantage of a simple tunnel layout (see figure 2). Both beams can be generated in a central injector complex and afterwards transported along the linac.

The proposed CLIC schematic layout is presented in the figure 3. In the center there are the two main beam linacs to boost electrons and positrons. The particle detectors will be installed in the interaction point (IP). There, the collisions will take place. After the beams will be delivered by two sophisticated beam delivery systems (BDS) and focused to dimensions of 1nm rms size in the vertical plane and 40nm horizontally, in order to achieve the required luminosity.

In parallel to each main linac, there are the twenty-four decelerator lines, consisting of PETS and focussing quadrupoles. Figure 2 shows the two-fold drive beam generation complex consisting of two drive beam linacs, followed by a delay loop and two combiner rings (CR) used to create the high current beam. At injection into the decelerators, the drive beam current is 100 A, the energy 2.38 GeV, and bunch spacing by 2.5 cm (12 GHz) and train duration 239 ns.

The main beam generation complex consists of the electron and positron injector linacs, the damping rings used to reduce the beam emmitance and achieve the required luminosities, and the booster linac which accelerates the electrons and positrons from 2.4GeV to 9GeV before their subsequent injection into the main linacs.



Center-of-mass energy	3 TeV
Peak Luminosity	7 10^{34} cm^{-2} s^{-1}
Peak luminosity (in 1% of energy)	2 10^{34} cm^{-2} s^{-1}
Main linac RF frequency	12 GHz
Loaded accelerating gradient	100 MV/m
Repetition rate	50 Hz
Beam pulse length	200 ns

Average current in pulse	1 A
Bunch charge	4 10^9
Hor./vert. normalized emittance	660 / 20 nm rad
Hor./vert. IP beam size before pinch	53 / ~1 nm
Overall two-linac length	41.7 km
Total site length	48.25 km
Total power consumption	390 MW

Table 1: CLIC parameters

	Energy range	Rep rate	Pulse length	Bunch frequency	Bunch charge	Bunches per train	**Electrons per train**
Drive Beam	**2.4 -> 0.24 GeV**	50 Hz	239ns	12 GHz	8.4nC	2922	**1.53e14**
Main Beam	**9 - > 1500 GeV**	50 Hz	156ns	12GHz	0.6nC	312	**1.16e12**

Table 2: Main Parameters of CLIC Beams

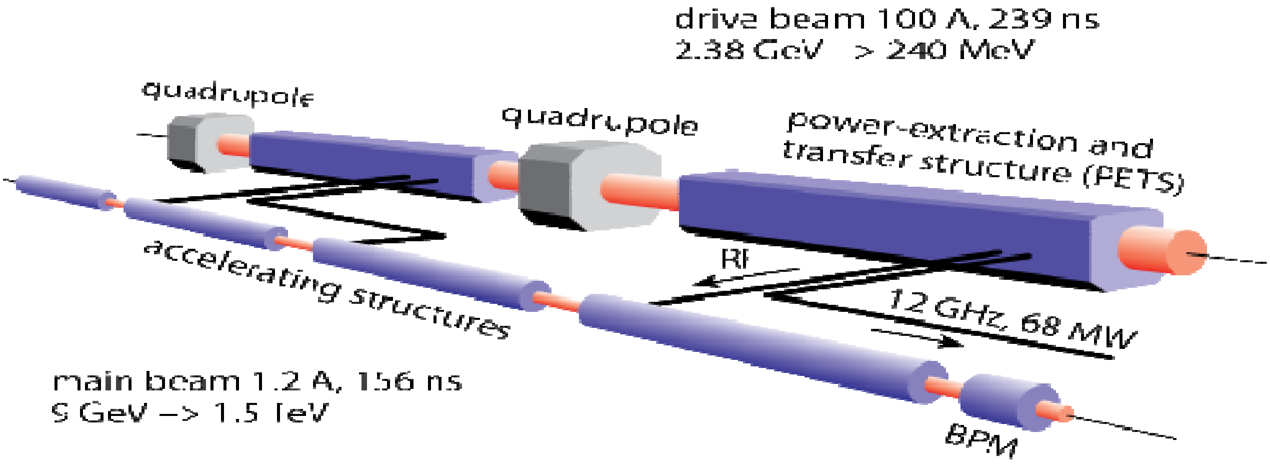


Figure 2: Two –beam Acceleration method

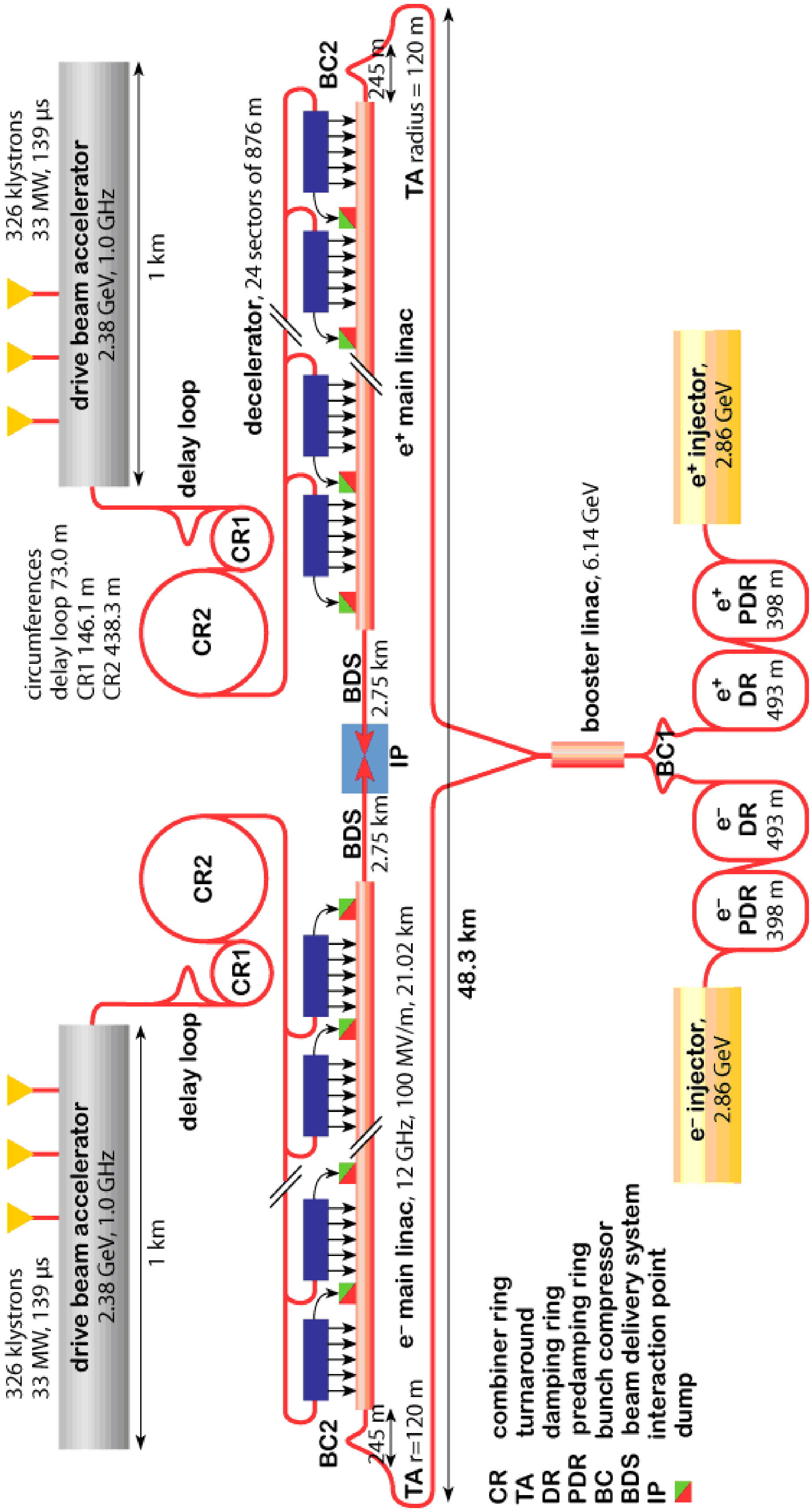


Figure 3: CLIC Layout

2.4 CLIC Test Facility 3 (CTF3)

It is foreseen that CLIC would be built in stages, starting at the lowest energy required by the physics, with successive energy upgrades that can potentially reach about five times the energy of the ILC*. Many key aspects of the CLIC scheme have been experimentally validated already in different test facilities (CTF, CTF2 and CTF3).

The CLIC Test facility 3 (CTF3) was built at CERN is in the former LPI complex to address the key technological challenges of the two-beam acceleration scheme. The CTF3 main goals are to:

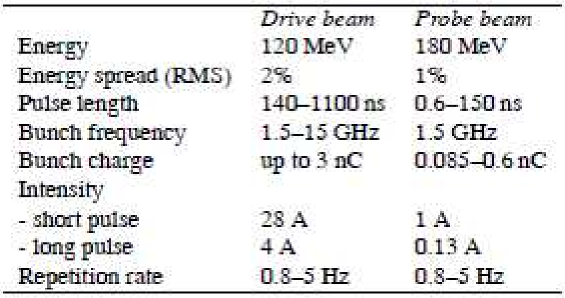
- Study and demonstrate the technical feasibility of drive beam generation with appropriate time structure and fully loaded acceleration
- Demonstrate the stability of a drive beam decelerator -including losses and the produced RF power by the PETS.
- Test the two-beam acceleration scheme with prototype CLIC accelerating structures.
- Benchmark the simulation tools in order to validate the corresponding systems in the CLIC nominal scheme.

The CTF3 facility consists of a drive beam linac, where the electrons are generated with a thermionic gun and accelerated to approximately 150MeV in a linac,then transferred to a 42m long Delay loop (DL) and a combiner ring (CR) which are used to create a 12 GHz bunch repetition frequency. The beam current is first doubled in the delay loop and then multiplied again by a factor of four in the combiner ring by interleaving bunches using transverse deflecting RF cavities.

In the CLIC Experimental area (CLEX), a 150 MeV “probe beam”, used for the two beam tests, is generated by a laser based electron gun or “photoinjector” (CALIFES). The high gradient acceleration is demonstrated at the Two Beam Test stand (TBTS), where the drive beam can be transported to lose its energy and produce RF power in PETS (Power Extraction and Transfer Structures). A transfer of RF power of 150MV/m to the probe beam was achieved in 2010.

Table 3 lists the main parameters of the drive and probe beams as delivered to the two-beam acceleration experiments, and decelerator test line (TBL).which run parallel to the TBTS in the CLEX hall. A photoinjector (PHIN) (alternative electron source for the Drive Beam) and 12 GHz klystron powered test stand complete the CTF3 facility.

**The International Linear Collider (ILC) is a proposed linear particle accelerator. It is planned to have a collision energy of 500 GeV initially, and, if approved after the project has published its Technical Design Report, planned for 2012, could be completed in the late 2010s.[1] A later upgrade to 1000 GeV (1 TeV) is possible. (from wiki)*



	Drive beam	*Probe beam*
Energy	120 MeV	180 MeV
Energy spread (RMS)	2%	1%
Pulse length	140–1100 ns	0.6–150 ns
Bunch frequency	1.5–15 GHz	1.5 GHz
Bunch charge	up to 3 nC	0.085–0.6 nC
Intensity		
- short pulse	28 A	1 A
- long pulse	4 A	0.13 A
Repetition rate	0.8–5 Hz	0.8–5 Hz

Table 3: Main parameters of the CTF3 Beams

	CTF3	CLIC
Beam Energy (GeV)	0.15	2.4
RF Frequency (GHz)	3	1
Multiplication Factor	8	24
Initial Beam Current (A)	3.75	4.2
Final Beam Current (A)	30	100
Initial Pulse length (us)	**1.2**	**140**
Final Pulse Length (ns)	140	240
Total Beam Energy (kJ)	**0.7**	**1400**
Repetition Rate (Hz)	5	50
Average Beam Power (MW)	**0.0034**	**70**
Charge density (nC/cm^2)	**$0.4\ 10^{6}$**	**$2.3\ 10^{10}$**

Table 4: The main differences between CLIC and CTF3 parameters.

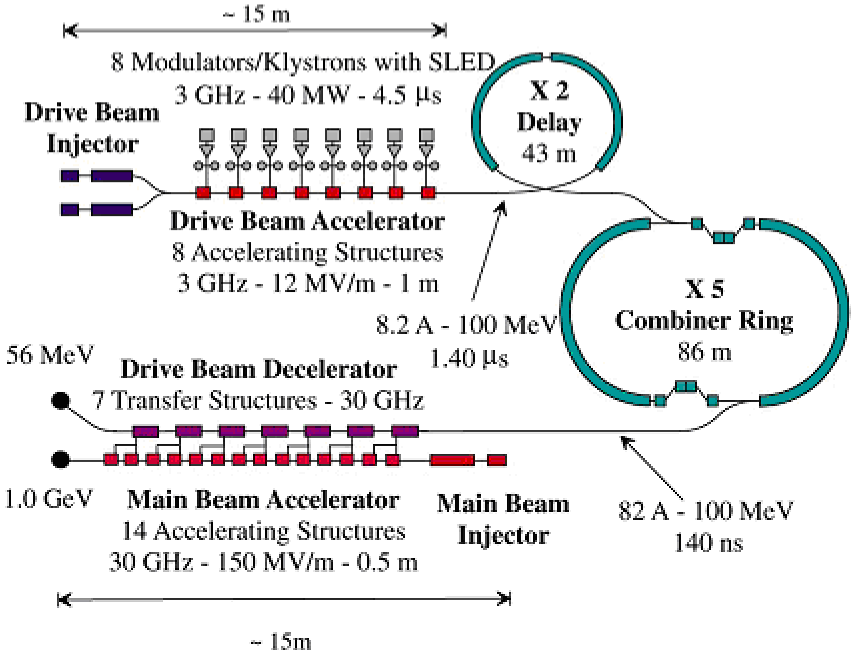


Figure 4: Test of Drive Beam Generation, Acceleration & RF Multiplication.

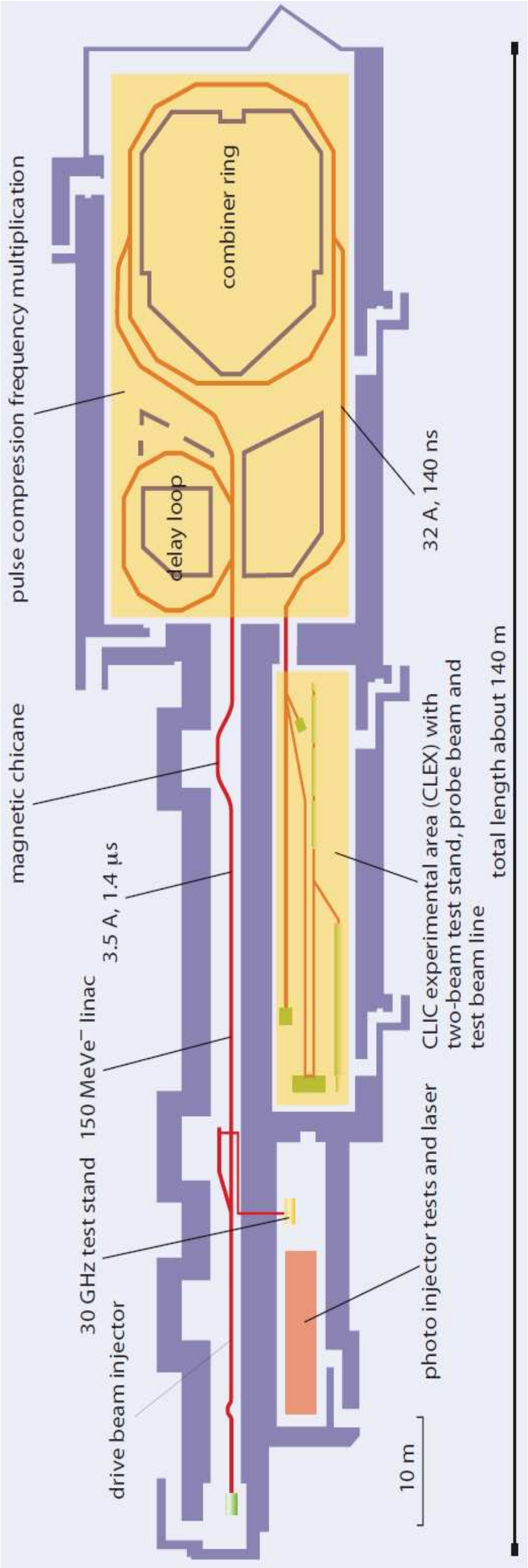


Figure 5: CTF3 Layout

CHAPTER 3. BEAM LOSS MONITORING SYSTEM

3.0 Introduction

The beam loss monitoring (BLM) system of a particle accelerator facility is one of the most widely distributed beam diagnostic systems. A proper understanding of the beam loss mechanism and a suitable design of a BLM system in combination with an appropriate location of the monitors can improve machine performance. This chapter reviews the tasks of this multi-faceted beam instrumentation tool in machine protection and loss location detection, the principles of beam loss detection as well as the properties of the devices used as "beam loss monitors" (BLMs).

3.1 The use of a BLM system

Beam loss monitors are sensitive tools designed for measuring beam losses around an accelerator or storage ring. A system based on an appropriate location of beam loss monitors and with fast response not only contributes to a fast machine protection but also enables a wide field of very useful beam diagnostics.

A BLM system with a good selectivity to the different loss mechanisms and with an adequate positioning of the BLMs can be a useful tool for machine diagnostic measurements such as tune scans, coupling studies and controlled beam blow-up as well as background measurements in detectors. A BLM system is used to optimize the beam lifetime and an efficient beam transport, by monitoring regular losses. In combination with the help in tuning the machine, it contributes to achieving high integrated luminosity for the particle accelerators.

Furthermore, BLM systems play an important role in localizing and studying any kind of loss at aperture restrictions and semitransparent obstacles in the accelerator, and in optimizing the injection, ejection or collimator scheme. Even a precise energy calibration can be performed with signals from a BLM system.

In addition to being a beam tuning machine, beam loss monitors, as front-line devices, provide an important role in machine protection. Actually, they are used to identify the position and consequently minimise any irregular, uncontrolled losses, protect the beam line components from radiation damage by signalling a beam abort system to fire if a certain loss rate is exceeded. In the case of superconducting accelerators, the fast abortion of a dangerous beam is of vital importance to protect the superconducting components from any possible induced quenches which may result in a shut-down of an accelerator.

It is crucial also to keep the radiation levels low in order to minimize the long-term activation for hands-on maintenance, personal safety and environmental protection of the accelerator by providing alarms when radiation from beam losses leads to excessive radiation levels.

**hands on maintenance: no more than 1mSv/hour residual activation (30cm from surface, after 4h cool-down)*

3.2 Types of losses

There are many different loss mechanisms. A beam loss can be intentional or not intentional. The unintentional beam losses can be either regular but also can become irregular and make damage. Generally, beam losses can be divided into two different classes:

- Fast / irregular / uncontrolled losses
- Slow / regular / controlled losses.

Fast losses, also referred to as irregular or uncontrolled losses, correspond to the total beam loss during one or very few turns. This type of loss is mainly distributed around the accelerator and not obviously on a collimation system. They are very often a result of a misaligned beam or a fault condition, e.g. operation failure or of a magnet power supply as well as errors in machine operation. This misaligned beam is able to destroy the beam, pipes or collimators and may break the vacuum.

However, sometimes such losses have to be tolerated even at a high level at low repetition rates during machine studies. A BLM system should define the allowed level of those losses and warn if an intolerable dose event occurs. A post mortem event analysis of this event helps to understand and analyze the fault condition.

Slow losses, which are also referred to as regular or controlled losses, correspond to a partial beam loss over a time (in circular machines) or distance (in linear machines) interval. Slow losses are usually due to many known reasons and they might occur continuously during operational running. They are mainly localized on the collimator system or other aperture limits and they are typically not avoidable. Some reasons for the regular losses in an accelerator structure are the residual gas scattering, not ideal beam alignment, several instabilities during operation and halo scraping (by beam pipe not dedicated scrapers).

As already mentioned, beam lifetime is defined by slow losses in storage-rings; vice versa the theoretical limitation of beam lifetime defines the lowest possible loss rate due to various effects such as residual gas, Touschek effect etc. Obviously, it is a great advantage to design a BLM system which is able to distinguish and deal with both loss modes.

3.3 Detection of a beam loss

3.3.1 Principles of loss detection

A particle is considered lost if it doesn’t follow its ideal trajectory and it interacts with matter. In case of a beam loss, a BLM system should be able to establish the number of lost particles, the loss location and the time interval of the event. A typical BLM is mounted outside the vacuum chamber, so that the monitor normally observes the shower caused by the lost particles interacting with the vacuum chamber walls.

One of the basic principles of loss detection is the proportionality between the signal from the BLM, the number of detected particles and the number of lost particles. This proportionality depends on the position of the BLM in respect to the beam the intervening material and on the properties of the lost particles (type, momentum). Together with the specification for acceptable beam losses as a function of beam momentum, this defines a minimum required sensitivity and dynamic range for BLMs.

3.3.2 Signal source of BLMs: ionizing radiation

The signal source of beam loss monitors is mainly the ionizing ability of the charged particles shower. There are different types of ionizing radiation produced from high-energy particle beam s depending on the type of the primary- lost particle and its energy. The produced radiation consists mainly of electromagnetic particles (electron-, positron- and gamma- shower), while the loss of a hadron (proton, ion) produces some hadronic particles (protons, neutrons), too.

Specifically, lost protons (for energies more than a few GeV) produce secondaries via hadronic showers, which include pions, neutrons and muons. Pi-zeros produce high energy gammas. Lost electrons produce electromagnetic showers via bremsstrahlung and gamma rays produced in electromagnetic showers convert back to electrons via Compton scattering and pair production. (more details for the particle –matter interactions in next chapter).

The primary mechanism by which a beam loss monitor detects a beam loss is from the transfer of energy from the incoming charged particle to the atomic electrons. This energy can be used to create electron/ion pairs or photons in the BLM detector material. The energy loss of incident charged particles scattering on atomic electrons is described by Bethe-Bloch formula (see chapter 4 for more details).

In accelerator structure it should also always be taken into account the existence of not beam related radiation, such as synchrotron radiation, cavity x-rays and background activation (which is beam related but not instantaneous).

3.3.3 Effects of Ionizing radiation – Hazards

The effects of ionizing radiation and the hazards are categorized below:

- Direct thermomechanical damage
    - material damage from overheating
    - thermal stress
        - heat load on vacuum chambers and components
- Indirect damage by showers and radiation field:
    - Radiation damage
    - Optics darkening (optical transmission)
    - Solid-state electronics single event upset

    - long-term damage (dislocations)
- Radio-activation of accelerator parts
    - personnel hazard for exposure specially to neutrons.
- Cryogenic systems excessive heat load
    - Quenching of super-conducting components
- Prompt radiation backgrounds in experiments

3.4 Beam loss monitors

When designing a beam loss monitor system, there are two crucial aspects which are analyzed below:

1. Type of loss monitor
2. Positioning of the loss monitor.

3.4.1 Criteria for the type of the beam loss monitor

Typical beam loss monitors detect beam losses by measurement of ionizing radiation produced by lost beam in real-time and with a certain position resolution. In addition, an ideal beam loss monitor should have a high dynamic range to be used for both regular (low) loss and irregular (high and fast) loss. It has to be also radiation hard enough in order to be suitable for an accelerator environment. It should be able to determine the amount of the beam lost by converting from radiation intensity to actual number of lost particles and also resolve the time structure of the loss and its spatial distribution.

There are many factors that must be considered in selecting a beam loss monitor design. There are some related to the type of radiation, some to the expected dose rates and some to other factors such as reliability, physical space and other. Systems based on differential beam current measurements have a very rough position. Dose measurements (or activation) have a very long time constant but they are not detailed described here.

The signal source of beam loss monitors mainly results from the ionizing capability of the charged shower particles. A list of the parameters that should be taken into account when selecting a beam loss monitor for a particular beam loss application is given below:

- Sensitivity (Coulombs or pulses per rad)
- Type of output (current-integrating or pulse type)
- Dynamic range (rads/sec and instantaneous rad doses)
- Ease of calibration and testability (online)
- System end-to-end online tests
- Uniformity of calibration (unit to unit)
- Calibration drift due to aging, radiation damage, outgassing, etc.
- Radiation hardness (material)
- Robustness (suitability for use in an accelerator enclosure environment)
- Cost (incl. Electronics)

- Sensitivity to backgrounds (RF cavity X-rays, synchontron radiation), to magnetic fields and to high voltage drift (photomultipliers)
- Physical size
- Spatial uniformity of coverage (e.g. in long tunnel, directionality)
- Bandwidth (temporal resolution)
- Response to low duty cycle (pulsed) radiation
- Instantaneous dynamic range (vs. switched gain dynamic range)
- Response to excessively high radiation level (graceful degradation or saturation)
- RAMI analysis (reliability, availability, maintainability, inspectability)

The list above gives a good guide for selecting the proper beam loss monitor when designing a BLM system.

3.4.2 Location of the beam loss monitor

The loss of a high-energy particle in the wall of a beam pipe results in a shower of particles. The closer to the beam loss location the BLM is installed, the higher the signal will be. Thus, the second crucial aspect that should be taken into account when designing a BLM system is the adequate positioning of the BLMs.

Usually, the optimum locations for the BLMs are based on estimates of the secondary particle shower using Monte Carlo simulations. The length of the shower depends strongly on the energy and the type of the lost particle. For high proton energies the length of the shower could be some meters, when for medium energy electrons only a few centimeters. Lost low energy beam particles do not even create a shower leakage outside the vacuum pipe wall.

Therefore, the expected location of lost particles should be studied in advance for the right positiong of the BLMs, especially in electron accelerators. This requirement is interpreted in understanding the loss mechanism and dynamics in the accelerator topology in order to predict the typical positions of losses. Thus, the different mechanisms of losses should be carefully studied when deciding where the more appropriate locations for the BLMs are.

3.5 Beam Loss Monitor types & Radiation detection methods

There are many different types of BLMs depending on the radiation detection method for detecting ionizing radiation. The most common method is to observe the interaction of charged particles with the atomic electrons in the detector by measuring:

- The Ionization charge (Ionization chambers, Solid State Ion Chamber).
- The Secondary emission current (SEM, Compton diode).
- The Fluorescence (Scintillators, Scintillating fibers).

- The Cherenkov light (Cherenkov radiators, Optical fibers).

3.6.1 Short Ionization Chamber

An ionization chamber consists of a sealed chamber and two parallel metallic electrodes (anode and cathode) separated by a distance D. The voltage between the electrodes is maintained by an external circuit. The sealed chamber, of width D, is filled with gas (air, argon, xenon, helium) or liquid of density ρ and defines the sensitive volume of the chamber.

When ionizing radiation enters the chamber through a foil-covered window, it ionizes one or more gas or liquid molecules by producing electron-ions pairs. The ions are attracted to the cathode and the electrons to the anode, due to the applied electric field E=V/D. Their presence causes a momentary drop in the voltage, which is recorded by the external circuit. The observed voltage drop helps identify the radiation because it depends on the degree of ionization, which in turn depends on the charge, mass, and speed of the photon.

The number of electrons reaching the anode depends on the applied voltage. If the voltage is very small, the electron produced by ionization recombines with its parent ion. If the electric field is larger than the Coulomb field in the vicinity of the parent ion, then electrons can escape this initial recombination. The number increases with voltage and the number of electrons collected at the anode increases with voltage up to saturation where all charges are collected. The region is called the Ionization Region.

Short ionization chambers are the most common BLM now in use. Whether a simple air filled chamber is adequate, or an Argon or Helium filled chamber, with superior higher dynamic range, must be used, depends on the conditions of the particular accelerator. They are more or less homogeneously distributed along the accelerator with additional units at special positions (e.g. aperture limits, collimators etc).

Unlike pulse-counting BLMs, current integrating ion chambers have a very high instantaneous dose limit. However, they correspond to synchrotron radiation. Another advantage of ionization chambers is that their calibration is determined by geometry and it is not strongly dependent on the applied voltage.

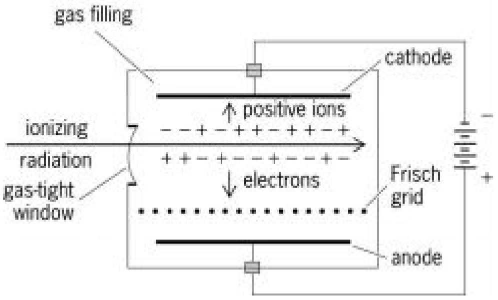


Figure 6: Parallel-plate ionization chamber filled with gas

3.6.2 Long Ionization-Chambers

In 1963, Panofsky proposed for a SLAC BLM system the later-called Panofsky Long Ionization_Chambers (PLICs) which was consisted of one long (3.5km) hollow coaxial cable - an industrial RG-319/U cable with a diameter of 4.1cm, filled with Ar/CO_2 gas. The outer conductor was grounded, the inner conductor was +HV and the output signal was ac-coupled. Because the electron beam was very short (<2μs), the up-beam PLIC signal (pulses traveling in the opposite direction to the beam) could be used to determine the loss point to a few meters.

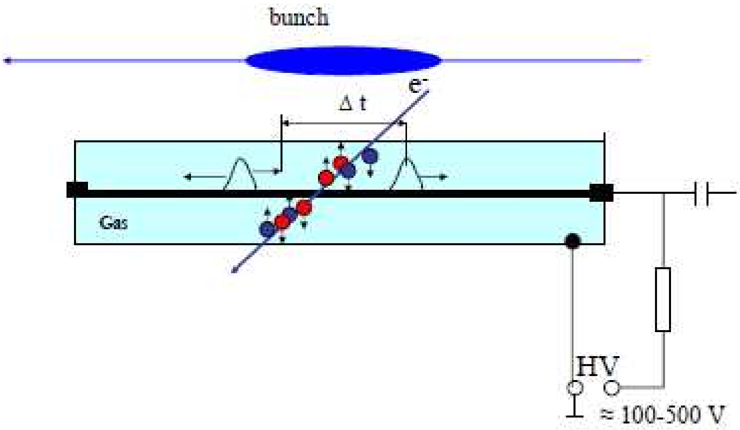


Figure 7: Panofski's Long-Ionization Chambers

Long ionization chambers using a single coaxial cable work well for one-shot accelerators or transport lines. There are currently, many variations of the original design installed at SLAC. They are very fast and can determine loss points by time-of-flight with roughly 1-meter resolution. To achieve spatial resolution of losses along an entire accelerator, two conditions must be fulfilled: The machine must be much longer than the bunch train, and the particles must be relativistic.
For particles travelling significantly slower than the signal in the cable (≈0.92c) and for circular machines it is necessary to split the cable. Each segment is read out separately, with spatial resolution approximately equal to the length of the unit .

3.6.3 PIN diode Pulse-Mode Coïncidence Circuit BLM

Circular electron accelerators suffer from hard synchotron – radiation emission (SR). The high irradiated environment makes practically impossible for the traditional BLMs to distinguish the radiation induced by beam loss and the synchotron radiation. Such an accelerator is the HERA which is an accelerator with an electron and a proton ring in the same tunnel, operating at the same time. The protection of the proton magnets from beam loss induced quenches should rely on a BLM system which sees only the proton beam losses and not the SR-background.

The hadronic shower created by beam losses includes a large number of charged particles, in contrast to the photons of the SR. The HERA BLM system consists of a combination of two PIN-diodes placed back-to-back and two pulse-output signals put into a coincidence circuit. Thus low energy photons interact in only one PIN diode, while charged particles crossing through the diodes give a coincidence signal. A limitation is the inability to distinguish overlapping counts, so that the count rate is proportional to the loss rate only for losses for which there is significantly less than one count per coincidence interval.

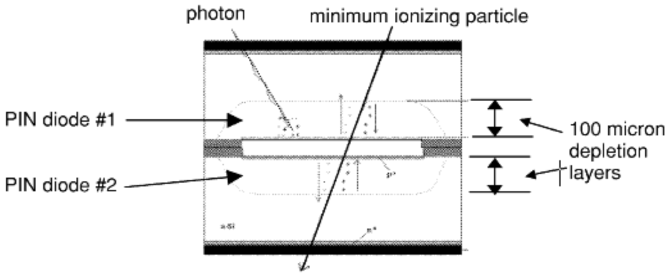


Figure 8: PIN diode coincidence circuit

3.6.4 Secondary Emission Monitors (ACEM)

An ACEM (aluminum cathode electron multiplier) is another type of BLM monitor which is useful at very high radiation areas. It is based on the concept of the photomultipliers and their enhanced sensitivity to ionized radiation. The difference is that the photocathode of the photomultipliers is replaced by an aluminum foil. This foil works as a secondary electron emitter when irradiated. This BLM is very fast and with an wide dynamic range. However, this BLM system is rather expensive as ACEM are not a standard tube of PM-suppliers.

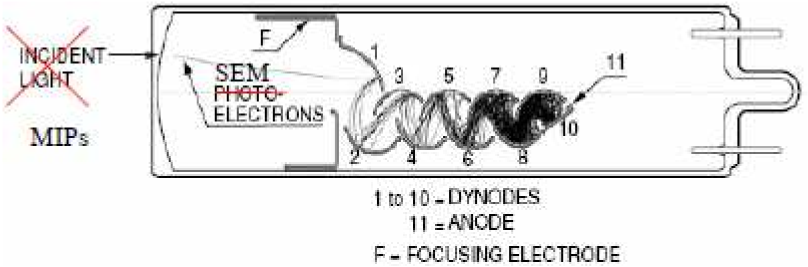


Figure 9: ACEM

3.6.5 Scintillators

Detection of beam-induced light in scintillators is another common method of monitoring beam losses. The scintillation process is also based on the Bethe-bloch dE/dx formula. The scintillators are usually liquid or plastic and read out by a photodetector usually a photomultiplier tube (PMT), a vacuum photodiode or a SiPM.

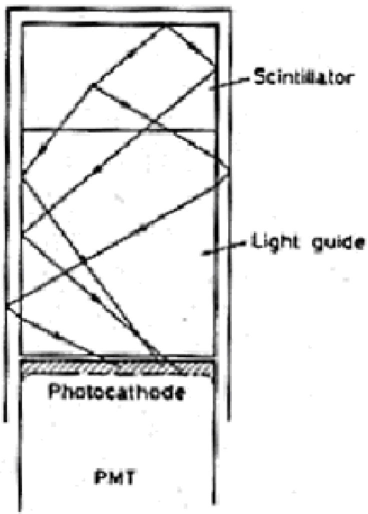


Figure 10: BLM based on a Scintillator read out by a photomultiplier

The light L produced by a particle of energy in the scintillator is for small dE/dx (e.g. for MIPs):

$$\frac{dL}{dx} = R_s \frac{dE}{dx} \qquad (3.1)$$

where R_s is the ratio of the average number of emitted photons to the energy of the incident radiation absorbed by the scintillator. The scintillator light is transported from the scintillator layers to the photosensitive device via light guides (internal multiple collections) or via a wavelength shifter (WLS). The cross section of the scintillator should not be larger than the cross section of the light guide. The shape of the scintillation pulse is characterized by a fast rise time of the order of 1ns and a decay time of a few ns.

3.6.6 Cherenkov radiators

Cherenkov light is the light emitted when a charge particle's velocity βc is greater than the phase light velocity c/n in a media with an index of refraction n>1. The charged particles polarize the molecules of that medium, which then turn back rapidly to their ground state, emitting radiation in the process. The light emitted has a characterizing blue colour. Photons are emitted at an angle defined by the velocity of the particle β and the refraction n of the medium.

The light from the Cherenkov radiator is read out by a photodetector (like scintillators) e.g. a PMT, ideally sensitive in the blue wavelength region. The photon yield dN emitted per dx for a Cherenkov radiator coupled with PMT with a photocathode sensitive in a wavelength interval of (λ_1, λ_2) is:

$$\frac{dN}{dx} = 2\pi a \sin^2\theta \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right) \qquad (3.2)$$

where $\cos\theta = 1/\beta n$ with $\beta > 1/n$.

Cherenkov light is instantaneous, unlike scintillators, and the threshold for light output ($\beta > 1/n$) is above Compton-electron energies of several hundred keV, making Cherenkov detectors useful where there is background radiation from RF cavity X-rays or synchrotron radiation, such as high-energy electron rings and superconducting RF cavities.

3.6.7 Scintillating Fibers and Optical fibers

Scintillating fibers are proposed as a possible beam loss monitors. An interesting suggestion is to use Ce-activated Li glass (with Li6) to detect neutrons. In general, the internal reflections and resultant attenuation are excessive unless a graded or stepped index fiber is used. A wavelength shifter must be used to limit self-absorption. For light produced isotropically in the fiber, only about 2% is in the cone that will be internally reflected. Lastly, the volume of the fiber is too small to produce sufficient light for most applications. However, scintillating fibers are very sensitive to radiation. Radiation darkening will probably limit the use to < 100 Mrad.

Another proposal for a BLM sensitive only to charged particles is a monitor based on optical fibers (fused silica) which produce Cherenkov light and allows real time monitoring of loss location and loss intensity like in PLICs. The fast response of the Cherenkov signal is detected with photomultipliers at the end of the irradiated fibers. There are two major issues to address when considering the Cherenkov effect in single quartz optical fibers:

1) the light yield caused by the passage of a single charged particle in a fiber.
2) the probability of survival of the emitted photons.

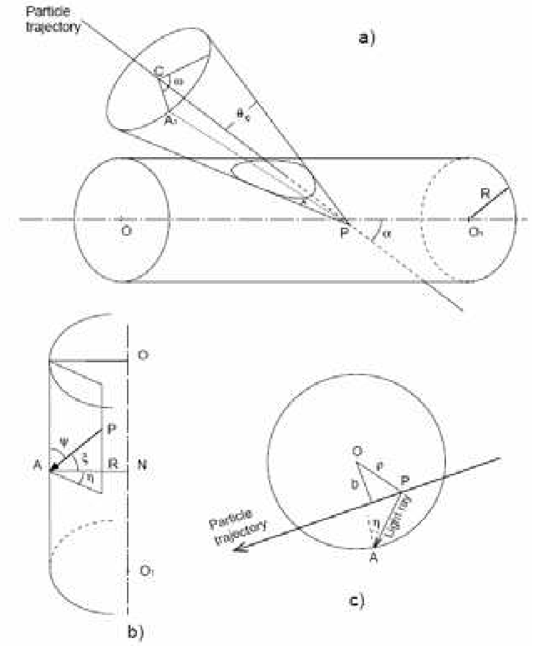


Figure 11: BLM based on an optical fiber

For an optical fiber with a numerical aperture ($NA^2 = n_{core}^2 - n_{clad}^2$) of 0.3 and assuming β=1 for all the charged particles hitting the optical fiber, the angle under which the Cherenkov light is emitted is $\Theta \approx 47^\circ$. The condition for capture and transport down the fiber is given by:

$$\xi \geq \arcsin\left(\frac{n_{clad}}{n_{core}}\right) \qquad (3.3)$$

where ξ depends on (shower) particle trajectory. Thus, the distribution of photons inside a fiber is not only a function of the numerical aperture of the fiber and the charged particle velocity, but also of the impacting particle's angle α and impact parameter b.

3.6.8 Conclusion

Most of the beam loss monitoring systems are based on current-integrating BLMs which produce a wide dynamic range response for the detection of both regular and irregular losses. Thus, the most widely used BLMs are either ionization chambers or the combination of scintillators coupled to photon-to charge converters (PMT, APD, PIN etc). In situations where there is low-energy background ionizing radiation (synchrotron radiation or RF cavity x-rays), photodetecor-Cerenkov radiators or PIN-diode pulse coincidence circuits can be used. Lead shielding can be used around any beam loss monitor to reduce the sensitivity to background x-rays.

CHAPTER 4. MATERIAL DAMAGE POTENTIAL SIMULATIONS

4.0 Introduction

In the problem treated in this study, a high energy electron beam is hitting a copper tube. When an electron impinges on the tube's face it penetrates the copper interacting with the matter in different ways and creating a particle shower. Each of all the possible interactions has a certain probability. The goal of these simulations is to find the longitudinal and lateral energy density distribution developed in the copper tube. This allows estimating the damage potential of high-energy electron beam.

4.1 The Monte Carlo Method

High speed computers and sophisticated computational methods have added a new dimension to scientific research. As many physical systems possess complex functionality that is hard to track the system's behaviour by formulas, the numerical modelling the system is often preferred to the analytical approach. However, the convergence to a good approximation of an accurate solution is likely to be slow and complicated. Therefore, simulation techniques in order to build an experimental model of a system have been developed by taking advantage the speed of electronic computers. Thus, the cost of the experiment in terms of time and money is reduced whilst the degree of flexibility is increased due to access to a wider range of the experimental conditions. The breadth and the extent of simulation models is extensive.

Monte Carlo method is one of the available general approaches to numerical modeling of systems with complex, non linear or involving more than just a couple uncertain parameters. It can be used to a wide range of problems from pure to finance. Furthermore, it finds wide application to almost all fields of physics to investigate both deterministic and primarily stochastic problems. A problem is called deterministic when it responds to forcing in one-to-one way ,even if it is not linear (determinable response). On the other hand, a statistic system behaves according to some statistical rules. The problem of particle transportation in matter, and hence the Monte Carlo modeling of electron-matter interactions to be reviewed in this study, belongs to the last category.

Monte Carlo methods are a class of computational algorithms that rely on iterative random sampling of probability distribution functions as model inputs to produce results that provide probabilities of different outcomes occurring (instead of a few discrete scenarios , method known as "what if " scenarios). The distribution functions that are used in any simulation process must reflect reality and be normalized over the appropriate domain. An approximation will be yielded by aggregating the results into a final one. Conspicuously, the main aim of this stochastic process is to evaluate a deterministic model and to determine the properties of some phenomenon.

Tracing back to the origin of its name, "Monte Carlo" was a code name for a secret project that John von Neumann and Stanislaw Ulam had been working at Los Alamos Scientific Laboratory in ‘40s. The two pioneers used that mathematical method to do hydrodynamic computations by using random numbers. The method was named after the famous Casino in the Mediterranean Principality of Monaco, in honour of an Ulam's relative having a propensity to gamble. However, the choice of the name does not mean to imply any "gambling" tinge in the method. Instead, it merely refers to the manner in which individual numbers are selected from valid representative collections of input data so they can be used in an iterative calculation process.

As pointed out above, Monte Carlo simulation involves calculating the average of probable behaviour of the system in question by observing the outcomes of a large number of repetitions. Apparently, that number has to be sufficiently large to ensure the accuracy of results. By the law of large numbers* this method will display 1/Sqr (N) convergence. Not only does that endeavour for a more accurate approximation depend on the number of trials. Instead, there is also a computational reliance on more uniformly distributed random numbers. Thus, that requirement of large amounts of random numbers for the statistical sampling used in Monte Carlo methods has spurred the development of pseudorandom number generators.

4.1.1 The Monte Carlo Method & Pseudorandom Numbers

The pseudo-random number generators (RNG) are mathematical algorithms that generate numbers and they are the essence of a Monte Carlo simulation .The numbers are not truly random in the sense that they can be predicted and have a deterministic behaviour. However, if the algorithm is good enough, a large set of pseudorandom numbers show the same statistical properties as true random numbers. Moreover, the use of pseudorandom numbers is much simpler and practical approach than the use of real random numbers, as they ensure the possibility of reproducing the performance of an exact same simulation and take advantage of the deterministic nature of the computers.

As pointed out above, the true stochastic or random nature of particle interactions is imitated by means of the pseudo-random number generator (RNG). In each trial in a MC simulation random numbers have to be produced as if they were single samples from of the random variable with a certain probability density function (PDF). A typical RNG algorithm is a periodic algorithm with a very large period (results repeat after millions of cycles) and it is initiated by a number called a seed. Two executions of the algorithm with the same seed give sets of equal pseudorandom numbers. Even a small variation of the seed gives completely different set of pseudo-random numbers.

4.1.2 The principle of Monte Carlo method

Monte Carlo method is following a series of steps in order to simulate particle interactions.

Firstly, a primary particle is generated with predefined fixed initial conditions (initial position, initial momentum). The particle travels inside the current medium for a certain distance before it interacts first time. The distance's value is sampled by a random number generator from a PDF which differs for each type of particle and material.

Afterwards, all the possible N interactions that can take place are considered and given a certain probability. The probability of occurrence π is assigned to the i^{th} interaction by a program internal algorithm depending on the cross sections from different processes. Evidently, Sum(Pi) must be 1. After sampling another uniformly distributed random number the type of interaction occurring is decided. Changes of the position and the momentum of the primary particle will be taken into account by the end of the particle step, as a result of the occurring interaction.

Additional samplings of random numbers may be needed in order to decide the value of other variables such as scattering angles etc. Moreover, there may be generation of secondary particles, due to interaction or decay, which will also have to be transported.

A program internal "step size" determines the space interval in which continuous effects on a travelling particle are taken into account. At the end of the step, it is computed a correction to the mean path length and also the lateral displacement. The step size may not be a variable depending on the particle type, its energy and current position. It is crucial that its value should be smaller that the smallest dimension of the adjacent regions. Volume boundaries, threshold for secondary particle production, discrete location interaction decide the particle step length.

All active particles at the moment that is a primary and its secondaries are placed on the computer stack with their properties (type, spatial coordinates, momentum, etc). There is also the possibility that a particle is annihilated during the interaction and it is no longer taken into account. Another way to “disappear” a particle is that it might also be stopped.

That annihilation may occur due to several scenarios such as if the particle undergoes an inelastic process- being absorbed, transformed or annihilated-, if it surpasses the geometry's boundaries, or if its energy is lower than predefined cut-off energy. In that cases it is removed from the stack of tracked particles, after the quantities of interest are recorded by the Monte Carlo code.

After the first interaction has finished, the primary particle is put on the stack with the others and the steps above are repeated for all given primaries and the Monte Carlo method records the quantities scored in regions of interest .The number of primary particles should be well-balanced between the time consumption of the simulation and the desired precision of the results.

4.1.3 The efficiency of Monte Carlo Method

The efficiency of a Monte Carlo simulation is a crucial factor and it is defined by:

$$E = \frac{1}{s^2 T} \qquad (4.1)$$

Where s and T are the variance on the quantity of interest and the computing time to obtain a variance s^2, respectively.

The statistical accuracy depends strongly on the number of particle histories which creates a balance issue between time and accuracy. In fact, there are several efficiency improvement techniques (so called EIT) in order to increase the statistical accuracy of Monte Carlo calculations without increasing the number of particle histories. These are used to reduce the calculation time with a given statistical uncertainty. The relationship between variance s^2, number of histories N and the time T:

$$s^2N = \text{const.}$$

$$s^2T = \text{const.}$$

As seen from the equations above, particle histories N is directly proportional to time T.

Efficiency improvement techniques are distinguished between techniques that achieve the improved efficiency through the use of approximations and techniques that do not alter the physics in any way when they increase the efficiency. The former are called approximate efficiency improving technique (AEIT), when the latter are called true variance reduction technique (VRT).

The most common VRTs are the splitting and Russian Roulette techniques which were originally proposed by J.von Neumann and S.Ula. The splitting technique can vary depending on the routines followed for splitting. There also other techniques such as particle weight. Depending on the combination and the splitting routines chosen, the simulation code can be very efficient.

On the other hand, there are three more frequently used AEITs. The first one, the condensed history technique (CHT) consists of two main components which influence the simulation speed and accuracy, the “electron-step algorithm” and the boundary crossing algorithm. Because of the importance of this AEIT, it is described in more details in the next chapter. Another AEIT is the Range Rejection where an electron is discarded if its residual range is smaller than the distance to the nearest boundary. In addition, there is the Region Rejection where an electron is discarded when it is relatively far from the region of interest. Finally, the third AEIT is the Transport Cutoffs. According to that technique, particles are not further transported and they do not create secondaries as soon as their energy drops below some certain threshold (ECUT & PCUT and AE & AP, respectively).

4.1.4 Electron transport modeling in Monte Carlo method

A particle passing through matter interacts with electrons and nuclei, possibly with the medium as a whole. This chapter summarizes how the transport problem of particles in matter can be modelled by the analog Monte Carlo technique and some of the principal interaction mechanisms.

Monte Carlo simulation of particle transport consist a faithful simulation of physical reality. However, the dependence of the desired accuracy or the complexity and size of the geometry lead to very long calculation times. In addition to that, another difficulty occurs in case of the simulation of electrons. Charged particles like electrons and positrons and hadrons lose their energy in a considerably more complicated way from photons. A relativistic electron slowing down in a material must undergo a number of the order of 10^5-10^6 interactions with surrounding matter before losing all of its kinetic energy. However, while few of these interactions cause significant energy loss and directional change, most of them cause minor changes in the particle's energy or direction of motion. Because of the large number of histories of individual small-effect interactions that are taken into account at the end of each electron step, a simulation of electrons transport could be often unpractical in terms of calculation effort.

To circumvent this difficulty, Berger in 1963 developed the "condensed history" technique, the first of the Approximate Efficiency Techniques that we mentioned before. According to that technique, large number of transport and collision small-effect processes of electrons is "condensed" to a single electron step. In other words, it sums the effect of up to thousands of discrete, small momentum transfer elastic and inelastic collisions into single larger-effect quasi events. Recently, Larsen has provided a mathematical proof that the condensed history technique is an approximate solution to the Boltzmann transport equation in the limit of small step size. M. Berger divided electron transport into two basics classes as class 1 and class 2 schemes. Geant4 uses class 2 algorithms and hence the electron steps are stochastic. In this scheme, the thresholds for interactions which cause energy loss and deflection of the primary electrons are introduced. Processes under these thresholds are all taken into account in a model called continuous slowing down model (CSDA). On the opposite, for inelastic collisions having higher energies than the thresholds, photons are generated by the Bremmstrahlung process and Moller knock on electrons set in motion are treated discretely by creation and transport (?). In the paragraphs below, more details and definitions of the possible interactions are given.

In general, at electron energies considered here the electrons striking on a target interact by developing an avalanche of electrons, positrons and gamma rays. This avalanche is known as an electromagnetic (EM) cascade or shower. Respectively, when hadrons are impinging on a material they interact by developing an hadronic cascade. The two categories of cascades are distinguished by the main interaction channel of the primary particles. For electromagnetic showers the main interaction channel is the electromagnetic force, when for the hadronic showers it is the nuclear interaction. In both showers newly produced particles, so-called, secondaries particles, have sufficient energy to interact themselves.

An accelerated charged particle will emit electromagnetic radiation. The intense accelerations can produce further gamma ray caable of producing more electron-positron pairs. The cycle of pair production and gamma ray generation continues with the original gamma ray energy eventually manifesting itself as many particles. The process of a typical electromagnetic cascade initiated by an electron, as the ones simulated in this study, is depicted in the image below:

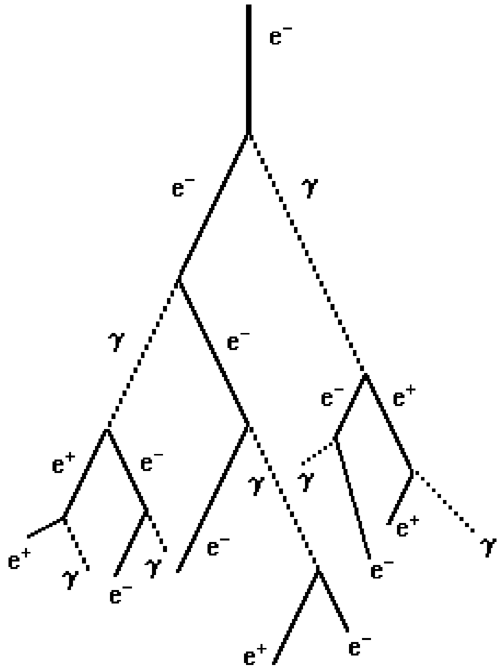


Figure 12: Electromagnetic cascade. The cascade generation ceases when the shares of energy get sufficiently small such that the electrons are no longer capable of radiating efficiently; the relatively slow moving electrons are then brought to treat by ordinary ionization processes. The electron energy at which the main energy loss mechanism changes from radiation losses to ionization losses is known as the critical energy Ec.

Electron and photon beams in the GeV and more range are capable of initializing a high energy hadronic shower too. The most important source of high energy hadrons is the hadronic interaction of photons from the photo-nuclear reactions. Thus, hadronic cascades are electromagnetic to large extent. Photons that will be produced as part of the electromagnetic cascade will be subject also to hadronic interaction whose cross sections are typically a factor 100 or smaller than the pair production and Compton scattering cross sections. The sufficient energy for an inelastic hadronic interaction of a photon is 200MeV, Nevertheless, as CLIC beam losses concern large amounts of electrons, the yield of hadrons observed is significant [CLIC Note] The sufficient energy for an inelastic hadronic interaction of a photon is 200MeV.

As already mentioned, one of the characteristics of a typical hadronic shower is the production of secondaries, such as nucleons and pions. Neutral pions may afterwards decay to gamma-rays and remaining nuclei are often left in an excited state and may also decay. Most of the secondary particle production takes place at energies below 1GeV. The hadronic multiplication process is measured at the scale of nuclear interaction length (λ_{inel}) . The nuclear interaction length is defined as the mean free path of a particle required to reduce the energy of relativistic charged particles by the factor 1/e or 0.368, as they pass through matter. For copper, that value corresponds to 15.32cm.

Dependent on the energy of the primary particles, an electromagnetic component rises not only in electromagnetic showers, but also in a hadronic shower. That happens because the mean free path for strong interactions, that is the collision length, is large; hence the electromagnetic energy loss of hadrons is important. To distinguish the nuclear collision length from the interaction collision length: the nuclear collision length is defined as the mean free path of a particle before undergoing a nuclear interaction that is neither elastic nor quasi-elastic (diffractive), in a given medium. The collision length is smaller than the nuclear interaction length (σ_{total}) because the elastic and the diffractive reaction are excluded from its definition, as the relevant cross-section is σ_{total}- $\sigma_{elastic}$- $\sigma_{diffractive}$.

Stopping power is used in order to express the ability of a material to reduce the kinetic energy of the particle passing through it and it is treated as property of the material. The linear stopping power corresponds to the energy loss per unit path length such as described from the formula below and its unit is energy per unit length ,such as MeV/cm :

$$S(E) = -\frac{dE}{dx} \qquad (4.2)$$

where E is the energy and x is the path length. The minus sign makes the S positive.

The mass stopping power is the linear stopping power divided by the density (ρ) of the substance and it is the most commonly used. It is divided into the "collision stopping power" and "radioactive stopping power". The former is based on the energy loss resulting from the sum of collisions, while the latter is based on Bremmstrahlung production alone. Without regarding the radioactive part, it is described by the formula:

$$\left(\frac{S}{\rho}\right)_{col} = \frac{1}{\rho}\frac{dE}{dx} \qquad (4.3)$$

The unit for mass stopping power is MeV cm^2 g^{-1}.

Finally, the atomic stopping power is the energy loss of a particle per distance travelled (MeV/cm) per number of atoms per unit area perpendicular to the particle's motion ,that is volume, (atoms/cm^3) and it is proportional to the square root of the atomic weight (Bragg-Kleeman rule). The respective formula is given below:

$$\frac{1}{n}\frac{dE}{dx} = A\frac{dE}{dx} \qquad (4.4)$$

Where n is the number of atoms in unit volume of substance,
N is the Avogadro number,
A is the relative by the Bethe-Bloch formula (A= ρN).

Except for electrons, charged particles which are not highly relativistic lose energy primarily lost by electromagnetic interaction between the incoming particle and the electron cloud of the atom. Hence ionization by liberating a single electron, as shown in the picture below, is the main electromagnetic contribution to the energy loss.

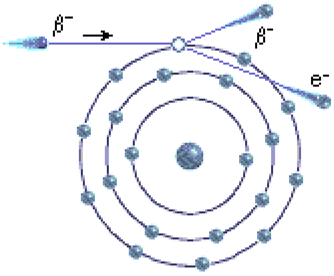


Figure 13: Ionization

The stopping power due to ionization (reglecting the radiative part) is given by the Bethe-Bloch formula:

$$-\frac{dE}{dx} = \frac{4\pi n z^2}{m_e c^2 \beta^2} \cdot \left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \cdot \left[\ln\left(\frac{2m_e c^2 \beta^2}{I \cdot (1-\beta^2)}\right) - \beta^2\right] \quad (4.5)$$

where:

= v / c
velocity of the particle
energy of the particle
distance travelled by the particle
speed of light
particle charge
charge of the electron
rest mass of the electron
electron density of the target
mean excitation potential of the target
vacuum permittivity

The stopping power and hence, the density of ionization, increases toward the end of the range and reaches a maximum, the Bragg peak, shortly before the energy drops to zero, when it comes to protons, α-rays, and other ion rays. The curve that describes this is called the *Bragg curve* .In the picture below the Bragg curve of 5.49MeV alpha particles in air is shown:

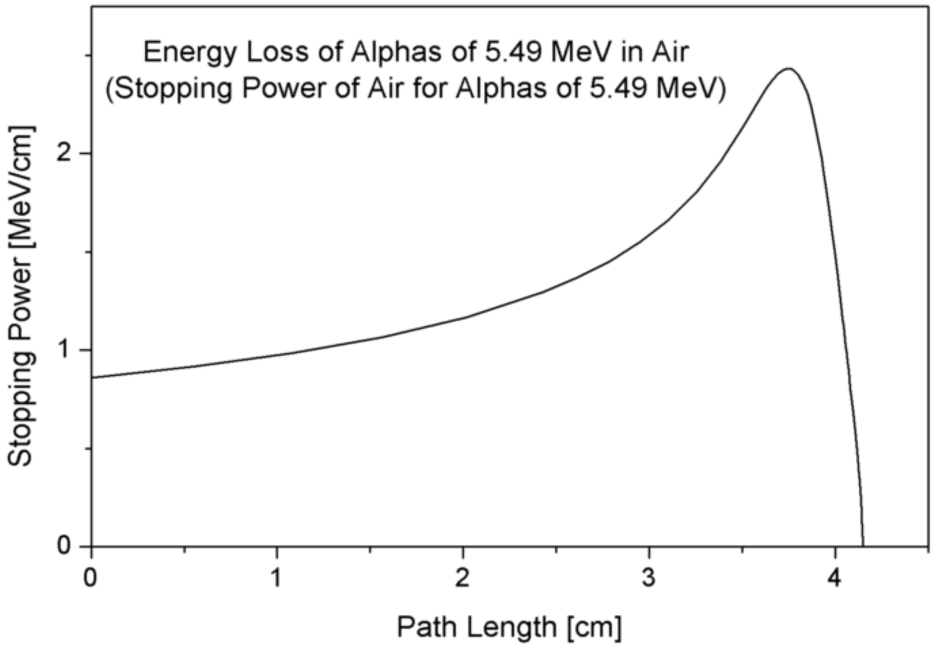


Figure 14: The Bragg Curve of 5.49MeV alpha particles in air

Now that more definitions and details were given about the electromagnetic component of a cascade we can proceed to the way Monte Carlo code models it. So, Monte Carlo code models the electromagnetic cascade by dividing the electron/positron interactions with matter in two categories:

"Catastrophic" interactions

- Large energy loss Moller ($e^- e^- \rightarrow e^- e^-$) and Bhabha ($e^+ e^- \rightarrow e^+ e^-$) scattering
- from atomic electrons
- Hard Bremmstrahlung emission ($e^\pm N \rightarrow e^\pm \gamma N^*$)
- Positron annihilation with "in-flight" and at-rest with atomic electrons

"Soft" events

- Low-energy loss Moller and Bhabha scattering ,modeled as a part of the collision stopping energy
- Atomic excitation ($e^\pm N \rightarrow e^\pm N^*$) ,modeled as another part of the collision stopping energy
- Soft Bremmstrahlung emission ,modeled as radiative stopping power
- Elastic multiple scattering of electrons and positrons from nuclei
- Excitation of atoms and molecules by electrons and positrons

Bremmstrahlung production is the creation of photons by electrons (or positrons) in the field of an atom. There are two possibilities. The predominant mode is a two-body interaction where the nucleus recoils. This effect dominates by a factor of about Z^2 over the three-body case where an atomic electron recoils ($e^\pm N \rightarrow e^\pm e^- \gamma N^*$). The two-body effect can be taken into account through the total cross section and angular distribution kinematics. The three-body case is conventionally treated only by inclusion in the total cross section of the two-body process. In Feynman diagram below is shown the predominant mode of the Bremmstrahlung production.

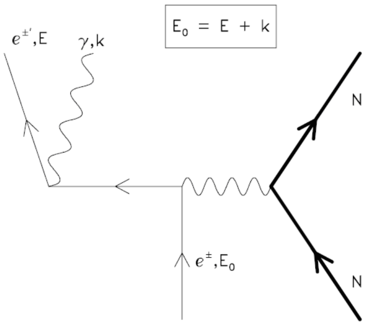
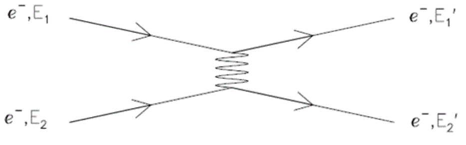
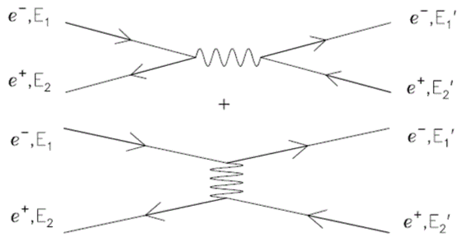


Figure 15: Bremsstrahlung

Moller and Bhabha scattering are collisions of incident electrons or positrons with atomic electrons. They appear to be similar at first glance but there is a difference between the participant particles, as depicted in Feynman diagram 16 and the thresholds for these interactions is also different. We adapt the convention that the atomic electrons are “free”, by ignoring their binding effects. The main difference between the two interactions is the extra interaction channel to the cross section for the Bhabha scattering which is contributed by the possible annihilation and recreation of the e^+e^- pair; hence there are two leading-order Feynman diagrams contributing to this interaction: an annihilation process and a scattering process. In fact, the positron in the e^+e^- case can give up all of its energy to the atomic electron, while the respective electron can only give at most half of its energy to the target electron as the two electrons are indistinguishable. The cross section of the Moller and Bhabha interactions scales with Z for different media and approximately as $1/\upsilon^2$, where υ is the velocity of the scattered electron.



Møller



Bhabha

Figure 16: Moller and Bhabha interactions

Electron-positron annihilation occurs when an electron and a positron collide. The result of the collision is the annihilation of the electron and positron, and the creation of gamma ray photons or, at higher energies, other particles. In figure () a two photon annihilation is depicted. For reasons of convenience, we make the assumption that the atomic electrons are free , by ignoring their atomic bending energy, as before. In addition, three and higher photon annihilations ($e^+ e^- \rightarrow n \gamma$ [n>2]) as well as one-photon annihilation ($e^+e^- N \rightarrow \gamma N^*$) can also be ignored. Annihilation is generally not treated as a positron slowing down process and is treated discretely as a “catastrophic event”.

Multiple scattering through predominantly small angles with the occasional large-angle effect scattering event occurs when a charged particle traversing a medium is deflected. Most of this deflection is due to Coulomb scattering from nuclei, and hence the effect is usually called Coulomb scattering. However, for hadronic projectiles, the strong interactions also contribute to multiple scattering. The cross sections are nonetheless very large. There are several statistical theories that deal with multiple scattering. The Moliere theory was originally designed to well represent small angle scattering by being roughly Gaussian for small deflection angles. However, with small modifications it can predict large angle quite successfully by ignoring differences between positrons and electrons, and uses the screened Rutherford cross sections instead of the more accurate Mott cross sections by having larger tails than does a Gaussian distribution.

So far , we have analyzed most of the processes how high-energetic electrons and positrons interact primarily through Bremmstrahlung process, , for lower energy electrons (below few tens of MeVs), ionization is still the predominant way of energy loss. Nevertheless, photons also contribute in an electromagnetic cascade as high energetic photons predominantly lose energy in matter by pair production (e^+e^-) and lower energy photons by Compton scattering.

For reasons of consistency, all of the photon interactions with matter are also presented below with an order depending on their energy (starting for the higher energy interaction):

- Compton effect (incoherent scattering)
- Photoelectric effect
- Pair production
- Rayleigh (or coherent) scattering

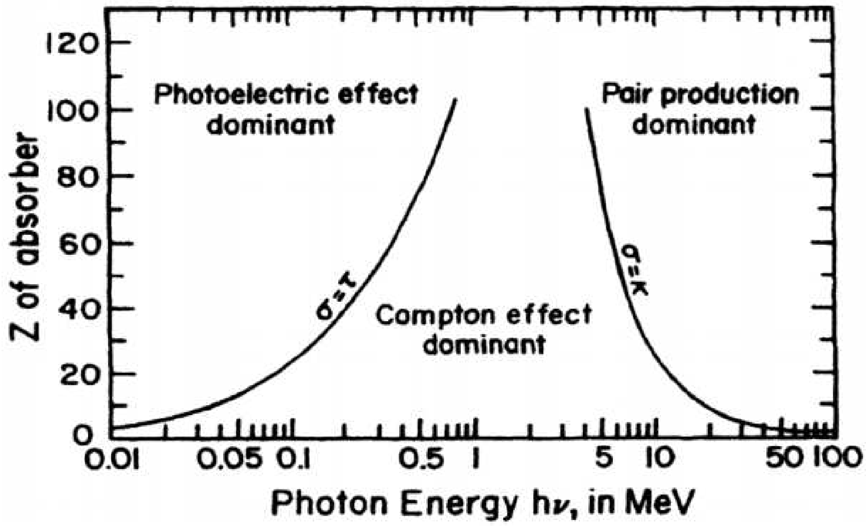


Figure 17: The relationship between predominant types of photon interactions, atomic number (Z) and photon energy.

For example, for a simplified photon transport for a photon with less than 1.022MeV energy, it considers only Compton Effect and pair production. Therefore, the total probability of a photon interaction with matter per centimetre is:

$$\sum tot = \sum Compton + \sum pairproduction\,. \tag{4.6}$$

For energies in the range of several MeV, the most common event is Compton scattering. In this process, the photon scatters from an atomic electron and transfers part of its energy to this electron that is ejected from atomic shell. (figure 18)

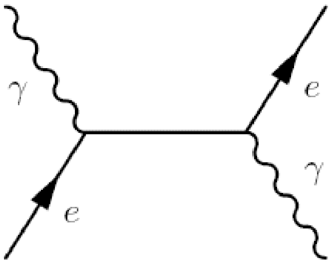


Figure 18: Compton Effect

For lower energy photons, the photoelectric effect is the most likely type of photon interaction. Here the photon is absorbed by the atom and a photoelectron is ejected. (figure 19)

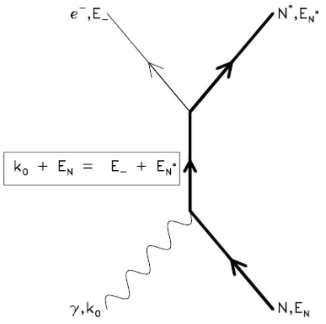


Figure 19: Photoelectric effect

Pair production dominates at photon energies above a few MeV. Pair production is an absorption process in which the photon interacts with the electric field of the nucleus. The photon disappears (is absorbed) and creates an electron-positron pair.

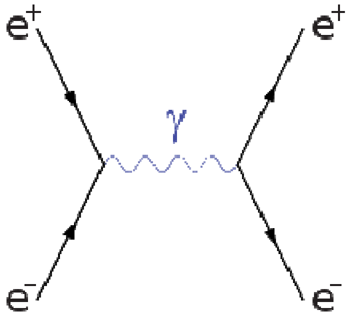


Figure 20: Annihilation and pair production

The last interaction process is Rayleigh (or coherent) scattering, in which the photon scatters elastically from atomic electrons. This interaction only leads to a change in the direction of the photon, and not to energy losses. However, it is not taken into account in the simulations performed.

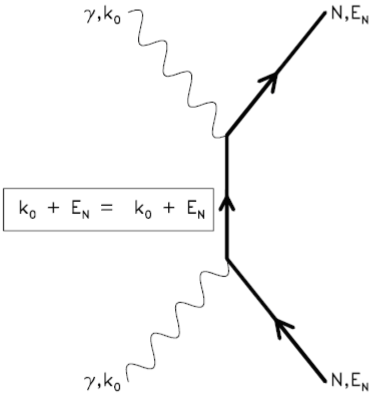


Figure 21: Rayleigh scattering

4.2 Geant4

Geant4 is a toolkit for Monte Carlo simulations of the passage of particles through matter and their interaction. Geant 4 is the successor of Geant3 and takes into account the requirements for a set of different fields. Thus, its multi-disciplinary nature allows its application in high energy, nuclear and accelerator physics, space, cosmic ray, heavy ion and radiation studies as well as medical science. It was proposed to and approved by the Detector Research and Development Committee (DRDC) of CERN at the end of 1994 and a first prototype was delivered at the end of 1995. Nowadays, it is maintained, developed and distributed by a worldwide collaboration of physicists and computer scientists.

In High Energy Physics Geant4 is being extensively used not only for detector and charged particle transport simulation, but also the graphical representation of the setup and of the particle trajectories. Its acronym stands for "Geometry and Tracking" and it actually condenses the functionality that the code provides for its applications. Firstly, it allows the description of the geometry and material of the detector elements. The particles are generated and then tracked through the detector, by simulating their physics interactions in matter and the effect of fields and boundaries on their trajectories. The response of the sensitive detector yields the event data which are stored. Afterwards, the code provides the options of visualization capture and post analysis in different levels of detail and refinement.

Geant4 is written in C++ and it exploits advanced software techniques and object-oriented technology to achieve transparency for the produced physics results. In addition, CLHEP is a mandatory part of underlying software for the Geant4 toolkit. It is a class of libraries for High Energy Physics specific and features such as random generators, physics vectors, geometry and linear algebra.

Multiple implementations of physics processes and models are possible and available. Charged particle transport includes all relevant processes for both electromagnetic and hadronic showers. It is the user’s responsibility to decide which physics processes are required and how they are modelled, and then to include them in the physics list. For that reason, each physics list has been specialized for a given area of application. The physics lists in Geant4 should be pre-declared.

Taking into account the degrees of freedom that the Geant4 code gives to the user and its versatility and extensibility, the Geant4 toolkit was the choice for the simulations presented below.

4.3 Simulations

Detector geometry in Geant4 is introduced by volumes and each of them is defined by three conceptual layers. Firstly, to describe a volume's shape and size the concept of a solid (G4VSolid class) is used. In order to add physical characteristics to the geometrical properties of a solid and describe a volume's full properties such as material, daughter physical volumes, sensitivity, magnetic field, the concept of a logical volume is used. Finally, a physical volume is created for the definition of the position and the rotation of a volume and it is a copy of the chosen logical volume inside a large, containing one called mother volume. The largest volume is called World. It is the only exception that a physical volume is not contained in a mother volume and it must contain all other volumes in the detector geometry.

In the present simulation , World represents the environment where the experimental hall is chosen to be and preferably it follows the detector's shape, therefore it is a tube with bigger dimensions and filled in with Vacuum at 3.e-18Pascal and 2.73K pressure and temperature conditions. The target of interest is a homogeneous cylinder filled in with Copper – single element material- and placed at the center of the hall respect to the origin of the world volume. The choice of the copper as a material for the simulations was based on the fact that most of the CLIC accelerating structures, which are exposed to potential damage by high energy beams, are made of Copper. The dimensions are 2m length and 10cm radius. The density is 8.96g/cm3. The logical volume of the copper tube becomes sensitive by having a pointer to a sensitive detector.

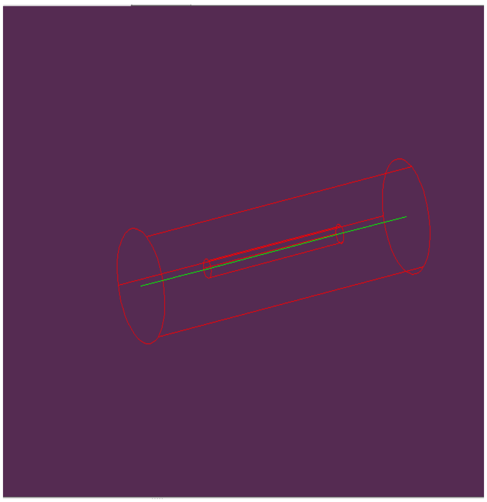


Figure 22: Copper tube geometry introduced in Geant4.

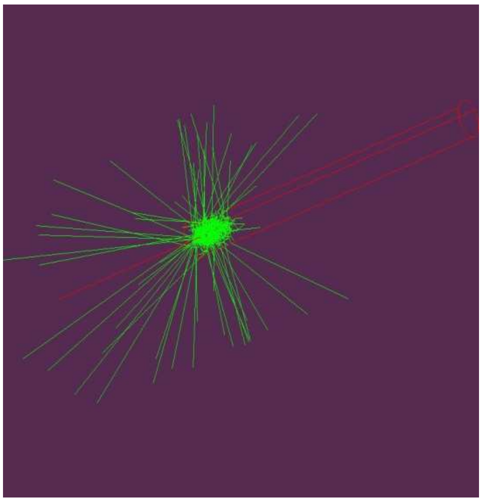


Figure 23: Beam interacting with the copper tube – Particle shower.

A sensitive detector is an abstract base class which represents a detector. The principal mandate of a sensitive detector is to create hits (hit objects) which are snapshots of the physical interaction of a track or an accumulation of interactions of tracks in the sensitive region of the detector. The class has three major virtual methods: Initialize(), ProcessHits() and EndOf Event(). The first one is invoked at the beginning of each event and the hits produced in this particular event are store in the Hit Collection. A hit collection is a dedicated collection instantiated from G4THitsCollection template vector class in order to store the pointers of hit objects. The quantity of interest to store per each hit object in the present study is the energy deposition. The EndOfEvent() is invoked at the end of each event and during the event the ProcessHits() is invoked.

As pointed out above, in order to describe the read out configuration of the detector a readout geometry was associated to the tube's sensitive detector. Tracks are traced in the tracking geometry, the “real” one, and the sensitive detector will have its own readout geometry Geant4 will message to find which “readout” cell the current hit belongs. That requires a previous voxelisation of the experimental tube. A read out geometry is a virtual and artificial geometry which is defined in parallel to the real detector geometry (there is no material associated with it). Parallel geometries in Geant4 are mostly used for event biasing (variance reduction), hit scoring and shower parameterization. As extendibly mentioned in the previous chapter, variance reduction is a general term that refers to a strategy to minimize statistical errors in a Monte Carlo calculation for a given number of iterations.

Figure 24 shows how the association of the tracking geometry and the readout geometry through a sensitive detector object is done in Geant4. The first step is to associate a sensitive detector to a volume of the tracking geometry, in the usual way. The next step is to associate the G4ReadoutGeometry object to the sensitive detector.

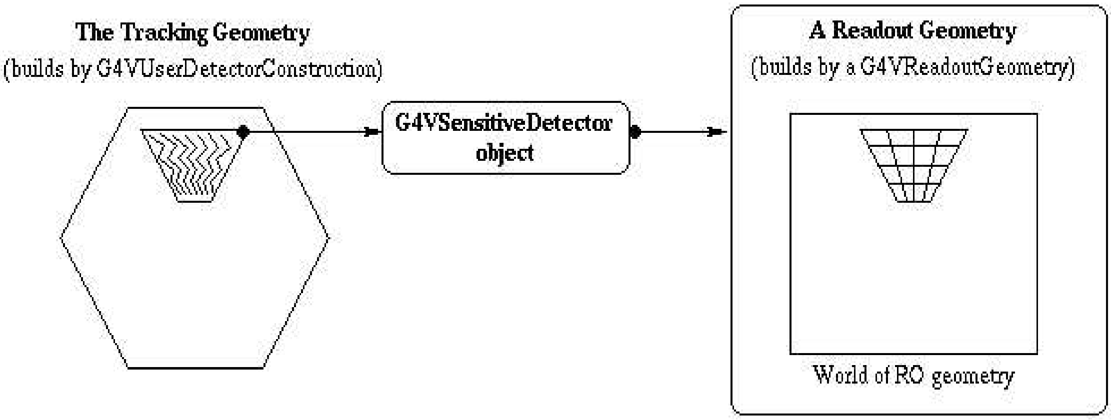


Figure 24: The tracking geometry and the readout geometry are associated through a sensitive detector object . [26]

The read out segmentation of the experimental tube is implemented by three levels of replica. G4PVReplica class is used for the repeated volumes in order to simplify the repetition along one axis and also reduce the use of memory. Therefore, the daughters of same shape will have to be aligned along one axis and fill the mother completely without gap in between. In reality, there is only one physical volume object for each level and its position is parameterized by its copy number.

For the voxelisation of the readout geometry of the copper tube replication occurred along z, rho and phi axes. Thus, the tube is virtually divided into 200 cells along z-axis,100 cells along rho-axis and 4 cells along phi-axis that corresponds to each cell's dimensions of 1mm radius, 1cm width and 90degrees angle.

Another mandatory class that is introduced in each Geant4 simulation, as this, is the PrimaryGenerator class. Each particle has a process manager, thus each particle has a set of many processes at rest, along step and post step. The target of solid copper is irradiated by a monoenergetic, parallel pencil electronic beam centered on the cylinder axis. The primary particles of interest for CLIC studies are electrons, but there were also performed simulations for LHC energies for protons and also for muons presented at the appendix for reasons of comparison and complicity of the study. There are six cases of primary particles which are 100GeV, 500GeV, 800GeV, 1.0TeV, 1.25TeV, 1.5TeV electrons. The number of events simulated is set to 10000 and it is introduced by the macro file that is taken as an input argument during the simulations. In addition the simulation, were performed with 1mm cut on secondary particles range which corresponds to the 1/10th of the cell's size along Z-axis.

As far as the physics list chosen for the results of the present simulations that is the QGSP_BERT_HP. QGSP_BERT is the basic physics list applying quark gluon string model for high energy interactions and Geant4 Bertini cascade for energies below ~10GeV for protons, neutrons, pions, and Kaons . The HP part of the list's name stands for the usage of the data driven high precision neutron package (NeutronHP) to transport neutrons below 20MeV down to equithermal energies (i.e.~1eV).

The interpretation of the data gained from the simulation provides understanding of the phenomenon under investigation. The results of the simulations they are all printed in root file by filling a ntuple (Ntuple in ROOT is a tree with a specially constructed Event object). Afterwards, the post analysis is based on histograms plotted in ROOT which use as an input file the .root output file of the simulation. ROOT is an object-oriented framework for large scale data analysis.

Geant4 provides the implementation of an interface to multifarious graphic systems (OpenGL, OpenInventor, DAWN, HepRep,RayTracer). WIRED4 is another external package that was used in the present study in order to visualize the detector geometry and the particles' trajectories, feature very helpful for the debugging of the geometry. WIRED4 works as a plug-in module in JAS3 (Java Analysis Studio) which is a general purpose, open-source, analysis framework. WIRED4 uses as its input format HepRep and it supports viewing of events using either conventional 3D projections or specialized projections. The user can choose the objects (geometry parts, hits, tracks) displayed and has also options such as changing parameters, scale and rotate as he wishes.

4.4 Results

Hadronic and electromagnetic cascades are induced by the high energy electron beams as electrons interact with target particles and generate new particles as a result of their interaction. The generated secondary, particles deposit a part of their energy in the material through various processes but mainly through electronic energy loss. However, the deposition of energy leads to a temperature rise and hence the probability of damage increases. Damage in the target is a function of the longitudinal coordinate (parallel to the beam axis) as well as transverse coordinates (perpendicular to the beam axis). The outcome of the simulations and their comparison with previous results from simulations, experiments and bibliography gives confidence that beam induced damage can be adequately predicted with simulations.

As it was described at the previous section, for the presented simulations, an electron beam is impacting on a homogeneous cylindrical target of solid copper. Geant4 simulations for energy deposition were used to define six different electron energies: 100GeV, 500GeV, 800GeV, 1.0TeV, 1.25TeV, 1.5TeV. The choice of the energies corresponds to various stages of beam acceleration in CLIC.

The target's dimensions are 200cm length and 10cm radius. Energy deposition was scored as a function of distance from the beam axis, of depth in target and of angle due to the read out voxelisation of the geometry. Because the angular is uniform distribution only the longitudinal and the lateral profiles are shown below.

If the measured values are integrated over the radius, the longitudinal distribution of the shower ("transition curve") is obtained. Figure 25 contains the plots of the energy density for the most exposed azimuth as a function of the longitudinal distance for the six different beam energies. The development of the cascade is a question of balance between the absorption of low-energy particles and their reproduction by high energy particles, which process is described by the parameter λ .The maximum energy density is reached after about a certain distance (about 10cm) which depends on the energy of the primary particles. The horizontal error corresponds to the bin error given by root while the vertical error is a function of the energy (0.01%) (i.e. statistical).

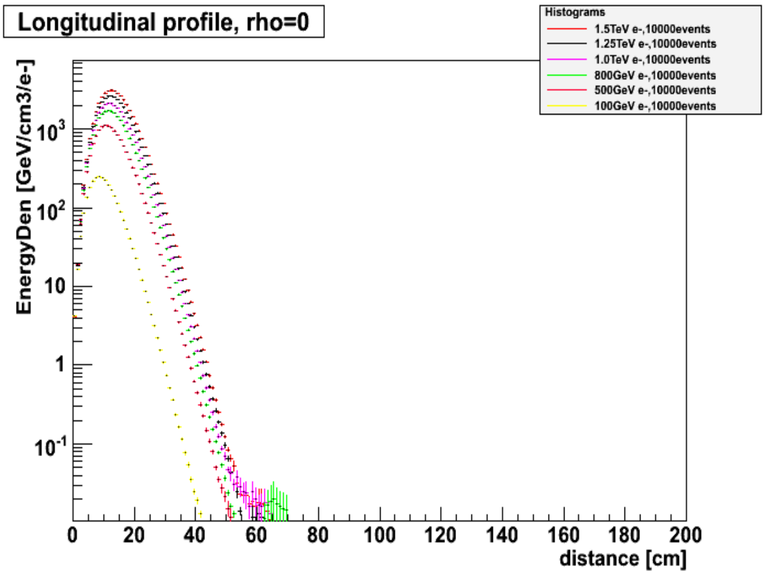


Figure 25: The energy density longitudinal profile in the center of the target estimated over 4 orders of magnitude

.It is obvious from the figure below that more statistics are needed in order to have smaller errors in more than six orders of magnitude. Figure 27 shows that the maximum of energy density is as close as 10-20 cm to the end f the impact of primary electros and 99% of the deposited energy is contained within 60cm of the target depth.

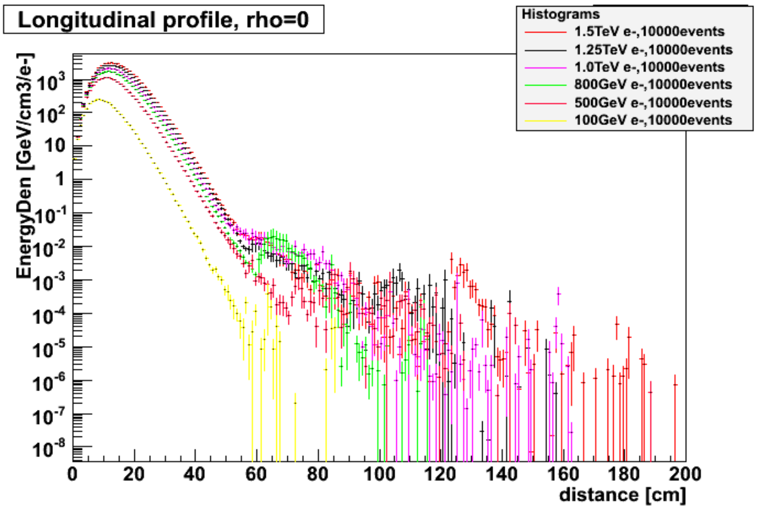


Figure 26: Energy density longitudinal profile in the center of the target

The average differential longitudinal energy deposit over the volume of the cascade , according to bibliography [50] can be fitted by the following formula:

$$\Delta E = kt^{(a-1)}e^{-bt}dt \qquad (4.7)$$

with a, b as fitting parameters, t corresponds to the depth starting from shower origin in units of Xo(1.436cm for Cu) and k a normalization factor. In Figures 27 it is impacted that fitting which seems to correspond well for the beginning of the cascade as after a certain distance the big errors contribute badly in the value of the chisquare of the fitting.

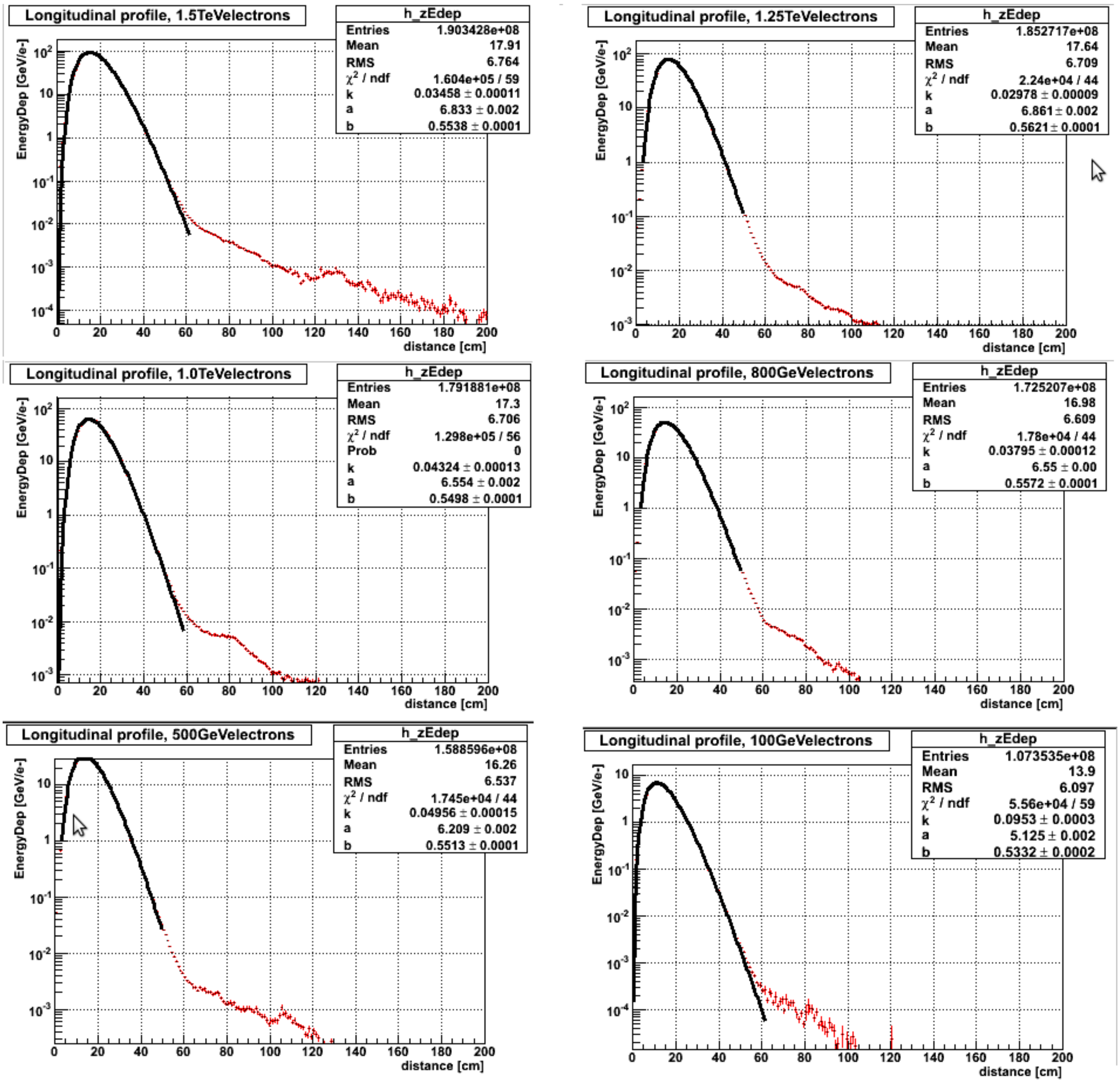


Figure 27: Average differential longitudinal energy deposit for six different energies of the electron beam.

The lateral profile of the developed cascade at the six different energies is shown in Figure 28. As expected, it follows an exponential shape for all of the energies and most of the energy deposition is concentrated close to the beam axis.

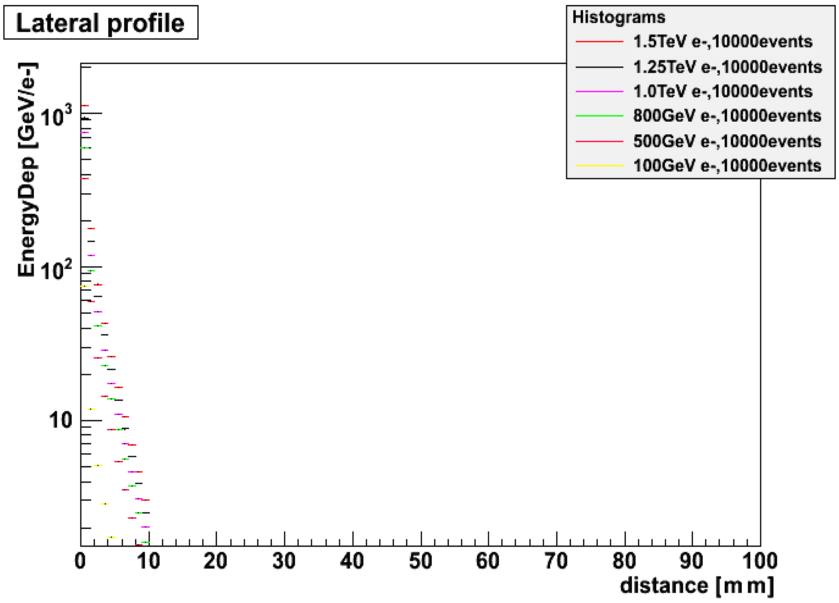


Figure 28: Lateral profile of the developed cascade at the six different energies.

As the impacting electron energy increases, so does the maximum energy deposition and hence energy density. That dependence is fitted with a logarithmic or a power law .Figure 29 and 30 provide two fittings of the maximum energy densities deposited by the beam as a function of the beam energy. According to the chisquare value of these two fittings, the power law fits better. The horizontal errors have value equal to 0.01% of the energy and the vertical errors have the error values derived from the histograms.

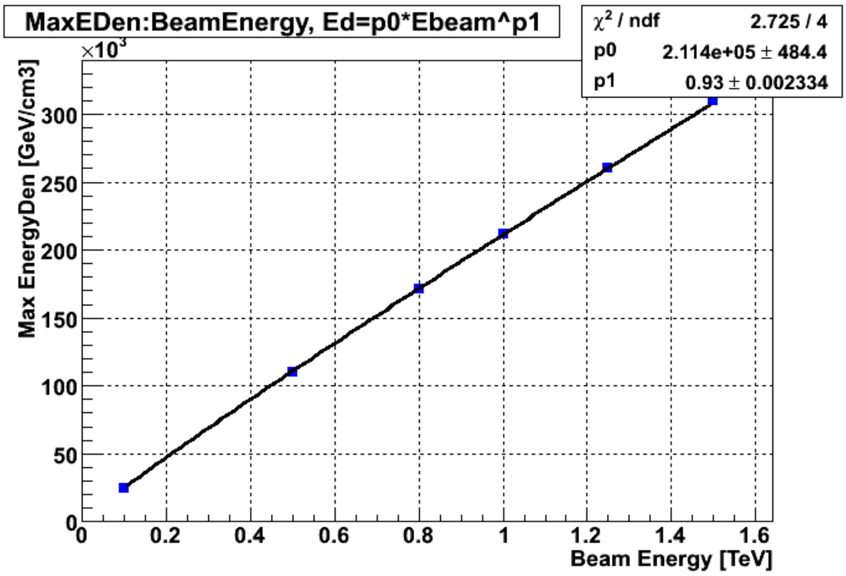


Figure 29: Fitting of the maximum energy densities deposited by the beam as a power function of the beam energy

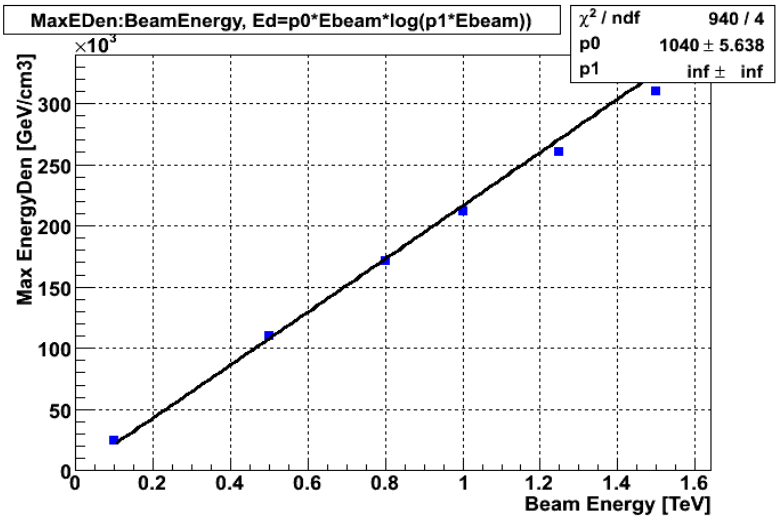


Figure 30: Fitting of the maximum energy densities deposited by the beam as a logarithmic function of the beam energy

Finally , the shower maximum ,with the largest number of particles, is reached when the average energy particle becomes low enough to stop further multiplication. Following the formula of the bibliography [50] the shower maximum can be fitted by the : t_{max}~log(E/ ε)- α (in units of Xo), with α =1 (for e-), E the energy of incident particle and ε the critical energy ,which for Copper and e-cascades it's found to be 19.42MeV. The fitting is shown in figure 31 below.

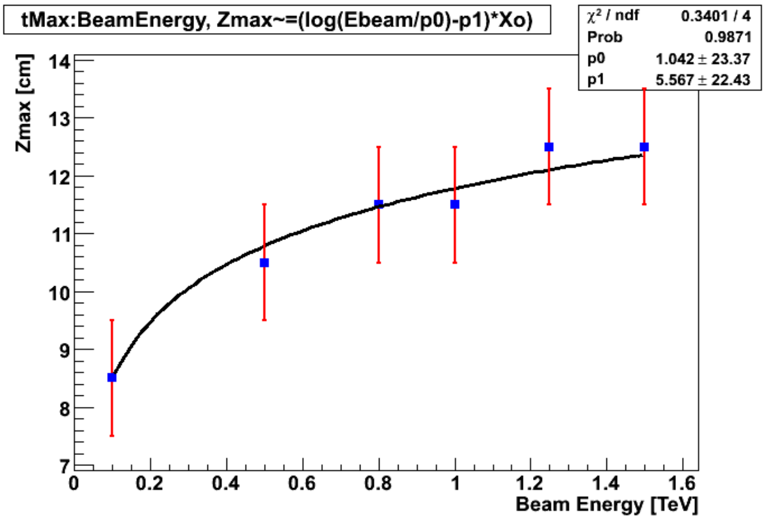


Figure 31: Shower maximum fitting.

CHAPTER 5. PHOTODETECTION THEORY

5.0 Introduction

Detectors efficient for photon counting and low light level detection (LLL) are required in a large variety of next generation experiments for high energy and astroparticle physics. For such measurements, the photon detectors typically employed are vacuum photodetectors (photomultiplier tubes –PMT, micro-channel plate photomultiplier tubes –MCP-PMT, or hybrid photodetectors – HPD), PN or PIN photodiodes and avalanche photodiodes (APD). More recently, arrays of Geiger mode APDs (GM-APDs) have been developed as a new promising photon detector to overcome the drawbacks of all the above mentioned devices.

Photomultipliers have been developed over the past 100 years. The first photoelectric tube was produced by Elster and Geiter in 1914.Photomultipliers have high internal gain (10^6-10^7), very good timing resolution (ps) and good single photoelectron resolution. However, they also have low quantum efficiency limited by the photocathode materials, require high operation voltages and are sensitive to magnetic fields. In addition, the high price, the bulky shape and the sensitivity to handling due to the vacuum technology, led to the development of alternative detectors such as solid state detectors.

The step-by-step evolution of solid-state photon detectors was mainly determined by their internal gain: a PIN photodiode has no gain, an Avalanche Photodiode has a gain of few hundreds and, only in recent years, with the invention of the Geiger-mode avalanche photodiodes with 10^5–10^6 gain, have the semiconductor photo detectors reached sensitivity comparable to that of photomultiplier tubes (single photon sensitivity). The following chapter explains the basic principle of the photo-detection in semiconductors and a brief history of Silicon Photomultiplier development.

5.1 Photodetectors based on Semiconductors

The purpose of any photodetector is the conversion of electromagnetic radiation (light) into an electronic signal- ideally one that is proportional to the incident light intensity. Photodiodes are fabricated from semiconductor materials. Semiconductors are subject to optical charge carrier generation and therefore most are packaged in light blocking material. In photoconductors, the semiconductor is packaged so that it can be exposed to light. The most popular semiconductor choices in photodetection are silicon (Si) or gallium arsenide (GaAs), and others include indium antimonide (InSb), indium arsenide (InAs), lead selenide (PbSe), and lead sulfide (PbS). Each material absorbs light over a characteristic wavelength range.

The basic principle of photodetection in semiconductors is based on the pair creation effect by the absorption of light. Each light quantum (photon) that is absorbed by the material and whose energy is greater than the band gap, produces a single pair of charge carrier- an electron and a hole. In that way, the number of electrical carriers of the semiconductor is increased and hence its conductivity, either by increasing the concentration of electrons in the conduction band and/or holes in the valence band. This effect is called photoconductivity.

When the charge carriers separate and move in opposite directions, current passes through the semiconductor. In a photodiode, in order to achieve collecting the photon-induced charge carriers at the electrodes before they have the chance to recombine, a pn or p-i-n diode junction structure is used – hence the term *p-i-n* photodiode. The process of the electron-hole creation due to photoelectric interaction of photon in a semiconductor structure as well as the transport of the carriers due to the introduced by the p-n junction electrical field (explained in the following paragraph) is shown schematically in the figure below.

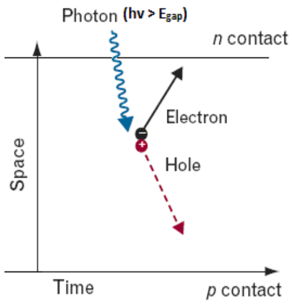


Figure 32: The photoelectric effect.

5.2 Photodiodes: PN junction

An n-type material is a semiconductor that has been doped in such a way to produce an excess of electrons as charge carriers, whereas a p-type material has an excess of holes. A p-n junction is formed by joining a n-type and a p-type semiconductors together in very close contact. The term *junction* refers to the boundary interface where the two regions of the semiconductor meet. This disparity in a pn-junction creates a concentration gradient that causes electrons to diffuse into the p-layer and holes to diffuse into the n-layer.

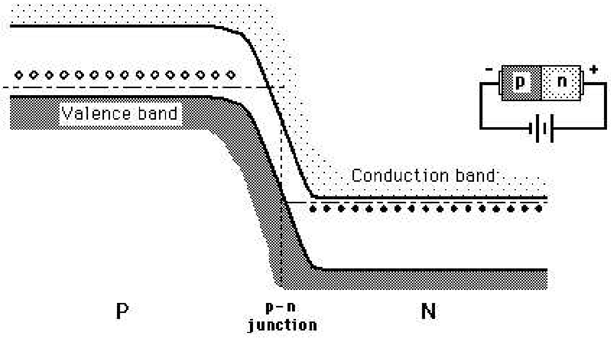


Figure 33: A p-n junction

The application of a reverse voltage to the p-n junction will cause a transient current to flow as both electrons and holes are pulled away. When the potential formed by the widened depletion region equals the applied voltage, the current will cease.

Due to this diffusion of carriers, the regions nearby the p–n interfaces lose their neutrality and become charged by forming a region spanning both sides of the junction. This electrical force causes any charge carriers to be rapidly swept to the appropriate layer and charge carriers cannot reside in that region, so called space charge or depletion region/layer. The space charge region has the same magnitude of charge on both sides of the p–n interfaces, thus it extends farther on the less doped side (the n side in figures).

The electric field created by the space charge region opposes the diffusion process for both electrons and holes. Therefore, when no external bias voltage is applied in a pn-junction at equilibrium is established by two concurrent phenomena: the diffusion process that tends to generate more space charge, and the electric field generated by the space charge that tends to counteract the diffusion. Due to the existence of the electric field, a potential difference is built across the junction, usually called internal or built-in bias/ potential, V_{bi}.

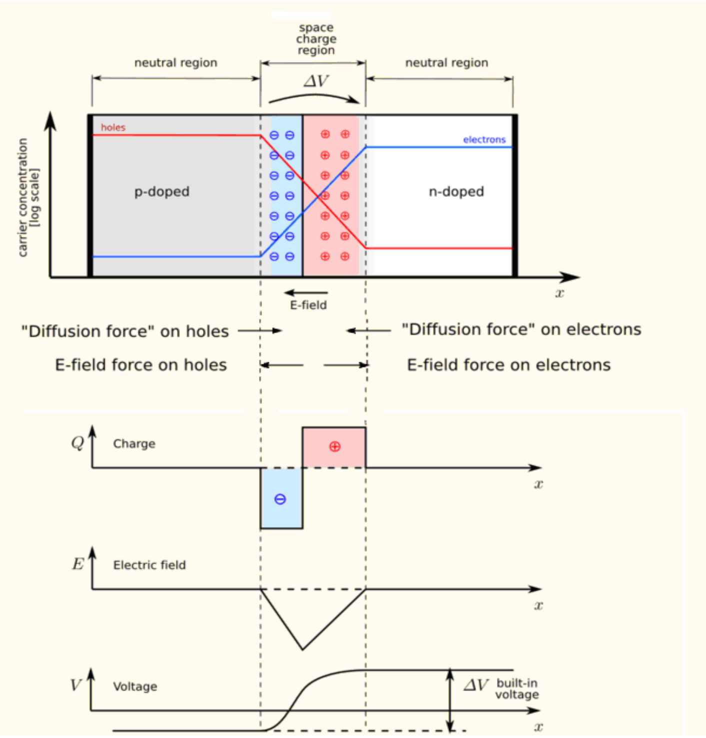


Figure 34 : A p–n junction in thermal equilibrium with zero bias voltage applied. Electrons and holes concentration are reported respectively with blue and red lines. Gray regions are charge neutral. Light red zone is positively charged. Light blue zone is negatively charged. The electric field is shown on the bottom, the electrostatic force on electrons and holes and the direction in which the diffusion tends to move electrons and holes. (from wiki)

5.3 PIN photodiode

A PIN diode has the p and n-type regions separated by an intrinsic region, which is a major absorption layer. This region is made by a lightly doped 'near' intrinsic semiconductor (also called undoped or i-type semiconductor) which is a pure semiconductor without any significant dopant species present. Therefore, the number of charge carriers is determined by the properties of the material itself instead of the amount of impurities. In the intrinsic region of a PIN diode the number of excited electrons is equal to the number of holes (n=p), whereas the p and n type regions are typically heavily doped because they are used for ohmic contacts.

Holes diffuse from the p side a thin layer of the i-region, and similarly do electrons from the n side. There is still a depleted region at the center, but it is reduced from the i layer width. The reason of the existence of the intrinsic reason is that it extends the region where electron-hole pairs can be produced, but still allows charge carriers to drift to anode and cathode.. The diode conducts current once the flooded electrons and holes reach an equilibrium point, where the number of electrons and the number of holes in the intrinsic region are equal. A small reverse bias, known as the "punch-through" voltage, will sweep these charges out and fully deplete the i-region width.

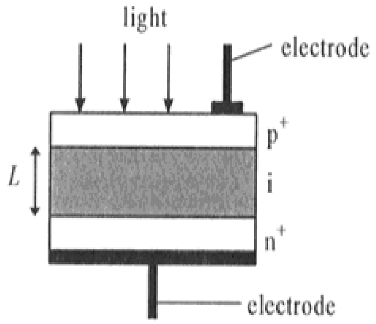


Figure 35: A PIN photodiode

In a generic *p-i-n* photodiode, light enters the device through the thin *p*-type layer. Absorption causes light intensity to drop exponentially with penetration depth. Photons that have been absorbed in the depletion region produce charge carriers that are immediately separated and swept across the junction by the natural internal bias. Charge carriers created outside the depletion region move randomly, many of them eventually entering the depletion region to be swept rapidly across the junction. Some of them will recombine and disappear without ever reaching the depletion region. This movement of charge carriers across the junction produces a small photocurrent, which can be detected at the electrodes.

The development of solid state photodetectors started with PIN photodiodes in Sixties. They were very successfully used in high energy physics experiments (CLEO, L3, BELLE, MAMAR, FERMI) and a large number of other applications like radiation detection and medical imaging. However, due to the absence of internal gain a charge sensitive amplifier is necessary (noise, slow) and the minimal detectable light pulses need to contain several hundreds of photons.

5.4 The classical APD operated in the proportional Mode

The next step in the evolution of photodetectors were the avalanche photodiodes (APD) which have also been used in big experiments (CMS at LHC was the first). Avalanche photodiodes use the same process as PIN photodiodes, but they achieve internal gain using an avalanche multiplication process. The internal gain of the APDs improves the signal-to-noise ratio by internal multiplication and thus a substantial reduction of noise but still about 20 photons are needed for a detectable signal, that is two order of magnitude improvement of the sensitivity of a PIN photodiode.

5.4.1. APD Description

An APD is based on a asymmetric doped p-n junction. Basically, an APD is very similar in design to a silicon p-i-n diode operated at very high reverse bias voltage, slightly below the breakdown voltage. However, the depletion layer in an avalanche photodiode is relatively thin, resulting in a very steep localized electrical field across the narrow junction. The physical mechanism which avalanche gain depends, is the “impact ionization”.

Minimal energy which is required for the impact ionization is called threshold ionization energy. This value is one of the main parameter of the theory of avalanche multiplication in semiconductor materials. To characterize the dynamic of avalanche process is used the impact ionization coefficients , α for electrons and β for the holes. The definition of these parameters is the inverse value of average distance , along the electric field, which is necessary for each type of carrier to produce secondary ionization and create a secondary electron-hole pair.

“Impact ionization” occurs only when the electric field in the depletion region is strong enough. When an electron, either generated by an interaction with a photon or thermally generated, in the conduction band, moves into the avalanche region, the electric field has to be strong enough to accelerate it to the point which it can cause “impact ionization” and liberate another electron. Both of these electrons can be accelerated as well, each creating an avalanche multiplication, which is basically independent of all other avalanches formed from the other electrons associated with the original ionizing event.

This "avalanche" of electrons eventually results in electron multiplication that is analogous to the process occurring in one of the dynodes of a photomultiplier tube. The amplification process adds though more noise to the signal than in case of PMT. The collected charge remains proportional to the number of original electrons. For that reason, that operating mode of an avalanche photodiode is also known as “proportional mode”.

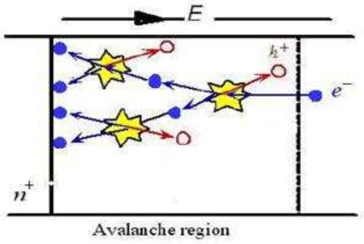


Figure 36: A schematic of the principle of proportional mode of operation of an APD. Incident photons generate charge carriers in the depletion region. These charges are then accelerated to high speeds by an applied electric field and ionize atoms within the avalanche region. Self quenching avalanche process

In figure (above) a self quenching avalanche is shown, .For a low electric field (10^4) impact ionization coefficient of holes is much lower than the one of the electrons and the avalanche process is created practically by one type of carriers – electrons. The avalanche process is one directional and self quenched when carriers reached the border of the depleted area of the semiconductor. Thus, the gain of the multiplication is limited by the thickness of the depleted area.

The process of multiplication results in detector internal gain. Typical gains for an APD are in the range of ten to a few hundred. However, their useful range of gain is limited by the excess noise due to statistical nature (fluctuations) of the multiplication process. That means that it is not possible to separate the first photoelectron peak from the pedestal, a constraining factor for the use of the APDs in single photon counting applications. Single photon counting is possible only at low temperatures (T~77K) and with slow electronics (Photon Detection Effiiciency PDE~20%) (see A.Dorokhovet, al , JournalMod, Opt v51 2004 p.1351). Another requirement for the operation of the APDs at high gain i.e. close below breakdown voltage is very good temperature and bias stability as there is a strong dependence of the gain on the temperature and bias voltage.

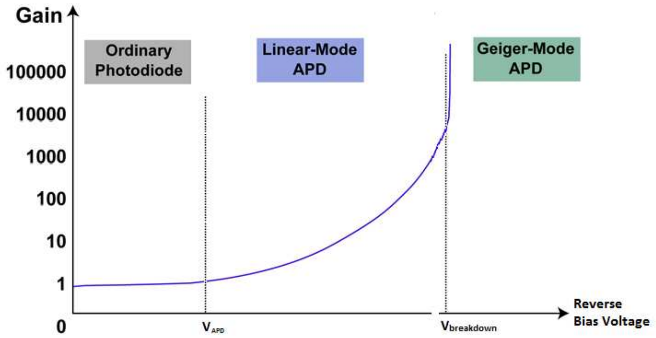
However, the stochastic nature of the avalanche multiplication process in an APD causes also an multiplication in the noise (R.J. McIntyre, IEEE Tr. ED-13 (1972) 164) – the number of charge carriers created per avalanche has a wide distribution . Therefore, the excess noise factor for an APD is defined by the two ionization coefficients for the carries (α for electrons, β for holes) and by the average multiplication coefficient through the formula:

$$F = k \cdot M + (1-k)(2-\frac{1}{M}) \qquad (5.1)$$

where k=β/α (k-factor), M is an average multiplication coefficient, α, β is an ionization coefficient of electrons and holes, respectively.

5.5 Small area APDs operated in the limited Geiger Mode

At the end of the millennium, semiconductor detectors evolved to overcome the limitations in gain and most of the stability problems of the APDs operating in the proportional gain mode. The novel type of photon detectors is small area APDs which can operate in limited Geiger mode by increasing the bias voltage slightly (10%-20%) above breakdown voltage, where the semiconductor junction breaks down and the APD will become a conductor. Geiger-mode operation can increase the modest gain of an APD to a more significant level. They are highly sensitive devices, which have an internal gain comparable to the gain of photomultiplier tubes and a response to single photons.



For a high electric field (10^5-10^6), above a critical value that corresponds to the breakdown voltage of the photodiode, impact ionization coefficients are now coming close to each other and both type of carriers , electrons and holes, contribute to the avalanche process. A self propagating chain reaction occurs, so that carriers rise exponentially with time and the breakdown conditions are reached. The gain of the multiplication is not any more limited by the depletion region thickness and becomes infinite.

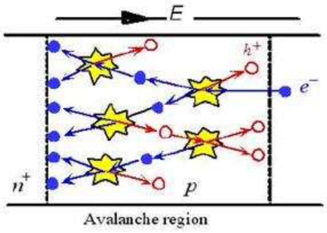


Figure 37: Geiger mode: With a higher electric field, a situation is created, in which one avalanche trigger itself a second avalanche at a different position. Self sustaining avalanche breakdown process

As pointed out before, there are two different working regimes for reversed biased avalanche photodiodes. By operating below the breakdown voltage, an APD works as an amplifier in a linear-proportional mode, where only electrons generate additional electron-hole pairs. In contrast, by operating a few volts above the breakdown voltage , in the so called Geiger mode, a single electron can trigger a diverging avalanche multiplication process where both electrons and holes actively participate in the multiplication process. Due to the diverging nature of the multiplication, any information about the primary signal i.e. the number of photons, respectively photoelectrons, that initiated the breakdown is lost . For that reason GM-APDs are operating as trigger devices, in a binary mode.

5.5.1 Close-up in an GM-APD

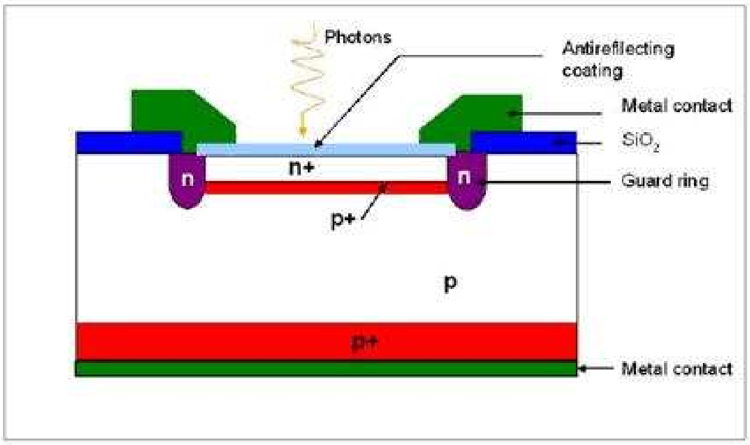


Figure 38: Planar APD structure

An APD consists of three p-type layers of different doping next to a n^+ type layers. The first, respective to the n^+ layer, is a thin p^+ type layer, the second is a thick layer of p type with a very small number of concentration of impurities (almost intrinsic) and the third is a strongly doped layer type p^+ with small resistance. When a reverse bias voltage is applied, the depletion region widens so much that reaches until the second layer of p.

Figure 38 shows a guard ring type n around the APD. This is constructed to keep the electric field stable in the pn junction and to protect the diode from tunnelling effects with adjacent APDs when used as part of a SiPM (described in next paragraph). The active area of the junction is covered with an anti-reflective coating optimized for short wavelength light.

The electric field increases in the thin n^+ layer due to the positive charged impurities (donors) and is maximized in the pn-junction. Then, it is slowly decreased in the first p^+ layer, while in the second the reduction is very small, something expected from the low number of the N dopping in that layer. Finally, the field fades out in the very thin depleted region of the last p^+ layer, Consequently, the conditions for the Geiger mode operation exist only in between the two strongly dopped layers n^+ -p^+ (usually 0.7-0.8 μm) where the electric field is very high (~ $5*10^5$ V/cm).

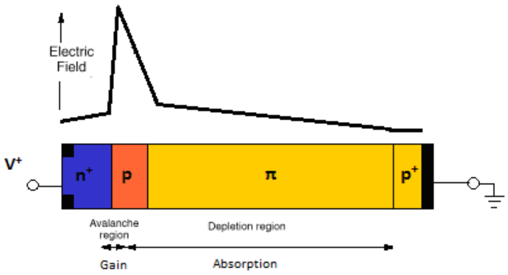


Figure 39: Electric field in a reverse bias diode

The absorption of the visible photons and thus the electron-hole pair creation takes place mostly in the middle p layer, which is the thickest one. There, because of the almost normal electric field, the electrons and the holes are separated and accelerated to opposite directions. The holes migrate towards the third p^+ layer where the anode (metal contact) and the electrons towards the n^+ layer where the cathode (metal contact).

Once the migrating electrons reach the first p^+ layer, they are accelerated moredue to the stronger electric field. In this way the electrons have sufficient kinetic energy to cause more ionization. The new electrons are also capable of new ionization and an avalanche process is provoked.

5.5.2.Quenching mechanism

When an APD is operating in Geiger mode, a single carrier entering the depletion region is enough to initiate avalanche multiplication process and produce a constant self-sustaining current which flows through the pn-junction. This avalanche is also known as the Geiger discharge and was studied 40 years ago by Haits and McIntyre. The initiation could be as result of incoming photon interaction or thermal created carrier inside depleted area. A device that triggers once, however, is not a very useful detector, so a quenching mechanism is necessary to stop the avalanche breakdown process and “reset” the APD and allow the detection of another photon

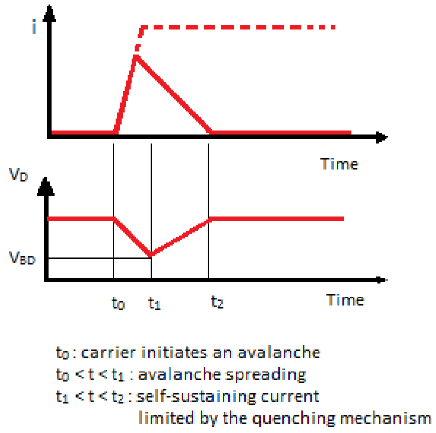


Figure 40: Current and Voltage in a diode biased above V_{bd}.[44]

A simple way to achieve the quenching of an avalanche breakdown is by inserting a high Ohmic resistor in series to the diode. This type of quenching is called passive quenching. The quenching element is acting in the following way: after the initiation of the avalanche breakdown process, the current is rising in the eternal circuit and causes the drop voltage drop on the quenching resistor and accordantly of the voltage applied to the pn-junction . The process of quenching is started when the dropping of the voltage on the quenching resistor brings the voltage applied to the pn- junction to value to lower then breakdown voltage. After quenching, the resistor prevents an instantaneous recharge of the diode capacitance

(C_D) and an instantaneous reset to the initial bias above breakdown. A disadvantage of the passive quenching, that is the slow recharge of the APD.

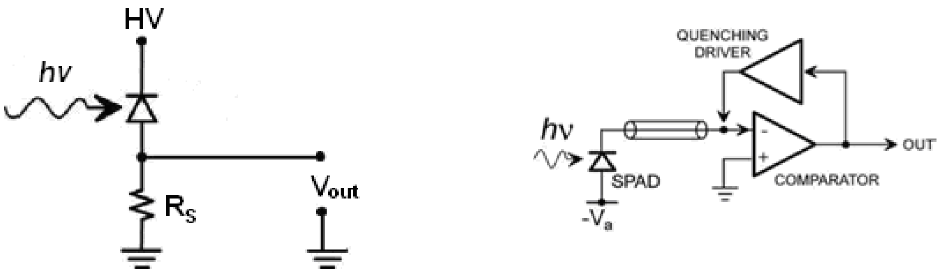


Figure 41: Electrical circuit of passive (left) and active (right) quenching mechanism of an APD

That limitation is overcome by the substitution of the quenching resistor Rs by an electronic circuit. This type of quenching is called active quenching and it uses a dedicated electronic circuit that lowers the bias voltage below the breakdown voltage for a certain period of time until the avalanche breakdown is quenched. In that way the APD is faster recharged than in the case of passive quenching. However, active quenching is more expensive to design, and requires more space.

The Geiger-Mode APD quenched by a quenching resistance can be modelled with an electrical circuit and two probabilities. The electrical circuit consists of the quenching resistor R_Q (>300kΩ), the external applied V_{BIAS} (>breakdown voltage) and the two elements in parallel for the diode. Firstly, the diode is always shown by its capacitance C_D and secondly, by a conducting micro-plasma which is only created during the breakdown of the APD V_{BD} with a series resistance R_S (~1kΩ). The microplasma is shown as a switch on the diagram below and the two probabilities of the operation of a GM-APD are the switch turn on /off probabilities, P_{01} and P_{10}. The P_{01} corresponds to the turn-ON probability of the switch that is the probability that a carrier traversing the high field region triggers an avalanche breakdown. The P_{10} corresponds to the 'turn-OFF probability', i.e. the probability that the number of the carriers in the high field region fluctuates to 0.

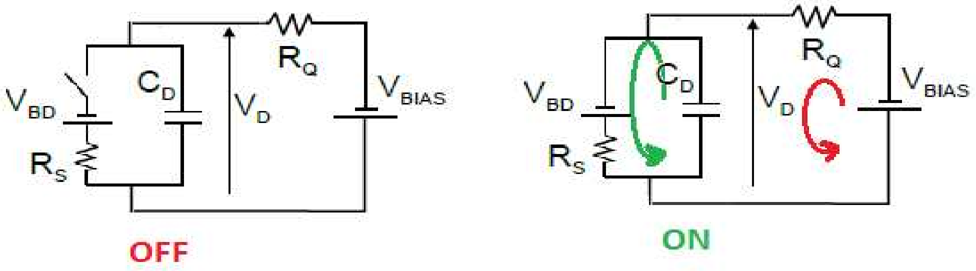


Figure 42: The Geiger-Mode APD quenched by a quenching resistance modelled with an electrical circuit and two probabilities. [44]

Supposing the initial condition is an OFF condition. Any previous avalanche has been quenched, the switch is open, the diode capacitance has been charged until no current is flowing from V_{BD} to V_{BIAS} with a time constant $\tau_{Quenching} = R_Q * C_D$. An avalanche is triggered by a carrier in the high field region with probability P_{01}. The switch now then closes and the diode capacitance C_D discharges to V_{BD} with a time constant $\tau_{Discharge} = R_S * C_D$. At the same time, the external current grows asymptotically to $(V_{BIAS}-V_{BD}) / (R_Q + R_S)$. The process of quenching is following as long as the number of carriers traversing the high field region fluctuates to 0 (P_{10}) until the avalanche is totally quenched and the diode fully recovered. The discharge-and-reset cycle described above is also known as the Geiger mode and is shown in the figures below.

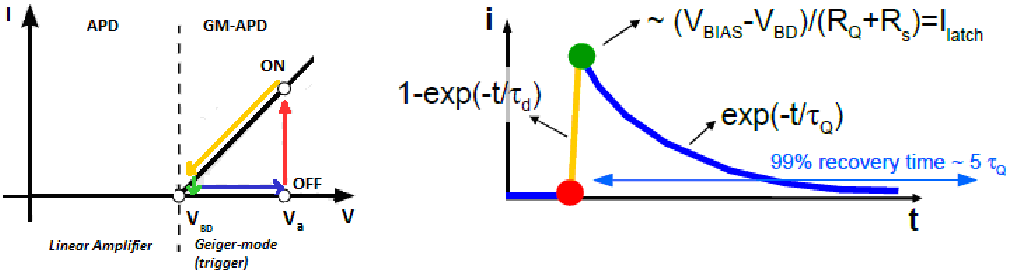


Figure 43: The linear and the Geiger mode of an APD (left). Current output pulse for an APD operating above V_{bd} (right).[44]

APDs operated in the Geiger mode have the advantage of large and well defined output pulses electrons, depending on the overvoltage and diode capacitance per breakdown. The leading edge of the current signal is much faster than trailing edge ($\tau_d = R_S * C_D << \tau_Q = R_Q * C_D$). The charge collected per event is the area of the exponential decay of the current for the recovery time (τ_Q) shown in the figure 2 above.

The limited Geiger mode is only useful for very small area avalanche diodes because besides free electrons being generated by the photo-electric, electron–hole pairs are constantly generated thermally. Moreover, traps are another source of charge carriers. In addition to the recovery time of the diode, these effects set an upper limit to the area of the diode.

Their impressive charge and timing resolution with simple electronics at room temperature make the APDs operating in Geiger mode useful for "single photon counting". Therefore, they are also often referred as Single Photon Avalanche Diodes (SPADs). Another important aspect of the Geiger mode operation of a APD is the significant reduction of statistical fluctuation of the signal. The amplification factor is defined not by the statistics of the avalanche process as in the proportional APD, but only by the pn-junction characteristics and the quenching circuit.

However, a GM-APD has the disadvantage that it acts as a binary device, having the inability to resolve the number of primary photons, photoelectrons. Therefore, such GM-APDs have been commercially produced for approximately the last three decades but have not achieved widespread use. In fact, they are found only in applications that require low rate single photon counting, and where a small size detector is sufficient.

5.6 The multicell APD operated in limited Geiger mode

Since the late1980’s a new photon detector concept was invented in the former Soviet Union by Golovin and Sadygov to overcome the inherent limitation of the binary nature of the GM-APD. The new photon detector was called Silicon Photomultiplier (SiPM) and made use of the advantages of the GM-APD but also allowed to retain over a large dynamic range the number of primary photoelectrons.

5.6.1 SiPM Description

Silicon photomultiplier is a silicon microstructure that consists of:

a densely packed array of typically 100 to 10000 GM-APDs per mm^2 fabricated in the same substrate and connected in parallel through the aluminum layer on the photosensitive side

a miniature integrated quenching resistor for each GM-APD

a common electrode system used as a bus for the common output, where all GM-APD- resistor combinations are connected to .

An equivalent electric schematic is presented in the figure below. A reverse bias voltage is applied to each pn-junction through the common substrate electrode to deplete the n$^+$-p junctions and the induced current is read on the resistor side electrode.

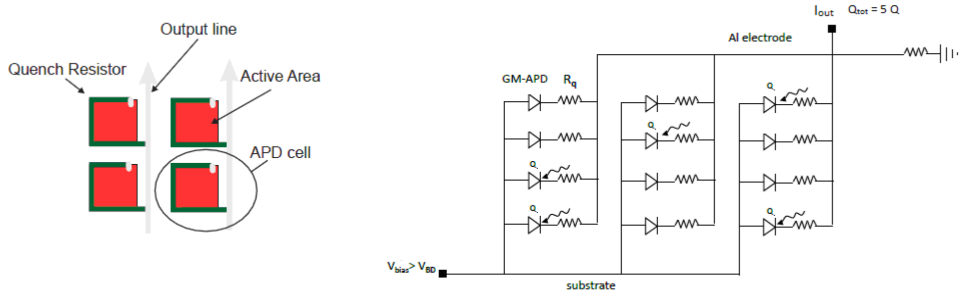


Figure 44: The left panel shows a sketch of 4 cells of a SiPM. Each cell consists of a photo diode and a quenching resistor that is connected in series between the diode and the readout line. The right panel depicts the simplified replacement circuit of a SiPM. [41]

The basic element for a SiPM is the shallow n+-p junction (GM-APD) in series with its polysilicon quenching resistor and from now on it will be called either cell or pixel. Each element is independent and gives the same signal when it is “fired” by a photon. A SiPM acts as an analogue device: the output signal represents the sum of the signals coming from the fired cells and thus the output charge is proportional to the number of the triggered cells (that is the number of incident photons). This allows the counting of single photons or the detection of pulses of multiple photons and makes an SiPM a suitable solid state device for LLL detection and photon counting applications, including HEP experiments, medical imaging and astroparticle physics.

Nonetheless, high-yield production of arrays has not been easily achievable. It is not trivial to achieve the same gain-voltage characteristic for all the APDs of the array. Sometimes that difference can be quite substantial, which means that for the same bias voltage applied in the common substrate the gain of each pixel will be different. That creates the requirement for an individual amplifier circuit for each pixel, something that increases the complexity of the readout circuit when the array consists of a large number of pixels.

Since the late 1990's the development has diversified and many prototype devices exist nowadays. Today more and more institutes and companies are involved in SiPM development and production : CPTA, Hamamatsu Photonics, SensL, INFN, STMicroelectronics, Zecotek, RMD, JINR, MePhI/Pulsar Enteprise, Photonique, Amplification Technologies, FBK-irst, KETEK, MPI Semiconductor Laboratory, Novel Device Laboratory, Excelitas Technologies. Whilst the device is named differently depending on where it has been developed or produced, e.g. Metal Resistive layer Semiconductor (MRS-APD), Multi Photon Pixel Counter (MPPC), Multi Photon Pixel Detector (MPPD) Digital Pixel Photo Diode (DPPD.) etc., in this document the term used to refer to all such devices is Silicon Photomultiplier, SiPM.

5.6.2 SiPM Photon Detection Efficiency (PDE)

The efficiency of detecting photons with a photodetector is known as Photon Detection Efficiency (PDE) and its definition is the percent of the incident photons detected:

$$PDE = \frac{N_{\det ectedphotons}}{N_{incidenphotons}} \tag{5.2}$$

For SiPMs, the PDE typically refers to the overall PDE, in contrast to the case for photomultipliers case, where usually only the quantum efficiency of the photocathode is quoted and the remaining losses are neglected (e.g. the non perfect collection of photoelectrons onto the first dynode of the multiplier). The PDE for a SiPM is a convolution of three factors: geometrical efficiency, which corresponds to the geometrical active area (fill factor), quantum efficiency and the probability of triggering an avalanche breakdown. The first two components are fixed for a particular device but the third varies as a function of the bias voltage applied. These factors are described in further detail the following paragraphs.

$$PDE_{SiPM} = \text{Fill Factor} * \text{QuantumEfficiency} * \text{AvalancheBreakdownProbability}$$

$$PDE_{SiPM} = F * \eta(\lambda) * P_{01}(V).$$

5.6.3.1 Fill Factor

The limiting factor for the SiPM PDE is the effective area. The so-called fill factor “F” (or sometimes referred to as “ε_{geom}”) corresponds to the ratio of the sensitive area, that is the surface area capable of detection of single photons, divided by the total area of the SiPM, including technological border. As expected, the Fill Factor is always less than unity as the physical separation of the GM-APD cells introduces considerable dead space (see pic). Except for the formation of independent cells, dead space is also dedicated in the electrodes and the polysilicon quenching resistors which are implemented upon the active surface of each cell. However, the electrical separation between cells is necessary as well as the metal trench in between them in order to reduce crosstalk effects (triggering avalanche causing an avalanche in an adjacent cell). Existing devices have effective areas from 25% up to 60%, but even highers fill factors (up to 80%) seem feasible in the foreseeable future.

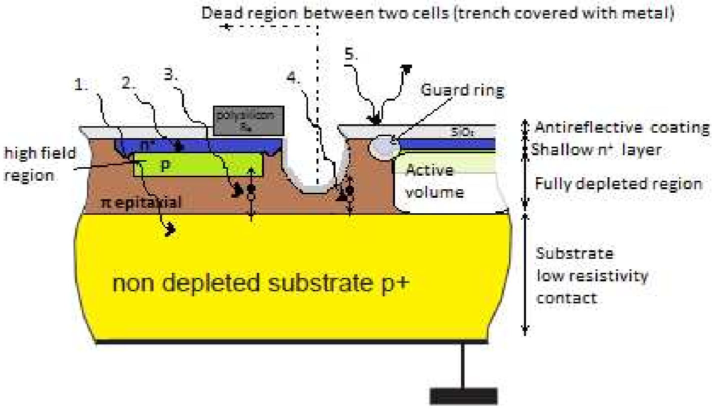


Figure 45: Diagram to show the most important scenarios for the incident photons: 1.Absorption of the photon in the depleted substrate, 2. Absorption in the silicon dioxide or non-depleted implantation below the surface, 3.Absorption in the depleted region and subsequent drift of the photoelectron in the high field region, 4.Absorption in the SiO2 or non-depleted implantation below the surface, 5.Reflection on the surface .

5.6.3.2 Quantum Efficiency

The quantum efficiency is usually defined as the average number of electron-holes pairs created by the conversion of one photon in the depleted area of a semiconductor or the ratio of created electron-hole pairs to the incoming photon flux. For photon energies above the band gap of the semiconductor (1.1.eV in Si) the QE is unity and rises above unity if the energy of the photoelectron is sufficient for impact ionization (>3.6eV in Si). An alternative definition of quantum efficiency is as the probability of a photon to generate a carrier that will reach the high field region.

There are two factors that contribute to the quantum efficiency of a SiPM. Firstly the external quantum efficiency which depends on the losses on the entrance window. Secondly the internal quantum efficiency which is defined as the probability of a photon inside to generate an electron-hole pair in the active layer.

The external quantum efficiency is also often called transmittance of the entrance window where there may be losses due to reflection and absorption (1 and 5 scenarios). In real detecting structures, part of the photon flux is reflected in the border of air/sensitive area and another part is absorbed in the non-depleted area of the structure. These losses could be efficiently minimized by proper engineering such as the use of optically pure materials and an anti-reflective coating (ARC) (see pic).

The internal quantum efficiency is a main factor of the overall QE of a semiconductor detector and it is mostly defined by the characteristics of the absorption process of photons in the depleted area of the structure. A photon flux of intensity $I(\lambda,z)$ will be absorbed in silicon according to the Beer-Lambert law creating an equivalent number of electron-hole pairs:

$$I(\lambda,z) = I(\lambda) * \exp(\alpha(\lambda)*z)$$

where $I(\lambda)$: the initial photon flux,

$I(\lambda,z)$: the photon flux on the distance z from the SiPM surface,

$\alpha(\lambda)$: the optical absorption coefficient

z: the penetrated thickness in Silicon.

This process, defined as photon attenuation, is a fundamental process of all silicon detectors i.e. photons entering a silicon layer travel a characteristic distance before giving up their energy to create a photoelectron. In addition, optical absorption coefficient is strong function of the wavelength (or energy) of the incident photons for every material. The figure below shows the absorption coefficient as function of wavelength of the photons in silicon.

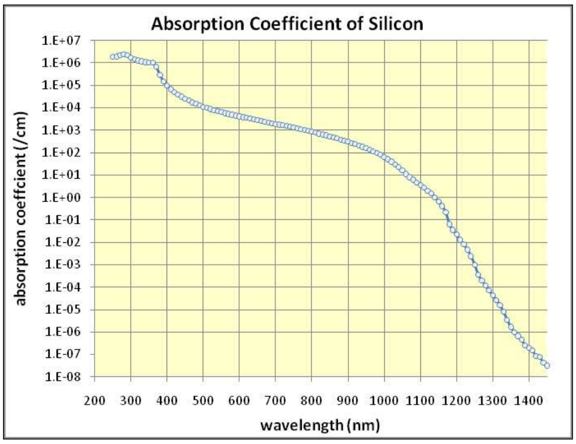


Figure 46: Absorption Coefficient of Silicon Vs Wavelength

Generally, the quantum efficiency of SiPM $\eta(\lambda)$ could be calculated as :

$$\eta(\lambda) = (1 - R)(1- \exp(-\alpha(\lambda)*L_d)$$

where R :reflection Frennel coefficient,

$\alpha(\lambda)$: optical absorption coefficient

L_d: thickness of the depleted area of a cell.

Photons with short wavelengths (<400nm) are mostly absorbed just beneath the silicon dioxide surface within less than 100nm (scenario 2). If the absorption takes place in the highly doped top implantation layer below the surface, the generated electron-hole pair will probably be lost due to the very short recombination times. Only the depleted region is fully active to efficiently photo-generate because of high recombination probability in the un-depleted regions. The very shallow $p^+(n^+)$ top layer should be relatively thin, especially for blue, respectively UV sensitive photon detectors, and the thickness of the sensitive layer (depleted area) L_d should agree to the condition : $\alpha(\lambda)*L_d >>1$. If, on the other hand, the photon energy is low, the photon penetrates deeply into the silicon and is mostly absorbed in the non-depleted bulk or traverses the detector without interaction. Therefore, red, respectively IR sensitive sensors need thick depletion layers.

5.6.3.3 Avalanche Breakdown Probability $P_{01}(V)$

The avalanche breakdown probability is the probability that a carrier triggers an avalanche breakdown, resulting in discharge of the whole pn-junction. This depends very much on the electric field strength in the junction and on the type of the charge carrier (electron/hole) that is entering the high field region. A saturation of the P_{01} (~1) is observed with increasing voltages applied. There is also an increase of the P_{01} in lower temperatures because of increased impact ionization effect.

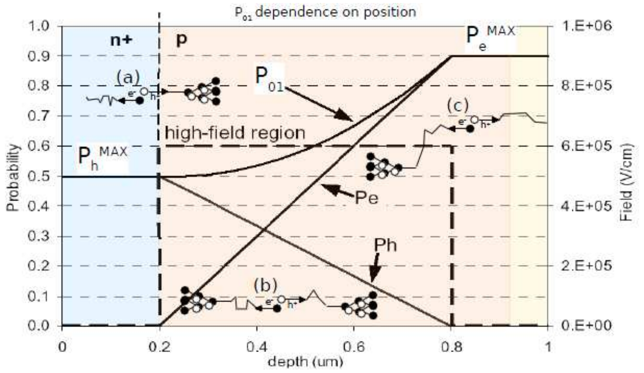


Figure 47: Example with constant high-field: (a) only holes may trigger the avalanche (b) both electrons and holes may trigger (but in a fraction of the high-field region) (c) only electrons may trigger [50]

5.6.4 Time performance

The time performance of the SiPM is defined by two parameters: the rise time of the avalanche breakdown signal and the recovery time, which is defined by the reconstruction of the pn-junction state after quenching the avalanche breakdown process and recharging through the quenching resistor. The time performance of the SiPM is more easily explained after taking into consideration the equivalent electronic circuit (where a GM-APD is now part of an array of GM-APDs and more elements are added in the circuit in comparison to the one describing the single GM-APD):

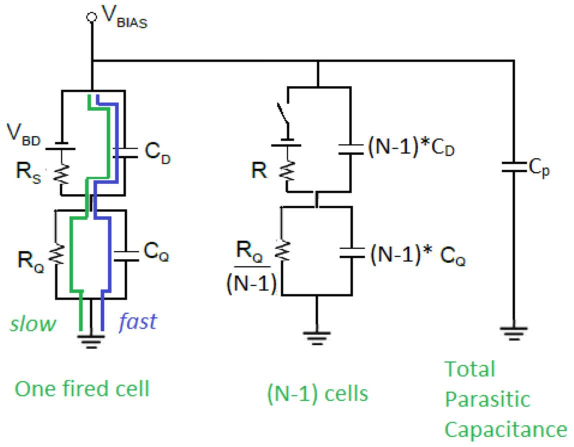


Figure 48: Except for the already known elements of the circuit, there are also two new parasitic capacitances Cq and Cp : the first one is in parallel to the Rq polysilicon quenching resistor lies on top of the junction area and forms a direct capacitive coupling between the resistor itself and the diode(Cq<Cd) and the second one represents the total parasitic capacitance of the whole device including the Al-lines which connect all cells in parallel capacitance Cd and series resistance Rs , which determine the response function of each GM- APD. [51]

5.6.4.1. Rising Time & Timing Resolution

The signal rise time is defined by the time creating the avalanche breakdown process and characterized by the drift time of carriers under the high electric field.

The constant fraction Q of the charge delivered during the avalanche of a fired cell is instantly collected on all of the capacitors of the circuit $C_{tot} = C_c + C_d + C_q$.

Generally, the photon detection probability is less than unity and the transmitted optical pulse has finite time duration. If photons arrive simultaneously entering different pixels at the same time, the output pulse from each pixel will not be necessarily the same time, so that a jitter occurs. The statistical variation of the time of the detection of a photon along a pulse contributes to the jitter and characterizes the rise time of the SiPM.

When photons enter the SiPM with smaller time difference than that jitter , then that difference is not detectable and therefore, the time resolution of the detector is the minimum time difference that can be detected by the pixels and it is defined as FWHM of the distribution of the time jitter. To reduce the variation the output pulses between the pixels, the trace resistance should be reduced. There is also an improvement of the timing jitter at low T due to higher mobility.

5.6.4.2 Recovery Time

The recovery time is mainly defined by the recharge process of the GM-APD and thus is mainly controlled by the passive quenching element of the circuit. The recovery time could be estimated from the values of R*C of the quenching resistor R_Q and the equivalent cell capacitance $C_{eq} = C_d + C_q$.

The circuit has two time constants for the output of a fired pixel: a fast component (~few 10ps), where the avalanche current is reproduced mainly at the output by the parasitic capacitor of C_q.). Therefore, the fast component is determined by the resistance of the silicon in the breakdown channel R_S and the capacitance C_{eq}, $\tau_d = R_S * C_{eq}$. The C_q capacitance is responsible for the peak in the signal shape -often called spike effect. In order to have a more obvious peak in the signal and thus better signal shape, the C_q should be increased (capacitance between a pixel and readout trace). That peak is followed by a slow tail due to the recharging of the cell. This represents the slow component of the recovery time (99% recovery time ~100ns) and it is slow because of the big value of the quenching resistor R_q in the $\tau_q = R_q * C_{tot.}$ time constant.

The recovery time of a cell should not be confused with the recovery time of the device. The situation is different in case of intense light, flashes or high rate applications when the average time between consecutive events becomes comparable to the recovery time. The effect of the parasitic capacitance of the device should not be neglected in such cases of saturation.

5.6.5 Gain

The absolute gain of a photodetector is the number of charges which have been created at the output of the device when one photon has hit it. Hence, the SiPM gain is determined by the charge (Q) that is released from a GM-APD after the breakdown $Q_{out} = V_{over} * C_{diode}$, as

$$Gain = \frac{Q}{e} = \frac{V_{over} \cdot C_{diode}}{|e|} \quad .$$

The time integration over the output pulse allows the measurement of the gain as explained in the paragraph of the GM-APD gain.

The gain of a SiPM is strongly depending on the temperature and on the applied overvoltage Vbias-Vbd. In addition, the breakdown voltage of a photodiode also depends on the temperature. Therefore the dependence of the gain on the temperature is double.

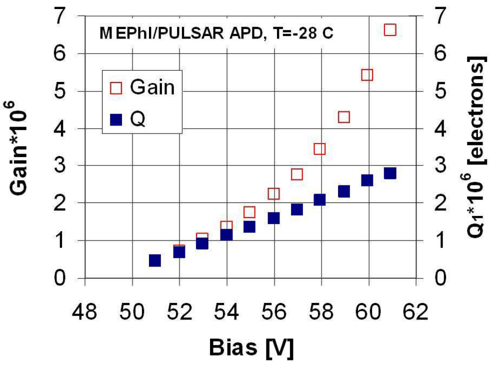


Figure 49: Gain and Single Pixel Charge of a SiPM [45]

5.6.5.2 Single photon Spectrum

The value of the gain (also called multiplication factor) could be precisely calculated from single photon spectra where the signal represented is the charge generated by the avalanche breakdown process and the peaks correspond to the number of detected photons. The resolution of the SiPM is enough to distinguish the signals with discrete numbers of photons, which shows the quantum nature of the light.

The resolution of SiPM allows the precise analysis of the detecting photon flux up to single photon. First peak corresponds to amplitude noise of electronic channel (pedestal), second peak corresponds to the amplitude of detecting one single photon, third peak corresponds to the amplitude of detecting two photon in the same time etc, The statistical behavior of the photon spectrum is seen in the figure below. In SiPM, it is possible to get the absolute calibration of the gain using the position of the single photon peak in the spectrum, because this position is exactly correspondent of creating on electron-hole pair in a cell.

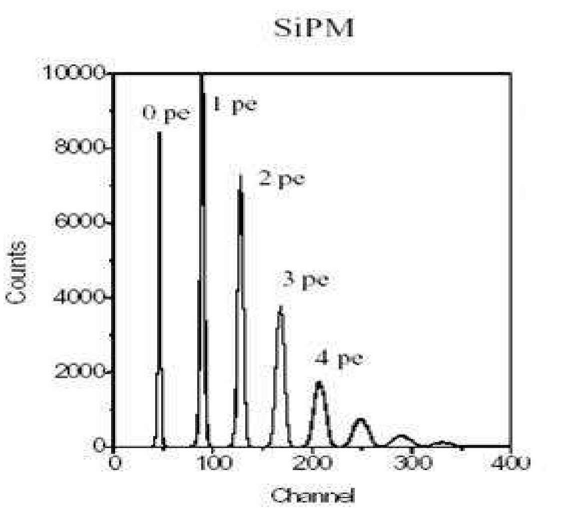


Figure 50: Single photoelectron spectrum recorded with a SiPM*[52]

**The statistics followed in photon counting (signal-processing technique that converts the output signal generated by a single photon into a digital pulse that is counted) are Poisson statistics. That means the standard deviation of a number of counts is equal to the square root of the number of counts. Nevertheless, the signal-to-noise ratio is the standard deviation of signal counts so it is simply the square root of the number of signal counts.*

5.6.6 Dynamic Range

The term Dynamic Range refers to the maximum number of photons that can be detected by a SiPM before it passes in a situation of saturation. The detection of photons by a SiPM is a a statistical process based on the probability of detecting randomly distributed photons by the limited number of sensitive pixels. Therefore, the photon detection efficiency and the total number of pixels determined the dynamic range of a SiPM. The number of detected N_{det} as function of the number of incident photons Ninc is approximated by the following expression:

$$Ndet = Navail\,(1 - \exp(-PDE*Ninc/\ Navail))$$

where Navailable : number of pixels of a SiPM
Nincident: number of incident photons
Ndetected: number of detected photons
PDE: photon detection efficiency of the SiPM..

The output signal of an array of GM-APDs is proportional to the number off fired cells, provided that the number of the photons in a pulse (Nincident) times the photon detection efficiency of the detector, which represents the number of photons that will probably be detected by the SiPM) is significant smaller than the number of the available cells. In the following figure the threoretical curbe of the dynamic range of a SiPM for two different values of the PDE as a function of the number of incident photons for a silicon photomultiplier is shown.

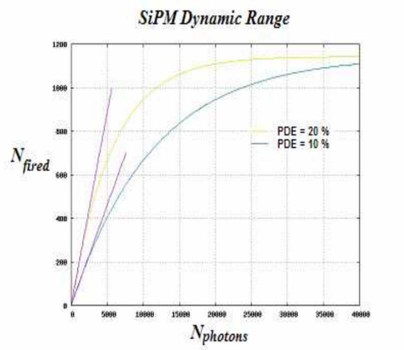


Figure 51: SiPM Dynamic Range (SensL ,SPMMini with 1144 pixels).

From the figure 51, it can be seen that the silicon photomultiplier response is linear when the number of photons is much less than the total number of pixels and becomes nonlinear when they are activated more than a quarter of the total number of pixels. This is the beginning of the saturation of the SiPM response.

The effective dynamic range depends also on the recovery time and the time scale of signal burst. When the pulses of photons have less duration than the recovery time of a pixel from the avalanche breakdown process, then the optical pulse can be considered simultaneous for all the pixels, as none of them can detect more than one photons. For pulses of larger duration the SiPM pixels have the time to recover and return to their operating voltage and as a result the dynamic range of the detector is increased.

The figure also shows how the SiPM dynamic range and linearity can be extended to handle higher photon fluxes by lowering the operating voltage and hence reducing the PDE of the SiPM. The statistical behaviour of the linearity and dynamic range curves give the possibility to calibrate this curve to improve the characteristic of the SiPM. Nevertheless, the main way to improve the linearity and increase the dynamic range of SiPMs is to increase the number of pixels.

5.6.7 Noise

5.6.7.1 Dark Count Rate (DCR)

One of the main factors limiting the performance and the size of the silicon photomultiplier is the high dark count rate. The dark count rate is defined as the average frequency of the dark pulses, which are signals with amplitude equivalent of single photon signal (or sometimes more due to crosstalk effect) produced not by photon-generated electron-hole pair but from thermally-generated one. The thermally generated carriers in the depleted area can also initiate a avalanche breakdown and result in a current pulse that is indistinguishable from a pulse produced by the detection of a single photon. This is particularly important for room temperature operation where the dark rate increases linearly versus the applied voltage. The dark count rate is thermally activated and hence it is strongly depending on the temperature. Although increasing the reverse voltage improves photon detection efficiency, it also increases the dark count. The dark count can be reduced by lowering the temperature.

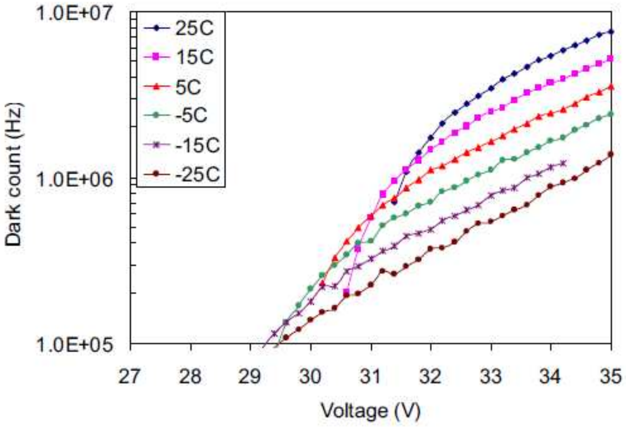


Figure 52: Dark Count Rate as a function of the applied voltage at different temperatures: -25, -15, -5, 5, 15, 25C) [49]

The amplitude of the dark rate pulses is equivalent of the single photon signal amplitude, that for many application deal with tens-hundreds photons it could be nneglected. Nevertheless, for LLL and single photon counting applications where the single photon detection is very important and can cause detection errors. In that cases the dark rate cannot be subtracted and constitutes a noises source that determines the minimum detectable signal.

5.6.7.1.1 Dark Current

Usually conventional photodetectors are characterized in terms of noise by the dark current. In case of the SiPM, the very high gain is practically neglecting the contribution of the dark current or leakage currents in the silicon structure to the output signal, because the leakage current carriers are not affected by the amplification process. Significant for the noise consideration becomes the multiplication process.

As long as the applied bias voltage is lower than the breakdown voltage, rhe current drawn by the device is mainly due to generation in the surface region around a diode and is linear to the V_{bias}. When the bias voltage is up to few volts above the breakdown voltage (Geiger mode) the current due to dark events is :

$$\mathrm{I}_{dark-post-breakdown} = q \cdot G \cdot N_{dcr} \approx q \cdot V_{bias} \cdot V_{bias} \tag{5.3}$$

and the dependence from the applied voltage becomes quadratic. The dark current of a SiPM is given by the formula:

$$I_d = I_s + M \cdot I_b \tag{5.4}$$

where I_s and I_b are the surface leakage current and the bulk generation current respectively and M the multiplication factor of carriers.

The multiplication factor can be approximated by the formula:

$$M = \frac{1}{\left(1 - \left(\frac{V_{bias}}{V_{bd}}\right)\right)^{n}} \tag{5.5}$$

where n depends on the type of the material of the pn junction and has typical values from 3 to 6.

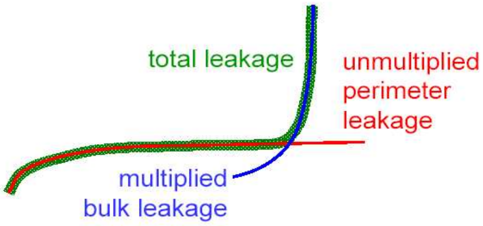


Figure 53: The dark current components of a SiPM

5.6.7.2 Optical Crosstalk

A well known process in semiconductors is the avalanche multiplication process. This phenomenon becomes very critical for SiPM structures due to the very tiny geometrical pattern. The hot carrier luminescence gives rise to an effect called optical crosstalk. Due to the optical crosstalk effect, the dark count signals often have more than one photoequivalent signal amplitude.

Optical crosstalk occurs when photons created by an avalanche during the avalanche breakdown process can propagate unhampered within the device and be absorbed by the sensitive volume of adjacent cells, thus triggering additional breakdowns. The ratio of the number of secondary created photons, detected in a SiPM to the number of detected incoming photons is determinate as the optical crosstalk of a SIPM. Optical crosstalk can also occur when luminescence photons trigger a neighboring cell if the conversion of the photon takes place in the non-depleted detector volume. A third way that it may occur optical crosstalk is through reflections.

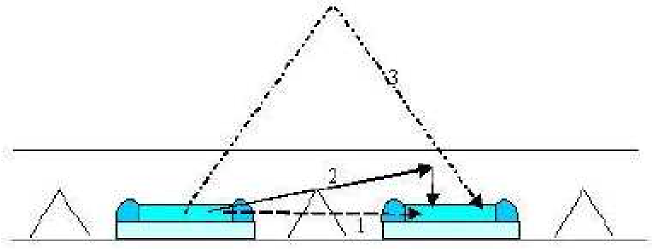


Figure 54: Three different scenarios of the optical cross-talk effect 1. Direct cross-talk 2. Inside the depletion layer 3.Through reflection [52]

Optical crosstalk is strongly dependent on the distance between the neighbouring cells due to the total geometric cross section for the interaction between two pixels. In addition, the crosstalk probability increases linearly with over-bias voltage as the product of optical generation inside a triggered cell. However , the optical crosstalk can be significantly reduced by proper engineering such as optical isolation of the pixels by etching a trench between and filling it with an optically opaque material. Nevertheless, this separation results also in a much lower fill factor and hence reduces the PDE of the SiPM.

5.6.7.3 Afterpulse Effect

Another source of noise in an SiPM is the afterpulsing processes which are breakdown processes that occur due to release of carriers captured by traps (crystal defects) after a certain delay time of a previewed breakdown avalanche process. These spurious pulses are following the signal and may cause detection errors. Afterpulsing is described as a probability of having a afterpulse effect after a given pulse and it is a function of the mean lifetime if the carriers in the traps.

The probability of afterpulsing effect is given by the formula below and has quadratic dependence on the overvoltage applied on the SiPM:

$$P_{afterpulse}(t) = P_c \cdot \frac{\exp\left(-{}^{t}\!/_{\tau}\right)}{\tau} \cdot P_{01} \approx \Delta V^2 \qquad (5.6)$$

where Pc : trap capture probability, proportional to the number of N of traps and to the number of carrier flux(current) during the avalanche which is proportional to the ΔV overvoltage, P_{01} : the avalanche breakdown probability which depends on the recovery time and is proportional to the overvoltage ΔV.

Afterpulsing is strongly dependent on the trap properties, the type of carrier trapped and the temperature. The lower the temperature, the higher the probability that carriers may be trapped by crystal defects, so after pulses will increase.

5.7 Conclusion

SiPM consists a new era in photon detection technology which has been developed and suitable for many critical applications such as LLL application , photon counting , medical imaging, astroparticle and nuclear physics and military applications. SiPM takes advantage of the Geiger mode of operation of APDs without losing information on the intensity of the light. SiPM have become widely accepted as a promising photon detector which can be coupled with scintillators, wavelength shifters or with direct radiation such as Cherenkov radiators. The current stage of development is sufficient for some applications but still far from others and the improvement of their performance is a major challenge in the future. Their main advantages and disadvantages are presented in the next page.

SiPM ADVANTAGES

- Low sensitivity against temperature and bias fluctuations compared to classical APDs
- Low sensitivity against pickup because of high internal gain (>10^5)
- Relaxed requirements on the preamplifier compared to that needed for classical APDs
- Very low response to passing ionizing particles (low "nuclear counter effect", e.g. even a heavy ion would only produce a signal equivalent to one photon)
- Ultra compactness
- Mechanical robustness
- Cheap
- Insensitiveness to very strong magnetic fields
- Radiation Hardness
- No aging over the years
- High internal gain (10^5 -10^6-10^7)
- Low operating voltage (<100V)
- Single photon response
- No damage from accidental and prolonged light exposure
- Low dark count rate (<1MHz/mm2)
- Potential of high Photon Detection Efficiency (PDE = QE (λ)*Pa(V)*GE)
- Superior time resolution (<<1ns)
- Low intrinsic power consumption (50µW per square millimeter sensor area)

SiPM DISADVANTAGES

- High intrinsic single dark counting rate
- Optical crosstalk
- Limited sensor areas

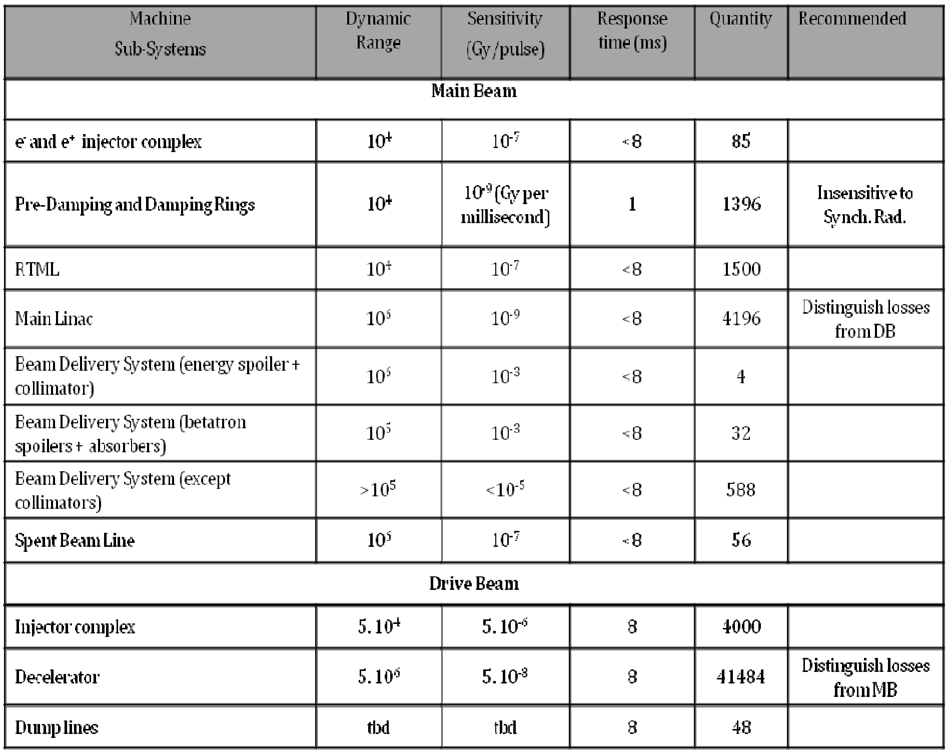
CHAPTER 6. CHERENKOV RADIATOR & CHARACTERIZATION OF A SiPM

6.0 Introduction

An overview of the CLIC study and key parameters are described in Chapter 1. CLIC is a large facility with many subsystems. Table 5 lists the requirements of BLMs for each machine subsystem, where the dynamic range, sensitivity and quantity are based on the signal produced in ionization chambers at loss limits specified in . [10][11] Ionization chambers used at LHC meet the requirements listed below (and specified as a suitable baseline technology choice except for the damping rings where a Cherenkov radiator coupled to a PMT is proposed). The two subsystems that are identified as requiring more careful consideration in the post CDR phase are the two beam modules and the damping rings :

The "two beam modules" contain the main beam (9-1500 GeV) and the RF providing drive beam decelerators (2.4-0.24 GeV). There are approximately 40 km of "2 beam modules" accounting for approximately 45000 of the BLMs in table 5. It is therefore desirable to find an alternative cost effective solution for beam loss monitoring.

In the damping and pre-damping rings, a detector insensitive to synchrotron radiation (such as a Cherenkov radiator) is required. Beam loss monitoring is important due to tighter timing requirements for protection of the superconducting damping wigglers.



Machine Sub-Systems	Dynamic Range	Sensitivity (Gy /pulse)	Response time (ms)	Quantity	Recommended
Main Beam					
e^- and e^+ injector complex	10^4	10^{-7}	<8	85	
Pre-Damping and Damping Rings	10^4	10^{-9} (Gy per millisecond)	1	1396	Insensitive to Synch. Rad.
RTML	10^4	10^{-7}	<8	1500	
Main Linac	10^6	10^{-9}	<8	4196	Distinguish losses from DB
Beam Delivery System (energy spoiler + collimator)	10^6	10^{-3}	<8	4	
Beam Delivery System (betatron spoilers + absorbers)	10^5	10^{-3}	<8	32	
Beam Delivery System (except collimators)	$>10^5$	$<10^{-5}$	<8	588	
Spent Beam Line	10^5	10^{-7}	<8	56	
Drive Beam					
Injector complex	5.10^4	5.10^{-6}	8	4000	
Decelerator	5.10^6	5.10^{-3}	8	41484	Distinguish losses from MB
Dump lines	tbd	tbd	8	48	

Table 5: Requirements for BLMs in the CLIC complex [CLIC CDR]

Long cherenkov fibers are currently under study as an alterbitvie technology choice for the CLIC Two beam modules where an SiPM is proposed for light readout, (ref) .SiPMs coupled to fibers will therefore be tested at CTF3 in the post CDR phase. [31]

Therefore, it is of interest to understand and introduce techniques used to fully characterise SiPM detectors. As a starting point for investigations, two quartz bars coupled to SiPMs (described in the following section) were borrowed from Fermilab (M.Albrow).

The first part of this work reports on the electrical (static and dynamic) characteristics of the performance of a SiPM, the experimental setups that were used. The second part describes the preliminary results from the CLEX area (CTF3) with the combination of the Cherenkov radiator and the SiPM.

Figure 55: Picture of the box that includes the SiPM and the quartz bar Cherenkov radiator.

6.1 Detector Description

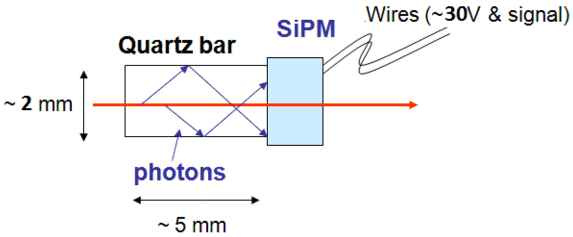


Figure 56: Schematic of the detector

The Cherenkov radiator is a quartz bar of 2mmx2mm surface and 5mm length. It is directly coupled to a SiPM presents important advantages such as short rise time, high internal gain, lower power consumption, compactness, insensitivity to magnetic fields and single photon counting resolution.

The prototype silicon photomultiplier that was tested was fabricated by STMicroelectronics in Catania, Italy in 2006 and it is not commercially available yet. It has an area of 3.5x3.5mm2 and it is composed of 60x60 = 3600 connected in parallel. Each pixel is active about over 45um x 45um and it is composed by a swallow n^+-p junction (GM-APD) in series with its integrated polysilicon quenching resistance (Rquench). All microcells are connected in parallel through an aluminum layer on the photo-sensitive side and the substrate on the other side.

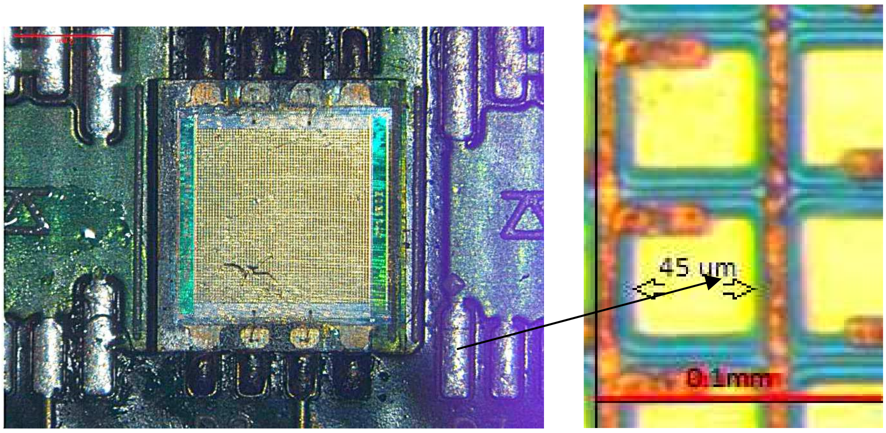


Figure 57:A photograph of the SiPM prototype from STMicroelectronics (left) and a zoom in the pixel area (right).

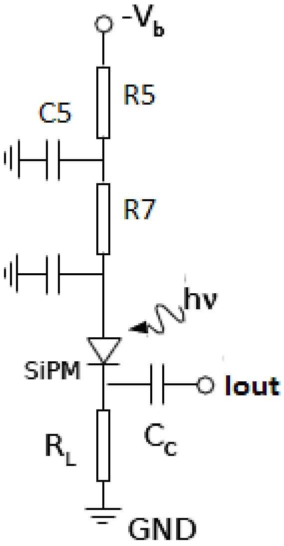


Figure 58: A simple equivalent circuit of the SiPM is shown above. A reverse bias voltage (V_{bias}) is applied to each junction through the common substrate electrode to deplete the n+–p junctions and the induced current is read on the resistor side electrode.The output signal is ac-coupled by the Cc capacitor. Two low-pass filters can also be recognized in the circuit.

The data related to the fill factor is unavailable. However form figure (a zoomed black and white image of part of the SiPM array), the fill factor is estimated to be approximately 40 %. The concept of estimating the fill factor is: if the value of the colour is below the 50% of the max value (dark) then the pixel is considered as dead area, if the value is above 50%of the maximum (light) then the area is considered as active.

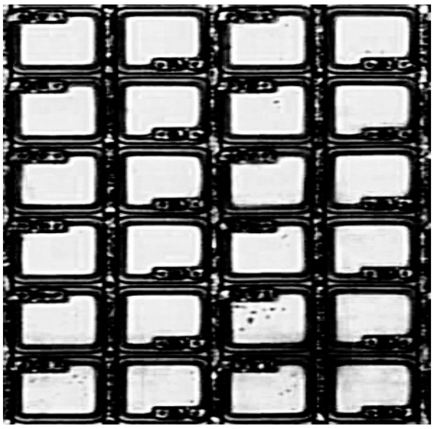


Figure 59: Black and white zoom in the SiPM in order to have a preliminary estimation of the fill factor.

6.2 Measurement set-ups

The characterization of the SiPM consists of the static test for the I-V curve measurement and the dynamic test. A black light tight box was used in order to protect the device from external light. All tests have been performed at room temperature.

6.2.1 Electrical measurement set-up for Static Characteristics

The electrical measurement set-up used for the static tests is illustrated in figure 60. It consists of the connection of the SiPM with a Sodilec DC power supply and a KEITHLEY 2002 multimeter used in a ammeter mode in order to measure the current drawn by the device . All the connections were through coaxial 50Ω cables.

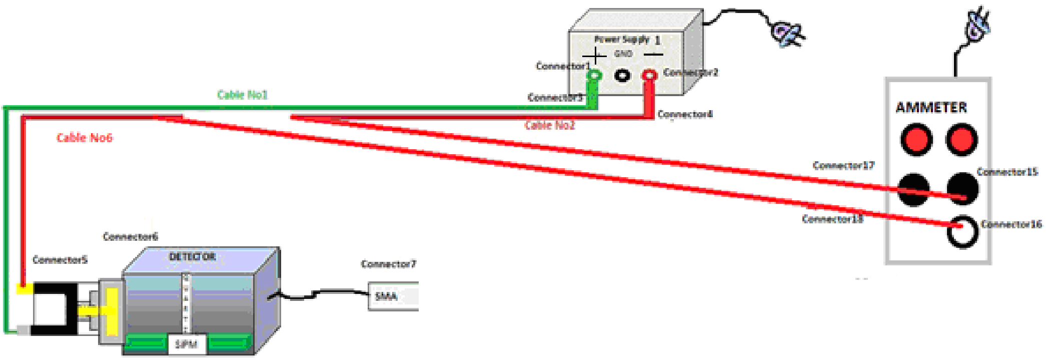


Figure 60: The electrical measurement set-up used for static tests.

6.2.2 Electrical and optical measurement set-up for Dynamic Characteristics

The electrical measurement set-up used for dynamic tests has two configurations. The first one consists of the connection of the SiPM to a single channel wide-band commercial DHCPA transmpidence amplifier read-out by a Tektronix TDS 7254B (2.5GHz, 20 GS/s) digital oscilloscope. The gain in the amplifier during all tests was fixed at 10^6 (I/V) with bandwidth ***. The experiment was focused on signal shape issues, and for that reasons, green LED giving a light pulse to the SiPM connected to a pulse generator Fluke Precision PM5786-001 was used. The second dynamic set-up includes the connection of the SiPM to a counter Agilent Universal Counter 53131A (225MHz) in order to count the frequency of the dark count rate.

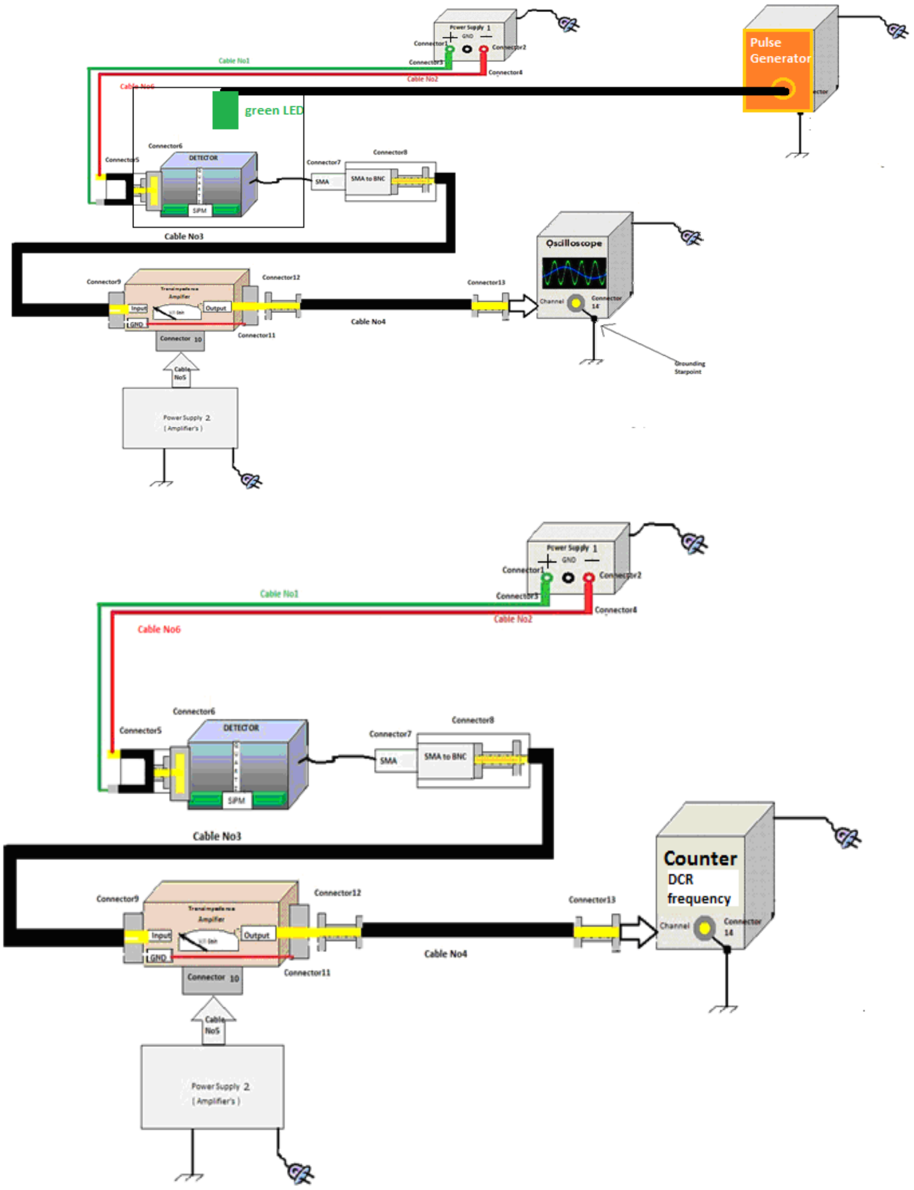


Figure 61: The two configurations of the electrical measurement set-up used for dynamic tests.

6.3.4 Characteristics of the prototype SiPM

6.3.4.1 Static characteristics

The static characteristics (reverse and forward IV plots) of the 16 SiPM's of the matrix are presented in figure (). From the reverse and forward current-voltage (I-V) characteristics, the values of the breakdown voltage V_{bd} and the Rquench of the device.

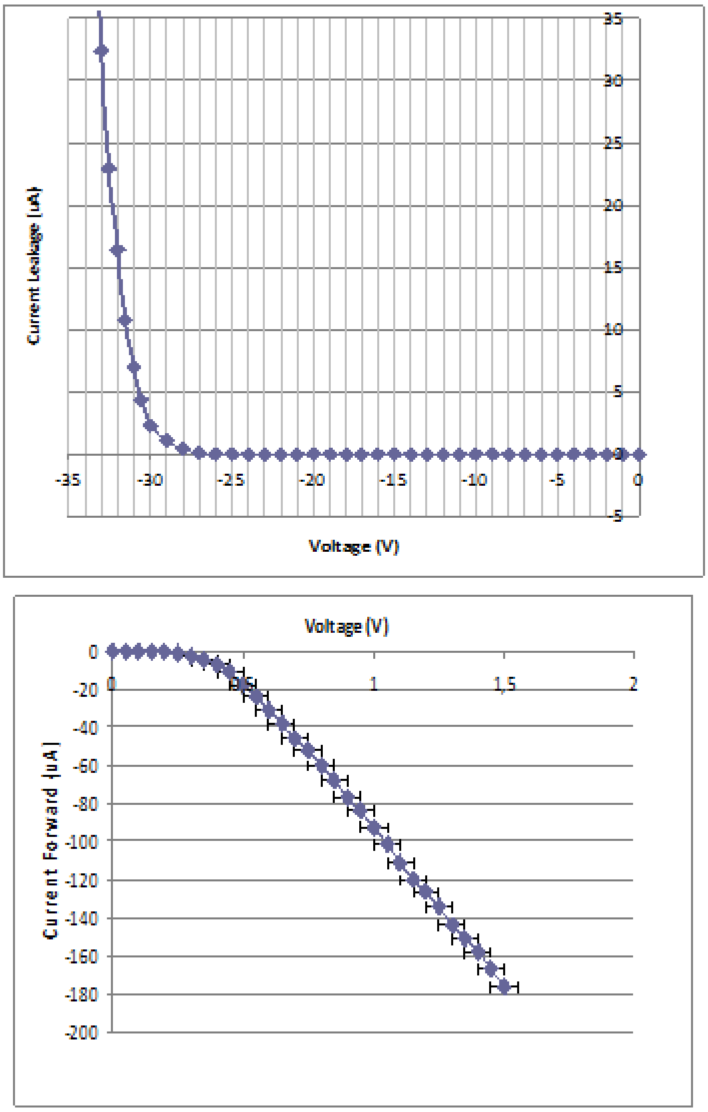


Figure 62 : a) Reverse and b) forward IV characteristics of the SiPM.

From the reverse IV plots, a mean value of V_{bd} ~29V has been determined. From the forward IV plots, a mean value of $R_{quenching}$ SiPM ~ 125 kΩ has been found over all the device. Considering that this represents the equivalent resistance of 3600 pixels connected in parallel $R_{quenchingSiPM} = R_{quenching\ Cell}$ / NumberOfCells, a mean value of R_{qpixel} ~450GOhm has been calculated.

6.3.4.2 Dynamic characteristics

The analysis of any SiPM of the matrix in dynamic conditions, revealed different types of signals as shown in figure below:

- signals coming from single pixels (e.g. primary pulses); they are generated by thermal carriers and represent the main source of the SiPM noise;
- signals following the primary pulses at short interval of time (e.g. afterpulses);
- signals coming from one or more neighbours pixels overlaid with the primary pulses (e.g. optical cross-talk.

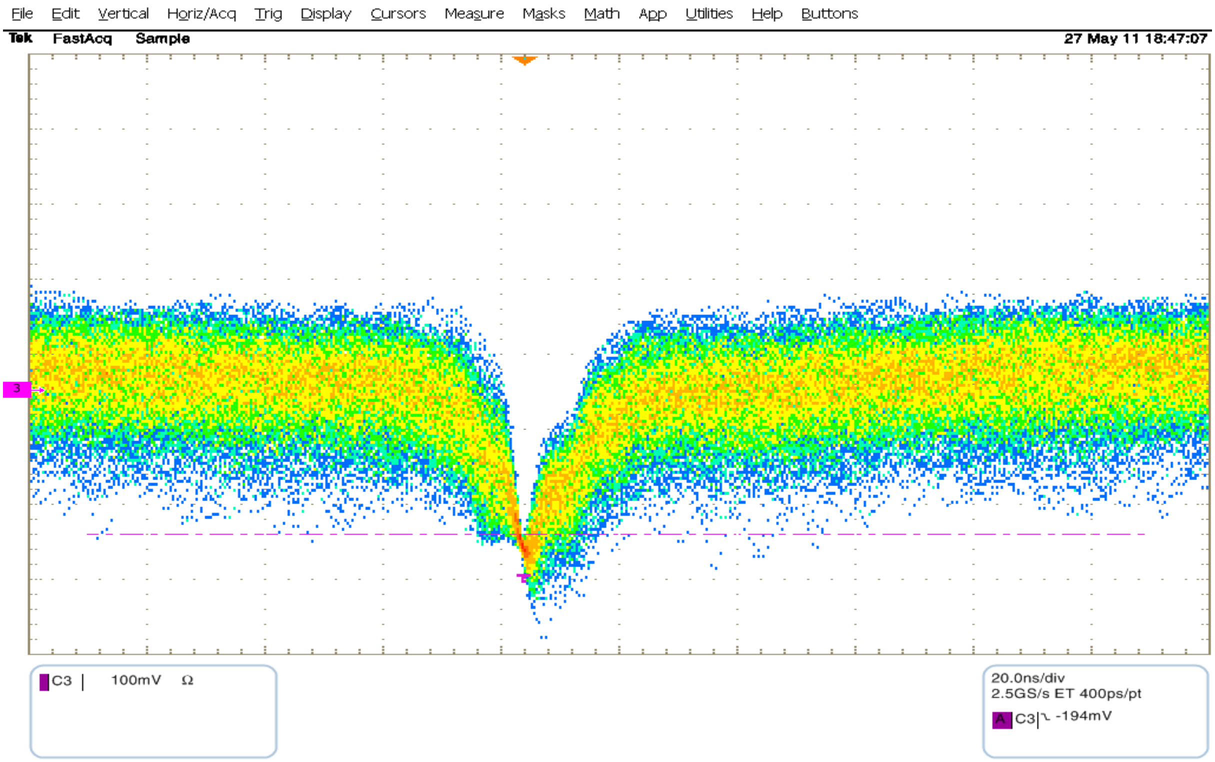


Figure 63: SiPM dark signals shape at 31V (2V overvoltage), Amplifier gain – 10^6

A dark pulse from a SiPM pixel is identical to one photo-equivalent signal and there is no way to distinguish such a signal whether its coming from a thermally generated carrier or from a photon generated one. Nonetheless, the dark pulses may have also more than one photo-equivalent signal due to optical crosstalk. The picture shows the signal on the scope overlaid on top of each other, dark noise signals from a SiPM. Most of the time, only one cell of the SiPM gives a signal. With lower probability, 2, 3, or even more cells can fire simultaneously due to optical crosstalk.

The SiPM dark count rate is determined by the threshold chosen during the measurements and it increases with temperature, detector volume and bias (since increasing bias increases the extent of the depletion volume). As the rate increases, the chance of observing multiple dark count together also increases. A typical figure for the characterization of the SiPM Dark Count rate is the so-called staircase figure plotting the dark count rates as a function of the pulse amplitude (thresholds) for different overvoltages .

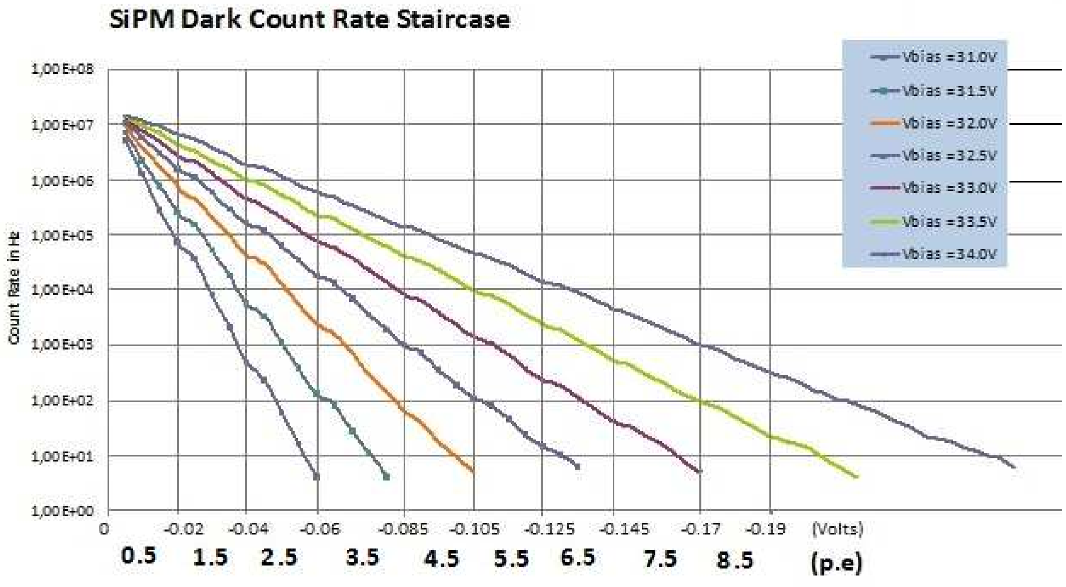


Figure 64: SiPM Dart Count Rate Staircase

Plotting the pulse frequency as a function of threshold position produces a step-like curve. The purpose of the staircase plot was to show the signal height of one p.e. at a certain voltage - (when 1 p.e it could not be observed possibly due to the noise). The flat regions of the curve correspond to the frequency of pulses at the single, double, triple photoelectron level. In the example shown the pulse frequency was measured at room temperature. At a low threshold the frequency corresponds to the single photoelectron level rate or dark rate of the silicon photomultiplier. As the threshold increases above the peak height of the single Geiger pulses, the rate falls as only pulses at the second photoelectron level, i.e. optic crosstalk, will now be measured. The rate then remains constant until the threshold increases above the peak height of two coincident pulses and then falls to the rate of the third photoelectron level. Pulses at the third and more photoelectron level are also mostly due to crosstalk.

The staircase plot is crucial for the reliability of the measurements with a SiPM. Supposing that for a measurement the threshold for the noise is not big enough to cut off any dark signals, the dark signals may be taken into account as real signals and result in detection errors. For LLL and photon counting applications are concerned, the choice of the right threshold becomes even more important. (picture of SiPM illuminated by the LED with 200ns pulse).

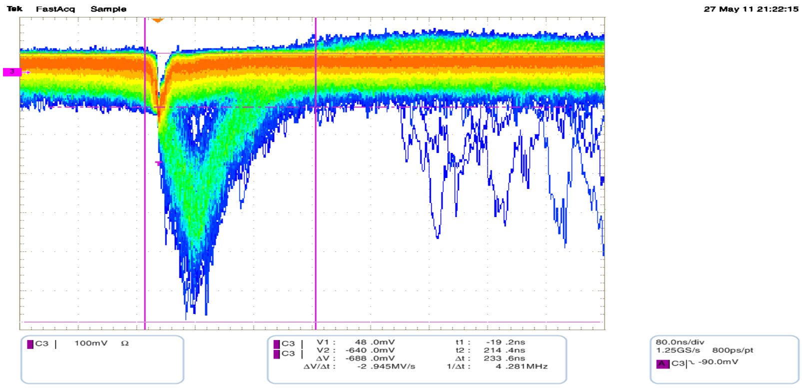


Figure 65 : SiPM illuminated by the LED with 200ns

In the figure below the SiPM was illuminated by the green LED with a light pulse of 500ns duration. The afterpulse effect is obvious. In addition, one can see the two components of the recovery time of the SiPM that were explained in previous chapter. This is a case of saturation of the SiPM as the response of the detector does not follow the command given by the LED. The pixels are all fired by once and then the recovery time of the device takes more time than the time interval between two consecutive light pulses.

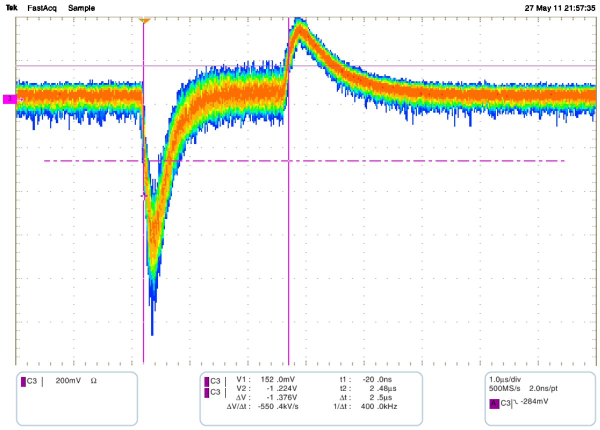


Figure 66: SiPM pulse illuminated by the green LED with a light pulse of 500ns duration. The afterpulse effect.

6.4 Test performed at CLEX area

The tests of the BLM detector were performed in the CLEX hall of CTF3. The location chosen for the installation was near the OTR screen (MTV.CC.0970) which is situated at the end of the TL2 -that is along the drive beam. Usually, the OTR screen is used for measurements of various beam parameters. However, the in the case of the tests performed in order to study the detector response and the background signal level., the OTR screen was used to create artificial losses, which is ideal for benchmarking, against MC simulations as losses within the CLEX hall are generally unknown.

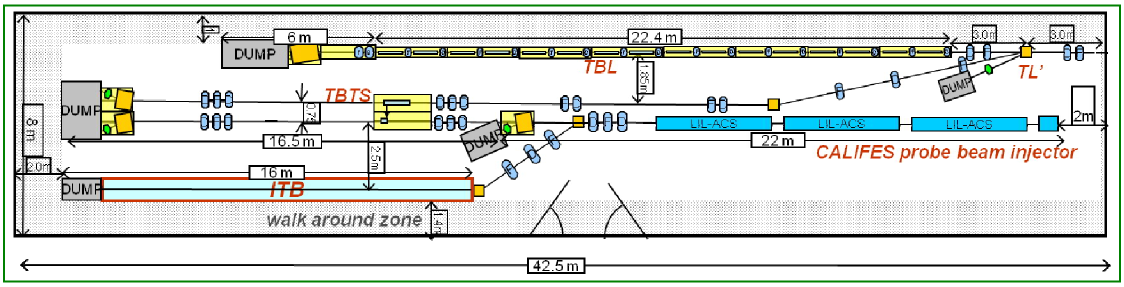


Figure 67: Layout of the CLEX area

Therefore, Monte Carlo simulations were performed by Marco Pionello and from their results, the fluence of particles (which are charged and over ~0.2 MeV), in the region upstream of the OTR is approximately **1.5e4cm^{-2} per bunch**, (@12GHZ this is 2.6e7cm^{-2} per bunchtrain). This number of particles for the volume of a Cherenkov radiator of 2mmx2mmx5mm as in our case is very high and saturation effects might be a problem (especially on SiPMs). Plus, the SiPM was directly in the radiation field and that the signal was from the directly irradiated SiPM.

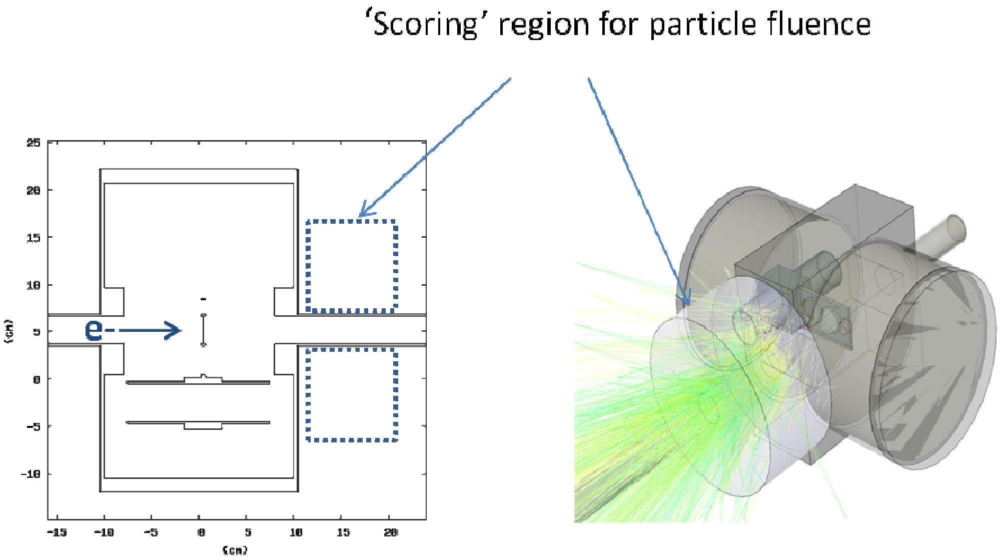


Figure 68: FLUKA representation of OTR Screen Chamber indicating regions where particle fluences are estimated. Left: Horizontal cut though screen. Right: 3D representation including visualization of particle tracks

The beam parameters when the tests were performed at the OTR screen (4/7/2011) were

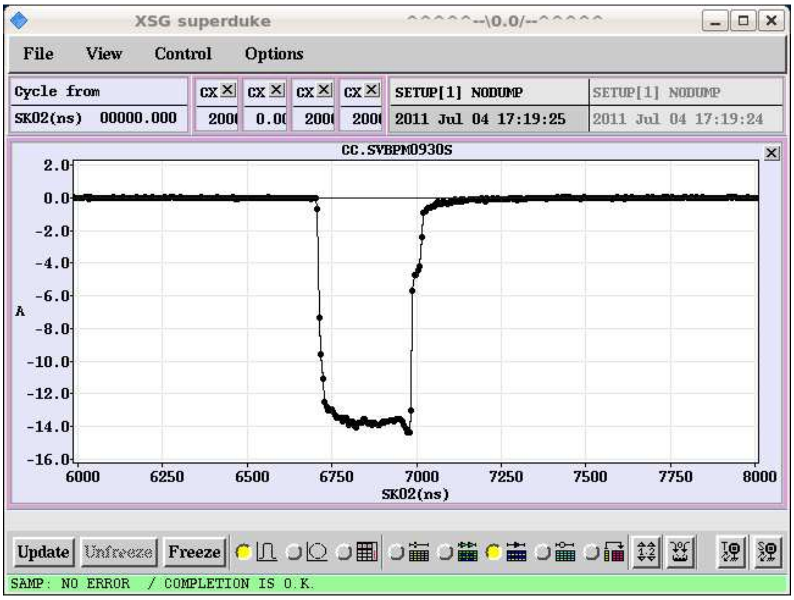
```
Energy: 112MeV

Rep Rate:  0.8333Hz

Pulse duration: 250ns
```

Bunch charge: 1.15nC

Bunch Frequency: 12 GHz



Cycle from SK02(ns) 00000.000

SETUP[1] NODUMP 2011 Jul 04 17:19:25

SETUP[1] NODUMP 2011 Jul 04 17:19:24

SAMP: NO ERROR / COMPLETION IS O.K.

Figure 69: The beam pulse had duration 250ns.

Average Current Reading – BPM upstream of OTR

The detector was installed in the same height of the beam pipe and as close as the narrow area permitted (picture below)

Figure 70: The OTR screen upstream side

For benchmarking reasons, an already calibrated ACEM was installed symmetrically from the other side of the beam pipe. The first step was the comparison of the raw signals (shape, saturation, etc). In the picture below one can see how the response from the two BLMs installed symmetrically. The ACEM signal is channel 2 (100mV/div) and the Cherenkov radiator-SiPM combination is channel 3 (20.0mV). Channel 1 corresponds to the trigger (400ns/div, 100mV/div).

As expected, one can see that the Cherenkov radiator+SiPM BLM is saturating. However, it has a fast rise time for the not saturated part. The ACEM is not saturating but in order to conclude to unbiased results an absolute calibration of the ACEM should be done in order to compare with the MC simulations results and their reliability.

As seen from above, it is possible that the tested combination of the Cherenkov bar and the SiPM is not appropriate for a 35A beam and for the detection of so high losses like these inserted by the OTR screen in the CLEX area, as it suffers from saturation.

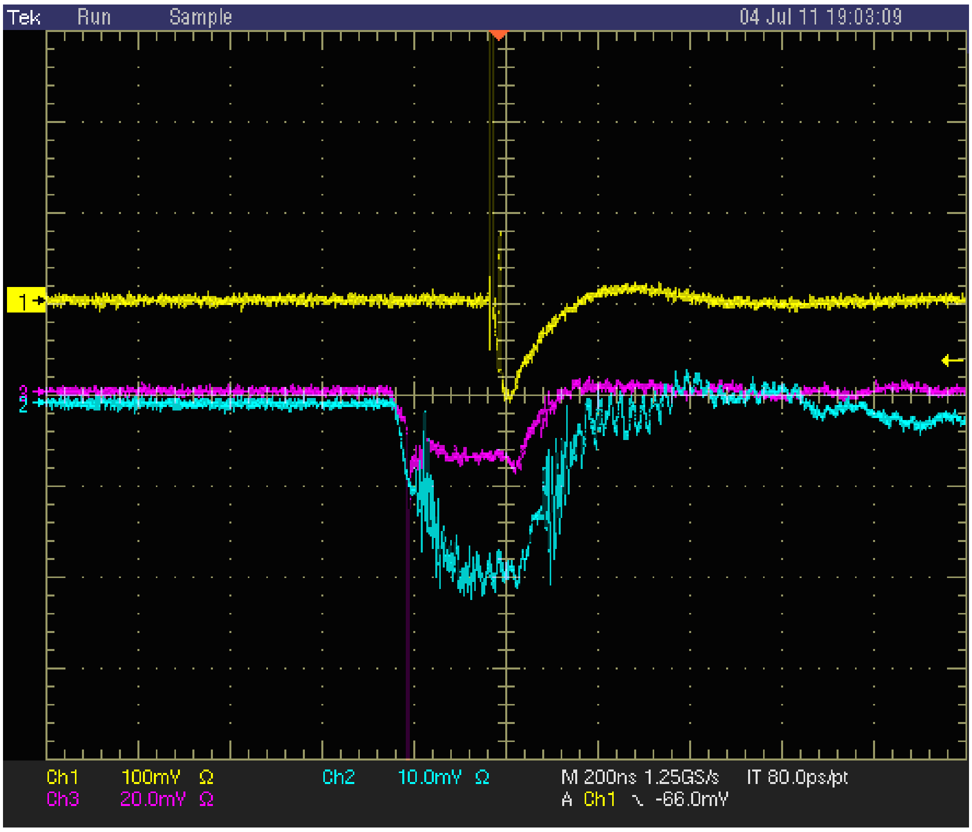


Figure 71: The response from the two BLMs installed symmetrically, triggered by a 400ns pulse (pink colour, 100mV/div). Blue pulse corresponds to the ACEM signal (100mV/div), yellow to the Cherenkov radiator-SiPM combination (20.0mV). (Date of the experiment: 04/07/2011 18:03)

It is proposed for the next measurements the device is removed further from the source for benchmarking purposes. In addition, the effect of radiation impacting directly the SiPM should be investigated. Should this effect prove to contribute significantly to the signal, a redesign of the detector geometry would be necessary such that the SiPM is shielded or removed from the secondary particle shower.

CHAPTER 7. CONCLUSIONS

7.0 Final remarks

Simulations and tests have been performed as part of the beam loss monitoring system design for CLIC study.

The first part was dedicated to the performance of extensive Monte Carlo calculations with Geant4 of the cascade in a copper target using Geant4 toolkit. There are six cases of primary particles which are 100GeV, 500GeV, 800GeV, 1.0TeV, 1.25TeV, 1.5TeV electrons. The number of events simulated is set to 10000. All the calculations have been made for cylindrical geometry and assuming target length of 2m and radius of 10cm. Results on the maximum energy deposition density in the target occurs at the downstream about z = 10cm and it depends on the energy of the primary particles. The lateral development of the cascade follows as expected an exponential shape for all of the energies and most of the energy deposition is concentrated close to the beam axis. As the impacting electron energy increases, so does the maximum energy deposition and hence energy density. Finally, the shower maximum, with the largest number of particles, is reached when the average energy particle becomes low enough to stop further multiplication.

The second part was dedicated to the characterization of a Silicon photomultiplier and the testing of a Cherenkov radiator read out by the SiPM. The response of the Cherenkov radiator was not ideal for the case of artificial losses that was tested. The signal was saturating and there are two possible explanations for that. The dynamic range of the SiPM is limited by the number of its pixel and the photons produced by the Cherenkov radiator correspond to a value of photons orders of magnitude higher than the number of pixels of the SiPM. Thus the signal produced of the Cherenkov radiator could not properly be read out by the used photodetector. In addition, the BLM was positioned very close to the beam pipe and the particle shower and the photodetector was possibly irradiated directly by gammas and electrons (electromagnetic shower) which could produce a secondary not Cherenkov correlated signal. A solution to the saturation problem would be to move the SiPM away from the beam pipe and have the Cherenkov radiator attatched to an optical fiber which will be at the end read out by the SiPM. In that way the SiPM would not be directly irratiated.

Further experimental plans for the CLIC BLM system, include testing of Cherenkov fibers read out by photodetectors and the are currently tested at CTF3. Other alternative technologies to be tested are scintillating fibers and long ionization chambers. The SiPM have the great advantage of cheapness and they are the primary present choice of photodetector for reading out the signal from optical fibers. Many tests were foreseen at the CTF3 facility CLEX in 2011 and their promissing results have already started to appear.

VI. REFERENCES

1. *http://press.web.cern.ch/public/en/LHC/LHC-en.html*

2. *http://clic-study.org/*

3. *http://project-clic-cdr.web.cern.ch/project-CLIC-CDR/Drafts/Beam_Instrumentation_v19.pdf*

4. *http://www.fnal.gov/pub/inquiring/matter/madeof/index.html*

5. *http://www.bnl.gov/bnlweb/facilities/LINAC.asp*

6. *S. Schlögl, K. Wittenburg; A Beam Loss Monitor System for HERA,Proc. XVth Int. Conf. on High Energy Accel., Hamburg 1992, Int. J. Mod. Phys. A(Proc. Suppl.) 2A (1993), Vol. 1, p. 254-256*

7. *K. Wittenburg; Beam Loss Detection, Proc. Ist European Workshop on Beam Diagnostic and Instrumentation for Particle Accelerators, Montreux 1993, CERN PS/93-35 (BD)*

8. *K. Wittenburg; Radiation Damage in PIN-Photodiodes, Nucl. Instr. and Meth. A270 (1988)*

9. *F Tecker 2008 J. Phys.: Conf. Ser. 110 112005*

10. *M. Sapinski et al., "Requirements of Beam Loss Monitoring System" Proceedings of IPAC'10, Kyoto, Japan*

11. *M. Jonker et al. "The CLIC Machine Protection", Proceedings of IPAC'10, Kyoto, Japan*

12. *G. Llosa et al. "Novel Silicon Photomultipliers for PET Applications", IEEE Transactions on Nuclear Science, 55(3):877 - 881, June 2008.*

13. *D.J.Herbert et al. "The Silicon Photomultiplier for application to high-resolution Positron Emission Tomography", Nucl Instr and Methods A 2007, A573, Issues 1-2, Pages 84-87.*

14. *G. Collazuol et al. "Single timing resolution and detection efficiency of the ITC-irst Silicon Photomultipliers", Nucl Instr and Methods A 2007, A581, 461-464.*

15. *F. Corsi et al. "Modelling a Silicon Photo Multiplier (SiPM) as a signal source for optimum front-end design", Nuclear Instr Meth A, 2007, 572, 416-418.*

16. *N.Dinu et al. "Development of the first Silicon Photomultipliers at ITC-irst", Nuclear Instr Meth A, 2007, 572, 422-426.*

17. *D.J.Herbert et al. "Study of SiPM as a potential photodetector for scintillator readout", Nuclear Instr and Methods in Phys Res A 2007, 567(1),356-359.*

18. *V.Bindi et al. "Preliminary Study of Silicon Photomultipliers fro Space Missions", Nucl Instr and Meth A 2007, 572, 662-667.*

19. *C.Piemonte et al. "Characterization of the first prototypes of Silicon Photomultipliers fabricated at ITC-irst", IEEE Trans Nucl Sci. 2007, 54(1), 236-244.*

20. *S.Moehrs et al. "A detector head design for small animal PET with silicon photomultipliers (SiPM)", Physics in Medicine and Biology, 51(2006), 1113-1127.*

21. *D.J.Herbert et al. "First Results of Scintillator Readout with Silicon Photomultiplier", IEEE Trans. Nucl. Sci. 2006, NS53(1), 389-394.*

22. *K. Wittenburg "Preservation of beam loss induced quenches, beam lifetime and beam loss measurements with the hera beam-loss monitor system"*

23. *J. Dusatkoet al. " The LCLS undulator beam loss monitor readout system"*

24. *R. E. Shafer " A tutorial on Beam Loss Monitoring",10th Beam Instrumentation Workshop 2002, Brookhaven, Mai 2002.*

25. *K. Wittenburg, "The PIN-Diode Beam Loss Monitor System At HERA" Beam Instrumentation Workshop 2000, Boston, USA, Mai 2000*

26. *http://geant4.cern.ch/*

27. *http://en.wikipedia.org/wiki/Pseudorandom_number_generator*

28. *J. Cohen-Tanugi "Optical Properties of the DIRC Fused Silica Cherenkov Radiator" SLAC-PUB-9735*

29. *K.K. Hamamatsu Photonics, Photomultiplier Tube Handbook, Electron Tube Division, third ed., 2006.*

30. *K.K. Hamamatsu Photonics, Photodiode Technical Guide. http://sales.hamamatsu.com/assets/html/ssd/si-photodiode/index.htm*

31. *A.Intermite et al. "Feasability study of an optical fibre sensor for beam loss detection based on a SPAD array"*

32. *http://www-zeuthen.desy.de/physics_seminar/transparencies/erika_garutti.pdf*

33. *Performance study of Silicon Photomultipliers as photon detectors for PET R. Verheyden, Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE*

34. *N. Dinu ,Characteristics of a prototype matrix of Silicon PhotoMultipliers (SiPM) PIXEL 2008 INTERNATIONAL WORKSHOP FERMILAB, BATAVIA, IL, U.S.A. 23–26 SEPTEMBER 2008*

35. *Paolo Finocchiaro, "Features of Silicon Photo Multipliers: precision measurements of gain, cross-talk, afterpulsing, detection efficiency" SORMA West 2008 June 2-5, Berkeley, California, USA talk c1a163*

36. *Kazuhisa Yamamura 1 " Production and Development status of MPPC| ", International Workshop on New Photon Detectors (PD09) Shinshu University Matsumoto Japan, 24-26 June 2009*

37. *H. Kume, S. Sawaki and M. ITO, "20 inch diameter Photomultiplier", Nuclear Instruments and Methods 205 (1983) 443-449 443, North-Holland Publishing Company*

38. *Ronzhina, M. Albrow "Tests of Timing Properties of Silicon Photomultipliers" Fermilab pub-10-052-PPD*

39. *Nicholas Simos "Irradiation damage studies of high power accelerator materials" IWSMT8*

40. *Barry W. Baumbaugh et al, Studies of SiPM and Scintillation Plates with Waveshifter Fiber and SiPM Readout, 2009 IEEE Nuclear Science Symposium Conference Record N13-199*

41. *N. Otte , The Silicon Photomultiplier - A new device for High Energy Physics, Astroparticle Physics, Industrial and Medical Applications.. Max-Planck-Institut fur Physik*

42. *Stoykov & R.Scheuermann, Silicon Avalanche Photodiodes, Laboratory for Muon Spin Spectroscopy, Paul Scherrer Institut*

43. *G. A. Greene C. C. Finfrock and A. L. Hanso , Energy deposition in a thin copper target downstream and off-axis of a proton-radiography target BNL-52668 Formal Report, Energy Sciences and Technology Department, Brookhaven National Laboratory*

44. *G.Collazuol Review of Silicon photo-multiplier physics and applications, including study at low temperature Overview IPRD08*

45. *Y. Musienko State of the art in SiPM's CERN, SiPM workshop, 16.02.2011*

46. *N. Dinu et al. Development of the first prototypes of Silicon PhotoMultiplier (SiPM)at ITC-irst426 / Nuclear Instruments and Methods in Physics Research A 572 (2007) 422–426*

47. *487R Shimizu and Ding Ze-Jun , Monte Carlo modelling of electron-solid interactions 1992 Rep. Prog. Phys. 55*

48. *C. Piemonte "Recent Developments on Silicon Photomultipliers produced at FBK-irst"*

49. *C.Piemonte NIM A 568 (2006) 224*

50. *The Particle Detector Brief Book (isbn 9783642083839)*

51. *D. Renker "Geiger-mode avalanche photodiodes for Cherenkov Detectors" workshop on fast Cherenkov detectors-photon detection, DIRC design and DAQ,May 11-13, 2009, Giessen, Germany*

52. *B. Dolgoshein Int. Conf. On New Developments in Photodetection, Beaune, France, 2002*