Diss.ETH No.13585

The Silicon Strip Tracker for the CMS experiment at LHC

A dissertation submitted to the

Swiss Federal Institute of Technology, Zürich

for the degree of

Doctor of Natural Sciences

presented by

Roberto Della Marina

Laurea in Fisica from Trieste University, Italy

born 20th February 1965

in Udine, Italy

accepted

on the recommendation of

Prof. Dr. F. Pauss, examiner, and

Prof. Dr. G. Viertel, coexaminer, and

Prof. G. Tonelli, coexaminer

2000

Abstract

This thesis is a contribution to the conception, the fabrication and the optimization of the silicon microstrip detector for the Silicon Strip Tracker (SST) of the Compact Muon Solenoid (CMS) experiment at the future Large Hadron Collider (LHC).

The physical motivation of this work are can be summed up by the High Energy Particle community's need to identify the mechanism of the electroweak symmetry breaking, taking the studies done at LEP and Tevatron into a much higher energy domain. LHC will allow experiments to cover the entire mass range foreseen in the Standard Model (SM) for the Higgs boson and most of the parameter space for its supersymmetric (MSSM) extensions. The CMS detector has been designed to fulfill these tasks: the combined use of a very precise silicon inner tracker and highly hermetic outer calorimeters is essential in order to successfully identify b–jets coming from $t \rightarrow h^{+}b$ or directly from Higgs decays such as $h \rightarrow b\bar{b}$. In particular, the goals of the SST are to ensure high quality momentum resolution, precise e/γ separation and excellent charged track isolation.

At LHC, the irradiation level is expected to be so high that silicon, the innermost detector construction material, will be affected during the sensor life, so that its performance will be strongly dependent on irradiation time and distance from the beam. Therefore, the detector's design and technology are constrained not only by the physics requirements but also by the difficult operating conditions. The key to fast and effective handling in this complex situation is close collaboration with silicon manufacturers. This thesis is a summary of the work I had the opportunity to perform at the *Centre Suisse d'Electronique et de la Microtechnique* SA (CSEM), a research centre in Neuchâtel, Switzerland.

The refinement of an existing process already used at CSEM for the fabrication of the detectors for ALEPH, the L3 double-sided microvertex detectors at LEP and the AMS satellite experiment, has led to the definition of a new silicon technology well-suited for the SST silicon detectors. This new technology features great improvements in radiation hardness, direct integration on-chip of a decoupling capacitor for each channel and high voltage operation. A profound understanding of silicon technology is necessary to achieve the planned goals, since problems such as dielectric breakdown, leakage current, metal resistance or integrated capacitor yield can only be solved by integrating ad-hoc process modifications with an optimized design.

The logical continuation of this work is the detector characterization. The results of such tests have been directly fed into a TCAD (technology computer aided design) framework in order to test the possibility of optimizing both design and process to give the required performances, without waiting for a new integration. The effectiveness of

electrically good detectors was verified by means of test beams before and after irradiation. The understanding of their integration in the read-out chain (i.e. coupling with the electronics and the DAQ) gave information about position resolution. As an example of this complex process, the results obtained in September 1997 in building the prototypes for the Silicon Barrel (SiB1) and Silicon Forward (SiF1) Milestones are presented.

Once the feasibility of such a silicon tracker system has been confirmed, it is important to test the global CMS performances by studying the simulation of an interesting physics channel. A study on the possible observation of the lightest SUSY scalar Higgs boson h via its decay $h \rightarrow b\bar{b}$ is presented in this thesis. The strategy is to exploit the strong interaction production of gluinos and squarks. Their typical cascade decay products are neutralinos and charginos (primarily $\tilde{\chi}_2^0$) which decay abundantly into h. The identification of the Higgs bosons is done via the reconstruction of the $b\bar{b}$ pairs once the large $b\bar{b}$ backgrounds are suppressed using missing energy $\not{E}_T$ and jet multiplicity cuts. In order to reduce the number of free parameter of the MSSM theory, we have chosen to work within the SUGRA approximation driven by supergravity-inspired theories. The conclusion is that the $h \rightarrow b\bar{b}$ peak was already visible in an interesting region in the SUGRA parameter space during the low luminosity period of LHC.

In Appendix 2 is described a study about the possibility to align several plans of Silicon microstrip detectors by means of infra-red (IR) laser tracks measured by the detectors themselves. We performed this study for the AMS (Alpha Magnetic Spectrometer) collaboration, being the silicon wafers designed and produced by CSEM. We show that by suitably modifying the dielectric thickness already present atop the sensor, the detector reflectivity can be reduced from 35% down to the 1% range: IR laser tracks can thus cross the six planes of detectors without significant intensity losses.

Riassunto

Lo scopo di questa tesi e.g. di descrivere il processo di concezione, progetto, fabbricazione e ottimizzazione dei detectors a microstrip in silicio, in vista della loro utilizzazione nel Silicon Strip Tracker (SST) dell'esperimento Compact Muon Solenoid (CMS) concepito per il futuro Large Hadron Collider (LHC).

Le motivazioni fisiche all'origine di questo lavoro sono riassunte dalla necessità di identificare il meccanismo della rottura della simmetria elettrodebole, estendendo cosí gli studi svolti a LEP e a Tevatron ad uno spettro di energie molto più vasto. LHC offre la possibilità ai differenti esperimenti di coprire l'intero spettro di massa previsto per il bosone di Higgs nel Modello Standard cosí come gran parte dello spazio dei parametri previsti per le sue estensioni supersimmetriche (MSSM). Il rivelatore CMS e' stato concepito proprio per soddisfare queste esigenze: un sofisticato sistema di tracking al silicio di altissima precisione posto nelle zone piu' prossime alla zona di interazione combinato con un sistema di calorimetri quasi completamente ermetico posto nelle zone piu' esterne, permette di identificare correttamente i b–jet provenienti da $t \to h^{\perp} b$ o direttamente da decadimenti di Higgs quali ad esempio $h \to b\bar{b}$. In particolare, gli scopi dello SST sono di assicurare la risoluzione in momento migliore possibile, la separazione precisa tra elettroni e gamma ed un eccellente isolamento nella ricostruzione delle tracce da particelle cariche.

Ad LHC il livello di radiazione previsto e' tale da causare l'inversione effettiva del tipo di drogante usato correntemente per la fabbricazione dei detectors al silicio, destinati a equipaggiare gli strati piu' interni del rivelatore. Questo fenomeno rende le prestazioni dei detector al silicio dipendenti dal tempo di irraggiamento e dalla loro distanza dal punto di interazione.

Al fine di trovare le soluzioni piu' rapide ed efficaci al progetto di un tracciatore al silicio per LHC si e' rivelata fondamentale la stretta collaborazione con i produttori dei sensori al silicio. In questa tesi ho raccolto parte del lavoro che ho effettuato per il rivelatore a microstrip di CMS a CSEM, *Centro Svizzero di Elettronica e Microtecnologia* SA, un centro di ricerca e sviluppo basato a Neuchâtel, in Svizzera. All'interno di questo osservatorio privilegiato, ho contribuito a sviluppare una versione resistente alla radiazione del processo precedentemente usato a CSEM per la fabbricazione dei detectors doppia faccia installati in ALEPH ed L3 a LEP e nel satellite AMS. Le modifiche riguardano ovviamente gli aspetti precipui dell'impiego di tali dispositivi in ambienti ad alto tasso di radiazione, quali ad esempio lo spostamento della soglia di rotture dei diodi ad altissima tensione o l'integrazione delle capacita' di disaccoppiamento direttamente sopra le strip. Cruciale si e' rivelata la mia partecipazione alla fabbricazione dei dispositivi: la mia

esperienza nel progetto dei detectors unita a dettagliate conoscenze tecnologiche hanno permesso giudicare a priori l'interesse di una modifica progettuale o di processo, riducendo quindi sensibilmente i tempi di implementazione di nuove soluzioni. Continuazione logica di questo lavoro e' il collaudo dei sensori, i cui risultati sono stati utilizzati come verifica delle simulazioni tecnologiche e di dispositivo che ho messo a punto in parallelo alla fabbricazione. Un corretto utilizzo dei simulatori permette di accorciare ulteriormente il processo di ottimizzazione sia del disegno che del processo, senza dover aspettare la fine della fabbricazione di una produzione di test. L'efficacia dei sensori giudicati buoni al test elettrico e' stata verificata utilizzando fasci di particelle, principalmente al CERN, sia prima che dopo l'irraggiamento dei singoli detector. Lo studio del loro comportamento con e senza l'elettronica di lettura mi ha permesso di estrarne le caratteristiche di risoluzione in posizione. Come esempio di interazione tra tutti questi differenti aspetti, presento qui i risultati che abbiamo ottenuto nel settembre 1997 costruendo i prototipi per le "Milestones" del Barrel Tracker (SiB1) e del Forward Tracker (SiF1).

Se la fattibilita' del progetto di un tracciatore al silicio e' confermata dai test, l'attenzione deve spostarsi sulle prestazioni dell'intero esperimento CMS misurate, ad esempio, studiando un canale di fisica particolare. In questa tesi presento lo studio della possibilita' di rilevazione del piu' leggero bosone di Higgs supersimmetrico h attraverso il suo decadimento $h \rightarrow b\bar{b}$. La strategia utilizzata per l'osservazione sfrutta l'importante sezione d'urto di produzione di gluini e squarks presente a LHC. I loro prodotti tipici di decadimento sono neutralini e chargini (principalmente $\tilde{\chi}_2^0$) i cui decadimenti producono soprattutto h. L'identificazione dei bosoni di Higgs e' data dalla ricostruzione delle coppie di $b\bar{b}$, una volta che i principali fondi sono stati eliminati da tagli sull'energia mancante $\not{E}_T$ e sulla molteplicita' dei jets. Per poter definire piu' facilmente lo spazio dei parametri della teoria MSSM in cui effettuare l'analisi, si e' scelto di utilizzare l'approssimazione SUGRA, derivata da considerazioni di supergravita'. La conclusione di questa analisi e' che il picco del segnale $h \rightarrow b\bar{b}$ e' visibile in una regione importante dello spazio dei parametri definiti da SUGRA gia' durante il periodo di operazione a bassa luminosita' di LHC.

Acknowledgments

In writing this thesis I have been fortunate to have had input from a large range of colleagues and friends. The manuscript has been through the sieve of their knowledge, each one culling out some dross. Many people have clarified concepts or relationships to me. Others have helped to solve the puzzle of english translation of italian thoughts. Still others, through their enthusiasm have encouraged me to persist in this project in spite of my job at CSEM taking over all my time.

I want to thank the following people without those help this work could not have been complete: Daniele Passeri, Alberto Messineo, Otilia Militaru, Martina Da Rold, Elisabetta Catacchini, David Lucchetti, Francesco Rizzo, Concezio Bozzi, Andrei Starodumov, Suchandra Dutta, Piero Giorgio Verdini, Richard Wheadon, Roberto Dell'Orso, Ettore Focardi, Nicola Bacchetta, Roberto Battiston, Gert Viertel, Wolfgang Wallraff, Gianmario Bilei, Marcello Mannelli and the whole CMS Tracker Group whose innumerable contributions have been essential for me and for this work. To Luciano Bosisio, my silicon detector mentor, who knew how to "migrate" a high energy particle physicist into the great world of semiconductor detectors, my deepest gratitude. To Guido Tonelli, thanks for your wonderful enthusiasm and all physics advices which drove me to the end of this work. Thanks to Felicitas Pauss, for giving me the possibility to fulfil my Ph.D. work at ETH and for your patience and helpfulness always proffered. To Hans van den Vlekkert, great spirit and good friend, who at CSEM guided me to features and subtleties of the industrial world, nevertheless pushing me to finish my thesis work, a big thank you. And to Paul Weiss, great scientist and friend, my guardian angel at CSEM, always prone to help and sustain me but firm in pushing me to the end of this thesis, all my gratitude and amity. Special thanks to Monica Federico e Simondavide Tritto, unique friends and colleagues, for your careful support with science, humor and food in the last years at Neuchâtel. To André Perret and all the colleagues at CSEM for your assistance and suggestions along the way, I extend a special thanks. Thanks to John Walsh and Giuliana Rizzo, Nilo Segura Chinchilla, Stefano Buono e Daniela Macina, Alessandro Variola e Fiorella Cavaniglia for your invaluable hospitality and inspiring discussions in Geneva. And finally to Paola, friend and partner, wife for life, for your steadfast and unconditioned support, incredible help and patience, all my love.

Contents

Chapter 1

Introduction

The purpose of this thesis is to describe the conception, the fabrication and the optimization of a silicon microstrip detector for the CMS experiment at the future Large Hadron Collider (LHC). Such a device must be able to survive 10 years in the difficult LHC environment.

At present, the silicon detector technology has been adapted to the purposes of the LEP and Tevatron experiments. At LHC, the irradiation level is expected to be so high that the effective doping of silicon, the innermost detectors construction material, will change type during the sensor lifetime, causing its performances to be strongly dependent on irradiation time and distance from the beam.

These dramatic changes to working conditions for silicon detectors in high energy experiments have triggered numerous developments in the field of radiation hardness. This has led to a full revision of the formerly standard fundamental parameters of a microstrip sensor, such as design choice, integration technology, acceptance criteria and cooling scheme.

The problems related to radiation damage are many: the increase of depletion voltage and of leakage current, the decrease of charge collection efficiency and of isolation between strips as well as the increase in their capacitive coupling. For the first time in the history of silicon detectors, their design and technology are constrained not only by the physics requirements but also by the difficult operating conditions.

The key to a fast and effective solution in this complex operation is a closer collaboration with silicon manufacturers. This thesis is a summary of the work I had the opportunity to perform for the CMS silicon microstrip detector at the *Centre Suisse d'Electronique et de la Microtechnique* SA (CSEM), a silicon detector producer based in Neuchâtel, Switzerland. In this privileged environment I became acquainted with the silicon detector fabrication process, which is a derivation of the known planar process, currently employed at CSEM for the production of integrated circuits. We refined an existing process already used in the fabrication of the detectors for the ALEPH and the L3 double-sided microvertex detectors at LEP and the AMS satellite. The modifications we studied are related to the different requirements of the LHC silicon detectors, namely radiation hardness, direct integration on-chip of a decoupling capacitor for each channel

and high voltage operation. In addition, the need of CMS to cover tens of square meters with microstrip detectors requires the development of simpler devices with simpler fabrication processes, which, however, have no detrimental effects on sensor performance. With such large instrumented surfaces the reliability of the devices becomes a crucial issue since no replacement of silicon microstrip detectors are foreseen.

The performances of the CSEM detectors developed with the support of the CMS Silicon Strip Tracker community are sufficiently good to allow the drafting of the Central Tracker Technical Design Report with only marginal modifications with respect to the Technical Proposal. This demonstrates the suitability of the choices made by the CMS collaboration and its shrewd confidence in the technological development.

In Chapter 2 I will briefly introduce the physical motivation behind the conception of the CMS detector. The starting point is the need to identify the mechanism of electroweak symmetry breaking, extending the studies done at LEP and Tevatron into a much larger energy domain. LHC will allow experiments to cover the entire mass range foreseen in the Standard Model (SM) for the Higgs boson and most of the parameter space for its supersymmetric (MSSM) extensions. The CMS detector has been designed to fulfill these tasks and discover the Higgs boson(s). In this Chapter I will summarize how each CMS subdetector is optimized and I will summarize their main characteristics. Particular emphasis will be given to the Silicon Strip Tracker, the main subject of this thesis.

Chapter 3 is specifically devoted to the study of the most fundamental element of the silicon tracker, the microstrip detector. Starting with a short summary of the energy loss mechanism in solids I will discuss the physics of microstrip silicon devices. I will show how the design can be optimized for long exposure to heavy irradiation. The biasing method, the decoupling from the electronics, the strip isolation and the importance of the leakage current are all items I will address with a view to optimizing the CMS tracking performances. Each piece of this complicated patchwork must be tuned in order to obtain the best spatial resolution, lowest noise and cheapest price possible. This chapter will finish with the definition of a set of design guidelines, a "cookbook", for the CMS silicon microstrip detector, ready to be applied in fabrication.

The details of the fabrication of the CMS silicon detector are addressed in Chapter 4. Often the final users of silicon detectors, high energy particle physicists, have a limited knowledge of the steps followed in designing silicon sensors. The details of the fabrication are often neglected. My thesis demonstrates the effectiveness of a good understanding of the basics of Silicon Technology. Problems such as dielectric breakdown, leakage current, metal resistance or integrated capacitor yield can be solved integrating ad-hoc process modifications with an optimized design. The strong interconnection between these two aspects has made possible the development of a "CMS process" for radiation-hard silicon detectors which is now employed at CSEM.

In Chapter 5 the characterization of the CMS detector prototypes performed in the last three years is described. First, electrical measurement on chip were performed in order to understand the design specificities and the process effectiveness. The results of such tests were directly fed into a TCAD(technology computer aided design) framework. The detailed process and device simulations may help in understanding fuzzy behaviours

or possible failures and may also be used to check the possibility of optimizing both design and process to the required performances, without being obliged to wait for a new production. After completion of electrical characterization, the detectors have been tested under particle beams, before and after irradiation. The understanding of the detector's integration in the read-out chain (i.e. coupling with the electronics and the DAQ) give information about position resolution for a track passing through the detector. Different detector geometries were tested. At the end of this chapter I will report the results obtained in September 1997 in building the prototypes for the Silicon Barrel (SiB1) and Silicon Forward (SiF1) Milestones.

Once the feasibility of such a silicon tracker system has been confirmed, it is important to test the global CMS performances by looking into the simulation of an interesting physics channel. In Chapter 6 a study on the possible observation of the lightest SUSY scalar Higgs boson h via its decay $h \rightarrow b\bar{b}$ is presented. The strategy is to exploit the strong interaction production of gluinos and squarks. Their typical cascade decay products are neutralinos and charginos (primarily $\tilde{\chi}_2^0$) which decay abundantly into h. The identification of the Higgs bosons is done via the reconstruction of the $b\bar{b}$ pairs once the large $b\bar{b}$ backgrounds are suppressed with missing energy $\not{E}_T$ and jet multiplicity cuts. In order to reduce the number of free parameter of the MSSM theory, I have chosen to work within the SUGRA approximation driven by supergravity-inspired theories. I will show that the $h \rightarrow b\bar{b}$ peak is already visible in an interesting region in the SUGRA parameter space during the low luminosity period of LHC.

Chapter 2

The CMS Detector at LHC.

2.1 Introduction

In the beginning of the new century a new large accelerator, the Large Hadron Collider (LHC), will become operational in the existing 27 km long tunnel of the LEP electron-positron collider, actually in operation at CERN. This new collider will accelerate protons in order to get proton-proton collision at centre-of-mass energy of 14 TeV and luminosity up to $\mathcal{L} \sim 10^{34}$ cm^{-2}s^{-1}. These performances will extend the centre-of-mass (c.m.) energy attained by the Tevatron (the $p\bar{p}$ collider at Fermilab (Chicago, USA)) by a factor ten and luminosity by more than two order of magnitude.

The main motivation for building such machine is to investigate the mechanism responsible for electroweak symmetry breaking. Both Standard Model (SM) [1] and its minimal supersymmetric extension (MSSM) [2] suggest models for the origin of particle masses [3, 4]. The present experimental mass constraints on the SM Higgs boson are loose ($77.5 < m_H \lesssim 450$ GeV/c^2 at 95% C.L.), with bounds fixed by LEP and SLD experiments [5, 6] Theoretical arguments on weak Higgs self-coupling fix in any case at 1 TeV/c^2 the upper bound for an elementary Higgs mass [7].

The design and optimization of an LHC detector like CMS (Compact Muon Solenoid) [8] must be focused in the possibility to explore the entire SM Higgs mass range taking into account all the Higgs decay modes and their experimental signatures. Since the Higgs mechanism is foreseen as very likely but it is not an observed phenomena yet, the detector must allow the study of the electroweak symmetry breaking mechanism even if it is not mediated by the Higgs boson. The requirement of detecting a variety of different signatures predicted by physics beyond Standard Model lead to the concept of a general purpose detector.

Recently it has been pointed out that the introduction of supersymmetric theories (SUSY) may cure the mass divergence problem of the SM Higgs, but at the expense of introducing several Higgs bosons [2]. In the minimal extension of the SM there are five such states and they can eventually be investigated by CMS in a large portion of parameter space, since the requirements for Higgs hunting in MSSM are similar to SM. The search for new physics in an hadronic environment is furthermore enhanced by noting

that the production cross section for strongly interacting gluinos and squarks are large. The detector requirements for SUSY studies are similar to what is foreseen for SM: the missing momentum measurement, crucial tag for MSSM Higgs decay modes as well as for gluino and squark detection, asks for highly hermetic detectors. Although high luminosity is essential to cover the entire range of mechanisms of electroweak symmetry breaking and to explore a significant fraction of SUSY parameter space, LHC will start at a significant lower luminosity. The CMS detector is well suited to benefit of this period ($\mathcal{L} \leq 10^{33}\text{cm}^{-2}\text{s}^{-1}$) allowing the study of large cross section processes as those involving beauty and top quarks. The important ingredients for b-physics are good momentum and effective mass resolution and high quality secondary vertex reconstruction. The combined use of a silicon tracker and a pixel detector allows high b-jet tagging efficiency and purity, especially useful for top physics ($t \rightarrow H^{(\pm)}b, H^{\pm} \rightarrow \tau\nu$) and some Higgs production ad decay modes, both at low and high luminosity. For example the observation of $H \rightarrow b\bar{b}$ with $m_H \sim 100\,\text{GeV}$, both in SM and SUSY may be possible only through the tagging of accompanying b-jets.

The advantages and the physics potential of the precise tracking performances provided by the micro-strip silicon detectors have been extensively demonstrated both by LEP experiments at CERN and CDF at Fermilab. In CMS goal is to get similar sophisticated performances on a larger scale of apparatus and in a much more difficult environment.

In the following sections the CMS structure will be summarized. In Sect. 2.3 all the subdetectors will be introduced and briefly described. In Sect.2.4 the structure of the silicon tracker will be studied in more detail.

2.2 The LHC collider

The construction of the Large Hadron Collider (LHC)[9] has been approved by CERN in late 1994. The machine will provide proton-proton collisions with a c.m. energy of 14 TeV and an unprecedented luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$. In order to achieve the design energy with the constraint of the 27 km circumference of the LEP tunnel, a superconducting magnet system must be operational at a temperature of superfluid helium below 2 °K, yielding a dipole field of 8.4 T. The LHC is designed as a double beam proton-proton collider: the two rings are inserted in the same magnet and cryostat in order to save space and costs. The machine will provide also heavy ions collisions with a luminosity of $10^{27}\text{cm}^{-1}\text{s}^{-1}$ using the existing CERN heavy ion production facility. Space will be available above the LHC tubes for the eventual reinstallation of LEP components in view of a possible $e - p$ physics program. Figure 2.1 shows a sketch of the beam facilities at CERN.

2.3 The CMS Detector

The Compact Muon Solenoid (CMS) is one of the two general purpose detectors foreseen at LHC.

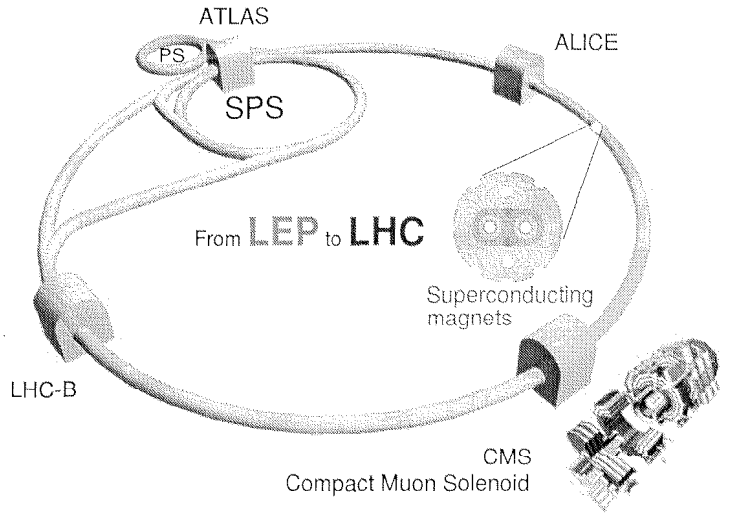


Fig. 2.1: Pictorial view of the accelerator complex at CERN. The LHC machine will be build in the same tunnel of LEP.

CMS has been designed with particular emphasis to hermeticity (covering as much of solid angle as possible), to momentum resolution up to the highest energies (by means of a large tracking system in a 4 T solenoidal magnetic field), to precision measurement of photon and electron (by means of a very segmented, short radiation length, high resolution crystal electromagnetic calorimeter placed inside the coil), to muon trigger and identification (by means of a redundant and hermetic system with sufficient iron to absorb hadrons which serves as the magnetic return yoke) [10].

In the following sections a summary of the characteristics and the performances of the main CMS sub-detectors is presented. Special attention is devoted to the silicon microstrip detectors, part of the central tracker, which are the main subject of this work.

2.3.1 The Magnet

The "engine" of CMS is a long superconducting solenoid (l= 13 m) with an inner radius of 2.95 m generating a uniform magnetic field of 4 T [8]. The magnetic flux is returned through a 1.8 m thick saturated iron yoke (1.8 T) instrumented with muon chambers. A single magnet thus provides the necessary bending power for precise inner and muon tracking, and efficient muon detection and measurement up to rapidity of 2.5. Measurements within the iron yoke provide the muon stand-alone capability. In Fig. 2.2 the CMS experiment is shown in a 3-dimensional view. The overall dimensions are: a length of about 21.6 m, a diameter of 15 m and a total weight of about 12500 tons.

2.3.2 The Muon System

Muons from p-p collisions are expected to provide clean signatures for a wide range of new physics processes. The muon detector should fulfill three basic tasks: muon identification, trigger and momentum measurement. The 4 T magnetic field and the return yoke, which

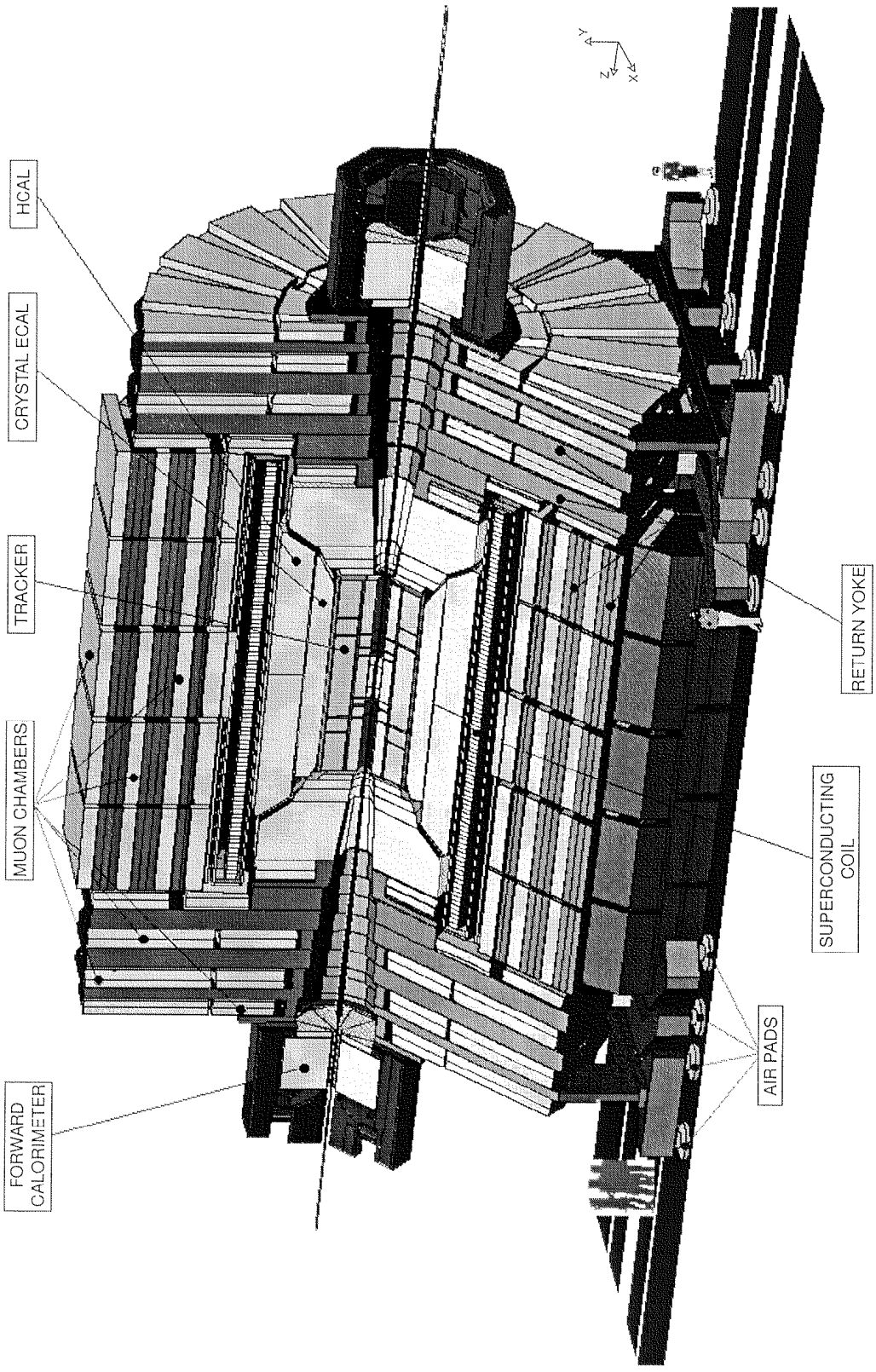


Fig. 2.2: Three dimensional view of the CMS detector.

also serves as the absorber for muon identification, permits to fulfill these requirements. Starting from the primary vertex, muons are first measured in the inner tracker, then traverse the calorimeters, the coil and the return yoke. They are identified and measured in four muon stations inserted in the return yoke in the barrel and end-cap regions. The four stations provide redundancy and optimize the geometrical acceptance. The four muon stations include also triggering planes that identify the bunch crossing and enable a cut on the muon transverse momentum at the first trigger level.

In a barrel muon station, about 40 cm deep, two groups of 4 layers of aluminum drift tubes are used in the bending plane to measure precisely the $r-\phi$ coordinate and the local trajectory slope. Another 4 layers are dedicated to the z coordinate. There are about $2 \cdot 10^5$ drift tube channels in the barrel region. The goal is to achieve a space resolution of better than 200 μm per layer, giving a position accuracy of about 100 μm and an angular precision of about 1 mrad per station.

The solenoidal field bends tracks in the $r - \phi$ plane. The muon momentum may be measured in three ways:

1. sagitta measurement in the inner tracker;
2. bending angle measurement immediately after the coil;
3. sagitta measurement in the return yoke.

The last two methods allow a muon stations stand-alone measurement. Best results are attainable combining the three methods, as shown in Tab. 2.1.

A time resolution of much better than 25 ns is needed to identify the bunch crossing. Six layers of RPC (resistive plate chambers) with a time resolution of 2 ns and sufficient space resolution are foreseen for the first-level muon trigger.

The four end-cap muon stations which must work in a magnetic field use six-layer cathode strip chambers (CSC) for precision measurement of muon position ($\sim 100\,\mu$m per layer) and momentum. The CSCs are also used for trigger purpose and are supplemented by the RPC trigger system which extends up to $|\eta| = 2.1$. This double triggering scheme insures robustness of muon triggering. In CMS the lowest values of the muon thresholds for 90% trigger efficiency in various rapidity range are listed in Table 2.2.

These thresholds will be useful for multi-muon final states at low luminosity or in B-physics studies. The two-muon final states in Z or Higgs decays are less demanding ,

p_t	$\Delta p_t/p_t$ stand alone	$\Delta p_t/p_t$ combined
10 GeV/c	$6 \div 10\%$	$0.5 \div 1.0\%$
100 GeV/c	$7 \div 20\%$	$1.5 \div 5\%$
1 TeV/c	$15 \div 35\%$	$5 \div 20\%$

Table 2.1: Momentum resolution for different muon transverse momenta. The first column accounts for muon chambers only; in the second column the three methods are combined in $0 < |\eta| < 2.4$ with a vertex constraint of 15 μm.

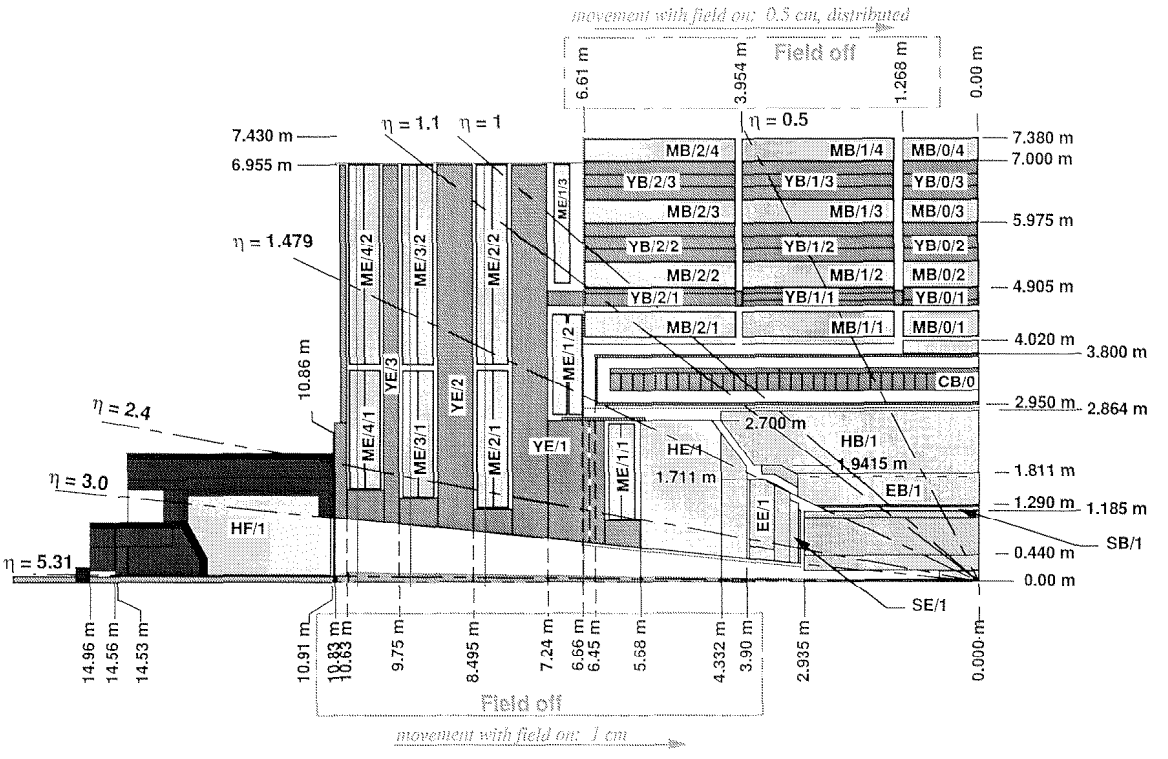


Fig. 2.3: Transverse view of the CMS detector. The four muon stations are indicated as MB and ME both in the barrel and in the forward part. The inner volume is devoted to the calorimeters and the tracker. The hadronic calorimeter as HB, HE and HF whether ECAL as EB and EE. The coverage of the subdetectors is expressed in η units.

Rapidity	Muon p_t threshold
$0.0 < \|\eta\| \leq 1.5$	$4.3\,\mathrm{GeV}/c$
$1.5 < \|\eta\| \leq 2.0$	$3.4\,\mathrm{GeV}/c$
$2.0 < \|\eta\| \leq 2.5$	$2.4\,\mathrm{GeV}/c$

Table 2.2: Muons transverse momentum threshold for 90 % trigger efficiency as a function of pseudorapidity range. The p_t^μ thresholds are evaluated for $\mathcal{L} = 10^{33}\,\mathrm{cm}^{-1}\mathrm{s}^{-1}$.

requiring $p_t^{\mu} > 7 \div 10\,\mathrm{GeV}/c$. At high luminosity ($\mathcal{L} \sim 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$), the di-muon trigger would have a threshold at 10 GeV and the more demanding single muon trigger with a threshold at 20 GeV would result in a 6 kHz first level trigger output rate [8].

2.3.3 The Calorimeter System

The calorimetric system of CMS is made of a crystal electromagnetic calorimeter [11] (ECAL) with an hadronic calorimeter (HCAL) [12] behind it. The primary function of the electromagnetic calorimeter is to measure precisely electrons and photons. In conjunction of the hadronic calorimeter it measures also jets. The ECAL system is made of a high resolution lead tungstate ($PbWO_4$) crystals. In the end-cap region the ECAL extends up to rapidity $|\eta| = 3.0$. This central calorimetric system is complemented in the forward region ($|\eta| < 5.0$) through its very forward extensions (cf. Fig. 2.2 and Fig. 2.3). Their function is to insure detector hermeticity for good missing transverse energy resolution, and to measure forward "tagging" jets signing Higgs production through WW or ZZ fusion. Hermeticity is particularly important for process where the missing E_t is on the order of few tens of GeV as in h, H, $A \to \tau\tau$, $W \to l\nu$, $t \to l\nu b$, $t \to H^{\pm}b \to \tau\nu b$ and for sparticle searches connecting LEP2 and Fermilab with LHC search ranges.

The desired performances, the choice of detection techniques and the design of the ECAL are mostly tuned on the requirements imposed by the $H \to \gamma\gamma$ channel. This is the most appropriate channel to search at a hadron collider for the SM Higgs boson or the lightest MSSM Higgs boson h in the $80 \div 130\,\mathrm{GeV}/c^2$ mass range. The natural width of the Higgs in this mass window is very small ($\ll 1$ GeV) thus the observed signal width is entirely determined by the experimental $\gamma\gamma$ effective mass resolution. This resolution and the level of the large and irreducible $\gamma\gamma$ background will determine the signal significance. To achieve a high mass resolution requires first an excellent electromagnetic energy resolution σ_E/E. However the $\gamma\gamma$ mass resolution depends also on the two-photon angular separation θ. At low luminosity the event vertex is designated by a hard track in the event and the angular term has only a minor effect on the $\gamma\gamma$ mass resolution (a contribution < 200 MeV)

At $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ there are in average 20 minimum bias events superimposed on the triggered $\gamma\gamma$ event. If a mean longitudinal vertex position is used in calculating the $\gamma\gamma$ effective mass, the error introduced in the mass resolution is very large (about 1.5 GeV for $\mathrm{m}_H = 100$ GeV). Thus, the shower position measurement in the calorimeter and in the silicon preshower in front of it insure a directional precision of < 8 mrad which is sufficient to keep the contribution to the mass resolution below 500 MeV. The goal is to have an overall resolution $\sigma_M = 800$ MeV for $m_H = 100\,\mathrm{GeV/c^2}$ at 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

The use of $PbWO_4$ for the crystals is due its short radiation length (0.9 cm) and small Moliere radius (2.0 cm). The low light yield of those crystals can be overcome using silicon avalanche photodiode as readout elements. The arrangements of the crystals is shown in Fig. 2.5. In the barrel the crystal have a length of 23 cm (25.8 X_0 deep) and the lateral granularity is $2.2 \times 2.2\,\mathrm{cm}^2$, i.e. $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$. The total volume of crystals is about 11 m^3 and the total number of crystals channels is ~ 82700.

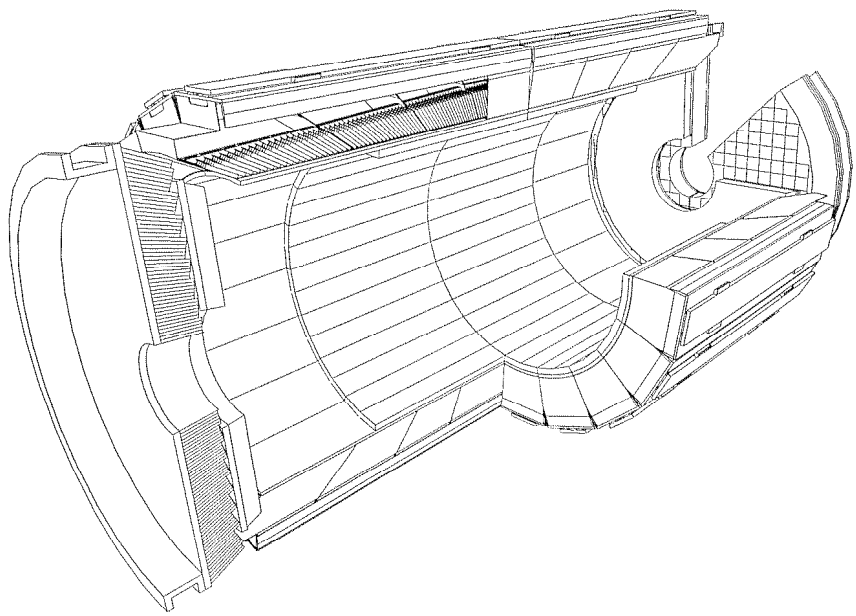


Fig. 2.4: A 3-D view of the electromagnetic calorimeter.

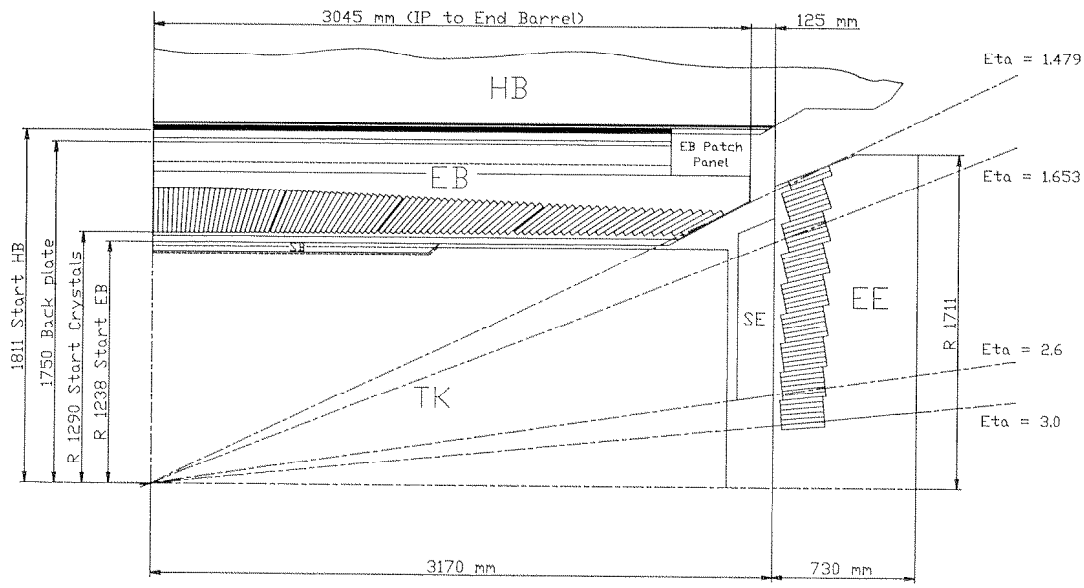


Fig. 2.5: Longitudinal section of one quadrant of ECAL. The arrangement of the crystals is visible as well as their pseudorapidity coverage.

The expected energy resolution for electrons or photons of $E_t = 120\,GeV$ is $\sigma_E/E = 0.6\%$. The in-situ calibration and online monitoring of the calorimeter is insured using E/p matching using isolated electron from W or Z decays measured with $< 1\%$ momentum resolution [8].

Hadron calorimetry with large geometrical coverage for measurement of multi-jet final states and missing transverse energy is essential in both squark and gluino searches, and for detection of the Higgs boson at $m_H > 700\,GeV/c^2$. The hadronic calorimeter plays also a fundamental role in the study of QCD jets as well as in top and τ physics. In CMS, HCAL is subdivided in a barrel and a forward part. Both parts experience the 4 T field of the CMS solenoid and hence require a non-magnetic material (copper alloy and stainless steel). HCAL is a sampling calorimeter: it is made of copper absorber plates interleaved with scintillator tiles readout with embedded wavelength shifting fibres. The minimal sampling thickness in the barrel part will be ~ 6.8 nuclear interaction lengths (λ) at $\eta = 0$. In order to insure adequate sampling depth in the entire pseudorapidity range, $|\eta| < 3.0$, the first absorber layer of the muon chamber is instrumented, as shown in Fig. 2.6. Due to their placement in a very high radiation and rate environment, the HCAL forward extensions use as active medium quartz fibres embedded in a iron absorber volume. This makes their having big sensitivity to Cherenkov light from neutral pions in hadron showers in a very forward region, extremely important for missing energy measurements.

The expected hadronic resolution is $\sigma_E/E = \sim 100\%/\sqrt{E} + 4.5\%$. The tiles are organized in towers, see Fig. 2.6, giving a lateral segmentation of $\Delta\eta \times \Delta\phi \approx 0.09 \times 0.09$, adequate for good di-jet separation and mass resolution. Particularly demanding is the reconstruction of highly boosted W or Z decaying in two jets, produced in decays of a $\sim 1\,TeV/c^2$ mass Higgs.

2.3.4 The Tracker System

The central part of the CMS detector is dedicated to the tracking systems. The advantages and the physics potential of the precise tracking performance provided by pixels and microstrip silicon detectors have been extensively demonstrated both by LEP experiments and CDF. In CMS we aim at a similar sophisticated performance on a larger scale of apparatus and in a much more difficult environment.

The Central Tracker (CT) is radially subdivided in three main subdetectors, namely the Pixel Detector (PD) in the innermost region closer to the interaction region, the Silicon Strip Tracker (SST) in the intermediate one and the Micro-Strip Gas Chamber (MT) in the outer part, close to the calorimeters (see Fig. 2.7).

The detector planes are distributed in the cylindrical tracking volume of CMS with dimensions $|z| < 2.935\,m$ and $R < 1.185\,m$.

A track originated in the interaction region encounters in the barrel part of the tracker, first three layers of pixel detectors providing accuracy of $15\,\mu m$, then 5 layers of microstrip silicon detectors providing points with $15\,\mu m$ precision, followed by 7 layers of $200\,\mu m$ pitch, 125 or 250 mm long gas microstrip chambers (MSGC) giving a measurement with $\sim 50\,\mu m$ precision at normal incidence [13]. In the forward part, the CT end-caps are made

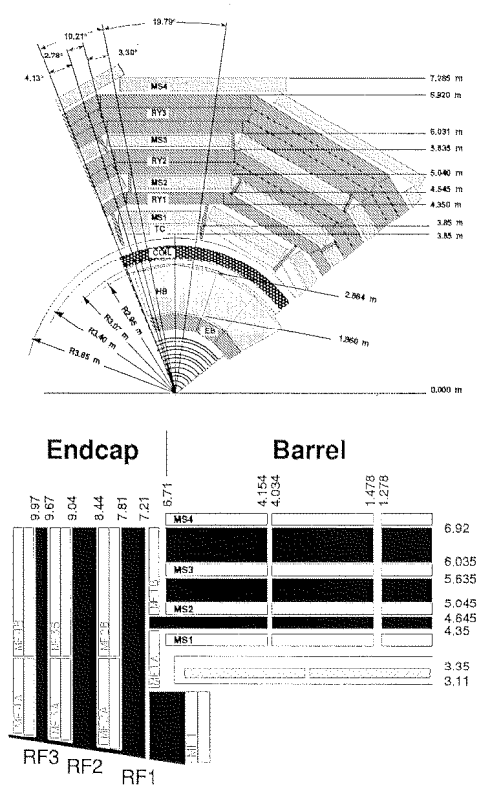


Fig. 2.6: Side view of the tower structure of HCAL for the barrel and end-cap regions.

of three pixel disks, 10 forward silicon microstrip disks and 10 MSGC planes.

The CT is designed to ensure high quality momentum resolution, precise e/γ separation and excellent charged track isolation.

In particular, isolated high p_t muons and electrons should be reconstructed with efficiencies greater than 98%, the fake track rate should be lower than 1% and the momentum resolution should be better than $\Delta p_t/p_t = 0.15 p_t$ (in TeV/c) over the full rapidity range of $|\eta| \leq 2.5$. This will allow measurement of the lepton charge up to $\sim 2\,\mathrm{TeV}/c$ and an in situ calibration of the electromagnetic calorimeter crystals through energy-momentum matching using $Z \to e^+e^-$ and $W \to e\nu$ decays.

Once the tracks have been correctly identified, the measurement of the momentum is simplified by the 4 T magnetic field and the large lever arm of the tracking system. In the case of non-isolated tracks inside jets of transverse energy up to a few hundred GeV, the Central Tracker is required to reach a track finding efficiency of better than 90% for tracks of $p_t \geq 1\,\mathrm{GeV}/c$ and ghosts at the level of $\leq 1\%$.

The pattern recognition at high luminosity requires severe constraints on the granularity of the subdetectors. At 10^{34} cm^{-2}s^{-1}, interesting events will be superimposed on a background of about 500 soft charged tracks from about 15 minimum bias events occurring in the same bunch crossing. Their vertices are distributed along the beam direction with a r.m.s of 5.3 cm. To solve the pattern recognition problem, the CMS approach for the CT is based on very high granularity, strong segmentation in z, high hit efficiency and

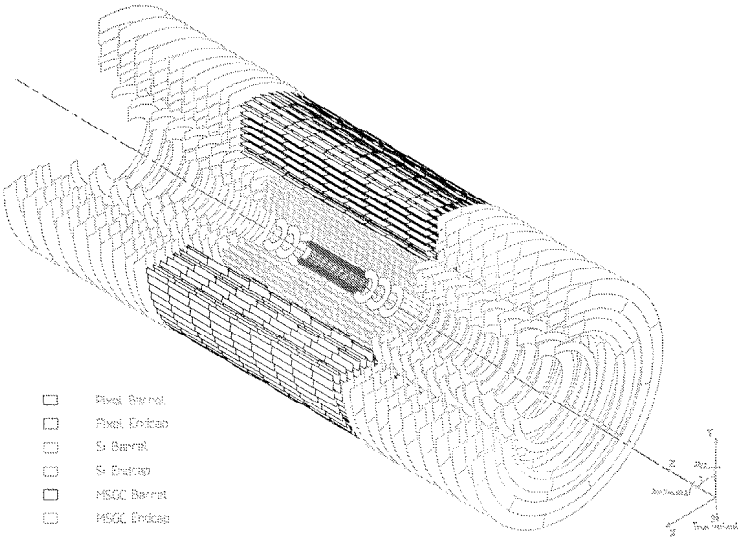


Fig. 2.7: Layout of the CMS Central Tracker. From inside out, resp. the Pixel Detector (PD), the Silicon Strip Tracker (SST) and the Micro Strip Gas Chamber Tracker (MT). In the barrel volume the detectors are arranged in coaxial cylinders, whether in the forward part the detectors are placed in many concentric disks ring shaped.

single-cell occupancy at the level of 1%.

In the CT the combination of pixel and strip detectors will provide the necessary granularity and precision. Pixel detectors provide three dimensional space points close to the interaction vertex and guarantee high precision for secondary vertex reconstruction: having points measured with 15 μm precision in both $r - \phi$ and z so close to the primary vertex significantly reduce combinatorial ambiguities.

Microstrip detectors are the natural choice for the layers following the pixels: high spatial precision and time resolution combined with radiation hardness make silicon strip detectors ideal for the intermediate region. The fast collection time of silicon allows the single bunch-crossing identification. The previous running experience of Silicon Vertex Detectors and a lot of test beam data show that with S/N ratio around 10:1 it is possible to obtain single hit efficiency compatible with 100%.

Extensive simulation work shows that 13 precise hits in the bending plane, reconstructed on average per high p_t track in the full tracking system, are enough to perform good pattern recognition. This fact implies that with 5 layers of silicon in the SST it is possible to achieve good efficiency in linking the hits produced by the Pixel Detector in the vertex region and the hits reconstructed in the less congested region of the outer MSGC layers. Moreover it is possible to perform track segment reconstruction and momentum measurement within the silicon tracker alone: this allow a preliminary SST internal alignment, facilitating then the alignment among the different sub-detectors of the Central Tracker. In any case, the requirements for an effective pattern recognition translates for the SST in a single hit resolution of better than 20 μm and a two track resolution in the inner layers of better than 200 μm. These requirements determine basically the read-out pitch while a maximum strip length of about 12 cm is necessary to maintain the cell occupancy at the 1% level. The combined effect of the requirements

Detector	Channels	Element Size	(r, ϕ) resolution	z resolution
PD Barrel	3.3×10^7	$150 \times 150\,\mu$m	$8 - 10\,\mu$m	$15 - 20\,\mu$m
PD Forward	2.4×10^7	$150 \times 150\,\mu$m	$8 - 10\,\mu$m	$20\,\mu$m
SST (ax)	2.1×10^6 + FD	$61 \div 122\,\mu$m	pitch/$\sqrt{12}$	strip length/$\sqrt{12}$
SST (st)	0.6×10^6 + FD	$52 \div 266\,\mu$m	pitch/$\sqrt{12}$	$\sim 500\,\mu$m
MSGC (ax)	3.5×10^6	$200\,\mu$m	$30 - 40\,\mu$m	strip length/$\sqrt{12}$
MSGC (st)	3.0×10^6	$180 \div 250\,\mu$m	$30 - 40\,\mu$m	2 mm

Table 2.3: Summary of sizes, number of channels and intrinsic resolutions of detectors in the Central Tracker (PD: pixel detector, SST: silicon strip detector, MSGC: microstrip gas chambers).

on resolution and segmentation leads to many thousands of detector modules for PD and SST and $\sim 10^7$ electronic channels.

The presence of b-hadrons may be tagged with reasonable efficiency from displaced vertices produced by their weak decays. Impact parameter measurements play an essential role in B physics and in tagging b-jets in high p_t events. The system of pixel and micro-strip silicon detectors provides vertexing capability in both coordinates not only to distinguish among different interaction vertices at full luminosity but also to provide a precise measurement of the impact parameter of tracks in both coordinates and over a large range of momenta. The goal is 3D vertex reconstruction, a tagging efficiency for b-jets above 20% with a mistagging probability below 1%. Further constraints come from the requirement of K^0 reconstruction at low luminosity with good mass resolution. For tracks with $p_t > 10\,\text{GeV}/c$, the impact parameter resolution in the transverse plane must be better than $40\,\mu$m over the full η-region while the corresponding impact parameter resolution in the r-z plane is required to be better than $100\,\mu$m in the central region.

These requirements on vertexing demand the requested accuracy for the z measurements provided by the silicon layers to be better than 500 μm.

Table 2.3 summarizes the main characteristics of the different detector elements of the central tracker.

2.4 The CMS Silicon Strip Tracker

Once the physical goal of the Silicon Strip Tracker is specified, the definition of the details of the system must be set. This section addresses the constraints due to its integration in the Central Tracker and those imposed by the outer detectors, especially ECAL.

2.4.1 System Constraints

Several experimental constraints must be taken into account in the definition of the design. The most relevant ones are the radiation environment and the limits on the material budget.

2.4.1.a The radiation environment

The most critical issue of the silicon tracker is the long-term survival after heavy irradiation. The system must be designed to guarantee stable operating conditions for several years of running at the highest luminosity.

The levels of radiation due to primary interactions will be very high around the collision region. In addition, a high flux of neutrons will be present in the tracking volume due to albedo neutrons evaporated from nuclear interactions in the material of the electromagnetic calorimeter. Assuming three years of running at low luminosity followed by seven years of high luminosity, it is possible to fix a design value for the radiation resistance of the system based on the flux of radiation intersecting the innermost silicon layer: the detector should maintain reliable performance after a maximum fluence of 5×10^{13} neutrons per cm^2 and 2×10^{14} charged particles per cm^2.

As will be demonstrated in Chapter 3, the design of the micro-strip devices for the SST will be based on the use of single-sided p^+ segmented implants on an initially n-type bulk silicon. This option implies the coupling of two single-sided detectors back-to-back to equip the double-sided layers (cf. Sec. 3.3.6). The use of simple devices offers a great advantage in terms of both cost and industrial production capacity (cf. Sec. 4.12). It does, however, present a significant challenge in that, after the inversion of the bulk induced by radiation, the detectors must be substantially over-depleted for having good performance (see Sec. 3.3.5 and Chapter 5). Since the expected depletion voltage, after type inversion, increases very rapidly as a function of the fluence, the single detector elements and the system as a whole must be designed to allow for high voltage operation of the silicon devices.

Managing the high radiation environment of the LHC will require the silicon wafers to be kept cold. The silicon operating temperature becomes increasingly important as the effect of the radiation becomes more pronounced (cf. Sec. 3.3.5). The entire volume of the silicon tracker must be permanently kept well below 0 °C during running and only for limited periods of time it will be allowed to reach room temperature for maintenance purposes.

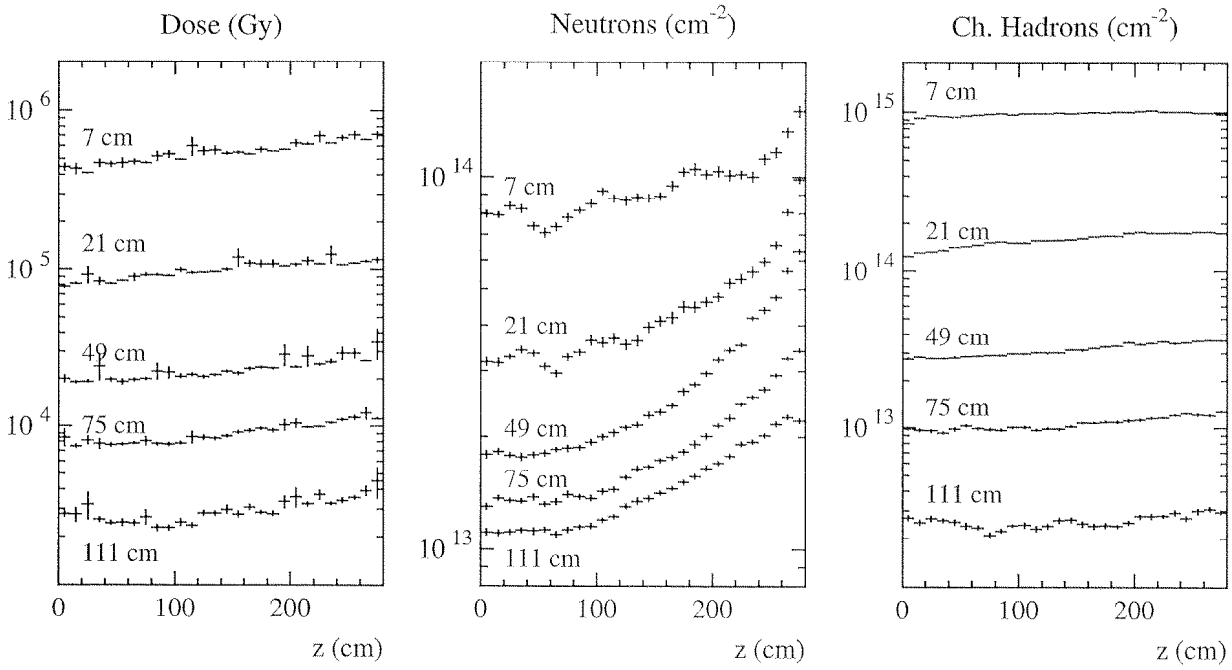


Fig. 2.8: Neutron and charged hadron fluxes along the beam direction and at different radii in the tracking cavity [8].

2.4.1.b Material Budget

The amount of material within the tracker volume will affect the tracker performance in various ways due, i.e. multiple scattering, delta rays, photon conversions and strong interactions. Even more critical is the effect of the tracker material on the ECAL performance. Strong emphasis has been put in the CMS design to allow efficient reconstruction of the $H \rightarrow \gamma\gamma$ decay mode. This places stringent requirements on the performance of ECAL, which in turn seriously constrains the material of the tracker. In order to limit the fraction of photons from Higgs decay converting within the tracking volume to less than 45%, the tracker must have a minimum radiation length compatible with a robust and stable construction. This constraint reflects in a limit of 40% of radiation length for the SST over a significant η range, see Fig 2.9.

2.4.2 System Characteristics

In Chapter 3 the choice of detectors are described in detail. For the purposes of this section it is important to note that the constraints imposed by radiation resistance, easier handling and overall system costs, strongly suggest the option of single–sided detectors. This is a major decision. In fact, once single–sided detectors are chosen, the progress in the definition of the readout electronics, the support structures and the cooling can be done consistently: the optimization of the design or the technology can be postponed in a later stage. Pros and cons of this choice will be summarized in Chapter 3.3.7.

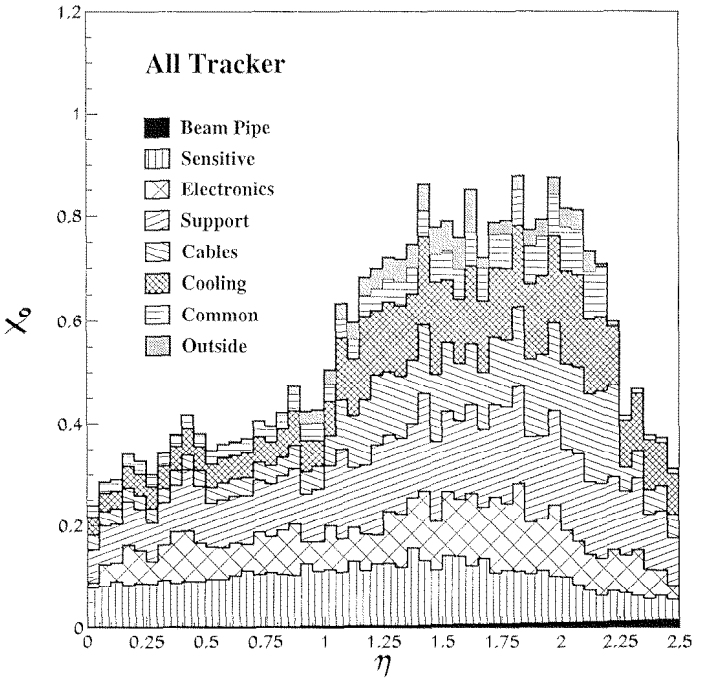


Fig. 2.9: Material in terms of X_0 as a function of pseudorapidity for SST and PD. [8].

2.4.2.a Read-out electronics

In accordance with the single–sided device option, a fully analog architecture based on a fast amplifier followed by an analog pipeline and a deconvolution circuit has been chosen. The analog read-out allows an efficient common mode noise subtraction. In addition, it allows a better control of the behaviour of the detectors, giving larger margins for understanding and following the change of the system performance induced by irradiation. Diagnostics of the failures will be easier and a fine tuning of the on-line and off-line thresholds will be possible to obtain the best performance for different running conditions in different layers. Multi-layer aluminised kapton hybrids will be used to reduce the overall material budget.

High speed optical links, based on laser technology, will feature radiation hardness, small size, low power dissipation and high analogue transmission quality.

2.4.2.b Detector Modules

The choice of very simple detector modules allow easy assembly and easy replacement. With 12.5 cm as maximum strip length both for the barrel and forward modules one expects to maintain an average signal-to-noise ratio of 12:1 at the end of the detector lifetime (10 years of LHC running).

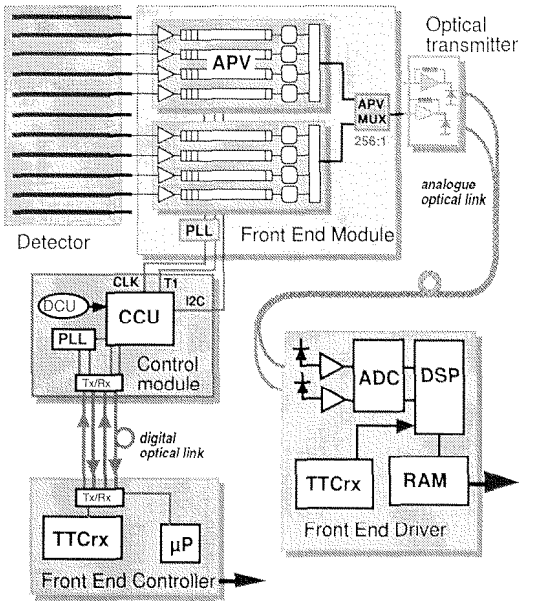


Fig. 2.10: Readout electronics scheme.

2.4.2.c Mechanical structure

Low mass supporting structures are foreseen. The mechanics will incorporate the cooling system, cables and optical links and all other services (alignment, monitoring and survey systems). High performance carbon fibre composites are foreseen as basic material of the mechanical structure. Since a temperature range of about 40 °C is expected in between the assembly and running temperature of the most critical parts the structure will be fabricated with the constraint of having a thermal expansion coefficient (CTE) close to zero. Symmetric structures will be used for all mechanical elements incorporating materials with CTE different from zero (i.e. the detector modules) to avoid bi-metallic effects.

2.4.2.d Cooling

The operating temperature of the silicon tracker will be -10 °C. The cooling system will be based on fluids used to cool down by direct conduction of all elements producing power (electronics, optical links and service cards; cables and silicon detectors) and on cold dry nitrogen flowing through the system. A thermal shield will isolate the volume occupied by silicon and pixel detectors from the volume of the MSGC running at +18 °C. Small, thin walled metal pipes will be used for the cooling system.

2.4.2.e Alignment and mechanical stability

The final position of each detector strip in space will be determined by using tracks. Survey data will be used to constrain the initial information and to speed up the iterative procedure. A monitoring system will be used to control the stability with time of the

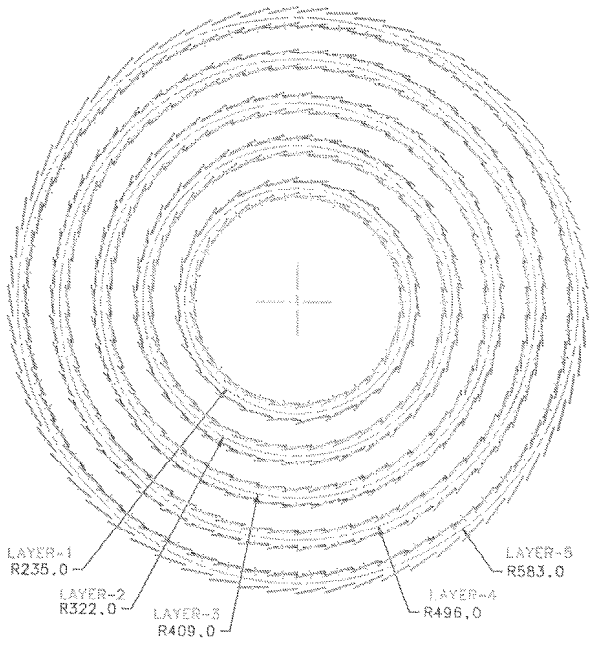


Fig. 2.11: Transverse view of the silicon barrel detector.

detector.

The single detector modules will be fabricated with an internal precision of 5 μm in the detector plane and of about 30 μm in the coordinate perpendicular to the plane. The mechanical structure will allow a survey procedure during the assembly of the modules to know the absolute initial position of each strip in space with an accuracy better than 110 μm. Maximum deviations from the nominal positions due to the combined effects of machining precision, gravity and thermal deformations will be kept at the level of tens of microns. Stability with time under stable operating conditions will be at the level of 10 microns.

2.4.3 System Organization

The Silicon Strip Tracker (SST) consists of approximately 75 m^2 of instrumented microstrip silicon detectors. It is composed of a barrel region consisting of five equally spaced cylindrical layers and six mini-disks and a forward/backward region each one composed of ten disks (cf. Fig. 2.12).

The silicon sensors are organized in detector modules formed by one or two detectors for an overall strip length between 6 and 12.5 cm. Detectors and on-board electronics are assembled on low mass supporting elements. The detector modules are mounted on very light and stiff supporting structures incorporating cables, cooling system and all other services. In the following paragraphs is given a short description of each part of the SST.

The Barrel Part

The barrel part instruments the radial region between 21 and 63.5 cm (cf. Fig. 2.11). The length of the entire system is 173 cm. All cylindrical layers are equipped with rectangular detectors and provide axial information; in addition layers 1-2 and 5 provide also small-angle stereo measurements. Three small disks on each side complement the first two barrel layers and are equipped with wedge-shaped detectors and provide a ϕ and a radial measurement.The use of this geometry reduces the number of detector modules while optimizing the detector performance for very inclined tracks.

Tracks at $|\eta| \leq 1.2$ are fully contained in the barrel detector and cross only cylindrical layers. For $|\eta| > 1.2$ tracks cross both the cylindrical layers and the mini-end-cap disks. The pseudo-rapidity coverage of the last disk of the mini-end-cap extends down to $|\eta| \leq 2$. Table 2.4 lists the pitch-size and the number of channels per module foreseen for each layer. The strips in the stereo view are tilted by 100 mrad with respect to the beam line, thus allowing a measurement of the z-coordinate.

layer #	r-ϕ read-out pitch [μm]	# of channels per module	stereo read-out pitch [μm]	# of channels per module
1	61	1024	122	512
2	81	768	122	512
3	81	768	-	-
4	122	512	-	-
5	122	512	244	256

Table 2.4: Read-out pitch and number of channels per module in the barrel detector. (Each time there are 128 channels per chip).

The dimension of the rectangular detectors are set to $64 \times 64\,\mathrm{mm}^2$ in order to use all the available surface on a 4″ wafer. The detectors are tilted by 9° to compensate for the Lorentz angle and are arranged to allow ϕ and z overlap to avoid dead regions. This overlap also takes into account the spread of the primary vertex. The structure is slightly asymmetric in z to allow active detector elements for perpendicular tracks at $\eta = 0$. The acceptance of each detector layer is listed in Table 2.5 (due to the overlap, acceptance numbers $> 100\%$ are listed).

A total number of 1372 single-sided and 1360 double-sided modules will be used to equip the cylindrical layers of the barrel detector. 8200 silicon detectors will be needed. In total, an area of $33.6\,\mathrm{m}^2$ of silicon sensors and $\sim 2.2 \cdot 10^6$ analog channels have to be produced.

The Barrel mini-end-caps

The so called "mini–end-caps" have been introduced in order to optimize the hermeticity of the barrel part of the SST keeping the number of detectors as small as possible. Each disk is an independent unit containing two rings of detectors. In each disk the two rings

	layer 1	layer 2	layer 3	layer 4	layer 5	total
r_{in} [mm]	217	304.5	391.5	479	566	
r_{out} [mm]	249	336.5	423.5	511	598	
N_{mod}/tot	208 (ds)	288 (ds)	616 (ss)	756 (ss)	868 (ds)	1364 (ds)
						1372 (ss)
N_{waf}/mod	4	4	2	2	4	
N_{waf}/tot	416 (a)	576 (a)	1232 (a)	1512 (a)	1736 (a)	5472 (a)
	416 (s)	576 (s)	-	-	1736 (s)	2728 (s)
$N_{ch}(\phi)$	212608	221184	473088	387072	444416	1738368
$N_{ch}(z)$	106496	147456	-	-	222208	476160
Acc.(ϕ)[%]	116	115	110	110	107	

Table 2.5: Radial positions, number of modules and detectors and geometrical acceptance (Acc.) of each layer for the barrel silicon tracker [14]. (ss) Single Sided detector, (ds) Double sided detector, (a): axial geometry, (s) stereo geometry, (ch) readout channel

of detectors are positioned on the two sides of the supporting structure. This allows free space for the optical links, services and connections.

The basic detector element of the mini–disks is a pair of wedge shaped dctectors, coupled back-to-back. In the front device, radial strips point to the beam–line (ϕ-view). In the device on the back, the strips are tilted by 100 mrad (stereo-view), thus pointing to a concentric ring around the beam line.

The most important parameters of the mini–end-cap detectors are listed in Table 2.6.

	disk $\pm$ 1	disk $\pm$ 2	disk $\pm$ 3	Total
r_{in} [mm]	218	218	218	
r_{out} [mm]	360	360	360	
z-coord [mm]	$\pm$ 523.77	$\pm$ 657.54	$\pm$ 758.11	
$N_{mod}(\phi)$	60	60	60	360 (DS)
N_{waf}/mod	6	6	6	
N_{waf}/tot	180	180	180	1080
$N_{ch}(a)(\phi)$	53760	53760	53760	322560
$N_{ch}(s)$	15360	15360	15360	92160

Table 2.6: Radial positions and relevant numbers of the barrel mini-disk detectors. (a) axial geometry, (s) stereo geometry, (ch) readout channel.

The 6 disks of the mini end-cap contain 1080 silicon detectors (540 with radial strips and the other half with stereo strips and double-metal connections) resulting in a total area of 4.1 m^2 and 415000 analog channels.

The grand total for the barrel detector amounts to 9280 silicon sensors (37.7 m^2) and $\sim 2.6 \cdot 10^6$ electronic channels.

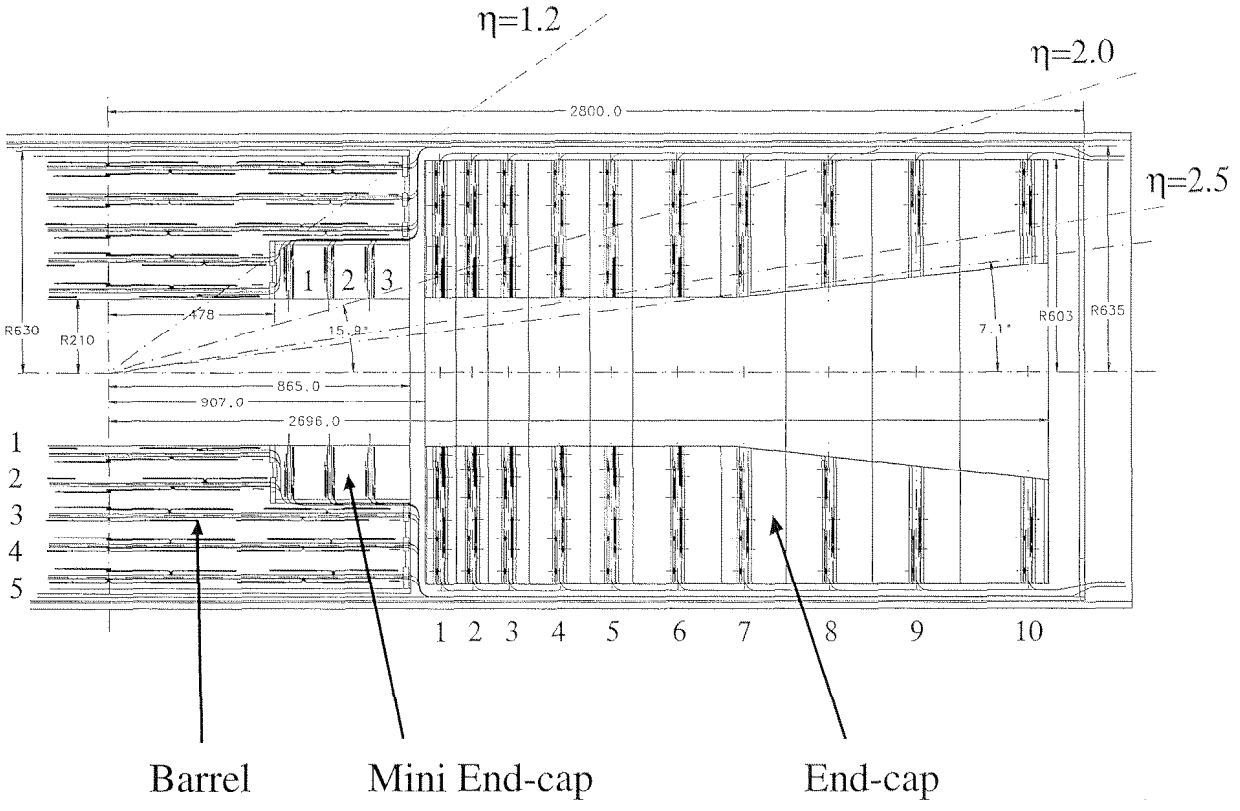


Fig. 2.12: Side view of the silicon barrel detector, including the mini end-caps.

The Forward/Backward Part

The forward/backward parts reconstruct tracks in the radial region between 22 and 59 cm. Each of the twenty disks is made out of concentric rings of wedge-shaped detectors. On each side disks 1 to 6 contain four rings, disks 7-8 three rings and disks 9-10 two rings only. Rings 1 and 4 provide double-sided information while disks 2 and 3 are single-sided.

The z positions of the disks have been defined to guarantee full coverage and to optimize the track reconstruction performance over the full pseudo-rapidity region covered by the forward-backward detector: $1.2 \leq |\eta| \leq 2.5$.

The SST extends longitudinally for about 5.5 m. Each silicon detector is read-out by front-end electronics on a pitch, varying for different layers and disks, between 52 and 266 μm.

The detectors in each disk are organized in rings. Table 2.7 shows their dimensions and characteristics.

The detector modules are assembled on both sides of the supporting structure of each disk. Overlap in both coordinates is obtained by mounting adjacent modules which are slightly offset in z.

The forward-backward detector contains 3600 detector modules for a total number of 9600 detectors split into 7 different designs. The overall silicon area needed for the forward-backward detector is 37.5 m^2.

ring type	r_{min} mm	r_{max} mm	detector type	w_{min} mm	w_{max} mm	h mm	minimum pitch μm	maximum pitch μm
1	220	300	double-sided	42	56	80	52	70
2a	290	400	single-sided	56	68	70	106	141
2b	290	400	single-sided	68	74	33	106	141
3a	390	490	single-sided	54	64	73	101	121
3b	390	490	single-sided	64	68	35	101	121
4a	480	590	double-sided	56	63	73	105	133
4b	480	590	double-sided	63	70	66	210	266

Table 2.7: Dimensions of the end-cap rings.

Chapter 3

Conception: the Silicon Microstrip Detector

A short review of the energy loss mechanism on silicon introduces to the problem of the efficiency of a silicon detector: both energy and position resolution depend on how well the detector is suited to respond to ionisation. For the specific application of silicon detectors described in this thesis, most of the particles crossing the detectors are minimum ionizing particles and the goal will be to reconstruct their passage through the system with the best precision possible. The objective can be attained only if the entire chain of detector, amplifier and readout system is able to keep the signal well above the intrinsic noise level. In this chapter, after a brief summary of semiconductor physics with particular attention to phenomena directly affecting the silicon detectors, the most important characteristics related to the S/N ratio are studied: design details, leakage current, high voltage operation, capacitive coupling and biasing methods are reviewed in order to optimize the CMS microstrip detector for safe operation in the LHC environment.

3.1 Energy Loss of High Energy Charged Particles in Silicon

High energy charged particles traversing crystalline silicon can lose energy either by ionisation of the atomic electrons or by the non-ionizing energy loss (NIEL), i.e. displacement of silicon atoms from the crystal lattice, mostly by the process of Coulomb nuclear scattering. The atom will be displaced from its lattice site only if the energy transfer to the nucleus exceeds approximately 25 eV. Below that, the energy is dissipated by lattice vibrations.

Displacement of silicon atoms to interstitial positions is one of the main radiation damage mechanisms. For a high energy particle, the fraction of energy loss going into the NIEL mechanism is relatively small, but the cumulative effects on the detector performance can be severe. A detector placed in a neutron flux experiences no signal from primary ionisation, but the interactions can cause a high level of NIEL in view of the large

neutron-silicon scattering cross-section. For both charged hadrons and neutrons, other mechanisms of energy loss and radiation damage exist, notably neutron capture followed by nuclear decay, and inelastic nuclear scattering. In this section, the attention is set on the ionisation energy loss [15], whereas non-ionizing energy loss and more generally radiation induced mechanisms will be analyzed in Section 3.3.5.

3.1.1 Rutherford scattering and Bethe-Bloch equation

If all the atomic electrons were free, as if the crystal consisted of the silicon nuclei neutralized electrically by a homogeneous electron plasma, when a charged particle traverses the material, it loses energy by collisions (Coulomb scattering) with the electrons. Close collisions, while rare, will result in large energy transfers, while the much more probable distant collisions give small energy transfers.

The process can be thought of classically in terms of the impulse generated by the attractive or repulsive Coulomb interaction between the projectile and the electron.

The greater probability of remote collisions arises simply from the greater volume of material available for collisions with a given impact parameter range, as the corresponding cylinder (of radius equal to the impact parameter) expands.

In this simple case, the probability for a collision to impart an energy E to an atomic electron is given by the Rutherford cross-section:

$$\frac{d\sigma_R}{dE} = \frac{2\pi q^4}{m\beta^2 c^2} \frac{1}{E^2} \tag{3.1}$$

where q and m are the charge and mass of the electron.

The fact that the mass of the struck particle is in the denominator explains why scattering off the silicon nuclei, which are much more massive, causes very little energy loss, even if these collisions cause the major contribution to the multiple scattering. Moreover, for sufficiently large momentum transfers, these nuclear collisions contribute to the NIEL referred to above.

The relevant information are the mean energy loss and also the fluctuations, for traversal of a given thickness detector. However, the integral of Eq. (3.1) over all E diverges like $1/E$. The stopping power of this free-electron plasma would indeed be infinite, due to the long-range Coulomb interaction. In practice, the electrons are bound and this prevents very low energy transfers to the vast number of electrons which are distant from the particle trajectory.

This divergence is conventionally avoided by introducing a semi-empirical cutoff (binding energy) E_{min} which depends on the atomic number Z of the material. This is necessarily an approximate approach, since it ignores the fact that the outer electrons are bound differently in gaseous media than they are in solids. However, a slightly corrected Rutherford formula is extremely useful as regards the close collisions, which are most important in defining the fluctuations in energy loss in 'thick' samples (greater than approximately 50 μm of silicon, for example). The required correction is the upper cutoff

E_{max} in energy transfer imposed by the relativistic kinematics of the collision process. If the projectile mass is much greater than m_e, we have $E_{max} = 2mc^2\beta^2\gamma^2$.

Keeping into account the quantum mechanical nature of particles and energy transfers, Bethe and Bloch [16, 17] modified the semiclassical expression of the specific energy loss, so that for a heavy, spin 0, incident particle is

$$-\frac{dE}{dx} = \frac{4\pi n_e Z^2 q^4}{m\beta^2 c^2}\left(\ln\frac{2m\beta^2\gamma^2c^2}{I} - \beta^2 - \frac{\delta(\gamma)}{2}\right) \tag{3.2}$$

where n_e is the electron density per unit volume, q and m are charge and mass of the electron, Z and β are charge and velocity of the incident particle, I is the mean ionisation potential for the medium and δ is a correction function which takes into account the dielectric screening in dense material, ("density effect") [18], responsible for the decrease of energy loss after the relativistic rise [1].

The most important features of Eq. (3.2) are its independence from the mass of the incident particle and its material dependence which is linear through n_e and logarithmic through the mean ionisation potential I. In Table 3.1 are resumed the most important electromagnetical properties of Silicon.

Z	n_a	n_e	I (eV)	X_0 (g/cm^2)	Density (g/cm^3)
14	$0.5 \cdot 10^{23}$	6.99	173	21.82	2.33

Table 3.1: Electromagnetic properties of silicon [20].

It is clear that for low momentum particles the dE/dx curve show the $1/v^2$ drop up to a broad minimum for $\beta\gamma \simeq 3 \div 3.5$. As β continues to increase the $\ln\gamma^2$ factor begins to dominate and dE/dx starts to increase (relativistic rise) up to a point where the screening of the atomic electrons reduce the increase from $\propto \ln\gamma^2$ to $\propto \ln\gamma$, flattening the energy loss for very high γ to a constant value, also referred as the Fermi plateau. Most of the relativistic particles have energy loss close to the minimum, and are said minimum ionizing particles, or mip. The Eq. (3.2) can be integrated to find the total range R of a particle loosing energy only through ionisation. It is clear that most of the ionisation is done when the particle slows down and it is almost stopped within the medium, i.e. at very low β. In Fig. 3.1 the range for different particles in Silicon is shown as a function of the particle energy.

The amount of energy lost by a charged particle that has traversed a fixed thickness of absorber will vary due to the statistical nature of its interactions with individual atoms in the material. Therefore the dE/dx should be intended as an averaged value.

The Poisson statistics on energy transfers in this range gives rise to fluctuations on the overall energy loss for each traversal. Thus the overall energy loss distribution consists of an approximately Gaussian core plus a high tail, populated by traversals for which a few close collisions occurred, each generating several times the mean energy loss. Due to the $1/E^2$ term in the Rutherford formula, for each sample thickness, there is an energy

[1] For a more detailed description see [19] and related references.

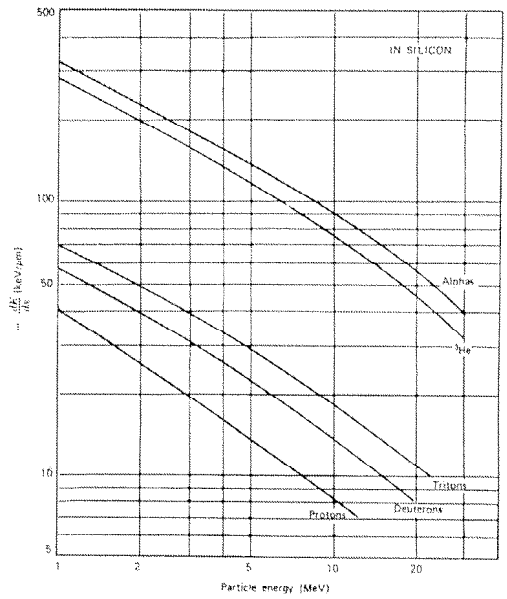


Fig. 3.1: Range-energy curves for different charged particles in Silicon [21].

transfer range in which the integrated probability of such transfers through the sample falls from almost unity to nearly zero. While the energy transfer region in which the probability function falls almost to zero is dependent on the sample thickness, this merely introduces an overall scale factor, so the form of the overall energy loss distribution is constant (the Landau-Vavilov distribution) over a wide range of detector thicknesses.

The rare close collisions with energy transfer greater than approximately 10 keV generate δ-electrons of significant range, which will release electron-hole pairs along their path with a rate of approximately one per 3.6 eV of energy loss[2].

This effect may be important in tracking detectors due to its potential for degrading the precision.

For these close collisions, all atomic electrons behave as if they are free and the Rutherford formula may be used with confidence. For thin samples, the energy loss fluctuations are not adequately handled by the Rutherford formula with cutoffs E_{min} and E_{max}. In this case, the bulk of the energy loss arises from low energy transfer collisions for which the binding of the atomic electrons must be handled in detail.

Qualitatively, the effect of the binding of the atomic electrons is to generate resonance-like enhancements in the energy loss cross-section, above the values expected from the Rutherford formula, see Fig. 3.2. The K-shell electrons produce an enhancement in the 2 to 10 keV range, the L-shell in the 100 eV to 1 keV range, and the M-shell a resonance at around 20 eV. Below this resonance, the cross-section rapidly falls to zero, in the region around 15 eV where the Rutherford formula would be cut off by the empirical ionisation

[2]It is interesting to note that the electron-hole generation is an "inefficient" process, since the silicon bandgap at ambient temperature is close to 1.1 eV: approximately 2/3 of the energy transferred from the hot electrons gives rise to phonon generation.

threshold energy. The most satisfactory modern treatment proceeds from the energy-dependent photo-absorption cross-section. This is of closely linked to the energy loss process for charged particles, which fundamentally proceeds via the exchange of virtual photons. Combining photo-absorption and EELS (electron energy loss spectroscopy) data, Bichsel [22] has made a precise determination of the mip energy loss cross-section for silicon. The most subtle effects are connected with the valence (M-shell) electrons. These behave as a nearly homogeneous dense gas (plasma) embedded in a fixed positive charge distribution. The real or virtual photons couple to this by generating longitudinal density oscillations, the quantum of which is called a plasmon and has a mean energy of 17 eV. The plasmons deexcite almost entirely by electron-hole pair creation. These somewhat energetic charge carriers are referred to as 'hot carriers'. Like the δ-electrons produced in the close collisions, they lose energy by thermal scattering, optical phonon scattering and ionisation.

The plasmon excitation is responsible for the extremely large cross-section in the UV region followed by a long tail for lower energies, allowing therefore silicon to be optically sensitive even in the visible region.

The closely related energy loss cross-section for a MIP is shown in fig. 3.2.

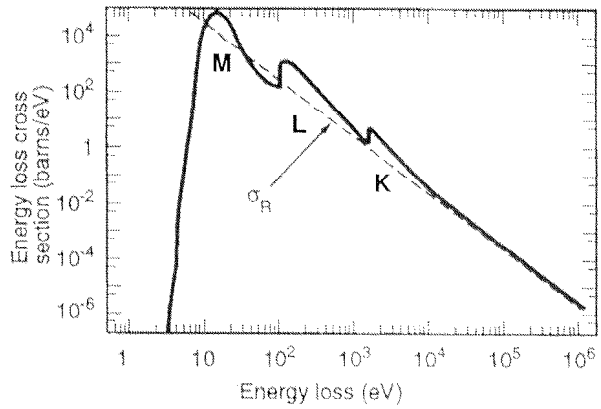


Fig. 3.2: Energy loss cross-section for MIPs in silicon, versus energy loss in primary collisions. Rutherford cross-section σ_R is superimposed.

Note that on average it exceeds the Rutherford cross-section in the energy range 10 eV to 5 keV. Above 10 keV, it is very close to the Rutherford value. By integrating the different components of this cross-section, it is possible to deduce the total mean collision rates associated with the different processes (see Tab. 3.2). Thus, despite the fact that on average a slice of silicon 1 μm in thickness will yield 80 electron-hole pairs, the Poisson statistics on the primary process (on average 3.8 collisions per micron) clearly implies a very broad distribution, with even non-negligible probability of zero collisions, i.e. absolutely no signal. For thin samples, correct statistical treatment of the primary process is essential if realistic energy loss (straggling) distributions are to be calculated.

Si shells	σ_e
$K(2)$	8.8×10^{-3}
$L(8)$	0.63
$M(4)$	3.2

Table 3.2: Electron collision probability σ_e per μm as a function of electronic shells in silicon

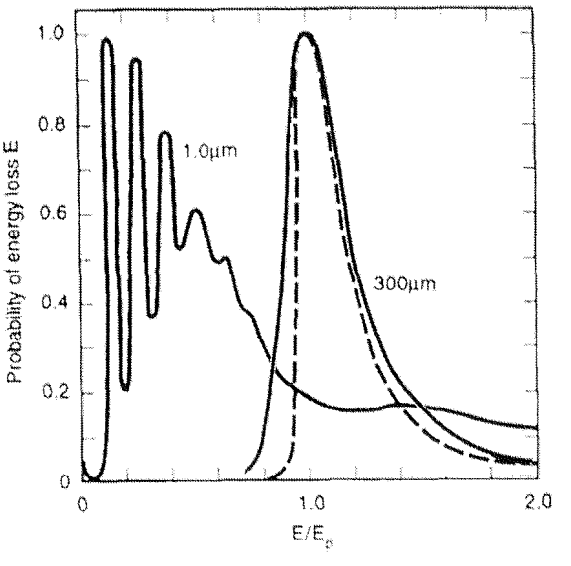


Fig. 3.3: Energy loss distributions for 1 μm and 300 μm silicon detector thicknesses. The Landau distribution is shown as a reference [22].

Implications for Tracking Detectors For high precision tracking, there are clear advantages in keeping the silicon detector as thin as possible. A thin detector is optimal as regards multiple scattering. A detector with the thinnest possible active region (which may be significantly less than the physical thickness) is optimal as regards point measurement precision, for two distinct reasons.

1. For non-orthogonal incidence tracks, the concern arises from δ-electrons of sufficient range to pull the centroid of the charge deposition significantly off the track. The range becomes significant for high-precision trackers for E greater than approximately 10 keV, for which the generation probability is less than 10% per micron. Thus, a detector of thickness 10 μm is much less likely to yield a 'bad' coordinate than one of thickness 300 μm . If the magnitude of the energy deposition in the detector can be measured some of the bad coordinates will be apparent by the abnormally large associated energy. They could then be eliminated by a cut on the energy deposit, but this usually leads to unacceptable inefficiency and is rarely implemented.

2. The second reason for preferring detectors to be as thin as possible applies to the case of angled tracks. In principle it may be possible to infer the position of such a track by measuring the entry and exit points in the detector, but more usually the best one can do is to measure the centroid of the elongated charge distribution, and take this to represent the track position as it traversed the detector mid-plane. In this case, large fluctuations in the energy loss (due to ejection of K- and L-shell electrons and δ-electrons) may be sufficient to cause serious track pulls for thick detectors. In the thin detector there is a 10% probability of producing a δ-electron which, if it occurs near one end of the track, pulls the co-ordinate from its true position by 4 μm . In the thick detector, there is the same probability of producing a δ-electron which can pull the co-ordinate by 87 μm .

However, the requirement of high detection efficiency suggest to come to a compromise. In Fig. 3.3 are shown the energy loss distributions for very thin (e.g. 1 μm Si) and "normal" detector thicknesses. It is clear that in the first case the energy loss distribution is very broad, with peaks corresponding to 0, 1, 2,... plasmons excited, followed by a long tail extending to very large energy losses. An efficient tracking detector could never be built with such an active layer. At 300 μm , the Landau distribution gives an adequate representation, thus strongly reducing the tails at lower energy.

Therefore, while very thin detectors are ideal from the viewpoint of tracking precision, great care must be taken to assure that system noise allows a sufficiently low threshold to achieve the desired detector efficiency.

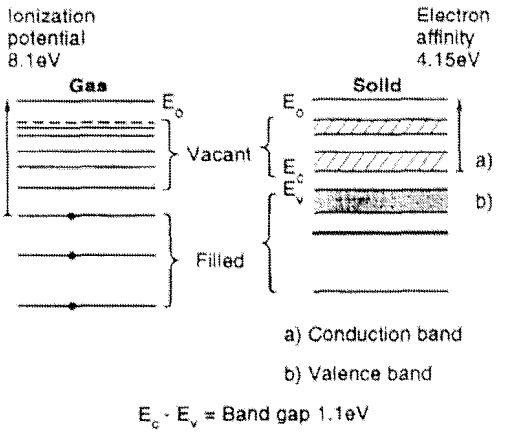


Fig. 3.4: Gaseous silicon allowed energy levels compared with energy bands in the solid material.

3.2 Physics and Properties of Silicon Devices

Electrons under the influence of the Coulomb potential of an atomic nucleus have only certain allowed levels. As more atoms are added to form a crystalline structure, the discrete energy levels of the individual atoms merge into a series of energy bands in which the individual states are so closely spaced to be essentially continuous. These allowed bands are bounded by maximum and minimum energies and they may be separated from adjacent allowed bands by forbidden-energy gaps or they may overlap other bands. The detailed behaviour of the bands fundamentally determines the electronic properties of the material: it is the essential feature differentiating conductors (no gap), insulators (big gap) and semiconductors (small gap). In Fig. 3.4 are shown the atomic energy levels and the energy bands for gaseous and solid silicon.

The levels previously occupied by valence electrons are merged in a valence band and those previously unfilled become the conduction band. Due to its original atomic energy level structure, silicon at absolute zero temperature has a small forbidden energy band gap ($\sim$ 1.1 eV), typical of semiconductor materials. It should be noted here that the large band gap in silicon dioxide is 9 eV, much larger than typical thermal or field imparted energies, thus making it an excellent insulator and, coupled with the ease with which the surface of the silicon can be oxidised in a controlled manner, accounts partially for the pre-eminence of silicon in producing electronic devices. At any temperature above 0 °K the valence band is not entirely filled because of small number of electrons possess enough thermal energy to be exited across the forbidden gap into the next allowed band, leaving an equal amount of vacant states (holes) in the valence band. The electrons in the conduction band can easily gain small amounts of energy and can respond to an applied electric field thus producing a current flow. At the same time in the valence band there is a re-arrangement of states which can be pictured as the backward movement (with respect to electrons) of holes in response of the applied field. A material in which the number of electrons in the conduction band n equals the number of holes in the valence band p

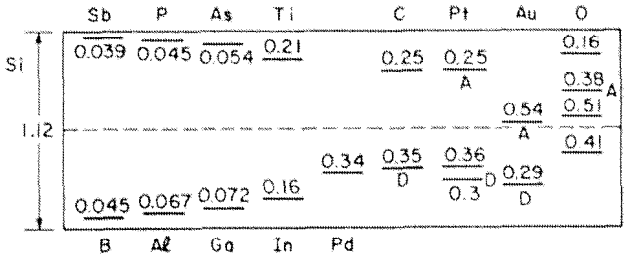


Fig. 3.5: Measured ionisation energies from various impurities in Silicon. The levels above (below) mid-gap are measured from the bottom (top) of the conduction (valence) band and are donor (acceptor) levels unless indicated by A for acceptor level (by D for donor level). From [23].

is called *intrinsic* semiconductor. When a semiconductor is doped with impurities, the semiconductor becomes *extrinsic* and impurity energy levels are introduced. The most useful means for controlling the number of carriers is by incorporating substitutional impurities, that is, replacing a silicon atom with an impurity atom with, for example, five valence electrons (like phosphorus): four impurity electrons will fill the covalent bonds between impurity atom and neighbors silicon atoms while the fifth electron is only weakly bound to the impurity atom by the excess of positive charge on the nucleus. Only a small amount of energy is sufficient to free this electron so that it can wander about the crystal and contribute to electrical conduction. Since the substitutional group V impurities donate electrons to the silicon which becomes $n-$type, they are known as *donors.* If most of the impurities are of the donor type, the number of electron in the conduction band is much greater than the number of holes in the valence band. Electrons are then called *majority* carriers, and holes are called *minority* carriers. Analogous arguments can be applied to Group III impurities (like Boron) in which one covalent bond per atom is left unfilled thus causing a hole conduction due to internal lattice's electron rearrangements. The silicon becomes $p-$type and these impurities atoms are called *acceptors.* In Fig. 3.5 are shown the measured ionisation energy of various impurities in Silicon.

It is clear that for shallow donors, there usually is enough thermal energy to supply the energy E_D to ionize all donor impurities at room temperature and thus provide an equal number of electrons in the conduction band. This condition is called *complete ionisation.*

Thermal equilibrium analysis. The free-carrier densities in semiconductors are related to the populations of allowed states in the conduction and valence bands. The densities are dependent upon the net energy in the crystal which is stored in crystal-lattice vibrations (phonons) as well as by the electrons. Although a semiconductor crystal may be excited by external sources such as incident photoelectric radiation, there exist many situations where the total energy is function only of crystal temperature. In this case the semiconductor will spontaneously reach a state known as *thermal equilibrium*, a dynamic situation in which every process is balanced by its inverse process. Most of the silicon properties will be described at the thermal equilibrium conditions unless stated differently.

In an intrinsic material all carriers result from thermal excitation across the forbidden band gap, so then the number of electrons in the conduction band n will be equal to

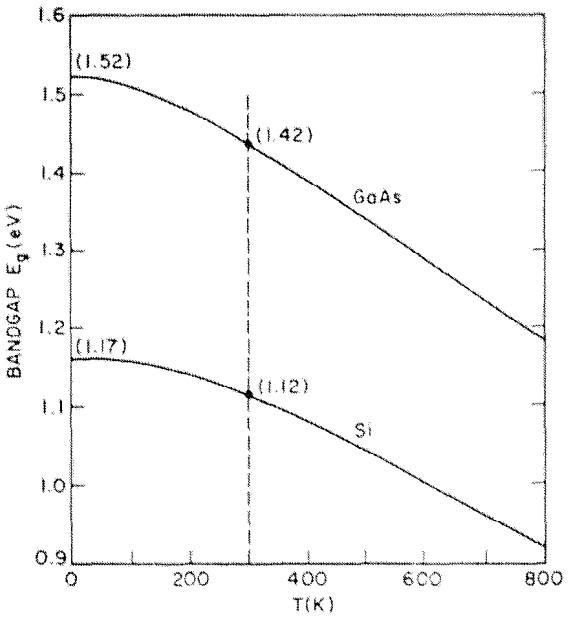


Fig. 3.6: Silicon energy bandgap as function of temperature.

the number of vacancies (holes) in the valence band p, $n = p = n_i$, where the subscript i reminds the use of intrinsic material. The intrinsic concentration of carriers clearly depends on temperature and is also function of the energy band gap. In Fig. 3.6 and Fig. 3.7 both dependencies are highlighted in the case of Silicon.

A first important consequence of thermal equilibrium is expressed in the so called *mass-action* law which holds for both intrinsic and extrinsic semiconductors:

$$np = n_i^2. \tag{3.3}$$

It shows that increasing the number of electrons in a sample by adding donors causes the hole concentration to decrease, since the intrinsic carrier concentration is constant (at room temperature for silicon $n_i = 1.45 \times 10^{10}\text{cm}^{-3}$).

The electron density in the conduction band in an intrinsic semiconductor is given by

$$n = \int_{E_c}^{E_t} N_E(E) f_D(E) dE \tag{3.4}$$

E_c and E_t are the bottom and the top of the conduction band, N_E is the density of allowed states per unit volume; $f_D(E)$ express the probability that a state of energy E is occupied by an electron and is expressed by the Fermi-Dirac distribution function

$$f_D(E) = \frac{1}{1 + \exp{(E - E_f)/kT}} \tag{3.5}$$

where k is the Boltzmann constant, T is the absolute temperature in °K and E_f is the Fermi level. The Fermi level is the energy at which the probability of occupation by an electron is exactly one half. For the common situation in which Boltzmann statistics

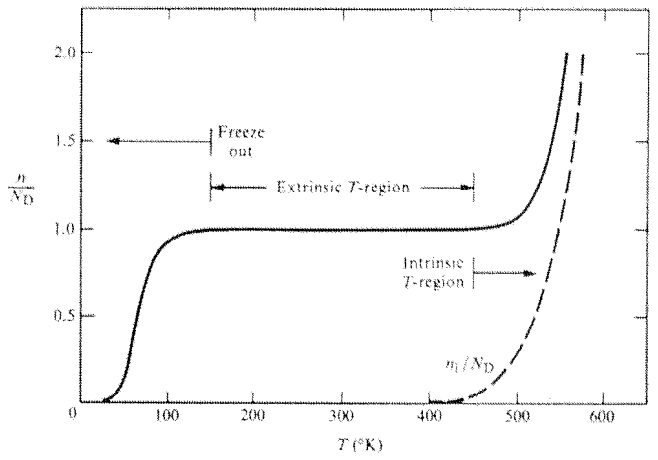


Fig. 3.7: Electron concentration in Silicon as a function of temperature in a sample with a donor concentration of 10^{15}cm^{-3}. From [24].

applies (i.e. $(E_c - E_f) \gg kT$ meaning that only few allowed states in the conduction band are filled as in a not too highly doped n–type semiconductor) Eq. (3.4) can be approximately evaluated to yield

$$n = N_c \exp\left(-\frac{E_c - E_f}{kT}\right) \tag{3.6}$$

where N_c is called the effective density of states at the conduction-band edges, given by the expression

$$N_c = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2} \tag{3.7}$$

where m_n^* is the effective mass of the electron[3]. The above expressions can be easily adapted for holes by means of indices change.

For an intrinsic semiconductor, the neutrality condition obtained equating n expressed by Eq. (3.6) with its analogous for holes p, leads to

$$E_f = E_i = \frac{E_c + E_v}{2} + \frac{kT}{2}\ln\left(\frac{N_v}{N_c}\right). \tag{3.8}$$

At room temperature the Fermi level E_f is close to the middle of the band-gap and is called *intrinsic* Fermi level E_i. As n_i is useful in relating carrier concentrations in both intrinsic and extrinsic semiconductors, E_i is used as reference level when discussing doped semiconductors. In fact, since Eq. (3.6) and the analogous equation for holes may be rewritten in terms of n_i and E_i,

$$n = n_i \exp\left(\frac{E_f - E_i}{kT}\right) \tag{3.9}$$

[3]The influence of the crystal forces are incorporated in an effective mass that differs somehow from the free-electron mass. For more details see Ref. [24]

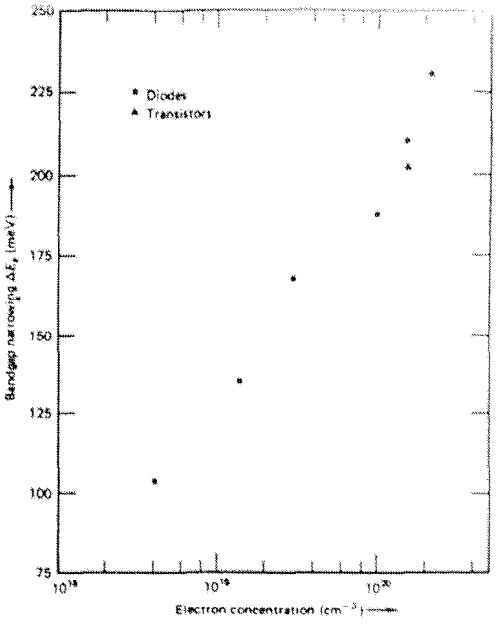


Fig. 3.8: Energy gap narrowing ΔE_g as a function of electron concentration. From [25].

and

$$p = n_i \exp\left(\frac{E_i - E_f}{kT}\right) \tag{3.10}$$

the energy separation from the Fermi level to the intrinsic Fermi level is a measure of the departure of the semiconductor from intrinsic material.

When the density of dopant in the crystal approaches N_c, the approximation of Fermi-Dirac statistics for electron density by the simpler Maxwell-Boltzmann statistics becomes invalid. Moreover, in a heavily doped semiconductor the band structure itself begins to be perturbed, since the density of dopant atom is such that individual impurity atoms interact with one another. The energy levels they introduce are no longer distinct and broadens into bands. These impurity bands can overlap the adjacent conduction or valence bands so that no energy is required to ionize the dopant atoms and provide free carriers. The most significant perturbation is the reduction of the silicon bandgap, giving rise to the so called *band-gap narrowing* effect (BGN). In Fig. 3.8 is clear that the effect can be noticeable from 10^{18} electrons cm^{-3} where ΔE_g is more than 10% of the bandgap energy.

This behaviour can be observed in microstrip detectors since the definition of the strips and guard-rings is done via ion implantations. Doses are such that the impurity concentrations may excess the threshold described above. The combination of BGN with a local high electric field may reduce significantly the thresholds for impact ionisation and junction breakdown (cf. Sec. 3.3.3). Therefore this phenomenon will cause at least higher leakage current flowing through the contacts, increasing the overall noise and hence, degrading the detector performances.

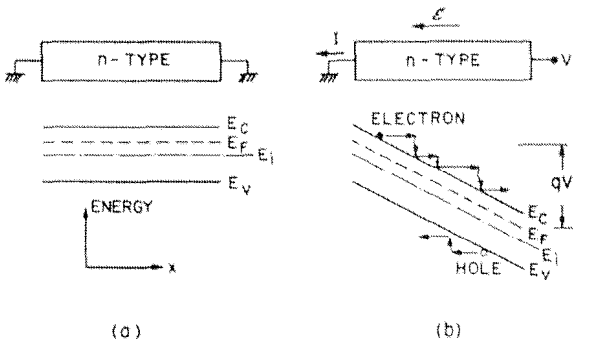


Fig. 3.9: Effect of an electric field $\mathcal{E}$ on the conduction process (a) at thermal equilibrium and (b) under biassing condition. From [23].

3.2.1 Carrier Action

In the previous section are defined the static properties of intrinsic and extrinsic semiconductor materials at the thermal equilibrium. If the semiconductor system is perturbed (e.g. by the passage of charged particles through a microstrip detector), a net carrier response is generated, giving rise to currents that can flow within and external to the semiconductor system. Under normal operating conditions there are three primary types of carrier actions occurring inside semiconductors: *drift*, *diffusion* and *generation-recombination*.

3.2.1.a Carrier Drift

In thermal equilibrium conditions, electrons within a lattice of a $n-$type semiconductor have a certain energy giving rise to almost random thermal motion of the electrons around equilibrium positions. Typical electron thermal velocity at room temperature (300 °K) in silicon is of the order of 10^7 cm/sec. If a small electric field $\mathcal{E}$ is applied to the semiconductor, each electron will experience a force $-q\mathcal{E}$ from the field and it will be accelerated along the field during the time between thermal collisions.

The net effect will be the superposition of a field induced velocity component to the thermal motion induced one. This additional component is called *drift velocity* and can be defined for electrons as $v_d = \mu_n \mathcal{E}$, where μ_n, the *electron mobility*, describes how strongly the motion of an electron is influenced by the applied electric field.

$$\mu_n = \frac{k\tau_n}{m_n^*} \tag{3.11}$$

The mobility, defined in Eq. (3.11) in the case of electrons, depends on the effective mass of the carrier and on the mean free time between collisions, τ_n, which is determined by various scattering mechanisms, like the scattering with impurities and the interaction with the lattice vibrations (phonons scattering). In Fig. 3.10 the electron and hole mobility as function of the dopant concentration at room temperature is shown: below 10^{15}cm^{-3}, mobilities are essentially independent of the doping concentration, since ionized

Parameter	Boron	Phosphorus
μ_{min}	44.6	68.7
μ_{max}	470.6	1415
N_{ref}	$2.23 \cdot 10^{17}$	$9.18 \cdot 10^{16}$
α	0.72	0.71

Table 3.3: Relevant parameters for mobility definition in Eq. (3.12)[26].

impurity scattering can be neglected compared with lattice scattering, which in turn is independent of N_A or N_D, the total acceptor or donor concentrations. For doping in excess of $\sim 10^{15}\text{cm}^{-3}$, ionized impurity scattering can no longer be neglected, correspondingly decreasing the carrier mobilities. An important observation is the dependence of mobility on both density and type of dopant impurities. Note that greater electron mobility is due mainly to the smaller effective mass of electrons. Figure 3.11 highlights the dependence of the electron mobility on the temperature for silicon with different donor concentrations. In all the calculations in this thesis, the mobility is parameterized following Masetti's description [26]:

$$\mu = \mu_{min} + \frac{\mu_{min} - \mu_{max}}{1 + (N/N_{ref})^{\alpha}} \tag{3.12}$$

where N is the dopant density of the silicon and the other parameters are different for different dopants and are resumed for Boron and Phosphorus (the impurities used in the CSEM process) in Table 3.3.

For example, in lightly doped silicon (as the usual bulk material used for silicon detector fabrication) at room temperature, the electron mobility is 1350 $\text{cm}^2/\text{V s}$ and hole mobility is 480 $\text{cm}^2/\text{V s}$. In much more highly doped regions, like the diodes building the micro-strips, the mobility of carriers drops to much lower values, thus influencing the conduction mechanism with important consequences on leakage current, signal development and response time of the device.

Resistivity is an important material parameter closely related to carrier drift. Quantitatively resistivity ρ is defined as the proportionality constant between electric field and the total particle current per unit area flowing in the material; that is

$$\mathcal{E} = \rho J. \tag{3.13}$$

In an homogeneous material, $J = J_n = q\mu_n n\mathcal{E}$ for the electron current density (analogous equation holds for holes), and therefore

$$\rho = \frac{1}{q(\mu_n n + \mu_p p)}. \tag{3.14}$$

In Fig. 3.12 the silicon resistivity at room temperature is shown as function of impurity concentration. For pure silicon at 300K, $\rho_i = 235$ KΩ·cm. Usually micro strip detectors are fabricated on $2 - 20$kΩ·cm substrates: as will be described later on, the higher the resistivity, the lower the bias voltage needed to fully deplete the substrate, thus simplifying

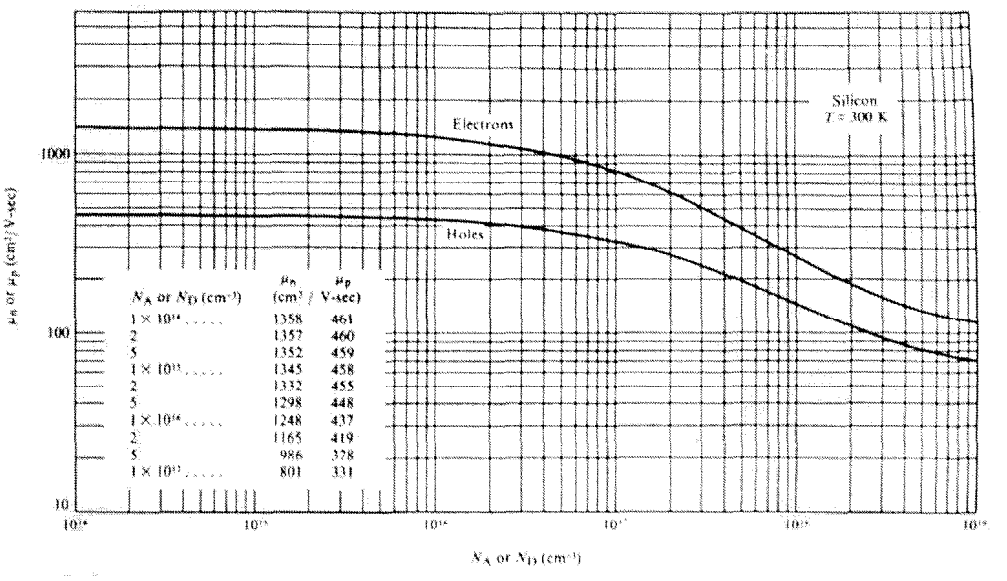


Fig. 3.10: Room temperature carrier mobilities in silicon as a function of dopant concentration. From [24].

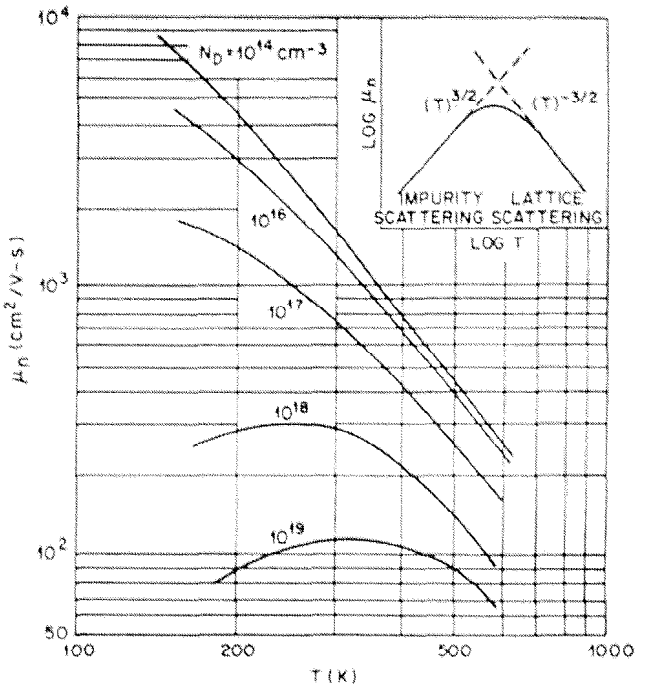


Fig. 3.11: Electron mobility in silicon versus temperature for various donor concentrations. Insert shows the theoretical temperature dependence of electron mobility. From [23].

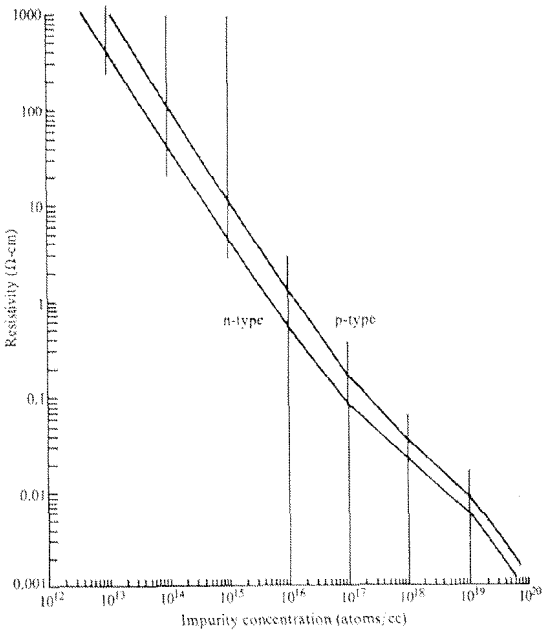


Fig. 3.12: Silicon resistivity as a function of doping concentration[27].

biasing and electronics. The total current density due to drift phenomena can be written as the sum of electrons and holes components:

$$J_{drift} = J_n + J_p = (qn\mu_n + qp\mu_p)\mathcal{E}. \tag{3.15}$$

3.2.1.b Carrier Diffusion

In semiconductors, the low conductivity coupled with the possibility of spatial variation of carrier densities and energies, held to diffusion of carriers which can affect very much the current flow. This contribution to the current is usually known as diffusion current. It can be easily demonstrated that the current induced by those processes in the one-dimensional case has the form, [28]:

$$J_n = qlv_{th}\frac{dn}{dx} \tag{3.16}$$

where x is the direction of carrier density, $l = v_{th}\tau_n$ is the mean free path of an electron, expressed in terms of thermal velocity and average time between scattering events. It is possible to write Eq. (3.16) in a more useful form by applying Eq. (3.11) with the theorem of equipartition of energy:

$$J_n = q\left(\frac{kT}{q}\mu_n\right)\frac{dn}{dx} = qD_n\frac{dn}{dx} \tag{3.17}$$

The quantity D_n is called *diffusion constant* (or diffusivity) and his definition is known as the Einstein relation: it relates free-carrier transport by drift and by diffusion.

In presence of an electric field the total current for electrons will then be:

$$J_n = qn\mu_n\mathcal{E} + qD_n\frac{dn}{dx} \tag{3.18}$$

and analogous for holes:

$$J_p = qp\mu_p\mathcal{E} - qD_p\frac{dp}{dx}. \tag{3.19}$$

3.2.1.c Carrier Generation and Recombination

In this section is presented a brief summary of the mechanisms of generation and recombination which are of direct interest for the microstrip detectors. For further details refer to [23, 28].

Wherever the thermal equilibrium is disturbed, e.g. $np \neq n_i^2$, processes take place to restore it. If excess carriers are introduced in the semiconductor, $np > n_i^2$, the mechanism which restore equilibrium is the recombination of injected minority carriers with the majority carriers present in the semiconductor.

External carriers can be introduced e.g. by the passage of an ionizing particle through the semiconductor: if the transferred energy exceeds E_g, the band-gap energy, electron holes pairs can be produced and they will contribute to the conduction. These additional carriers are called excess carriers. Usually the concentration of excess carriers does not significantly compare with the doping concentration present in the semiconductor: this condition, that is, $\Delta n = \Delta p \ll N_D$ is referred to as *low-level injection.*

Recombination processes can involve the emission of a photon (radiative recombination) or not, and can be defined as *direct* or *indirect.* Direct recombination is frequently encountered in direct semiconductors like gallium-arsenide rather than silicon where indirect recombination is dominant. Direct recombination reflects the fact that an electron in the valence band of a direct semiconductor does not require a change in the crystal momentum to make the transition in the conduction band, which is actually the case for silicon: such an electron needs a simultaneous three particle reaction (the electron, the hole and a lattice phonon) to be able to jump in the conduction band, thus making such a reaction very unlikely. Indirect recombination makes use of two-particle interactions, such as those between a free carrier and a lattice phonon, that can takes place if there are localized allowed energy states into which electron and holes can make transitions. In practice localized states in the forbidden band gap are always present, due to lattice imperfections or because of impurity atoms, thus both acting as stepping stones for carriers.

The formulation of generation and recombination currents through localized states or recombination centers was originally derived by Shockley and Read [29] and by Hall [30] and the process is commonly referred as SHR recombination. In Fig. 3.13 are sketched the various transitions which takes place through recombination-generation (RG-) centers. If N_t is the RG-centers density in the silicon, R'_{abcd} are the spontaneous recombination rates due to the four different processes, G' is the spontaneous generation rate under

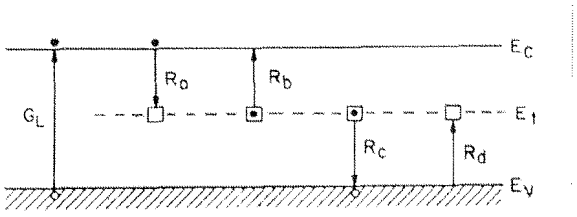


Fig. 3.13: Indirect generation and recombination processes under illumination.

equilibrium conditions the net recombination rate U can be defined as follows [23]:

$$U \equiv R' - G' = v_{th}\sigma_0 N_t \frac{(p_n n_n - n_i^2)}{p_n + n_n + 2n_i \cosh\left(\frac{E_t - E_i}{kT}\right)} \tag{3.20}$$

assuming that the capture cross sections for electrons and holes are equal ($\sigma_p = \sigma_n = \sigma_0$). If the thermal equilibrium is disturbed by a sudden creation of equal number of excess carriers, e.g. the passage of a mip, under low-injection conditions we get ($p'_n = p_n - p_{n0}$ is the excess density of vacancies injected in a $n-$type material at thermal equilibrium and p_{n0} is the carrier equilibrium density):

$$U \equiv R - G = v_{th}\sigma_0 N_t \frac{(p_{n0} + n_{n0})p'_n}{p_n + n_n + 2n_i \cosh\left(\frac{E_t - E_i}{kT}\right)} \tag{3.21}$$

which can be rewritten as:

$$U \simeq \frac{p_n - p_{n0}}{\tau_r} \tag{3.22}$$

with τ_r is the *recombination lifetime* given by:

$$\tau_r = \frac{1 + \left(\frac{2n_i}{n_{n0}p_{n0}}\right)\cosh\left(\frac{E_t - E_i}{kT}\right)}{v_{th}\sigma_0 N_t}. \tag{3.23}$$

These relationships are related to an excess carrier injection driven by an external illumination source, i.e. $np > n_i^2$. The same arguments can be applied to the case of a carrier extraction as is the usual situation of a $pn-$junction under reverse bias conditions ($np < n_i^2$): to restore the system to equilibrium, carriers must be generated by the RG-centers, giving:

$$U = \frac{v_{th}\sigma_0 N_t n_i}{2\cosh\left(\frac{E_t - E_i}{kT}\right)} = \frac{n_i}{\tau_g}. \tag{3.24}$$

where τ_g is called *generation lifetime*. The physical meaning of those lifetimes is related to the exponential decay of the excess carriers density with time as the silicon returns to thermal equilibrium conditions. The generation lifetime depends strongly on the trap energy level and it determines the leakage current level of a microstrip detector since its mode of operation requires full depletion of the bulk, as it will explained later. The closest

the trap to the mid-gap, the shortest the generation lifetime. Heavy metal contamination, like gold, even at low concentration, or radiation induced defects are setting trap level close to the midgap: this situation has to be avoided. Usually τ_g before irradiation is of the order of some milliseconds, allowing minority carrier generated all over the silicon bulk to be fully collected to the electrodes, avoiding dangerous charge loss.

Another source of RG-centers is the surface of the semiconductor. Because of the abrupt discontinuity of the lattice structure at the surface a large number of dangling bonds will increase the density of extra localized states having energies within the forbidden gap. The presence of a passivating oxide at the surface ties up many of the bonds that would otherwise contribute to surface states and protects the surface from foreign atoms. Such an oxide can reduce the density of surface states[4] from about 10^{15} to less than 10^{11} cm^{-2}. The kinetics of RG at the surface are similar to those considered for the bulk, except for the fact that the density of traps is an area density, thus expressed as $N_s(\text{cm}^{-2})$. In low-injection conditions the surface RG-rate is:

$$U_s \simeq v_{th}\sigma_s N_s \frac{(p_s n_s - n_i^2)}{p_s + n_s + 2n_i \cosh\left(\frac{E_s - E_i}{kT}\right)} \tag{3.25}$$

where p_s and n_s are the superficial density of electrons and holes at the surface Si-SiO$_2$. The product $v_{th}\sigma_s N_s$ is called *surface recombination velocity*, s, and depends on physical nature and density of the surface RG-centers as well as on the potential at the surface. If the surface is depleted of mobile carriers, n_s and p_s are small and s is large. If we compare s with the definition of minority carrier lifetime τ_g, given in Eq. (3.24), is evident that s is related to the rate at which excess carrier will recombine at the surface just as $1/\tau_g$ is related to the rate at which they recombine in the bulk. In the matter of facts, s is very sensible to the potential of the silicon surface and can determine the overall behaviour of the surface conduction in a MOSFET-like device. An average value for s in a microstrip detector made on high resistivity silicon is typically centered around 10 cm/sec.

The combined effect when drift, diffusion and generation-recombination phenomena are present is described in the *continuity equations* which are given below [23, 31]:

$$J_n = q\mu_n n\mathcal{E} + qD_n \frac{\partial n'_n}{\partial x} \tag{3.26}$$

$$J_p = q\mu_p p\mathcal{E} - qD_p \frac{\partial p'_p}{\partial x} \tag{3.27}$$

$$\frac{\partial n}{\partial t} = \frac{1}{q}\frac{\partial J_n}{\partial x} + (G_n - R_n) \tag{3.28}$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x} + (G_p - R_p) \tag{3.29}$$

where it is assumed that neither diffusion constants D_n and D_p nor mobilities μ_n and μ_p are function of position. Together with the continuity equations, the Poisson equation

[4]The theoretical RG-density on the surface depends from the crystal orientation of the bulk. A $< 100 >$ orientation allows a stiffer packing (with respect to $< 111 >$) of oxygen atoms in the silicon thus decreasing the surface density of traps. Typically the surface density of recombination centers in $< 111 >$ silicon ranges from $2 - 3 \cdot 10^{11}$ to $4 - 5 \cdot 10^{12}$ cm^{-2} where it saturates, to be compared with values ~ 10 times lower in $< 100 >$.

must be satisfied,

$$\frac{d\mathcal{E}}{dx} = \frac{\rho}{\epsilon} \tag{3.30}$$

with ρ is the spatial charge density (given by the algebraic sum of carrier densities and ionized impurity concentrations) and ϵ is the dielectric constant of the semiconductor ($\epsilon_{Si} = 11.7$). With appropriate boundary conditions, this system of differential equation gives an unique solution and it allows to describe most of the electrical characteristics of a silicon device, like a microstrip detector.

3.2.2 Toward the microstrip detector: the $p-n$ junction

Up to now the discussion took into consideration the properties of semiconductor materials and the action of carriers in the hypothesis of uniform density of dopants. It is interesting here to briefly resume the behaviour of two different semiconductors once they are kept in contact. A $p-n$ junction is produced by forming a single crystal semiconductor such that part of the crystal is doped $p-$type and the other part is doped $n-$type. Junctions are then classified as *abrupt* when the transition from $p-$type to $n-$type is extremely narrow (some atomic distances), and *graded* if this transition is more spread out over a larger distance. The fundamental characteristic of a $p-n$ junction is that it rectifies the current flowing through it, i.e. it allows the carriers to easily pass in one direction only, depending to the bias polarity applied to the junction. It behaves then diode-like. At equilibrium, assuming internal charge neutrality, when the two materials are joined, the carrier concentration difference between the $p-$ and $n-$regions causes the carriers to diffuse. The charge imbalance produces an electric field, which counteracts the diffusion so that in thermal equilibrium the net flow of carriers is zero. The space charge region near the metallurgical junction is referred to as *depletion zone* since majority carriers have been removed. In Fig. 3.14 are summarized the characteristics of the junction electrostatic.

The general form for the electric field $\mathcal{E}$ can be inferred from the Gauss theorem, using as the net charge density the charge imbalance between charge carriers and ions, $\rho(x) = q(p - n + N_D - N_A)$. It follows that, even at thermal equilibrium, it exists a potential difference across the junction, the built-in potential:

$$\phi_i = \frac{kT}{q} \ln\left(\frac{N_D N_A}{n_i^2}\right). \tag{3.31}$$

In the case of high resistivity silicon for MSSD with the typical doping concentration for bulk and implanted areas, ($N_D = 10^{12}\text{cm}^{-3}$ and $N_A = 10^{19}\text{cm}^{-3}$), $\phi_i = 0.639$ V.

In the depletion approximation (i.e. the density of mobile carriers in the depletion region is small compared to the acceptors and donors concentration), it is possible to solve the Poisson equation (3.30) in order to get the electric field in both regions:

$$\mathcal{E}(x) = -\frac{qN_A}{\epsilon_{Si}\epsilon_0}(x_p + x) \qquad \text{for } -x_p \leq x \leq 0, \tag{3.32}$$

$$\mathcal{E}(x) = -\frac{qN_D}{\epsilon_{Si}\epsilon_0}(x_n - x) \qquad \text{for } 0 \leq x \leq x_n. \tag{3.33}$$

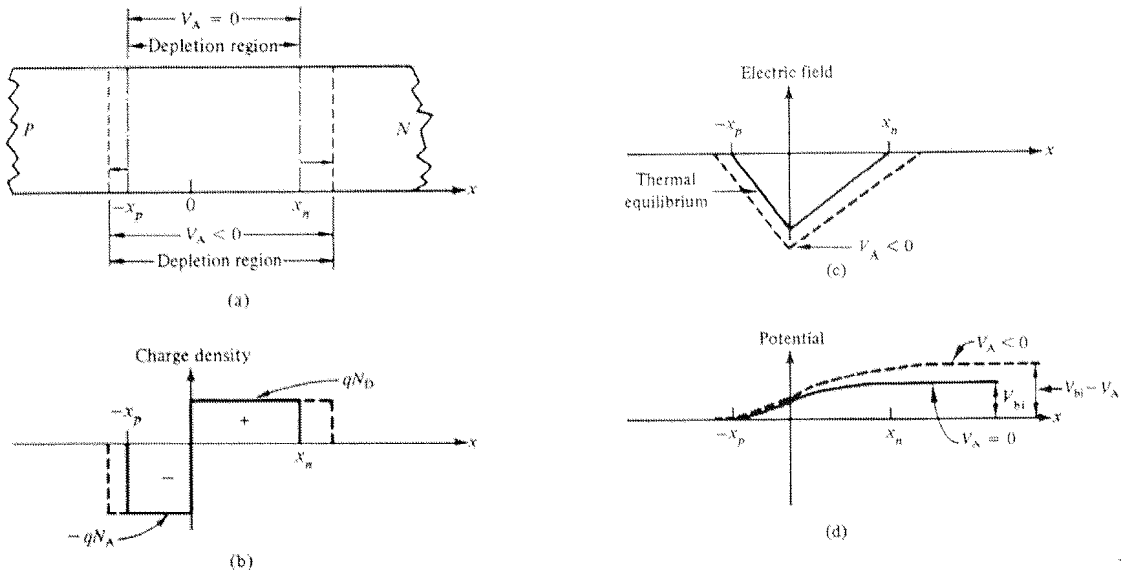


Fig. 3.14: Summary of a $p-n$ junction electrostatics at equilibrium (filled line) and under reverse bias (dashed lines). (a) Depletion region, (b) Charge density, (c) Electric field, (d) Potential.

For the potential $V(x)$ within the depletion region, one gets:

$$V(x) = -\frac{2qN_A}{2\epsilon_{Si}\epsilon_0}(x_p + x)^2 \qquad \text{for } -x_p \leq x \leq 0; \tag{3.34}$$

$$V(x) = -\frac{2qN_D}{\epsilon_{Si}\epsilon_0}(x_n - x)^2 + \phi_i \qquad \text{for } 0 \leq x \leq x_n. \tag{3.35}$$

Since the potential function $V(x)$ must be continuous at the interface, $(x = 0)$, equating there Eq. (3.34) and Eq. (3.35), one can solve for x_n thus getting an expression for the depleted zone in the n–side of the junction:

$$x_n = \sqrt{\frac{2\epsilon_{Si}\epsilon_0\phi_i}{q}\frac{N_A}{N_D\,(N_A + N_D)}}. \tag{3.36}$$

As N_D increase, x_n will decrease for a fixed value of N_A, that is the higher the resistivity of the bulk material, the lower the depletion width. A completely analogue expression can be derived for x_p, by swapping in (3.36) N_D with N_A. The total depletion width is then given by the sum of x_n and x_p,

$$W_d = \sqrt{\frac{2\epsilon_{Si}\epsilon_0\phi_i}{q}\frac{(N_A + N_D)}{N_D N_A}} \tag{3.37}$$

Using high resistivity material as in the former example ($N_D = 10^{12}\text{cm}^{-3}$), the depletion region extends dramatically, $W_d = 28.8\mu$m, since it depends almost exclusively upon the lower charge density, in this case N_D. In the case of an external potential difference is

applied across the junction, Eq. (3.37) becomes:

$$W_d = \sqrt{\frac{2\epsilon_{Si}\epsilon_0}{q}(\phi_i - V_b)\frac{(N_A + N_D)}{N_D N_A}} \tag{3.38}$$

and the diode is known as *forward* biased if $V_b > 0$ on the $p-$region and *reverse* biased if the polarity is opposite. In the first case the width is reduced. This fact yields an increase in the hole diffusion current component over the thermal equilibrium value, since a large number of holes have energies greater than the barrier height and are able to diffuse into the $n-$material. The hole drift current component remains the same as the thermal equilibrium value, since a change in the barrier height has no effect on the number of holes p_n. The holes able to diffuse from the $p-$region into the $n-$region are called *injected minority carriers*. Same arguments can be applied to the electrons. The net effect of a forward bias is a large increase in the diffusion current component while the drift components remain fixed near their thermal equilibrium values. Since the Fermi distribution function distributes carriers nearly exponentially with increasing energy, the number of carriers able to diffuse across the junction will increase exponentially when the potential barrier is reduced, as is the case of the forward biased diode characteristic curve.

If the diode is reversely biased, the width of the depletion region, the electric field and the potential will increase, implying a wider depletion region and an higher electric field due to the increase of the potential drop across the junction. In Fig. 3.15, the slope of the band edges in the depletion region has increased, reflecting the increase in the electric field. The number of holes (p_p) having sufficient energy to get over the potential barrier into the $n-$region is reduced, thus reducing the hole diffusion current to less its thermal equilibrium value. The drift component remains at its equilibrium value, limited by the supply of minority carriers electrons (n_p) in the $p-$region. The reverse current is therefore small, negative, limited by the supply of thermally generated minority carriers and independent of V_b after few volts of reverse bias. Increasing further the reverse bias voltage does not affect the minority carrier supply, therefore keeping the reverse current to a constant value. Its temperature dependence is important and it will be studied later in this chapter.

3.3 The microstrip detector

At this point there are enough information to derive the basic characteristics of a typical high energy particle detector. In Fig. 3.16 is sketched a double-sided silicon detector of thickness l made of high purity, high resistivity $n-$type silicon. In order to obtain position sensitivity for the silicon detectors, the conventional front side $p^+ - n$ junction is patterned, subdividing the diode electrodes into many smaller diodes which can be readout separately. On the backside a contact structure is created. This contact consists essentially in a high dose of type n dopant either diffused through the surface or implanted into and it can be uniform or structured. The backside doping profile has a twofold function. First, it assures an ohmic contact at the Silicon-Aluminum interface, and secondly, as it will be described in Sec. 3.3.2.c, it decreases the diode leakage current by reducing

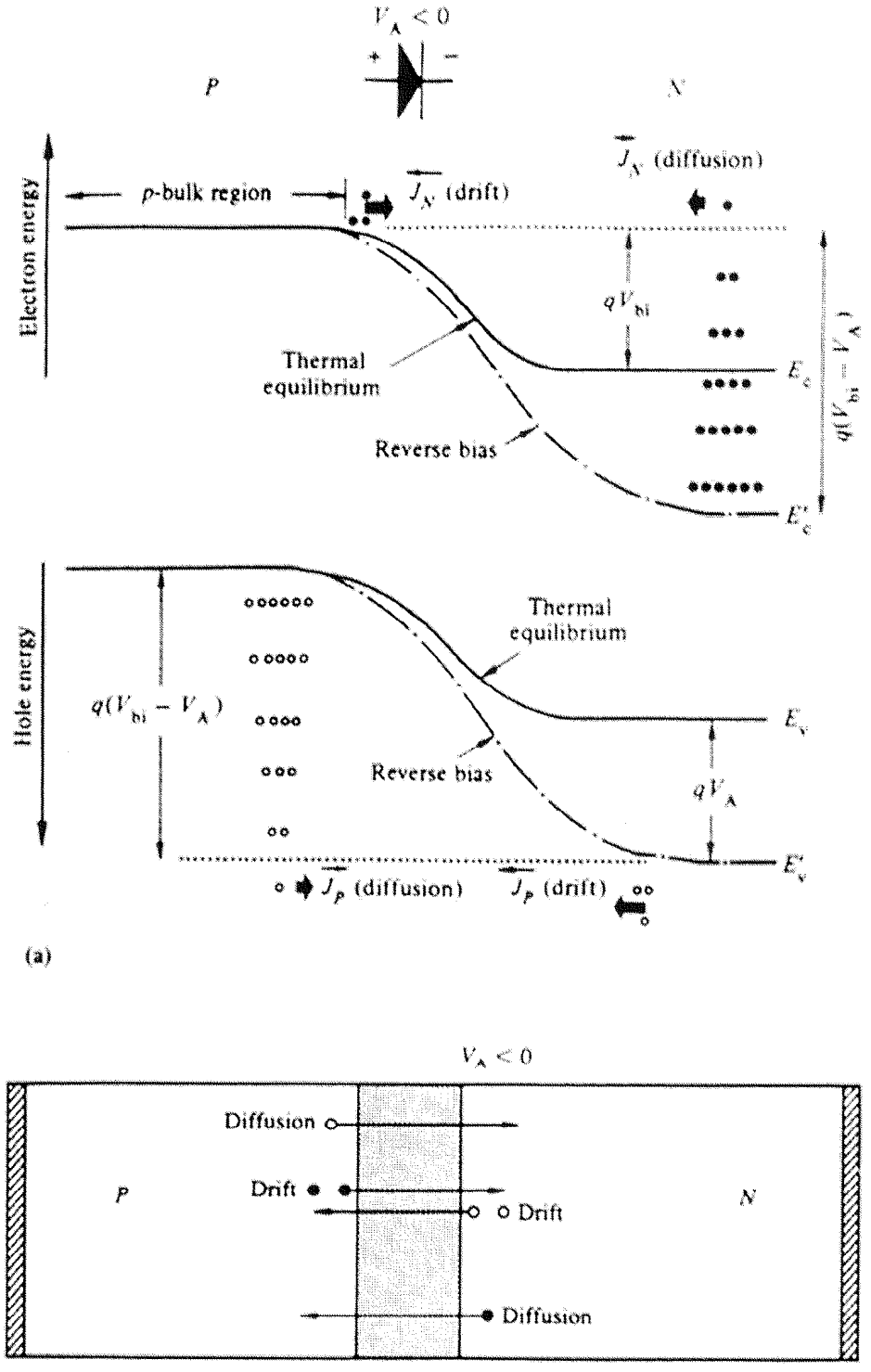


Fig. 3.15: Energy band diagram for reverse bias (dashed line) and at thermal equilibrium (continuous line) [31].

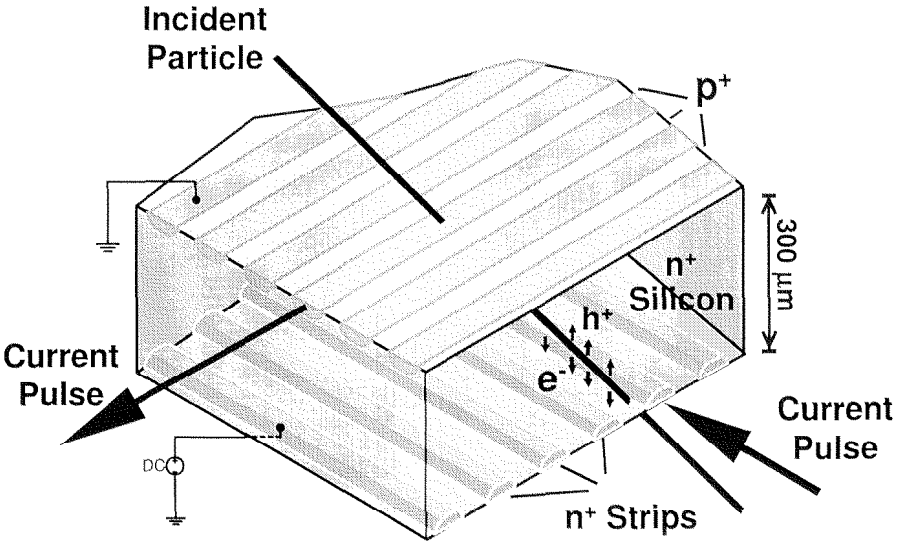


Fig. 3.16: Cross section of a double-sided silicon microstrip detector: operative principle.

the number of minority carrier reaching the n^+ layer from the back contact. Whereas the p^+n junction is permeable for minority carries coming from one region and entering the other region and blocking for majority carriers, the high-low "junction", meaning with this a $n^+ - n$ zone, is permeable for majority carriers and blocking for minority ones. Because of this fact the high-low "junction" is referred to as minority carrier *reflection junction* [32].

A positive voltage V_b applied to the n–side of the detector, will deplete the high resistivity silicon bulk up to the depletion region will reach the n^+–region of the back contact. The entire thickness of the silicon bulk is now entirely depleted of majority carriers and the only current flowing through the detector is due to the thermal generation of electron-hole pairs in the space charge zone (SCR).

Charge Collection When a charged particle pass through the detector, it loses energy by Coulomb scattering with the electrons and it will free an electron from the lattice, thus creating an electron-hole pair, with an average rate of one per 3.6 eV[5]. It follows that, with a typical thickness of 300 μm, around 24000 electron-holes pairs are released in the bulk. Since the detector is fully depleted, full collection of the electron-holes pairs generated by the passage of a charged particle is assured. Electrons will drift toward the n^+–contact, holes toward the p^+-strip following the internal electric field.

While the carriers are drifting, the process of diffusion spreads out the original charge. In the case of a minimum ionizing particle (m.i.p.) the diffusion of charge can be described by simple diffusion theory[6]. In this framework, the cluster charge will drift through the

[5]The precise value depends on temperature since the E_g depends on temperature

[6]In the case of non-m.i.p.'s or alphas, the density of electrons and holes is so high that space-charge effect should be added to the description.

detector towards electrodes and at the same time will diffuse radially. The RMS radius of the diffused charge distribution is described by

$$\sigma = \sqrt{D\tau_d}, \tag{3.39}$$

where D is the diffusion constant from Eq. 3.17 and τ_d is the mean drift time, as previously described in Sec. 3.2.1.a.

It should be noted that for a given applied reverse voltage V_b, the electric field is linear in the bulk rather than constant (see Fig. 3.14) and therefore one should expect to find differences in respect to conventional ionisation chambers. First, outside the depletion region the electric field $\mathcal{E}$ is zero, therefore the charge produced in this zone cannot be collected. Then, even in a totally depleted detector, the electric field is almost zero on at least one electrode, thus, strictly speaking, causing an infinite charge collection time. In practice, because of charge motion, an induced signal appears on the electrodes:

$$q(t) = Q\left(1 - \frac{x_0}{W}\right)\left(1 - e^{-\tau/\epsilon_{Si}\rho}\right) \tag{3.40}$$

where W is the thickness of the detector and x_0 is the distance of the produced charge Q from the junction. About 86% of the charge is collected in $\tau = 2\rho\epsilon_{Si}$, which is typically 20 nsec for 10 kΩ·cm silicon detector.

In an overdepleted substrate, collection time is determined mainly by the drift velocity calculated at the electrode, where the electric field is not anymore zero. The previous estimation is confirmed by assuming a typical field of 1 kV/cm and a 300 μm thick detector. Following Eq. 3.39, electrons created by a mip impinging orthogonally the detector will be fully collected after 20 nsec thus fixing the minimum response time of the device, assuming that full charge collection efficiency is required. Holes collection time is connected to electron collection time through the ratio of mobilities, $\tau_p = \tau_n(\mu_n/\mu_p)$: in the case of silicon at room temperature the ratio $\mu_n/\mu_p \simeq 3$ and collection is dominated by electrons. Of course, collection is strongly affected by carrier trapping and re-emission, which can be considerably lengthen the pulse, as in heavily doped lithium compensated crystals. After 20 nsec the electronic cloud will have a diameter of about 18 μm. Assuming a typical implant pitch of 25 μm, thus fixing the elementary cell dimensions, a charge released all over a cell gives a signal on the same strip. Clearly, if the trajectory of the ionizing particle is not orthogonal with respect to the plane of the detector, the transit time will increase and the electron cloud corresponding to a tacking angle of 45° will have a diameter of 25 μm. It should be noted that here we are not considering the case of $\delta-$electrons emission or the presence of a magnetic field: both causes would systematically displace the charge centroid, thus introducing an error in the position measurement.

Such cloud widening may spread the charge over the neighboring strips due to the mutual capacitive coupling of microstrips, eventually increasing the precision on position measurement beyond its intrinsic upper limit described by $\sigma_x = p/\sqrt{12}$, where p is the strip pitch [33]. This can be seen as an attractive option for improving spatial precision beyond the limits of the detector granularity if and only if the electronics meant to readout the single channel is able to measure the quantity of charge deposited on its referring strip and this charge is above the acceptable noise threshold.

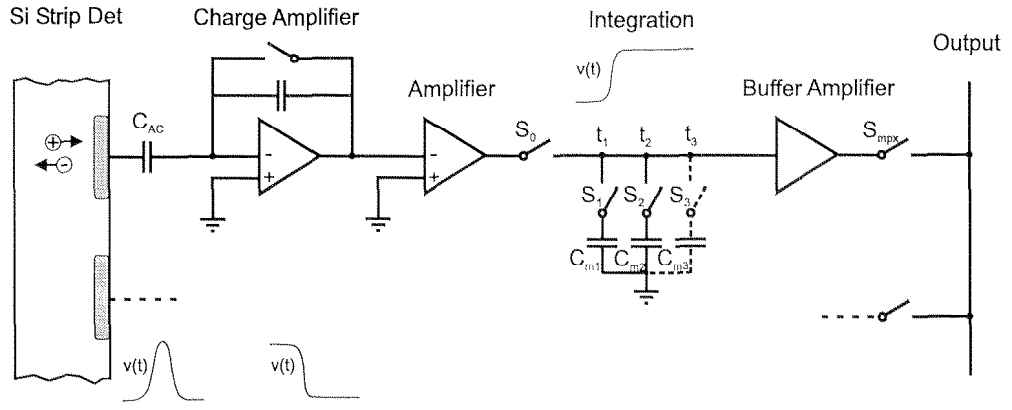


Fig. 3.17: Signal charge measurement in microstrip detectors by low noise amplification, sampling, storage and multiplexed output readout: functional system.

With a typical implant pitch of 25 μm some authors [34] reported one micron precision by centroid finding on the basis of measured charge collection on adjacent strips and by using ultra low noise amplifiers. More detailed description on charge division and position resolution are given in Sec. 5.3.3.

3.3.1 The optimization of the Signal to Noise ratio

In the process of optimizing the detector system it is very important to have a model of how the ideal system should look like [35]. In the most general case the system consists of a solid-state detector with biasing network, a preamplifier which conditions the shape of the signal a first time before putting it on the signal bus, and a shaper, which recovers the signal from the added noise and makes signal processing more accurate (cf. Fig. 3.17). In most cases a charge sensitive preamplifier is preferred to a voltage or current amplifier. The main reason for that is the fact that this type allows the output signal to be made independent of detector capacitance, which behaves non-linear as a function of the applied bias. The shaper consists usually in a combination of RC(integration) and CR(differentiation) circuits [36]. In this chapter the emphasis is put on the detector part; more details of the amplification circuits can be found in the Tracker TDR [14].

The spatial resolution and the hit finding efficiency in a microstrip detector depend on the strip pitch and on the signal-to-noise ratio. In order to study the expected S/N ratio it is useful to decouple the various noise sources in different contributions and evaluate them separately, whenever possible.

The most important noise sources for a microstrip detector can be summarized as in Fig. 3.18[7]: a voltage generator in series with the detector and the amplifier and a current generator in parallel with the detector: those two contribution are then identified respectively as *parallel* and *serial* noise.

In the first category are included the shot-noise created by the leakage current in the

[7]Here the noise induced by the rest of the readout chain is not taken into account

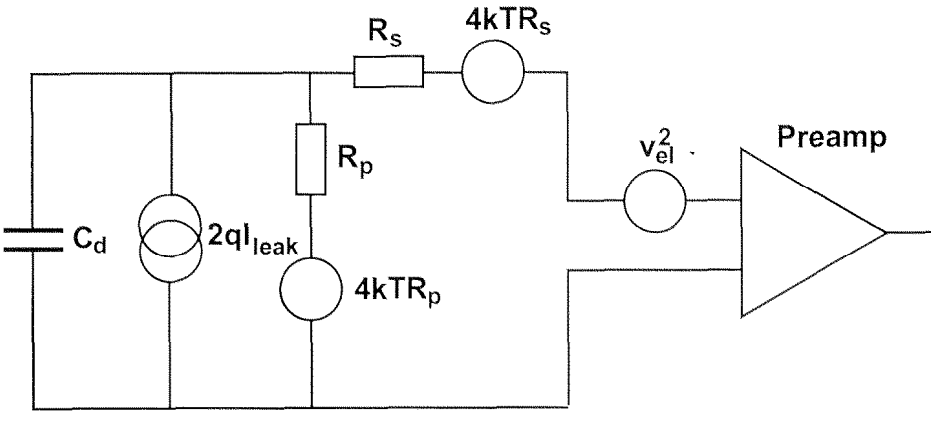


Fig. 3.18: Noise sources in a silicon detector system. For symbols refer to Tab. 3.4. From [37]

detector and the thermal noise from the biasing resistors. On the other hand, the series resistance noise in the detector and in the front of the input-transistor of the pre-amplifier, the flicker noise, the channel thermal noise and the bulk series-resistance noise in the input transistor are considered in series with the input transistor. All the above noise sources are white except the flicker noise which has a $1/f$ dependence on the frequency.

For low leakage current, the parallel noise is mainly Johnson-type thermal noise generated in the biasing resistors of the detectors and/or the feedback resistor of the pre-amplifier. The serial noise can be classified as external or internal to the preamplifier. The only external source is a Johnson-type noise due to the resistance of the metal strip connected at the input of the amplifier. The internal sources come from fluctuations in the drain current.

Noise is usually expressed in terms of equivalent noise charge (ENC), defined as the charge that, generated in the detector would give a signal equivalent to the noise at the output of the readout chain and related the root mean square (RMS) noise voltage at the output of the shaper directly to the signal height at the input. The definition of ENC is dependent on the type of shaper in the amplifier, and the definition have been adapted to the CR-RC chain of the APV6.

In Table 3.4 are resumed the different noise contribution expressed in ENC. Since the noise contributions from the amplifier are independent we can regroup them in a more familiar form:

$$ENC_{APV} = a + b \cdot C_{tot} \tag{3.41}$$

with a and b functions of the quantities described in Tab. 3.4, knowing that the $1/f$ and bulk resistor noise factors are very small and can be neglected. Hence, for the APV6 case without applying deconvolution yet [14], $a = 512\,e^-$ and $b = 41\,e^-$. After deconvolution for three sampling in 50 nsec this is equivalent to multiply the series noise by a factor ~ 1.45 and the parallel one by 0.45.

It is interesting now to make the point about the different weights those contribution to the noise have.

Noise Source	Kind	ENC
Leakage current	Series	$\frac{e}{q}\sqrt{\frac{qI_{tot}\tau}{4}}$
Bias resistor	Series	$\frac{e}{q}\sqrt{\frac{kT\tau}{2R_p}}$
Metal strip	Parallel	$\frac{e}{q}C_{tot}\sqrt{\frac{kTR_s}{6\tau}}$
Trans. flicker noise	Parallel	$\frac{e}{q}C_{tot}\sqrt{\frac{F_k}{2WL_{eff}}}$
Trans. channel	Parallel	$\frac{e}{q}C_{tot}\sqrt{\frac{kT}{3g_m\tau}}$
Trans. bulk resistor	Parallel	$\frac{e}{q}C_{tot}\sqrt{\frac{R_{bulk}\eta^2kT}{2\tau}}$

Table 3.4: Summary of noise sources for a microstrip detector [38]. e is the Euler constant, τ is the peaking time of the amplifier, $I_{tot}$is the detector dark current, R_s and R_p are resp. the value of biasing resistors and the resistance obtained by the parallel of the metal contacting the strips and the feedback resistor across the amplifier; C_{tot} is the total capacitance of the detector, F_k is a process dependent constant ($\sim 10^{-22}$) [39], W is the transistor width, L_{eff} is the effective transistor length, g_m is the transconductance, R_{bulk} is the amplifier bulk resistance and finally η is another process dependent parameter, (~ 0.1).

Main contributors are the leakage current I_{tot}, the total capacitance at the input of the amplifier C_{tot}, the shaping time τ and both R_s and R_p. The straightforward recommendations would then be: low leakage current, short shaping time, high polarization resistance, low metal resistance and, finally, the smallest strip capacitance. Some of these parameters are already constrained, as the shaping time τ, which is strongly dependent on the LHC beam crossing rate and thus is "freezed" to 50 nsec in order to maximize the collected luminosity.

The purpose of the rest of the chapter is to demonstrate that there are viable solutions to the S/N problem, understanding the impact of all the formerly mentioned physical quantities on the signal in the CMS microstrip detector. The study begins with leakage currents, continues with capacitances and ends up by looking into the resistances. During the analysis I took the opportunity to stress on certain arguments like junction breakdown or edge stabilization, not strictly linked to the optimization of the S/N ratio but rather to the correct operation of microstrip detectors. The final part is devoted to the study of radiation induced problems and a final estimation of the noise contributions is given.

3.3.2 Leakage currents

In the following sections the major sources of dark current for a detector will be investigated; the effect of the back contact to the leakage current will be described, ending with an analysis of the breakdown performances for such devices.

One of the biggest issues concerning detector quality is its leakage current. In those cases where the detector is followed by a DC-coupled integrating preamplifier, the leakage current is integrated together with the signal and degrades the S/N ratio and detector resolution. If a coupling capacitor is used instead, the signal is decoupled from the detector dark current, allowing thus much better S/N performances. Even so, the leakage current still has to be small in order to limit power dissipation and avoid the detector from heating, which in turn tends to increase the leakage current, as demonstrated later.

Assuming that optical generation and avalanche multiplication do not occur, we can identify four main sources of leakage current in a reverse biased diode type detector (cf. Fig. 3.19):

1. Thermal generation in the bulk depletion layer,
2. Thermal generation at the depleted $Si - SiO_2$ surface,
3. Diffusion from neutral regions towards the edge of the depletion layer,
4. Injection from the (metal) contact.

Generally speaking, the contribution to the dark current of the detector of these four leakage sources can be reduced by taking care of some details. In fact, the formerly mentioned high-purity material can help since it presents a lower amount of impurities and thus of generation-recombination centers (contribution (1)); the passivation of the surface, usually by oxidation, reduces the number of surface states and thus the surface generation (cause (2)).

An accurate design, making use of guard rings surrounding the device (cf. Sec. 3.3.3.a), strongly suppresses the contribution of surface generation; an n^+ implantation at the cutting edges (cf. Sec. 3.3.3.c) shield the active zone of the detector from lateral diffusion of free carriers coming,for example, from the strongly damaged cutting edge (causes (2) and (3)).

The contact to the bulk with the high-low junction reduces the effective diffusion length and avoid carriers injection from the contact(cause (4)) (cf. Sec. 3.3.2.c). Moreover, as generation is a thermal process a further reduction can be achieved by cooling the device (cf. Fig. 3.21). Last but not least, an adequate processing sequence has to be chosen, as will be shown in Chapter 4.

In order to understand the different processes responsible for the leakage current of a detector, the diode structure is divided into four regions as shown in Fig. 3.19.

The outer two regions (A and D) represent the high-doped regions (resp. p^+ and n^+), the inner two the depleted (space-charge, region B) and undepleted parts (quasi-neutral

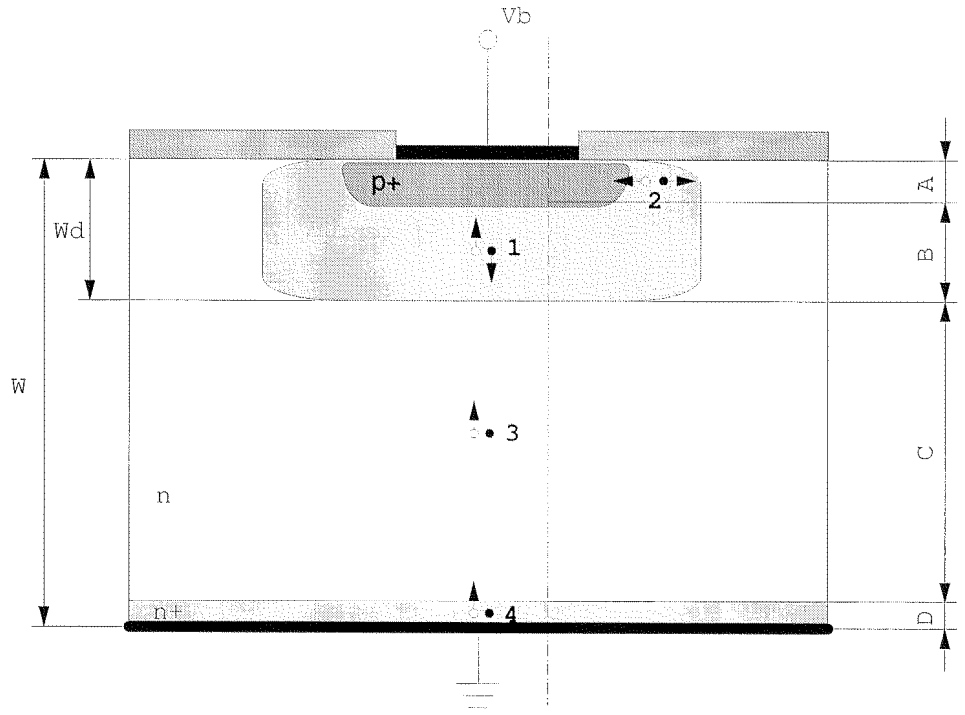


Fig. 3.19: Leakage current sources in semi-depleted silicon detector:(1) generation in space-charge region, (2) surface generation, (3) diffusion in the quasi-neutral region, (4) diffusion from the contact. The bulk is $n-$type silicon, $V_b < 0$ is the reverse bias.

region C) of the low-doped bulk. A reverse bias voltage which is applied to the metal contacts will appear across the pn junction since the potential drop across the *high-low* junction is negligible (cf. Sec. 3.3.2.c).

The thickness of the depleted bulk (see Fig. 3.19 for symbols and later for mathematical details, Eq. (3.52)),

$$W_d = \sqrt{\frac{2\epsilon_{Si}(V_b - \phi_i)}{qN_d}} \tag{3.42}$$

is proportional to the squared root of the applied bias voltage, V_b (cf. Fig 3.20), N_d is the doping of the bulk material and ϕ_i is the built-in potential at the metal-semiconductor interface. W is the physical thickness of the wafer.

In each region electron-holes pairs are thermally generated, which contributes to the total leakage current when collected at the contact, either by diffusion or drift.

In order to calculate the current in a fixed area, one has to solve the current and continuity equations, given in Eq. (3.26)-(3.29).

In this Section it will be assumed that both p^+n and n^+n junction are abrupt (cf. Sec. 3.2.2) with uniform doping densities together with full ionisation of dopants. Again a low-injection regime is assumed if not stated differently and detailed balance between majority and minority carriers across junction is assumed to be valid in steady state as well as under bias.

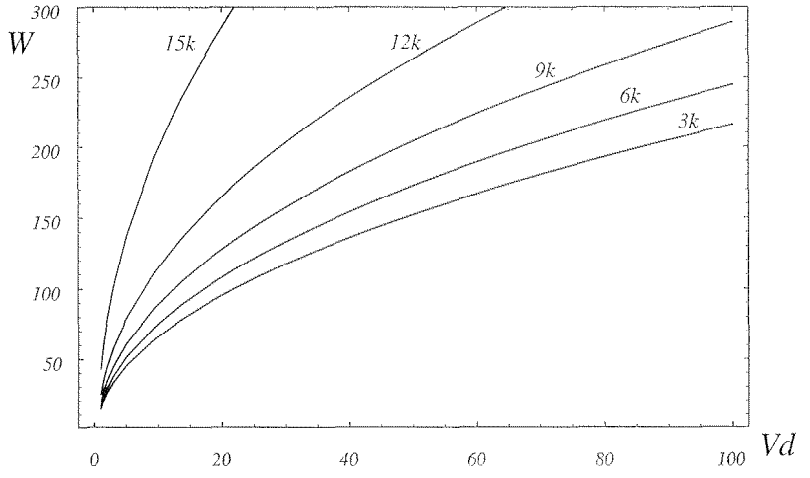


Fig. 3.20: W_d as a function of depletion voltage V_d. Different curves corresponds to different N_{dl}, here expressed in terms of resistivity, [Ohm/cm^2].

In this case the total current density in the structure sketched above is given by the sum of contributions due to J_{np}, diffusion current due to electrons in the p^+ region, J_{pnl} and J_{pnh}, the diffusion currents due to holes in the n and n^+ regions and J_g, diffusion current in the p^+n depletion layer.

For the sake of completeness, it is more interesting to derive the leakage current contribution of each region separately, cf. Fig 3.19.

3.3.2.a Region A: the diffusion current

The region A is undepleted, quasi-neutral. Thus, the current in this region is determined by the diffusion of minority carriers (here electrons) into this region. The general form of a diffusion based current for this region is:

$$J_{diff} = qs\Delta n_A \tag{3.43}$$

where Δn_A is the variation of the minority carrier concentration in region A with respect to its equilibrium value, and s is the surface recombination velocity. Assuming infinite surface recombination velocity ($s \to \infty$), the minority carrier concentration is always equal to its equilibrium value. This corresponds to a perfect ohmic contact and leads to $\Delta n_A = 0$ at the surface.

Using Eq. (3.27), assuming a constant doping profile and no degenerative effects, the contribution of this region to the leakage current of region A can be directly calculated [40] as the standard solution of Eq. (3.26)-(3.29):

$$J_A = q\frac{n_i^2}{N_A}\frac{D_A}{L_A \tanh(W_A/L_A)} \tag{3.44}$$

where D_A is the diffusion coefficient for this region, expressed by the Einstein equation (cf. Eq. (3.17)), $L_A = \sqrt{D_A \tau_{gA}}$ is the diffusion length and μ is the minority carrier mobility (electrons in region A); τ_{gA} is the generation lifetime in this region.

In most cases it is sufficient to use Eq. (3.44) for an estimation of current contribution from region A. However, in the case of very thin layers, this simple estimate for the current contribution is far too high, due to the simplifying assumption of a infinite surface generation velocity. In these cases the equation should be substituted by that of Schottky barrier [28].

3.3.2.b Region B: the generation current

The depleted part of the diode is referred as part B. For the sake of simplicity the generation rate is assumed to be the same in the entire layer [41], even if some caution must be exercised if depletion of the bulk is only partially developed: in the usual mode of operation for silicon detectors, i.e. full depletion or even overdepletion, the approximation is justified [42].

Given that, the current contribution of the space charge region B can be seen as the standard (thermal) generation current:

$$J_{gen} = -q \int_0^{W_d} U(x)dx \tag{3.45}$$

where $U(x)$ is the net balance of generation and recombination in the space-charge zone (see Eq. (3.20)). The integral can be approximated by:

$$\begin{aligned} J_B &\simeq qU_B W_d \\ &\simeq q\frac{n_i}{2\tau_d} W_d \end{aligned} \tag{3.46}$$

where τ_d is the generation lifetime in the space charge zone which has been assumed to be equal to the recombination one. From this formula it is clear that the generation lifetime depends on n_i as compared with the diffusion components which have a n_i^2 character. This characteristic is fundamental in the process of understanding the main causes of dark currents in a silicon detector. Fig. 3.21 represents the current at two different temperatures as a function of reverse bias voltage, τ_d =1 msec. As expected, the curves depend to V_b as squared root. This behaviour is observed in $I-V$ plot of real devices, especially at sufficiently high reverse voltage ($V_b \gg 2$Volt) and can be verified studying its behaviour as a function of $1/T$, as explained by Arrhenius [43]. If the conduction process is mainly due to thermal generation, then plotting the current, defined in Eq. (3.46), versus $1/T$, a slope corresponding to half the forbidden gap should be seen, see Fig. 3.22. This can be derived from the temperature dependence of the intrinsic doping concentration n_i:

$$n_i(T) = n_i(T_0)\left(\frac{T}{T_0}\right)^{\frac{3}{2}} \exp\left[-\frac{qE_g}{2kT}\left(1-\frac{T}{T_0}\right)\right]. \tag{3.47}$$

On the other hand, if the current is dominated by diffusion, according to Eq. (3.44) depending on n_i^2, then a change with slope E_g is to be found. We will come back to these considerations in the chapter devoted to the test of the devices.

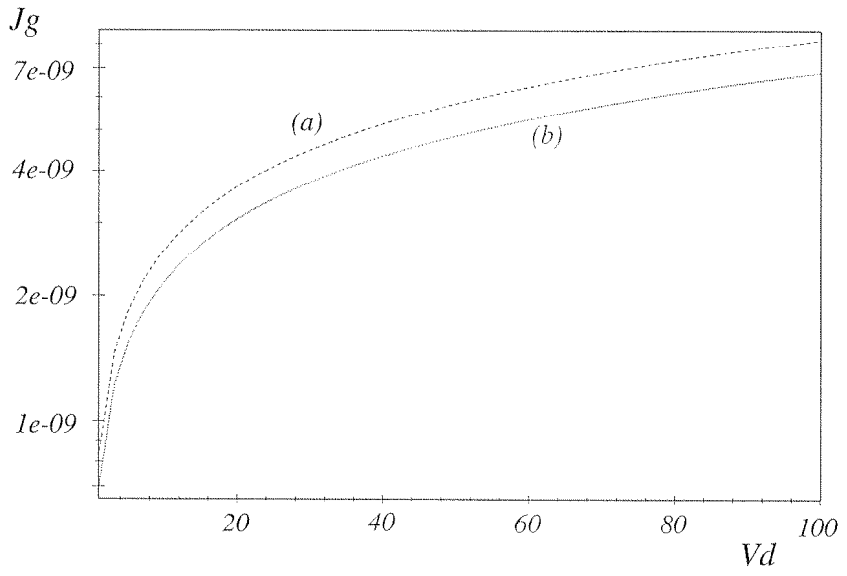


Fig. 3.21: Generation current ($[A/cm^2]$) as a function of reverse bias voltage at $380K$ (curve (a)) and $300K$ (curve (b)).

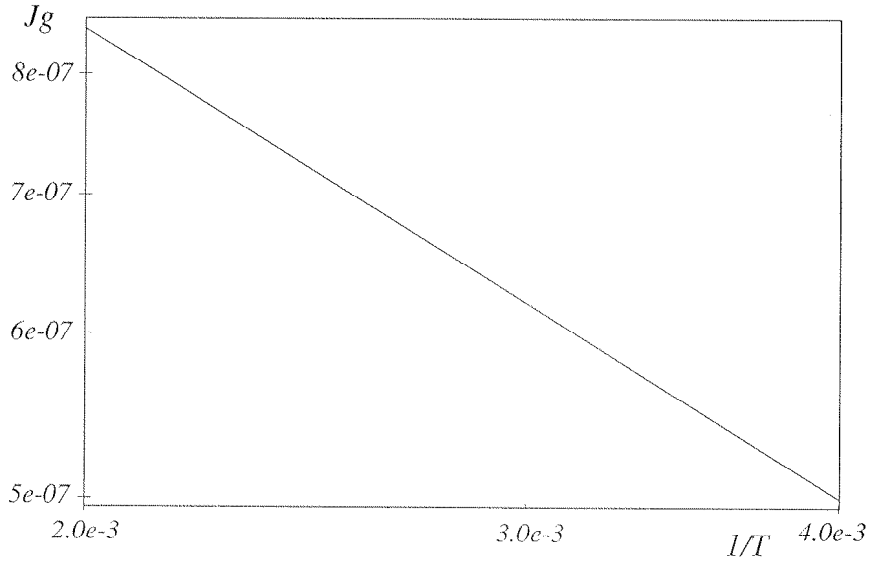


Fig. 3.22: Arrhenius plot for the generation current in the space-charge zone. Current is expressed in $[A/cm^2]$,temperature T in Kelvin.

3.3.2.c Region C and D: the high-low junction

Regions C and D are undepleted. As in region A, the diffusion of the minority carriers (here holes) is the dominant contribution to the leakage current. But because of the different doping levels of the two zones, it is convenient to model the contributions of these regions separately. The contact of two regions of the same type of doping but different densities (like region C and D) is called *high-low* junction and it has been object of extensive studies especially in the field of solar cells [44, 45].

In this section[8] it is shown that the *high-low* junction has similar effect on lowering the level of the leakage current of reverse biased diodes like silicon detectors.

The potential barrier present at the interface between the bulk n and the heavily doped n^+ region inhibits the flow of minority carriers coming from the depleted bulk towards the ohmic contact and then reduces the dark current: this is the cause of its denomination of "reflecting junction".

In fact it can be demonstrated (cf. Appendix A) that the minority carrier concentration in the "low" zone, p_C, relates with p_D in the following way (see Eq. (A.9)):

$$p_D \simeq p_C \frac{N_C}{N_D}.$$

This expression shows that the minority carrier concentration is reduced by a factor N_C/N_D in going from n to n^+ side of the *high-low* junction.

If the extra doping profile for region D was not present, the leakage current J_C (cf. Eq. (A.18)) would become very large as full depletion is approached. In fact, at full depletion, the diffusion current across the *high-low* junction J_{hl} saturates at (cf. Eq. (A.23)):

$$J_{hl}^{max} = -q \frac{n_{ie}^2 D_D}{L_D N_D} \coth\left(\frac{W_D}{L_D}\right).$$

This current density can be made smaller by simply increasing the doping density of region D, N_D. It is possible to compare the effect of the D region on the current, by comparing this situation with the case of no "high" region (i.e. no D region, only C-like): the ratio of the hole current in the same structure with or without a n^+ layer, measures then the effectiveness of the backside impurity concentration (cf. Eq. (A.24)):

$$\frac{J_{hl}}{J_C} = \frac{n_{ie}^2 N_C D_D W_C}{n_i^2 N_D D_C L_D} \coth\left(\frac{W_D}{L_D}\right)$$

It is clear that J_{hl} is minimized by increasing N_D and W_D/L_D, using therefore a highly doped, deep n^+ layer: physically it means that one has to avoid any holes from reaching the low region by making the width of the high region much larger than the hole diffusion length in this region. The presence of such a high doped layer avoids a strong increase of the leakage current driven by J_l (cf. Eq. (A.18)) and furthermore such *high-low* junction is responsible for the suppression of any carrier injection from the backside contact.

[8] In Appendix A are reported the analytical derivation of the main properties of *high-low* junctions.

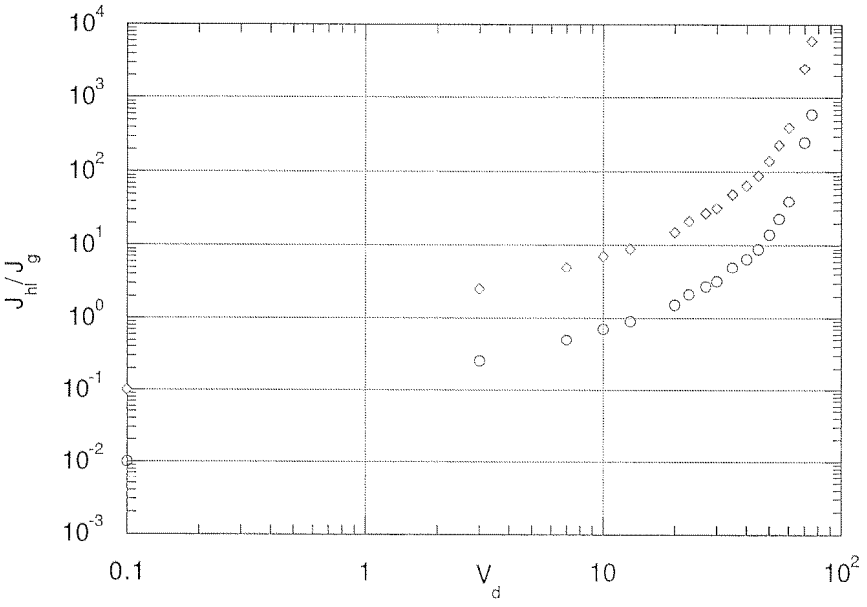


Fig. 3.23: Theoretical ratio between diffusion and generation current as a function of depletion voltage and recombination lifetime. Diamonds ($\tau_g = 1$ msec) and circles ($\tau_g = 10$ msec).

One should be conscious that, in a structure like that sketched in Fig. 3.20, at full depletion, assuming a bulk generation lifetime of $1 \cdot 10^{-3}$ sec, the generation current is $J_g \simeq J_B = 5$ nA/cm^2, still two order of magnitude larger than the diffusion current J_{hl}. In Fig. 3.23 is shown how the diffusion current can become dominant over the generation in the space-charge zone, especially at low bias (i.e. $W_d \ll W$) in case of high generation lifetime. In the section devoted to the generation currents (region B) has been highlighted how increasing temperature the current would increase (cf. Fig. 3.21); we should thus expect an interval of low depletion voltages in which $J_g > J_{hl}$ with a similar way of the previous case of higher τ_g.

3.3.2.d Conclusions

One important conclusion which can be drawn is that the effectiveness of the n^+ layer in reducing the hole diffusion component to the current strongly depends on the degree of "ohmicness" of the back contact (cf. Appendix A).

A practical consequence of this fact should then be that the quality of the surface of the silicon (i.e. the wafer lapping) should be not too good! A worse surface will have more trapping centers which will then increase the probability of recombination and then s_b. It seems a non-sense situation, especially in the semiconductor industry where purity and cleanness are pushed to their physical limits. However, since at the surface, both recombination and generation are controlled by these interface states, which do not have well defined energy level, the control of s_b, like the optimization of the backside doping profile, is a very challenging task. Nowadays there are some techniques used in semiconductor industry which deliberately makes use of backside damages, like scratching

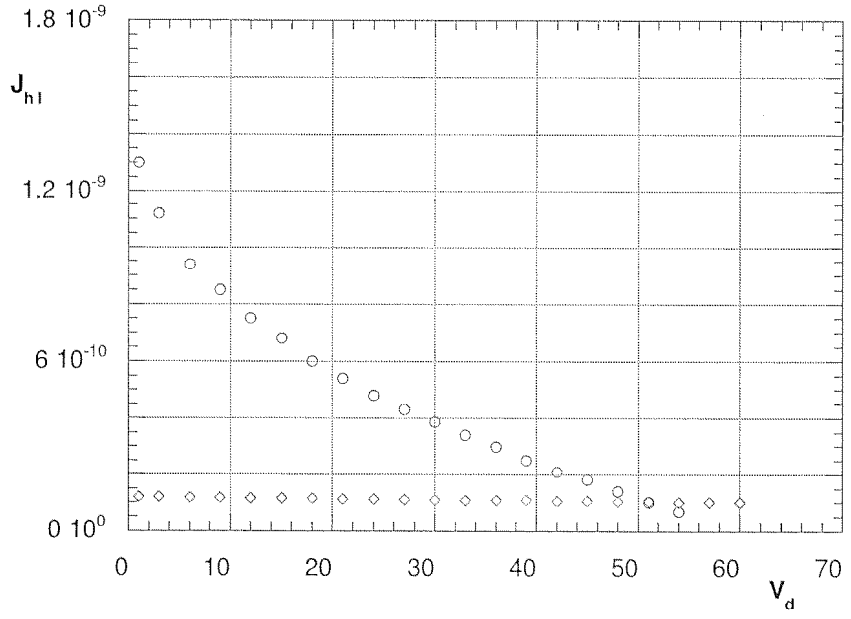


Fig. 3.24: $J_{hl}\,[\mathrm{A/cm^{-2}}]$ as a function of depletion voltage $V_b\,[\mathrm{V}]$. The circles represent the case of $N_C = 8 \cdot 10^{11}\mathrm{cm}^{-3}$ and $N_D = 1 \cdot 10^{20}\mathrm{cm}^{-3}$; the diamonds $N_C = 8 \cdot 10^{12}\mathrm{cm}^{-3}$ and $N_D = 1 \cdot 10^{20}\mathrm{cm}^{-3}$. Ideal ohmic contact is assumed.

the surface or depositing "aggressive" material directly on surface, in order to stimulate an higher recombination of minority carriers. Other possibilities are to use as metallic contact an element with an higher potential barrier, stopping diffusion. In the Chapter 4 I will analyze in detail some of these techniques. In any case, the leakage current is always reduced by the presence of a *high-low* junction.

	$N_i[cm^{-3}]$	$D_i[cm^2/sec]$	$\tau_i[sec]$	$W_i[\mu m]$
A	10^{19}	10	$1 \cdot 10^{-6}$	1.2
B	10^{12}	12.5	$1 \cdot 10^{-3}$	W_d
C	10^{12}	12.5	$1 \cdot 10^{-3}$	$W - W_d$
D	10^{20}	2	$0.5 \cdot 10^{-6}$	1.7

Table 3.5: Comparative table for main parameters used in leakage current definition in the four detector regions.

In Table 3.5 are listed some typical values for the main parameters in the four detector regions [46]. In case of full depletion, the most common operative condition for detectors, the contributions of region B to the current is the most important. In fact for a typical value for L_d of $1 - 2mm$, J_B is of the order of $5nA/cm^2$ to be compared with J_A which is in the range of $\approx pA/cm^2$. Assuming for the *high-low* region the situation of s_b finite as more realistic, the diffusion current contribution of region C, as explained at page 208, will be lower by some order of magnitude with respect to the ideal situation, sketched in Fig. 3.23. It follows that the contribution of region D, the *high-low* junction, can be significantly lowered, once the ratio N_{dl}/N_{dh} is kept as low as possible, descending to values well below of those from even a partially depleted region B. This effect, already

known as Back Surface Field [44], is enhanced in devices with diffusion lengths which are large compared to the substrate thickness, as in microstrip silicon detectors. It can be demonstrated that even in the case of partial depletion of the substrate $W_d < W$, the presence of the *high-low* junction helps to keep the contribution to the current of region C is below J_B.

In conclusion, the leakage current of a diode made on high-resistivity substrate can be modeled as the sum of a generation and diffusion current:

$$J_{tot} = q\frac{n_i}{\tau_d}W_d + q\frac{n_i^2}{N_{dl}}\frac{D_{pl}}{L_{pl}}\tanh\left(\frac{W - W_d}{L_{pl}}\right). \tag{3.48}$$

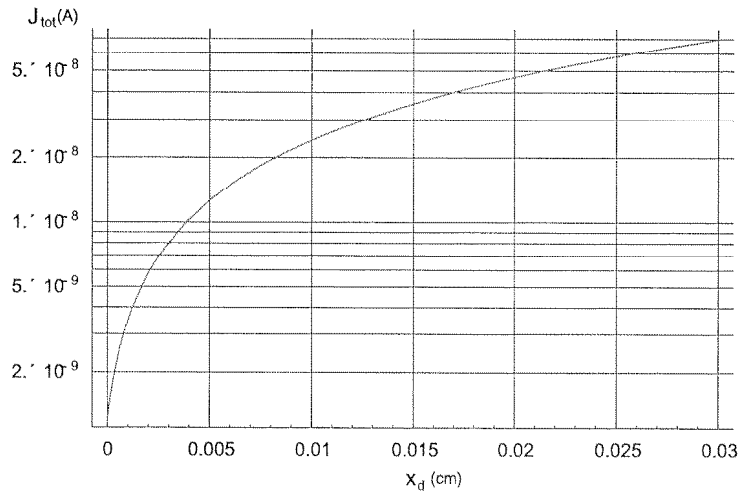


Fig. 3.25: Example of ideal behaviour of the total leakage current J_{tot} in a 1 cm^2 diode.

Up to now multiplication and breakdown effects are not taken into account due to their higher voltage threshold: we can only say that, following our ideal description, the current will increase with depletion voltage as far as full depletion occurs and then it will saturate.

It is useful to distinguish leakage current contributions as depletion layer dependent instead of diffusion and generation ones:

$$J'_{tot} = J_d \cdot W_d + J_0. \tag{3.49}$$

where J_0 is the W_d–independent contribution.

The diffusion current contributes to both terms since it is lowering with depletion voltage, whereas the generation current only contributes to the first one. If the generation in the depletion layer dominates, the leakage current is proportional to the depletion layer volume $A_g \cdot W_d$ and inversely proportional to the generation lifetime, (Eq.3.46). In Chapter 5 I will give some practical examples by using CMS diodes for extract both current contributions.

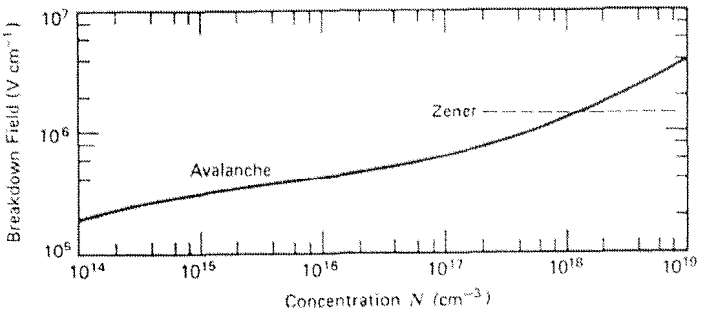


Fig. 3.26: Critical electric field for avalanche and Zener breakdown in Silicon as functions of dopant concentration [28].

The contribution J_0 can also be described as the slope, m, of the double logarithmic plot of the $I-V$ curve where a value of $\frac{1}{2}$ means no contribution, whether for $m \ll 1/2$ this contribution may be sizable. The factor m may be also greater than $\frac{1}{2}$ in the case of high reverse bias toward depletion and over, where the effective generation width approaches W_d with $W_d \simeq W$ and of course due to non-ideal junction phenomena. It is probably useful recalling Eq. (3.47) and the discussion made about the origin of generation current: it is clear now that if the current stems mainly from a diffusion process, according to their dependence on n_i^2, then a change in the Arrhenius plot with a slope E_g is to be found , instead of $E_g/2$ as it should be due to pure generation (cf. Fig. 3.22).

3.3.3 Breakdown Phenomena

When the electric field in the Silicon is increased above a certain value, called *critical field*, $\mathcal{E}_{bkd}$, the carriers gain enough kinetic energy to generate electron-hole pairs.

At high field one out of two breakdown (BKD) processes may be dominant: *Zener* breakdown or *avalanche* breakdown. The first case implies that the high field within the silicon can exert sufficient force on a covalently bound electron to free it: the electron tunnels through the energy barriers without interacting with any particle. This event has a consequence the creation of two carriers which contribute to the dark current of the diode. In the second breakdown process, free carriers may gain sufficient kinetic energy during their drift in the electric field that they can break covalent bounds in the lattice when they collide with it (*impact ionisation*). Every carrier create two more carriers which may accelerate and become themselves a cause for avalanching collisions, leading to a multiplication of carriers in the space-charge region when the field is large enough. In the case of silicon microstrip detectors operated in full depletion the impact ionisation is the major cause of breakdown.

An obvious requirements for silicon detectors is that their breakdown voltage is much higher than the voltage needed for fully deplete the detector volume. To analyze the

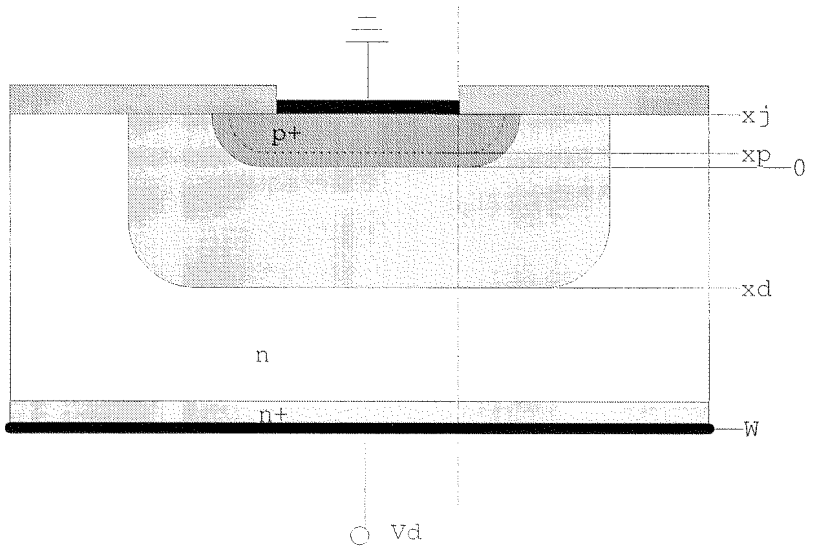


Fig. 3.27: Details of a $p^+ - n - n^+$ structure used in the breakdown calculations.

breakdown behaviour, the two cases of a parallel plane and of a cylindrical junction are reviewed: it will be clear that the breakdown characteristics are determined by the curved part of the junction. It should be noticed that in reality the theoretical breakdown voltage is never reached due to far-from-ideal conditions of the silicon surface and its internal structure: trapped charge at the $Si - SiO_2$ interface as well as the presence of impurities or defects in the lattice act as a combined source of midgap defects which are lowering the strength of the junction barrier. For the moment we will develop the ideal case, delaying the effects of impurities and lattice damage to a later section.

In the following, is assumed the step character of the junction, reflecting better the abrupt distribution of implanted dopants in a high resistivity diode even after thermal annealing.

The structure used for the breakdown computation is pretty simple and is shown in Fig. 3.27: it can be represented as a vertical p^+nn^+ diode, with the silicon bulk, intrinsic or not, acting as a bridge. For sake of calculation simplicity, the origin of $x-$axis is at the main junction, which stays at a distance x_j from the silicon surface and has a curvature r_j. $x_p(< 0)$ and $x_d(> 0)$ are the depletion region widths in respectively p^+ part and n^+ part of the junction. In the first case we assume an infinite flat junction. With regard to a reverse bias voltage V_b, we can solve the Poisson equation [23] using as boundary conditions the fact that the electric field is zero in the neutral regions. This holds to an expression for electric field across the junction region:

$$\mathcal{E}(x) = \frac{qN_{dl}}{\epsilon_{Si}}(x - x_d) = -\frac{qN_p}{\epsilon_{Si}}(x - x_p) \tag{3.50}$$

which depends linearly on the thickness of depleted silicon in both regions and having its maximum at $x = 0$, i.e. at the junction. Here the electric field must be continuous so that $N_p x_p = N_d x_d$. It is clear from this that the width of the depleted region on each side of the junction is inversely proportional to the dopant density: the higher

the dopant concentration, the narrower the space-charge region. In the usual case of microstrip detectors, the ratio N_p/N_d is of the order of $10^6 \div 10^7$, causing an highly asymmetric depletion mostly developed in the lowly doped bulk. Integrating again the field expression Eq. (3.50), we obtain the potential variation across the junction, mostly referred as built-in potential, ϕ_i:

$$\phi_i = \phi_n - \phi_p = \frac{kT}{q} \log\left(\frac{N_p N_d}{n_i^2}\right) \tag{3.51}$$

Typically in a silicon detector the built-in potential is of the order of $0.27 - 0.28V$ and the major portion of potential change occurs in the region with the lower dopant concentration. The extension of the space charge zone will be as in the previous section (cf. Eq. (3.42)) therefore:

$$W_d = x_p + x_d \simeq x_d = \sqrt{\frac{qN_d}{2\epsilon_{Si}}(\phi - V_d)} \tag{3.52}$$

In case of full depletion ($x_d \geq W$) it is useful to rewrite the field as the sum of two parts, the first referring to the voltage V_{rt} necessary to extend the space charge zone up to the back contact (process often referred as *reach-through*) and the latter to the eventual overdepletion voltage $V_{od} = V_d - V_{rt}$:

$$\mathcal{E}(x) = \frac{qN_d}{\epsilon_{Si}}(x - W) + \frac{V_{od}}{W}. \tag{3.53}$$

From this equation is evident that after full depletion the field is growing linearly with the applied bias V_d, up to its maximum,

$$\mathcal{E}_{max} = -\frac{qN_d}{\epsilon_{Si}} W \left(1 + \frac{V_{od}}{2V_{rt}}\right). \tag{3.54}$$

We can extend these results to the case of a cylindrical junction, assuming that the lateral depletion is independent from the vertical one: after full depletion the vertical depletion stops whether the lateral one does not. In [47] the Poisson equation is solved in cylindrical coordinates, giving:

$$\mathcal{E}(r) = -\frac{qN_d}{2\epsilon_{Si}}\left(\frac{r_d^2 - r^2}{r}\right) \tag{3.55}$$

valid for $r_d \geq r \geq r_j > r_p$. Here r_d is the radial development (radius of curvature) of the depletion region in the low resistivity bulk, r_j is referred to the main junction and r_p to the p^+- implantation. The radial development is

$$r_d = \frac{W}{\left[\log\left(\frac{r_d}{r_j}\right)\right]^2}. \tag{3.56}$$

which is an iterative equation which can be solved by means of a suitable normalization around an approximated guess.

The maximum field, which is the most interesting situation for breakdown calculations, occurs for small r, more precisely at the junction $r - r_j = x_j$, giving [48]:

$$\mathcal{E}_{max}(r) \approx V_d \frac{r_d^2}{rW^2}. \tag{3.57}$$

In both cases, the depletion-region width and the maximum electric field in a $p-n$ diode increase as reverse bias increase. Physically, one expect that these quantities should reach a maximum somehow. At high voltages, some of the material making up the device, such as the insulating layers of silicon dioxide or nitride, may rupture, or the current through the junction may increase fast. In the first case, the damage will be irreversible and the device stops functioning. In the case of breakdown of the reverse-biased junction barrier, impeding current to flow through, the damage is generally not destructive, at least unless the high current involved melt a portion of the silicon. The voltage at which the breakdown will take place,V_{bkd}, depends on the structure of the junction and on the doping distribution.

As stated before, the electric field is low at the edges of the space charge region, so the avalanche phenomena is confined to the portion of the depleted bulk closer to the junction. Following [49] we can define n_i and n_f as the number of carriers respectively entering and leaving the depletion region after the multiplication had occurred:

$$n_f - n_i = n_f \int_{x_i}^{x_f} \alpha dx \tag{3.58}$$

where α is the ionisation coefficient, expressing the density of probability that a carrier creates an electron-hole pair passing through the high field region; x_i and x_f are respectively the limits of the depletion region. It is possible to define a multiplication factor M as the ratio of n_f/n_i:

$$M = \frac{n_f}{n_i} = \frac{1}{1 - \int_{x_i}^{x_f} \alpha dx}. \tag{3.59}$$

This parameter tends to infinity, implying infinite multiplication, as the integral at denominator tends to one. Avalanche breakdown is defined then as:

$$\int_{x_i}^{x_f} \alpha dx = 1 \tag{3.60}$$

Following [50] α is defined as:

$$\alpha = K \exp\left(-\frac{B}{\mathcal{E}}\right) \tag{3.61}$$

with K ad B dimensional constants: for electrons $K = 7 \cdot 10^5\,\mathrm{cm}^{-1}$ and $B = 1.23 \cdot 10^6\mathrm{V}$ cm^{-1}. Here is stated the dependence from the electric field, which is reasonable due to its influence on the carriers drift velocity and then on their kinetic energy, enhancing the probability of carrier multiplication. Same authors [51] prefers an approximated expression for Eq. (3.61),

$$\alpha \simeq 1.8 \cdot 10^{-35} \mathcal{E}^7 \tag{3.62}$$

Derivation of the breakdown characteristics In this paragraph a derivation of BKD performance characteristics of an ideal junction is done: a gradual approach is chosen, looking first to uniformly doped zones, then to implantation edges and finally to implant corners.

1. One dimensional case: the planar junction.

   In the case of an infinite plane junction on a substrate of finite W, we can study the BKD characteristics using Eq. (3.53) with the help of Eq. (3.60): following Kurana [52],

   $$V_{bkd}(x) = V_{rt} + \frac{bW}{\ln\left[KW\left(\frac{V_{bkd}-V_{rt}}{2V_{rt}}\right)\right]} \tag{3.63}$$

   where it is clear that the breakdown voltage depends in good approximation linearly from the detector thickness.Using the approximation (3.62), we can extract V_{bkd} and $\mathcal{E}_{bkd}$ as

   $$V_{bkd} \simeq 5.34 \cdot 10^{13} N_d^{-3/4} \tag{3.64}$$

   $$\mathcal{E}_{bkd} = 4010 N_d^{1/8}. \tag{3.65}$$

   In the case of a detector with a high-low junction at the backside, once the depletion voltage reach the reach-through value V_{rt}, the depletion zone does not extend significantly any more. Since W is fixed, increasing the bias voltage will just rise the electric field at the main junction, x_j, and also its value at the high-low junction. Whether $\mathcal{E} = \mathcal{E}_{bkd}$ the expression (3.64) should be modified as [51]:

   $$V_{bkd}(x) = \mathcal{E}_{bkd} W - \frac{qN_d W^2}{2\epsilon_{Si}} \tag{3.66}$$

   In Fig. 3.28 is shown V_{bkd} as a function of the substrate doping density N_d.

2. Two dimensional case: the cylindrical junction.

   Up to now the approximations made are consistent with one dimensional calculations of the BKD characteristics, setting the best possible scenario and its asymptotic conditions. In a microstrip detector the patterning of the implanted doping for the strip structure definition, causes several discontinuities of the implantation in the plane parallel to the surface. The implanted ionic species will spread over their mask defined zone due to the high temperature thermal activation (see Fig. 3.27). It can be shown that the lateral diffusion of the junction r_j on the CSEM process is close to 80% of the junction depth x_j: in order to simplify calculations it will be is assumed that $r_j = x_j$. Following [48, 51], it is possible to derive V_{bkd}:

   $$V_{bkd}(r) = \frac{Br_j W^2}{r_d^2 \ln\left(\frac{r_d^2 K V_{bkd}}{BW^2}\right)} \tag{3.67}$$

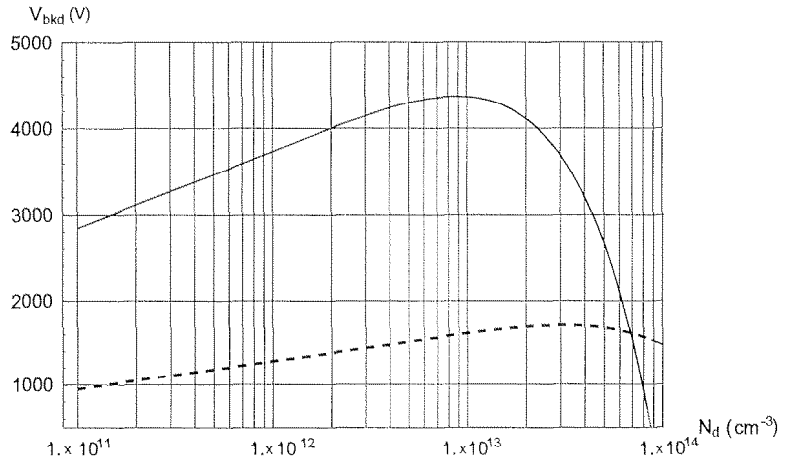


Fig. 3.28: Breakdown voltage for p^+ on $n-$type silicon detectors with high-low junction on the backside as a function of bulk doping density. Infinite plane junction is assumed. Comparison between 100 μm (dashed line) and 300 μm (continuous line) thick wafers.

This formula shows the same behaviour as in the planar case: V_{bkd} is proportional of $r_j \simeq x_j$. The radial breakdown is independent of the substrate doping concentration since its contribution at full depletion cancels in the ratio W/r_d. In Fig. 3.29 is plotted V_{bkd} as a function of the lateral displacement of the junction: it is interesting to see how the breakdown improves when the junction is smoother. It is possible to compare both cases of planar and cylindrical junctions through the ratio of the maximum electric fields [53]:

$$\frac{\mathcal{E}_{max}(r)}{\mathcal{E}_{max}(x)} \approx \frac{r_d}{2r_j} \tag{3.68}$$

Increasing the depletion voltage, r_d increases to values much bigger than $2r_j$, thus increasing the electric field peak value. Moreover, the shallow the junction, the lower the breakdown, since $\mathcal{E}_{max}$ will rapidly increase. Assuming for $\mathcal{E}(r)$ a dependence $\propto \frac{1}{r}$ it is possible to relate the two critical voltages [53] as follows:

$$\frac{V_{bkd}(r)}{V_{bkd}(x)} = \left(1 + \frac{1}{2}u^{7/3}\right)\log\left(1 + 2u^{3/4}\right) - u. \tag{3.69}$$

where $u = \left(\frac{r_j}{W}\right)^{6/7}$ and it is plotted in Fig. 3.30.

3. The 3D case: the junction corners.

An obvious extension of the cylindrical junction is needed whenever the patterning of a strip or a guard ring needs to be bent. Wherever there is a corner, the junction can be seen as a portion of an hemispherical profile. Following Da Rold [53] it is possible to adapt Eq. (3.69) to the spherical case as

$$\frac{V_{bkd}(\Phi)}{V_{bkd}(r)} = u\left(2.14 + u^{4/3}\right) - u^{13/9}\left(1 + u^{4/3}\right)^{2/3}. \tag{3.70}$$

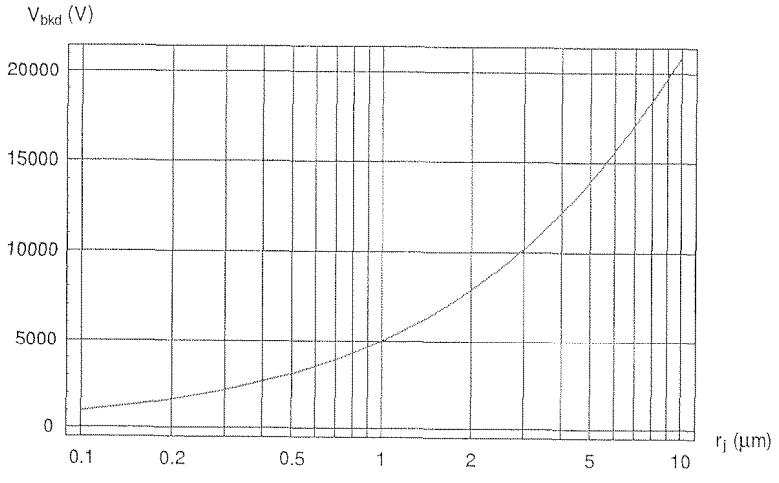


Fig. 3.29: Breakdown voltage in the cylindrical case as a function of the radial development of the junction.

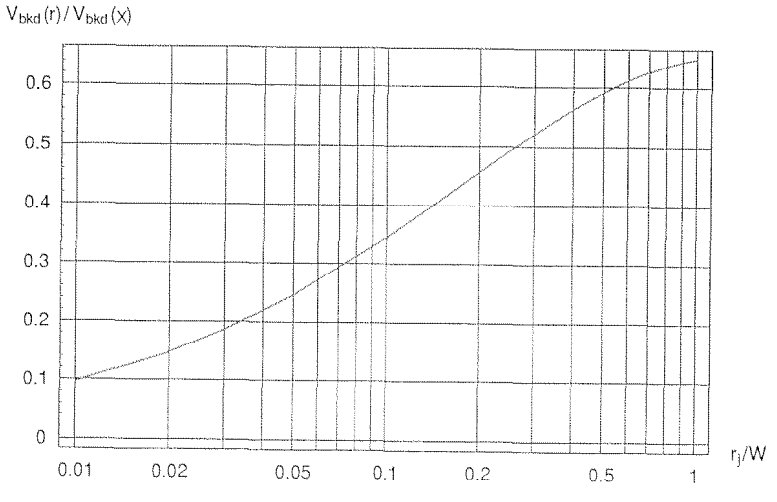


Fig. 3.30: Normalized breakdown voltage of cylindrical junction as a function of the normalized junction radius of curvature $\frac{r_j}{W}$.

In Fig. 3.31 is plotted the comparison between the normalized breakdown voltage V_{bkd} in the spherical and cylindrical case as a function of the curvature of the junction. It is clear that especially at low radii, i.e. shallow junctions, the critical voltage is a factor ten lower in the spherical junction case, meaning that the breakdown in ideal conditions will start at the implant corners. The relative balance between these two junction terminations can be seen in Fig. 3.32: for $r_j < 100\mu$m (assuming 300 μm thick substrate) the spherical termination will be the primary cause of a structure breakdown. Since in most practical cases r_j is much lower than 100 μm , we expect most of detector breakdowns happen at the corners or in the bent zones.

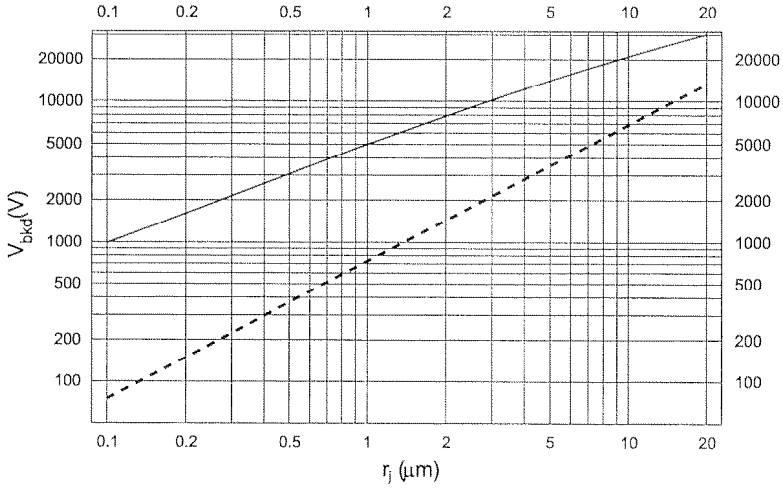


Fig. 3.31: The breakdownV_{bkd} (in V) for spherical (dashed line) and cylindrical (continuous line) junction termination as a function of the junction curvature r_j (in μm).

Breakdown instabilities An interesting effect often seen in measurements is the so called *walk-out* breakdown [54]. The overall effect is that after successive operations in avalanche mode the breakdown voltage increase. This is due to the injection of electrons in the oxide close to the junction edges: this accumulation of negative charges temporarily acts as a local MOS effect, thus relaxing the field lines in the most critical area. This effect will saturate due to the increase of the barrier height at the interface Si-SiO_2: after a period of annealing (no voltage applied across the junction, no temperature changes) the breakdown voltage will drift again back to its initial threshold. This phenomena is mainly caused by the trapping of electrons in water molecules which are present within the oxide layer with a concentration which depends on the type of oxide and on its deposition e/o growth method. Even if it can be seen as a beneficial effect, this can causes severe instabilities in the behaviour of the device and it should be avoided. In Fig. 3.34 is shown the drift of the breakdown voltage of a $p^+ - n$ junction at time-delays from 0 to 10 hours after the application of 100V bias: the effect of the drift of negative charges towards the

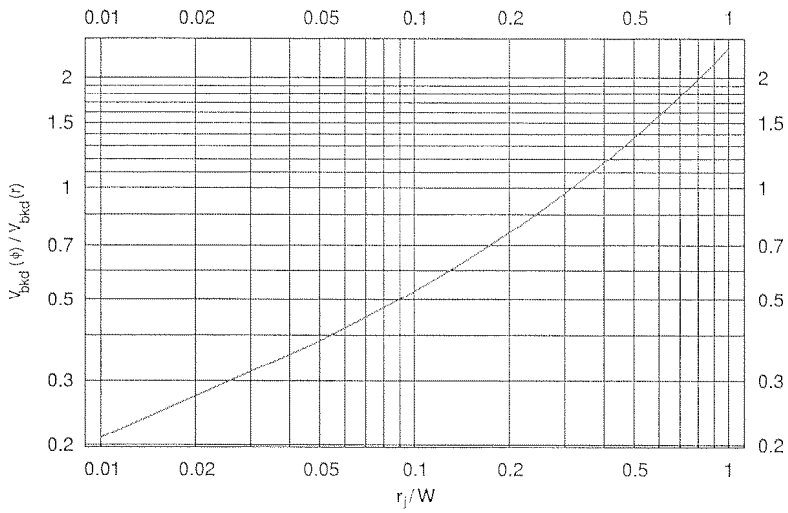


Fig. 3.32: Ratio of spherical and cylindrical breakdown voltages. In abscissa the junction curvature normalized to the thickness of the substrate W.

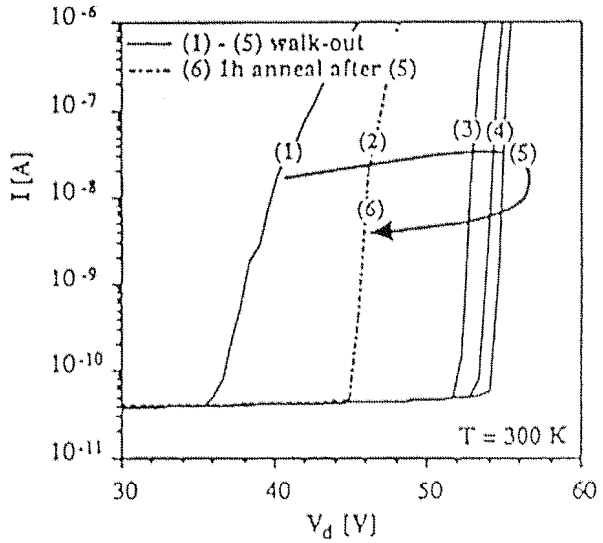


Fig. 3.33: Breakdown instability in a $1mm^2$ diode. From [54].

Si-SiO_2 interface is evident. Should be noted that the drift in V_{bkd} is only relevant in conditions of sufficient humidity: Longoni [55] shows that in dry or ambient humidity conditions ($\leq$ 40% humidity) the carrier migration through the oxide is frozen out or at least is characterized by a very long time constant (many hours). Rising the humidity ($\geq$ 80%), the system relaxes to equilibrium conditions, for a given bias condition, in a reasonable time and BKD improvements take place.

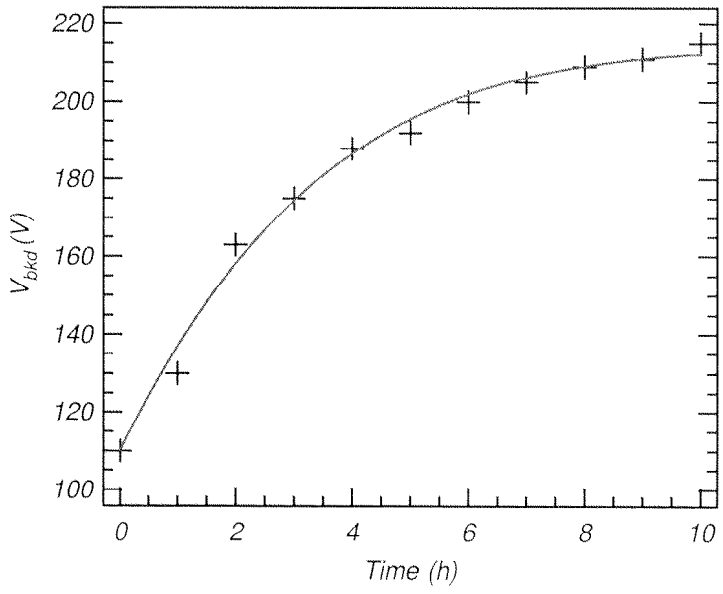


Fig. 3.34: Drift of the BKD voltage of a p^+–junction on 10 $k\Omega$·cm n–type substrate. The dots refer to BKD values after 0 to 10 hours after the application of 100 V reverse bias. Ambient humidity conditions.

Conclusions Previous calculations show clearly that the most delicate region with respect to BKD in a $p^+ - n$ silicon device are the implantation corners. Here the critical voltage may be hundred time lower than in the planar case and unsafe conditions may be reached even before full depletion. It is therefore clear that the layout rules used in designing the device must define a lower limit for implantation width, so that in the case of the shallower junction profile, the corner effect does not transform in a tip-effect! Good habit is to smooth implantations corners in the design, rounding them whenever the design rules allow it.

As mentioned before, the real values for breakdown voltages are lower than theory suggests: defects, impurities and radiation damage acts as sinker for carriers and because at those sites the field is higher, breakdown occurs when the field there is high enough to free carriers from associated traps. After being released these carrier can retain enough energy to start avalanching. At the surface the exact shape of the depletion layer determines the shape of the surface components of the electric field. As will be demonstrated

later, cf. 4.4 and 5.1, at the surface of a detector made of $n-$type silicon, there is an accumulation layer of electrons which prevents the silicon surface from depleting, bending the potential lines towards the diode. In a $p^+ - n$ detector this causes a higher gradient for the field, with lines which terminate almost perpendicular to the surface, increasing the possibility of impact ionisation, lowering the breakdown threshold. One possible solution would be trying to smooth the behaviour of these field lines close to the surface, making them parallel to the surface, as in the case of planar junctions. Two solutions are proposed up-to-now: guard ring structures and field plates.

3.3.3.a Guard rings

The first idea is to surround the active diode with other diode-like structures left mostly floating which act as a voltage divider. By increasing the reverse bias, punch-through

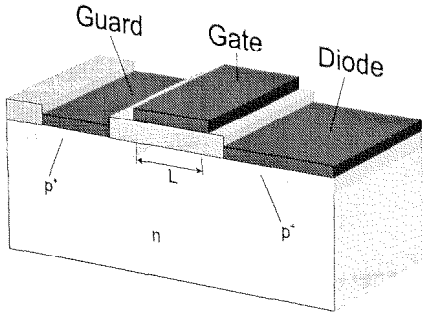


Fig. 3.35: Pictorial view of a guard ring and an active diode. A gate structure is drawn.

(PT) conduction is established between the active diode and the guard[9], self biasing the floating ring at an intermediate potential with respect to that of both sides of the $p^+ - n$ junction. The depletion region reaches then the floating ring, acting as an equipotential region and spreads the depletion region to a wider area: the main junction (e.g. the active strip) has to resist the voltage difference with the ring and the ring has to resist the remaining part of the applied voltage. The electric field at junction edges (i.e. close to the surface) is thus decreased and consequently the avalanche breakdown voltage is increased. A straightforward extension of this principle is to multiply the number of guard-ring structures surrounding the active diode. In order to achieve a real improvement in the

[9]The punch-through effect is very well known in the MOSFET devices field. In fact this is one of the limiting factor in scaling down device dimensions. As the channel length L is reduced, the depletion layer widths of the source and the drain junction of the transistor become comparable to the channel length. Using Eq. (3.38) the width of the drain junction W_D and that of the source junction W_S are

$$W_S = \sqrt{\frac{2\epsilon_{Si}}{qN_A}(\phi_i + V_b)}$$
$$W_D = \sqrt{\frac{2\epsilon_{Si}}{qN_A}(\phi_i + V_D + V_b)}$$

where V_b is the substrate bias.When $W_S + W_D = L$, the two depletion layers will merge and PT takes place, causing the gate to lose the control of the current.

V_{bkd} it is necessary to carefully study the geometry and the working conditions of the multi-guard structure. In fact, the potential barrier between each pair of diffusions which set the PT threshold depends strongly on the relative distance, on the biasing conditions and on the device leakage current. For more details and analytical calculations see [51, 53].

Recently at CSEM we participated to the design, production and test of a series of test structures devoted to the optimization of the spacing between the detector and the guard rings. Results are very encouraging and they will be illustrated in Sec. 5.2.1.a.

In devices developed to resist high doses of radiation, a tremendous depletion voltage should be applied in order to fully deplete the substrate after type inversion, see 3.3.5: this implies the design of many concentric guards at precise distances each other in order to "digest" up to $1kV$ over $300\,\mu$m without breakdown.

3.3.3.b Field plate

The use of field plates is useful as well. A field plate is thought as the metal contact to the active part of the diode, which often is the readout metal, left extending on top of the oxide outwards, well beyond the projected distance of the implantation, i.e. the junction (cf. Fig. 3.36).

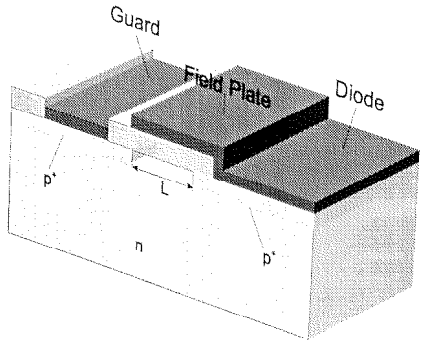


Fig. 3.36: Pictorial view of a field plate structure.

The surface of a silicon device is always passivated with a SiO_2 layer (see later in Sec. 4.4): at or near the interface resides a variety of fixed and mobile charges. In the most common case of p^+ on $n <111>$ silicon detectors the net contribution of these charges translates in a density of $10^{11} - 10^{12}\text{cm}^{-2}$ mobile electrons at the silicon surface. Having an electron accumulation layer so close to the edges of the $p-n$ junctions, prevents the silicon surface from depleting, narrowing the space charge region. Therefore the electric field is higher increasing the probability of impact ionisation. In Figs. 3.37 and 3.38 the development of the high field zone at the interface Si-SiO_2 of the strip junction edge (highlighted in the figures by the thicker line in the silicon bulk) by rising the depletion voltage, is shown.

If properly grounded or polarized, the field plate deplete the surface beneath the oxide just close to the junction (also known as *MOS effect*), smoothing the potential lines and relaxing the electric field.

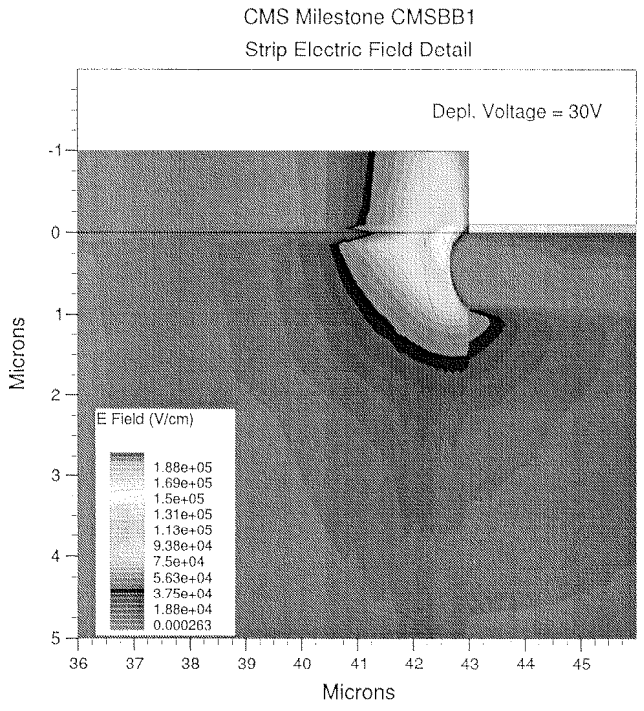


Fig. 3.37: Simulation results of a detail of the electric field at the strip edge. Situation at 30V, before full depletion. Trapped charge density of $3{\cdot}10^{12}\mathrm{cm}^{-2}$ has been used.

In high voltage operation this kind of structure may encounter some problems due to dielectric breakdowns or can induce noise on the signal, often referred as micro-discharge problem [56], but if some design recipes are respected, this potential problem may be avoided, at least its threshold is shifted to higher voltage regimes.

3.3.3.c Edges

One of the critical regions is the cut edge of the detectors. Polarizing the active zone with an increasing bias voltage, the space charge region extends laterally even after full depletion and the injection in the SCR of an amount of charge carriers might be a source of leakage current increase.

The cut border results in a heavily damaged surface: here a large quantity of dangling bonds, traps and defects acts as a source of free carriers which can be injected within the active part of the detector if the SCR reach the cut border. An obvious but impractical solution would be to keep as large as possible the distance from the active part to the cut line. A viable solution is to perform n^+ implants across the scribe line: here the SCR is obliged to "shrink" itself within a slightly smaller volume beneath the junctions, without spreading out towards the damaged surface.

The best solution for the sake of BKD performances would be a surface depleted of mobile charges (electrons) up to the scribe edges. This can be attained by means

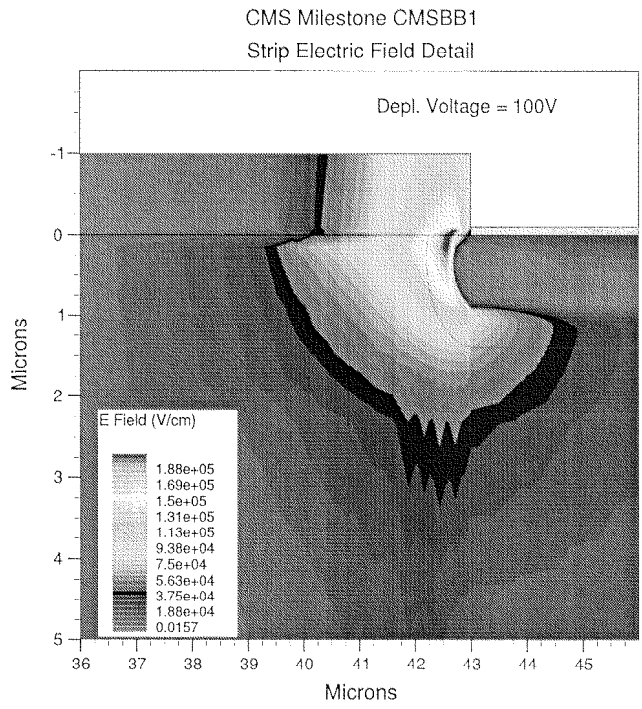


Fig. 3.38: Simulation results of a detail of the electric field at the strip edge. Situation at 100V, after full depletion. Trapped charge density of $3{\cdot}10^{12}\mathrm{cm}^{-2}$ has been used.

of a low dose shallow p^+implant which superimposes to those of the multiguards and extends towards the dicing line. This option combined with the n^+implantation across the cut as described above, with an accurate study of the geometrical disposition increase significantly the high-voltage performance of the device, as will be proved later in the Chapter devoted to the test of real devices.

Conclusions Nowadays in designing a silicon strip sensor suitable for high radiation environments, a combination of these techniques is currently used. In Chapter 5 will be resumed the results of an important program of prototypes design and developed at CSEM for the CMS SST purposes.

3.3.4 Microstrip Capacitance

The capacitance of silicon microstrip detectors is an important quantity which is a major limiting factor to the noise, especially for fast systems. As described in the section devoted to the position resolution of MSD, charge sharing and its improvement of spatial resolution of systems with analogue readout, are based on the delicate balance among capacitances of a strip against the bulk (i.e. the back contact), the neighboring strips and the eventual decoupling capacitance (if the detector is AC coupled with the electronics). Therefore a careful design of the geometry of the strips is necessary to ensure good detector performances. The contribution of non-geometrical sources to the detector capacitance, as surface charge or environmental humidity, can be determined only once the geometrical value is established. In the following are reported the results of a geomet-

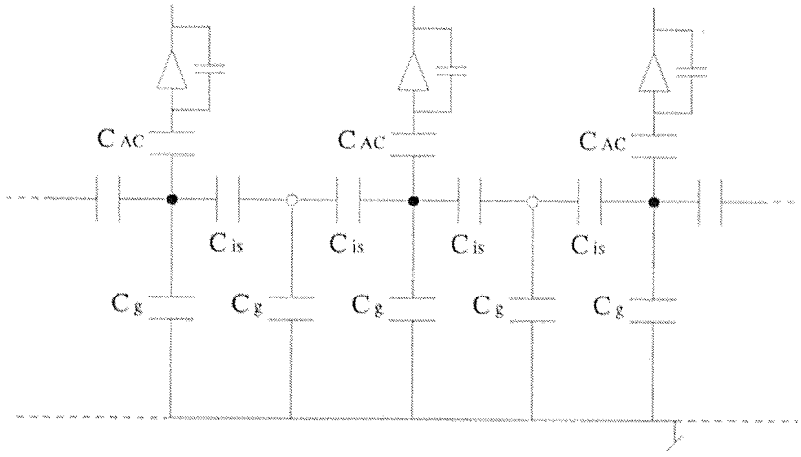


Fig. 3.39: Electrical scheme of the different capacitance present in a microstrip detector. Filled points represents readout strips, empty points floating strips(not directly connected to the electronics). C_{AC} is the decoupling capacitance of a strip, C_{is} is the interstrip capacitance and C_g is the capacitance to the bulk.

rical calculation of microstrip capacitances. These results will be confronted later with a 2D simulation performed by means of full resolution of Poisson and transport equations for a strip-like structure. This approach takes into account (i) a suitable description of carrier mobilities, by means of a concentration dependent model [57] and (ii) an accurate parameter characterization of the SHR generation-recombination model. This allows a complete description of both geometrical and non-geometrical contributions to the total strip capacitance. Knowing the geometrical results, the effect of surface charge is studied.

3.3.4.a Geometrical capacitances

The calculations of the wiring capacitances in the integrated circuit field has been well exploited in the seventies in order to propose viable design guidelines for the design of microwave circuits in VLSI technology [58, 59]. Since those numerically calculations are

mathematically very rigorous and CPU intensive, a number of approximated methods, well suited to be integrated in a CAD program have been proposed [60]. Sakurai [61] makes use of an empirical formula tuned on the numerical data formerly cited. The total capacitance per unit length of the central of three lines of width w, thickness t, spaced s (thus with a pitch p), lying on a silicon substrate of thickness W, is

$$\frac{C_{tot}}{\epsilon_{Si}} = \frac{C_g}{\epsilon_{Si}} + 2\left[0.03\left(\frac{w}{W}\right) + 0.83\left(\frac{t}{W}\right) - 0.07\left(\frac{t}{W}\right)^{0.2222}\right] \cdot \left(\frac{s}{W}\right)^{-1.34} \tag{3.71}$$

$$\frac{C_g}{\epsilon_{Si}} = 1.15\left(\frac{w}{W}\right) + 2.80\left(\frac{t}{W}\right)^{0.222}. \tag{3.72}$$

The formula is compact, easy to handle and the error[10], especially for very small w/W and $w/p \simeq w/s$ as in the case of microstrip detectors, is well below 10%: the problem is that can not differentiate among the different contributions of the strip capacitance, e.g. C_{is} and C_g. Kötz *et al.* [62] studied the capacitive charge division in a microstrip detector by means of a purely electrostatic method: every strip is approximated through three equally spaced wires with radius ρ so that the distance between the first and the third one is equal to the strip width, while the ground plane, i.e. the back plane, is substituted through the method of images by the same array of wires lying symmetrically respect to the plane itself at a distance of $2W$. With a suitable choice of initial conditions, it is possible to solve the linear system correlating the charge on each strip to the voltage of all the others:

$$\Psi_{ij} = \frac{\lambda_j}{2\pi\epsilon_0\epsilon_{Si}} \ln\left(\frac{r_{ij}}{\sqrt{r_{ij}^2 + 4h^2}}\right) \tag{3.73}$$

$$\Psi_{ii} = \frac{\lambda_i}{2\pi\epsilon_0\epsilon_{Si}} \ln\left(\frac{\rho}{2h}\right) \tag{3.74}$$

where Ψ_{ij} is the potential induced on the $i-$strip by a charge λ_j at the jth wire and its image charge.

Cattaneo [63] confirmed later the previous results proposing the application of the conformal transformations [64] to solve the capacitance problem in a system with translational symmetry along the z axis. This method allow to transform in the complex plane a complicate configuration of conductors into another that can be analytically solved, like the case of two concentrically cylinders or two parallel planes. The results for the interstrip capacitance are the following:

$$C_{is} = \epsilon_0(\epsilon_{Si} + 1)K(k)/K'(k) \tag{3.75}$$

[10]Here the error is calculated as the distance of the approximated method from the numerical calculations [58, 59]

where

$$\frac{K'(k)}{K(k)} = \begin{cases} \frac{1}{\pi}\ln\left(2\frac{1+\sqrt{k'}}{1-\sqrt{k'}}\right) & \text{for } 0 \leq k \leq 0.7, \\ \frac{\pi}{\ln\left(2\frac{1+\sqrt{k}}{1-\sqrt{k}}\right)} & \text{for } 0.7 \leq k \leq 1 \end{cases} \tag{3.76}$$

$$k = \frac{\tan\left(\frac{\pi}{4}\frac{w}{p}\right)}{\tan\left[\frac{\pi}{2}\left(1-\frac{w}{2p}\right)\right]} \qquad \text{and} \qquad k' = \sqrt{1-k^2}$$

in Fig. 3.40 is shown C_{is} as a function of w/p compared with the calculation made for the total strip capacitance C_{tot} (i.e. including the contribution of C_g, the capacitance to backplane).

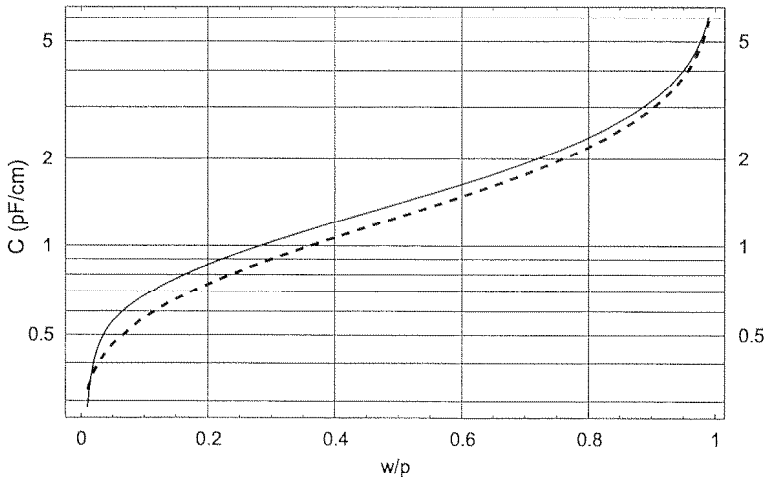


Fig. 3.40: C_{is} (dotted line) and C_{tot} (full line) as a function of w/p.

Hall *et al.* [65] approach is based on the variational determination of the line capacitance in the Fourier domain, using the charge density function as trial function. The general result is

$$\frac{1}{C} = \frac{1}{\pi Q^2 \epsilon_0}\int_0^\infty [F(\beta)]^2 G(\beta) d\beta \tag{3.77}$$

where $F(\beta)$ is the Fourier transform of the charge distribution $f(x)$ and Q provides the normalization. $G(\beta)$ and $f(x)$ are defined as

$$G(\beta) = \frac{1}{|\beta|\left(1+\epsilon_{Si}\coth[|\beta|h]\right)} \tag{3.78}$$

$$F(\beta) = \frac{q}{\beta}\left[2T\cos(\beta p) - 1\right]\cdot\left[\frac{2}{\beta}\left(1-\cos\left(\frac{\beta w}{2}\right)\right) - w\sin\left(\frac{\beta w}{2}\right)\right] \tag{3.79}$$

where $F(\beta)$ is suited for a multistrip structure. The best estimate of the capacitance of the central strip to all other conductors is obtained by maximizing Eq. (3.77) by varying

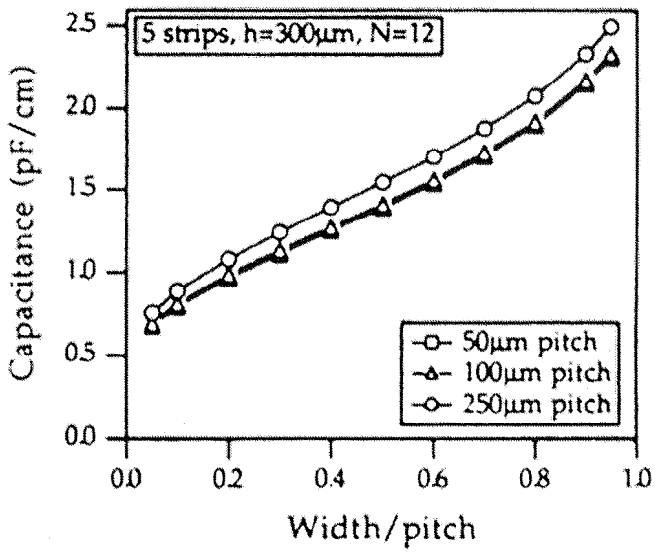


Fig. 3.41: Total strip capacitance as a function of w/p for different pitch values. From [65]

the charge fraction on each strip; the charge fractions on each strip allow the calculation of the neighboring capacitances. Requiring charge neutrality allows the computation of the strip versus the backplane. In Fig. 3.3.4.a is shown the total strip capacitance as a function of the ratio width over pitch w/p.

All the mentioned methods are consistent and sufficiently precise to estimate the capacitance coupling of a strip with respect to all other electrodes. It should be noted that none of the above takes into account the effect of the oxide charge on the interstrip capacitance. The electron layer present at the surface of a $p^+ - n$ detector creates a floating electrode and it will not vanish even after the full depletion of the substrate. As will be explained in the next section, the electrons will create a charge pocket in between the strips which shrinks if the reverse voltage is risen: the exact distance from the p^+strip and the electron layer depends not only on reverse voltage but also on the electron density.

In order to take into account the effect of this floating electrode, the solution (3.75) has been suitably modified. In the ideal case of no interstrip charge, $C_{is} \simeq C_{geom}(\frac{w}{p})$, i.e. it can be simulated quite accurately by means of a pure geometrical calculation (see Eq. (3.71), Eq. (3.75), Eq. (3.77)) and all the parameterisations can be expressed as functions of the ratio w/p. If we include the floating electrode (see later, Fig. 3.44), the interstrip capacitance becomes $C_{is} = C_{geom} + C_{float}/2$. In order to define C_{float} it is useful to introduce the gap s from the junction to the floating electrode and by means of that, rewrite the expression of C_{float} as:

$$C_{float} = C_{geom}(\frac{w'}{p'}) \tag{3.80}$$

where

$$w' = \frac{p}{4} - \frac{s}{2} \qquad\qquad p' = \frac{p}{2}.$$

In this simple description the value of s is taken from detailed simulation of such a structure as a function of V_d and $N_i s$.

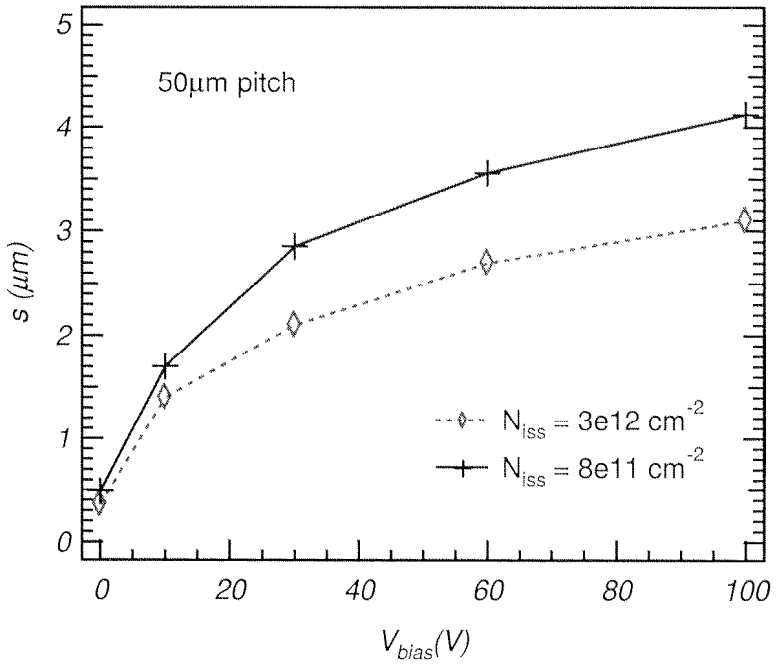


Fig. 3.42: Depletion voltage dependence of the strip-floating electrode gap (in μm). Results for two typical trapped charge densities are shown.

3.3.4.b Simulated capacitances

In the following the combined use of process and device simulations for the description of microstrip detector capacitances is described [66]. This kind of approach make it possible to establish a close correlation between fabrication process parameters and detector performance. The analysis has been carried out starting from the knowledge of process recipes and parameters, and accounting for the same photolithographic CAD mask layout used for the actual device production.

First, a comprehensive tuning of the simulator against experimental results has been carried out, in order to select (among several possible alternatives) the proper set of physical models and numerical algorithms to be used for the simulation. To this purpose, device testing has been performed at CSEM and at INFN laboratories in Pisa and Perugia, Italy. Then, the effect of changes in some process parameters (resulting in different values for the oxide charge and doping concentration, for instance) have been investigated.

Simulation outline All of the simulation activity described in the following has been carried out by using the Virtual Wafer Fab (VWF, [67]) developed by SILVACO©. Such a package provides process- and device-simulators, as well as database functions (allowing data exchange among simulators and/or experimental data) and pre- and post-processing graphic utilities. In particular, 2D process-simulation has been performed by the ATHENA program, predicting actual device cross-section, doping profiles, oxide-trapped charge distribution, etc.; a detail of the obtained device structure is illustrated by Fig. 3.43.

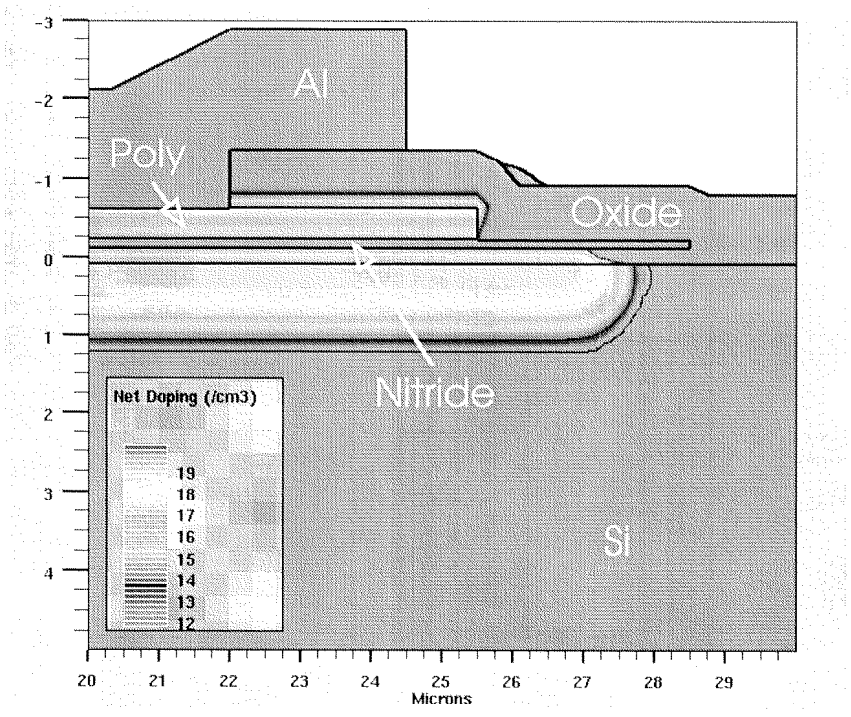


Fig. 3.43: Net doping distribution as obtained from process simulation: view of the implanted strip and of the coupling capacitor.

Such data are then forwarded to the device simulator ATLAS, which in turn works out the electrical behaviour of the device under investigation; "external" responses (I-V curves, terminal capacitances, for instance) are predicted, which can be compared with measures; moreover, "internal" quantity distributions (e.g., electric field, carrier concentrations) that, instead, can hardly be obtained through experimental techniques, are obtained as well.

Simulation results The simulated device has been characterized with respect to two fundamental parameters: depletion voltage and strip capacitance.

Depletion voltage V_{depl} can be extracted by looking at the strip capacitance to backplane (C_b) which, in turn, can be obtained from AC simulations. Depletion voltage can be inferred from the knee position in the customary $1/C_b^2$ vs. V_b plot (Figs. 3.46- 3.47). As expected, depletion voltage is quite sensible to the substrate doping concentration,

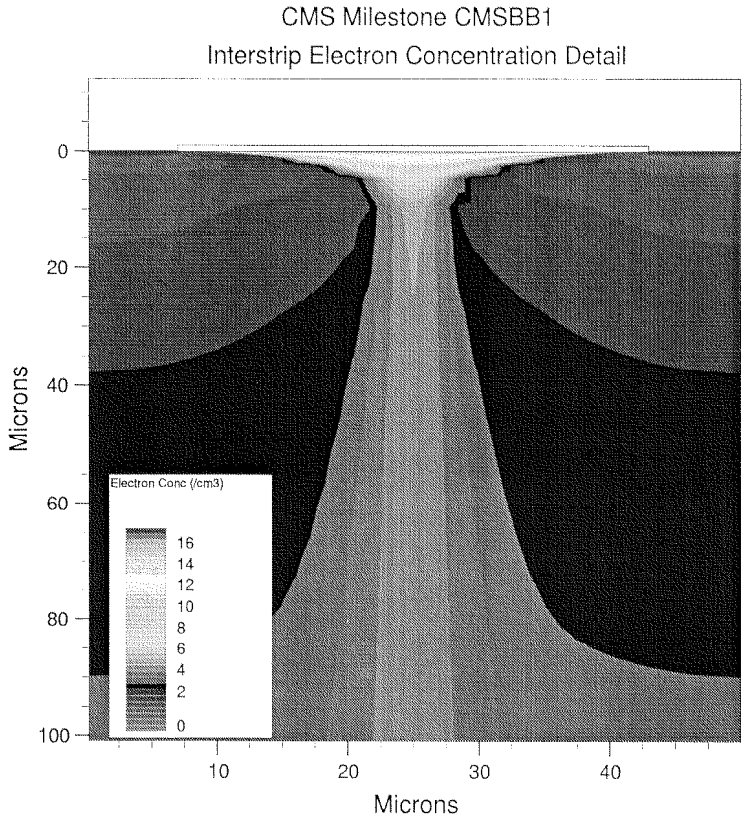


Fig. 3.44: Electrons concentration in the interstrip volume as computed by device simulation, at $V_b = 100$V. Log of density is shown.

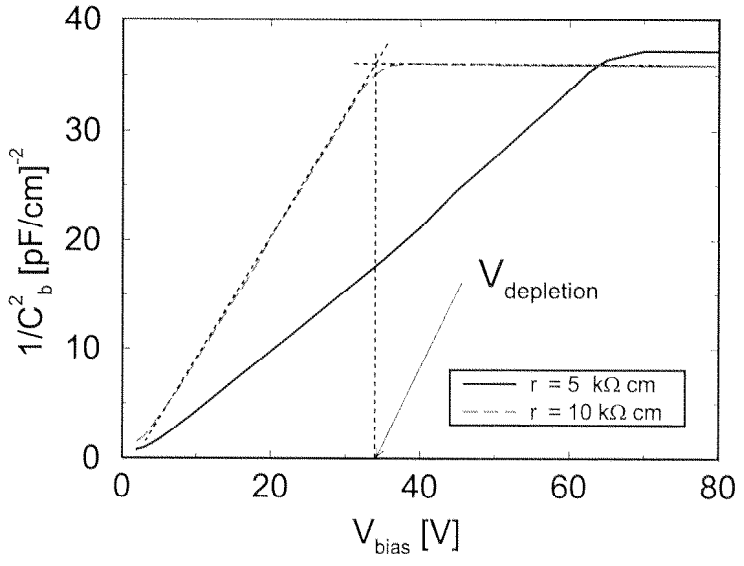


Fig. 3.45: $1/C_b^2$ plot for different substrate resistivity.

whereas it exhibits a much smaller sensitivity on the amount of oxide-trapped charge. It is worth mentioning that, at first order, radiation damage effects on Si microstrip detectors can be taken into account by proper adjustments of both bulk- and oxide-impurity concentration [68], so that the same kind of analysis can account for irradiated detectors as well.

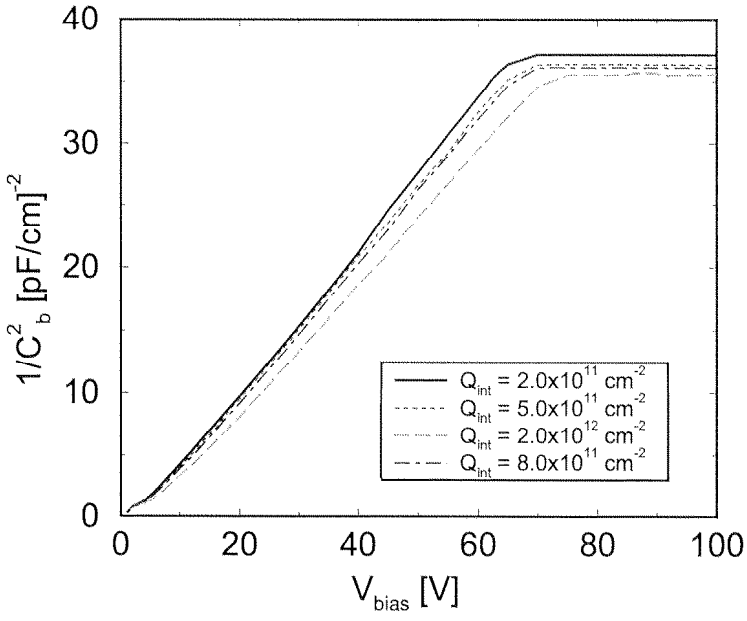


Fig. 3.46: Effect of the oxide charge on the capacitance to backplane.

As already mentioned, strip capacitances plays a dominant role in settling the noise performance of Si microstrip systems. Optimal strip geometry comes from a trade-off between detector resolution (which calls for minimum pitch, i.e., large capacitance values) and noise performance which, on the contrary, require to lower the capacitance value. AC simulations have been used to investigate the dependency of the strip capacitance upon process parameters. Although strip capacitance depends essentially on device geometry, significant non-geometrical contributions come also from the surface charge layers accumulated at the Si/SiO_2 interface (see Sec. 4.4.3). Simulation allows for analyzing the correlation between the amount of oxide-trapped charge and the electron layer build-up. In this particular case I used a $< 111 >$ silicon substrate, in order to match with what is commonly used in production. The possibility of using $< 100 >$ substrates for radiation hardened device is encouraged by Brews [69] since the amount of charge trapped at the surface is ~ 10 times lower (see Sec. 4.4.3). Its effect on the charge build-up will be ten times lower and therefore easily mastered by means of V_d augmentation. Some experimental evidence of this effect will be shown later in Chapter 5.

An almost logarithmic dependency on the interface charge-density can be inferred from the plot in Fig. 3.47.

As expected, the nearest-neighbors interstrip capacitance is much more sensitive to

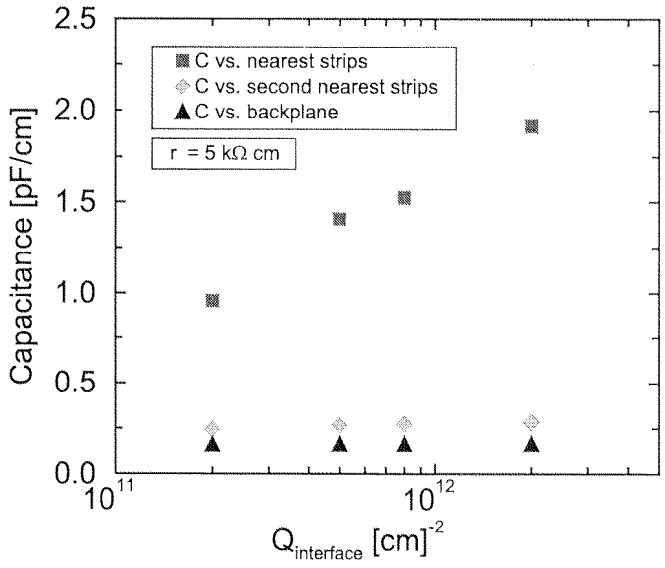


Fig. 3.47: Oxide charge effect on separate strip contributions.

the charge surface layer than other contributions. Dependency of the total strip capacitance C_{strip} ($C_{strip} = C_b + 2(C_{is,1} + C_{is,2})$, where $C_{is,n}$ denotes the capacitance to n-th neighbors) upon the applied strip bias is illustrated by Fig. 3.48 (contributions of 3rd and higher-order neighboring strips turn out to be negligible, since $\frac{C_{is,3}}{C_{is,1}} < 1\%$). Again, different values of the oxide-trapped charge are accounted for: the increase of the latter is shown to result in a capacitance increase and, thus, in a worse S/N ratio. To this respect, hence, simulation allows to predict the operational limits, in terms of accumulated oxide charge, within which the detectors can provide satisfactory performance for their whole lifetime.

In order to compare simulation results with actual silicon microstrip detectors, measurements have been performed on a full-scale AC-coupled detector coming from a CMS test production. Separate capacitance components have been individually measured, following the scheme of Fig. 3.39. Comparisons with data are given in Fig. 3.49. both from a quantitative

Conclusions An extensive set of numerical simulations has been carried out, to investigate several issues concerning design and optimization of integrated silicon microstrip detectors. Joint process- and device-simulation allowed for establishing close correlations between fabrication process parameters and device electrical response. Simulation results have been validated by comparing them with actual device measurements: good agreement has been found in all cases. This illustrates both practicality and reliability of the TCAD approach to the efficient design and optimization of silicon microstrip detectors: the application of conventional numerical simulation technique to wafer-scale IC design appears to be an attractive way to reduce prototyping time and expenses, as well as providing a mean for physical insight of detailed detector operation.

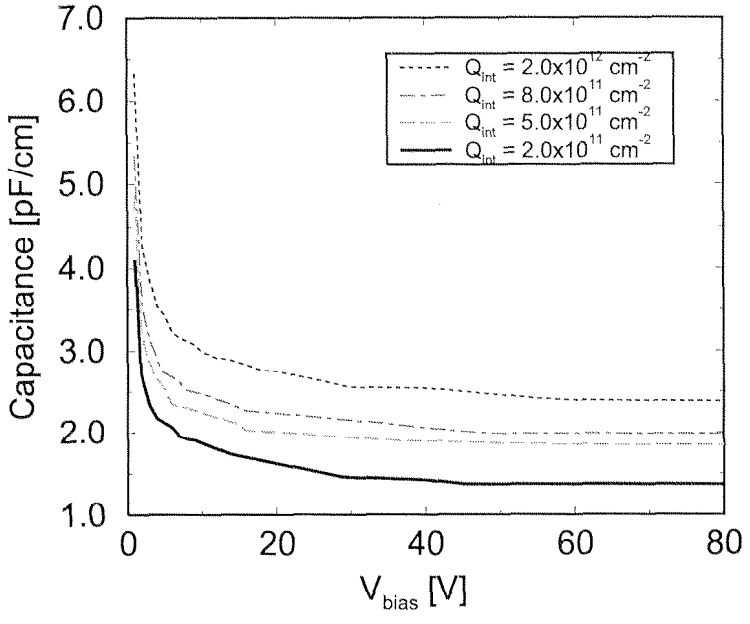


Fig. 3.48: Oxide charge effect on total capacitance.

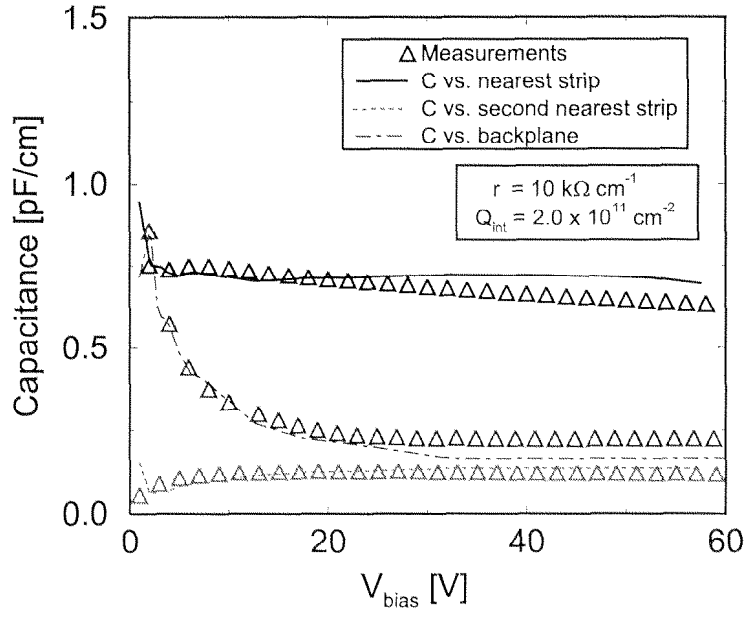


Fig. 3.49: Interstrip capacitance: measurement and simulation result comparison.

3.3.5 Radiation resistance

The survival of silicon detectors in the LHC radiation environment depends strongly on careful detector design and wise choice of the material. The design value for the radiation resistance of the system is set by the flux of neutral and charged particles received by the innermost barrel layer. Already mentioned in Chapter 2 a total fluence of $1.6 \cdot 10^{14}\text{cm}^{-2}$ neutrons is the prediction at this radius. Broadly speaking, the radiation damage suffered by the detectors can be divided into two classes; effects which are due to surface damage and those which are due to bulk damage. As will be shown in the following sections, the effects of bulk damage are considered to be the greatest source of concern and it is here that every effort must be made to guarantee reliable performance for the lifetime of the experiment.

3.3.5.a Surface damage

Surface damage occurs when the holes produced by ionizing radiation in the surface oxide layer either become trapped in the oxide or interact with atoms at the silicon-oxide interface to form interface states (see Sec. 4.4.3). Fixed positive charge in the oxide layer modifies the oxide field, while interface states give rise to new energy levels in the forbidden gap which can modify the device behaviour. These changes can lead to a decrease in inter-strip isolation, causing unwanted signal charge sharing, and an increase in inter-strip capacitance, which is a major factor in determining the electronic noise of the system. However, it has been demonstrated that a careful choice of fabrication technology and geometry can minimize the changes in device behaviour to an acceptable level. The capacitive coupling between each strip and its neighbors is dominated by the quality of the oxide at the interface (which depends on process and crystal orientation) and by the ratio of strip width to strip pitch. After irradiation, the value of the coupling can be increased by as much as a factor of ~ 3, when the trapped charge in the oxide saturates, as can be appreciated in Fig. 3.47 and Fig. 3.48.

It is possible to reduce the damage-induced coupling by substantially over-depleting the device. This produces high fields on the strip side and confines the oxide charge in the region between the strips, thus reducing the capacitance. Hence a technology which allows high voltage operation of irradiated devices will keep the impact of these changes on system performance under control. The use of AC coupling is not believed to present any particular problems with regard to radiation damage. Polysilicon bias resistors have been shown to be relatively insensitive to ionizing radiation (see Fig. 5.25) while the use of high quality dielectric layers for the coupling capacitors should guarantee stable operation for the duration of the experiment.

3.3.5.b Bulk damage

Leakage current The bulk leakage current in detectors is described in terms of the quantity J_v,which is the leakage current per unit volume. It is well established that the

increase in J_v, during irradiation is directly proportional to received fluence:

$$\Delta J_v = \alpha\Phi \tag{3.81}$$

where the proportionality constant, α, is known as the *damage constant*. The bulk current exhibits temperature-dependent annealing effects which must be taken into account. The usual quantity for measuring the leakage current increase is α_∞, the value of α after all annealing has ceased. At room temperature, α_∞ is achieved $\sim$ 20 days after the end of the irradiation. At the projected operating temperature (-10 °C), the current will fall to α_∞ within 3 years. The annealing effect will be considered in the following. The current itself is strongly dependent on temperature. The data are found to fit the form:

$$I \propto T^2 \exp\frac{-E_a}{kT} \tag{3.82}$$

The average value of E_a, from the various measurements taken is (0.62 ± 0.03) eV. The value of α_∞ for irradiation with 1 MeV neutrons at 20 °C is $(2.9 \pm 0.2) \cdot 10^{17}$ A cm^{-1}. At -10 °C, α_∞, falls to $(1.4 \pm 0.2) \cdot 10^{-1}$ A cm^{-1}. Hence, for the innermost micro-strip layer in the barrel region, a current density of $\sim 6\mu$A cm^{-2} is expected after 10 years of operation (see Fig. 3.50).

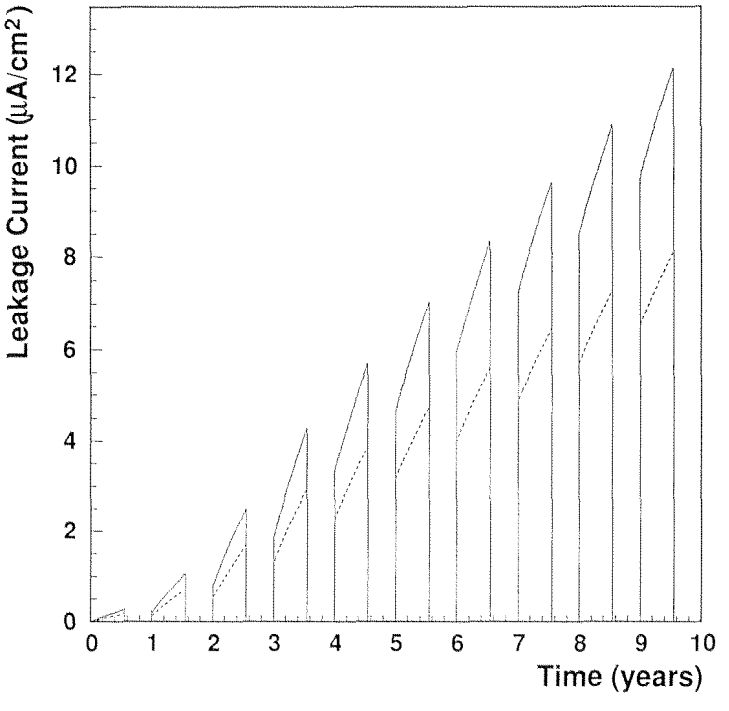


Fig. 3.50: I_{leak} versus time for the most irradiated detectors (dotted line). The simulation was done for a 4 kΩ·cm substrate. The solid line shows I_{leak} for a maximal fluence of $2.4 \cdot 10^{14}$n cm^{-2}.

The AC coupling between strips and amplifiers allows to avoid baseline shifts in the amplifier. The fast shaping time of the front-end amplifiers (50 nsec or less, with deconvolution) minimizes the noise contribution of the leakage current. It should be noted

that, with heavily irradiated silicon detectors, one must take into account the possibility of thermal runaway.

Charge collection efficiency The signal to noise ratio is also affected by decrease in charge collection efficiency, which are caused by the trapping of charge carriers at radiation-induced defects in the silicon bulk. It has been shown by several groups that the resulting signal loss is moderate and has a value of $< 10\%$ after a fluence of $1 \cdot 10^{14}$ n cm^{-1}. Recent results demonstrate that $\sim 40\%$ of the signal remains after $1 \cdot 10^{15}$ n cm^{-2}. On the basis of the references quoted there, it is estimated that the ballistic deficit suffered by the innermost micro-strip layers - for an effective shaping time of 25 ns - will not exceed 15%, which is considered tolerable.

Effective doping concentration The depletion voltage of a silicon detector depends upon the effective doping concentration of the substrate material:

$$V_{depl} = \frac{qW^2}{2\epsilon_{Si}} N_{eff} \tag{3.83}$$

where d is the depth of the diode. The value of N_{eff} is determined by the concentration of space charge in the depletion region. In an unirradiated device, the space charge arises predominantly from the phosphorus dopant.

Irradiation results in an accumulation of negative space charge in the depletion region due to the introduction of acceptor defects which have energy levels deep within the forbidden gap. $n-$type detectors therefore become progressively less $n-$type with increasing hadron fluence until they invert to effectively $p-$type and then continue to become more $p-$type beyond this point, apparently without limit (see Fig. 3.51).

The inversion fluence, depends strongly on the initial resistivity of the substrate material (see Fig. 3.51) and has been parameterized for 24 GeV protons as:

$$\phi_{inv} \approx (18.0 \pm 0.6) N_{eff} \tag{3.84}$$

Detectors still work beyond the inversion fluence as the junction moves from the p^+ strips to the n^+ back-plane contact. At high fluences, however, N_{eff} can be such that the depletion voltage exceeds the breakdown voltage of the device and efficient operation is no longer possible. In the period after irradiation, the annealing behaviour of N_{eff} displays two distinct phases. There is an initial reduction in negative space charge (beneficial annealing), which is later dominated by a slower, but much larger, increase in acceptor concentration. This second phase is known as "reverse annealing". The rate at which reverse annealing proceeds is highly temperature dependent; the changes, up to saturation, can take many years at -10 °C but are accelerated to a matter of weeks at room temperature. Consequently, reverse annealing imposes strict limits on the operation temperature of the SST and requires that maintenance and warm-up periods be kept to a minimum. It follows that the changes in N_eff can be written as the sum of three components:

$$\delta N_{eff} = N_{eff0} + N_{eff} = N_a + N_c + N_Y \tag{3.85}$$

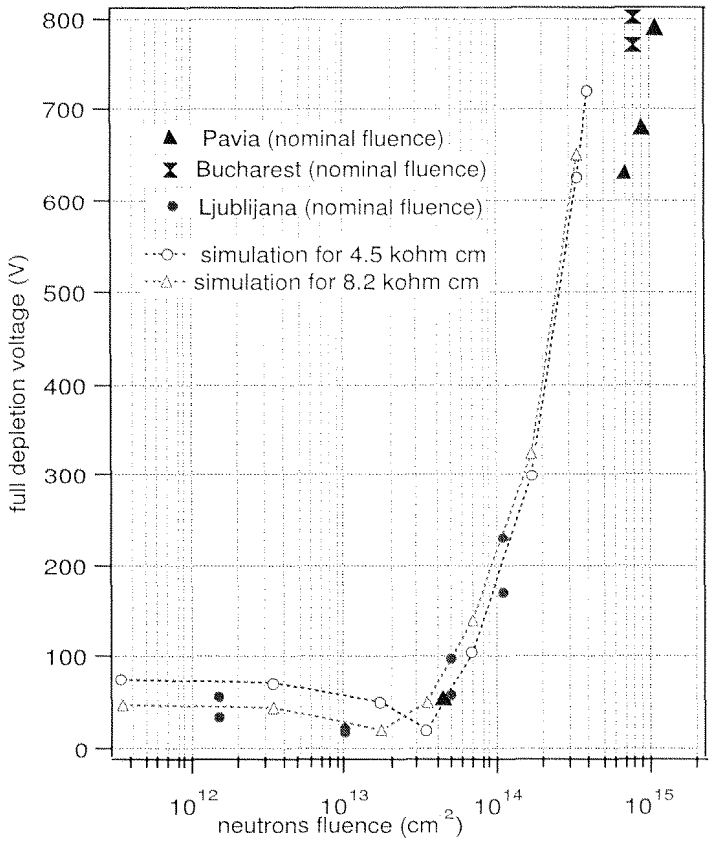


Fig. 3.51: Evolution of the depletion voltage with fluence. Data refers to edge structures described in Sec. 3.3.3.c.

where

$$N_a = g_a \Phi \exp\left(-k_a t\right) \tag{3.86}$$

corresponds to the beneficial annealing,

$$N_c = N_{c0}\left(1 - \exp\left(-c\Phi\right)\right) + g_c \Phi \tag{3.87}$$

to the stable damage and

$$N_Y = g_Y \Phi \left(1 - \left(1 + k_{Y1} t\right)^{-1}\right) \tag{3.88}$$

to the reverse annealing. The most recent and accurate parameterisations of these three terms has been carried out by Feick et al. [70]. The parameter values are shown in Table 3.6.

The predicted evolution of V_d, with time is shown in Fig. 3.52.

A conservative scenario comprising running operation at -10 °C and stand-by at 15 °C apart for periods of 21 days at 10 °C and 7 days at 20 °C per year for repairs and maintenance has been assumed. Results have been calculated for two different values of initial resistivity (4 kΩ·cm and 1 kΩ·cm). A worst case scenario using our safety factor of

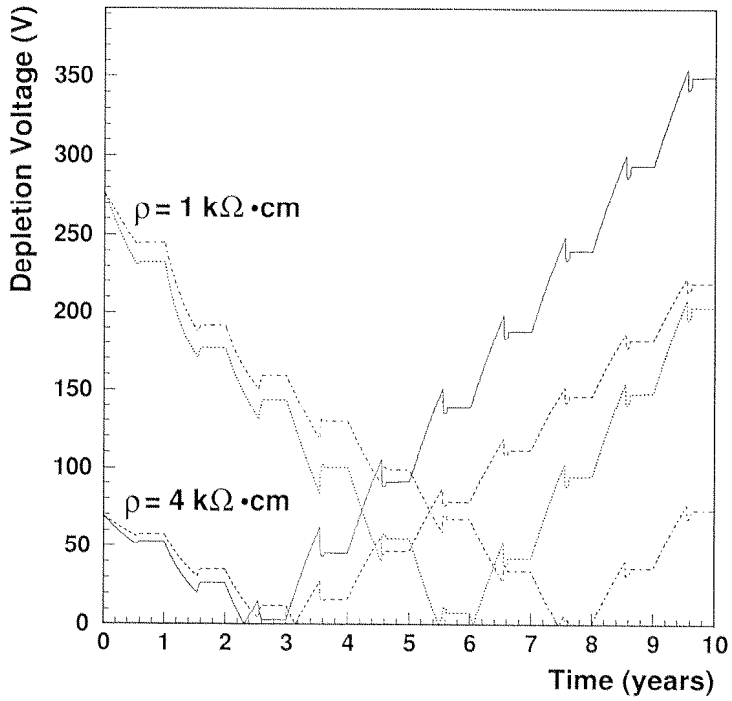


Fig. 3.52: Predicted evolution of V_{dep} with time for barrel layer 1 for two different initial resistivities. The worse scenario for a total fluence of $2.4 \cdot 10^{14}$n cm^{-2} is also shown for each initial resistivity.

1.5 is also shown. It is clear that efficient operation can be guaranteed for the duration of the experiment if the breakdown voltage is high enough. From these calculations and from signal-to-noise considerations which are discussed in section 3.3.1, we require that the detector can be safely operated up to 500 V.

Fig. 3.52 and Fig. 3.53 show the leakage current and the corresponding dissipated power. The worst case scenario including the safety factor of 1.5 is shown as well. The same maintenance scenario as for Fig. 3.52 has been assumed. The power dissipation has been calculated with an operating voltage which is $1.5 \cdot V_d$. From these figures we conclude that safe operation conditions can be obtained.

Table 3.6: Measured values of the relevant parameters for bulk damage.

Parameter	value or expression
N_{c0}	$(0.3 \times N_{eff0}) + 10^{11}$ cm^{-3}
N_{eff0}	1×10^{12} cm^{-3}($\rho = 4$ $k\Omega$cm); 4×10^{12} cm^{-3}($\rho = 1$ $k\Omega$cm)
c	2×10^{-13} cm^2
g_c	1.77×10^{-2} cm^{-1}
g_Y	4.6×10^{-2} cm^{-1}
g_a	1.54×10^{-2} cm^{-1}
k_a	$k_{a0}\ exp(-E_{aa}/k\ T)$
k_{a0}	2.3×10^{13} s^{-1}
E_{aa}	1.08 eV
k_{Y1}	$k_{Y10}\ exp(-E_{aY}/k\ T)$
k_{Y10}	8.74×10^{14} s^{-1}
E_{aY}	1.31 eV

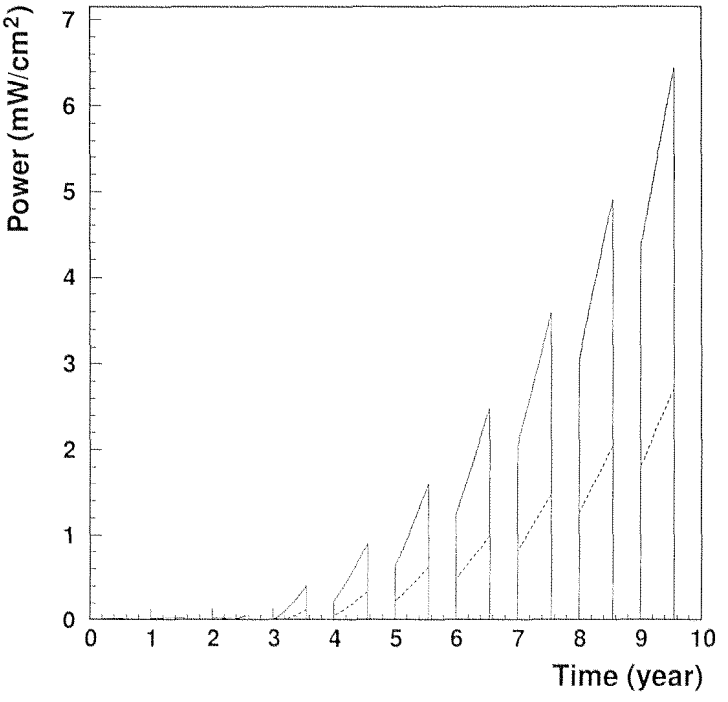


Fig. 3.53: Power density dissipated on a silicon module (dotted line). The simulation was done for a 4 kΩ·cm substrate. The solid line shows the power dissipated for a maximal fluence of $2.4 \cdot 10^{14}$ n cm^{-2}.

3.3.6 The Single or Double Side Choice

Conventionally the most common type of microstrip silicon detectors is single-sided (SS) $p^+ - n$. The reason for that is straightforward: the fabrication process used in the fabrication of such devices is a direct derivation of the planar process used in the IC industry. By means of that, the front side only of the silicon wafer is patterned, being the backside damaged by the automatic handling.

In developing the technology suited for silicon detectors, particular care have been devoted to avoid serious damages on the back of the wafer: mechanical scratches on the back surface may weaken the high-low junction suited to prevent charge injection in the bulk from the back-side contact.

I participated in the past to the fabrication of a fully double-sided (DS) silicon detector, meaning active patterning on both sides, i.e. both detector faces able to collect charges from ionizing particles [71, 72]. This work led to the production of the detectors for the first fully double-sided silicon vertex detector, the ALEPH VDET [73].

Radiation resistant double-sided devices have been produced and tested within the CMS R&D activities [74].

Double-sided detectors may provide two advantages over (for example) a pair of SS detectors:

- Easiness in solving the ambiguity problem in multi-hit events.
- Less material.

In principle is possible to measure the signal charges in the $p-$ and the $n-$side clusters and use the correlation between them to rule out fake associations. In practice, operative conditions of silicon detectors greatly reduce this possibility (i.e. low S/N ratio on wide clusters on one or either sides).

The effectiveness of DS detectors in reducing the multiple scattering by means of support structures, cables and cooling, should be carefully analyzed since multiple scattering is proportional to the squared root of the thickness of the material and by using two SS detectors mounted back-to-back, many items are shared, thus the added material do not necessarily doubles. Moreover DS devices are intrinsically more expensive because of the special handling explained above and of their intrinsic lower production yield.

For the double-sided layers CMS plan to use two single-sided back-to-back sensors, one of which will have small angle stereo strips. This allow an easy routing of the signal to the electronics located at one end of the module without making use of flexible circuites (copper on kapton or upilex thin foils [72]). Even in the case of small angles, in order to avoid dead regions, the use of double metal connections is envisaged, since the number of second metal lines crossing the detector is very limited, thus the S/N ratio is not significantly degraded by the increased strip capacitance. Another reason to prefer back-to-back sensors is driven by the unavoidable high voltage operation: the electronics can be operated without an offset ground which is necessary when truly double-sided devices are used.

The n^+ on n-bulk devices, although interesting, do not seem to offer definite advantages in the SST system. They are more expensive than truly single-sided devices since they need a pattern on the front side, making them "quasi" double-sided. Finally, optimization of the performance of these devices in the 4T field of CMS would be highly problematic due to the higher Lorentz angle of the electrons.

3.3.7 Conclusions

In the previous sections, the main sources of noise for a SSD have been discussed.

Fig. 3.54 reports the expected performance of non-irradiated detectors as a function of the width/pitch ratio.

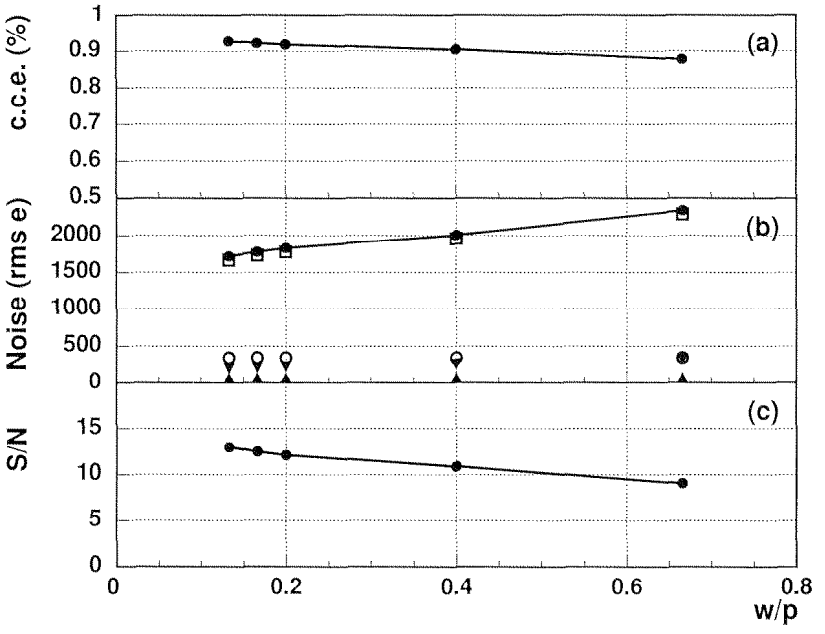


Fig. 3.54: Predicted charge collection efficiency (c.c.e.) (a), noise (ENC) b) and S/N ratio (c) versus w/p for CSEM detectors. In the noise plot, the solid line represents the sum of all contributions, whereas the unconnected points represent, from top to bottom, noise from: the amplifier, the strip resistance, the bias resistance and the reverse bias current. From [37]

An essential feature of Fig. 3.54b is that noise is totally dominated by the front-end electronics.

Leakage current will not affect substantially the detector performances until the final stages of high luminosity operation, when strongly irradiated detectors will show total leakage currents in the mA range. An acceptable level of performance will be reached only by means of depletion voltages significantly higher than the depletion one (see Sec. 5.3.1).

Strip bias resistors in the $M\Omega$ range are safe enough to keep parallel Johnson noise contribution very low since even after strong irradiation their values do not lower significantly.

The compact geometry chosen for the modules (1 or 2 detectors/module) will keep the series resistance contribution under control but the capacitive load must be kept as low as possible, which will also ensure the highest charge collection efficiency (Fig. 3.54a). Low values of the w/p ratio are therefore preferred. The predicted signal-to-noise ratio would be above 12 for $w/p \leq 0.2$.

From Sec. 3.3.1, it is clear that some items may be solved by means of a careful design only, others combining a good design and a fine technological tuning and the rest by process choices alone. Pure technological choices will be described in the next chapter.

It is possible to draft a guideline for the best CMS silicon strip detector design, by summing up the most important items not exclusively process related:

- The thickness of the silicon substrate should be chosen as a compromise between signal-to-noise and process easiness. Thin substrates translates in low signal charge, exacerbated by the reduction in the signal voltage due to an increased capacitance from strip to backside, and thus poor signal to noise performances but are less sensitive to multiple scattering problems[11].

- Carefully controlled dose and doping profile will limit the occurrence of critical fields at the junction edge.

- Good ohmic contact is needed to avoid leakage current increase by charge injection from the back contact. The effectiveness of the high-low junction on the backside is proven.

- Edge stabilization by means of an n-well implantation on top of the cutting area on the junction side will avoid charge injection from the area damaged by the cutting procedure. The presence of such a wide and smooth n^+-implant relaxes the high fields close to the cut thus improving the breakdown limit of the detector.

- Multi-guard structures between the detector active area and the n^+ edge implant are important to distribute the voltage drop on the junction side over a larger region, enhancing the breakdown performances of the device.

- The thickness of the decoupling oxide of the strip which fixes the decoupling capacitor value should be chosen by looking to the aimed noise performances and the design width of both implantation and readout metal.

- The read-out pitch being defined by the requirements on spatial accuracy and two-track resolution, the strip width will be determined by a compromise between wider implants and safety in terms of field gradients close to the implants. Narrower implants which minimize the coupling to the adjacent implants will reduce the system noise too. The inter-strip capacitance should then be such that it greatly exceeds the strip-to-backside capacitance in order to avoid serious signal loss from floating strips.

[11] It should be noted that multiple scattering depends from the squared root of the substrate thickness.

- The value of the poly-silicon resistor needed to bias the implants is a compromise guaranteeing good uniformity and a negligible voltage drop, minimizing the contribution to the thermal noise.
- The thickness and the lateral dimensions of all metal electrodes are chosen as high as possible to reduce the total series resistance of the read-out strips.
- The intermetal dielectric for the detectors with double metal signal re-routing (for back-to-back modules) should be as high as possible in order to lower the stray capacitance between the strip of the two metal layers which contributes to the series Johnson noise of the detector (or module).

Chapter 4

Fabrication

This chapter is meant as an introduction to the fabrication procedure of silicon microstrip detectors. The basic concepts come from the planar process, widely used in IC industry for circuit fabrication. But some important differences emerge, some of them being responsible for the specific complexities of detector fabrication. The individual processes required to produce silicon detectors and the manner in which they are combined to make the detector will be briefly described. In particular I will summarise the major steps we have chosen at CSEM for the fabrication of single sided, AC-coupled silicon detectors, the latter being the final choice for the SST of the CMS experiment.

4.1 Introduction

The complexity and scale of many new applications for charged particle detectors require well-controlled fabrication techniques. Hence, a modern detector process should offer minimum sensitivity to process-induced contaminants and ideally would be compatible with contemporary integrated circuit (IC) fabrication techniques. However, typical fabrication processes for radiation detectors do not fulfill either of these goals. While in one notable case outstanding results have been achieved [75], the process used is incompatible with conventional IC fabrication and is very sensitive to the temperature at which the implanted dopants are annealed. This is due to the extreme sensitivity of detector leakage current to impurities that introduce energy levels near the centre of the silicon bandgap.

At CSEM, a combination of fairly conventional integrated circuits processing techniques are used to realise low-leakage, radiation-hard radiation detectors. By utilising gettering techniques that actively counteract the presence of harmful impurities in the active volume of the device (see Sec. 4.11), the sensitivity of the process to variations in the fabrication environment is greatly reduced. In addition, the use of relatively high processing temperatures makes the process described here fully compatible with conventional IC processing.

Silicon detector as well as IC wafer fabrication starts with the careful cleaning of a bare silicon slice of appropriate crystallographic orientation and the growing of a layer

of thermal oxide[1] on each surface. As fabrication proceeds, other layers covering only one side of the wafer such as CVD oxide, poly-crystalline silicon, silicon nitride, and various metals will be added. Intermixed with these steps will be patterning (lithography), etching, diffusion, and ion implantation.

Each of the wafer fabrication operations listed below have many steps and usually many variations, depending on exactly the kind of structure to be built. A wafer may circulate back through some of them several times, defining the order in which the steps are performed, the so called "process flow", is a key task of the detector technology designer.

4.2 Substrate choice

Basic characteristics of the silicon substrate are of capital importance in detector fabrication. Crystal orientation, resistivity, uniformity in doping and thickness differentiate detectors from mainstream CMOS or bipolar IC technologies.

As Table 4.1 shows, the orientation affects nearly all of the wafer-fab processing steps.

Operation	Effect of Orientation
Thermal oxidation	Rate varies with orientation; $\langle 100 \rangle < \langle 110 \rangle < \langle 111 \rangle$.
Etching	With some etchants rate is very low in [111] direction.
Diffusion	Diffusion is dependent in silicon when done in conjunction with thermal oxidation; $\langle 100 \rangle > \langle 111 \rangle$.
Ion implantation	Range is greatest in [110] direction; also is greater directly in other low-indices directions as compared with slightly off-orientation directions.

Table 4.1: Effect of crystal orientation on the main processing steps.

The material specifications for silicon detector fabrication are quite different from what is needed in ordinary IC technology[76]. In Table 4.2 are summarized the main substrate requirements.

Resistivity characteristics are very different both in absolute value and in uniformity. This fact translates in using exclusively Float-Zone (FZ) silicon, which ensure the highest purity, for detector fabrication. But FZ is not the standard for IC industry which prefers Teal-Little or CZ grown silicon, and this causes difficult supplies, long delivery times (several months) and higher prices.

Wafer thickness is another crucial difference: standard ICs are mainly based on surface phenomena. The silicon bulk thickness is therefore not of paramount importance from the electrical point of view. On the other hand, the necessity of fast automatic handling requires good mechanical strength, which translates in the thickness specifications.

[1]Thermal oxide is SiO_2 formed by oxidising the silicon surface.

Item	IC	Detector	Item	IC	Detector
Resistivity (Ω·cm)			Slice thickness (μm)		
Min	0.008	1000	Min	375 ± 25	200 ± 10
Max	20	> 8000	Max	710 ± 25	300 ± 15
Resistivity gradient	10%	5%	Diameter tolerance	$\pm$0.5mm	$\pm$0.5mm
Dislocation density/cm^2	< 100	< 10	Bow (μm)	35	15
Carbon Conc.	< 2	< 2	Oxygen Conc.	25–40 ppm	> 100 ppm

Table 4.2: Comparison between SEMI - a standardization organization in the semiconductor industry- standards for 4" $n-$type Si IC wafers and detector-grade wafers.

In the case of silicon detectors, the substrate is the active part of the device and thickness defines its energy response. Moreover, in high radiation environments like LHC, $n-$type silicon is subject to type inversion (see Sec. 3.3.5.b) and the depletion voltage may rise to unacceptable values if the thickness is too large: CMS chose $300\mu m$ thick wafers as a good compromise. Nevertheless innermost layer of pixel detectors may ask for thicknesses down to 200μm. Difficulties in processing such thin wafers may then arise:

- due to their lighter weight, they are difficult to handle with automatic equipment, tuned for standard IC wafers: this fact causes wafer breakage, lowering dramatically the so-called mechanical yield;
- thermal oxidation and dielectric layers grown or deposited during process do induce tensile and compressive stresses to the lattice, causing bowing or even breakages;
- wafer bowing may be too large to allow, i.e. bump bonding or support gluing with full size detectors;
- badly conditioned wafer edges may induce edge chipping: silicon particulates may fall back onto wafer surface during fabrication, affecting device integrity;
- procurement time for thin wafers is very long and may considerably slow down the duty cycle of the fabrication process.

Historically the crystal orientation used for detectors is $\langle 111 \rangle$. I tried to understand the motivation for this choice by looking into early silicon detector literature. Kemmer [75], which is considered as the founder of the modern way of processing silicon detectors, or Holland [77], who extended the Kemmer process to higher thermal budgets, do not give any explanation. My personal believe is that this choice which amazingly lasted for more than 15 years has been driven by one, or a combination, of the following facts: $\langle 111 \rangle$ wafers were at that time the only available in FZ at high resistivity, $\langle 111 \rangle$ wafers are

less prone to aluminum pitting (see Sec 4.9.1) and therefore junctions are safer, ⟨111⟩ wafers suffer less implant channelling (see Sec. 4.6) and therefore the control on the implanted doping species is better. Recently one important drawback in using ⟨111⟩ slices for strip detectors suited for operation in high radiation environments has been highlighted: the higher density of surface charge (see Sec. 4.4.3) with respect to ⟨100⟩ causes higher interstrip capacitances, especially after irradiation (see Sec. 5.2.4), rising the noise contribution of the detector [78].

Resistivity lack of uniformity on the wafers are not welcome: local variations in depletion voltage as well as electric field may cause weird detector behaviours, especially if operated just above full depletion. A good rule of thumb is to take a safety factor of $30\% - 50\%$ on the depletion voltage, when defining the operative voltage.

Last important difference is wafer's oxygen concentration. Since some months, the RD48 collaboration (ROSE) at CERN has performed in-depth studies of the effect of oxygen in detector grade silicon and its effect on detector performances after type inversion [79]. It comes out that oxygen may definitely help in lowering the effect of reverse annealing (Sec. 3.3.5.b and Sec. 4.11.2): in Fig. 4.1 are summarized the ROSE tests on diodes made with different substrates, highlighting the excellent behaviour of oxygenated materials. CSEM is currently processing a CMS silicon strip detector batch making use

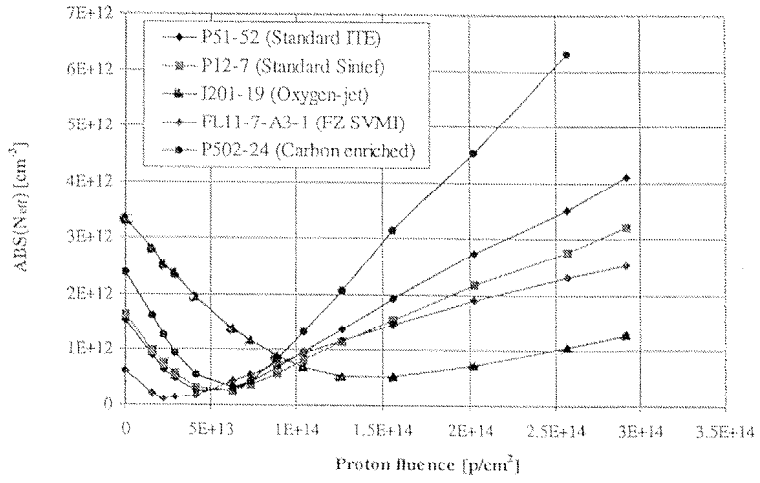


Fig. 4.1: Effective doping change as a function of 24 GeV/c proton fluence for various substrates. [79]

of a preliminary deep oxygen diffusion.

Substrate procurement for silicon detector manufacturing is a serious problem. Nowadays silicon manufacturers are concentrated in producing $6 - 8''$ silicon, CZ-grown, typically p-type, $700 - 800$ μm thick and with a resistivity close to 1-2 Ω·cm. Therefore the production of substrates suited for silicon detectors is not in their mainstream commercial interest: growing hyperpure silicon is more difficult, more expensive, the diameter (on $4''$ there will be the larger volume) has been almost abandoned by IC industry, lapping and polishing silicon to 300 μm thick wafers may case yield losses. Moreover, production volumes required by the HEP community are not large enough to convince the major silicon

producers to change their attitude and to rise the detector-grade silicon from a small "niche" market to something "out-of-the-catalog": the entire silicon surface required for both silicon trackers of ATLAS and CMS can be produced in less than one week!

4.3 Surface Cleaning

Optimal cleaning of semiconductor surfaces requires that particulates, organic films, and adsorbed metal ions are removed. Most cleaning procedures are based on immersion in liquid baths or liquid sprays. In addition, ultrasonic agitation or brush scrubbing may be required. In some cases, high-temperature vapor etching or low pressure sputter etching may be used. A good cleanup is complicated by the fact that unless great care is taken, the materials used for cleaning may contain (and leave behind) more particulates and metals than were on the surface in the beginning. Consequently, semiconductor-grade chemicals that are specially filtered and purified, must be used. Organic solvents are widely used, but since they sometimes leave residues themselves, high-purity water is ordinarily the last stage of a cleanup. This last stage is often immediately preceded by an acid to oxidise and remove any remaining organics.

At CSEM, the care in cleaning procedure during silicon detectors fabrication is taken very seriously since the requirements on defect density are extremely tough: one particulate or a small resist residue may translate in a pinhole in the decoupling oxide or in a implantation short between two strips. We experienced in the past that by simply increasing the number of cleaning stages before and after critical processing steps, we increased the devices final yield considerably: a conservative approach [80] (when shortening of process time is secondary) should be preferred.

4.4 Dielectrics: thermal oxidation and CVD deposition

4.4.1 Thermal Oxidation of Silicon

The operation of oxidising the silicon surface is performed in a temperature range of 800°C- 1250 °C and in an atmosphere containing oxygen (dry) or steam (wet) in an inert carrier gas.

The oxidation rate increases with increasing temperature and pressure and the oxide thickness increases, although not linearly, with oxidation time. Silicon is consumed during the oxidation process at a rate such that for every micron of oxide grown, the silicon surface recedes by 0.45 μm. Oxidation rates depends also on crystal orientation: since oxidation rate depends to the rate of incorporation of silicon atoms into the silicon dioxide network and the surface density of silicon atoms is higher on the $\langle 111 \rangle$-plane than on $\langle 100 \rangle$-plane, oxides will grow faster on $\langle 111 \rangle$ substrates.

To avoid growth of stacking faults during heat treatment, a small percentage of HCl

or TCE (trichloroethylene) is added to the O_2 gas. Chlorine has the ability both to getter heavy metals from the silicon bulk and to form an efficient sink for silicon interstitials. As a consequence, generation and recombination centers in the Silicon bulk are reduced, enabling high minority carrier lifetimes. Additionally chlorine is useful for neutralization of alkaline ions, which are responsible for device instabilities.

4.4.2 Chemical Vapor Deposition

CVD is used to deposit layers of materials such as silicon oxide, silicon nitride and polycrystalline silicon onto the wafers. CVD processes usually operate in a temperature range of 300 °C - 900 °C and proceed by reactions such as follows:

$$SiH_4 + 2O_2 \xrightarrow{\sim 450\ ^\circ C} SiO_2 + 2H_2O \qquad \text{(silicon oxide)}$$

$$SiH_4 \xrightarrow{\sim 600\ ^\circ C} Si + 2H_2 \qquad \text{(polysilicon)}$$

$$3SiCl_2H_2 + 4NH_3 \xrightarrow{\sim 800\ ^\circ C} Si_3N_4 + 6HCl + 6H_2 \qquad \text{(silicon nitride).}$$

Oxides deposited in this manner have a lower silicon interface quality than a thermal oxide and, in addition, may have a lower density. Consequently, they are not an alternative to thermal oxides but rather an adjunct to be used when thermal oxidation temperatures cannot be tolerated, like after a metal deposition a passivating oxide, or whether fast depositions of thick non-critical oxide layers, like intermediate, sacrificial oxides are needed. Furthermore, they can be used when an insulating layer is required and there is no silicon to be oxidised.

It should be noted that thickness control on deposited dielectrics is worse than on thermal oxides. By using a plasma to supply a portion of the energy required, CVD reactions such as those just listed can be done at substantially lower temperatures (PECVD - Plasma Enhanced CVD - silicon dioxide, also known as LTO, low temperature oxide). In particular, usable silicon nitride films can be deposited at 300 °C, whereas straight CVD requires over 800 °C. Basically LPCVD (Low Pressure CVD) depositions are made on furnaces tubes and the deposition is therefore double-sided; on the contrary, by using PECVD, wafers are laying on an electrode plate within the plasma reactor and the deposition is single-sided.

Polysilicon consist of small grains of single crystal material separated by highly disordered regions called grain boundaries. Those grain boundaries act as impurities sinks and hence polysilicon itself is an efficient gettering material (see Sec. 4.11.1). Polysilicon is extensively used both as electrode and as resistor, by exploiting the possibility of doping it either by diffusion from a gaseous or solid source, or by means of dopant implantation. It is interesting to note that the polysilicon resistivity strongly depends on deposition conditions which fix i.e. the grain size: once doped, conduction occurs through grain boundaries where trap density is very large and its variations defines resistance. In the MAXIS technology, developed at CSEM for rad-hard silicon detectors, polysilicon is used only for the strip bias resistors and its resistivity is tuned by means of dopant implantation.

4.4.3 Oxide charges

The presence of charges in the oxide and at the silicon-silicon dioxide interface is unavoidable in practical systems and may be the cause of strong instabilities in surface operation of MOS based device, like strip detectors. In fact, as already pointed out, interstrip resistance, interstrip capacitance, oxide breakdown voltage and also charge collection efficiency (through the detector depletion voltage) are affected by oxide charges. It is therefore worth to give some insights of the different kind of interface charges as well as to analyze the possibilities a silicon detector manufacturer has to keep those charges under control. The basic classifications of these traps and charges are shown in Fig. 4.2: they are interface-trapped charge, fixed-oxide charge, oxide-trapped charge and mobile ionic charge.

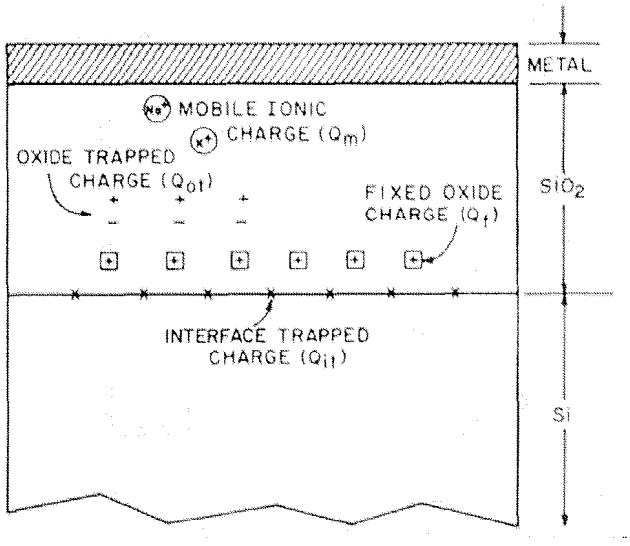


Fig. 4.2: Scheme of mobile and fixed charges at the interface between silicon and silicon dioxide [81].

Interface-trapped charges Interface-trapped charges Q_{it} are due to Si-SiO_2 interface properties and dependent on the chemical composition of this interface. The traps are located at the Si-SiO_2 interface with energy states in the silicon forbidden bandgap. The interface trap density N_{it} (i.e., number of interface traps per unit area) is orientation dependent. In $\langle 100 \rangle$ orientation, the interface trap density is about an order of magnitude smaller than that in$\langle 111 \rangle$. Present-day MOS diodes with thermally grown silicon dioxide on silicon have most of the interface trapped charge neutralized by low temperature (450 °C) hydrogen annealing. The value of Q_{it}, for $\langle 100 \rangle$-oriented silicon can be as low as $1 \cdot 10^{11}\text{cm}^{-2}$, which amounts to about one interface-trapped charge per 10^5 surface atoms. For $\langle 111 \rangle$-oriented silicon, Q_{it} is about $8 \cdot 10^{11}\text{cm}^{-2}$.

Fixed-oxide charge The fixed-oxide charge Q_f is located within approximately 30Å of the Si-SiO_2, interface. This charge is fixed and cannot be charged or discharged over a wide variation of the surface potential. Generally, Q_f is positive and depends on oxidation and annealing conditions and on silicon orientation. It has been suggested that when the oxidation is stopped, some ionic silicon is left near the interface. These ions, along with unfilled silicon bonds (e.g. Si-Si or Si-O bonds) at the surface, may result in the positive fixed-oxide charge Q_f. Q_f can be regarded as a charge sheet located at the Si-SiO_2 interface. Typical fixed-oxide charge densities for carefully treated Si-SiO_2, systems are about 10^{11}cm^{-2} for a $\langle 100 \rangle$ surface and about $5{\cdot}10^{11}\text{cm}^{-2}$ for a $\langle 111 \rangle$ surface. Because of the lower values of Q_{it}, and Q_f, the $\langle 100 \rangle$ orientation is preferred for commercial CMOS devices, where threshold variations in MOSFET devices may cause severe instabilities.

Oxide-trapped charge The oxide-trapped charges Q_{ot}, are associated with defects in silicon dioxide. These charges can be created, for example, by photon radiation or high-energy charged light particle (e.g. pions or electrons) bombardment. The traps are distributed inside the oxide layer. Most of the process-related Q_{ot}, can be removed by low-temperature annealing.

Mobile ionic charge The mobile ionic charges Q_m, such as sodium or other alkali ions, are mobile within the oxide under high-temperature and high-voltage operations. Trace contamination by alkali metal ions may cause reliability problems in semiconductor devices operated under high bias-temperature conditions. Under high bias-temperature conditions mobile ionic charges move back and forth through the oxide layer, depending on biasing conditions, and thus give rise to shifts of the C-V curve along the voltage axis. Special attention must therefore be paid to the elimination of mobile ions in device fabrication.

Irradiation effects Charges can also be introduced into MOS-like devices (i.e. strip detectors) by irradiation. Light particles, like pions or electrons, or photons hit the device during fabrication or during operation. Both charge in the oxide Q_{ot} and in interface trapping states D_{it} can be altered by irradiation. Such particles may create electron-holes pairs in the oxide exactly in the same manner they would do in silicon bulk (except for the minimum energy required, which is 8-9 eV in SiO_2 - its bandgap). Since oxide is generally much thinner than Si bulk and contains few free carriers, recombination rate is small. Instead, most of electrons are swept out of the oxide by any field that may be across it. However, here are many holes traps within the oxide and consequently holes may be immobilised, increasing the positive charge in the oxide. This fact affects the flat-band characteristic of MOS capacitors, as explained later in Sec. 5.1.2.b. Fig. 4.3 shows the energy-band diagram of a typical MOS system (e.g. the field plate region in our GR's): it is clear that an electron, easily excited from Si conduction band or aluminum by photons or electrons, can readily be driven into SiO_2. Energetic electrons are created if an avalanching field is present (as described in Sec. 3.3.3). Therefore, avalanching field in silicon provides another means for electrons to gain sufficient energy to surmount the

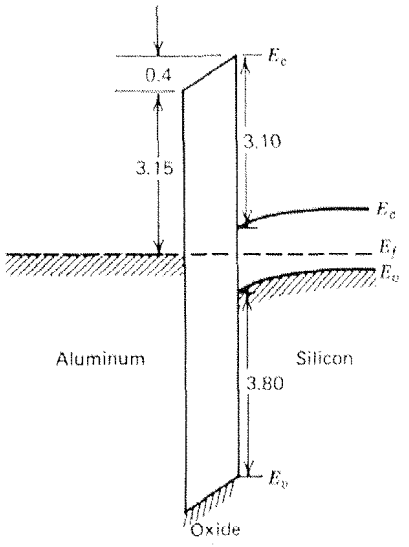


Fig. 4.3: Energy-band diagram for at thermal equilibrium for an MOS system made out of $p-$type silicon. The oxide here is assumed to be free of any charge. Energies in eV. [28]

barrier at the oxide interface. If an oxide presents very high density of trapped holes, that charge may be reduced by causing one of the two phenomena describe above: the photoemission of electrons from either the metal or the silicon or the avalanche in silicon: some electrons will recombine with trapped holes, thus reducing the net positive charge.

4.5 Diffusion

Thermal diffusion is a thermally activated process in which atoms or molecules in a material move from an high- to a low-concentration regions. If performed in an oxygen ambient, then diffusion is complemented by oxidation. Conversely, all oxidations are accompanied by dopant diffusion. It can be visualized as atomic movement of the dopant atom in the crystal lattice by vacancies or interstitials. Since heating makes lattice atoms vibrate around their equilibrium sites, there is a finite probability that an host atoms acquires enough energy to leave the lattice site and to become an interstitial, thereby creating a vacancy. This process is called *vacancy diffusion*. An other possibility is that the interstitial host atom moves from a site to another without occupying a lattice site, thus an *interstitial diffusion* takes place. Detailed description of the physical laws describing the diffusion processes can be found elsewhere [23, 76]. Diffusion of atomic species is affected by several parameters like temperature, the presence of an oxidising ambient, crystal orientation (the rate of non-equilibrium interstitials created by oxidation is different on different crystallographic surfaces, and on silicon the dopant diffusion in enhanced in $\langle 100 \rangle$ bulk with respect to $\langle 111 \rangle$), stress or strain, and heavy doping. During the process, the wafers are held at their edges, separated by 1-2 mm, in a quartz boat and heated in a long, quartz tube that has the necessary gases flowing through it. Temperature control is required to about ± 1 °C, and diffusion times range from some minutes to many

hours.

4.6 Ion Implantation

In order to introduce the required dopant into the silicon in a controlled way, today the standard solution is ion implantation. With this technique dopant atoms are accelerated at high velocity by means of an accelerator so that after striking the surface, they will continue on into the body of the semiconductor. The acceleration voltages used in ion implanters range from a few thousand to a few million volts. The penetration depth of the ions increases with increasing voltage and is typically a few hundred angstroms. In order to attain greater depths, a diffusion step follows the implantation. Even if no further depth is desired, some additional heat treatment is still needed because the implanting causes substantial crystal damage, which must be annealed out. Moreover, this thermal treatment allows most of implanted impurities to become substitutional states instead of interstitial, taking active part to the conduction mechanism. In very high dose implants, damage is so severe that amorphous layers are formed, but under appropriate heat treatment, such layers will regrow epitaxially to form high quality single-crystal layers. At CSEM in order to limit this lattice damage, implantation for silicon detectors are always performed through a thin sacrificial oxide layer: the acceleration voltages used for Boron and Phosphorus (60 kV and 140 kV respectively) are tuned consequently. For detectors, the amount of dopant on strips or GR's is ruled by the dose of impurities used in the implantation. In Fig. 4.4 and Fig. 4.5 are shown the p^+- and n^+-implantation profiles of the MAXIS process employed at CSEM for the CMS microstrip detectors [82], respectively. In Fig. 4.5 is highlighted a MAXIS process peculiarity: the Phosphorus edge implantation is tuned in such a way that overcompensates the Boron already present in the site (more details in Sec. 4.10). Fig. 4.6 shows the backside contact doping profile: both density and depth of doping are larger than those on the front side, following the considerations made on high-low junction in Sec. 3.3.2.c.

A full 2D-simulation takes into account all the details of the fabrication process; we can extract implant profiles as well as their details after patterning, at each stage of the fabrication process. In Figs. 4.4 and 4.5 full thermal budget is encountered, thereby showing the implant depths as they appear the end of the fabrication process. Lateral diffusion is slightly less important than the orthogonal to the silicon surface one due to anisotropic lattice plane effect.

4.7 Photolithography

Photolithography is the means by which a pattern of masking material (usually organic) is applied to the surface of a wafer. That pattern then provides protection to the desired portions of the wafer and allows material to be removed from the remaining area by a suitable etchant. The steps in patterning oxide on a wafer are shown in Fig. 4.7. The photosensitive resist is liquid and is applied by dropping a controlled amount onto the middle

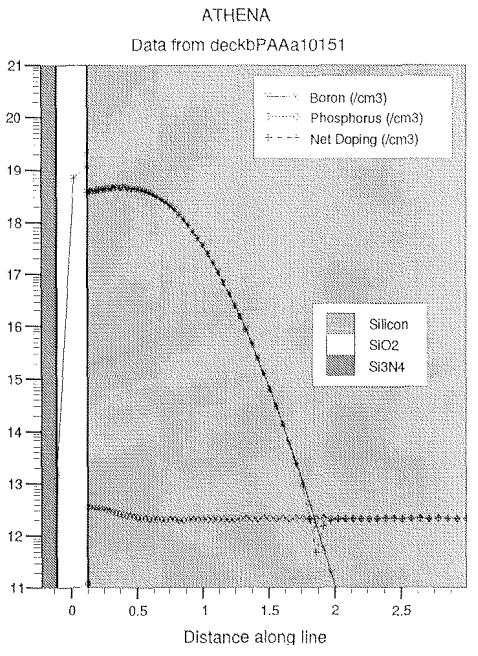


Fig. 4.4: Simulated profile of Boron implantation for the MAXIS process on the front side: strips and guard rings diodes definition. SILVACO ™'s ATHENA package has been used. Ref. [82]

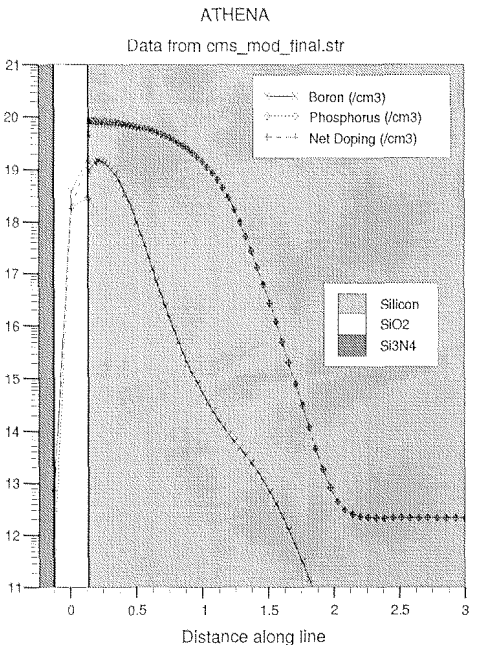


Fig. 4.5: Simulated profile of Phosphorus implantation for the MAXIS process on front side: cutting region. SILVACO ™'s ATHENA package has been used. Ref. [82]

of a slowly rotated wafer. The speed is then increased to a few thousand revolutions per minute so that centrifugal force can force the resist to flow out over the wafer surface and uniformly coat it. The rotational speed and resist viscosity are adjusted to give the required thickness (usually about $1 - 2$ μm). Exposing the resist to a similar pattern of high-intensity ultraviolet light and then developing it produces the pattern in the resist. The light pattern is defined by the pattern on the mask mentioned earlier. The mask is thus the counterpart to the conventional negative of photography. Photolithographic equipment can be easily grouped in two classes: proximity and projection machines. In proximity machines, the mask sits some micrometers above the wafer to be exposed. A light flash will expose the resist: the closer the gap between resist and mask, the smaller the pattern distortions. The critical issue is clearly the cleanliness of photolithographic masks: if some particulate or a small resist residue sticks on the mask, this will translate in a permanent modification of the resist patterning on the wafer underneath, thus affecting the device integrity. Projection equipments interpose a lens between photolithographic mask and wafer. The quality of the patterning depends on lens transparency. At CSEM we have at our disposal for detector fabrication both equipments but since the introduction of the new 1X projection equipment, we noted that the number of defects caused by photolithography dropped dramatically. In proximity machines, defective parallelism between mask and wafer or loose control in resist thickness uniformity may resort in a

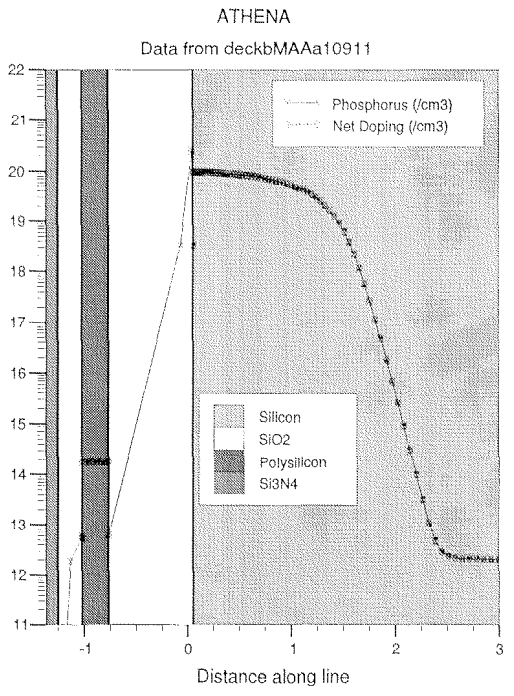


Fig. 4.6: Simulated implantation profile of Phosphorus for the backside contact using MAXIS process. SILVACO ™'s ATHENA package has been used. Ref. [82].

contact and make the mask dirty. Projection machines are less prone to those kind of problems, since masks and lenses are far from wafers and therefore more protected from potential pollutants.

In the developing process, the unwanted resist is removed, thus leaving a pattern of bare oxide or whatever other material was under the resist.

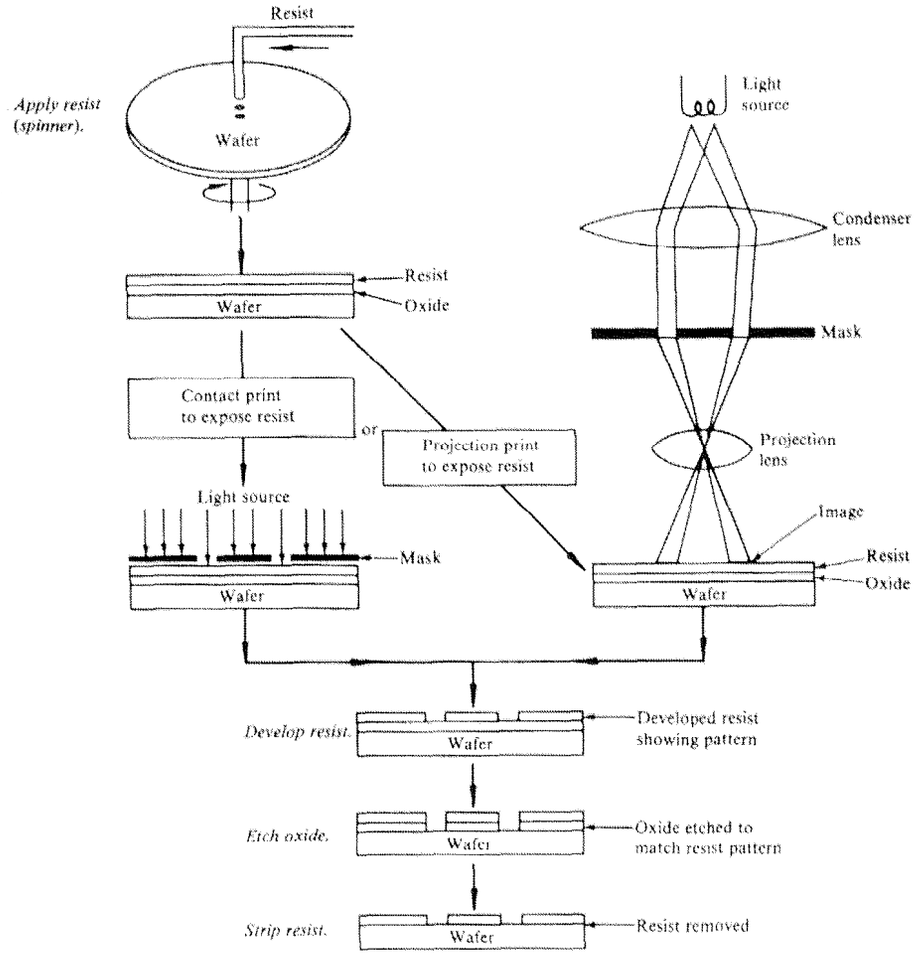


Fig. 4.7: Steps in patterning oxide in a wafers. Both projection and proximity techniques are illustrated. From [76].

4.8 Etching

After the resist pattern is defined, the underlying material must be removed by etching, either by using aqueous-based etches or by plasma etching. In either case, an etch stop condition should be reached at the bottom of the layer being etched. It is dangerous to depend on etch rate and time. Rather, the etch must be chosen so that the rate differential between the layer being etched and the layer immediately beneath it is great enough so that the desired layer can always be completely removed and without significantly affecting the one below. For fine geometries, plasma etching is preferred since it is considerably more anisotropic than wet etching [2]. Typical wet etchants used at CSEM are HF-HNO_3 and for silicon and HF for SiO_2. Plasma etching gases usually contain fluorine or chlorine (i.e. CF_4).

4.9 Metal Deposition

At CSEM metal depositions for silicon detectors are accomplished either by evaporation or by sputtering. Sputtering is preferred for detector fabrication, because it yields better step coverage and allows the deposition of alloys.

Aluminum, with its alloys, is the most commonly used metal, but titanium, tungsten, and their alloys are also used.

In order to achieve good metal-to-semiconductor contact after deposition, heating the wafer to a high enough temperature to allow some interdiffusion and alloying to take place is usually necessary (*sintering*). For aluminum, the temperature is usually between 400 °C and 450 °C. The operation is performed in a diffusion-like tube in forming gas (80% N_2, 20% H_2) atmosphere. Moreover, there are indications [83] that forming gas anneal reduces the component of current arising from the surface since it is known to reduce interface densities in MOS devices [84].

The final resistivity of aluminum or its alloys (i.e. 2.7$\mu\Omega$·cm for Al and up to 3.5$\mu\Omega$·cm for alloys) allows very low resistance (also by choosing a thick enough aluminum sheet, typically 1-1.5μm) for strips, keeping serial Johnson noise at acceptable levels (see Sec. 3.3.1).

4.9.1 Pitting and Diffusion Barrier

Silicon has an appreciable solubility at elevated temperatures in several of the possible contact materials. Aluminum, which has been used since the inception of the planar silicon IC for contacts and metallisation, is an example. Its ability to dissolve thin layers of SiO_2 helps ensuring good physical contact to the silicon even if surface cleaning is not complete. However, silicon will dissolve in the aluminum during the contacting operation and may form pits at the silicon surface. The pits are filled with aluminum, and the phenomenon

[2] Anisotropic etching has two meanings. In reference to single-crystal silicon, it means crystallographic orientation dependent etch rates. In most other contexts, it refers to etching straight down through the pattern window without any lateral undercutting.

is often referred to as *aluminum spiking*. Fig. 4.8 shows the way the dissolution (alloying) takes place.

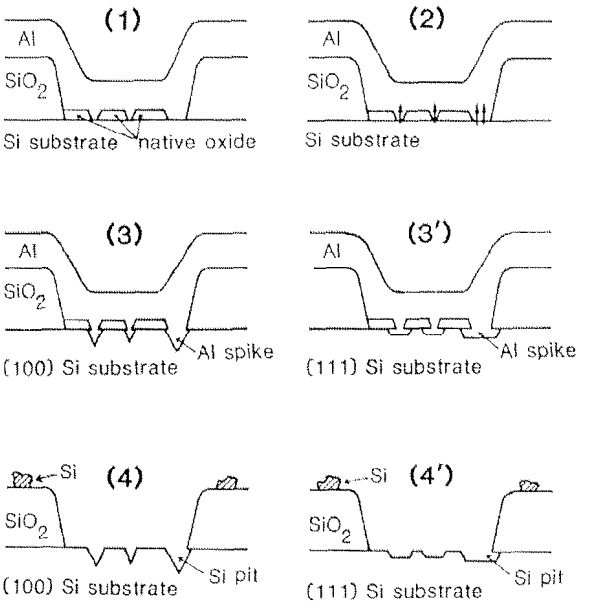


Fig. 4.8: Schematics of formation of Al spikes and Si pits caused by Si dissolution in Al lines. Both ⟨111⟩ and ⟨100⟩ Si orientations are shown.

This pitting is seen well below the aluminum-silicon eutectic temperature of 577 °C. The high diffusivity of silicon in aluminum accounts for the observed effect. Like aqueous etching, the etching of silicon by metals is crystallographic orientation dependent, and as in aqueous etching, the slow etching planes are the {111}s. Because of the slow dissolving ⟨111⟩'s, it is possible to get very flat alloy fronts when alloying into surfaces ⟨111⟩. Thus, silicon alloy transistors, where shallow junction are mandatory, were all made in ⟨111⟩ oriented wafers. The equilibrium form of patterns formed by the aluminum-silicon interaction will appear as flat-bottomed, triangular pits in ⟨111⟩ and square-topped, pyramidal pits in ⟨100⟩ material. Since no ⟨111⟩ plane is parallel to a ⟨100⟩ surface to act as an etch stop, the aluminum-silicon interface of contacts on ⟨100⟩ wafers will not be smooth. Of more importance, because the thin layer of interfacial oxide that must first be dissolved will not fail uniformly, different regions will begin reacting at different times. The first ones can allow enough silicon into the aluminum to saturate much of the contact and thus limit dissolution over the remainder of the contact area. The unfortunate result is that the first pits will be much deeper than the rest and, in shallow junction devices, may extend through a junction and cause shorting. Fig. 4.9 shows scanning electron microscope photographs of ⟨111⟩ and ⟨100⟩ silicon surfaces after the aluminum contacts have been removed. The important pitting of the ⟨100⟩ surface is clearly visible.

Experimentally, it has been reported that pure aluminum sintering at 300 °C produces discernible pitting with some depths to 0.2 μm. At 350 °C, pits of 0.75 μm have been observed, and at 450 °C, pits of 2 μm. Individual behaviour will, of course, depend on the amount and placement of the aluminum over and around the contact window.

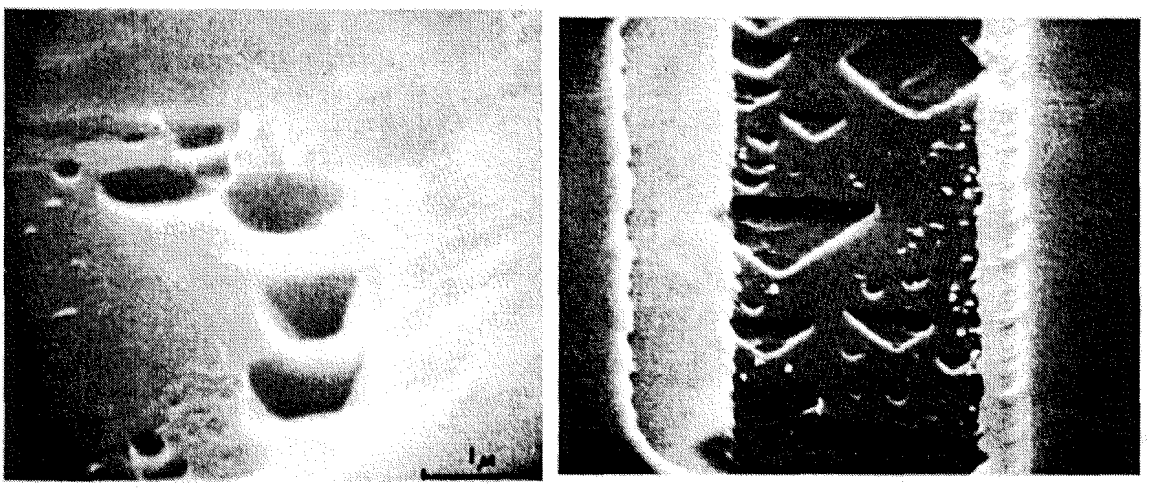
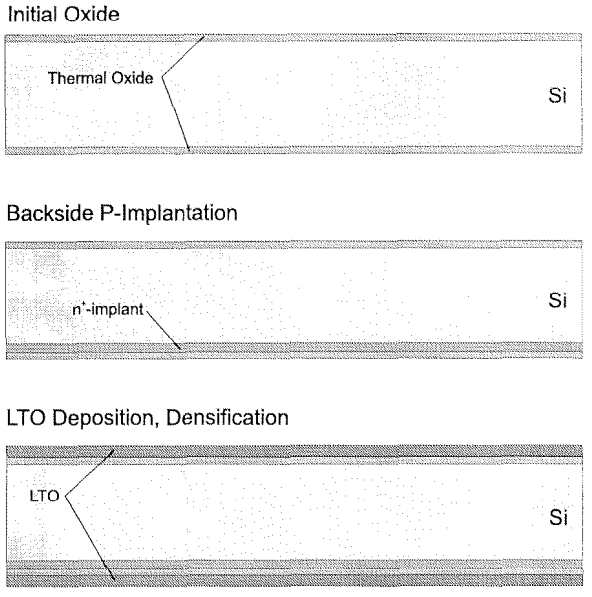


Fig. 4.9: SEM photographs of evidence of pitting: on ⟨100⟩ silicon (left) and ⟨111⟩ silicon (right). Notice the typical triangular shape in the second case.

In order to prevent such pitting, the aluminum may be deposited saturated with silicon (in the order of 1% in volume) so that it is unable to absorb any more. A alternative method is to introduce a barrier metal layer between the aluminum and the silicon substrate. This barrier metal layer must meet the following requirements: a) it forms low contact resistance with silicon, b) it will not react with aluminum, c) its deposition is compatible with the overall process. Typical barrier metals are W, Ti, TiW or TiN. For CMS microstrip detectors there are no particular reasons to resort to shallow junction. Furthermore, using ⟨111⟩ substrates, the problem could be alleviated. For these reasons, pitting has up to now be handled by resorting to Al-Si(1%) alloy. Of course, if ⟨100⟩ is used, or if for some reasons the junction depth has to be made thin enough for pitting to occur with a statistically non-negligible probability, then further methods should be considered.

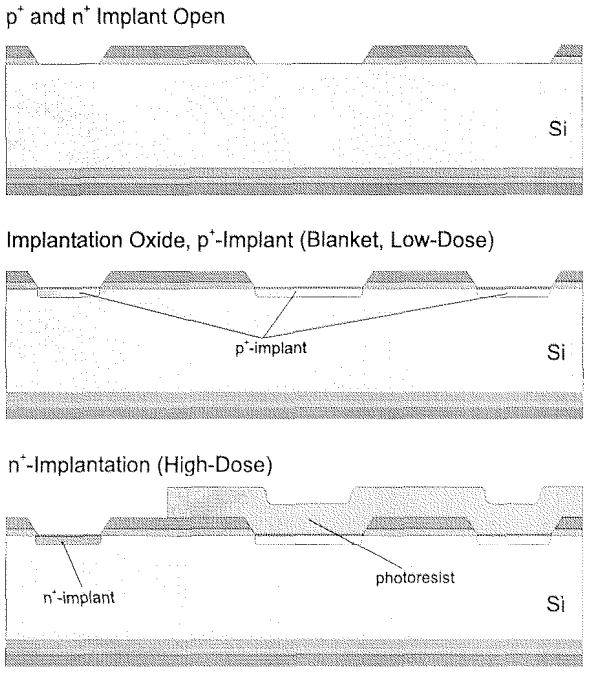
4.10 The MAXIS detectors fabrication process

At CSEM we developed a special variation of a commercial CMOS fabrication process, internally called MAXIS, in order to fulfill all the requirements previously described in the Chapter 3 and successfully fabricate the CMS microstrip detectors.



The flowchart of the process can be followed in a graphical way through the figures embedded into the text. The process begins with a careful inspection of the bare wafers and a soft wet etching in order to eliminate particulates and the native oxide. Wafers are then oxidised in order to grow an implantation oxide. The rear contact high-dose Phosphorus implantation is performed, followed by the deposition of the field oxide, a thick layer (7000 Å) of LPCVD oxide, that will have a twofold role: protect the backside from mechanical scratches and chemical pollution during the entire fabrication and prepare the frontside to the implantations.

Its further densification through a drive in a high temperature furnace, will act as a first activation and diffusion for the backside implantation.

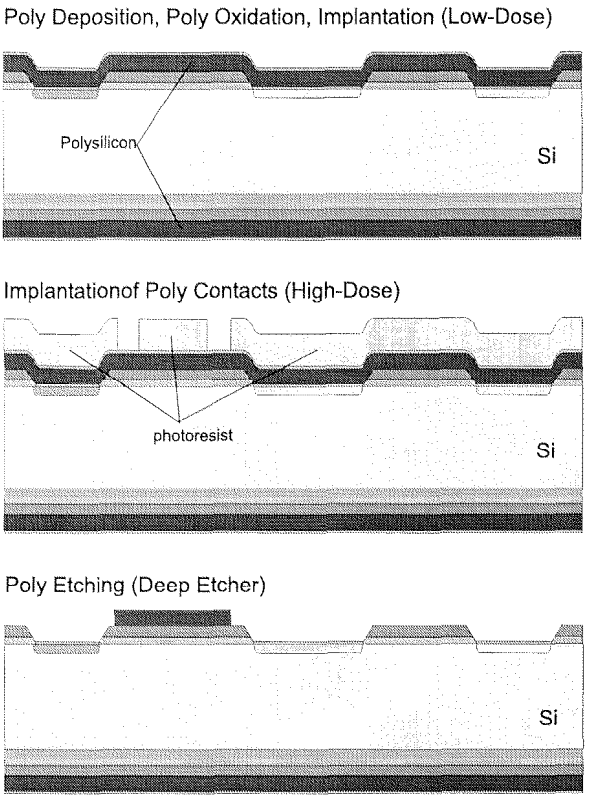


At this stage it is necessary to define via a photolithographic step the diodes (strips and GR's) patterning on the frontside. In the MAXIS process the patterning is made merging p^+- and n^+- geometries. The oxide opening is made via a dry (plasma) etching. A thin thermal oxide is regrown in order to protect the bare Si surface from implantation damage and then, as a first ion implantation, Boron is implanted. The strips and the guard rings are so defined.

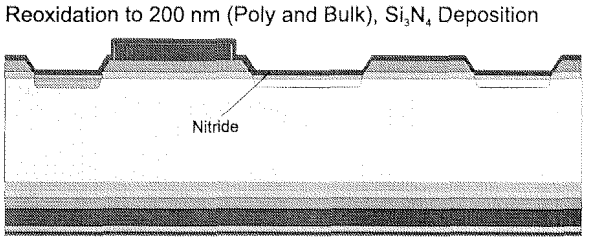
The detector edges need to be protected from eventual external carrier injection once they are cut (see Sec. 3.3.3.c): the edge region, which will be implanted with phosphorus, is defined by a further lithography step and the photoresist will shield the diode region from being implanted.

The dose and energy of this implantation have been chosen such that it will overcompensate the presence of boron (see Fig. 4.5), since parasitic junctions in this region would have a nasty influence in the detector performances.

After removal of the photoresist, we perform a LPCVD deposition of polysilicon, 3500 Å thick, in order to define the strip bias resistors. The low-dose Phosphorus implantation, necessary to define the required resistivity for the resistors, is made through a thin thermal oxide. This oxidation acts as densification for the dielectrics and diffusion for the implants. A further photolithographic exposure defines the resistors contact regions, i.e. where they will be connected through a metal layer to resp. the bias guard ring and the strip. In order to minimize the contact resistance, the MAXIS process resorts to another phosphorus implantation, at higher dose. The contact resistance so obtained amounts on average of 20Ω/□. Once implantations are done, polysilicon resistor are defined, by means of dry etching. Great care is necessary in performing this etching step, in order to not damage the thin oxide laying above the strip diodes, defining the first dielectric layer of the integrated decoupling capacitor.

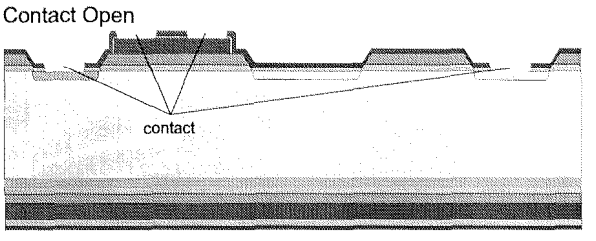


After strip cleaning a thermal oxidation is performed, in order to precisely define to 2200 Å the oxide thickness of the integrated capacitors. This thermal process acts as further diffusion for all the implantations. A second layer of dielectric, 1250 Å of Si_3N_4, is then deposited onto the wafer.



The latter acts as protection layer for all the structures present on the surfaces and it enhances the strength of the strip integrated capacitor, possibly decoupling an existing pinhole (a hole in the first layer of SiO_2), thus preventing that a conductive path could be established between the diode and the metal electrode on top. Moreover, the relative permittivity of silicon nitride is almost the double of that of silicon oxide ($\epsilon_{Si_3N_4} = 7.5$ vs. $\epsilon_{SiO_2} = 3.9$), allowing the deposition of thicker layers, yielding precise thickness control even for a deposited layer.

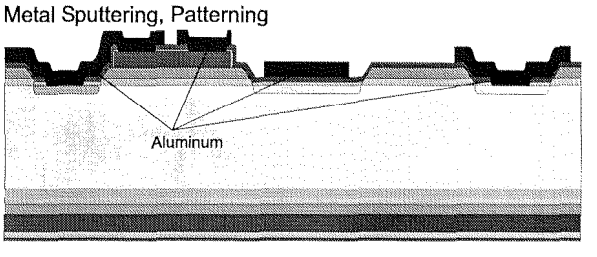
The contact holes are then opened. The photoresist defines the regions in which we need a DC contact to the implants: bias ring, spy contacts, etc.

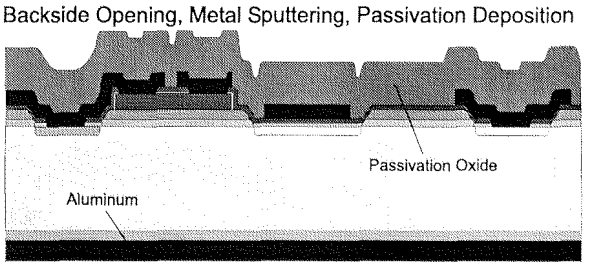


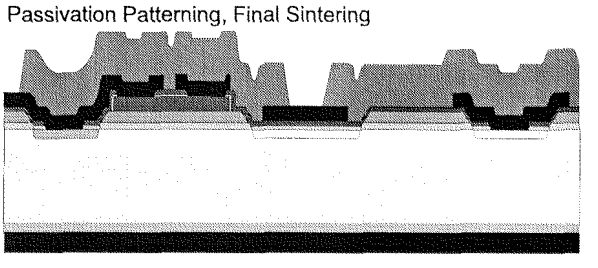
Just two more steps are missing: metallisation and passivation. A 1.5 μm thick Si-Aluminum alloy is sputtered on the frontside of the wafer, followed by its patterning, via dry etching. The deposition of 1 μm PECVD oxide on the frontside, as a passivation layer, ends the structuration of the front side. This oxide prevents accidental scratches to damage the surface and it helps in shielding the active surface from chemical pollution. Last but not least, manipulation of passivated detector is significantly alleviated.

Once passivation is in place, turning the wafer upside-down in order finish structuration on the back, it will not damage the front side.

On the backside, all the dielectric layers we deposited and removed on the front, are still in place. We have chosen this option to keep maximal protection of the high-low junction. Mechanical scratches on an exposed backside surface could have caused a weakening or even the drilling of the rear junction, compromising the electrical behaviour of the detector.







In fact, most of the soft-breakdowns visible in I-V curves at high reverse voltages (e.g. see Figs. 5.5 or 5.11) are caused by bad rear contacts[3]. For the sake of clarity, the rear contact in single-sided detectors (like CMS) covers the full wafer surface: the presence of a single defect may cause the failure of the device! The larger density of defects lays on wafer edges, where chipping or equipment supports may damage the wafer. This fact explains also why detector's BKD performance are often better after cut.

The contact opening on the backside is made via combination of dry (plasma) and wet (dip) etching. This approach allows to exploit the strong anisotropy of dry etching together with the excellent selectivity between Si and SiO_2 of wet etching. By this method we are sure that we do not dig into the doped silicon, weakening the high-low junction or, even worse, shortcircuiting the Si-bulk with the metal.

[3]It has been experimentally demonstrated that HV current breakdowns in devices fabricated with the MAXIS technology are not caused by strip implant defects but by local weakness of the backside contact. Photoemission microscopy did not revealed any hot spot on the detector front side [85].

Once dielectrics are removed, a suitable backside metallisation is deposited onto the rear surface.

Opening the passivation in selected areas, which allow the contact with the metal beneath, is the last photolithographic step.

A thermal treatment ends the fabrication: this step (called sintering) is needed to allow better adhesion of aluminum with silicon, lowering the contact resistance.

4.11 Gettering: the control of minority carriers lifetime.

Traditionally a major goal in detector fabrication is the minimization of the diode reverse-leakage current, which contributes shot noise and can limit the dynamic range of the input amplifier. For silicon detectors developed for operation in high irradiation environments, like LHC, leakage current is not the major concern. Nevertheless, low leakage current is considered as an important quality factor, defining a clean fabrication environment and a good mastering of the technology, preventing unforeseen yield loss and by-problems.

Leakage current contributions have been studied in detail in Sec. 3.3.2: minority carriers may arise from the undepleted neutral regions, the surface and the depleted bulk. In order to minimize reverse junction leakage current of a fully depleted silicon detector, high semiconductor bulk lifetime is desired. Depleted bulk generation is greatly enhanced by the presence of impurities that introduce energy levels near the centre of the silicon bandgap. Such impurities are often referred as lifetime killers, since they also reduce the generation lifetime: gold, nickel, copper, and iron are typical examples but all of the transition metals (Ti, V, Cr, Mn, Fe, Co, Ni) are reported to be deleterious [86]. The most common effect is that of lifetime reduction, but gold, which is the earliest lifetime killer identified [87], for example, also reduces mobility in MOS structures [88]. In addition, any of the metals with high solubilities at processing temperature and very low solubility at room temperature are prone to form precipitates and cause excessive leakage or even direct shorts between active elements. The presence of these metals will cause pits after oxidation and lead to the subsequent formation of stacking faults.

Great care should be given to the cleanness of the fabrication environment, the purity of chemicals, the way different products are managed into the fab and the handling of detectors before and after fabrication. However, the total elimination of these contaminants has proven to be impractical due to the pervasiveness of such elements and the sensitivity of the leakage current to extremely small concentrations, typically part per billion or less. This has led to the development of techniques, collectively known as *gettering*, which render any lifetime-killing impurities electrically inactive.

The basic concept of gettering relies on the fact that at typical integrated circuit processing temperatures in silicon (approximately 900-1100 °C) most harmful impurities have relatively high diffusivities and hence are very mobile. Therefore suitable impurity sinks can be incorporated in electrically inactive regions. Two general classes of gettering are commonly employed. In *intrinsic gettering*, the bulk of the wafer forms the impurity

sink with a relatively thin, defect-free region near the surface where the active devices reside. The technique is often applied to CZ-grown substrates and relies on the formation of oxygen precipitates in the bulk of the wafer. While perfectly adequate for conventional IC processing, intrinsic gettering presents some difficulties with usual silicon detectors, where the depletion region often extends to the backside contact of the wafer. The second class of gettering, *extrinsic* gettering, is more commonly used. Extrinsic gettering relies on the formation of a gettering layer on the back side of the wafer, away from the active region of the device, e.g. the depletion region of detector diodes. Typical gettering methods are listed in Table 4.3 and will be briefly described later on.

Method	Comments
Back-side phosphorus diffusion	
Back-side abrasion	Usually uses sandblasting.
Ion implant damage	Typically uses argon ions.
Silicon nitride deposition	$\sim$ 4000 Å layer deposited on the back of the wafer.
Back-side polysilicon deposition	Combination of polysilicon, residual oxide, and diffusion temperature introduces crystal damage.
Oxygen precipitates	Requires correct oxygen level in crystal and a prescribed heat-treat cycle to be effective.

Table 4.3: Overview of different gettering approaches in silicon processing

4.11.1 Extrinsic gettering

Heavy $n-$layers on the back of the wafer were the first gettering sinks used and are probably still the most common since their use requires no additional processing steps. Heavily n-doped regions have increased substitutional solubility for those lifetime killers that behave as acceptors when substitutional (for example, Au and Cu), so, in principle, heavily doped phosphorus, arsenic, or antimony layers should perform equally as well. Experimentally, however, it is found that phosphorus is much more effective than others are. The backside contact implantation of the CMS detectors acts perfectly as a sinker.

Backside ion implantation gettering is now used with regularity. By selective masking, it can also be used locally on the front of the wafer and possibly be closer to the volume needing gettering than if it were on the back-side: in fact, the MAXIS technology foresees a phosphorus implantation at the device edges, which can be effective for getter the regions close to the GRs. Methods of annealing and the point in the process at which the implant is done affect the final results, as does the implant species and the wafer orientation. Much more residual damage remains in $\langle 111 \rangle$ oriented silicon than in $\langle 100 \rangle$, and the gettering

efficiency is correspondingly higher.

Regions of misfit dislocations induced into ⟨100⟩ wafers have also been shown to be effective gettering sites.

Mechanical backside damage gettering has been studied at least since 1965. By using controlled sandblasting or other forms of light abrasion, the results are reasonably reproducible, and indeed wafers with such damage are now standard items of commerce. This sort of damage is applicable to processes where the formation of a heavy phosphorus layer presents problems. Within limits, the more abrasion, the more pronounced the gettering effect.

A silicon nitride layer deposited on the back of a wafer will provide enough stress during high-temperature processing to effectively getter [89]. Both the low-temperature plasma-deposited nitride and the higher-temperature CVD nitride appear effective. Thickness' required are in the $1000 - 4000$ Å range. Gettering is less for the thinner films, and wafer bowing can occur if the film is much above 4000 Å.

A thin layer of low-temperature polycrystalline silicon deposited on the back of the slice will provide for gettering sinks [90]. The polycrystalline layers are deposited at low temperature, such as 650 °C, to a thickness of about $0.5\mu m$. Even if the layers are thin enough to be fully oxidised during processing, stacking faults that continue to getter are propagated into the single-crystal silicon.

A different kind of gettering, and one not included in Table 4.3, is the use of an atmosphere in the furnace tube during diffusion (or oxidation) that will assist in transferring unwanted impurities from the wafer surface to the gas stream. It has been known since 1960 that the use of chlorine or a chlorine-bearing species such as PCl_3 or BCl_3, would improve lifetime. In the case of phosphorus and boron, the relative impact of the glassy layer, the heavy doping in the silicon, and the chlorine species in the tube ambient were not resolved. Since then, the use of HCl to clean tubes has become standard practice. The mechanism is one of forming chlorides volatile at the tube temperature, and the same mechanism can be used to clean wafers if etching of the surface can be prevented. Such conditions prevail during HCl oxidation, and substantial lifetime improvement is sometimes observed [91]. In this case, the effectiveness will also depend on how easily the impurity to be gettered can diffuse through the protective thermal oxide and whether or not it is trapped at the interface. In the case of gold, it appears that there is a substantial pileup at the Si-SiO_2, interface and that gold is not effectively gettered from wafers during an HCl oxidation.

4.11.2 Intrinsic gettering

The damage caused to silicon detectors by intense hadron fluxes can lead to the failure of these devices before the end of the experimental program (see Sec. 3.3.5). The radiation tolerance of silicon can be improved by the deliberate introduction of impurities into the silicon crystal [92, 93]. The impurity atoms can react with the primary induced defects (vacancies and interstitials) and effect the formation of electrically active centers. Certain

impurities can thus influence the changes of the electrical parameters of the devices, such as the increase of the full depletion voltage due to the changes in the effective spatial charge in irradiated detectors. In particular, oxygen atoms can capture vacancies, reducing the rate of formation of V_2O, divacancy (or multivacancy) complexes and therefore limiting the rate of degradation of the electrical properties of silicon detectors. The deliberate introduction of oxygen in the bulk of a silicon wafer generally refers as *intrinsic gettering.*

Usual IC integrations makes use of wafers made with the Teal-Little technique (often-called CZ, or Czochralski) from a fused silica container. The molten silicon dissolves a small part of the fused silica, and some of the oxygen from it is incorporated into the single-crystal silicon. Under ordinary circumstances, an as-grown CZ crystal is supersaturated with oxygen, since typical oxygen concentration specifications range from $1.25 - 2 \cdot 10^{18}$ atoms/cc. As the crystal comes from the puller, most of the oxygen is dispersed as interstitial atoms and is electrically inactive. However, heat treatments such as are involved in diffusions can cause it to aggregate and eventually form rather large precipitates. Treatment in the 1200 °C temperature range will cause a redissolving of the oxygen structures, but any crystallographic defects, such as stacking faults that formed because of the precipitates, will remain. The oxygen-precipitation-induced defects act as heavy-metal gettering sites, but if such sites are within the active volume of the IC devices, performance degradation results.

All the high-resistivity silicon crystals used for microstrip detector fabrication are grown by the Float-Zone process (FZ), since hyper pure silicon assure longest minorities lifetime and therefore lower leakage. A significantly high oxygen concentration ($> 10^{17}$atoms cm^{-3}) cannot readily be achieved during FZ silicon refinement; usually the quantity of oxygen in FZ wafers is 2-3 orders of magnitude lower than in CZ silicon. In order to increase the density of oxygen is then necessary to perform a thermal oxidation followed by deep diffusion. The ROSE Collaboration together with the ITE institute developed a dedicated process, growing ~ 3000 Å of dry thermal oxide on standard FZ detector grade silicon. The samples were then heated at 1150 °C in an inert ambient (N_2) and kept there for 24, 48 and 72 hours, in order to make oxygen diffuse.

Figure 4.10 shows the profile of the oxygen concentration as a function of the depth in the silicon, as measured by SIMS. The oxygen concentration achieved a value of $\sim 1.5{\times}10^{17}cm^3$ in the middle of the wafer (150 μm) after a diffusion time of 24 hours and a homogeneous value of $\sim 2.5{\times}10^{17}cm^3$ all over the silicon bulk after 48 hours diffusion time. This technique also guarantees optimal radial homogeneity of the oxygen concentration.

The general approach of typical IC processing to intrinsic oxygen gettering is to first form a region from the wafer surface toward the wafer interior that is thick enough to contain the devices made in that wafer and has so little oxygen that it will not precipitate. Later, during subsequent processing, this region will remain free of oxygen-induced defects and thus not adversely affect devices made in it. The high-temperature operation to allow oxygen near the surface to out-diffuse and also dissolve any oxygen aggregates that may be in the wafer is referred to as a "denuding" step. After denuding, the wafer is subjected to a low-temperature heat cycle in order to allow precipitate nuclei to form. With this preparation, the dissolved oxygen will form large precipitates during the heat

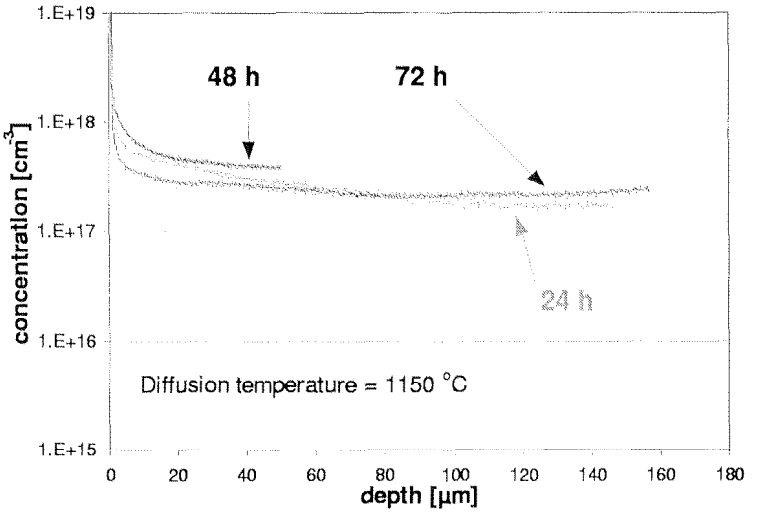


Fig. 4.10: Profile of the oxygen concentration as a function of the depth in the silicon, measured by SIMS by the Evans Europa Company. Samples processed by ITE for 24, 48 and 72 hours.[94]

cycles associated with subsequent oxidations and diffusions. These large precipitates, in turn, produce crystal defects that act as gettering sites.

In the case of silicon detectors, the wafer region occupied by the device (the detector) represents $\sim 50\%$ of the overall surface (on 4″ wafers) and the entire bulk is active part of the device. Therefore it is not possible to surround the active area with oxygen precipitates. The oxygen will diffuse through the whole wafer and thus oxygen precipitates will be present everywhere except close to the surface due to the denuding high temperature phase. Before irradiation, an overall leakage deterioration is expected, if compared with similar samples without increased oxygen content. Conversely after type inversion, the gettering action of the oxygen-induced defects will limit the increase of N_{eff} with fluence, thus improving its radiation hardness, allowing manageable depletion voltage until the end of LHC operations (see Fig. 4.1).

4.12 Conclusions

At the end of Chapter 3, I drafted the recommendations for an effective design of the CMS microstrip detector. Those hints should be combined now with the conclusions of the technological optimization: design and fabrication process requirements defines the road towards the CMS microstrip detector able to operate 10 years in the harsh LHC environment.

In the following I summarize the precautions we took at CSEM in defining the MAXIS process in order to match fabrication yield requirements, necessary to keep silicon detector an economically viable product, and the CMS physics requirements, settled by the LHC environment, the geometry of the SST and elementary particle physics goals. The result is a set of process recommendations suited for the silicon detector of the CMS SST.

The choice of the silicon substrate must reflect a good balance among signal-to-noise, depletion voltage after irradiation and mechanical yield. Its resistivity, defined as doping density, must be extremely uniform, in order to avoid variations in depletion voltage throughout the wafer. The quality of polishing (double-sided) should be good enough to eliminate local mechanical defects (i.e. surface scratches) and avoid silicon particulates detaching from edges, falling back onto the Si surface during processing, lowering overall device yield. Strips and GRs are defined by means of p^+-implantation. Dopant implantation energy, its dose and its further thermal diffusion should be tuned to assure the smoothest profile as possible, in order to limit the occurrence of critical fields at junction edges. Rear phosphorus implantation must assure an excellent ohmic contact, able to sustain several hundreds volts of depletion voltage after type inversion and stopping charge injection from the rear contact. Thermal treatment after implantations must assure the activation of most of the dopant. The definition of the strip integrated capacitors must be done by using a multi-layer of thin dielectrics. The combination of thermally grown silicon dioxide and deposited silicon oxide and nitride assures:

- high density, low trapped charge, homogeneous coverage and low intrinsic pinhole distribution, exploiting the excellent properties of thermally grown SiO_2;
- enhance the breakdown performance of the decoupling capacitor;
- decouple layer-intrinsic defects by means of dielectric multi-layer structure, strongly reducing the overall number of metal-implantation short circuit (pinholes);
- allow good precision and stability of decoupling capacitances. The high dielectric constant of silicon nitride allows easier control of the thickness, in spite of being deposited.

Polysilicon resistors can be used as biasing resistors: dopant implantation and deposition conditions will set size and boundaries of poly grains, thus defining its resistivity. Metallisation must avoid pitting and its thickness should be tuned in order to limit serial noise, 1.5μm thick aluminum is enough and mostly geometry independent. An LPCVD silicon dioxide passivation layer protect both sides from mechanical scratches, simplifies the handling and weaken the influence of ambient conditions (i.e. humidity) or other pollutants (i.e. alkaline metals) on the detector performances. Device instabilities and high leakage currents must be counteracted by means of efficient gettering techniques: both intrinsic and extrinsic ones could be effectively used, but extreme care should be given to their insertion in the process flow and their combination with the existing stages.

The MAXIS process allow us to reliably fabricate radiation hard, AC-coupled, poly biased silicon detectors: in the last two years this technology changed dramatically both CMS and CSEM expectations on silicon detector viability as matter of production yield and electric performances. A compact process like MAXIS, with a reduced number of photolithographic steps allows also rapid duty cycles, well suited for large productions and low cost. Mechanical yield close to 95% assure large number of devices available for testing. Very high electrical yield (as it will described in Chapter 5), mostly due to integrated capacitors yield higher than 99.95%, a value almost incredible until one year

ago, allows reliable predictions as far as delivery schedules, assembly organization and overall silicon tracker performances are concerned.

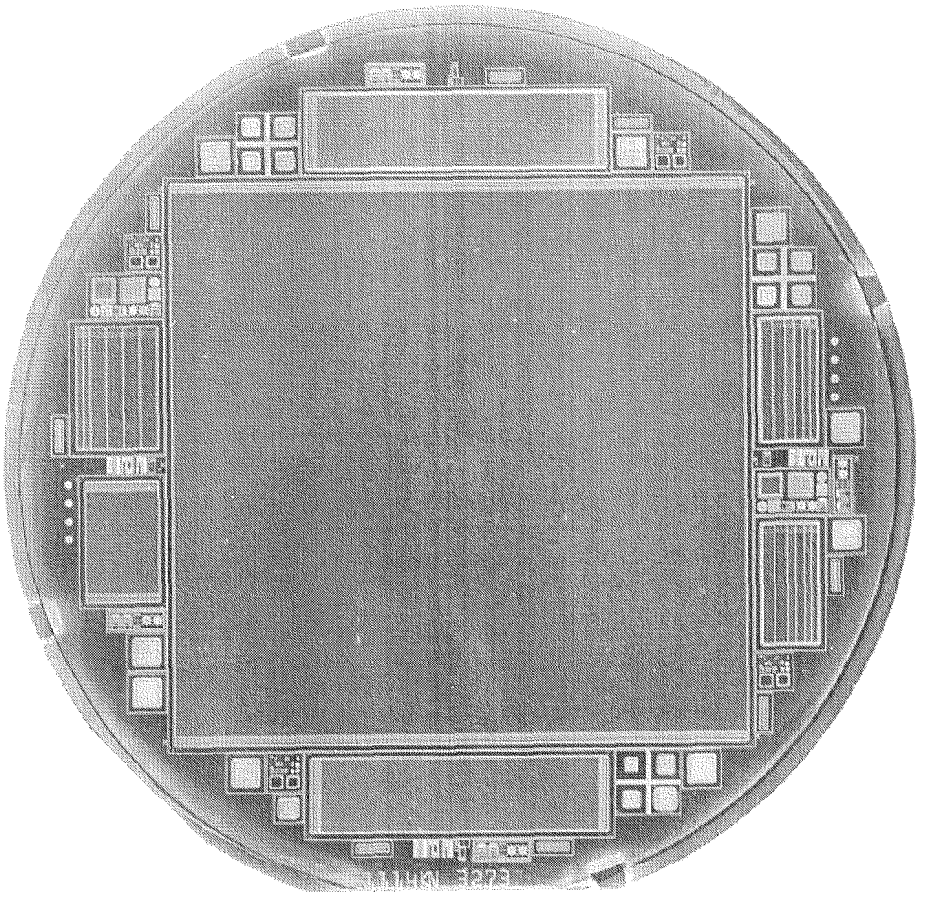


Fig. 4.11: CMS Barrel prototype photo

Chapter 5

Test and Measurements

In this chapter the electrical tests used to qualify the CMS SST detectors are described. A number of technological options have been implemented and tested in order to finalize the fabrication process definition before starting the extensive production.

Most of these electrical test have been performed at CSEM and in the INFN silicon detector laboratory of Pisa (Italy). The irradiation tests presented hereafter have been performed at CERN and at the ENEA centre of Casaccia (Roma, Italy).

5.1 Electrical Characterization

The electrical characterization of a complex circuit as a microstrip detector should be divided in two phases:

- fabrication process control,
- device specific electrical functionality,

the first meaning probing test structures specially designed to check unambiguously fundamental process parameters, like sheet resistance, thickness of depositions, width (CD's, i.e. critical dimensions) of implantations or metal lines, oxide characteristics (breakdown or capacitance) or transistor thresholds; the latter being related to the CMS specifications for the SST detectors in order to complain with the overall performance of the tracker as matter of noise, thermal dissipation, position resolution, etc.

5.1.1 Measuring apparatus

The electrical measurements have been performed in the test laboratory of CSEM and at the CMS test laboratory of INFN Pisa: both laboratories are equipped with a clean room, where cleanliness, temperature and humidity are strictly monitored in order to avoid wafer pollution and insure measurement repeatability.

The equipment is very similar in the two laboratory and functionally identical. At CSEM we have two different measurement setup for R&D studies and production tests. The first accounts for a manual probe station (Micromanipulator 6000) with 6 manipulators, a semiconductor parameter analyzer (HP 4145A), a voltage source and a LCR-meter system (HP 4284A). Both measurements and data acquisition are done manually.

The production setup consists in a fully automatic probestation (Electroglas 2001X) with a probecard with 60 tips coupled with a integrated circuit parameter analyzer (Keithley 450SX) whose performances are enhanced with a current/high-voltage source (up to 1200V) (Keithley 237), in order to allow precise breakdown tests. Data are collected and analyzed via a Keithley proprietary software running on a Sun UNIX station.

In Pisa it is available a semi-automatic probestation (Karl Suss PA150), two voltage/current monitor source (HP4142A and KT237) and an LCR-meter (HP4248A). Data acquisition is done on a Macintosh using Labview interface.

5.1.2 Fabrication Process Control

The characterization of the fabrication process is performed both during production and immediately after its end.

A certain number of tests is needed by the process engineers in order to monitor the performances of the most critical fabrication steps.

At CSEM I implemented a set of test structures (PCM, Process Control Monitor) organized in blocks, each one specific for a particular process step (implantation, dielectrics, polysilicon, metals). These structures are nowadays standard for every CSEM integration and design independent.

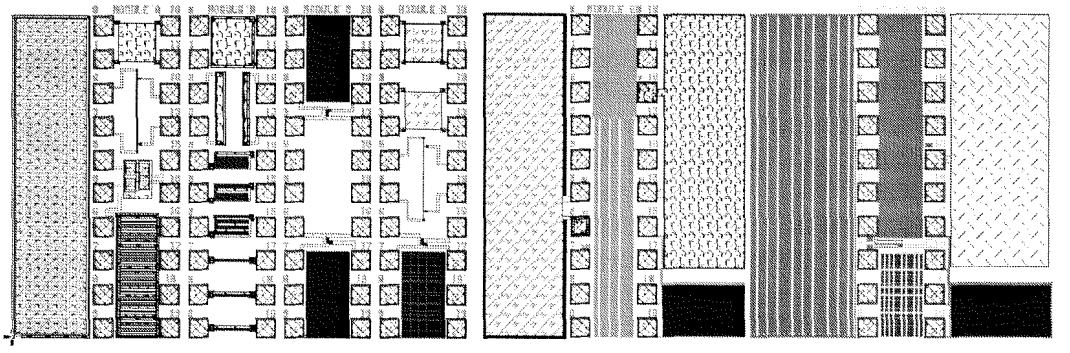


Fig. 5.1: Picture of the CSEM set of PCM (Process Control Monitor). Set A: p^+ implant monitor; Set B: n^+ implant monitor; Set C: contact opening; Set D: polysilicon; Set E: metal 1; Set F: via and metal 2.

At the end of the production, the wafers are passed through an automatic test machine which scans (by means of a specific probe card) all the sites of each test structure for every block integrated on the wafer, saving results on tape. The program takes into account

fiducial intervals for every parameter to be checked, labeling every single value out-of-bonds, allowing a very fast monitoring of potential problems. A preliminary scan of PCMs before functionality tests allow to skip bad wafers, sparing precious time.

Besides that, some other test structures are always inserted into the design of the detector within the wafer, in order to complement the information of PCMs. Those structures are simple diodes, guard ring diodes (GRD), multiguardring diodes (MGRD), gate-controlled diodes (GCD), MOS capacitors and transistors, Van der Pauw and Kelvin structures. In Fig.5.2 it is shown a typical example of one block of test structures.

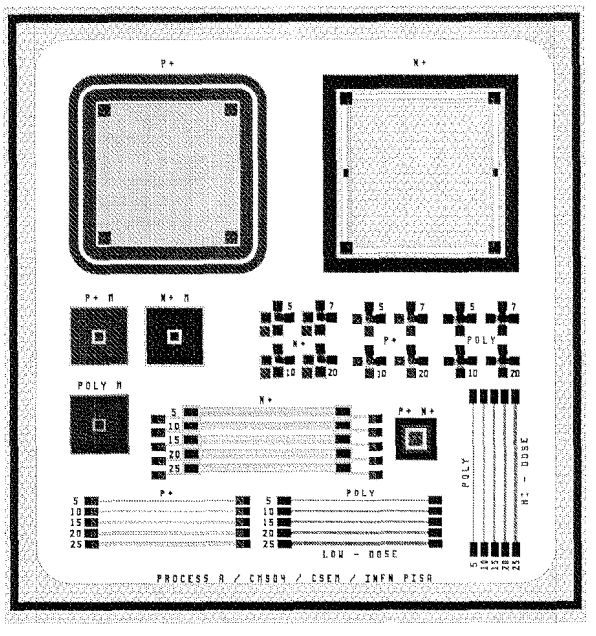


Fig. 5.2: Typical example of a test-structure block inserted in a CMS wafer layout.

Those structures are available in order to cross check information coming from PCMs and the detector. Their interest resides in the possibility to eventually characterize a physical behaviour, easily and more precisely than with PCMs.

In the following I will describe a certain number of fundamental tests that are necessary to study the behaviour of a detector.

5.1.2.a Simple and GR Diodes

With this kind of structure it is possible to extract fundamental information about the substrate in use (doping and thickness), the generation lifetime of the process and the relative weight of generation and diffusion leakage current. Once the geometry of the diode is known, both C-V and I-V curves should be performed.

- **Capacitance-Voltage (C-V) test.** In Fig. 5.3 typical C-V and $\frac{1}{C^2} - V$ curve of a diode is shown. In both graphs V_{bias} is the diode reverse voltage.

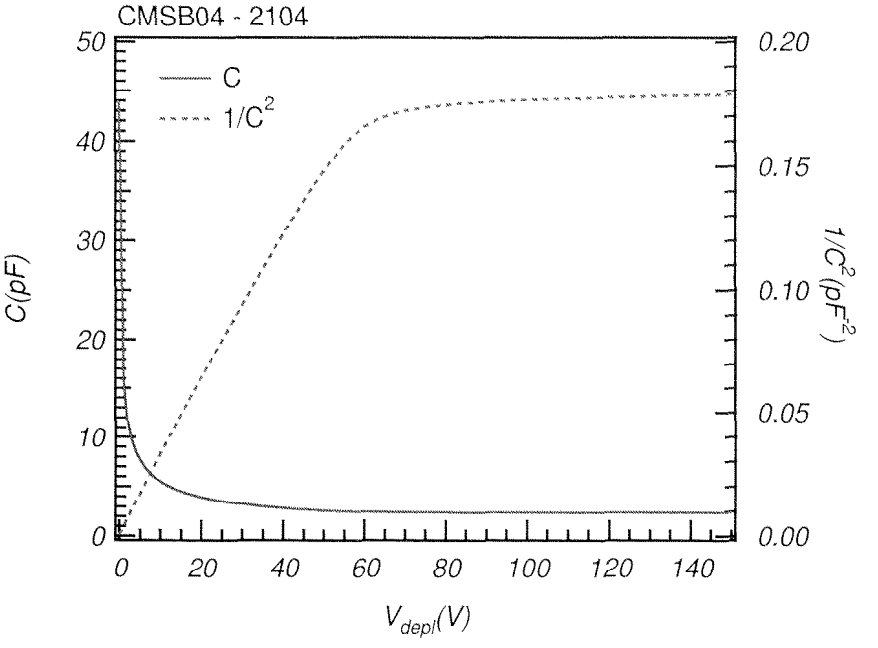


Fig. 5.3: C-V and $\frac{1}{C^2}$-V curves for a typical guard ring diode with A=0.07mm^2. Measurements are made at 100 kHz.

device	A (cm^2)	C_{min} (pF)	V_{depl} (V)	W (μm)	N_d (cm^{-3})
2104	0.0697	2.48	58	292	$8.45 \cdot 10^{11}$

Table 5.1: Typical results from diodes of CMSB4 series (milestone). C-V are made at 100 kHz.

Using Eq. (3.38), we can calculate the capacitance of the diode as:

$$C = \frac{\epsilon_{Si} A}{W} \tag{5.1}$$

where A is the device surface. By squaring Eq. (5.1) and taking the inverse, we get:

$$\frac{1}{C^2} = \frac{2}{qA^2 N_d \epsilon_{Si}} (\phi - V_{bias}). \tag{5.2}$$

The information we can extract from this plot are:

1. Substrate doping N_d, by taking the derivative of Eq. (5.2) in V.
2. Wafer thickness, by taking the maximum value of $1/C^2$ which corresponds to C_{min} and therefore to the maximum depleted thickness;
3. Full depletion voltage, the voltage corresponding to the knee of the $1/C^2$ curve.

For the device shown in Fig. 5.3, Table 5.1 resumes the results.

- **Current-Voltage (I-V) test.** Reverse current versus voltage curves are used to understand the different contributions to the detector leakage current: bulk generation, Si/SiO_2 surface generation and diffusion. The different contributions are better

disentangled by using a simpler device as a GR-diode with different geometries, e.g. circular with annular GR's for better understanding the bulk generation, comb-like to enhance the surface contribution. The polarization of the external GR's in both diodes geometries will shield from the lateral diffusion current contribution. The different current density contributions to the total leakage are described in Eq. (3.44), Eq. (3.46). In Fig. 5.4 are shown typical results for CMS-process diodes, where the behaviour is characterized, as expected, by a squared root dependence from the depletion voltage.

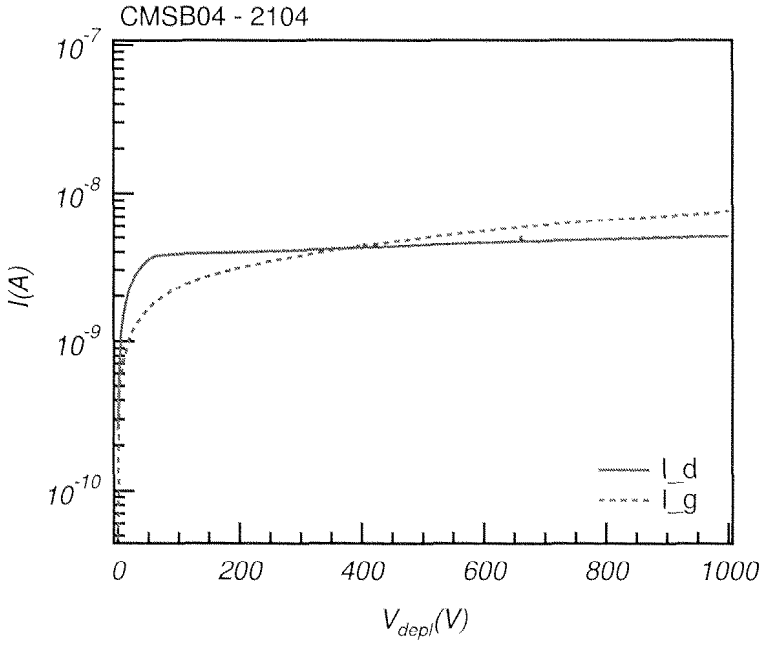


Fig. 5.4: Typical example of I-V curve of a GR-diode. Continuous line is the diode, dotted line the GR.

From the saturation current in Fig. 5.4 we can extract the generation lifetime τ_g (cf. Eq. (3.24)),which in this example is 641 μsec, quite short, highlighting a large density of RG-traps in the substrate.

In this example it is also possible to appreciate the stability of the device up to 1000 V, as compared with Fig. 5.5 where the GR runs into breakdown at 500 V.

At the end of Sec. 3.3.2 the idea of plotting the reverse current as a function of the depletion depth, in order to understand the relative weight of generation and diffusion components, is introduced. In Fig. 5.6 an example of this plot is shown.

In the insert of Fig. 5.6 we can appreciate the extraction of the diffusion and of the generation component: from Eq. (3.48) the generation component is proportional to W_d, whereas the diffusion component is independent of W_d. Consequently, in the I-W plot, I_{diff} corresponds to the extrapolated value of I at $W_d = 0$, and I_{gen} corresponds to the rate at which I increases with W_d, when a reverse bias is applied.

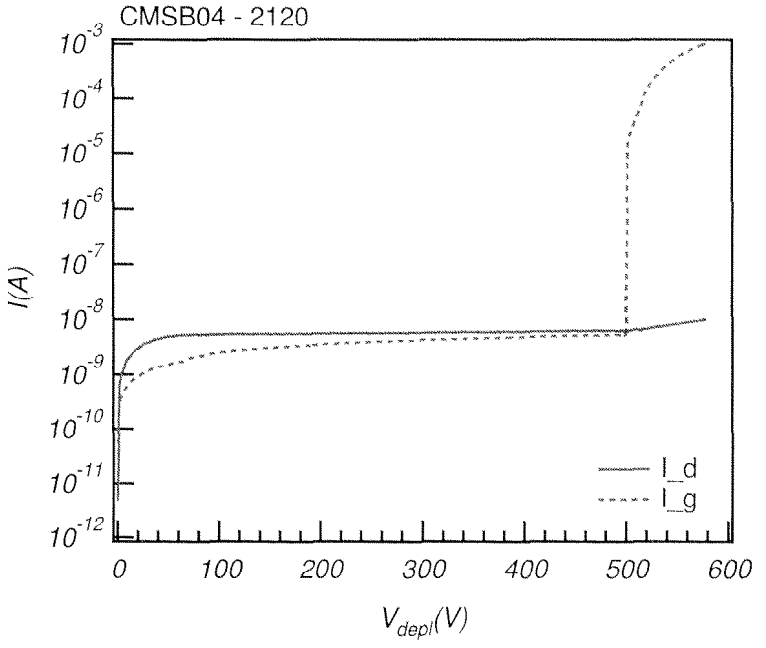


Fig. 5.5: I-V curve of a GR-diode. Continuous line is the diode, dotted line the GR. Breakdown of the GR is evident at around 500V.

In this particular case $I_{diff} = 2.918$nA comparable with the contribution of generation extracted from Fig. 5.5, consistent with the large generation and recombination expected for this bulk.

5.1.2.b MOS capacitor

This simple device allows the extraction of several fundamental parameters for the understanding of the fabrication process and the subsequent detector performance. The first measurement is a C-V plot at high frequency (typically in the range of 100kHz-1MHz) sweeping the gate bias from accumulation to inversion (see Fig. 5.7).

This measurement shows how the interface mobile charge beneath the gate react to an applied bias. For details about oxide charges, see ref. [69].

It is important to note that the use of high resistivity silicon will limit the maximum value for the measuring frequency: in order to understand this phenomenon one should measure C_s and R_s (series capacitance and series resistance) for a MOS capacitor in accumulation conditions (see later for definition). In Ref. [95] we found that thick oxides are necessary for interface charge determination since contact resistance effects may affect the precision in this measure for thinner oxides (< 200nm). Therefore thicker oxide MOSC have been designed for integration together with the CMS milestone detectors.

Three regions are seen: A) *accumulation* ($V_G > 0$): free electrons are accumulated beneath the gate. The measured capacitance is the oxide capacitance only $C_{acc} = C_{ox} = \frac{\epsilon}{d_{ox}} A$

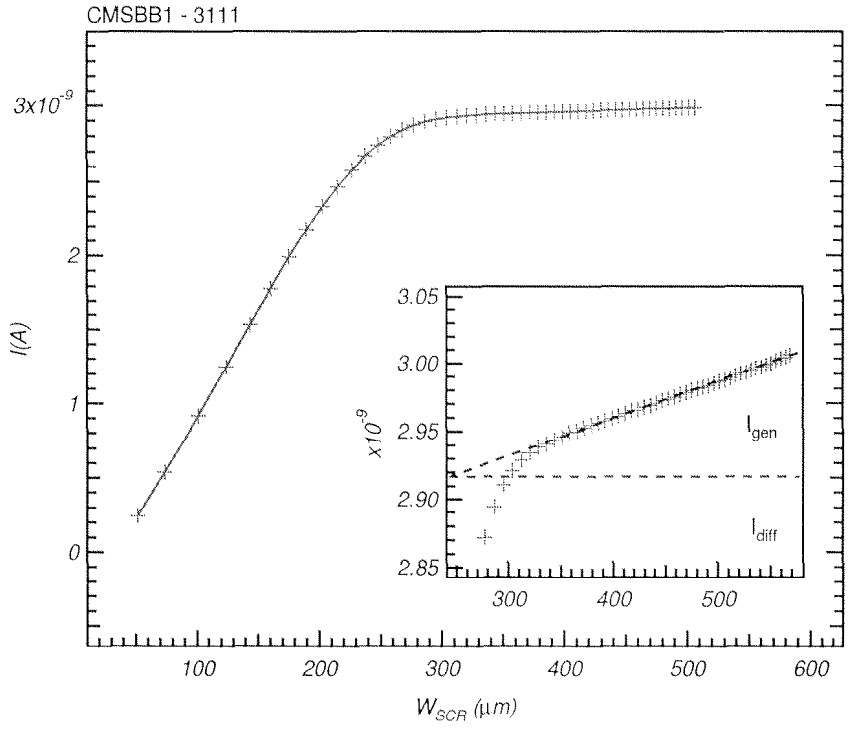


Fig. 5.6: I-W plot for a CMS standard diode (2.5mm^2). On the abscissa is plotted the depletion (and overdepletion) width $W_d = W_{SCR}$ (in microns). In the insert a detail of the central region is shown.

and therefore it is possible to extract the oxide thickness d_{ox}.
B) *depletion* ($V_G = V_{FB}$): rising the voltage in the negative domain the Si/SiO_2 surface progressively depletes of free electrons. The capacitance decreases up to the complete absence of charges on the armature of the capacitor, the *flat-band* condition. In a theoretical curve $V_{FB} = 0$.
C) *inversion* ($V_G < V_{FB}$): holes are accumulated beneath the metal gate and the depleted region extends progressively up to a maximum value, w_{inv}, which depends upon temperature, bulk doping concentration, ac-signal frequency and amplitude.
In general high frequency measurements are easier to make and give

$$C_{inv} = \frac{1}{\frac{1}{C_{ox}} + \frac{w_{inv}}{\epsilon_{Si}\epsilon_0}} A. \tag{5.3}$$

Humidity, chemical pollution, migration through the surface or light particle irradiation may cause an increase of the induced charge at the Si/SiO_2 interface. The consequent effect on the capacitance is clearly seen at the flat-band condition: the shift of the flat-band voltage will measure the trapped charge. Since

$$C_{FB} = \frac{1}{\frac{1}{C_{ox}} + \frac{L_d}{\epsilon_{Si}\epsilon_0}} A \tag{5.4}$$

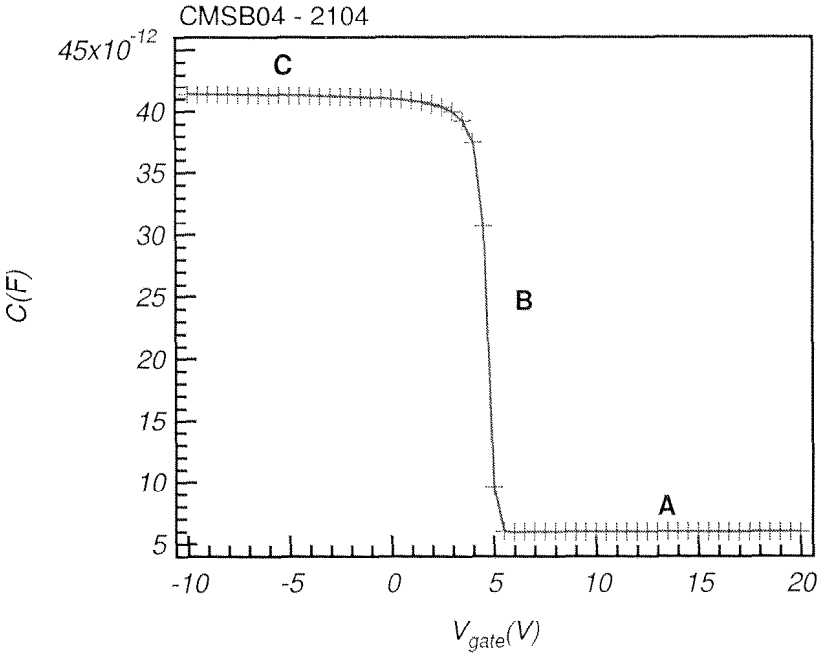


Fig. 5.7: Typical high-voltage C-V curve for a MOS capacitor. Superimposed to the DC gate voltage there is a small-amplitude ($\sim 10 \div 15$ mV) AC signal in order to measure the capacitance.

d_{ox} (nm)	N_d (cm^{-3})	C_{FB} (pF)	V_{FB} (V)	N_{ss}(cm^{-2}))
200	$4.2 \cdot 10^{11}$	36	7.46	$1.9 \cdot 10^{11}$

Table 5.2: Some typical values from MOSC measurement on CMS barrel prototypes.

where L_d is the silicon Debye length[1], therefore we have

$$N_{ss} = \frac{Q_{ss}}{q} = \frac{(\phi_{ms} - \Delta V_{FB})}{q} \cdot C_{ox}. \tag{5.5}$$

where ψ_{ms} is the working function difference between metal and semiconductor.

This information is extremely useful for understanding the interstrip capacitance of the detector (see Sec. 3.3.4.a). After irradiation we expect that the charge trapped into the oxide will grow (see sec. 3.3.5.a) and saturate at a value which is mostly dependent by the crystal orientation [66]. The corresponding interstrip capacitance will be maximized (independently from geometry, see Fig. 3.48) and will set up the maximum in C_{load} at the input of the amplifier, minimizing the S/N ratio.

[1]The Debye length is a measure of the distance over which a charge imbalance is neutralized by majority carriers under steady-state or equilibrium conditions and is defined as $L_d = \sqrt{\frac{kT\epsilon_o\epsilon_{Si}}{q^2 N_d}}$

5.1.2.c Gate Controlled Diode

Strongly linked with the MOS capacitor is the *gate controlled diode* (GCD). In this device we use to modify the charge status beneath a metal gate in order to study the different current contributions collected by the diode: a large gate area is interdigitated with the p^+implantation of the diode in order to get the best coupling between the generation region (beneath the gate) and the collection region (the diode). This design is optimized in order to measure the surface generation rate, s (cf. Eq. (3.25)), which is strongly linked to the density of traps at the Si/SiO_2 interface.

Figure 5.8 explains the link between the GCD and the MOS capacitor: condition A) corresponds to the former condition A) for a MOSC, i.e. accumulation, and similar for conditions B) and C), respectively depletion and inversion, but instead of measuring static charges we are interested in the transients i.e. the currents collected by the diode. The current of the diode is measured keeping a constant depletion voltage to the back, as a function of the gate voltage V_g. In accumulation conditions, (condition A) the mobile electrons induced by the fixed charge in the oxide, are preventing the reverse voltage to deplete completely the volume beneath the gate. The measured current will then be the generation current of the depleted volume below the diodes, J_g. Increasing the negative gate potential, the leakage current increase (condition B) since the depleted volume beneath the gate contributes to the diode current via J_g^{gate} together with the surface generated current J_s. The latter suddenly increases at the flat-band voltage and soon saturates. The next condition, the inversion (condition C), is reached by increasing again the negative gate bias: holes are recalled from the bulk just compensating the J_s current contribution, lowering the total current. The surface generate current is therefore extracted by subtracting the total current measured at inversion from its value at depletion, $I_s = I_B - I_C$.

Since J_s is defined by Eq. (3.43) and Eq. (3.44), the surface recombination velocity can be derived. From Fig. 5.9 we can derive $J_s = I_s/A_g$. Generally speaking, for the

V_{bias}	I_B	I_C	I_s	s(cm/sec)
10 V	915 pA	863 pA	52 pA	2.8

Table 5.3: Some values from GCD measurement on CMS barrel prototypes.

CMS silicon detectors the range of s is between 2 and 20 cm/sec. For this particular case the value for s is very small, causing a very small surface-generated leakage current contribution.

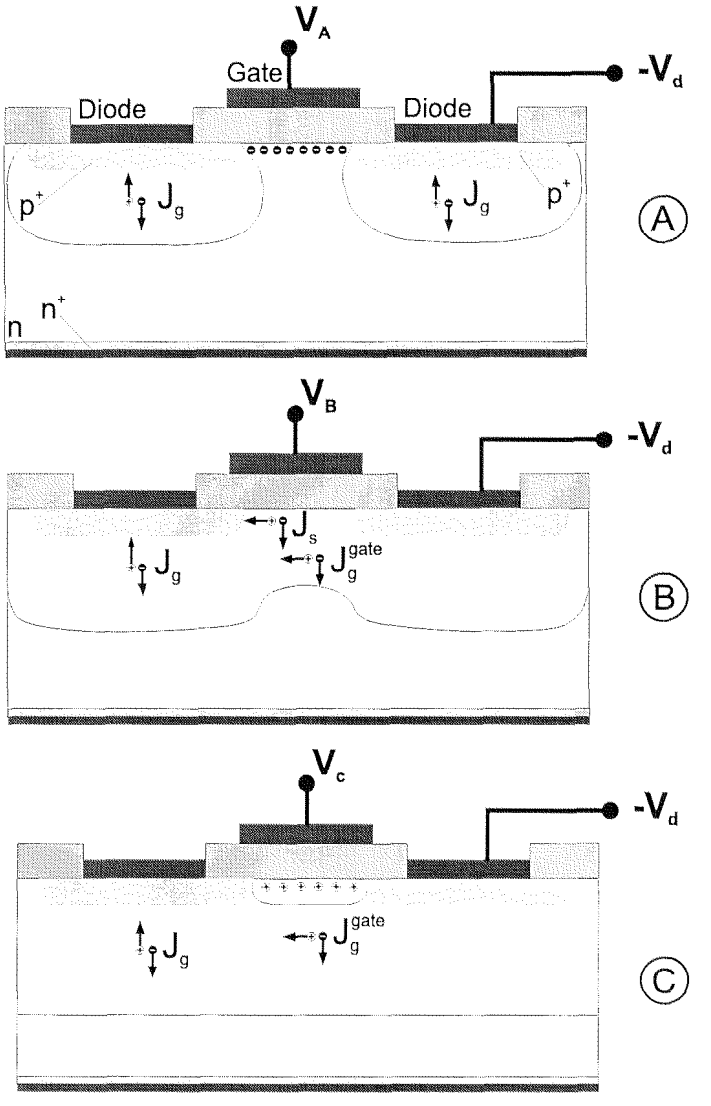


Fig. 5.8: Schematics of the GCD current contributions as a function of the gate polarization.

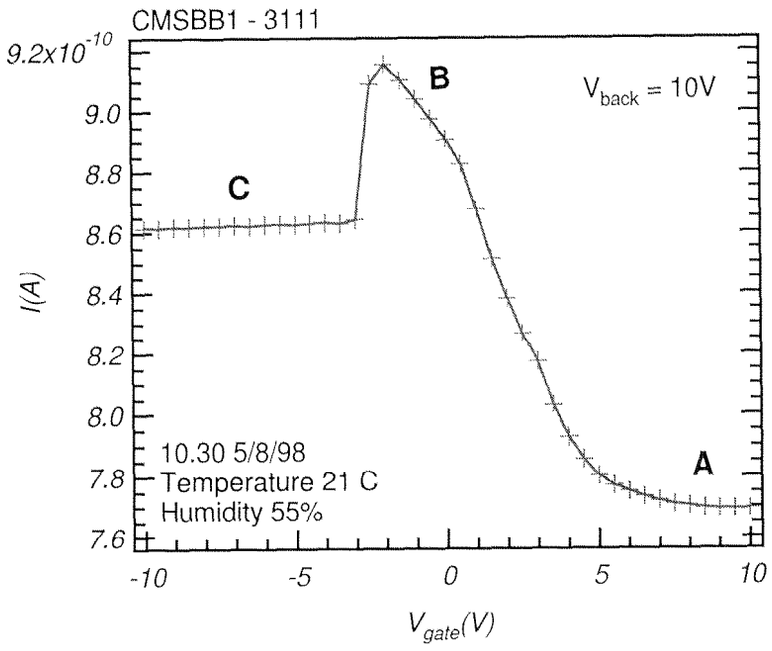


Fig. 5.9: Typical I-V curves for a GCD with large gate surface.Conditions A,B and C are highlighted.

5.2 Results on detector prototypes

In this Section the results of measurements on detector prototypes are reported. These results have been used to validate the design choices and processing technologies for the CMS silicon strip detectors.

5.2.1 Total leakage current and breakdown

We studied the breakdown characteristics of several series of devices which have been measured before and after irradiation. Variations within the same technology are due to the quality of the substrate and the back surface. In this paragraph I described the leakage and breakdown performances obtained with the standard edge configuration: n-well on the cutting area and two guard-rings only. In the following paragraph the improvements achievable with an optimized edge design will be described.

Fig. 5.10 shows the distribution of the breakdown voltage calculated for three different integrations.

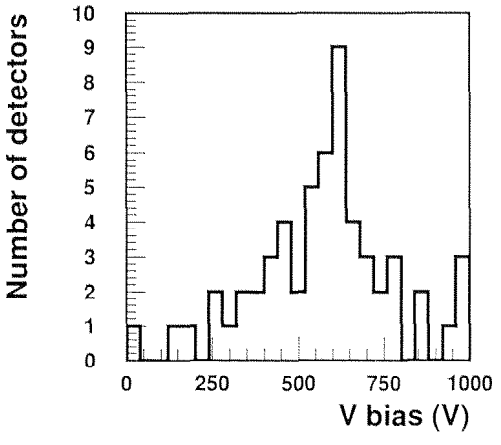


Fig. 5.10: Breakdown voltage values for full size detectors. The average value is at 580 V with a r.m.s. spread of 200 V.

Most detectors fulfill the requirement of break-down voltage $V_{bkd} \geq 500$ V (see Fig. 5.10). Some exhibit excellent performance ($V_{bkd} \geq 800$ V) and for a few it was not possible to measure any breakdown up to the limit of our measuring system (1000 V). It is worth noticing that these measurements were performed on full size devices (54×64 mm^2) with 1024 implants on a 50 μm pitch, and the results were stable after cutting.

In Fig. 5.11 I plotted the total and GR current of a CMS detector (milestone geometry) before and after the cut. It is interesting to notice that after cut the breakdown performances are much better. A possible explanation follows from the fact that often current is very much influenced by surface contributions. If the undiced detector is under bias, it collects carriers from the entire wafer surface, since the minority carrier lifetime is very high. Once cut, the active surface corresponds to the detector surface, without

extra contributions; moreover the n-well implantation on the scribe line shortcircuits the backside n^+ implantation with the front through the highly damaged lateral surface, which is injecting electrons into the bulk: as it will be described more deeply later, this fact enclose the field lines within the detector volume, fixing the volume. A carefully study of critical dimensions of n-well implantations and distances from guard-ring is required to avoid softer breakdowns and injections from the highly damaged cutting region.

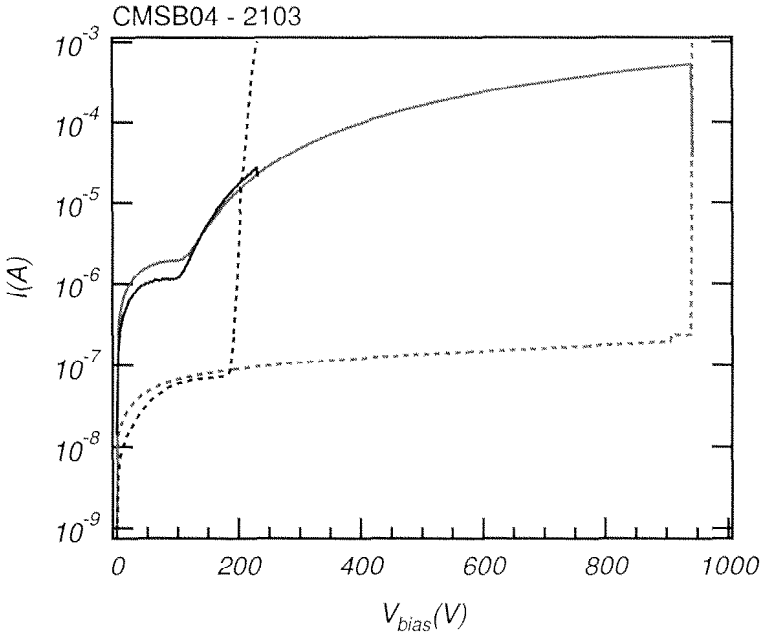


Fig. 5.11: Leakage current of a detector (continuous line) and its guard ring (dotted line) before and after cut. Softer breakdown before cut.

The stability of reverse currents with time has been studied carefully: in Fig. 5.12 I plotted the leakage current of the same detector of Fig. 5.11. After one hour at high bias, the current is stable, no sudden breakdown.

The stability of the breakdown performance under irradiation and different operating conditions has been carefully tested. Fig. 5.13 shows the breakdown performance before and after neutron irradiation at a fluence of 10^{14} $n\ cm^{-2}$ and for different temperatures.

In Fig. 5.14 is shown the breakdown behaviour characteristics after different neutron irradiations. The corresponding depletion voltage is shown.

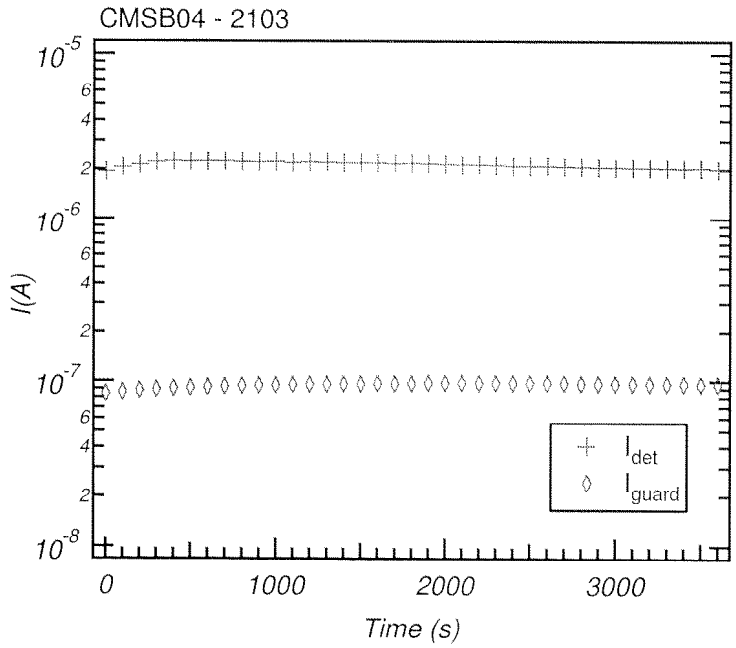


Fig. 5.12: Total leakage current of a detector and its guard ring after cut versus time. The detector is kept at $V_{bias} = 100$V.

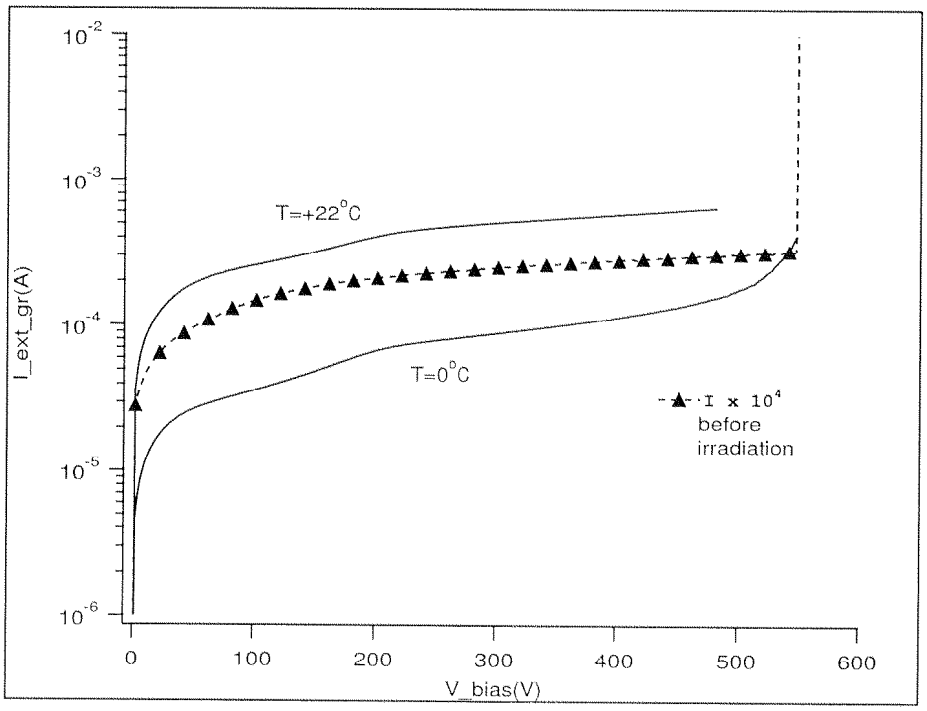


Fig. 5.13: $I-V$-curves showing the breakdown for full-sized devices at different temperatures after irradiation at 0.9×10^{14} n/cm^2

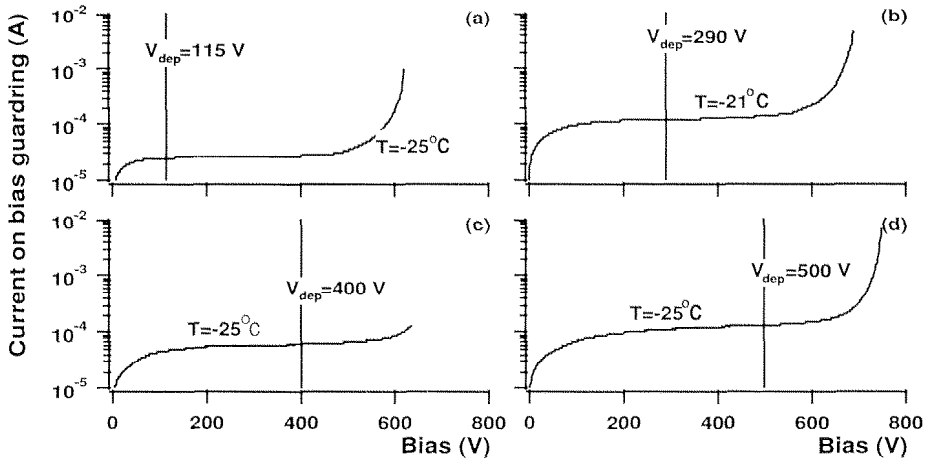


Fig. 5.14: $I-V$-curves for neutron irradiated detectors. a) at $0.9 \times 10^{14} n/cm^2$, b) at $1.8 \times 10^{14} n/cm^2$, c) at $2.7 \times 10^{14} n/cm^2$ and d) at $3.6 \times 10^{14} n/cm^2$.

5.2.1.a Multiguard structures

As previously explained in Sec. 3.3.3.a when high reverse bias is applied to both detector active area and guard-ring, intense electric fields may lead to avalanche breakdown on the external side of the guard-ring, enhancing detector noise and eventually causing physical breakdown of the device [96]. The electric field can be influenced by the position of the external n^+ well and can be controlled by a series of multiple floating p^+ rings around the guard-ring: the multi-guard structure (see Fig. 5.15).

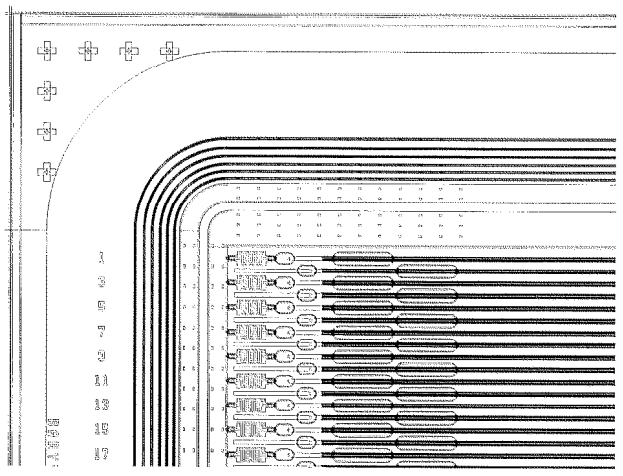


Fig. 5.15: Corner of a typical radiation resistant detector.

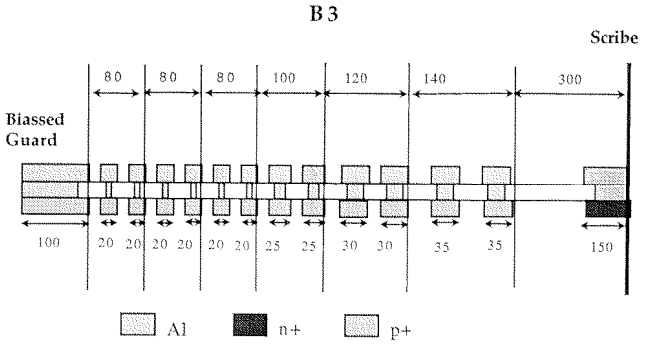


Fig. 5.16: Multi-guard structure layout B3, the biased guard is on the left, while the scribe line is on the right. Distances are expressed in microns.

The position of the n-well with respect to the most external p^+ implant has been thoroughly studied on test structures in which the distance was varied from 25 μm to 250 μm. Details of this study will be presented later (ref. Sec. 5.2.1.b).

The strong efficiency of MGs can be appreciated in Fig. 5.17. Here is plotted the total leakage current of two twins baby detectors ($\sim$5 cm^2) integrated on the same wafer: the first having multiguard structure, the latter two standard GRs. The effect is striking: breakdown voltage is increased systematically of several hundreds volts by using MGs.

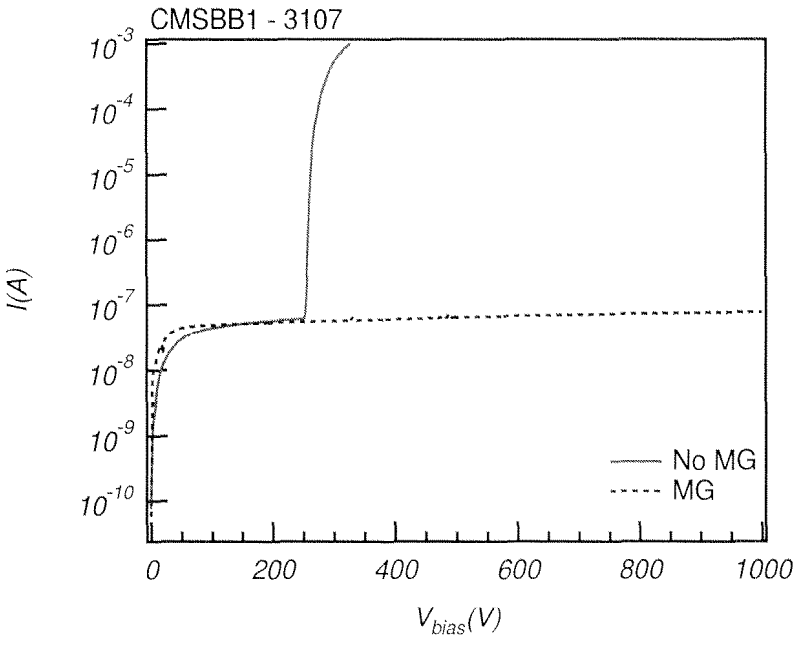


Fig. 5.17: Total leakage current of two baby detectors integrated on the same wafer: normal layout, no MG (continuous line) and Multi guard structure (dotted line).

Multi-guard optimization studies have been performed by means of device characterization, irradiation tests and simulations [97]. For this device, I-V characteristics before and after irradiation are shown in Fig. 5.18. The breakdown, above 1000 V before irradiation, is reduced to 650 V just after irradiation and recovers almost completely 21 days later due to annealing of the surface damage, which is responsible for the breakdown voltage reduction. Further tests with 1 MRad photons under bias show that no breakdown appears up to 1000 V in these devices.

5.2.1.b Detector Edges

Recently CSEM has been involved in the design and the production of several test structures devoted to the study of this phenomena and of the subsequent optimization of the main detector design. The EDGE-1 test structures account for circular p^+- implantations of 1 mm diameter surrounded by n^+- implantations placed at different distances from the diode, resp. $W = 25, 50, 75, 100, 150, 250$ μm; the EDGE-2 have circular diodes with different diameters placed at a fixed distance from the n^+-implantation ($W = 200$ μm) but different distances from the junction to the scribe line, resp. $D = 200, 300, 350, 400, 500$ and 600 μm. For both series of test structures the p^+- implants were covered by a metal layer overlapping them by 5 μm, thus creating a beneficial MOS effect through this field plate.

In Sec. 3.3.3 I studied the theoretical breakdown behaviour of a junction: it is clear that for all layouts the maximum electric field will be situated close to the rounded border of the $p^+ - n$ junction. Deviations to the ideal behaviour should be attributed

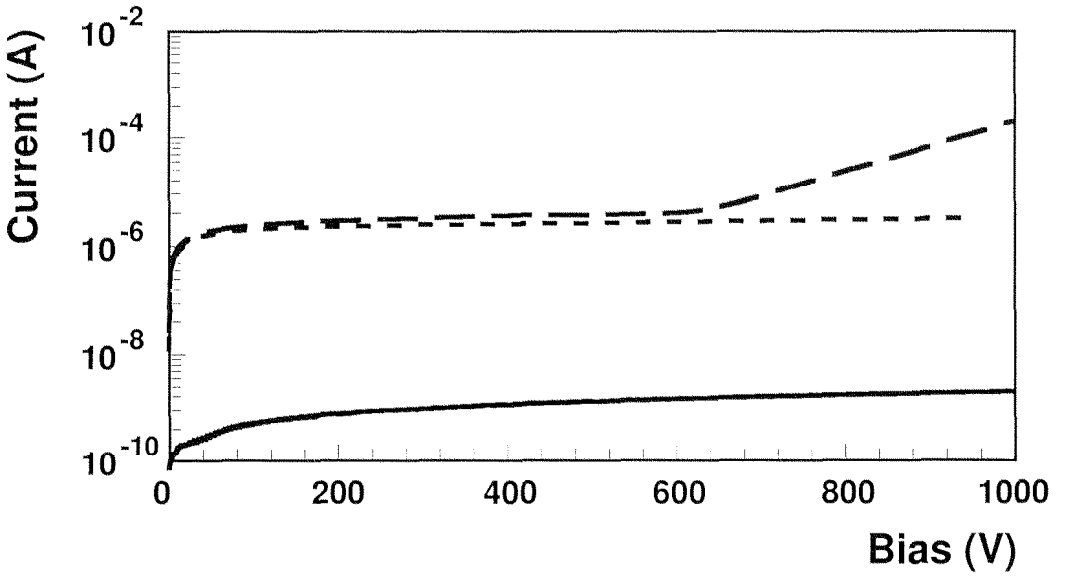


Fig. 5.18: I-V characteristics before irradiation (solid) after $1.8 \times 10^{13} cm^{-2}$ protons (dashed), and 21 days after irradiation (dotted).

to the fabrication technology (i.e. implants depth and lateral diffusion) and also to the presence of the n^+border. The aim of the measurements was to find the dependence of the bias voltage at which the avalanche breakdown process can start, with the geometrical parameters (W and D) described above. All the measurement have been made in Pisa [98].

All layouts have the physical cuts on all four sides (plain lines in Fig. 5.19). To separate the effects given by the two distances W and D, Edge-1 diodes are at large distance (500 μm) from the border. We studied the breakdown performance as a function of W and D taking into account also the variation of the breakdown voltage (V_{bkd}) for different electron accumulation layer concentrations, generated by the positive charge trapped on the SiO_2 during the production process.

Typical I-V characteristics of Edge-1 devices are presented in Fig. 5.20. The breakdown values increase with W and for W greater than 100 μm all the values are similar. The same behaviour can be seen in Fig. 5.21 where the breakdown voltage is plotted as a function of W for two different devices together with simulation results for two different silicon oxide charge values. A clear saturation of the breakdown voltage is seen for W greater then 150 μm, regardless of the silicon-oxide charge concentration. The absolute value of V_{bkd} depends on the oxide charge and is compatible with simulation data within the experimental error [55].

This result suggests that, to improve the breakdown performance, n^+- implant at the cut border should not be closer than $100 - 150\mu$m from the most external p^+- implant. As a consequence, for the final detector design this fact limits the number of p^+- implants external to the active area (guard-rings) by the minimum distance to the sensor border, considering that the electric field in the guard-ring region is strongly influenced by the

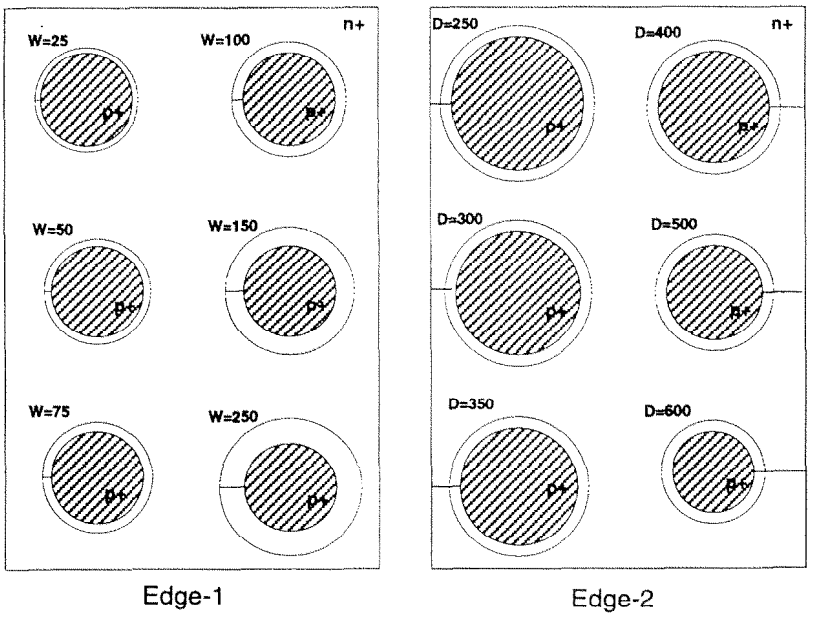


Fig. 5.19: Layout of Edge-1 and Edge-2 structures.

vicinity of the n^+- implant.

The I-V characteristics were measured for Edge-2 diodes before and after cut and are represented in Fig. 5.22. Two effects are visible on these plots:

- after mechanical cut there is a sizable improvement on the breakdown value regardless of the distance D between the diodes and the border. In the region highly damaged by the cut the density of traps is very high. The native oxide on the detector edges will induce a high density of mobile electrons within the bulk, de facto shortcircuiting the n^+- implantations on the front and the back. If D is long enough the depletion region does not reach the cutting region since the potential

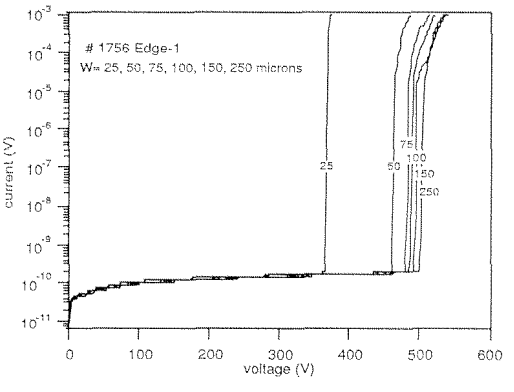


Fig. 5.20: Current voltage characteristics of a typical EDGE-1 device. Values plotted on curves define W.

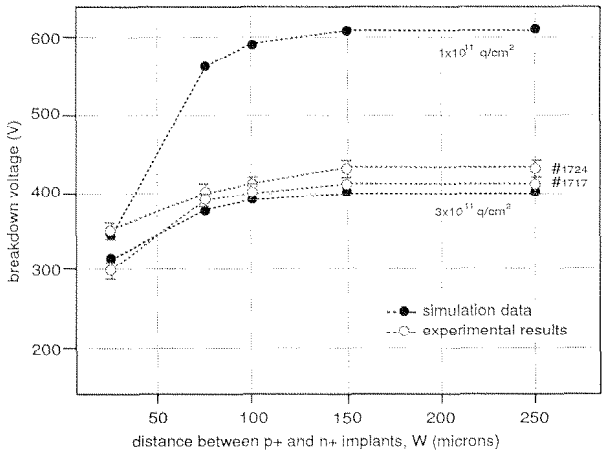


Fig. 5.21: Breakdown voltages as a function of W (open dots). Two simulation results for two different surface charge densities are superimposed (dots).

distribution is influenced by the electrons at the edges, lowering the value of the electric field in the space charge region (SCR);

- for distances D up to around 350 μm an increase of the bulk current as an effect of charge carrier injection from the cut region is seen. For D values greater than 400 μm this effect is not visible anymore. The current increase can be explained by the injection of carriers caused by the punch-through of the SCR with the heavily damaged dicing zone. This is an indication of the minimum distance E that must be left to avoid external current injection.

In Fig. 5.22 the I-V characteristics of all Edge-2 diodes of a wafer are given. It can be inferred that the breakdown is, for small D, influenced by the current rising and has a stable value for D greater than 400 μm. This behaviour is similar for all measured structures.

After irradiation, all structures were kept at -20 °C. All measurements were performed one week after irradiation and no corrections for self annealing were applied. Silicon test structures were exposed to neutron fluences from $1 \cdot 10^{13}$ cm^{-2} to $1.8 \cdot 10^{15}$ cm^{-2}.

After neutron irradiation the leakage current for all diodes increased by around three orders of magnitude (depending on the neutron fluence). As a consequence, the electrical power dissipated increases and the onset of the avalanche breakdown process might be influenced. For this reason, irradiated diodes were measured at room temperature (18-19 °C) and also at -5 °C and -10 °C.

For Edge-1 structures, as a general behaviour, it is observed that after irradiation the dependence of the V_{bkd} on the W values is less visible as the fluence increases. This can be explained as a consequence of the silicon dioxide charge increase during the irradiation (due to the surface damage produced by gamma irradiation background) which was reported by other experiments [99].

For Edge-2 diodes there are two main effects due to irradiation. The border injection

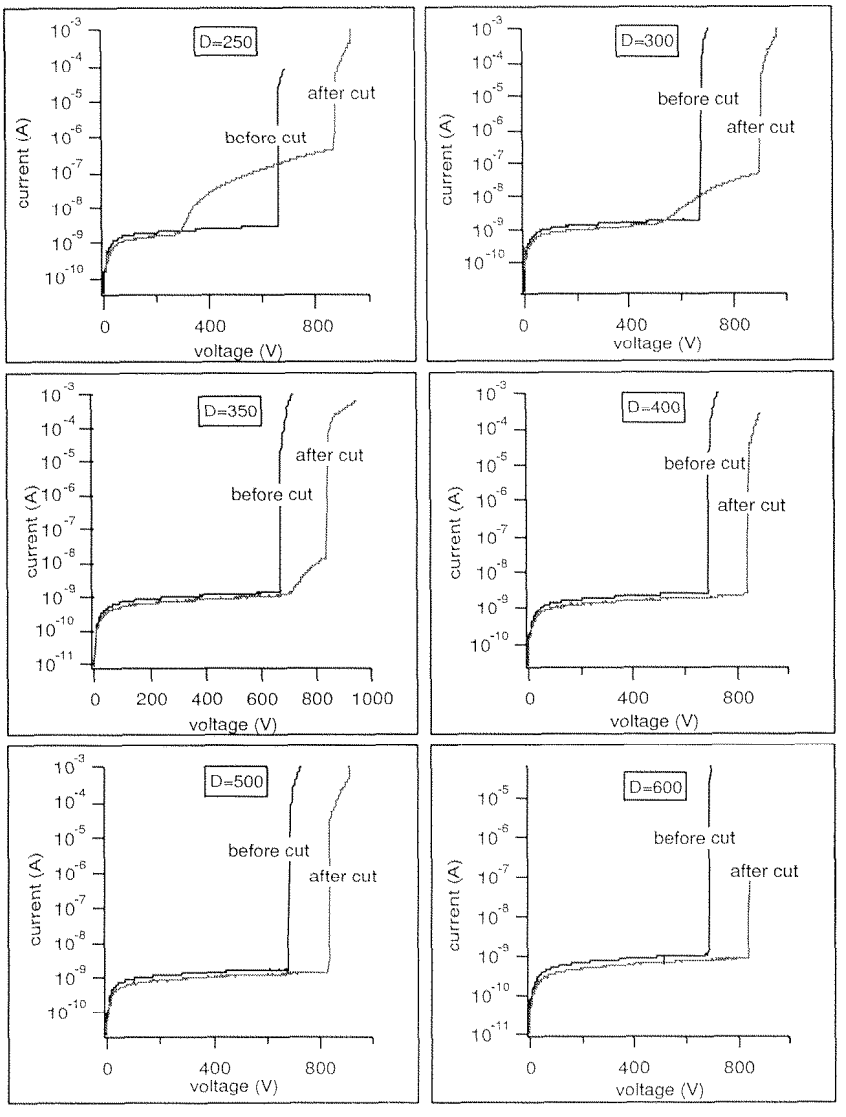


Fig. 5.22: Typical EDGE-2 breakdown performance (here wafer CMSB01 - 1707). Numbers marked on curves defines the distance D (in microns) of the p^+- implantation from the cutting edge.

current is not visible anymore and the dependence on distance D is much more reduced, the breakdown value being independent on this parameter. For high fluence values, a large increase of the bulk leakage current was detected and no avalanche breakdown process was observed (see Fig. 5.23).

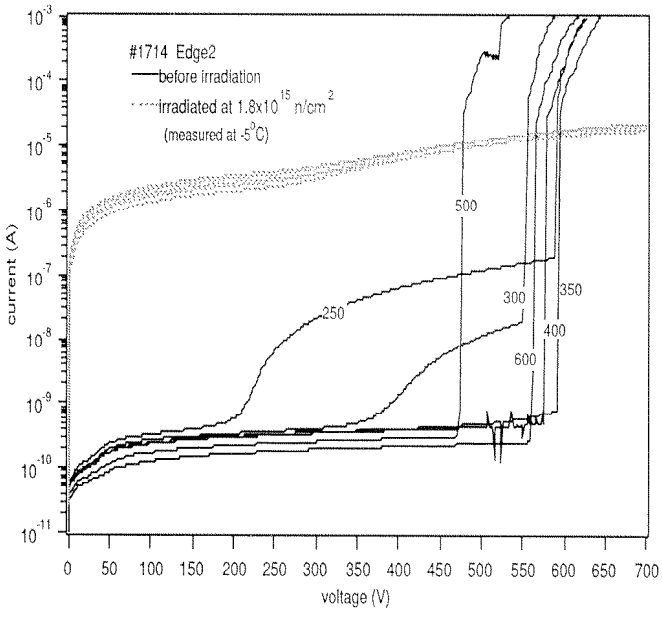


Fig. 5.23: Current-voltage characteristics of EDGE-2 structures before and after irradiation.

Conclusions Experimental data from Edge structures compared with simulation results, before and after irradiation, reveals some constraints to be fulfilled in the design of radiation hard silicon detectors. At the detector border the n^+-implant, necessary to protect the active area from the outside charge carriers currents, should not be closer than $100 \div 150$ μm. All diodes show an improvement of the breakdown voltage after the mechanical cut not depending by the distance of p^+-implant from the edge. The breakdown voltage does not change significantly and the effect of the bulk leakage current increase is not visible anymore for a distance of p^+-implant from the edge exceeding 400 μm.

For fluences up to $1.1 \cdot 10^{14}$ cm^{-2} the dependence of the breakdown voltage on the distance between p^+- and n^+-implants does not change. This indicates that the p^+-implants side remains the critical region for the electric field. Nevertheless the sensitivity on W and D is reduced as the fluence increases. For higher fluences, up to $1.8 \cdot 10^{15}$ cm^{-2}, the dependence of V_{bkd} on geometrical parameters disappears.

5.2.2 Poly-silicon bias resistors

CSEM has produced several series of prototypes in order to check the reproducibility of the poly-silicon bias resistors and their radiation resistance.

In general the resistivity of the poly-silicon layer depends mainly on the implant dose which is used to tune the conduction properties of the layer. Because of this strong dependence, a careful setup of this phase of the process is required.

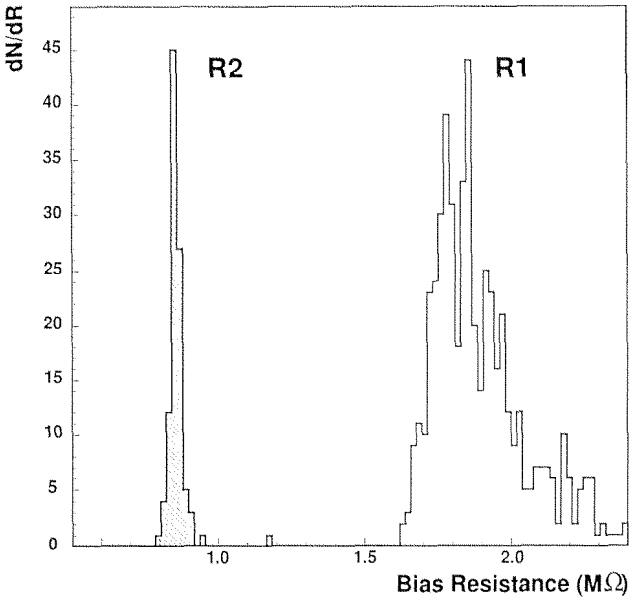


Fig. 5.24: A distribution of resistance values for different doping values. The mean value of R2 is $(860 \pm 20) k\Omega$. R1 has a mean value of $(1800 \pm 150 k\Omega)$.

Within a given technology several parameters may seriously affect the final results: the quality of poly-silicon (i.e. the grain size), the thermal cycle of the process, the thickness and uniformity of the deposition and the presence of a passivation layer. Fig. 5.24 shows how the obtainable uniformity is higher for lower resistivity.

Differences at the level of 5 % related to radial non-uniformities or up-down asymmetries in the deposition phase or in the lithographic steps have also been measured within the same wafer.

By using a large number of squares (50-60) and a standard poly-silicon thickness of 300 nm it is possible to obtain an uniformity better than 10 % for resistivity in the range of 10-20 k$\Omega/\Box$. The requirement of an average value of resistance of (2.0 $\pm$ 0.6) MΩ can be fulfilled.

The stability with radiation of the poly-silicon resistors has been directly tested in CMS prototypes up to fluences of 3.6 $\times$ 10^{14} neutrons/cm^2 and we observed variations within 10 %.

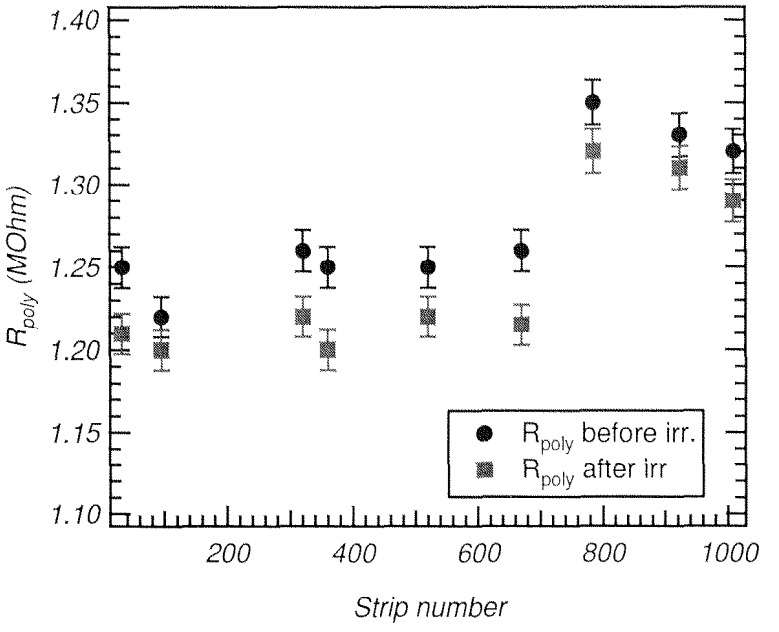


Fig. 5.25: Effect of polysilicon resistors before and after irradiation with a fluence of 3.6×10^{14} neutrons/cm^2.

5.2.3 Integrated decoupling capacitors

Several options for the multi-layer of dielectrics used for the coupling capacitors have been investigated: two thin layers of SiO_2, different combinations of SiO_2 and Si_3N_4 (oxide+nitride -ON-, oxide+nitride+oxide -ONO-), different processes (thermal growth or low temperature deposition), different thicknesses. When nitride was used it has also been tested the effect of patterning the nitride layer by removing it from the inter-strip region compared to the simple case where the dielectric layer is left all over the surface.

The use of nitride as a second dielectric is important since the higher dielectric constant increases the coupling of the signal to the read-out electrodes for a given overall thickness. Moreover, the use of multi-dielectric sampling allows to decouple defects presents in one layer from the following, thus increasing significantly the integrated capacitor yield.

The specific capacitance obtained for a given design depends on the control of the thickness and on the uniformity of the insulating layers. On test structures and full size devices I have tested the reproducibility of the coupling capacitors for different technologies. I measured the coupling capacity of the various strips within the same wafer and observed uniformity within a few percent. In general the thickness of a thermally grown oxide layer is much better controlled (thickness variations $\leq$ 5 %) with respect to the oxide or nitride layers deposition ($\leq$ 10 %).

Good control of the process is needed since variations in thickness may appear from batch to batch within the same process, from wafer to wafer within the same batch and also from strip to strip within the same wafer.

5.2.3.a Yield analysis

A good yield in coupling capacitors is one of the most challenging production goals.

Since the thin dielectric layers of the coupling capacitors cover a large fraction of the detector area, they may be an important source of defects. These capacitors are large elements of typical area around $0.7 \div 0.9\,\mathrm{mm}^2$ (i.e. the dimensions of the capacitors investigated for this work were $14 \cdot 62{,}500\,\mu\mathrm{m}^2$). The fabrication procedure should be virtually defect-free for wafer size devices. When a capacitor happens to be shorted, the corresponding strip is not properly read-out but still the adjacent strips may efficiently collect the charge. When a series of adjacent strips are shorted, the corresponding region becomes inefficient. Therefore an excessive number of defective capacitors can be the main limiting factor for the overall yield in manufacturing large area microstrip detectors.

This study have been triggered by a very bad capacitor yield of detector batch produced in 1996 at CSEM [100]. I participated to the redefinition of the production process and to the electrical characterization in Pisa before and after the technological modifications.

The origin of these defects may be various. The status of the substrate surface may play a role, as may any defect in the growing or in the deposition of the insulating layers. Any lithographic step involving the patterning of the strip may also affect these capacitors because of defective resist depositions or mechanical scratches. The use of testing procedures to identify the number and the position of defective capacitors is the most critical step in the general testing activity of complex devices. The standard techniques identify defective capacitors through a scan of all strips. The most common protocol foresees biasing the back of the detector with a constant voltage, grounding guard and bias-ring on the front, contacting the bias pad, which is in contact with the implant, and the corresponding bonding pad for each strip, and applying a modest potential difference across each capacitor to check the isolation. By measuring the current flowing through the circuit one can easily identify shorted capacitors. A typical result of this scanning procedure is shown in Fig. 5.26. For good capacitors one measures the zero current level of the instrument with the typical measurement fluctuations (in this case $1 \div 10$ pA). Defective capacitors reach the compliance value, set for this measurement at 1mA. Partially defective capacitors can also be identified and monitored.

Early application of the standard diagnostic techniques is very useful for a safe operation of the devices: if the number of defective capacitors is within the required specifications and the wafer is accepted for the assembly phase the information on the number and position of the defective strips is used to skip the shorted capacitors in the micro bonding phase and this information is added to the general database of the experiment as a possible source of local inefficiency. When one knows which capacitor is broken it becomes possible to perform further optical and electrical investigations to identify the origin of the defects. This procedure may be successful when the defects are due to macroscopic problems like mechanical scratches or serious lithographic mistakes. However, when the dimensions of a defect are small, $1 \div 2\,\mu$m or less, the standard optical inspection cannot be used anymore because it becomes practically impossible to localize such pin-hole defects, by scanning the full area of the defective capacitors.

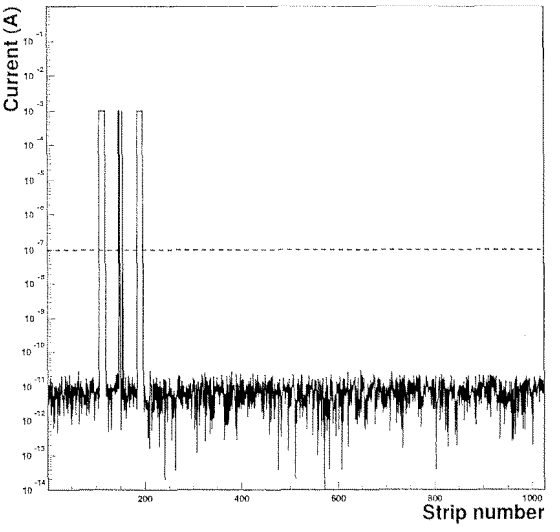


Fig. 5.26: Typical capacitor scan of a good device.

Defect characterization The standard diagnostic techniques give us only a projection of the position of the defect in the wafer while a more precise bidimensional identification of the defect is needed if one foresees the use of more sophisticated techniques (like SEM or TEM) to identify the origin of the problem. The need for more advanced tools was originated by the test of a particularly bad batch of detectors which was apparently produced following the same fabrication protocol, as other successful ones, but the capacitor yield for which was really marginal ($< 20\%$). Fig. 5.27 shows a typical capacitor scan for one of these wafers. In this particular case, the percentage of defective capacitors was higher than 80%. No correlation was found between defects and the results of the optical inspection. Therefore we investigated the possibility to develop further measurements to reach a better bidimensional defect localization. This defective batch will be used in the following as an example to describe how the combined use of conventional and new investigation techniques may lead to the identification of the processing steps responsible for the shorted capacitors. In this particular case the coupling capacitors were produced with a multi-layer of thermal oxide and CVD nitride while metal and poly electrodes were used as upper armature.

The basic idea is to exploit the different resistivity of the metal layer with respect to the implant layer. If a short across a capacitor appears in a given position along the strip one should be able to identify the position of the short by comparing the total resistance between bias and bonding pads measured at both ends of the strips.

Defects closer to one end of the strips should exhibit the largest difference in total resistance while defects exactly in the centre of the capacitor should give the same resistance within errors. The method is obviously based on a precise determination of the resistivity of the metal and implant layer in the actual devices. In our case this information is obtained using standard Van der Pauw test structure.

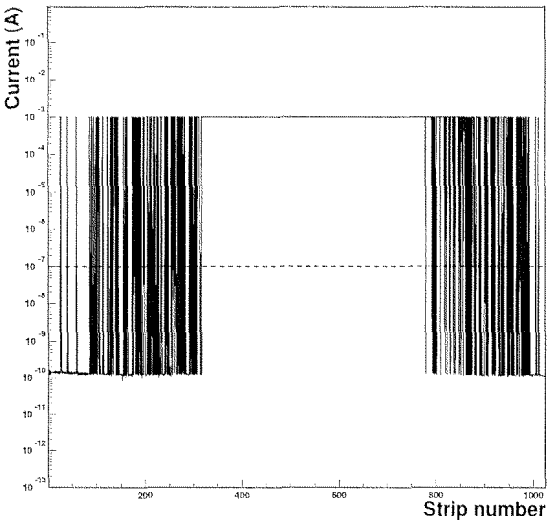


Fig. 5.27: Typical capacitor scan of a bad wafer.

The method is better described by the simple model shown in Fig. 5.28.

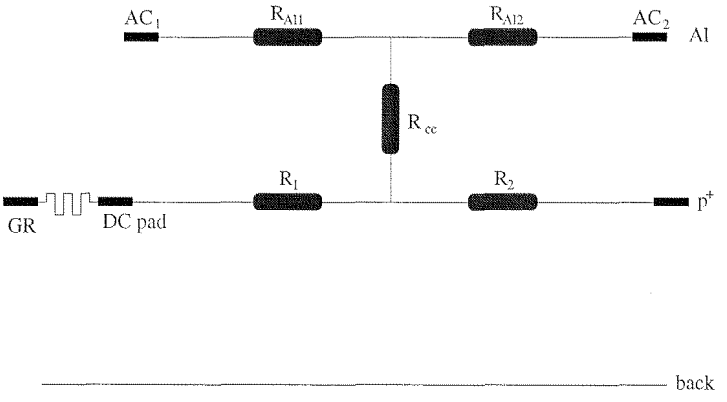


Fig. 5.28: Simplified model for a shorted capacitor.

If we describe the short circuit as a low value resistor, R_{cc}, connecting the aluminum electrode with the p^+ implanted strip, we can use the following set of simple equations to

identify the position of the defect along the strip:

$$I_1 = \frac{1}{R_{Al1} + R_{Al2} + R_{cc}}$$

$$I_2 = \frac{1}{\frac{(R_{Al1}+R_1+R_{cc})\cdot(R_{Al2}+R_2+R_{cc})}{(R_{Al1}+R_1+R_{cc}+R_{Al2}+R_2+R_{cc})}}$$

$$I_3 = \frac{1}{(R_1 + R_2)}$$

$$R_{Al1} + R_{Al2} = \frac{L}{s} \cdot \rho_{Al}$$

$$R_{Al1} = \rho_{Al} \cdot \frac{\Delta x}{s}$$

$$R_{Al2} = \rho_{Al} \cdot \frac{L - \Delta x}{s}$$

where R and R_{Al} are the resistance values for the implanted strip and the metal electrodes respectively, L the strip length, S its cross section and Δx the distance of the short circuit from the DC pad or bias pad. The method is based on the comparison between the measurement of the total resistance along the path AC_1 - R_{cc} - DC pad and AC_2 - R_{cc} - DC pad. Taking into account the measurement precision of the different parameters, the method allows a reconstruction of the position along the strip, Δx, with an accuracy of about 5% of the strip length L. In the standard devices used for this study the error was approximately 3 mm. By scanning each strip and plotting the position Δx of the defect versus the strip number one obtains a bidimensional reconstruction of the position of all defects in each wafer. This information is very useful to identify different classes of defects. Scratches for example may be easily recognized and the optical inspection may be focused on the critical region. Fig. 5.29 shows an example of a mechanical scratch as seen by the new defect localization technique and the corresponding outcome of the optical inspection.

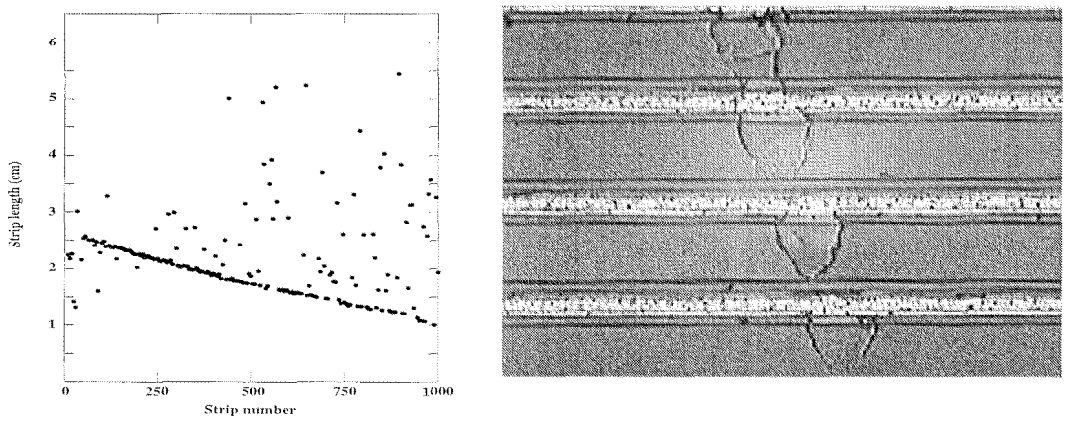


Fig. 5.29: Bidimensional plot of the position of defects in a wafer (left); Enlarged view of a few strips in the defective region (right).

The method is very powerful in identifying other possible sources of defects. Fig. 5.30

shows the outcome of the test on the same wafer with more than 80% of defective capacitors already described in Fig. 5.27. The distribution of defects exhibits an elliptical pattern with a local density of defects in the wafer increasing toward the centre of the wafer. Similar circular or elliptical shapes were observed in all defective wafers of the same batch suggesting as possible explanations a serious non uniformity in the deposition of the nitride layer or loss of control in the etching procedure of the polysilicon layer.

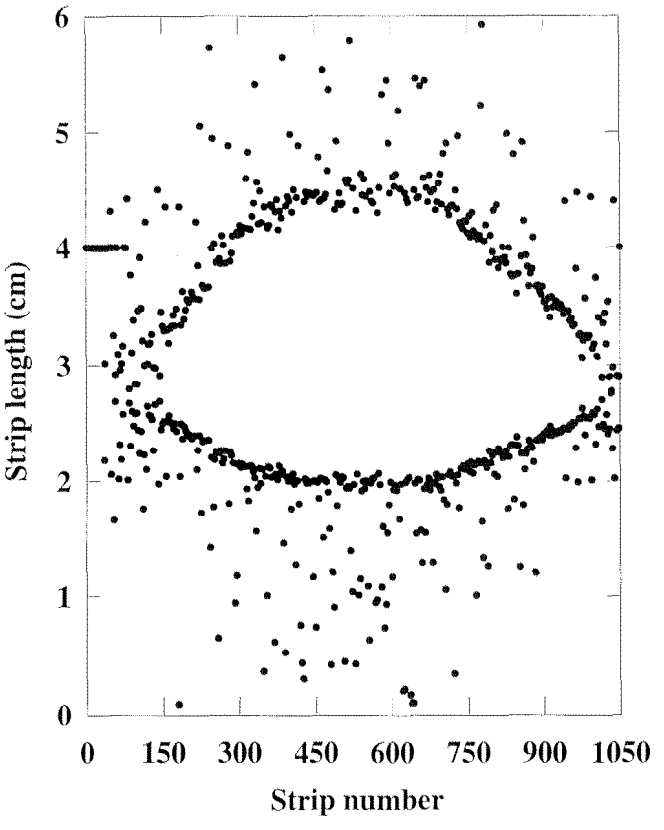


Fig. 5.30: Radially symmetric distribution of damaged capacitors. The wafer is the same of Fig. 5.27.

From the same set of equations one can extract the value of R_{cc} and, with simple assumptions, it becomes possible to evaluate the dimensions of the defects. If the short is due to aluminum we obtain:

$$R_{cc} = \frac{D}{s} \cdot \rho_{Al} \tag{5.6}$$

where D is the nominal dielectric thickness, s the area of the pinhole and ρ_{Al} the measured resistivity of the Al/Si layer. Fig. 5.31 shows the distribution of the defect diameter in the same badly damaged device already described in Figs. 5.27 and 5.30. Typical values are in this case 2-3 μm. The surface of the damaged area increases in going from the wafer edge toward the central region.

To efficiently use more direct investigation techniques like SEM, the 3 mm accuracy so far achieved in the localization of the defects is still marginal. A new method was

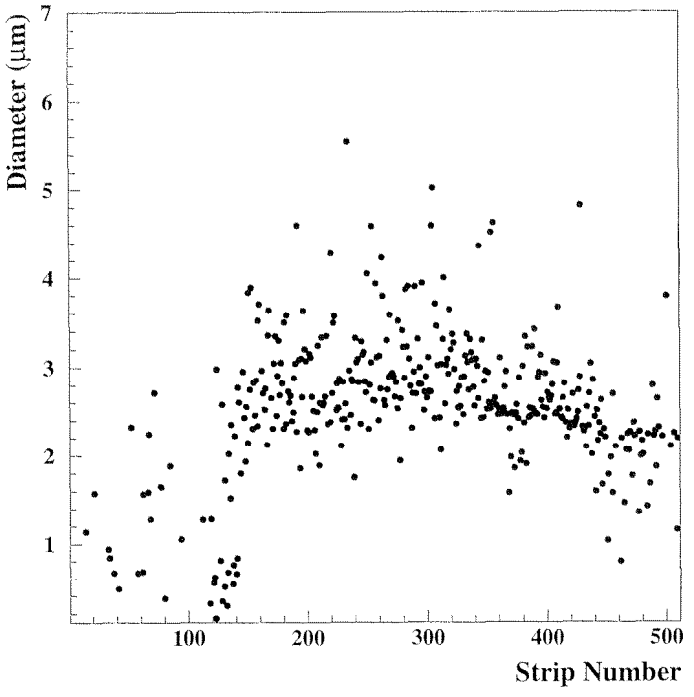


Fig. 5.31: Distribution of the defect diameter along the strips.

therefore developed to increase the resolution. The technique is based on an approach similar to the one so far used for the low resolution localization, the only difference being the trick of repeating the measurements along the metal electrodes. More specifically the back is biased at 1V, guard and bias rings are kept floating and the strip under study is contacted to ground. By measuring the current flowing through the detector and the shorted capacitor and collected by the probe tip it becomes possible to identify the position of the short by simply looking at the point where there is a change of slope. Contacting the strip every 50 μm the accuracy of the measurement becomes very high. In principle one could also make two series of steps, a first one at low resolution (say every millimeter) and a second one at a finer pitch. The main drawbacks of this method are that the procedure is very time consuming and its destructive activity cannot be repeated several times. The reason of this is evident from Fig. 5.32 where the photograph of the strip after the investigation shows clear signs of damage due to the tips.

The change of slope which identifies the position of the short is the point of minimal resistance found by the probe during the scan. Fig. 5.33 shows examples of this kind of measurement. It should be stressed that the technique is very sensitive to noise and several measurements are needed for each point to get rid of fluctuations.

From the raw data shown in Fig. 5.33, after numerical elaboration, it is possible to achieve a precision of about 20 μm.

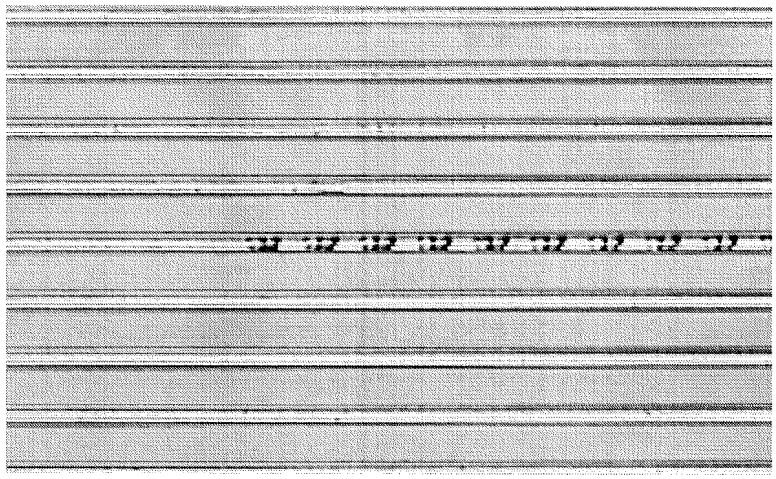


Fig. 5.32: Photograph of the metal strip after the high resolution measurement.

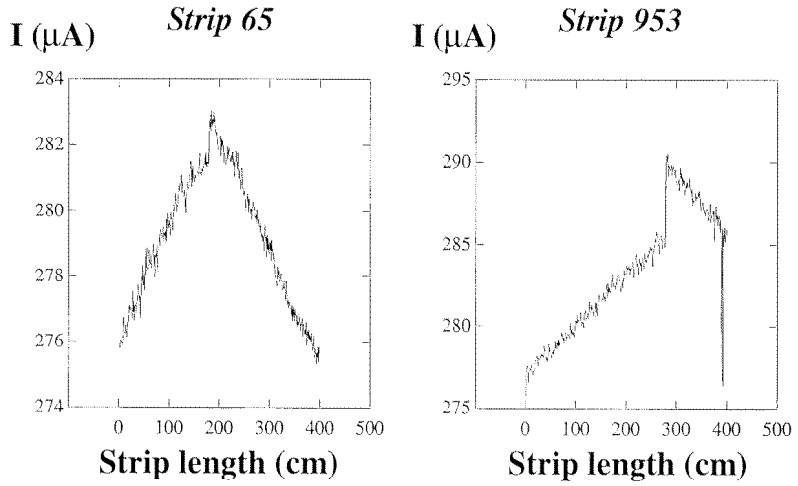


Fig. 5.33: High resolution defect localization.

SEM analysis With a localization of the defects at the level of 20 μm it becomes possible to investigate the shorts through Scanning Electron Microscopy. To this purpose a device was cut in proximity of the measured positions of the shorts and, after careful lapping, the cross section of the detector was examined. The photograph in Fig. 5.34 (left) shows the cross section of one strip in correspondence of the short. The aluminum layer is in contact with the substrate over a width of about 3 μm as was pointed out by the preliminary evaluation of the diameter of the defect. The poly layer is completely etched away in the middle of the strip as the two layers of dielectric (thermal SiO_2 and CVD Si_3N_4 in this case). Tracing back the origin of the problem with a careful review of the flow chart used in this particular production it was discovered that the origin of the anisotropic over-etching (0.5 μm in depth) was a consequence of a mistake in the step of Reactive Ion Etching used for the opening of the metal vias. The effect was more pronounced in the centre of the wafer and this was the origin of the elliptical patterns of

defects. To confirm the analysis, metal and poly layers were etched away, in preparation for the SEM analysis in plane, in samples of detectors belonging to the heavily damaged region in the centre of the wafer. Fig. 5.34 (right) shows the nitride layer of one of these samples completely damaged by the over-etching of the metal vias.

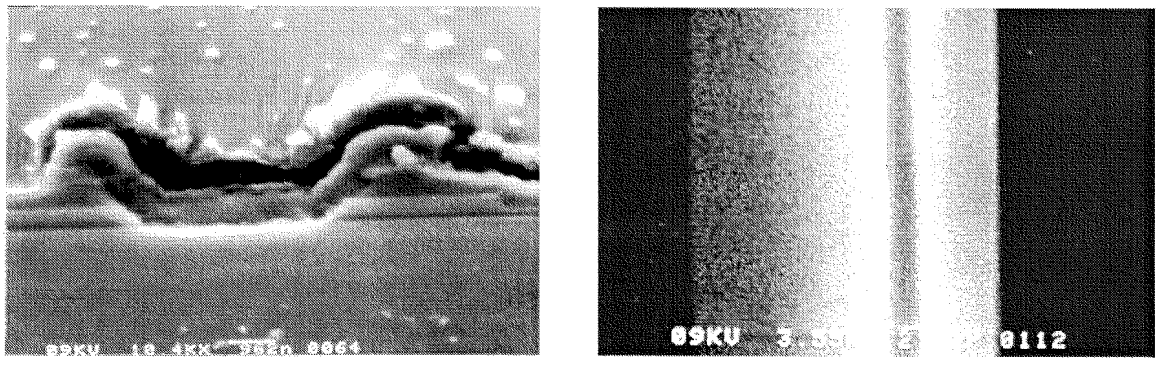


Fig. 5.34: Cross section of a shorted capacitor as seen in SEM (left); Plan view of the nitride layer in the central part of the wafer: notice that it is fully removed (right).

Improvements in the fabrication technology A complete revision of the process of fabrication of AC coupled microstrip devices, including the simplification of a few steps, was triggered by the analysis of these failures. The MAXIS technology described in Sec. 4.10 was born.

As outcome of this activity the yield of good capacitors in the MAXIS technology has significantly improved, exceeding 99%. Fig. 5.35 shows the scan of the coupling capacitors made on one full-size detector featuring 1024 strips integrated with the new technology: no defective capacitors are detected.

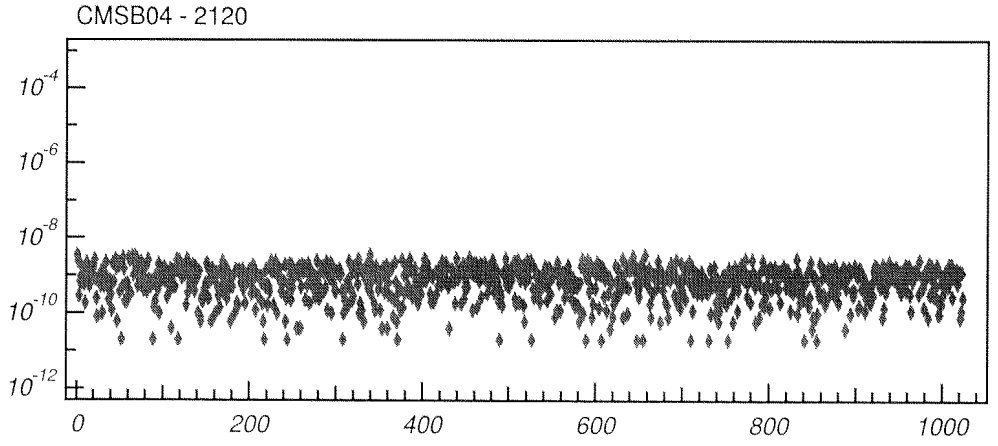


Fig. 5.35: Current flowing through each strip coupling capacitor versus the number of strip. The scan refers on a full size CMSB04 detector (milestone geometry).

Fig. 5.36 shows the comparison between the results obtained in terms of percentage of bad strips before and after the optimization of the process. Most of the detectors

are now with zero or one defective capacitors. The average value of shorts per detectors is 0.34%, well within the required specifications while the problems affecting the worst devices, including the one with 3% bad capacitors, are mainly due to defects in the resist deposition and to mechanical scratches.

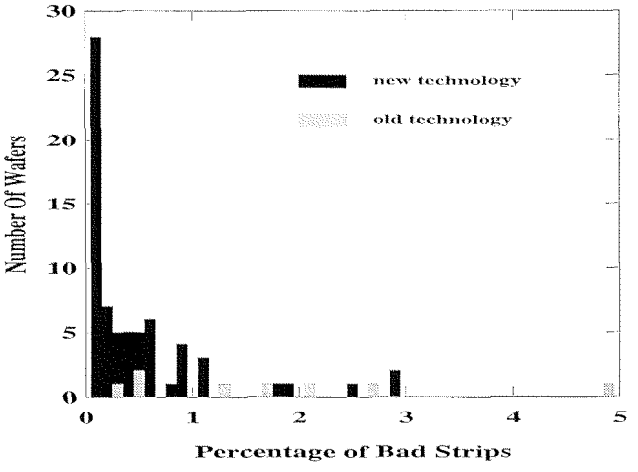


Fig. 5.36: Improvement in the yield of capacitors.

5.2.3.b Integrated capacitor breakdown performances

Although in the final detector modules, for normal running conditions, the potential difference across the coupling capacitors will be negligible, good breakdown characteristics of the insulator layers is required. In fact, if a fraction of the beam is lost in one detector, the accumulated charge may discharge the detector with a time constant depending on the total capacitance of the bias circuit. Under these extreme conditions, a potential difference as high as 100 V may appear across the coupling capacitors. Several series of prototypes for CMS have been tested both at CSEM and at the CMS home laboratories in different conditions: Fig. 5.37 shows the results obtained for the most promising configuration of multi-layer of oxide and nitride. Breakdown performance exceeding 150 V has been obtained after testing several hundreds capacitors under different conditions of electrical stress. The stability of the coupling capacitance with time and irradiation has been verified both in terms of possible changes in the effective coupling due to charge trapped at the interface and in terms of degradation of the breakdown performance after irradiation.

5.2.4 Interstrip capacitance

Since it is the dominant source of electronic noise, the inter-strip capacitance has been carefully measured on test structures and prototype devices. The coupling of each strip

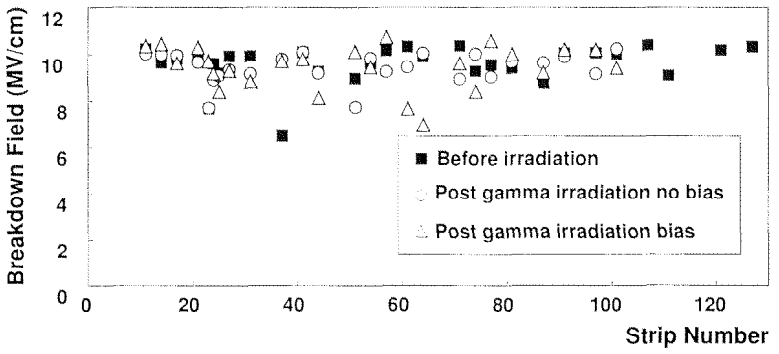


Fig. 5.37: Breakdown field for coupling capacitors before and after irradiation with Co^{60}, 10 Mrad with and without bias during irradiation. The minimum breakdown value of 6 MV/cm is equal to 120 V across the coupling capacitor section.

to the adjacent ones is dominated by the field lines connecting the p^+ implants through the silicon and by the charge at the Si-SiO_2 interface.

Fig. 5.38 a) shows the behaviour of the inter-strip capacitance (one neighbor) as a function of the bias voltage for four different devices. The limiting values (after full depletion) for different wafers depends on the amount of trapped oxide charges, which is consistent with the measurement performed with the GCDs.

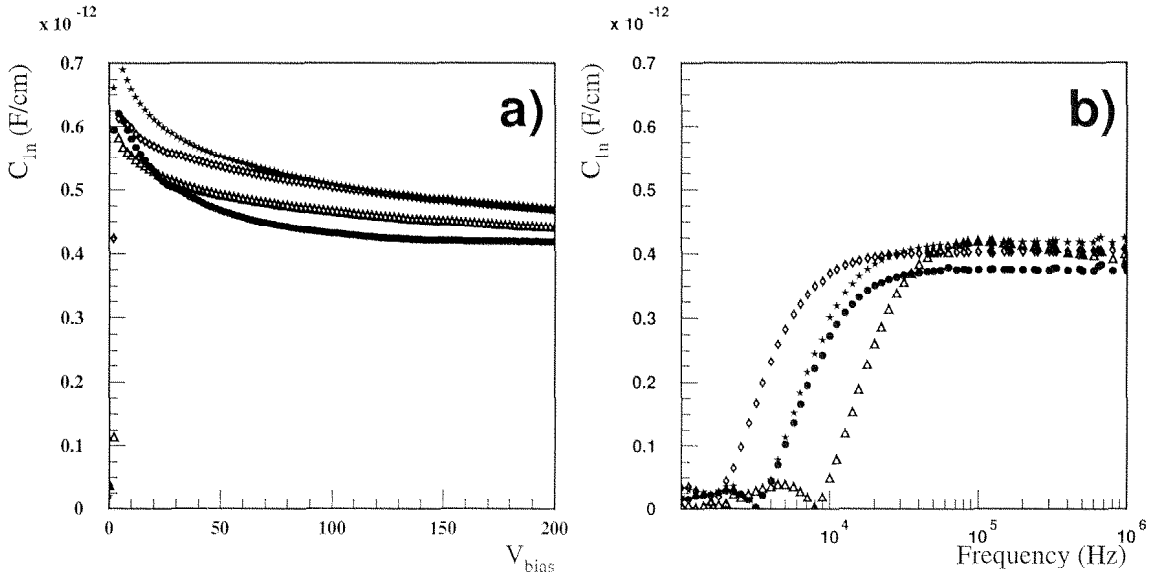


Fig. 5.38: Inter-strip capacitance (pF/cm) before irradiation: a) vs backplane bias voltage (V) for 4 different wafers; b) vs frequency at full depletion bias voltage.

Fig. 5.38 b) shows the behaviour of the inter-strip capacitance as a function of the frequency at full depletion. At high frequency the limiting value shows a plateau at a typical value of 0.4 pF/cm.

The role of various designs and technologies has been investigated before and after

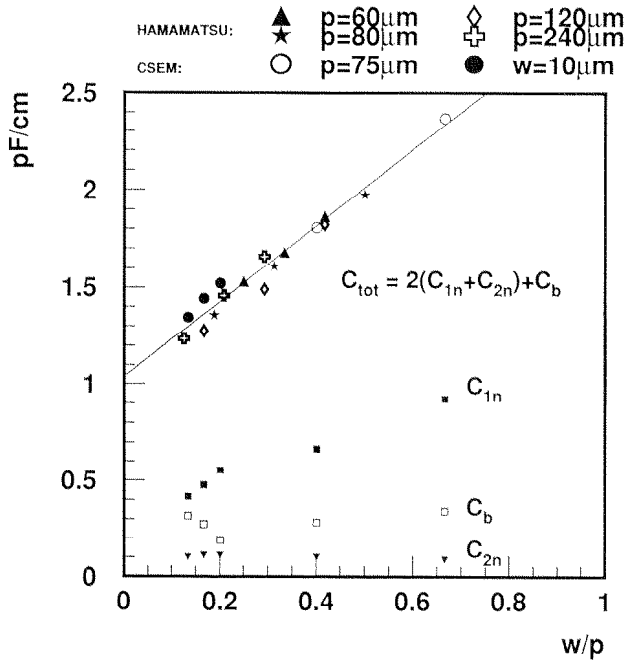


Fig. 5.39: Total capacitance per unit length as a function of strip width/strip pitch ratio (w/p) for detectors from CSEM and Hamamatsu [14].

irradiation with measurements performed at 100 V bias voltage and 1 MHz frequency. Fig. 5.39 shows for a given technology the value of the total capacity coupling

$$C_{tot} = C_b + 2(C_{1n} + C_{2n}) \tag{5.7}$$

which is a function of ratio of the width over pitch of the electrodes (see Sec. 3.3.4.a).

As described in Sec. 3.3.4.a the total capacitive coupling can be approximated with a linear function of the strip over pitch ratio when w/p is around 0.5.

By plotting the total interstrip capacitance per unit length as a function of the width/pitch ratio divided by the gap, a good parameterisation can be obtained for different designs (see Fig 5.40).

A width/pitch ratio of about 0.2 is chosen for an optimized detector design since it allows a low interstrip coupling with a still manageable number of readout channels.

With irradiation the concentration of the positive fixed charge in the oxide will increase to the saturation value. The higher concentration of negative charge at the interface will result in a stronger inter-strip coupling. A reduction of inter-strip capacitance can be obtained in all devices by increasing the bias voltage, since a higher field on the p^+ side confines the electrons at the interface to the middle of the gap between the two implants. Fig. 5.41 shows the inter-strip capacitance versus the bias voltage after different irradiations.

Inter-strip capacitances measured after irradiation show a very strong bias dependence (Fig. 5.41). At full depletion the coupling is still high but it sharply decreases by increasing

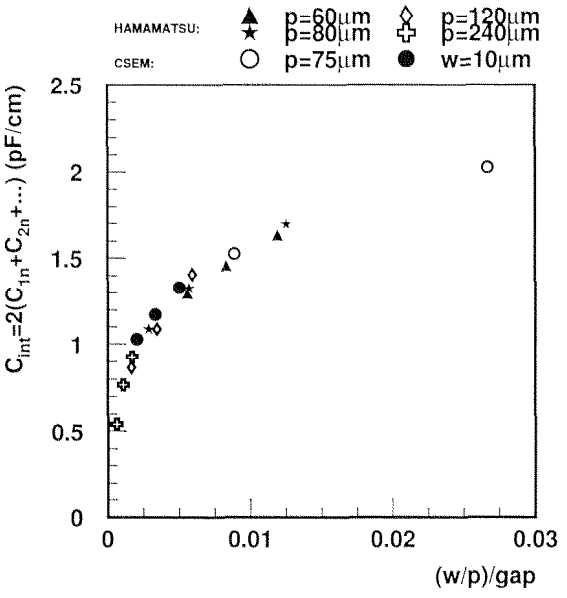


Fig. 5.40: Total interstrip capacitance $C_{int} = 2\sum_i C_{in}$ per unit length, as a function of the w/p over gap ratio. The gap is the difference between strip pitch and width. Same detectors as Fig. 5.39 [14].

the bias voltage. Asymptotic values at most 25 % higher with respect to the pre-irradiation values can be reached for a bias voltage 1.5 higher than V_{depl}. Recent studies show the influence of the substrate crystal orientation on the interstrip capacitance: using ⟨100⟩ substrates seems be more adequate since the interstrip capacitance after irradiation does not increase significantly [78].

5.2.5 Interstrip resistance

All strips, being fabricated as p^+ on n devices, are highly isolated before irradiation. In order to estimate the interstrip resistance I made use of a device simulator by means of DC (steady-state) simulations. Since interstrip resistance depends on the bias voltage, a constraint can be placed on the operating voltage, such that the interstrip resistance is guaranteed to be larger that a specified lower limit. Simulation results are plotted in Fig. 5.42, and show that at any bias voltage exceeding the depletion voltage an interstrip resistance well in excess of 1 GΩ/cm (i.e., a factor 1000 greater than the bias polysilicon resistance) is obtained, well in agreement with measurements.

After type inversion the radiation induced increase of surface charge will affect the interstrip resistance. We have measured lower values for the inter-strip resistance but they are still high enough to guarantee strip isolation and acceptable noise figures (see Fig. 5.43).

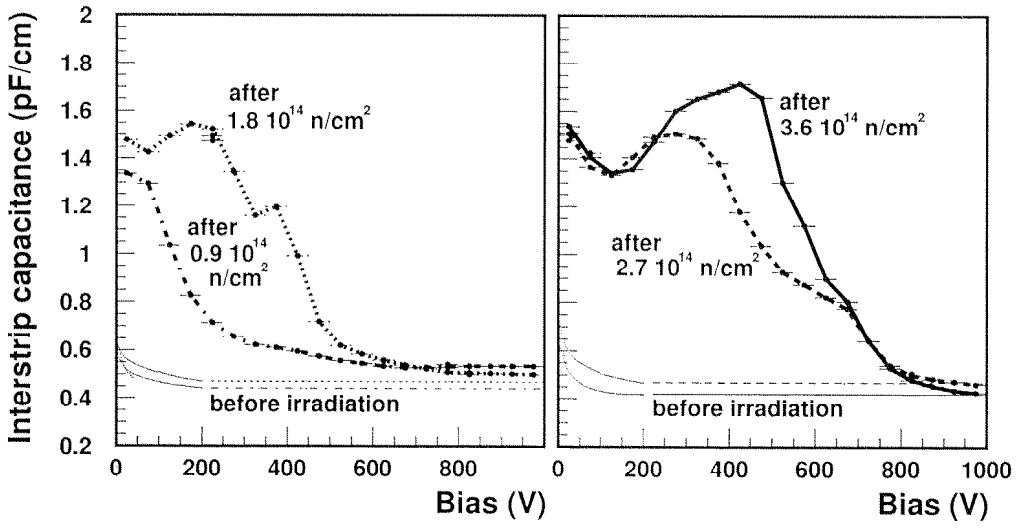


Fig. 5.41: Capacitance between first neighboring strips before and after different neutron irradiations as a function of the bias voltage.

5.2.6 Strip series resistance

The strip series resistance depends on the geometry and the resistivity of the metal line. Using standard aluminum-silicon I have measured resistivity around 27 $m\Omega \cdot \mu$m, which results in a series resistance of the order of 90 Ω for metal electrodes 12 μm wide, 1.5 μm thick and 62.5 mm long.

5.2.7 Double-metal read-out

The use of a double-metal layer to read-out the stereo strips without introducing dead areas has been validated by several series of prototypes. The quality of contacts between the two metal layers may affect the number of dead channels in the system. A series of test chain structures was implemented to find the best solution to this problem. With a contact area of about 16 μm^2 and a 4 μm thick insulator in between the two electrodes a failure rate lower than 10^{-4} was measured. The additional coupling capacitance introduced by the second metal layer has been estimated to add a 10 % contribution to the total input capacitance. Measurements performed in test devices confirm this assumption.

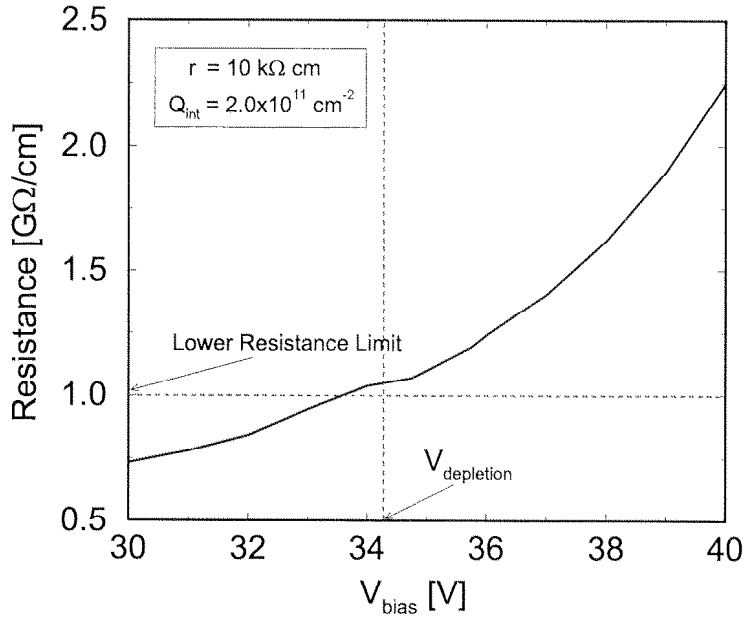


Fig. 5.42: Interstrip resistance as a function of V_{bias}.

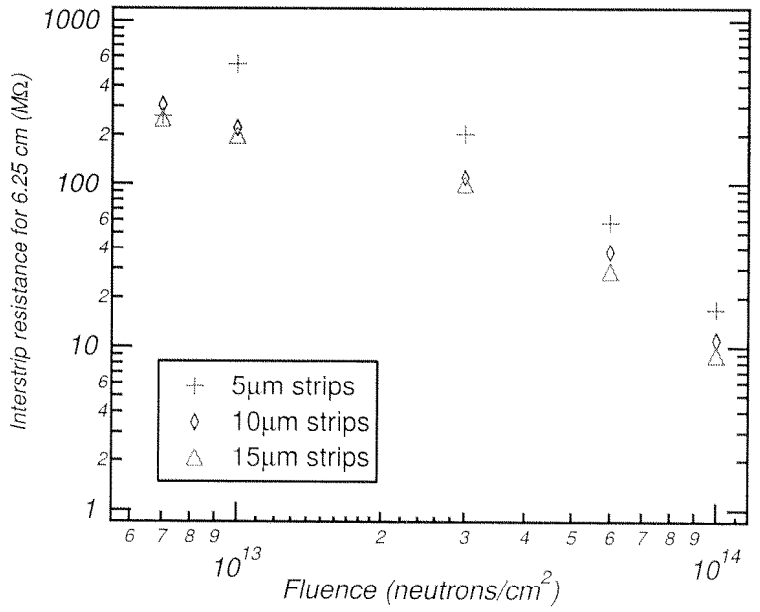


Fig. 5.43: Interstrip isolation before and after type inversion for three strip's implant width [101].

5.3 Test beam results

In the last three years I participated to several beam tests together with the SST group in order to carefully check the performances of barrel and end-cap devices in terms of signal-to-noise ratio, detection efficiency and space resolution under different conditions of bias voltage, temperature, radiation damage and incident angle. The devices used for the test were from different producers, so I will report on the whole work, not restricting myself to CSEM devices results.

Fast analog amplifiers (PREMUX) with peak mode readout and 50 ns shaping time, based on a similar input stage that is used in the final front-end chip, were used as readout electronics. Data were analyzed with a set of algorithms, described in [102].

5.3.1 Signal to noise ratio

Signal-to-noise (S/N) is defined as the ratio of the most probable value of the Landau fit to the signal distribution with the average single strip noise.

Signal-to-noise ratios as high as 25:1 were obtained for non-irradiated full-size detectors (Fig. 5.44).

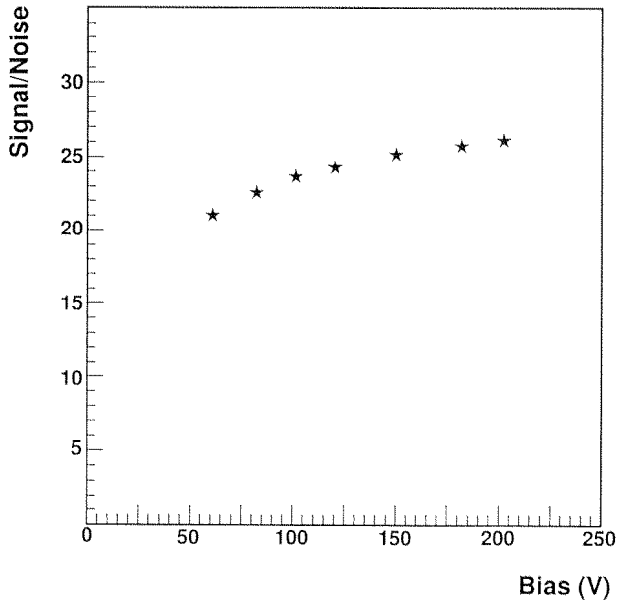


Fig. 5.44: S/N ratio vs. bias voltage for a detector, manufactured by CSEM, with 50μm strip pitch and 12.5 cm length, at 21 °C.

The effect on the S/N ratio due to the geometrical parameters of the sensor design has been studied widely in non-irradiated devices for both barrel and end-cap type detectors.

For the barrel detectors the relevant geometrical parameters are the strip pitch, the implant width and the strip length.

Figs. 5.45 and 5.46 show, as a function of V_{bias}, a similar behaviour in cluster charge and S/N for readout pitch values ranging from 60 to 240 μm and comparable width/pitch ratios (0.21-0.33).

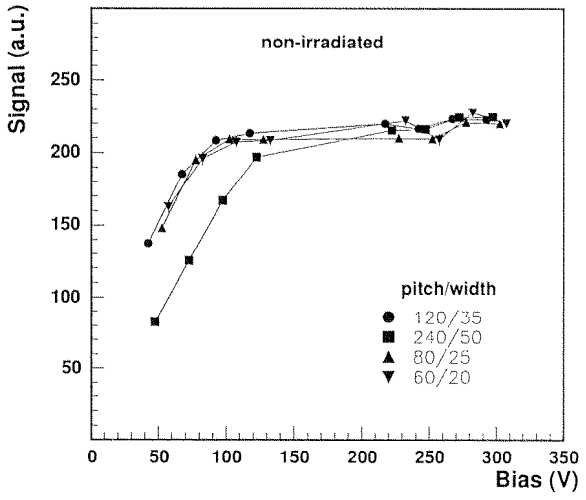


Fig. 5.45: Cluster charge as a function of the bias voltage. Hamamatsu detectors with strip pitches varying from 60 μm to 240 μm.

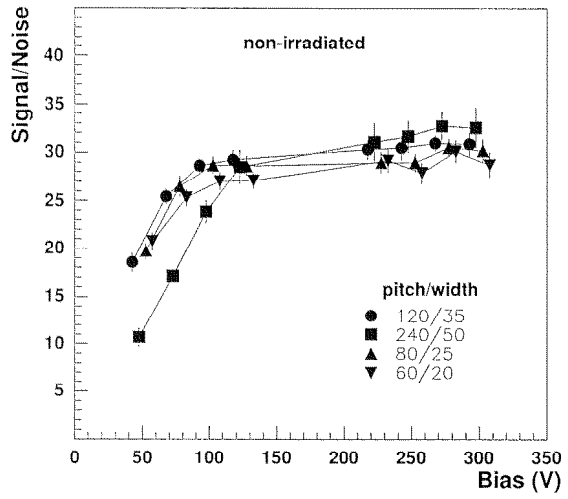


Fig. 5.46: Signal-to-noise ratio vs bias voltage for the same detectors as Fig. 5.45.

In order to optimize the strip design for the wedge detectors, we have carried out a preliminary comparative study on such devices where the strips on the p-side were subdivided into three groups: one based on a constant strip width of 14 μm, the second based on a constant width over pitch ratio of 0.37 and the last featuring a variable strip width from 14 to 26 μm maintaining a constant inter-strip gap of 24 μm along the detector. We tested one of these devices at a bias voltage of 100 V and in good conditions of signal to noise ratio. The three options imply, in principle, only small differences in the inter-strip and coupling capacitance. To check this, we subdivided the tracks hitting the device according to the incident region and for each we compared the relevant distributions, finding no significant differences except for a slight worsening of the noise in the constant-gap region (Fig. 5.47). Based on this result, the wedge detectors are designed with a constant width over pitch ratio.

Since a significant fraction of particles will cross the detectors at non-orthogonal incident angles, the effect of inclined tracks has been studied in detail. The amount of charge released in the silicon and the number of fired strips is expected to increase while the average single strip noise should remain unaffected. Fig. 5.48 shows that the detector behaves as expected: there is a clear widening of the clusters and an increase of the signal-to-noise ratio. This is compatible with the predicted $1/\cos\theta$ dependence of the cluster charge and a constant strip noise.

Another important parameter which determines the behaviour of the detectors is the

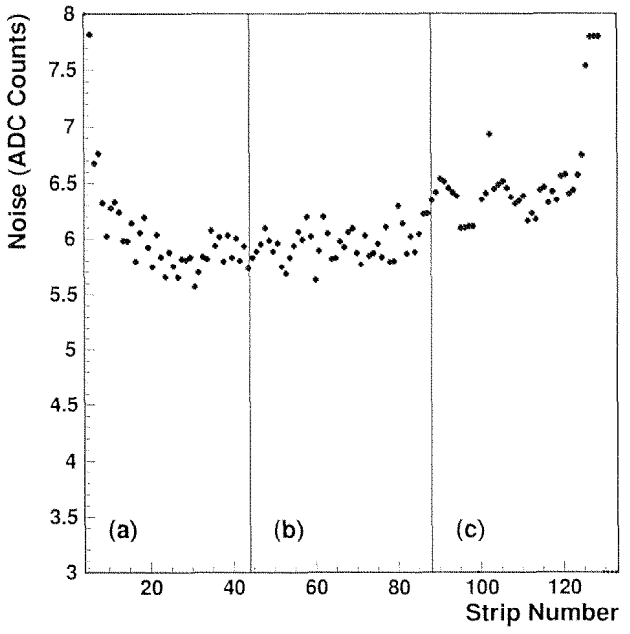


Fig. 5.47: Data for a wedge detector: strip noise for the three different regions: constant strip width (a), constant w/p (b), constant inter-strip gap (c)

temperature at which the silicon is operated. This can in fact influence sizeably both the amount of charge collected and the noise. We have studied the dependence of noise and charge collection efficiency as a function of temperature, [103] obtaining an increase in the signal-to-noise ratio up to 20% going from room temperatures to -10 °C, which is our choice as operating temperature (Fig. 5.49).

The last important parameter that influences the signal-to-noise performance is radiation damage. The irradiation produces both an increase of the noise, mainly as a consequence of the increased inter-strip capacitance (Fig. 5.41), and a decrease of the charge collection efficiency due to bulk damage [104].

These two effects can be partly reduced by increasing the bias voltage above the full depletion point, as shown in Figs. 5.50 and 5.51 for two detectors which have been irradiated with 1.0×10^{13} and 3.6×10^{13} n/cm^2 and then heated to produce an anti-annealing simulating an effective fluence of 1×10^{14} n/cm^2. We observe also (Fig. 5.51) that the collected signal grows with V_{bias}, reaching an asymptotic value for bias voltages significantly higher than the depletion voltage. Fig. 5.52 shows the collected signal for proton irradiated devices. For the highest dose of 1.6×10^{14} p/cm^2 the loss in charge collection efficiency is not more than 10%.

As a consequence, the signal-to-noise ratio for the irradiated devices approaches a plateau value at higher bias voltage (Fig. 5.53). There is still a 20% loss of signal-to-noise that is not recovered even going to the highest bias voltage. We have also investigated the performance of detectors irradiated up to $3.6\times10^{14}n/\mathrm{cm}^2$ [105]. The corresponding signal-to-noise ratio as a function of bias voltage is shown in Fig. 5.54. We see that for all fluences a similar dependence on V_{bias}/V_{dep} is observed.

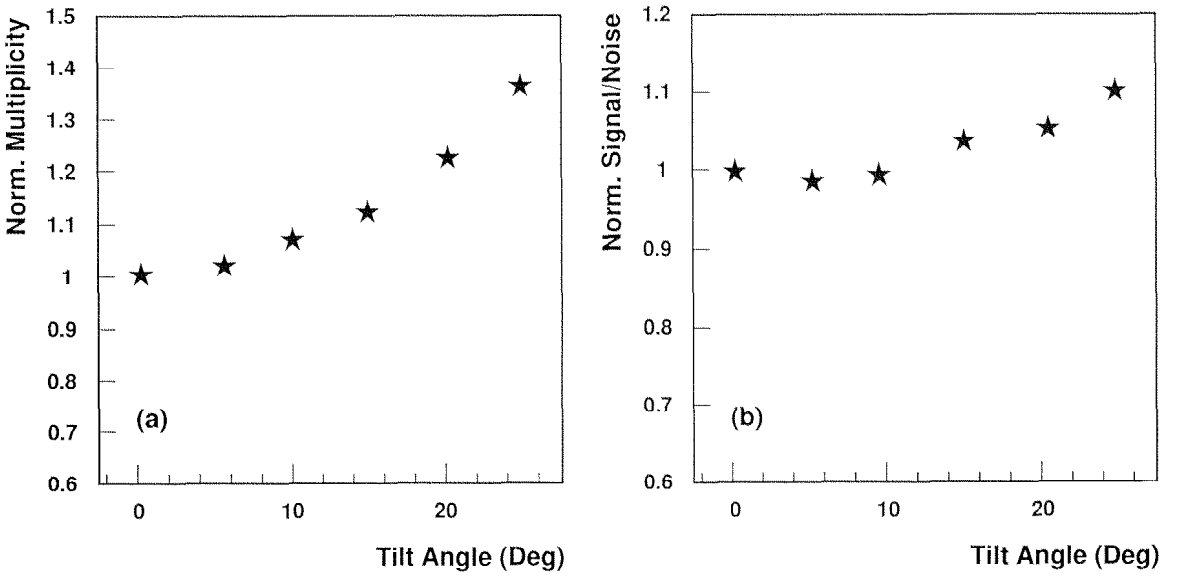


Fig. 5.48: Cluster multiplicity (a) and S/N ratio (b) as a function of the tilt angle. All points are normalized to the orthogonal incidence values. The readout pitch is 50 μm.

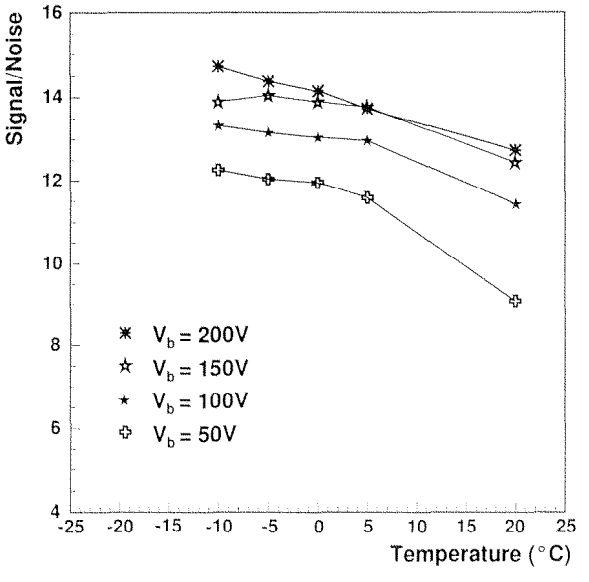


Fig. 5.49: Signal-to-noise ratio as a function of temperature at various bias voltages for a detector from CSEM irradiated at 1×10^{13} n/cm^2. The readout pitch is 50 μm.

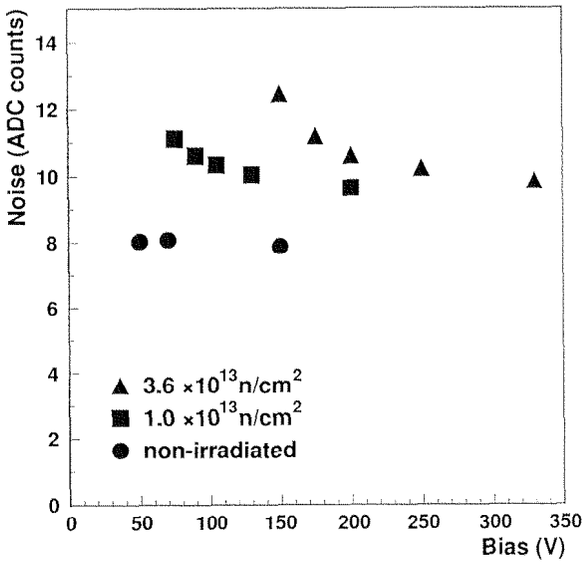


Fig. 5.50: Noise in ADC counts vs. bias voltage for detectors of 50μm pitch, 12.5cm length.

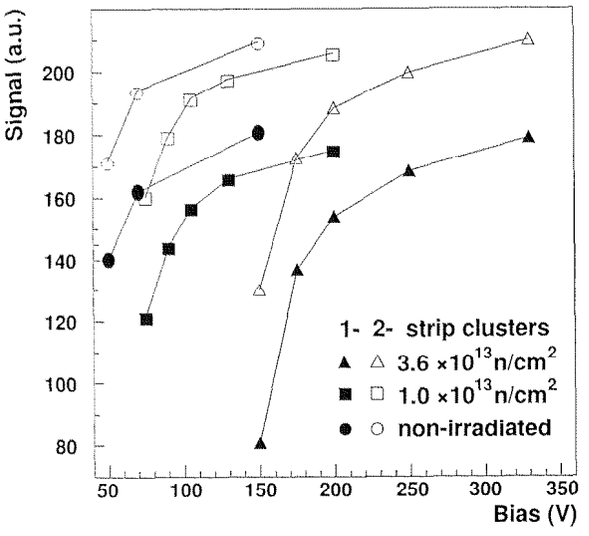


Fig. 5.51: Signal in ADC counts vs. bias voltage for the same detectors of Fig. 5.50, for 1- and 2-strip clusters.

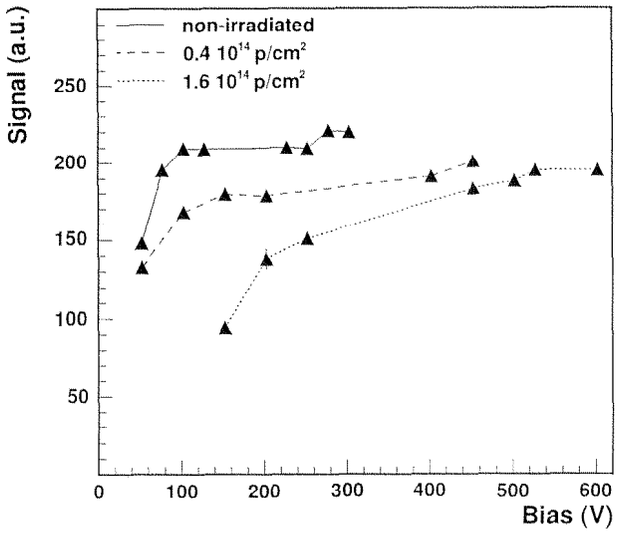


Fig. 5.52: Collected signal as a function of the bias voltage for proton irradiated devices, produced by Hamamatsu, strip length = 6.25 cm, strip pitch = 80 μm.

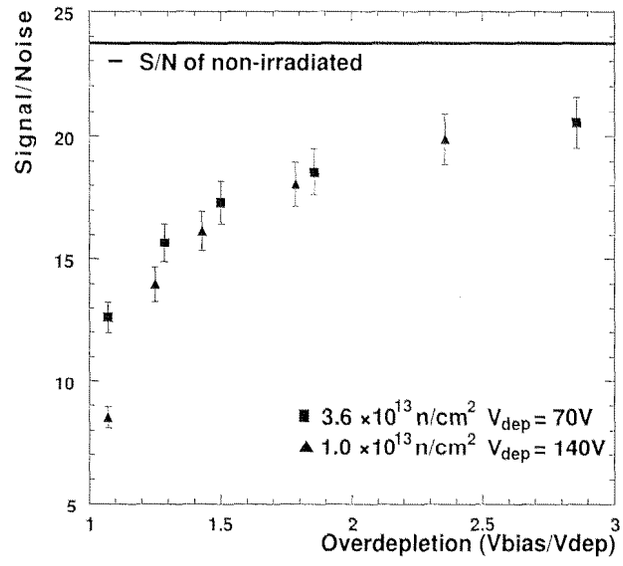


Fig. 5.53: Signal-to-noise ratio vs. bias voltage in over-depletion units for the same detectors of Fig. 5.50.

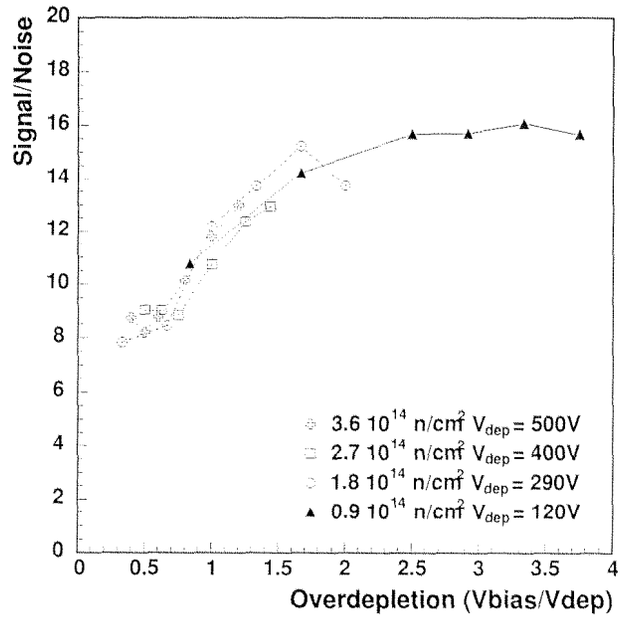


Fig. 5.54: Signal-to-noise ratio for heavily irradiated CSEM detectors of 50 μm pitch, 12.5cm length. Lower values with respect to Fig. 5.53 are due to a different, noisier experimental setup.

5.3.2 Detector efficiency

The global hit efficiency is defined as the ratio between the number of reconstructed hits and the number of tracks predicted to cross the detector under test within its geometrical acceptance and away from dead or noisy strips.

Fig. 5.55 and Fig. 5.56 show the dependence of the efficiency on bias voltage and signal-to-noise ratio, respectively.

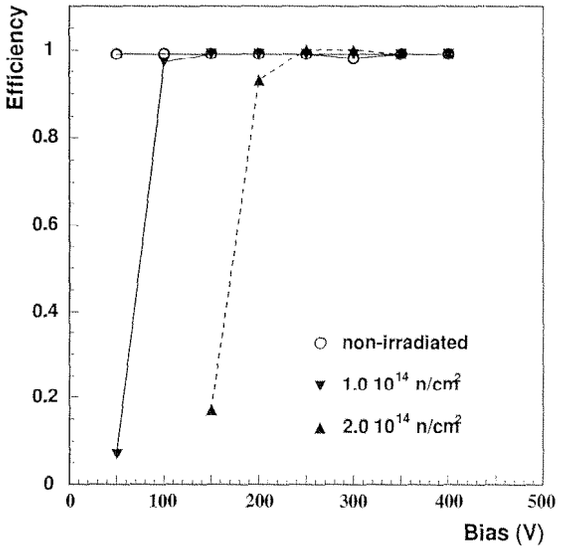


Fig. 5.55: Efficiency as a function of bias voltage. CSEM detectors of 75 μm pitch, 11 cm length, irradiated with neutrons.

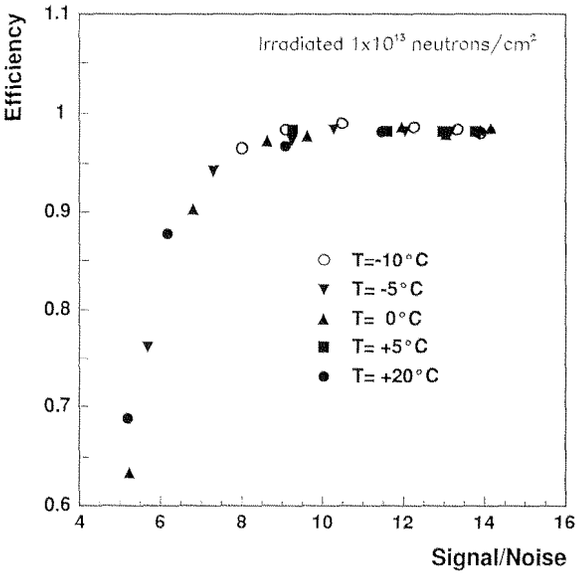


Fig. 5.56: Efficiency vs. S/N for irradiated CSEM detectors of 50μm pitch 6.22 cm length.

At full depletion both irradiated and non-irradiated detectors reach an efficiency compatible with 100%. It is worthwhile to note that in Fig. 5.56 even when S/N is 7.5 we have an efficiency greater than 95%.

The efficiency for an end-cap detector has also been checked before and after irradiation. Full efficiency is obtained also for non-uniformly irradiated devices at full depletion.

We have also measured the charge collection efficiency between strips as a function of the radiation damage. The cluster charge distribution versus the track position between two strips indicates good uniformity of charge collection even in heavily irradiated devices (Fig. 5.57).

5.3.3 Response function and spatial resolution

The particle impact point is usually given by the centre of gravity of the cluster, the weights being given by the charge collected on each strip. This way of reconstructing

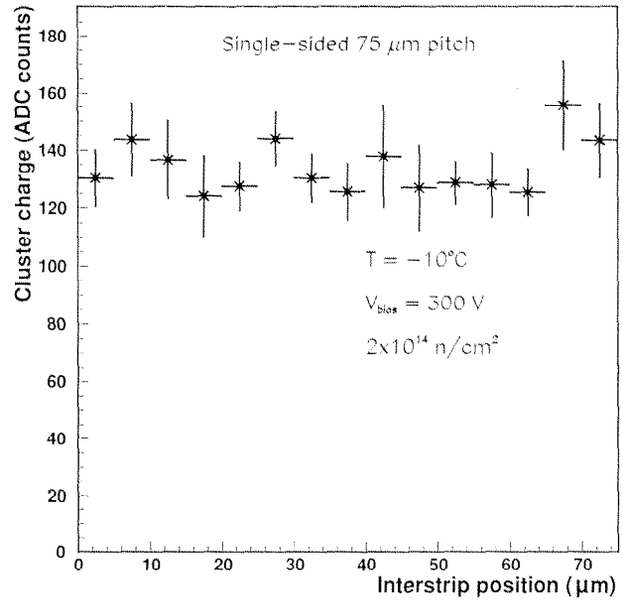


Fig. 5.57: Average cluster charge as a function of the inter-strip coordinate in an irradiated device.

positions assumes a linear sharing of the charge released on each strip. To study the deviations from the linear approximation and obtain a correction to the reconstructed position we have studied the response function η [102] and its dependence on radiation damage. The response function is defined as $\eta = Q_L/(Q_L+Q_R)$, where Q_L and Q_R are the charges collected on the strips to the left and to the right of the reconstructed hit position, respectively.

Fig. 5.58 shows that in the central region, populated by multi-strip clusters, the charge is linearly shared among the strips and hence the linear interpolation is correct, while in proximity of the strips single-hit clusters dominate and a digital response should be assumed [102]. Fig. 5.59 shows the same plot for a similar device irradiated to $1.8\times\ 10^{14}$ n/cm^2. We see that the shape is very similar.

A typical distribution of the residuals is plotted in Fig. 5.60 for a 50μm pitch detector. After subtracting the contributions due to multiple scattering and to the extrapolation error, we obtain a resolution of about 11μm in the coordinate orthogonal to the strips. This coordinate will be, in the final CMS tracker, the one measuring $r-\phi$ in the barrel and $z-\phi$ in the end-cap.

The dependence of the resolution on the signal-to-noise ratio, bias voltage and incidence angle is weak both for irradiated and non-irradiated modules. As an example, Fig. 5.61 shows the measured resolution for tracks at 20° incidence angle for irradiated and non-irradiated detectors as a function of the bias voltage.

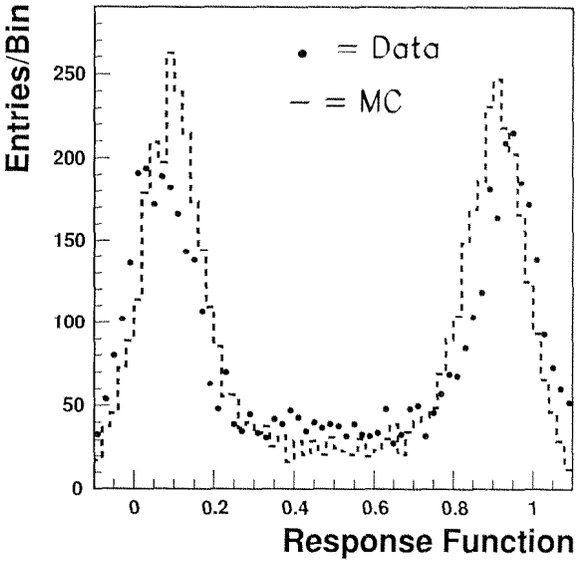


Fig. 5.58: Typical response function distribution. Monte-Carlo simulation (dashed line) is superimposed.

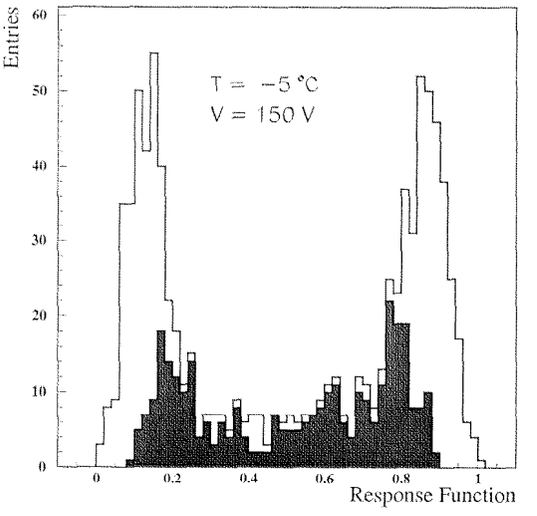


Fig. 5.59: Response function distribution for a CSEM detector irradiated to 1.8×10^{14}n/cm^2.

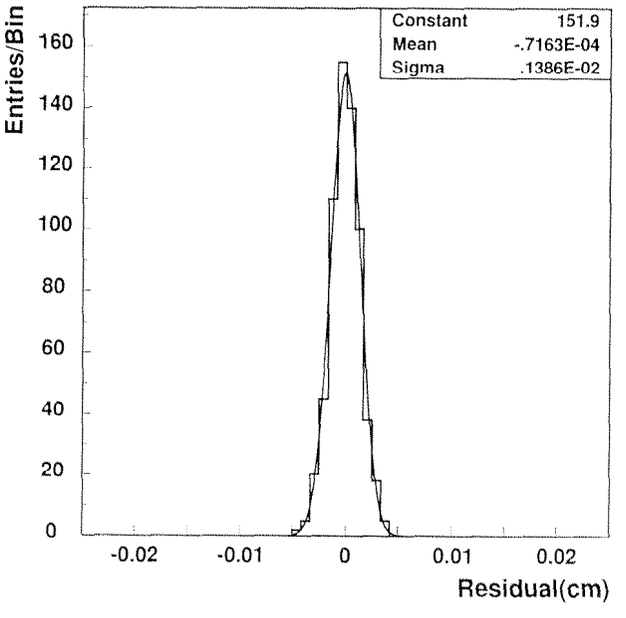


Fig. 5.60: Residual distribution for a 50μm pitch detector. From the width of this distribution we obtain a value of 11 μm for the intrinsic resolution.

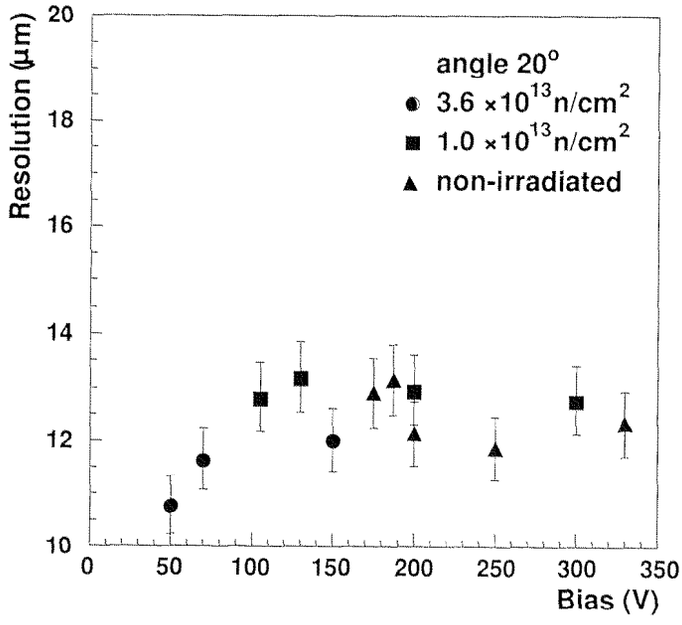


Fig. 5.61: Spatial resolution for tracks with an incident angle of 20°. Same detectors as Fig. 5.50.

5.3.4 Low momentum particles

Since a large fraction of particles at LHC will have momenta well below 1 GeV/c, we studied the detector response to low momentum particles. In a first beam test, data were collected with hadrons between 270 and 408 MeV/c momentum. Fig. 5.62 reports the charge collected on the p-side in double-side detectors and shows a clear separation between protons and pions at both energies.

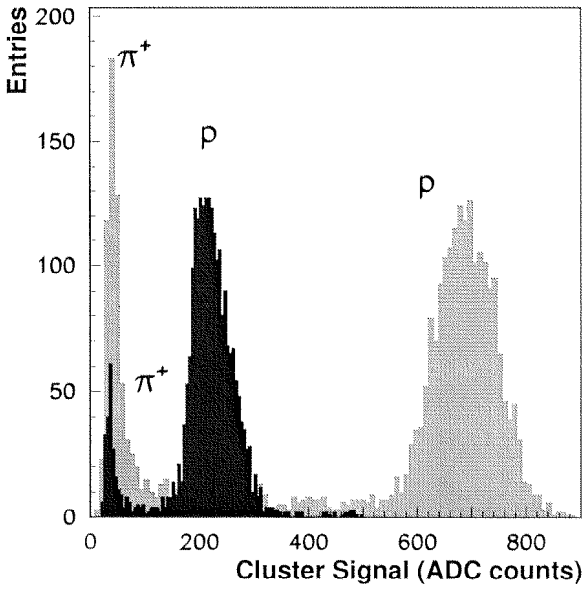


Fig. 5.62: Cluster signal at two different beam momenta of 270 (light area) and 408 (dark area) MeV/c. The contributions due to protons and pions are clearly visible.

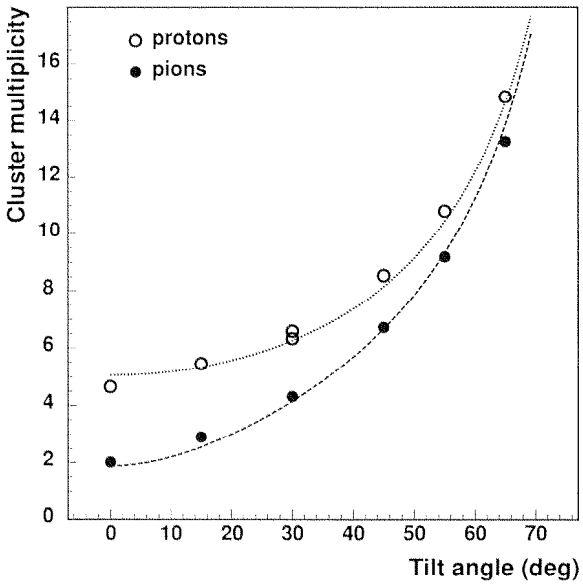


Fig. 5.63: Cluster width as a function of tilt angle for 310 MeV/c protons (empty points) and pions (filled points).

The cluster multiplicity associated to these particles may affect the occupancy. A careful study was performed with 310 MeV/c particles at different angles. Fig. 5.63 shows the cluster multiplicity for protons and pions. The experimental data have been used to simulate the SST response to low momentum hadrons.

5.4 The CMS Silicon Tracker September 1997 Milestones

During 1997 I collaborated to the assembly of a large scale prototype requested by the LHC Committee in order to address the main system problems of the SST. Two large system prototypes of the Barrel and of the Endcap part of the SST have been built for this purpose. CSEM delivered all the working detectors for both forward and barrel milestones. I have been involved especially in the barrel prototype, since most of the assembly has been performed in Pisa.

5.4.1 The SST Barrel Prototype

The features of the milestone prototype were based on a previous version of the tracker (V2) where 3 layers of silicon detectors were foreseen covering the radial region between 20 and 40 cm. The layout of the V2 geometry is shown in Fig. 5.64.

The detector modules are distributed in seven layers on a spiral geometry which leaves enough room for all services (cables, interconnect cards and cooling tubes). The detectors are tilted to partially compensate for the Lorentz angle. The detectors are organized to provide three detection points per track without dead regions in the tracking volume. Detectors overlap in $r - \phi$ over a few millimeters to allow for intermodule alignment.

A total number of 112 detector modules is used to equip the entire wheel. To reduce the costs there are only 14 working modules to be installed in the structure, whether the other 98 were produced with dummy components, mainly to address specific mechanical and thermal studies. Each dummy half-module consists of two dummy detectors, made out of 400 μm thick silicon wafers with aluminum strips patterned like the working ones. 50 Ω ceramic resistors are used in place of the front-end electronics for power dissipation.

The detector modules used in the structure were based on a preliminary engineered version of the barrel module approved as intermediate milestone in January 1997. The detectors are all single-sided. The hybrids are produced on a ceramic support with standard commercial connectors. The read-out electronics incorporate only the basic elements (preamplifiers, shapers and multiplexer) of the final chip and are implemented in radiation soft technology. An interconnect card, separated from the module, houses the components for the control of the read-out. The cables are made out of copper lines etched on thin kapton ribbons. No optical link was included in the read-out scheme.

5.4.1.a Support structure

Two machined carbon fibre disks are used as supporting elements coupled by thin inner and outer cylindrical skins. The two end-plates act as supporting elements for the detector modules, defining the exact position of the silicon detectors in space (see Fig. 5.64).

The radial position of the detector modules is defined by carbon fibre supporting ledges, while the other two coordinates are fixed by the coupling between precision pins in the modules and slots precisely manufactured in the ledges.

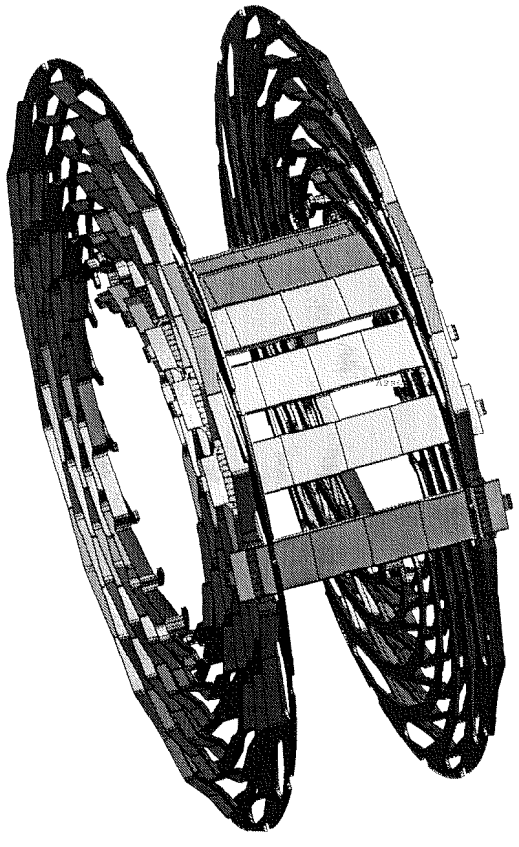


Fig. 5.64: Basic structure of the barrel wheel.

The first prototypes of this complicated structure were built in aluminum and carbon fibre. After an optimization of the fabrication procedure the two final disks were fabricated from a solid block of a composite material with the carbon fibres held together by a thermoplastic material cured at high temperature. The low coefficient of thermal expansion of this material prevents significant deformation and thermal stress when the structure is cooled down from the assembly temperature to the operating temperature. The machining was done in the INFN Pisa laboratory. The total weight of each disk is 2.1 kg. An accuracy better than 70 μm was measured for all of the critical elements.

5.4.1.b The cooling system

The heat generated by the detector modules in the wheel is removed by a cooling system based on a circulating fluid, with the goal of keeping the temperature of the silicon elements during running below -5 °C. The cooling tubes are arranged in layers in the supporting structure, in good thermal contact with the detector ledges. Several measurements have been performed to study the coupling interface between the cooling ledges and the modules. The heat generated inside the modules is estimated in 2 W for the electronics and 0.2 W for each silicon detector.

Tests and simulations show that at -15 °C and with a cooling pipe diameter of 3 mm,

using a mixture of water with 20 % of glycol ethylene, the maximum gradient between the cooling fluid and the silicon detectors is about 10 °C, with an average fluid flow per layer of about 0.6 l/min.

5.4.1.c Interconnect cards and cables

The interconnect cards (ICC) are used to distribute timing and input/output signals between the read-out hybrids and the DAQ system. Each module needs two interconnect cards, therefore 28 ICCs for the real modules and 196 dummy ICCs for the dummy modules have been produced. The cables are made of thin copper strips (35 μm thick) etched on a 75 μm kapton ribbon, long enough to reach the outer periphery of the wheel.

5.4.1.d Detector Module assembly

A barrel detector module is fabricated by coupling together two half-modules joined with stiff carbon fibre elements [2] (see Fig. 5.65).

Each half-module consists of two silicon detectors glued together head-on. The strips are daisy-chained between the two detectors, giving an effective length of 12 cm.

The assembly procedure consists of 5 steps: gluing of the precision positioning elements in the frame, alignment and gluing of the silicon detectors to the carbon fibre support, alignment and gluing of the hybrid, micro-bonding, and gluing the two half modules together.

A special jig was used to glue the positioning elements to the support giving a position accuracy of 50 μm. A specific jig has been used to glue the two detectors together: one of the detectors is held by vacuum in a fixed position while the other one can be moved in x and y and rotated in ϕ in order to align it with respect to the first. This alignment is done under a 3D measuring machine with a precision of a few microns. A gap of 60 μm is left between the detector edges and filled with glue. Then the pair of detectors is glued to the support by using another jig, which allows the 3D control of the planarity of detectors and carbon fibre support. After monitoring of the assembly procedure, it turned out that the strips can be aligned with an accuracy better than 0.1 mrad, whereas the detector position has an accuracy of 8 μm.

Once detectors are in place, the hybrid circuit is positioned by using the same jig and glued on the support. The hybrid circuit is made of ceramic and contains 8 PREMUX chips to read out 1024 silicon channels. The next step is bonding detectors and hybrid together. Approximately 2000 micro-bonds are needed for one half-module.

A full module is made by gluing together two half modules. Another dedicated jig is necessary. One half-module is held in a fixed position while the other one is on a x-y table that can be rotated around the axis normal to the silicon detectors.

The functionality of the modules was tested by measuring pedestals and noise as well

[2] High thermal conductivity carbon fibre is used for the mechanical frames of the modules. The fibre direction runs parallel to the side-rails that support the silicon detectors to optimize the heat removal.

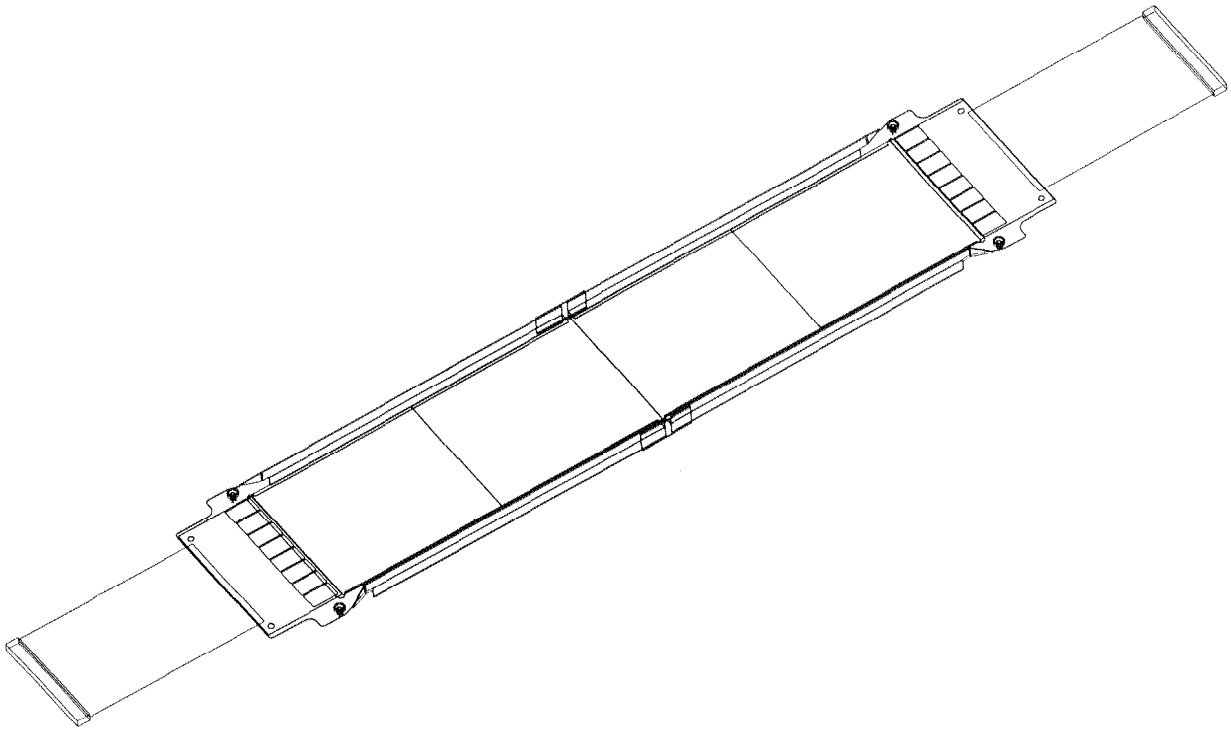


Fig. 5.65: Full detector module for the SiB1 milestone.

as signals generated by a pulsed LED. One example of the test results is shown in Fig. 5.66 and in Fig. 5.67.

Out of 31 fabricated half-modules, 30 passed the test and were used to build 15 modules for the milestone.

5.4.1.e System assembly

An assembly tool and a large coordinate measuring machine have been used to align the two disks together and to survey the insertion of the detector modules.

After the alignment of the disks, the module insertion was done manually by two operators. A total of 112 modules were installed in one week. At the end, the wheel position was measured again and no significant displacements was found. The assembled wheel is shown in Fig. 5.68.

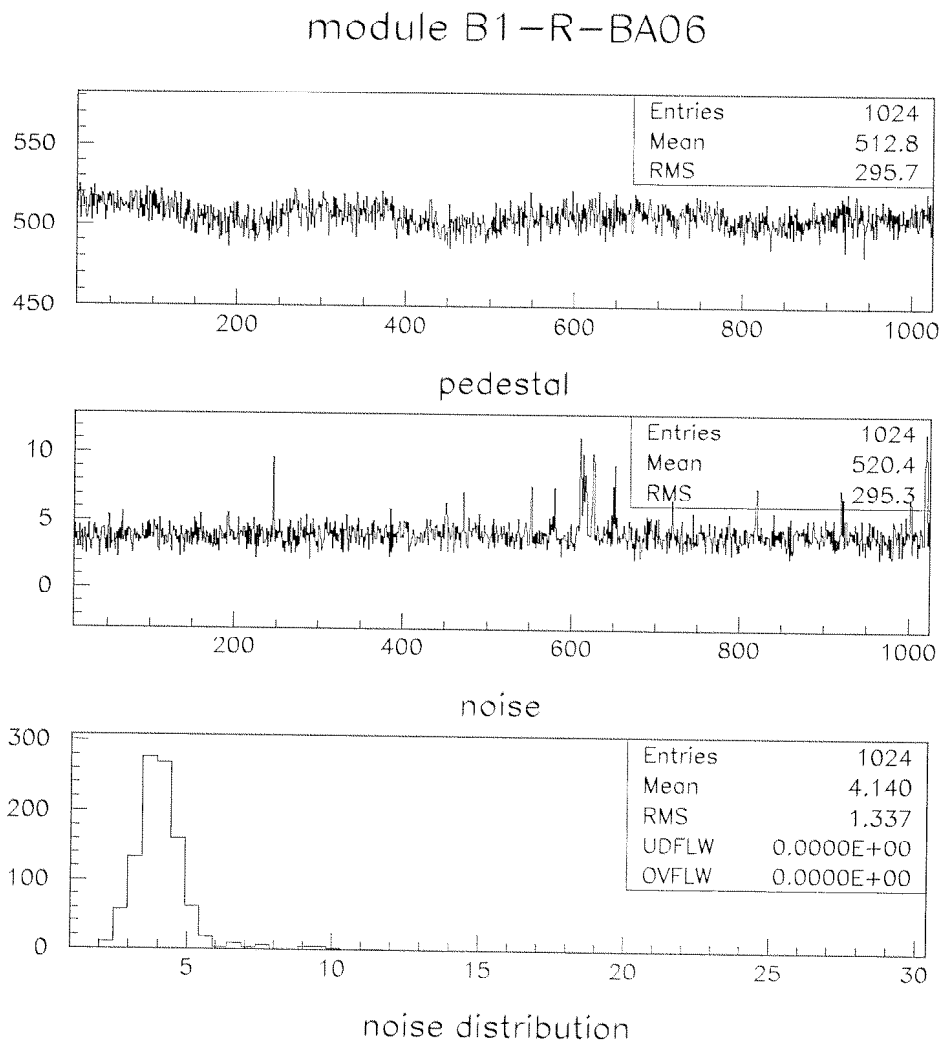


Fig. 5.66: Pedestal and noise of a module.

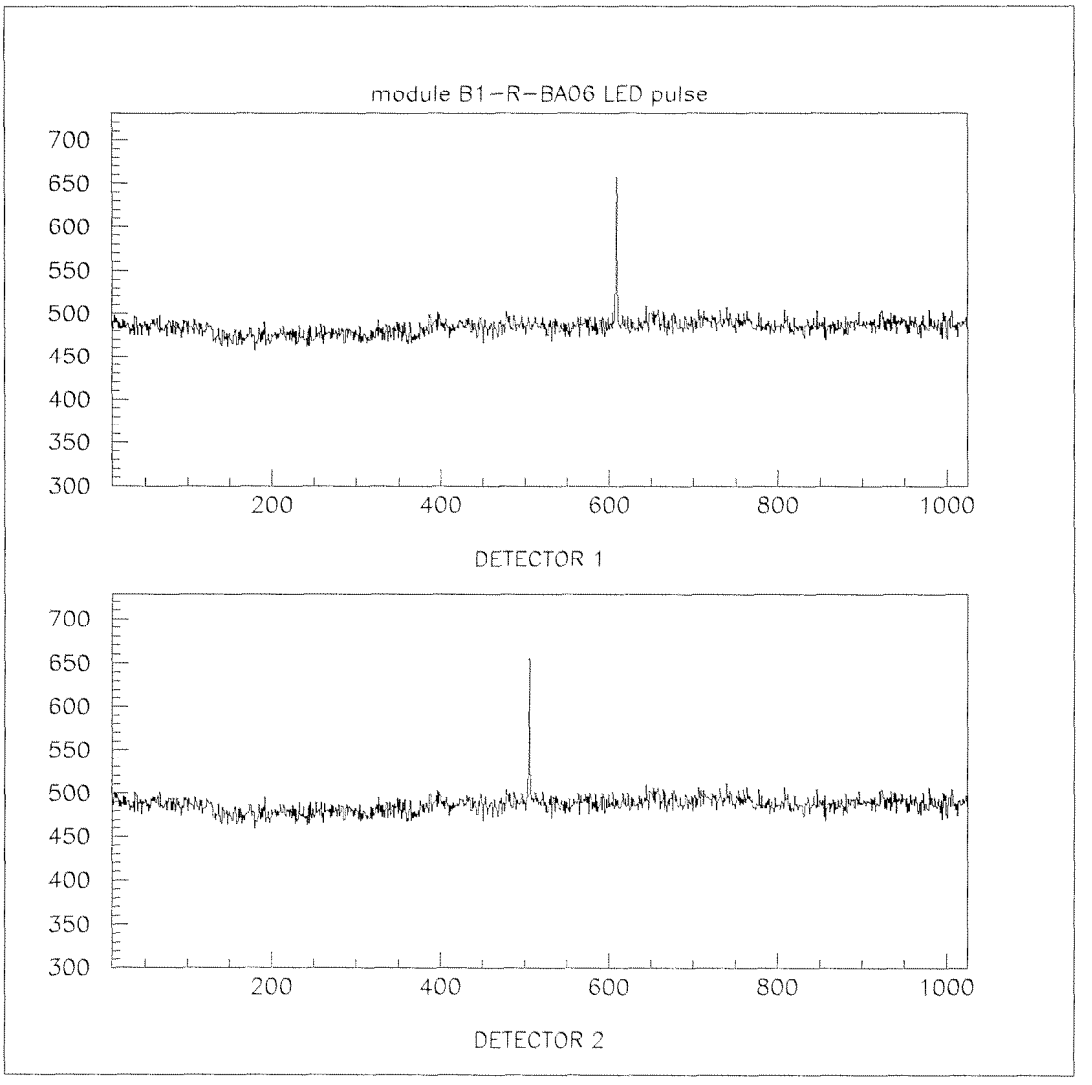


Fig. 5.67: Signals generated by LED pulses.

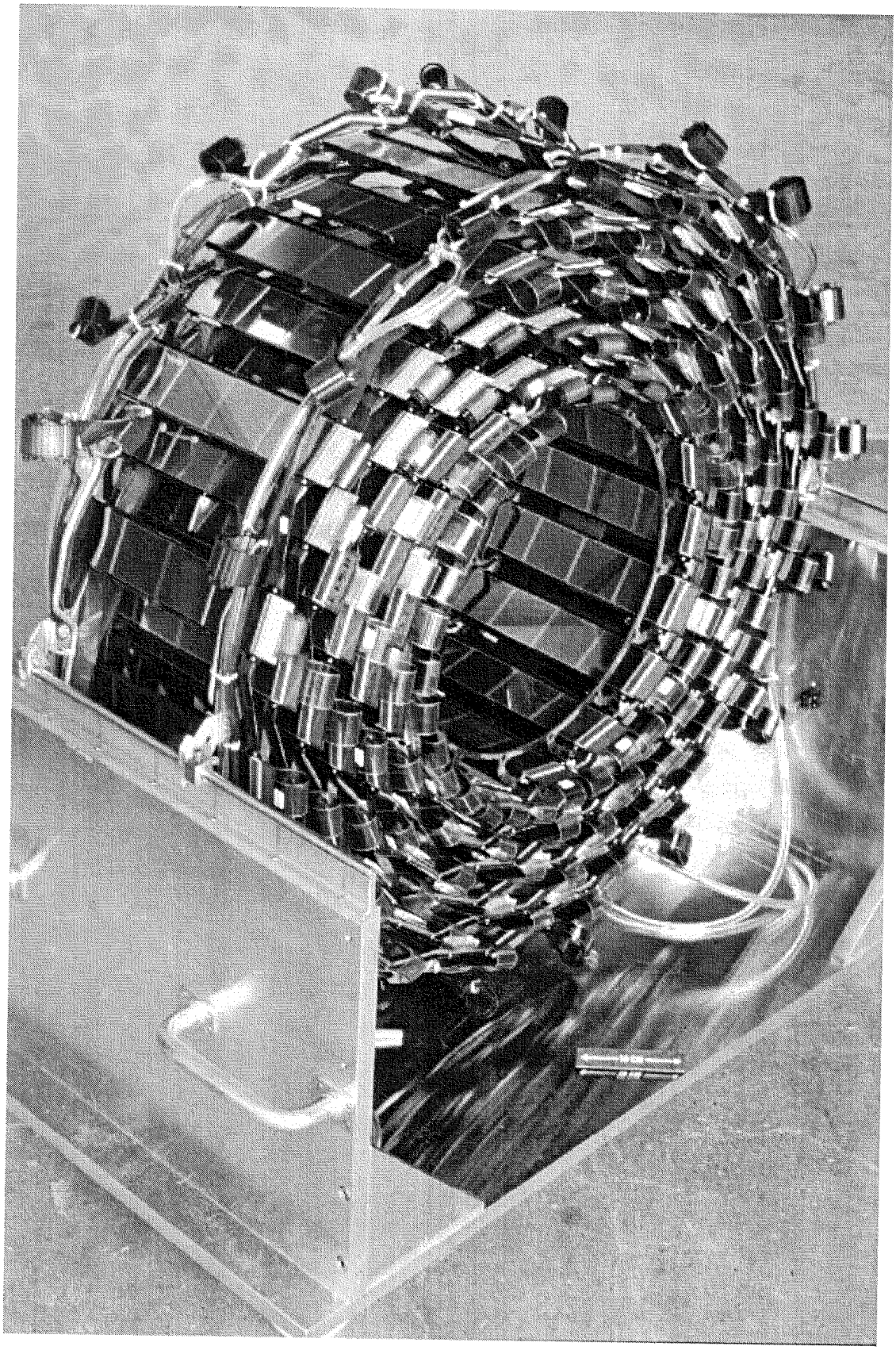


Fig. 5.68: The CMS SST barrel milestone.

Chapter 6

Search for SUSY Higgs boson h at CMS

The Standard Model (SM) of high energy physics provides a remarkably successful description of the presently known phenomena. The experimental high-energy frontier has advanced into the hundreds of GeV range with no confirmed deviations from Standard Model predictions and no unambiguous hints of additional structure. In particular LEP experiments (at the Z pole and above the WW threshold) showed an agreement at $< 0.1\%$ level of various observables with SM predictions. Despite its phenomenological successes, the SM suffers from various theoretical problems (*i.e.* naturalness, hierarchy) which make it seem unlikely to be a complete theory and will have to be extended to describe physics at arbitrarily high energies.

6.1 Supersymmetry

Supersymmetry (SUSY) is a novel type of symmetry that relates bosonic and fermionic degrees of freedom [2]. It therefore predicts new supersymmetric particles which are partners of all known particles and differs in spin by half a unit. In addition, the total number of fermionic and bosonic degrees of freedom must be equal: therefore to a given quark there are two complex scalar-quarks, one for each chiral degree of freedom. It should be noted that SUSY requires at least two Higgs doublets: by using only one Higgs doublet (as in SM), some of the Higgs-fermion Yukawa interaction terms, which are responsible for fermion masses, would violate supersymmetry, thus living some quarks (and leptons) massless.

Although no compelling supersymmetric model has yet emerged, and despite the fact that there is no direct experimental evidence for SUSY, the remarkable theoretical properties of SUSY theories have provided ample motivation for their study. Big interest is caused by the fact that SUSY leads to an amelioration of divergences in quantum field theory. This, in turn, protects the electroweak scale from large quantum corrections, and stabilizes the ratio $\frac{M_W}{M_X}$, when the Standard Model (SM) is embedded into a larger theory, involving an ultra-high energy scale M_X (*e.g.* M_{GUT} or M_{Planck}). In other words, SUSY

models do not require [106] the incredible fine-tuning (naturalness problem) endemic to the Higgs sector of the SM, provided only that the super-partners exist at or below the TeV energy scale.

To evaluate the experimental consequences of low energy supersymmetry, one must set up a Lagrangian including the various particles and partner sparticles, and their interactions [107]. Such a theory should reduce to the well-tested SM when the supersymmetric degrees of freedom are integrated over. The simplest possibility, the Minimal Supersymmetric Standard Model (MSSM), is a direct supersymmetrisation of the SM [2, 108]. It is a Yang-Mills type gauge theory based on the SM gauge group, with electroweak symmetry spontaneously broken via vacuum expectation values (*vev*) of two different Higgs superfields ($Y = \pm 1$) that respectively couple to $T_3 = \frac{1}{2}$ and $T_3 = -\frac{1}{2}$ fermions[1]. The effective scale of supersymmetry breaking is tied to the electroweak scale which is characterized by the SM Higgs vacuum expectation value $\langle H \rangle = 246$ GeV. The (renormalisable) superpotential that determines the Yukawa interactions of quarks and leptons is required to conserve baryon and lepton numbers; it is then possible to define a multiplicatively conserved R-parity quantum number which is +1 for ordinary particles and -1 for supersymmetric partners. The MSSM is thus minimal since it contains the smallest number of new particles and new interactions to be compatible with phenomenology. An important consequence of R-parity conservation is that the lightest supersymmetric particle (LSP) is stable. The LSP, which would have been abundantly produced in the early universe, is unlikely [109] to be colored or electrically charged since it would then be able to bind to nuclei or atoms to make heavy isotopes, for which searches [110] have yielded negative results. The LSP, which is the end product of every sparticle decay, thus escapes experimental detection, resulting in apparent non-conservation of energy/momentum in SUSY events.

6.1.1 SUSY Parameters

Supersymmetry must be a broken symmetry, otherwise ordinary particles and supersymmetric partners should be degenerate in mass. In the MSSM, supersymmetry breaking is accomplished by including the most general renormalisable soft-supersymmetry-breaking terms consistent with the $SU(3) \times SU(2) \times U(1)$ gauge symmetry and $R-$parity invariance. These terms parameterize the ignorance of the fundamental mechanism of SUSY breaking. For the MSSM, they consist of

- gaugino masses (M_1, M_2 and M_3 associated with the $U(1)$, $SU(2)$ and $SU(3)$ gauge groups of the Standard Model),
- mass terms for various left- and right- spin-0 (squark, slepton, Higgs) fields,
- trilinear Higgs-squark-squark and Higgs-slepton-slepton interactions amongst the scalars, the A-parameters,

[1] T_3 is the third component of weak isospin. It is related to the electric charge through the weak hypercharge Y as: $Q = T_3 + Y/2$

- analogous bilinear interactions (off-diagonal Higgs squared mass term), the B_0-term.

In addition to these soft-breaking terms, the ratio $\tan\beta$ of the two Higgs field *vev*'s and a supersymmetric Higgs-Higgsino mixing parameter μ (which also provides for the Higgsino mass terms as well as Higgs (mass)^2 terms) must be specified. Moreover it is useful to mention the parameters (other than μ) characterizing the conserving sector of the theory:

- the gauge couplings, g_s, g and g', corresponding to the SM gauge group $SU(3) \times SU(2) \times U(1)$ respectively,
- the Higgs-fermion-fermion Yukawa coupling constants, λ_i, corresponding to the coupling of the three generations of quarks, leptons and their superpartners to SM and SUSY Higgs.

6.1.2 SUSY particle spectrum

Aside from the particles of the Standard Model, the physical spectrum of the MSSM consists of the following additional states.

- squarks (spin-0): $\tilde{d}_L$,$\tilde{u}_L$,$\tilde{s}_L$, $\tilde{c}_L$,$\tilde{b}_1$,$\tilde{t}_1$, $\tilde{d}_R$,$\tilde{u}_R$,$\tilde{s}_R$, $\tilde{c}_R$,$\tilde{b}_2$,$\tilde{t}_2$;
- sleptons (spin-0): $\tilde{e}_L$,$\tilde{\nu}_{eL}$, $\tilde{\mu}_L$,$\tilde{\nu}_{\mu L}$, $\tilde{\tau}_1$,$\tilde{\nu}_{\tau L}$, $\tilde{e}_R$,$\tilde{\mu}_R$,$\tilde{\tau}_2$;
- charginos (spin-$\frac{1}{2}$): $\widetilde{\chi}_1^{\pm}$, $\widetilde{\chi}_2^{\pm}$;
- neutralinos (spin-$\frac{1}{2}$): $\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$, $\widetilde{\chi}_3^0$, $\widetilde{\chi}_4^0$
- gluino (spin-$\frac{1}{2}$): $\tilde{g}$;
- Higgs bosons: (spin-0): h, H, A, $H^{\pm}$.

Here, $\tilde{t}_i, \tilde{b}_i$, and $\tilde{\tau}_i$ $(i = 1, 2)$ are mixtures of the corresponding left- and right- chiral scalar fields.

As far as charginos and neutralinos are concerned, it should be noted that after electroweak symmetry breaking, the supersymmetric partners of the electroweak gauge and Higgs bosons (the gauginos and the Higgsinos) can mix. As a result, the physical mass eigenstates are linear combinations of these states, called charginos and neutralinos, which are obtained by diagonalising the corresponding mass matrices.

The intra-generational mixing of sfermions (being proportional to the corresponding *fermion* mass) is negligible for the first two generations.

An independent value for each one of the above masses and couplings leads to a proliferation of new parameters: the MSSM possesses 124 truly independent parameters [111], of these 18 corresponds to SM parameters, one corresponds to a Higgs sector (e.g. SM Higgs mass) and 105 are completely new ones (mixing angles, sparticle masses, CP-violating phases). It is clear that with those premises a phenomenological analyses becomes intractable! It is therefore customary to assume that higher symmetries, which are broken at some ultra-high scale, relate these parameters.

6.1.3 The minimal supergravity model (mSUGRA)

An especially appealing and economic class of models is based on minimal supergravity (mSUGRA) GUTs (often also called "constrained MSSM"), where it is assumed the weak, electromagnetic and strong coupling parameters α_i, $i = 1, 2, 3$ (roughly speaking $\alpha_i = g_i^2/4\pi$) unify at some ultra-high energy scale M_X (see Fig. 6.1). It is also assumed

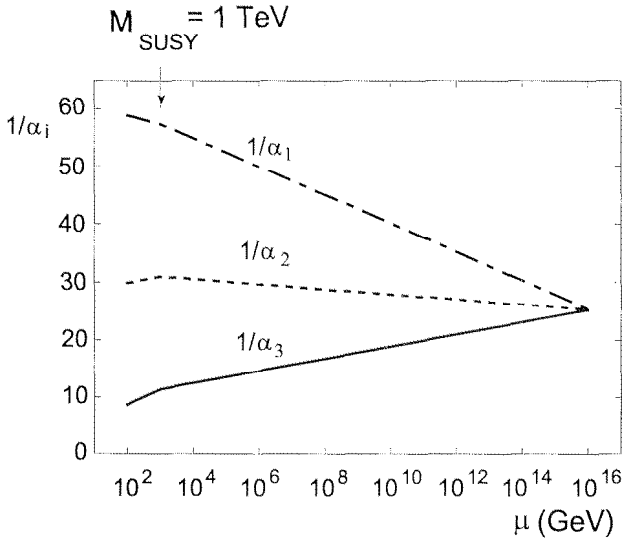


Fig. 6.1: RG evolution of the inverse gauge couplings $\alpha_i^{-1}(Q)$ in the MSSM. Sparticle masses threshold set to 1 TeV.

that SUSY breaking in the effective theory defined at M_X arises due to gravitational interactions which, being universal, allow [2] only a few independent soft SUSY breaking parameters, renormalised at M_X: these are

- a common gaugino mass ($m_{1/2}$),
- a common scalar mass (m_0),
- a common trilinear interaction (A_0), and
- the bilinear coupling (B_0).

The various MSSM masses and couplings have to be evolved through Renormalisation Group Equations (RGEs) from the common value at M_X to the electroweak scale to sum the large logarithms arising from the disparity between the two scales. This generally involves solving 26 coupled differential equations, with the values of the four GUT-scale parameters as boundary conditions. A bonus of this framework is that the same radiative corrections that give rise to these large logs also yield a mechanism for the breakdown of electroweak gauge symmetry, leaving colour and electromagnetic gauge symmetries unbroken. The electroweak symmetry breaking constraint allows one to eliminate B_0 in

favor of $\tan\beta$, and also to adjust the magnitude (but not the sign) of the μ parameter to get the measured Z^0 boson mass. Thus, the RG-evolution of these four parameters, renormalised at the GUT scale where the physics is simple, results in a rich pattern of sparticle masses and couplings at the weak scale relevant for phenomenology. The various SUSY parameters, masses and mixings are then determined in terms of the four plus a sign parameter set

$$m_0,\ m_{1/2},\ \tan\beta,\ A_0 \text{ and } sign(\mu).$$

In addition, as for the SM, the top mass m_t must be specified. The simplest supergravity model leads to (approximate) degeneracy of the first two generations of squarks, and so, is automatically consistent with constraints [112] from flavor changing neutral currents in the K-meson sector. When these supergravity constraints are incorporated, one finds (approximately):

$$\begin{aligned} m_{\widetilde{\chi}_1^0} &\approx \frac{1}{2} m_{1/2} & \qquad m_{\tilde{q}_{L,R}} &\approx \sqrt{m_0^2 + 6m_{1/2}^2} \\ m_{\widetilde{\chi}_2^0} &\approx m_{\widetilde{\chi}_1^\pm} \approx m_{1/2} & \qquad m_{\tilde{\ell}_L} &\approx \sqrt{m_0^2 + 0.52m_{1/2}^2} \\ m_{\tilde{g}} &\approx 3m_{1/2} & \qquad m_{\tilde{\ell}_R} &\approx \sqrt{m_0^2 + 0.15m_{1/2}^2} \end{aligned}$$

It follows that sleptons may be significantly lighter than the first two generations of squarks. Furthermore, the value of $|\mu|$ is generally large compared to the electroweak gaugino masses, so that the lighter neutralinos ($\widetilde{\chi}_{1,2}^0$) and the lighter chargino ($\widetilde{\chi}_1^\pm$) are gaugino-like, while the heavy chargino and the heavier neutralinos are dominantly Higgsinos.

If $m_{\tilde{q}} \simeq m_{\tilde{g}}$ so that sleptons are significantly lighter than squarks, the leptonic decays of $\widetilde{\chi}_2^0$, and sometimes also of $\widetilde{\chi}_1^\pm$ can be significantly enhanced relative to those of Z and W bosons, respectively; this has important implications [113] for detection of sparticles at hadron colliders.

Within the mSUGRA framework, the lightest SUSY particle is a viable candidate for dark matter [114], provided that the sfermions are not too heavy (the LSP, being mostly a gaugino, mainly annihilates via sfermion exchange, so that the annihilation rate is proportional to $\frac{1}{m_{\tilde{f}}^4}$). It is customary to identify the LSP with the lightest neutralino $\widetilde{\chi}_1^0$.

While minimal supergravity models indeed provide an economic and elegant framework, it should be recognized that the assumptions (about the physics at an ultra-high energy scale) on which they are based may ultimately prove to be incorrect. The point, however, is that these models lead to rather definite correlations between various sparticle masses as well as between the cross-sections for numerous signals. These predictions, which serve as tests of the underlying assumptions, can be directly tested at LHC.

6.1.3.a Simulation

For the event simulation of the various SUSY signals and the SM background reactions at hadron colliders two general purpose Monte Carlo programs are available: ISASUSY (ISASUGRA) [115] and SPHYTIA [116]. The approach to calculation of mass spectrum, branching ratios and decay ratio is slightly different in the two programs: ISAJET implements numerical solutions of RGE's. Iteratively evolves between GUT and electroweak scale including SUSY thresholds and electroweak breaking by means of a next-to-leading order potential. SPHYTIA makes use of approximated formulas instead of RGE's and provides external mass calculations. It includes a 2-loop effective potential for the Higgs. Both programs agrees within 10% using the same assumptions. Typical common features are e.g. full BR's spectrum calculation, generation of SUSY process with leading-order cross sections, decay sparticle using branching ratios and phase space and leading-log QCD radiation and hadronisation.

6.1.3.b Sparticle production and cascade decays

At hadron colliders, sparticles can be produced via the following lowest order reactions (particles/anti-particles not distinguished):

- $qq,\ gg,\ qg \to \tilde{g}\tilde{g},\ \tilde{g}\tilde{q},\ \tilde{q}\tilde{q}$, (strong production)
- $qq,\ qg \to \tilde{g}\widetilde{\chi}_i^0,\ \tilde{g}\widetilde{\chi}_i^{\pm},\ \tilde{q}\widetilde{\chi}_i^0,\ \tilde{q}\widetilde{\chi}_i^{\pm}$ (associated production)
- $qq \to \widetilde{\chi}_i^{\pm}\widetilde{\chi}_j^{\mp},\ \widetilde{\chi}_i^{\pm}\widetilde{\chi}_j^0,\ \widetilde{\chi}_i^0\widetilde{\chi}_j^0$ ($\tilde{\chi}$ pair production)
- $qq \to \tilde{\ell}\tilde{\nu},\ \tilde{\ell}\tilde{\ell},\ \tilde{\nu}\tilde{\nu}$ (slepton pair production)

In addition, the Higgs bosons can be produced via direct s-channel subprocess,

- $qq,\ gg \to h,\ H,\ A,\ H^-H^+$

in association with other heavy quarks and vector bosons, and in some cases, production via vector boson fusion.

At LHC, pair production of strongly interacting particles like $\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$ and $\tilde{q}\tilde{q}$ has the largest cross section. In Fig. 6.2 the sum of the total cross sections of $\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$ and $\tilde{q}\tilde{q}$-production at $\sqrt{s} = 14$ TeV is plotted as a function of m_0 and $m_{\frac{1}{2}}$ for $\tan\beta = 2$ and $\mu < 0$; the corresponding iso-cross-section curves are shown in Fig. 6.3.

The sum of the gluino-squark total production cross section for $m_{\tilde{g}} = m_{\tilde{q}}$ is shown in Fig. 6.4; for comparison, the cross section for associated production for $\tilde{g}/\tilde{q}$ with $\tilde{\chi}_i^{\pm}/\tilde{\chi}_i^0$ is plotted with dot-dashed lines, and $\tilde{\chi}_1^{\pm}/\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_1^{\pm}/\tilde{\chi}_2^0$ production cross-sections with dashed lines. The summed strong production of $\tilde{g}\tilde{g} + \tilde{g}\tilde{q} + \tilde{q}\tilde{q}$ is the dominant cross section over the complete range of $m_{\tilde{g}}$ all the way up to 2 TeV. Associated production is always a sub-dominant component of sparticle pair production.

Once produced, sparticles decay to other sparticles through a *cascade* process: squarks and gluinos can have strong decays (left) or weak decays (right)

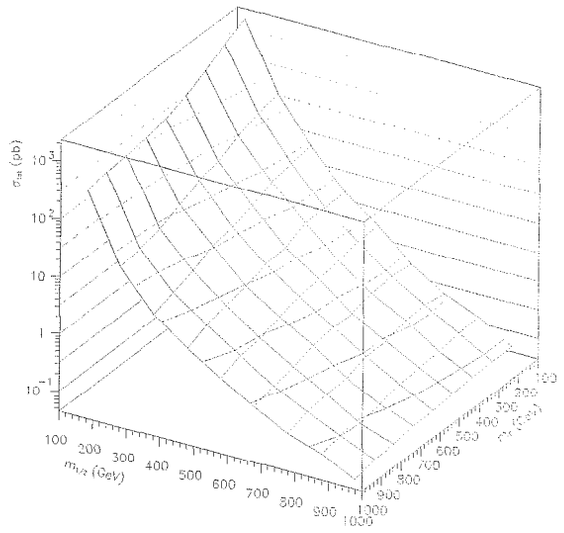


Fig. 6.2: $\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$ and $\tilde{q}\tilde{q}$ production cross sections at LHC for $tan\beta = 2$ and $\mu < 0$.

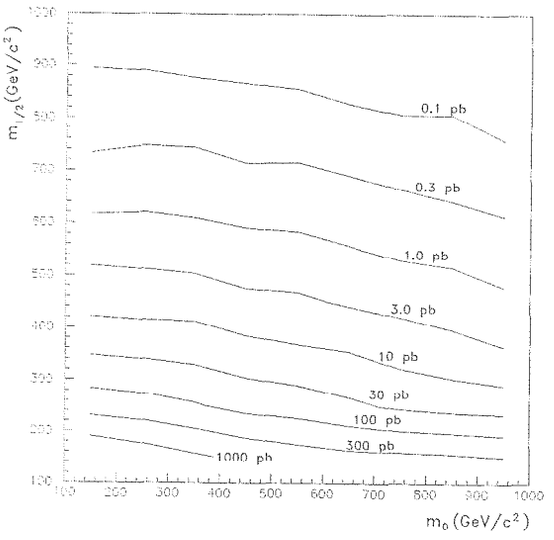


Fig. 6.3: Iso-contour curves of the production cross-section for the same parameters as in Fig. 6.2.

$$\begin{aligned} \tilde{q}_{L,R} &\to q\tilde{g}, \quad (m_{\tilde{q}_{L,R}} > m_{\tilde{g}}) \\ \tilde{g} &\to q\bar{\tilde{q}}_{L,R}, \bar{q}\tilde{q}_{L,R} \quad (m_{\tilde{g}} > m_{\tilde{q}_{L,R}}) \end{aligned}$$

$$\begin{aligned} \tilde{q}_{L,R} &\to q\tilde{\chi}_i^0, \\ \tilde{q}_L &\to q'\tilde{\chi}_i^{\pm}, \\ \tilde{g} &\to q\bar{q}\tilde{\chi}_i^0, \\ &\to q\bar{q}'\tilde{\chi}_i^{\pm}, \\ &\to g\tilde{\chi}_i^0. \end{aligned}$$

Strong decays are dominant if they are kinematically allowed. Charginos and neutralinos have both leptonic and hadronic decays via (virtual or real) $W^{\pm}$, Z^0, $\tilde{l}$, $\tilde{\nu}_l$ or Higgs particles:

$$\begin{aligned} \tilde{\chi}_i^{\pm} &\to l^{\pm}\nu_l\tilde{\chi}_j^0, \\ &\to q\bar{q}'\tilde{\chi}_j^0, \\ &\to l^{\pm}l^{\mp}\tilde{\chi}_1^{\pm}, \\ &\to \nu_l\bar{\nu}_l\tilde{\chi}_1^{\pm}, \\ &\to q\bar{q}\tilde{\chi}_1^{\pm}, \end{aligned}$$

$$\begin{aligned} \tilde{\chi}_i^0 &\to l^{\pm}l^{\mp}\tilde{\chi}_j^0, \\ &\to \nu_l\bar{\nu}_l\tilde{\chi}_j^0, \\ &\to q\bar{q}'\tilde{\chi}_j^0, \\ &\to q\bar{q}'\tilde{\chi}_j^{\pm}, \\ &\to l^{\pm}\nu_l\tilde{\chi}_j^{\mp}, \\ &\to q\bar{q}\tilde{\chi}_j^{\pm}, \end{aligned}$$

The cascade decays terminate when the $\tilde{\chi}_1^0$, assumed to be the LSP, is reached. For a large part of the parameter space most of the electroweakly interacting SUSY particles

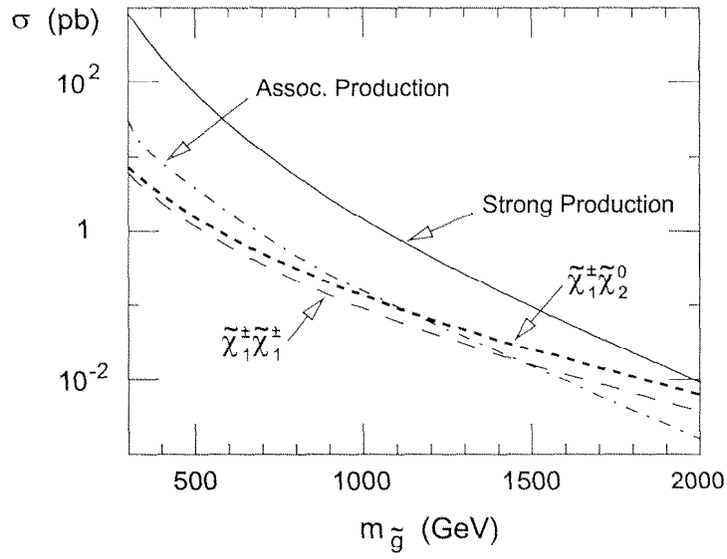


Fig. 6.4: mSUGRA production cross sections at LHC for $m_{\tilde{q}} = m_{\tilde{g}}$, $tan\beta = 2$ and $\mu < 0$.

come from cascade decays rather than direct production. SUSY events from cascade decays will therefore have the following characteristic signatures:

- large missing transverse energy $\not{E}_T$,
- high multiplicity of jets with large transverse momentum P_T,
- copious production of central b–jets, since they result from the decay of massive squarks and gluinos which are centrally produced.

The CMS detector is designed to clearly identify and analyze events which exhibit one or more of these characteristic features. An experimental signal for SUSY will show an excess of these events compared to the SM prediction.

6.1.3.c Choice of Parameters

Both CMS and ATLAS Collaborations are setting up a common phenomenological framework for the simulations within the mSUGRA space, by defining five points to compare results of analyses (see Tab. 6.1). Point 3 is the so-called "comparison point", as it is also used for SUSY studies at the proposed NLC and TEVATRON (TEV33) colliders.

Table 6.2 gives the total cross sections for pair production of SUSY particles at the five LHC points.

As an example the masses of the SUSY particles predicted for the point 1 points are shown in Table 6.3.

In Fig. 6.5 iso-mass contours are plotted in the $m_0 - m_{\frac{1}{2}}$ plane for $\tan\beta = 2$, $\mu < 0$ and $A_0 = 0$. The shaded region is excluded either because electroweak symmetry is not broken

Point	m_0	$m_{\frac{1}{2}}$	A_0	$\tan\beta$	sgnμ
1	400	400	0	2.0	-
2	400	400	0	10.0	+
3	200	100	0	2.0	-
4	800	200	0	10.0	+
5	100	300	300	2.1	+

Table 6.1: mSUGRA parameters for the five LHC points

	σ(pb) at LHC Point:				
Channel	1	2	3	4	5
$\tilde{g}\tilde{g}$	0.24	0.26	437.19	108.77	1.75
$\tilde{q}\tilde{q}$	1.65	1.01	176.83	1.35	5.20
$\tilde{g}\tilde{q}$	1.49	1.44	642.77	8.26	8.31

Table 6.2: Strongly interacting particle production cross section in *pb* for production of SUSY particles in the five LHC points

appropriately, or because the lightest neutralino is not the LSP. The cross-shaded region is excluded by experimental searches for SUSY at LEP and TEVATRON experiments. Referring to Fig. 6.5, the most interesting domain in the $m_0 - m_{\frac{1}{2}}$ parameter plane for mSUGRA Higgs search is that labeled as B: here the decay $\widetilde{\chi}_2^0 \longrightarrow \widetilde{\chi}_1^0 h$ is kinematically allowed with BR's greater than 50% (see Sec. 6.2). In either regions A and C this decay is either kinematically forbidden or has very low branching fraction: in region A the largest fraction of $\widetilde{\chi}_2^0$ decays are in two bodies $\tilde{\nu}\nu$ or $\tilde{l}l$ whether in region C, decay $\widetilde{\chi}_2^0 \longrightarrow \widetilde{\chi}_1^0 Z$ is enhanced.

Particle masses (in GeV) at LHC Point 1			
$m_{\tilde{g}} = 985$	$m_{\widetilde{\chi}_1^0} = 162$	$m_h = 85.2$	$m_{\tilde{\ell}_L} = 494$
$m_{\tilde{q}_L} = 933$	$m_{\widetilde{\chi}_2^0} = 317$	$m_H = 1092$	$m_{\tilde{\ell}_R} = 430$
$m_{\tilde{q}_R} = 900$	$m_{\widetilde{\chi}_1^\pm} = 316$	$m_A = 1089$	
$m_{\tilde{t}_1} = 648$			

Table 6.3: Masses for the main SUSY states as generated by SPYTHIA in LHC point 1.

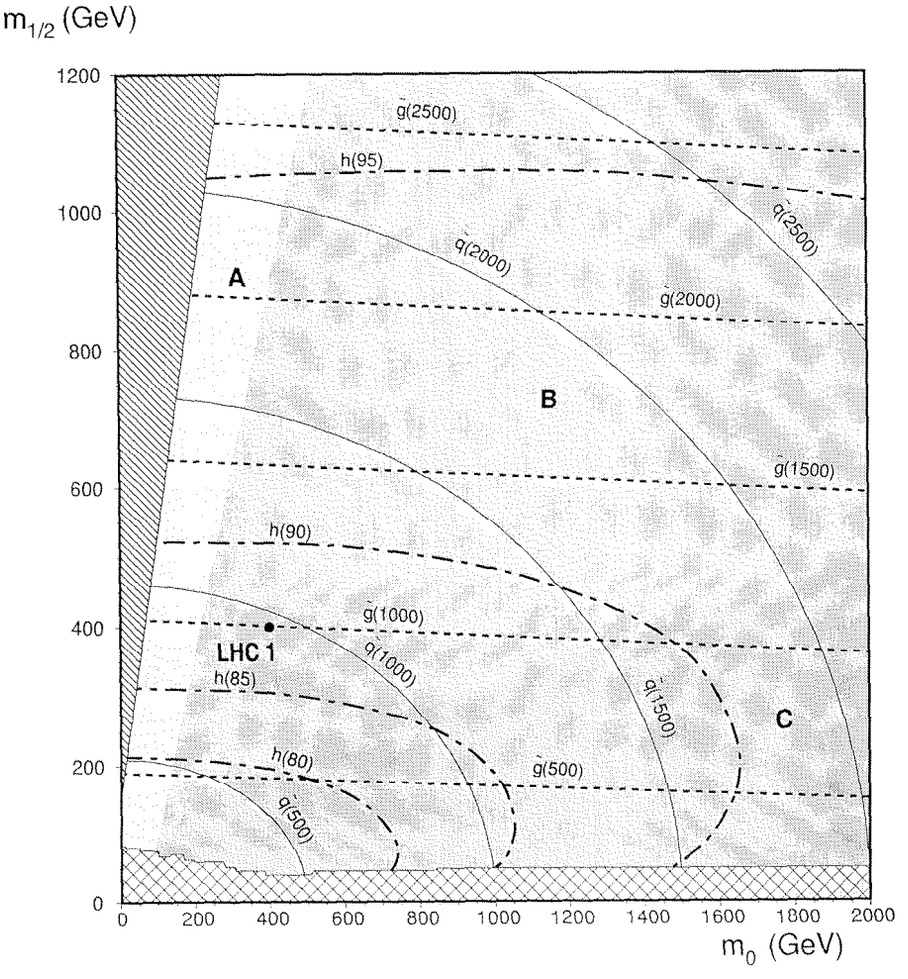


Fig. 6.5: Isomass contour in the $m_0 - m_{\frac{1}{2}}$ plane for gluinos, squarks and the lighter Higgs. Shaded regions in the parameter space are forbidden either by theoretical or experimental reasons. Regions are labeled with letters referring to different decay channels for the $\widetilde{\chi}_2^0$.

6.2 Search for SUSY lightest CP-even Higgs h

The MSSM Higgs sector requires the existence of two Higgs doublets, as pointed out in Sec. 6.1. Within this model at least one Higgs boson, h, should have a mass smaller than 125 GeV [117]. The upper mass limit depends on top and stop quark masses and on $\tan\beta$. Present status of lightest SUSY CP-even Higgs searches at LEP2 fixes the lower limit for its mass m_h to ~ 84 GeV [118]. One expects that such Higgs boson should be found at LEP200 if $|\tan\beta| \lesssim 4$. Current analyses of the SUSY Higgs sector within CMS show that the parameter space available for those bosons is already restricted [119]. The most favorable signature for the detection of the lightest Higgs appears to be the inclusive decay $h \longrightarrow \gamma\gamma$, where, for masses of $m_A \geq 400$ GeV, rates and sensitivity are of the same order of SM. For lower m_A the branching ratio falls below a statistically significant signal. In order to complement this analysis it is therefore necessary to highlight another h decay channel with good detection probability.

6.2.1 Method

As pointed out in Sec. 6.1.3.c and shown in Fig. 6.5, the production of h in decays of neutralinos and charginos (primarily $\widetilde{\chi}_2^0$) is sufficiently large in a significant portion of the mSUGRA parameter space (zone B), especially at low values of $\tan\beta$. This is due to the large strong interaction cross section for the production of squarks and gluinos (see Figs. 6.2-6.3), which in turn abundantly decay via cascades to $\widetilde{\chi}^0$ and $\widetilde{\chi}^{\pm}$ (mainly $\widetilde{\chi}_2^0$). If $M_{\widetilde{\chi}_2^0} - M_{\widetilde{\chi}_1^0} > m_h$, the decay $\widetilde{\chi}_2^0 \longrightarrow \widetilde{\chi}_1^0 h$ is kinematically allowed, the leptonic decay rate is reduced while the number of b–jets in the event is enhanced (as in LHC Point 1): this decay becomes the main source of h production. By looking to dominant decays of $h \longrightarrow b\bar{b}$ and large missing energy (due to $\widetilde{\chi}_1^0$ escaping the detector undetected) it would be possible to detect and measure the mass of this CP-even Higgs boson.

In Fig. 6.6 an example of a $\tilde{g}$,$\tilde{q}$ cascade decay chain ending up with a $b\bar{b}$ pair from a h is plotted. Here I will focus only on the region close to the aforementioned LHC Point 1 (for definition see Tab. 6.1). The branching ratios corresponding to the decay process in Fig. 6.6 have been calculated and are listed in Tab. 6.4. This point is typical for the domain of parameter space where $m_{\tilde{g}} > m_{\tilde{q}}$, where gluinos decays via squarks and $\sim 30\%$ of $\tilde{q}_L$ decay to $\widetilde{\chi}_2^0$+quarks. It should be noted that with increasing m_0 (at fixed $m_{\frac{1}{2}}$) squarks are becoming more massive than gluinos, decreasing significantly the yield of $\widetilde{\chi}_2^0$ from squark decays [120]. From Fig. 6.6 and Table 6.4 it should be clear which kind of event pattern we should expect: two (four) b–jets, five (three) non b–jets and $\not{E}_T$: two b's are coming from h and eventually two others, in brackets, from $t/\bar{t}$ decay. In any case the boost in p_T^{jet} of the $b\bar{b}$ pair from h is big, on average ~ 200 GeV, making jets very close in $\eta - \phi$ plane.

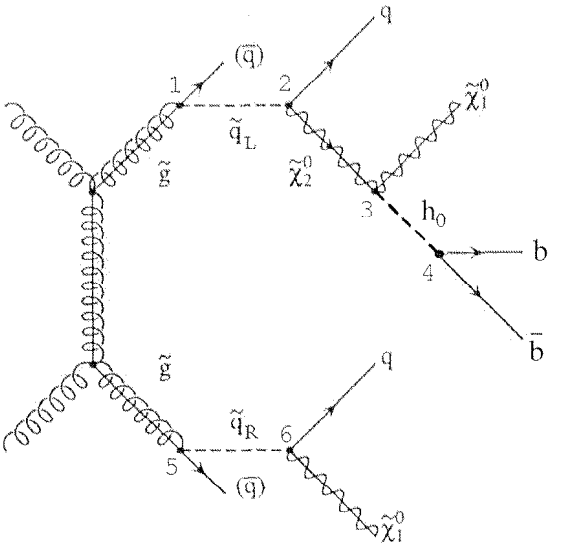


Fig. 6.6: Example of $\tilde{g}$,$\tilde{q}$ decay chain into $h \longrightarrow b\bar{b}$.

6.2.2 Event selection

In this analysis the SUSY event generators ISASUSY and SPYTHIA have been used, PYTHIA 5.7 for SM background events and the CMSJET [121] fast Montecarlo package for reproducing the detector response. The analysis starts with the "parton-level" simulation: ISASUSY or SPYTHIA generate the hard scattering process (e.g. $\tilde{g}\tilde{q}$, $Wb\bar{b}, t\bar{t}, \ldots$), and simulate the appropriate decays. The final state consists of charged leptons, neutrinos, neutralinos and partons on which selection criteria are applied. Main characteristics of the signal and background can be studied. But this procedure leads to too optimistic results since instrumental effects, hadronization/decay or inefficiencies in the jet-clustering algorithm are not taken into account. The more realistic "jet-level" procedure takes into account QED and QCD initial and final state radiation, fragmentation, hadronisation and decay of unstable particles. In the jet-level procedure jets are formed from calorimeter cells of granularity $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, where particle energies are stored with an energy smearing according to $\frac{\Delta E}{E} = 80\%/\sqrt{E} \oplus 6.5\%$, which is the nominal resolution of the hadron calorimeter. Only cells with total transverse energy E_T above a threshold of 1 GeV where used to collect the jet energy in a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.5$ around the cell with the highest E_T. A jet energy and a jet axis is therefore defined.

The Standard Model backgrounds to consider are: QCD $2 \rightarrow 2$ which also includes $b\bar{b}$, $t\bar{t}$, W+jets and Z+jets. These background channels can be divided in three categories:

- irreducible backgrounds: $WZ \longrightarrow q_i\bar{q}_j b\bar{b}$ which is resonant in terms of $m_{b\bar{b}}$ and therefore is of special concern if m_h is close to m_Z; Z+jets with two b jets.
- reducible background with at least two b–jets in the final state: mainly due to $t\bar{t}$ events.

Vertex	Decay Channel	Branching Ratio
1	$\tilde{g} \longrightarrow \tilde{q}_L \bar{q}$	28.4%
2	$\tilde{q}_L \longrightarrow \widetilde{\chi}_2^0 q$	32.5%
3	$\widetilde{\chi}_2^0 \longrightarrow \widetilde{\chi}_1^0 h$	99.6%
4	$h \longrightarrow b\bar{b}$	87.7%
5	$\tilde{g} \longrightarrow \tilde{q}_R \bar{q}$	71.6%
6	$\tilde{q}_R \longrightarrow \widetilde{\chi}_1^0 q$	99.9%

Table 6.4: Branching ratios of the main decay processes involved in strong production of the Higgs boson h. Values are generated using SPYTHIA for the LHC Point 1. Vertex numbers refers to Fig. 6.6.

- reducible background containing jets misidentified as $b-$jets, mainly due to $W+$jets production, but their contribution is small, provided a high $b-$tagging efficiency.

In Table 6.5 the production cross sections of signal and various backgrounds at $\sqrt{s} = 14$ TeV are shown.

Channel	Cross Section (pb)
$\tilde{g}\tilde{g}, \tilde{g}\tilde{q}, \tilde{q}\tilde{q}$	3.34
WW, WZ, ZZ	1.54×10^3
$t\bar{t}$	612
$Z+$jets	103×10^3
$W+$jets	250×10^3
QCD (100-250 GeV)	3.94×10^6
QCD (250-750 GeV)	7.78×10^3
QCD (750-1500 GeV)	88.9
QCD (1500-... GeV)	0.317

Table 6.5: Cross section for signal and SM background values generated by SPYTHIA in LHC Point 1. QCD events are generated in intervals of p_T in order to facilitate accumulation of statistics for the higher p_T ranges. A lower p_T cut for QCD events is set to 100 GeV.

The accumulated statistics for the background corresponds to 100 fb^{-1}, equivalent to 10 years of LHC operation, except for QCD events where a rescaling has been necessary due to its huge cross section, especially at low p_T.

By looking to the event topology of Fig. 6.6, the initial requirements for all the samples have been set as following: ≥ 4 jets with $E_T^{jet} > 20$ GeV within $|\eta^{jet}| < 4.5$, $\not{E}_T > 100$

GeV and sphericity in the transverse plane > 0.1[2]. No requirements on isolated leptons have been set.

In Fig. 6.7 some kinematical distribution of mSUGRA signal compared with SM background are shown. From this is clear that the most efficient background rejection cut is $\not\!\!E_T$. Significant difference comes also from jet multiplicity distributions. After a prelim-

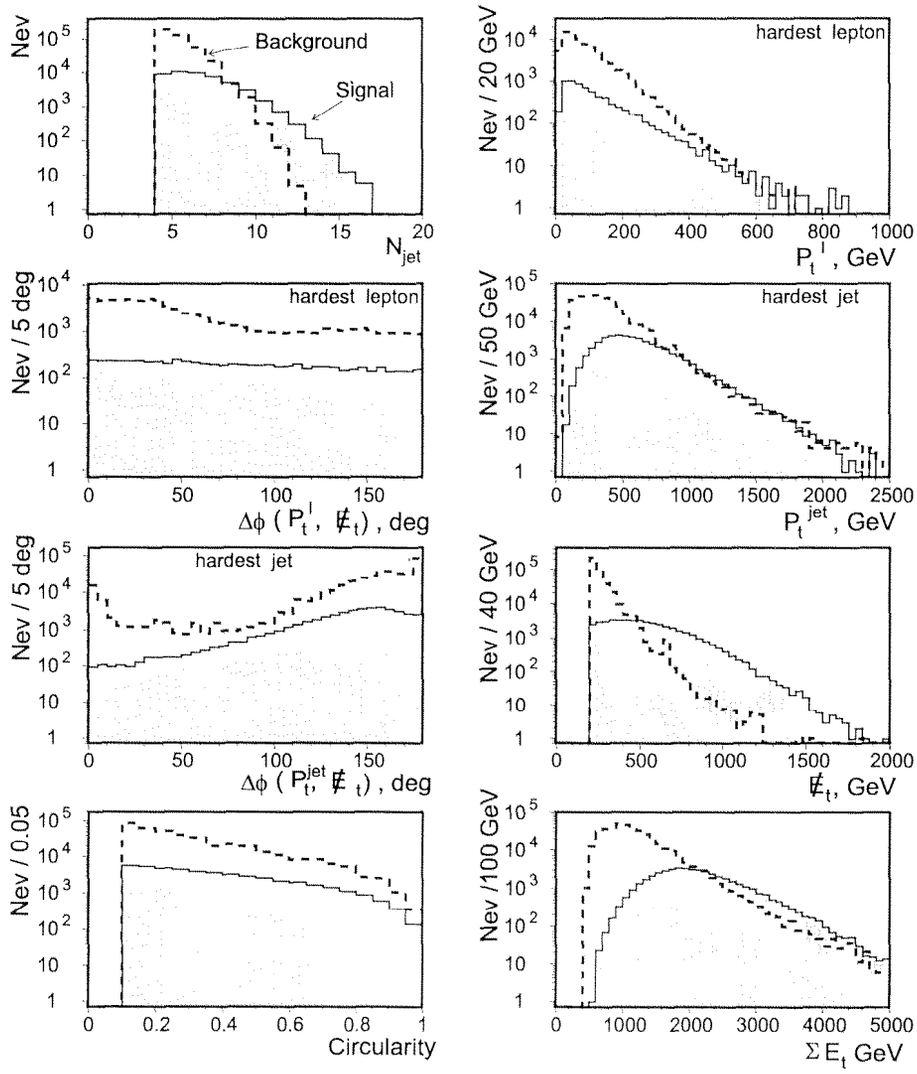


Fig. 6.7: Comparison of mSUGRA signal and SM background distributions. 100 fb^{-1} of integrated luminosity equivalent [122].

inary background analysis for Point 1, the cuts on $\not\!\!E_T$ and p_T^{jet} have been modified as: $p_T^{jet} > 40$ GeV and $\not\!\!E_T > 400$ GeV.

[2]The sphericity in the transverse plane is defined as the ratio $2\lambda_2/(\lambda_1 + \lambda_2)$ of eigenvalues of the sphericity tensor

6.2.2.a b-Tagging

Crucial ingredient for h detection is the efficiency in successfully identify and reconstruct $b-$jets from its decay. To this purpose excellent performances of the Central Tracker (CT) must be combined with a efficient jet energy reconstruction (and correction) in ECAL and HCAL. The key role of the tracking system of CMS for this analysis is summarized in the following requirements: momentum resolution for high-p_T muons (see Fig. 6.9) and isolated electrons and hadrons over its full pseudo-rapidity coverage ($|\eta| < 2.5$) of the order of $\Delta p_T/p_T \sim 0.15 p_T \oplus 0.5\%$ (p_T in TeV), minimum p_T reconstruction threshold set to $1 \div 2$ GeV, good impact parameter resolution (see Fig. 6.8), high granularity to improve pattern recognition and decrease occupancy. The combination of those factors translates

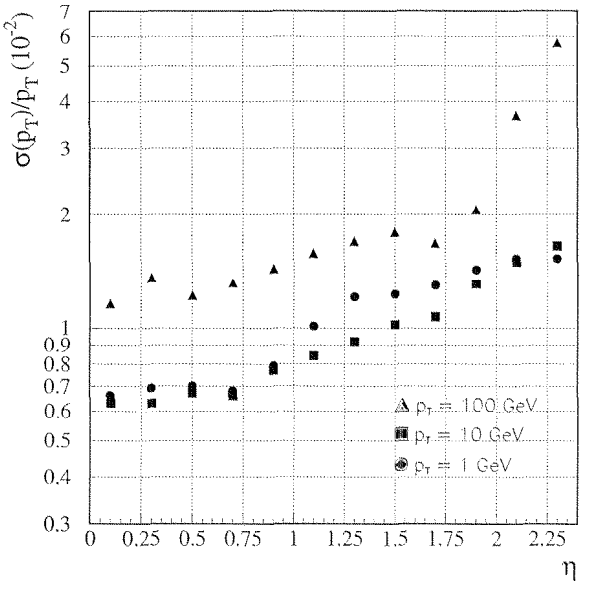


Fig. 6.8: Impact parameter resolution as a function of η. Phase 1 tracker. Ref. [14]

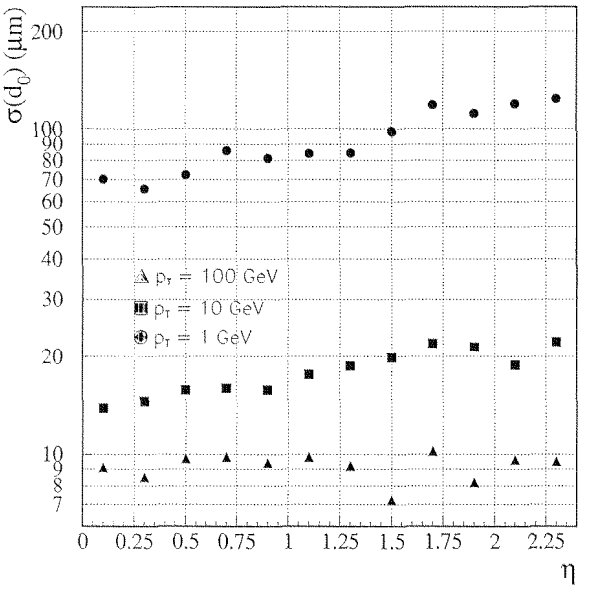


Fig. 6.9: Transverse momentum resolution vs. η for muons of different p_T. Phase 1 tracker. Ref. [14]

in an efficient impact parameter $b-$tagging (see Fig. 6.10 and Fig. 6.11). In these pictures the b-tagging probability and its purity for jets is obtained from a parameterized gaussian impact parameter resolution: crucial is the position of the innermost pixel layer as well as the number of high precision points at higher r (silicon and MSGC), since in $h \longrightarrow b\bar{b}$ decays, ϵ_{b-tag} enters quadratically. The effect of the multiple scattering on the tracker material is taken into account and non-Gaussian tails due to pattern recognition problems or noise are described rescaling the simulated CMS response at $\sqrt{s} = 1.8$ TeV in order to match the relative fraction between gaussian and tails as in present CDF data [123].

For the search of $h \longrightarrow b\bar{b}$ at Point 1 (or close by) the region allowed for $b-$tagging have been restricted to the barrel region, $|\eta| < 1.75$, and only one pair of $b-$tagged particles have been selected, chosen the closest in $\eta - \phi$ space. An eventual jet-charge analysis can

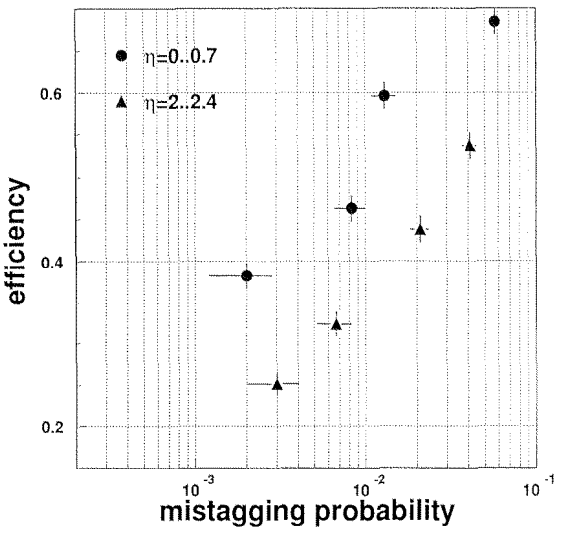


Fig. 6.10: b–tagging efficiency vs. purity for 100 GeV jets in two pseudorapidity regions. Phase 1 tracker. Ref. [14]

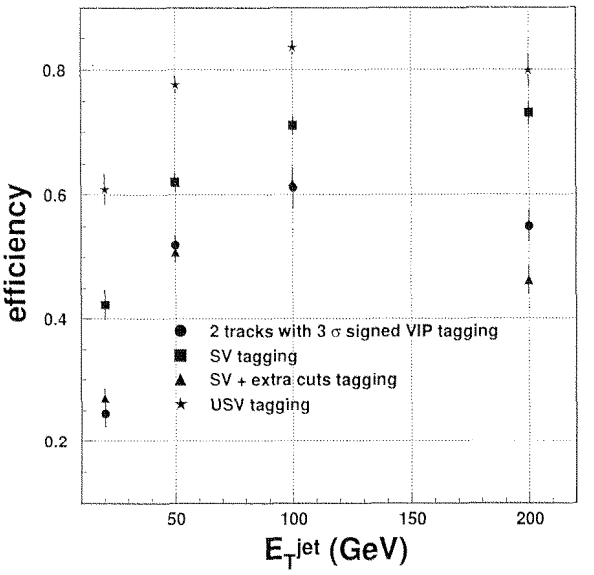


Fig. 6.11: b–tagging efficiency vs. E_T^{jet}, estimated with four different algorithms. Ref. [14]

help in enriching the $b\bar{b}$ sample, improving signal significance by reducing combinatorics [3].

6.2.3 Results

By applying cuts explained above, taking into account masses and branching ratios listed in Tab. 6.3 and Tab. 6.4, assuming 100% b–tagging performances, one find for the $b\bar{b}$ invariant mass distribution the distribution shown in Fig. 6.12. The distribution has been fitted with a sum of a gaussian and a quadratic polynomial. The Higgs peak is clearly visible above SUSY and SM background with a $S/B \simeq 3$ and a significance of $S/\sqrt{B} = 26.7$. The width of the signal peak is ~ 8 GeV, entirely due to the jet energy resolution. The main remaining background is due to SUSY events with wrong b-jet pair association.

The position of the Higgs mass peak is displaced of about 6 GeV towards lower values than the nominal one (theoretical value: $m_h = 89.7$ GeV) since it is not corrected for energy losses in jets (cone size, energy thresholds) and neutrino induced $\not{E}_T$ in b–jets.

A degradation of the HCAL energy resolution of the order of 20% can be tolerated

[3] It is possible to define a quantity called *jet charge* as follows: $Q_{jet} = \frac{\sum_i q_i(\vec{p}_i \cdot \hat{\alpha})}{\sum_i \vec{p}_i \cdot \hat{\alpha}}$ where q_i and $\vec{p}_i$ are respectively charge and momentum of the i^{th} track in the jet and $\hat{\alpha}$ is the versor defining the jet axis [124]. For b–quarks jets, the sign of the jet charge is on average the same as the sign of the b–quark that produced the jet [125], so the sign of the jet charge may be used to improve rejection against same sign b–jets pairs.

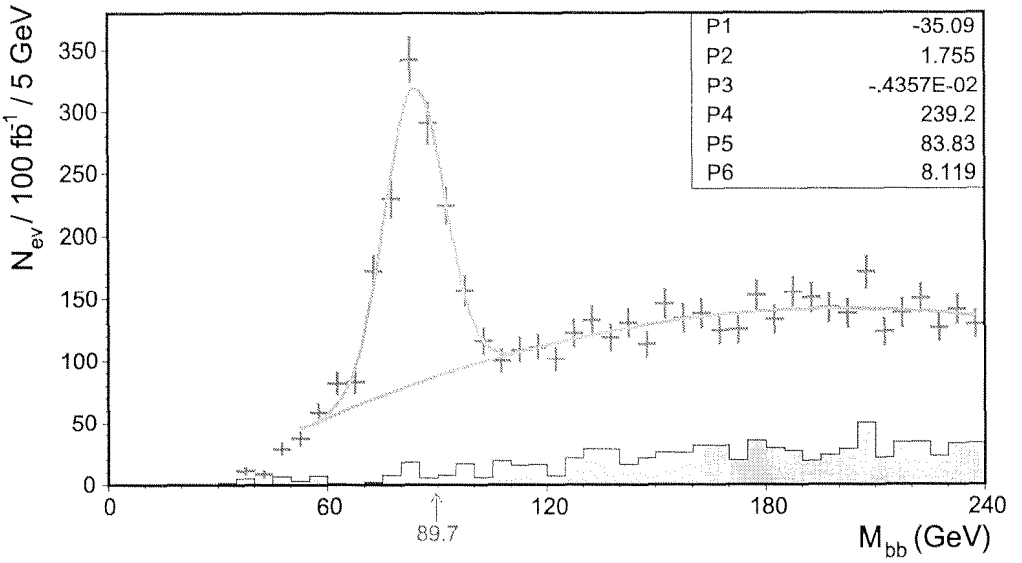


Fig. 6.12: $b\bar{b}$ pairs effective mass distribution for 100% tagging performances. Calculated for $m_0 = m_{\frac{1}{2}} = 500$ GeV, $\mu < 0$, $\tan\beta = 2$. [122]

without spreading out significantly the mass distribution and degrading too much signal quality($\Delta S/\sqrt{B} \simeq 7\%$ for $E_T^{jet} \simeq 100$ GeV), since $b-$jets from h are central, carry a significant p_T and calorimeter granularity assures good jet-jet opening angle measurement.

The influence of $b-$tagging has been studied by varying of $\pm 15\%$ its efficiency, keeping constant the purity at nominal values. Results can be appreciated in Fig. 6.13, keeping as a reference the signal of $S/B \simeq 2$ with a significance of $S/\sqrt{B} \simeq 18.3$ for 10^5 pb^{-1} [122], obtained under nominal condition of $b-$tagging and event selection.

The dependence of signal significance on $b-$tagging efficiency shown in Fig. 6.13 is important, close to $\pm 25\%$, but it is less critical on mistagging probability: it follows that the best recipe for h searches in CMS would be $b-$tagging with pixel points as close as possible to the interaction point even during high luminosity regime, a high number of very precise reconstructed points on silicon strip detectors with the largest lever arm possible.

6.2.4 Conclusions

A preliminary analysis of the possibility of discovery the lightest SUSY Higgs h in a large portion of the mSUGRA parameter space has been performed.

The large production of h in cascade decays of strongly interacting particles may resort in its detection through its decay in $b\bar{b}$ pairs even at early stages of LHC operations (i.e. low integrated luminosity). The $S/B \sim 1$ level of detectability may be reached by cutting almost only on $\not{E}_T$, exploiting the typical SUSY event characteristic. Furthermore, Fig. 6.14 shows that, after 10 years of data taking (i.e. ~ 100fb^{-1}), the 5σ discovery limit

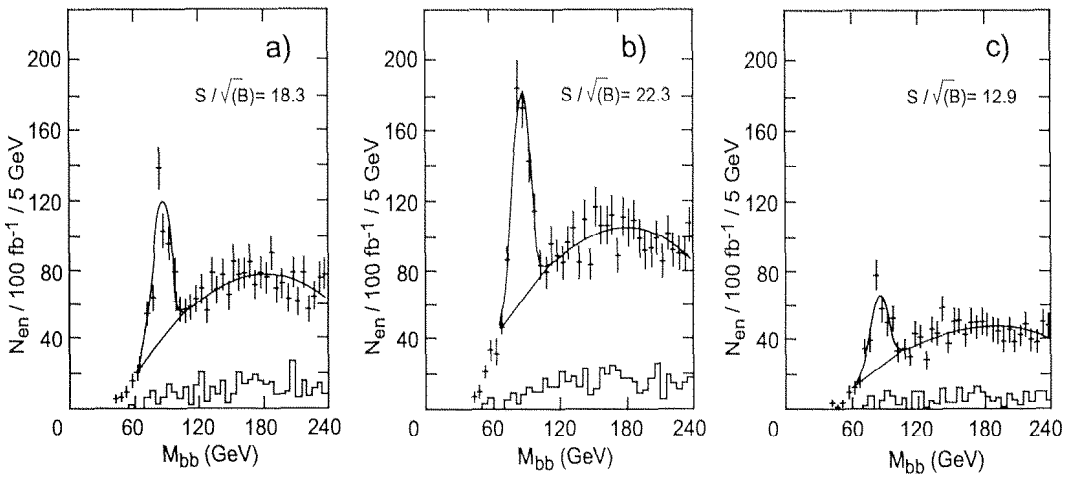


Fig. 6.13: $b\bar{b}$ pairs effective mass distribution for a) nominal, b) nominal+15% and c) nominal−15% b−tagging performances. Calculated for $m_0 = m_{\frac{1}{2}} = 500$ GeV, $\mu < 0$, $\tan\beta = 2$ for the same point of Fig. 6.12. [122]

for $\tan\beta = 2$ and $\mu < 0$ extends over a large part of the $m_0 - m_{1/2}$ parameter plane.

Unambiguous signal identification depends critically on b−tagging: efficiency and mistagging probability can be improved by exploiting the Central Tracker (CT) characteristics, i.e. first pixel layer as close as possible to the interaction region, and its optimal pattern recognition with the 5 layers of silicon strip detectors within the CT.

6.3 Current status of supersymmetry searches

Most of direct limits on sparticle masses come from the non-observation of any SUSY signals at high energy hadron colliders, from the precision measurements of the properties of Z and W bosons in experiments at LEP or at HERA. The agreement of the measured value of Γ_Z with its expectation in the SM, gives model-independent constraints on decays of the Z-boson into any new particles with known $SU(2) \times U(1)$ quantum numbers. Experiments at the Z pole (for each of the LEP experiments ~ 150 pb^{-1} collected data) excluded all visible SUSY particles up to half the Z mass [19]. The measurement of the invisible width of the Z-boson leads the lower bound of $m_{\tilde{\nu}} > 43$ GeV, even when the sneutrinos decay invisibly via $\tilde{\nu} \to \nu\tilde{\chi}_1^0$ [126]. These bounds are relatively insensitive to the details of the model.

The LEP experiments have also directly searched for charginos, sleptons and squarks. These searches assume that the charginos and sfermions decay directly to the LSP. The typical signature of SUSY events is a pair of acollinear leptons, acollinear jets or a lepton-jet pair recoiling against $\not{E}_T$. The non-observation of such spectacular event topologies at LEP1 have led to SUSY masses lower bounds very close to $\frac{M_Z}{2}$. Nevertheless, the new data at higher energies (LEP2) allow much stronger limits to be set [127], but the

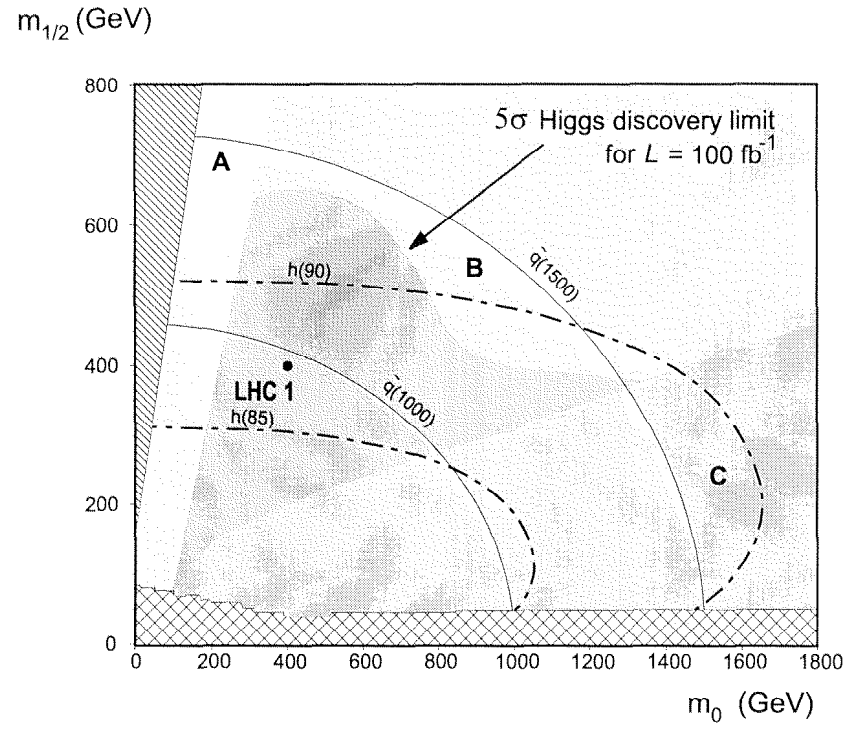


Fig. 6.14: 5σ discovery contours of the lightest SUSY Higgs h for 100 fb^{-1} of integrated luminosity for $\tan\beta = 2$, $\mu < 0$ and $A_0 = 0$. From [123].

complex interplay of masses, cross sections and branching ratios makes simple general limits impossible to specify. A short summary of the present status of SUSY searches at LEP and Tevatron experiments is shown in Table 6.6.

The gauge fermions generally are produced with large cross sections while the scalar particles are suppressed near threshold by kinematic factors. Charginos are produced via γ^*, Z^* or $\tilde{\nu}_e$ exchange. This translates to a lower limit of $\sim$ 85 GeV on the mass of the lightest chargino of the MSSM, independent of the decay patterns of these sparticles. In contrast, even within the MSSM framework, the corresponding bounds on neutralino masses are sensitive to model parameters. This is because in the limit $\mu \gg M_1, M_2$, the lighter neutralinos are dominantly gaugino-like, so that their couplings to the Z-boson are strongly suppressed by electroweak gauge invariance [126]. When Higgsino components dominate the field content of charginos and neutralinos, cross sections are large and insensitive to slepton masses. Present most conservative limit is 24 GeV, independently from m_0.

LEP experiments have also searched for neutralino production via $Z \to \widetilde{\chi}_1^0\widetilde{\chi}_2^0$ and $Z \longrightarrow \widetilde{\chi}_2^0\widetilde{\chi}_2^0$ decays, assuming that $\widetilde{\chi}_2^0 \longrightarrow \widetilde{\chi}_1^0 f\bar{f}$ ($f = q$ or ℓ). The non-observation of acollinear jet or lepton pairs from this process excludes certain regions of the parameter space, but does not (for reasons already explained) lead to an unambiguous lower bound on $m_{\widetilde{\chi}_2^0}$. The region of the $\mu - M_2(m_{\tilde{g}})$ plane excluded by LEP searches for charginos and

neutralinos is illustrated in Fig. 6.15 for two values of $\tan\beta$. The dashed line present in this plot shows the kinematic bound for charginos found from neutralinos searches. Plots on the right side show how limits on charginos weaken when sneutrinos masses are light and correspondingly cross sections weaken.

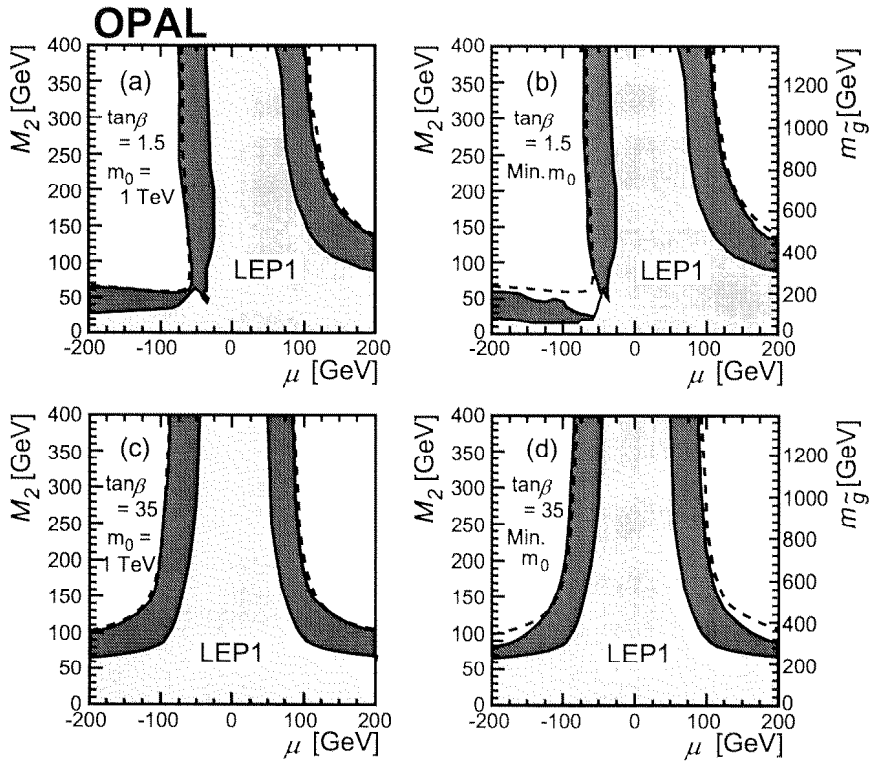


Fig. 6.15: Regions in the μ *vs.* $M_2(m_{\tilde{g}})$ plane excluded by charginos and neutralino searches performed by OPAL collaboration at LEP, for $\tan\beta = 1.5$ and 35. Light shaded region shows the limits derived from Z width, while the dark region shows the additional exclusion from LEP2. From [19].

Although LEP experiments have yielded a limit of $m_{\tilde{q}}$, the search for the strongly interacting squarks and gluinos is best carried out at high energy hadron colliders such as the Tevatron via $\tilde{q}\tilde{q}, \tilde{g}\tilde{q}$ and $\tilde{g}\tilde{g}$ production and their subsequent cascade decays. As pointed out in Sec. 6.1.3.b, the final state consists of several jets plus (possibly) isolated leptons (from $\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^0$ production via their primary decays) and $\not\!\!E_T$ from the two LSPs in each final state. For an integrated luminosity of $\sim$ 110 pb^{-1} that has been accumulated by each of the CDF and D0 experiments at Tevatron, the classic $\not\!\!E_T$ channel offers the best hope for detecting SUSY. The non-observation of an excess of $\not\!\!E_T$ events above SM background expectation has enabled the D0 collaboration to infer a lower limit of $\sim$ 212 GeV on gluino mass if $m_{\tilde{g}} > m_{\tilde{q}}$. For what concerns squarks, by imposing $m_{\tilde{q}} \leq m_{\tilde{g}}$, D0 set the lower limit to 212 GeV analyzing $\not\!\!E_T$ channels. The region of the $m_{\tilde{g}} - m_{\tilde{q}}$ plane excluded by these analyses depends weakly on other SUSY parameters, and is shown in Fig. 6.16 for $\mu = -400$ GeV and $\tan\beta = 4$. As the experiments at the

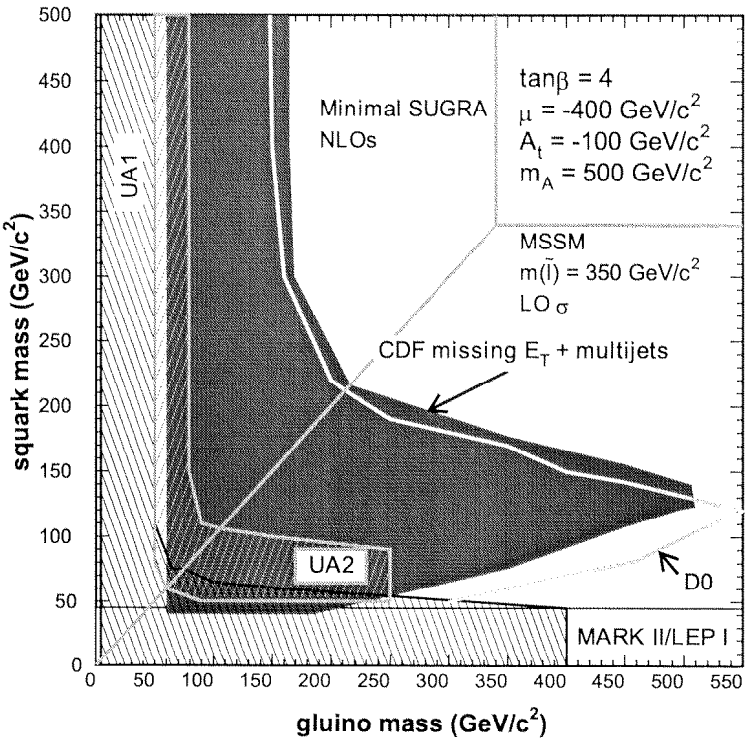


Fig. 6.16: Regions in the $m_{\tilde{g}}$ *vs.* $m_{\tilde{q}}$ plane excluded by searches for $\not\!\!E_T + jets$ events of CDF, D0 and older UA1 and UA2 results. LEP1 and SLC limits are also shown. From [128].

Tevatron continue to accumulate more data, they will also become sensitive to leptonic signals from cascade decays of squarks and gluinos. Although the single lepton signals are overwhelmed by the background from $(W \to \ell\nu) + jet$ events, the multilepton signals offer new ways of searching for Supersymmetry at Tevatron.

Some more remarks upon some constraints from "low energy" experiments and from cosmology [136]. Some judgment must be exercised in evaluating these constraints which, unlike the direct constraints from collider experiments, can frequently be evaded by relatively minor modifications of the model framework. For instance, an overabundance of LSPs produced in the early universe leads to significant restrictions on mSUGRA parameters. This bound can, however, be simply evaded by allowing a small amount of R-parity violation that causes the LSP to decay, although at a rate that has no other impact on particle physics phenomenology. Likewise, constraints from proton decay [137] are sensitive to assumptions about physics at the GUT scale. Supersymmetry also allows for new sources of CP violation in the form of new phases in gaugino masses or SUSY breaking trilinear scalar interactions [138]. Indeed, for sparticle masses ~ 100 GeV, these phases (which are usually set to zero in the MSSM) are limited to be $\sim 10^{-3}$ in order that the induced electric dipole moment of the neutron or electron not exceed its experimental upper limit. If, however, these phases are zero at some ultra-high unification scale, it has been checked that their values at the weak scale induced via renormalisation group evolution do not lead to phenomenological problems. This only pushes the problem to

Particle		Condition	Limit (GeV)	Experiment
$\widetilde{\chi}_1^{\pm}$	gaugino	$m_{\tilde{\nu}} > 200$ GeV, $\tan\beta = 1.41$	85.5	ALEPH 189[129]
	Higgsino	$\forall m_0$, $\tan\beta < 1.41$	69.2	L3 172[130]
$\widetilde{\chi}_1^0$		$\forall m_0$, $\tan\beta = 1$, $m_{\tilde{\nu}} > 43$ GeV	24.2	OPAL 196[131]
		$\tan\beta < 10$, large m_0	34.8	L3 196[132]
		$\forall A_0$, $\tan\beta < 10$	36	LEP 189 combined [132]
$\tilde{e}_R$	$e\widetilde{\chi}_1^0$	$\Delta M > 15$ GeV	89	LEP 189 combined [132]
$\tilde{\mu}_R$	$\mu\widetilde{\chi}_1^0$	$\Delta M > 15$ GeV	84	LEP 189 combined [132]
$\tilde{\tau}_R$	$\tau\widetilde{\chi}_1^0$	$m_{\widetilde{\chi}_1^0} < 20$ GeV	71	LEP 189 combined [132]
$\tilde{\nu}$			43	LEP1 combined [132]
$\tilde{t}_1$	$c\widetilde{\chi}_1^0$	$\forall\theta_{mix}$, $\Delta M > 10$ GeV	90	LEP 189 combined [132]
	$b\ell\tilde{\nu}$	$\forall\theta_{mix}$,$\Delta M >$ 7 GeV	90	LEP 189 combined [132]
$\tilde{b}_1$	$b\widetilde{\chi}_1^0$	$\forall\theta_{mix}$	81	LEP 189 combined [132]
$\tilde{g}$	jets+$\not{E}_T$	$m_{\tilde{g}} \geq m_{\tilde{q}}$	212	D0 [133]
	dileptons	$\forall m_{\tilde{q}}$	180	CDF [134]
	dileptons	$\tan\beta < 10$, $m_0 < 300$ GeV	129	D0 [132]
$\tilde{q}$	jets+$\not{E}_T$	$m_{\tilde{q}} \leq m_{\tilde{g}}$,$\tan\beta = 2$	212	D0 [133]
	dileptons	$m_{\tilde{q}} = m_{\tilde{g}}$ $\tan\beta = 4$, $\mu = -400$ GeV	216	CDF [135]
	dileptons	$\tan\beta < 10$, $m_0 < 300$ GeV	138	D0 [132]

Table 6.6: Lower limits on SUSY particle masses. Only R-parity conserving models are taken into account.

the unification scale where the physics is as yet speculative.

There are also constraints from the universality of the charged-current and neutral-current weak interactions. The Cabibbo universality between the μ-decay and β-decays put constraints only on rather light sparticles $\lesssim 20$ GeV [136]. Z-decay partial widths into different species of light fermion are more sensitive than low-energy experiments. Non-decoupling effects in ρ-parameter or $\Gamma(Z \to b\bar{b})$ are relatively sensitive to the virtual exchange of sparticles. Indeed, it has been claimed that the experimental value of $R_b = \Gamma(Z \to b\bar{b})/\Gamma(Z \to \text{hadrons})$ prefers a light top squark and chargino [136]. These measurements do not currently lead to any significant restrictions on sparticle masses.

Finally, some words about the flavor violating inclusive decay $b \to s\gamma$ recently measured by the CLEO collaboration [139]. Even within the minimal SUSY framework, there are several additional contributions to this amplitude. Of course, the agreement of the SM computations with the experimental data lead to an interesting limit (within theoretical and experimental errors) on the sum of various new physics contributions. Since it is possible for these new contributions to (partially) cancel over a significant range of model parameters, these measurements do not lead to unambiguous bounds on the masses of various sparticles. Like the neutralino search at LEP, they do, however, exclude significant regions of parameter space. In particular if one follows the implications for this branching ratio measurement and the LEP2 Higgs searches [140], one might find that either mSUGRA model is excluded or, if a Higgs is found at LEP2, this measurement may exclude $\mu < 0$ and constrain $\tan\beta$ between 2 and 4. Experiments at B-factories will decrease the current error on the branching ratio by at least a factor 4, thus clarifying the scenario.

In conclusion, a wide variety of empirical constraints have served to restrict the parameter ranges of supersymmetric models. It is, however, interesting that even the simplest, highly constrained supergravity GUT model, is consistent with all experimental data including those from cosmology.

Appendix A

The High-Low junction currents

In this Appendix are reported the main calculations for the derivation of the *high-low* junction properties. The effect of the parameterisation of the back-side recombination velocity will be highlighted as well as its impact on the modeling of the high region.

In Fig. (A.1) the device is sketched: all the dimensions and the parameters used later on are highlighted.

Following the SHR modeling for generation and recombination [29], in order to find the minority carrier concentration in the quasi neutral $n-$region (here C), one should solve the continuity equation obtained from Eq. (3.27) and Eq. (3.29) by means of low injection hypothesis (i.e. external disturbances of equilibrium does not appreciably change the total free carrier density from its equilibrium value). This implies the term $(G_p - R_p)$ in Eq. (3.29) to be equal to $\Delta p_n/\tau_p$, where Δp_n is the excess of minority carriers (holes) region C and τ_p is their recombination lifetime.

$$\frac{\partial p_n}{\partial t} = D_{p_n}\frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{\tau_{p_n}}$$

Putting $\Delta p_n = (p_n - p_{n0})$, assuming constant doping profiles along x, the steady state equation is:

$$D_{p_n}\frac{\partial^2 \Delta p_n}{\partial x^2} - \frac{\Delta p_n}{\tau_p} = 0 \tag{A.1}$$

where D_{p_n} is defined in Eq. (3.17).

From that we observe that the diffusion constant D_p is doping dependent and it will be different for the high and low region. Eq. (A.1) has the simple exponential solution

$$\Delta p_n(x) = Aexp\left(-\frac{x - x_n}{\sqrt{D_{p_n}\tau_{p_n}}}\right) + Bexp\left(\frac{x - x_n}{\sqrt{D_{p_n}\tau_{p_n}}}\right) \tag{A.2}$$

where A and B are constants determined by boundary conditions of our problem. In order to solve Eq. (A.1), we can apply the idea of Boltzmann statistics for the distribution of

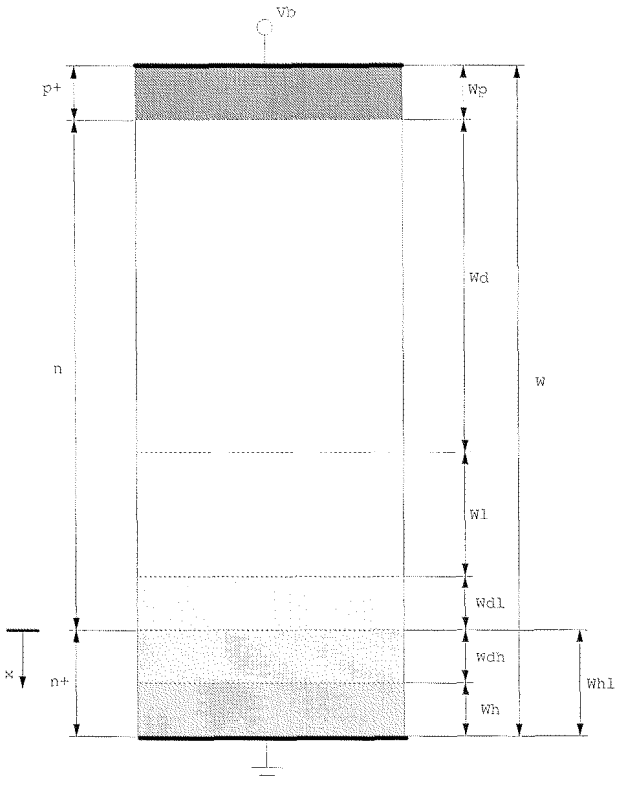


Fig. A.1: Detector Cross Section with regions symbols. W is the total width, W_{hl} is the width of the high-low region, W_d,W_{dl} and W_{dh} are the widths of the space charge regions belonging respectively to the p^+n, low and high region of n^+n junction, W_l and W_h are the thickness of neutral regions in the lowly and highly doped $n-$part. The origin of the x-axis is taken at the *high-low* interface.

carriers at the main junction:

$$-\Delta p_{nl}(-W_l - W_{dl}) = \left[1 - exp\left(-\frac{qV_{pn}}{kT}\right)\right] p_{nl0} = \beta p_{nl0} \tag{A.3}$$

where $p_{nl0} = \frac{n_i^2}{N_{dl}}$ represents the equilibrium minority carrier concentration at the edge of the space charge region of the *low* zone and V_{pn} is the reverse bias across the main junction. As other boundary conditions we can solve Eq. (A.3) at the edges of the space charge region of the high-low junction:

$$-\Delta p_{nl}(-W_{dl}) = \beta_{hl} p_{nl0} \tag{A.4}$$

$$-\Delta p_{nh}(-W_{dh}) = \beta_{hl} p_{nh0} \tag{A.5}$$

with $p_{nh0} = \frac{n_i^2}{N_{dh}}$. In these equations β_{hl} is unknown. The assumption of full ionisation of dopants, $n_{nl} \simeq N_{dl}$ and $n_{nh} \simeq N_{dh}$ helps in solving Eq. (A.4) and Eq. (A.5), once we assume in this region also the validity of Boltzmann distribution for holes, since:

$$\beta_{hl} = \left[1 - exp\left(-\frac{qV_{hl0}}{kT}\right)\right] = 1 - \frac{n_{nl}}{n_{nh}}. \tag{A.6}$$

where V_{hl0} is the built-in potential for the *high-low* junction at equilibrium, in the same way ϕ_i is it at the metal-semiconductor contact. Then:

$$V_{hl0} = \frac{kT}{q} \ln \frac{n_{nh}}{n_{nl}} = \frac{kT}{q} \ln \frac{N_{dh}}{N_{dl}} \tag{A.7}$$

From Eq. (A.7) we conclude that V_{hl0} can only change if there is a significant change in majority carrier concentration in one of the two sides of the high-low junction. But, assuming low injection regime from the depleted bulk, $p_{nl} \ll n_{nl}$, it results that n_{nl} and n_{nh} will remain unchanged, keeping V_{hl0} constant even in non-equilibrium situations. We can conclude that under depletion, the reverse voltage drop is entirely sustained by the main p^+n junction.

Using typical values for doping densities used in the silicon detector technology ($N_{dh} \approx 10^{20} cm^{-3}$ and $N_{dl} \approx 10^{12} cm^{-3}$) we get $V_{hl0} = 0.23$V. This is the potential barrier a hole coming from the depleted region B should overcome in order to be collected by the electrode behind.

At this point it is worth to note that the ratio of minority carriers (holes in this case) across the *high-low* junction is the responsible for its reflecting properties of the junction. In fact, assuming that the change in minority carrier quasi-Fermi level is negligible across the interface, then

$$p_{nh} = p_{nl} \exp\left(\frac{-qV_{hl}}{kT}\right). \tag{A.8}$$

If high injection does not occur in the lightly doped $n-$side, V_{hl} is approximately at equilibrium and then, putting Eq. (A.7) in Eq. (A.8), we get:

$$p_{nh} \simeq p_{nl} \frac{N_{dl}}{N_{dh}}. \tag{A.9}$$

This expression shows that the minority carrier concentration is reduced by a factor N_{dl}/N_{dh} in going from n to n^+ side of the *high-low* junction. Eq. (A.9) is valid only if the injected minority carrier densities are small as compared with the majority carrier densities: in order to extend the domain of application to high injection, Eq. (A.8) must be combined with the assumption of charge neutrality at the edges of the junction depletion region [45]:

$$p_{nh} = p_{nl} \left(1 + \frac{p_{nl}}{N_{dl}}\right) \frac{N_{dl}}{N_{dh}}. \tag{A.10}$$

Coming back to the steady-state continuity equation, Eq. (A.3), we can solve it with the help of Eq. (A.4) and Eq. (A.5), getting [141]:

$$-\Delta p_{nl}(x) = \frac{p_{nl0}}{\sinh\left(\frac{W_l}{L_{pl}}\right)} \left[\beta \sinh\left(\frac{x + W_{dl}}{L_{pl}}\right) + -\beta_{hl} \sinh\left(\frac{x + W_{dl} + W_l}{L_{pl}}\right)\right] \tag{A.11}$$

$$-\Delta p_{nh}(x) = \beta_{hl} p_{nh0} \frac{\frac{D_{ph}}{L_{ph}} \cosh\left(\frac{x-W_{hl}}{L_{ph}}\right) - s_b \sinh\left(\frac{x-W_{hl}}{L_{ph}}\right)}{\frac{D_{ph}}{L_{ph}} \cosh\left(\frac{W_{hl}}{L_{ph}}\right) + s_b \sinh\left(\frac{W_{hl}}{L_{ph}}\right)} \tag{A.12}$$

where s_b is the effective recombination velocity for the region and ρ_{hl} is:

$$\rho_{hl} = \frac{\beta}{\beta_{hl}} = \cosh\left(\frac{W_l}{L_{pl}}\right) + \frac{p_{nho} L_{pl} D_{ph}}{p_{nlo} L'_{ph} D_{pl}} \sinh\left(\frac{W_l}{L_{pl}}\right) \tag{A.13}$$

where L_{ph} and L_{pl} are the diffusion lengths in the high and low regions, defined as $\sqrt{D_{px}\tau_{px}}$, and L'_{ph} is the effective diffusion length for the high region defined as:

$$L'_{ph} = L_{ph} \frac{\cosh\left(\frac{W_l}{L_{ph}}\right) + s'_{hl} \sinh\left(\frac{W_{hl}}{L_{ph}}\right)}{\sinh\left(\frac{W_{hl}}{L_{ph}}\right) + s'_{hl} \cosh\left(\frac{W_{hl}}{L_{ph}}\right)}. \tag{A.14}$$

These results can be compared with the situation where no *high-low* junction is present [45]:

$$-\Delta p_n(x) = \beta_{hl} p_{nh0} \frac{\frac{D_{pl}}{L_{pl}} \cosh\left(\frac{x-W}{L_{pl}}\right) - s_b \sinh\left(\frac{x-W}{L_{pl}}\right)}{\frac{D_{pl}}{L_{pl}} \cosh\left(\frac{W-W_d}{L_{pl}}\right) + s_b \sinh\left(\frac{W-W_d}{L_{pl}}\right)} \tag{A.15}$$

where $W - W_d = W_l$ (we neglect here the thickness of the p^+−region).

Since in the quasi-neutral regions C and D the field is negligible, the minority carrier diffusion current in the high-low region can be calculated from Eq. (3.27) as:

$$J_p = -qD_{pl} \frac{\partial(\Delta p_{nl})}{\partial x} \tag{A.16}$$

Using Eq. (A.11) and Eq. (A.13), J_{hl} becomes:

$$J_{hl} = -q\beta \frac{p_{nlo} D_{pl}}{L_{pl}} \left[\frac{\cosh\left(\frac{W_l}{L_{pl}}\right) - \frac{1}{\rho_{hl}}}{\sinh\left(\frac{W_l}{L_{pl}}\right)} \right] \tag{A.17}$$

We can calculate this current in the case of the *low* region:

$$J_l = -q\beta \frac{p_{no} D_{pl}}{L'_{pl}} \tag{A.18}$$

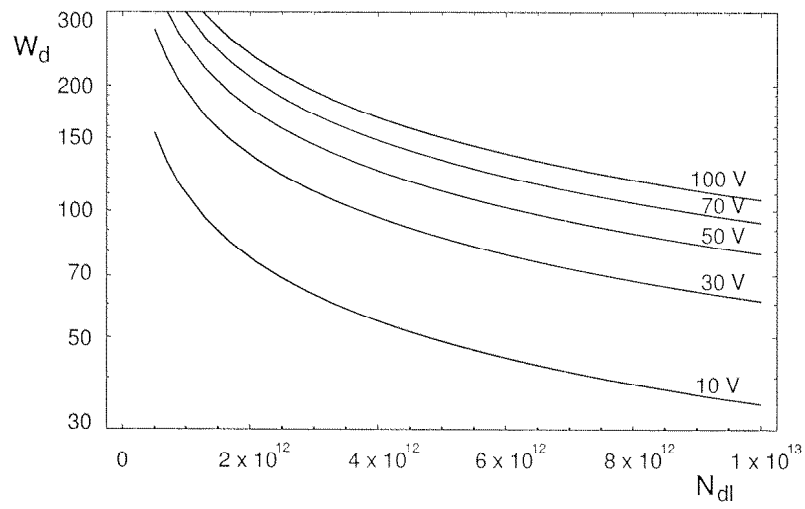


Fig. A.2: W_d as a function of wafer doping. Different curves correspond to different value for depletion voltage.

where L'_{pl} is the effective diffusion length for the low region [142], $W_b = W - W_d - W_p \approx W - W_d$ and it is defined as:

$$L'_{pl} = L_{pl} \frac{\cosh\left(\frac{W_b}{L_{pl}}\right) + s'_b \sinh\left(\frac{W_b}{L_{pl}}\right)}{\sinh\left(\frac{W_b}{L_{pl}}\right) + s'_b \cosh\left(\frac{W_b}{L_{pl}}\right)} \tag{A.19}$$

with $s'_b = s_b \frac{L_{ph}}{D_{ph}}$, the effective recombination velocity in the low region.

In order to compare the two current densities, we must know the width of the space charge region [1] [32]:

$$W_{hl} = \sqrt{\frac{V_{hl}\epsilon_{Si}}{2qN_{dl}}} \tag{A.20}$$

which can be compared with the Eq. (3.42) (cf. Fig. A.3).

From this formula it follows that the width of the depleted region in the high-low region depends only on doping concentration of the low region (thus on bulk resistivity), and not on applied bias, as happens in $p-n$ junctions.

For silicon microstrip detectors it is more interesting to look in the situation near full depletion (i.e. $W_l = W - W_d \approx 0$, then $W_l \ll L_{pl}$), where the full wafer thickness participate to the signal generation caused by an ionizing particle passing through it. We will analyze both cases in which the ohmic character of the contact is varied, passing from

[1] It should be noted that the space charge region in a high-low junction is not really a space-charge region as in the *pn* junction. Following the approach that Gunn used in his work, it is a kind of accumulation region in the low doped side compensated by a depletion in the high region.

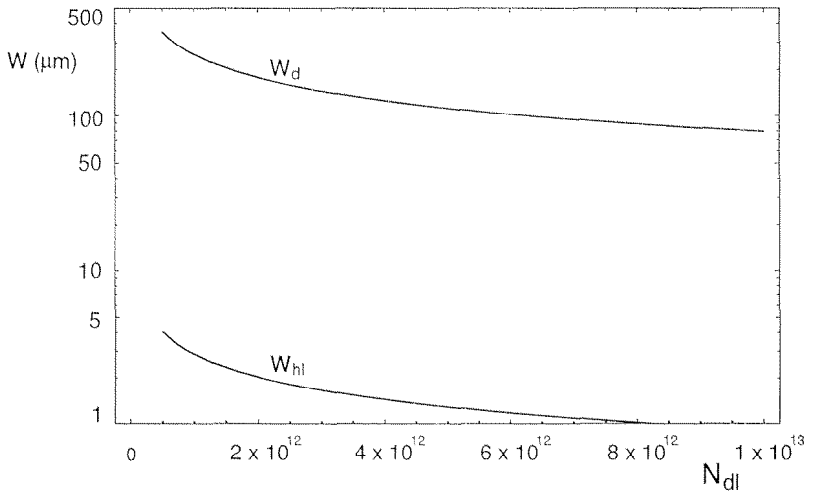


Fig. A.3: W_d (top curve) and W_{hl} (bottom curve) as a function of wafer doping. Reverse bias of $50V$ is assumed.

the perfect one ($s_b \to \infty$) to the one in which recombination occurs at the interface ($s_b =$ finite).

First the case in which we assume a perfect ohmic contact on the backplane. We can simplify Eq. (A.13) and Eq. (A.17) by using Taylor series expansion for sinh and cosh terms around $\frac{W_l}{L_{pl}} = 0$:

$$\rho_{hl} \approx 1 + \frac{1}{2}\left(\frac{W_l}{L_{pl}}\right)^2 + \frac{p_{nh0} D_{ph} L_{pl}}{p_{nl0} D_{pl} L_{ph}} \coth\left(\frac{W_l}{L_{pl}}\right) \tag{A.21}$$

$$J_{hl} \approx -q\beta \left\{ \frac{p_{nl0} D_{pl}}{W_l}\left(\frac{W_l}{L_{pl}}\right)^2 + \frac{p_{nh0} D_{ph}}{L_{ph}}\left[1 - \left(\frac{W_l}{L_{pl}}\right)\right]^2 \coth\left(\frac{W_h}{L_{ph}}\right)\right\}. \tag{A.22}$$

For $W_l \to 0$ (i.e. full depletion, $W_d \to W$), J_{hl} saturates at:

$$J_{hl}^{max} = -q\frac{p_{nh0} D_{ph}}{L_{ph}} \coth\left(\frac{W_h}{L_{ph}}\right). \tag{A.23}$$

Since $p_{nh0} = n_{ie}^2 / N_{dh}$[143], this current density can be made smaller by simply increasing N_{dh}. The same applies for diffusion length and diffusion constant used in the high region.

It is interesting to define a quality factor $\mathcal{M}$ defined as the ratio of the hole current in the same structure with or without a n^+ layer, measuring then the effectiveness of the backside impurity concentration:

$$\mathcal{M} = \frac{J_{hl}}{J_l} = \frac{n_{ie}^2 N_{dl} D_{ph} W_l}{n_i^2 N_{dh} D_{pl} L_{ph}} \coth\left(\frac{W_h}{L_{ph}}\right) \tag{A.24}$$

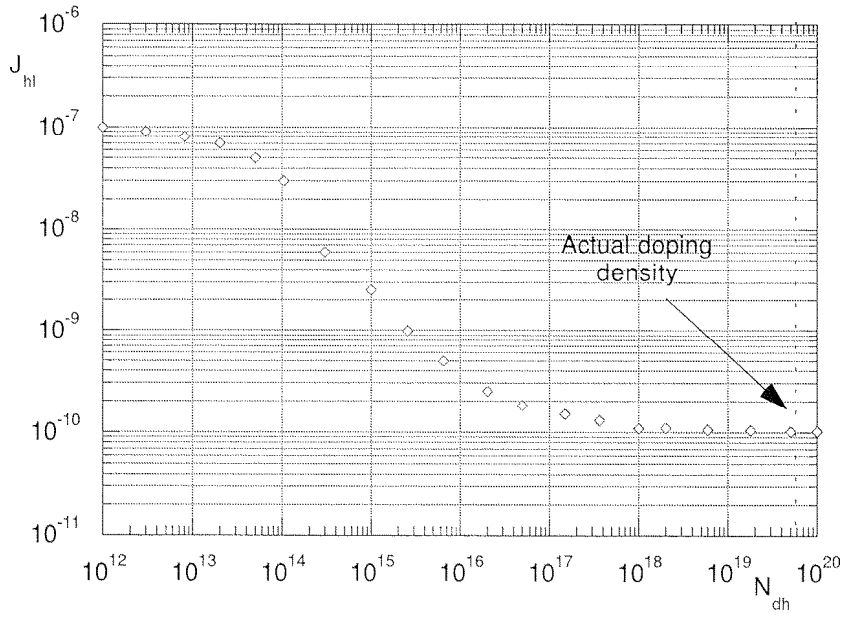


Fig. A.4: $J_{hl}[A/cm^{-2}]$ as a function of doping concentration of the *high-low* junction for an applied bias of $70V$. Ideal ohmic contact is assumed. The substrate concentration is assumed to be $8 \cdot 10^{11} cm^{-3}$ and its thickness is 300μm.

From Eq. (A.24) is clear that J_{hl} is minimized by increasing N_{dh} and W_h/L_{ph}, using therefore a highly doped, deep n^+ layer: physically it means that one has to avoid any holes from reaching the low region by making the width of the high region much larger than the hole diffusion length in this region. The presence of such a high doped layer avoids a strong increase of the leakage current driven by J_l and furthermore such *high-low* junction is responsible for the suppression of any carrier injection from the backside contact.

In Fig. A.4 is plotted J_{hl} as a function of the doping concentration of the n^+ layer for $300\mu m$ thick substrate assuming $V_d = 70V$. The plot show the increasing importance of the n^+ layer with increasing the doping concentration and the width of the depletion.

Let us analyze the case of a non-ideal ohmic contact, where the n^+-metal surface has a higher concentration of carriers as compared to equilibrium, enhancing the recombination. The first consequence is a smaller gradient in carrier concentration in the neutral regions which gives rise to a bigger diffusion length (effective diffusion length, L'_{ph} or L'_{pl}) and thus a smaller diffusion current, cf. Eq. (3.29). Considering again the situation near full depletion, the expressions for the minority carrier current (on backside, holes) are:

$$J_{hl} = -q\beta \frac{p_{nho} D_{ph}}{L'_{ph}} \tag{A.25}$$

$$J_l = -q\beta \frac{p_{nlo} D_{pl}}{L'_{pl}} \tag{A.26}$$

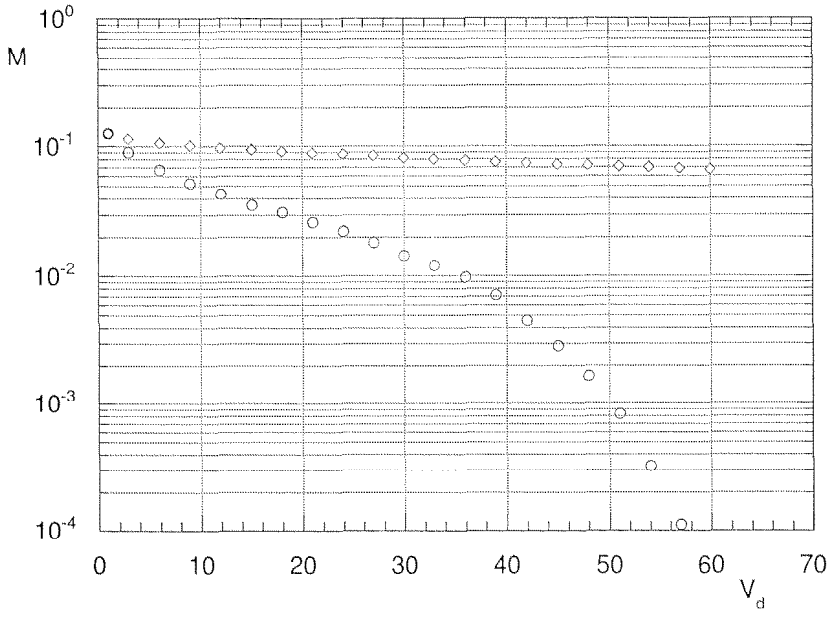


Fig. A.5: $\mathcal{M}$ as a function of depletion voltage V_d. The circles represent the case of $N_{dl} = 8 \cdot 10^{11} cm^{-3}$ and $N_{dh} = 1 \cdot 10^{20} cm^{-3}$; the diamonds $N_{dl} = 8 \cdot 10^{12} cm^{-3}$ and $N_{dh} = 1 \cdot 10^{20} cm^{-3}$. Ideal ohmic contact is assumed.

We can also update the expression of Eq. (A.24) with these conditions :

$$\mathcal{M}' = \frac{n_{ie}^2 N_{dl} D_{ph} L'_{pl}}{n_i^2 N_{dh} D_{pl} L'_{ph}} \tag{A.27}$$

In comparing Eq. (A.27) with Eq. (A.24) $\mathcal{M}'$ reach a minimum at a threshold which depends from N_{dh}[2] and is strongly influenced by the behaviour of the ratio L'_{pl}/L'_{ph}.

[2]For the sake of precision, depends from N_{dl}/N_{dh} which is a strong function of the backside surface recombination velocity s_b.

Appendix B

Silicon detectors optimized for optical transparency

In this appendix is described a study about the possibility to align several plans of Silicon microstrip detectors by means of infra-red (IR) laser tracks measured by the detectors themselves [144]. We performed this study for the AMS (Alpha Magnetic Spectrometer) [145] collaboration, being the silicon wafers designed [146] and produced by CSEM.

For the alignment of Si particle detectors, straight tracks can be simulated by IR light rays at $\lambda \approx 1100$nm. In order to use this technique, the silicon reflectivity has to be reduced considerably. The antireflective materials deposited atop a microstrip detector must satisfy several stringent conditions, and are in practice restricted to SiO_2 and Si_3N_4. For the AMS experiment, an antireflective process implying two supplementary masking steps has been defined. For blanket test wafers, the reflectivity has typically been reduced from 35% to less than 3%. Reflectivities of less than 0.5% have been measured for the best wafers. The optical and electrical effects of these changes on standard AMS detectors are presented.

B.1 Introduction

Within the frame of the AMS project, a way of significantly reducing the optical reflectivity of microstrip detectors in the infrared range has been developed. This should allow using laser beams for simulating straight tracks for detector alignment. This study concentrates on the technological issues affecting antireflective coating atop microstrip detectors. The related modifications of the fabrication process are fully modular and the tests in production conditions show that the detector reflectivity can be reduced down to the 1% range, without significantly affecting the electrical behaviour of the detectors.

B.2 Problem description

B.2.1 AMS solution

AMS is a space-physics experiment meant to detect antimatter and dark matter. The main detector is based on a silicon tracker made of 6 different silicon planes, with a total surface of 2.3 m^2. To reach the expected precision of about 10μm , the 6 silicon planes must be relatively registered with a precision of 2-5μm . This is obtained with the help of 4 laser beams which are shone through the 6 silicon planes (artificial straight tracks). The microstrip detectors are used to sense and to measure the position of the beams. The results are then fed to the correction software.

B.2.2 Antireflective coating

At each vacuum-matter interface, light will be partially reflected. For instance, the reflectivity of bare silicon is about 30% at 1083 nm. This means that after 5 planes nearly all the incident light will have been reflected one or several times, will interfere with other parts of the detector, and significantly degrade the alignment performance. Computations performed at Aachen show that to get a clean alignment, the reflectivity of the detector should be lowered in the 1% range.

B.3 Material & Processes

Multilayer antireflective coatings are well-known, and for standard optical devices it is customary to get reflectivities of less than 0.5% at a given wavelength. However, in the case of a microstrip detector the coating has to satisfy a number of supplementary conditions:

B.3.1 Material compatibility

First, the antireflective layer has to be compatible from the material point of view. Standard layers for optical coating would be for instance magnesium fluoride or lanthanum oxide. Such materials would require a lengthy characterization before being accepted in a production facility. The material has to be available. It means practically that it may be deposited with existing equipments and preferably within the clean room facility. The antireflective layer should be compatible with the existing process. For instance, etching the new layers should not destroy the underlying patterns. Finally device compatibility should be ensured. Therefore the basic device performances should not be significantly altered by the presence of the antireflective coating.

B.3.2 Proposed process

After studying all these conditions, we have chosen a simple double layer antireflective coating, made of silicon oxide and silicon nitride. At the considered wavelength (1083 nm), these materials have refractive indexes of respectively 1.45 and 2.05. The lower layer is a slight modification of something which is normally present on the microstrip detector, that is thermal and CVD oxide The second and supplementary layer is CVD nitride, with a thickness of the order of 85 nm. In the normal AMS process, which has up to now been used to produce some 800 detectors, the situation is reached at a certain point: the silicon wafer, with its implanted strips, is covered by a 100 nm thermal implantation oxide and by a 700 nm CVD oxide. The next process step consists in etching the contact holes down to the silicon. The process then goes on with the deposition and the definition of the metal layer. For the low-reflectivity devices, the operations are identical up to the oxide deposition point, except that the CVD layer thickness is adapted to the optical conditions. This happens to correspond to a modification of about 10%, and has no influence neither on the technology nor on the device. Then a thin layer of CVD nitride is deposited on the wafer. This nitride layer is removed from the wafer everywhere except where it is really needed, that is in the areas where the laser beam will illuminate the detector. We remove the nitride for two reasons:

- To minimize the effect on the nitride charges on the detector behaviour. These charges are known to increase leakage currents.
- To be able to use the same etching conditions to define the contacts, that is to ensure process compatibility.

It should be noted here that this approach is modular, in the sense that set of low-reflectivity operations are simply added to an otherwise standard process. The others manufacturing steps are not affected by this addition.

B.3.3 Thickness computation

The standard optical theory [147] yields an analytical formula allowing to compute the reflectivity of the deposited layers as a function of their thicknesses and refractive indexes (see Fig. B.1). The resulting function is periodical: the minimum which is most suited to the deposition conditions may be chosen. It should be seen that the major difficulty does not lie in finding the minimum position, but in getting a sufficient control. Optical layers require an absolute thickness precision, which relates to the wavelength and which proves difficult to reach for thick layers like the SiO_2. Here in practice the SiO_2 layer is measured on each wafer after CVD deposition, and the nitride thickness is adapted to obtained value.

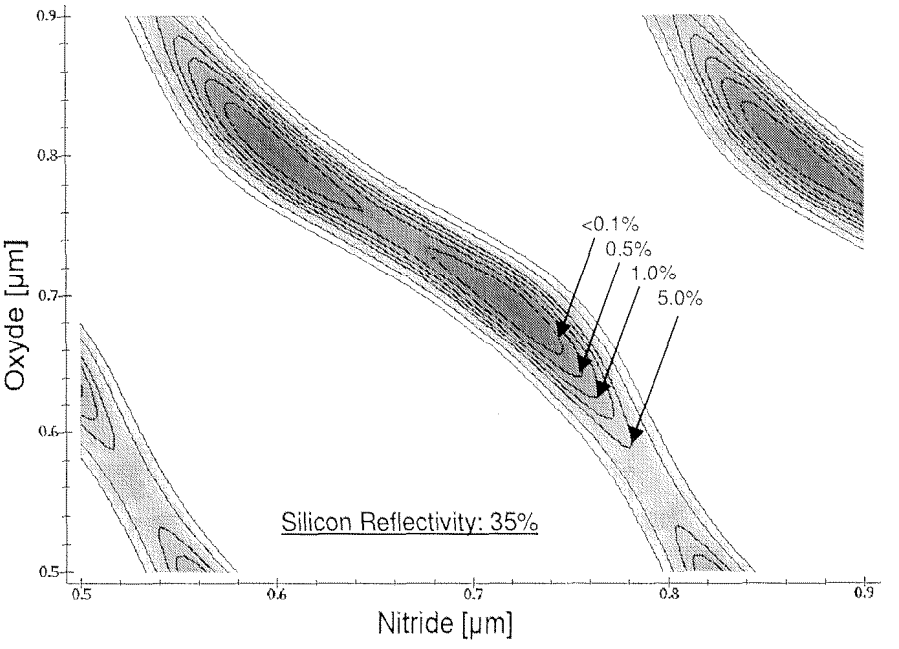


Fig. B.1: Reflectivity of the deposited layers as a function of their thicknesses and refractive index.

B.4 Optical results

To test the soundness of the concept, we have first integrated a blanket test lot, which comprised the right oxide thicknesses, but was otherwise not structured. The several layer thicknesses, together with the reflectivity were then measured at Aachen with a Perkin-Elmer Lambda 2 spectrograph. The results are shown on Fig. B.2. Two conclusions may be drawn from Fig. B.2:

- The targeted minimum was reached with a reasonable approximation,
- For most samples the reflection coefficient has been lowered to the 1% range.

In a next step, half of the wafers of an AMS production lot (lot AMS/63) has been submitted to the proposed low-reflectivity process, whereas the other ones have followed the standard flowchart. Among the wafers processed with the special process, different thickness of SiO_2 and Si_3N_4 have been chosen. The more efficient split resulted in five good wafers which have been measured in Aachen, and the results are listed in Table B.1.

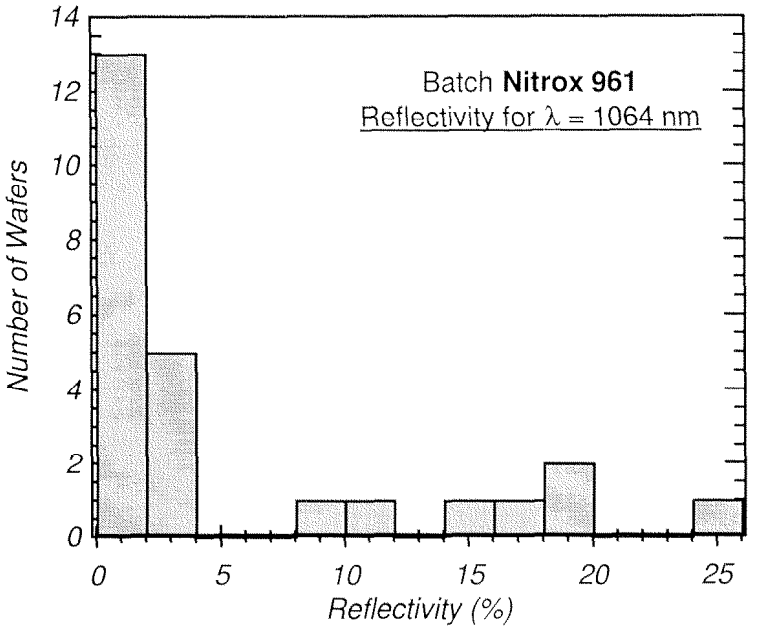


Fig. B.2: Measured reflectivity of 24 wafers of the blanket test lot "Nitrox 961". Samples measured at λ=1064 nm.

Wafer	Reflectivity
1238	1.76%
1244	1.05%
1245	0.52%
1246	0.46%
1247	0.58%

Table B.1: Reflectivity measured on five AMS/63 wafers.

B.5 Electrical results

Having good optical properties is not enough. It is also necessary to test that the electrical behaviour of the detectors has not been affected. To this end the leakage current of each detector of lot AMS/63 have been measured in full depletion conditions. The measurements have been performed at CSEM with an HP4145 up to 100 V. The detectors do totally deplete at about 60 V. Fig. B.3 shows the leakage curves corresponding to a low-reflectivity detector. The total leakage current of the detector amounts to some 240 nA at 75 V. A total of 9 low-reflectivity detectors have been measured, with similar results. The comparisons with the 11 standard detectors from the same lot does show not a significant increase of the leakage currents for the low-reflectivity process.

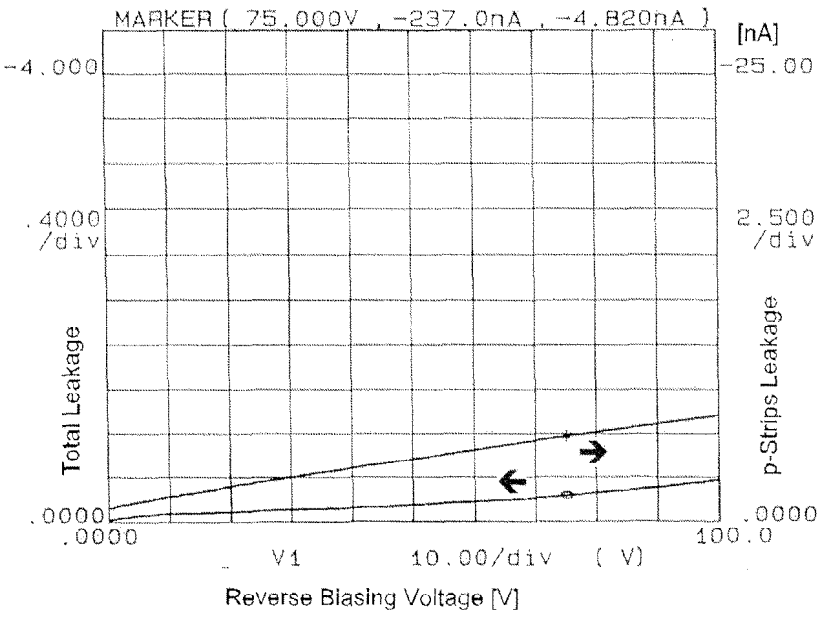


Fig. B.3: AMS detector total leakage current in low-reflectivity technology.

B.6 Conclusions

A satisfactory solution to the silicon reflectivity problem has been found: the reflectivity of the AMS detectors has been reduced to less than 1%, without significantly altering their electrical properties. The chosen solution, which is modular, resorts to standard silicon-technology material and is completely compatible with silicon manufacturing process flow.

Glossary

Abbreviations

c.m. center-of-mass, p. 5.

ENC Equivalent Noise Charge.

LEP Large Electron Positron collider, p. 4.

LHC Large Hadron Collider.

MSSM Minimal Supersymmetric extension of the Standard Model, p. 4.

SM Standard Model, p. 4.

SUSY SUper SYmmetry, p. 5.

Symbols

α damage constant.

η pseudo-rapidity, $\eta = -\log\tan(\theta/2)$.

$\mathcal{L}$ Luminosity.

$\mathcal{E}$ electric field.

$\mu_{n,p}$ mobility.

ϕ_i juction built-in potential.

ρ resistivity.

ρ spatial charge density.

$\sigma_{n,p}$ capture cross sections.

σ_s surface capture cross-section.

τ amplifier peak shaping time.

τ_g	generation lifetime.
τ_r	recombination lifetime.
$\tau_{n,p}$	mean time between collisions, recombination lifetime.
E_c	conduction band energy.
E_f	Fermi-level energy.
E_g	band-gap energy.
E_v	valence band energy.
G'	generation rate.
$J_{n,p}$	electron, hole current density.
k	Bolzmann constant.
M	avalanche multiplication factor.
n	electron density.
$n'_{n,p}$	excess electron density in a $n-$ (resp. $p-$)region.
N_c	effective density of states (conduction band).
n_i	intrinsic free-carrier concentration.
n_s	surface electron density.
N_t	RG-center density.
N_v	effective density of states (valence band).
$N_{A,D}$	total acceptor, donor concentration.
$n_{n,p}$	electron density in a $n-$ (resp. $p-$)region.
$n_{n0,p0}$	equilibrium electron density in a $n-$ (resp. $p-$)region.
p	hole density.
p	strip pitch (in μm).
$p'_{n,p}$	excess vacancy density in a $n-$ (resp. $p-$)region.
p_t	transverse momentum.
$p_{n,p}$	vacancy density in a $n-$ (resp. $p-$)region.
$p_{n0,p0}$	equilibrium vacancy density in a $n-$ (resp. $p-$)region.

p_s	surface vacancy density.
R'	recombination rate.
R_p	noise equivalent parallel resistor.
R_s	noise equivalent series resistor.
U	net rate of recombination, $U \equiv R - G$.
V_b	bias voltage.
v_d	drift velocity.
W_d	depletion region width.
X_0	radiation length.

References

[1] S.Weinberg. *Phys. Rev. Lett.*, 19:1264, 1967.

[2] H.E.Haber and G.L.Kane. The search for supersymmetry: Probing physics beyond the standard model. *Phys. Reports*, 117:75, 1985.

[3] G.Altarelli. In *LHC Workshop, Aachen*, volume CERN 90-10, 1990.

[4] S.Dimiopulos and M.Lindner. In *LHC Workshop, Aachen*, volume CERN 90-10, 1990.

[5] Patrick Janot. Search for new particles at present colliders. In *1997 EPS Conference in Jerusalem*, August 1997. Contribution.

[6] A.Boehm. Contribution. In *Rencontre de Moriond*. LEP/SLD Electroweak Working Group, 1997.

[7] M.Veltman. *Ann.Phys.*, B8:475, 1977.

[8] CMS Collaboration. The Compact Muon Solenoid. Technical Proposal. Technical report, CERN, 1994.

[9] The LHC Study Group. The Large Hadron Collider: Conceptual Design. Technical report, CERN, 1995. CERN/AC 95-05 (LHC).

[10] D.Denegri. The CMS detector and physics at the LHC. In *Conference on Elementary Particle Physics, Present and Future*, June 1995.

[11] CMS Collaboration. CMS ECAL Technical Design Report. Technical report, CERN, 1997. CERN/LHCC 97-33.

[12] CMS Collaboration. The CMS HCAL Project Technical Design Report. Technical report, CERN, 1997. CERN/LHCC 97-31.

[13] F. Angelini *et al. Nucl.Instr. & Meth. A*, A343:441, 1994.

[14] CMS Collaboration. The CMS Tracker Technical Design Report. Technical report, CERN, April 1998. CERN/LHCC 98-6.

[15] C.J.S.Damerell. Lectures presented at the XXIII Annual Summer Institute on Particle Physics, SLAC, Stanford, California, USA, July 1995.

[16] H.Bethe and J.Ashkin. Passage of radiation through matter. In E.Segre, editor, *Experimental Nuclear Physics*, volume 1, chapter 2. John Wiley & Sons, New York, 1959.

[17] S.Ahlen. Theoretical and experimental aspects of the energy loss of relativistic heavily ionizing particles. *Rev. Mod. Phys.*, 52:121, 1980.

[18] R.M.Sternheimer *et al.* Density effects fro the ionization loss of charged particles in various substances. *Atomic Data and Nuclear Data Tables*, 30:261, 1984.

[19] C. Caso *et al.* Review of Particle Physics. *European Physical Journal*, C3:1, 1998.

[20] J.A.Kerr *et al. Handbook of Chemistry and Physics.* CRC Press, 64 edition, 1983.

[21] D.J.Skyrme. *Nucl.Instr. & Meth. A*, 57:61, 1967.

[22] H.Bichsel. *Rev.Mod.Phys.*, 60:663, 1988.

[23] S.M.Sze. *Semiconductor Devices: Physics and Technology.* John Wiley & Sons, 1985.

[24] R.F. Pierret. Advanced semiconductor fundamentals. In *Modular Series on Solid State Devices*, volume VI, page 65. Addison-Wesley, 1987.

[25] A. Neugroeschel *et al. IEEE Trans. Electron Devices*, ED-29:894, 1982.

[26] G. Masetti *et al. IEEE Trans. Electron Devices*, ED-30:764, 1983.

[27] W.F. Beadle *et al. Quick Reference Manual for Semiconductor Engineers.* John Wiley & Sons, 1985.

[28] R.S. Muller and T.J. Kamins. *Device Electronics for Integrated Circuits.* John Wiley & Sons, 1986.

[29] W.Shockley and W.T.Read. Statistics of the recombination of holes and electrons. *Phys. Rev.*, 87:835, 1952.

[30] R.N.Hall. *Phys. Rev.*, 87:387, 1952.

[31] G.W. Neudeck. The P-N junction diode. In *Modular Series on Solid State Devices*, volume 2. Addison Wesley, 1989.

[32] J.B.Gunn. On carrier accumulation and the properties of certain semiconductor junctions. *J. Electro. Control*, 4:17, 1958.

[33] E.H.M.Heijne *et al. Nucl.Instr. & Meth. A*, 178:331, 1980.

[34] P.Weilhammer. Private Communication.

[35] G.F.Knoll. *Radiation Detection and Measurement.* John Wiley & Sons, 1979.

[36] E.Gatti *et al.* Processing the signals from solid-state detectors in elementary-particle physics. *Riv. Nuovo Cim.*, 9:1, 1986.

[37] C.Bozzi. Signal-to-noise evaluations for the CMS silicon microstrip detectors. Technical report, INFN Pisa, 1997. CMS Note 97-26.

[38] G.Baricchello *et al.* Performance of long modules of silicon microstrip detectors. *Nucl.Instr. & Meth. A*, 413:171, 1998.

[39] E. Nygard *et al. Nucl.Instr. & Meth. A*, 301:506, 1991.

[40] W.ten Kate. *The microstrip detector.* PhD thesis, Delft University, 1988.

[41] C.T. Sah *et al.* Carrier generation and recombination in $p-n$ junctions and $p-n$ junction characteristics. *Proc. IRE*, 45:1228, 1957.

[42] P.U.Calzolari *et al.* A theoretical investigation on the generation current in a silicon $p-n$ junction under reverse bias. *Solid-St. Electr.*, 15:1003, 1972.

[43] J.W.Chen and A.G.Milnes. Energy levels in silicon. In R.A.Huggins *et al.* eds., editor, *Annual Review of Material Science*, volume 10, page 157. Annual Reviews, 1980.

[44] J.Alamo *et al.* High-low junctions for solar cell applications. *Solid-St. Electr.*, 18:533, 1981.

[45] J.R.Hauser *et al.* Minority carrier reflecting properties of semiconductor high-low junctions. *Solid-St. Electr.*, 18:715, 1975.

[46] R. Della Marina. Process dependence of microstrip detectors electrical characteristics. Technical report, CSEM, 1997.

[47] V.Anantharam *et al.* Analytical solutions for the breakdown voltage of puntched-through diodes having curved junction boundaries at edges. *IEEE Trans. Electron Devices*, 27:939, 1980.

[48] S.K.Gandhi. *Semiconductor Power Devices.* John Wiley & Sons, New York, 1979.

[49] J.L.Moll *et al.* Charge multiplication in silicon $p-n$ junction. *Solid-St. Electr.*, 6:147, 1963.

[50] M.S.Adler *et al.* Theory and breakdown voltage for planar devices with a single field limiting ring. *IEEE Trans. Electron Devices*, 24:107, 1977.

[51] Baliga. *Modern Power Devices.* John Wiley & Sons, 1987.

[52] B.S. Kurana *et al.* Thermal breakdown in silicon $p-n$ junction devices. *IEEE Trans. Electron Devices*, ED-13(11):763, 1966.

[53] M. Da Rold. PhD thesis, Padova Univ., Italy, 1998.

[54] K.C.Saraswat *et al.* Breakdown walk-out in planar $p-n$ junctions. *Solid-St. Electr.*, 21:813, 1978.

[55] A. Longoni *et al. Nucl.Instr. & Meth. A*, 288:35, 1990.

[56] T. Oshugi *et al.* Microdischarges of AC-coupled silicon strip sensors. *Nucl.Instr. & Meth. A*, 342:22, 1994.

[57] Silvaco Inc. ATLAS User Manual, 1997.

[58] A.E. Ruehli *et al. IEEE J. Solid-State Circuits*, SC-8(8):289, 1973.

[59] J.W.Duncan. *IEEE Trans. Microwave Theory Tech.*, MTT-15(10):575, 1967.

[60] W.H.Chang. *IEEE Trans. Microwave Theory Tech.*, MTT-24(9):608, 1976.

[61] T.Sakurai *et al.* Simple formula for two- and three-dimensional capacitances. *IEEE Trans. Electron Devices*, ED-30(2):183, 1983.

[62] U. Kötz *et al.* Silicon strip detectors with capacitive charge division. *Nucl.Instr. & Meth. A*, 235:481, 1985.

[63] P.W.Cattaneo. PhD thesis, Milano Univ., 1987.

[64] K.J.Binns and P.J.Laurenson. *Analisys and computation of electric and magnetic field problems.* Oxford Pergamon Press, 1973.

[65] G. Hall *et al.* Calculation of the geometrical capacitance of silicon microstrip structures using a variational approach. *Nucl.Instr. & Meth. A*, A326:228, 1993.

[66] R. Della Marina *et al.* Silicon strip detectors for LHC: comprehensive process and device analysis. *Nucl.Instr. & Meth. A*, A392:178, 1997.

[67] *Virtual Wafer Fab, Silvaco User's Manual.*, 1997.

[68] Holmes-Siedle *et al. Nucl.Instr. & Meth. A*, A339:511, 1994.

[69] E. Nicollian and J. Brews. *MOS Physics and Technology.* John Wiley & Sons, 1982.

[70] H.Feick *et al.* Long term damage studies using silicon detectors fabricated from different starting materials and irradiated with neutrons, protons and electrons. *Nucl.Instr. & Meth. A*, A377, 1996.

[71] G.Batignani *et al.* The ALEPH silicon vertex detector. *Nucl.Phys.B*, 23A:291, 1991.

[72] G.Batignani *et al.* Recent results and running experience of the new ALEPH vertex detector. *Nucl.Instr. & Meth. A*, 360:98, 1995.

[73] The ALEPH Collaboration. Performance of the ALEPH detector at LEP. *Nucl.Instr. & Meth. A*, 360:481, 1995.

[74] R. Della Marina *et al.* Results on double-sided AC-coupled silicon strip detectors. *Riv. Nuovo Cim.*, 109(9):1319, 1996.

[75] J. Kemmer. Fabrication of low noise silicon radiation detectors by the planar process. *Nucl.Instr. & Meth. A*, 169:499, 1980.

[76] W.R. Runyan and K.E. Bean. *Semiconductor integrated circuit processing technology.* Addison-Wesley, 1990.

[77] S. Holland. Fabrication of detectors and transistors on high-resistivity silicon. *Nucl.Instr. & Meth. A*, 275:537, 1989.

[78] L. De Maria. Interstrip capacitance comparison between $\langle 111 \rangle$ and $\langle 100 \rangle$ silicon detectors, January 1999. Oral communication to CMS Silicon Tracker Meeting.

[79] Rose Collaboration. 2nd RD48 status report. Technical Report 98-39, CERN/LHCC, October 1998.

[80] A. Bloch. *Murphy's law and other reasons why things go wrong.* Putnam Pub Group, 1989.

[81] B.E.Deal. Standardized technology for oxide charges associated with thermally oxidated silicon. *IEEE Trans. Electron Devices*, ED-27:606, 1980.

[82] R. Della Marina. Doping implantations for MSD's with MAXIS technology. Technical report, CSEM, 1996.

[83] S.Holland. An IC-compatible detector process. *IEEE Trans. Nuclear Science*, 36(1), February 1989.

[84] B.E.Deal. The current understanding of charges in the thermally oxidized silicon surface. *J. Electrochem. Soc.*, 121(6):198, June 1974.

[85] G.Tonelli. Private Communication.

[86] E.R.Weber. Transition metals in silicon. *Appl. Phys.*, A(30):1, 1983.

[87] A. Goetzberger and W. Shockley. Metal precipitates in silicon p-n junctions. *J. Appl. Phys.*, (31):1821, 1960.

[88] P. Richmond. The effect of gold doping upon the characteristics of MOS field-effect transistors with applied substrate voltage. *Proc. IEEE*, (56):774, 1968.

[89] M.C.Chen and V.J.Silvestri. Post-epitaxial polysilicon and si_3n_4 gettering in silicon. *J. Electrochem. Soc.*, (129):1294, 1982.

[90] W.T.Stacy *et al.* The microstructure of polysilicon backsurface gettering. In *Electrochem. Soc. Ext. Abst.*, volume 83, pages 484–485, 1983.

[91] R.S.Ronen and P.H.Robinson. Hydrogen chloride and chlorine gettering as an effective technique for improving performance in silicon devices. *J. Electrochem. Soc.*, (119):747, 1972.

[92] Rose Collaboration. 1st RD48 status report. Technical Report 97-39, CERN/LHCC, June 1997.

[93] B.C. McEvoy. Investigations using a simple model for defect kinetics in silicon. TN 95-41, CERN-RD20, April 1995.

[94] G. Casse. Introduction of high oxigen concentration into silicon wafers by high temperature diffusion. TN 99-01, CERN-RD48, January 1999.

[95] R. Della Marina *et al.* Characterization of MOS transistors integrated on high-resistivity silicon with DSSD process. *Riv. Nuovo Cim.*, 110A:817, 1997.

[96] M.Da Rold *et al.* Radiation effects on breakdown characteristics of multiguarded devices. *IEEE Trans. Nuclear Science*, 44:721, 1997.

[97] N.Bacchetta *et al.* High voltage operation of silicon devices for LHC experiments. In *Proceedings of the 7th Pisa Meeting on Advanced Detectors*, 1997.

[98] O. Militaru *et al.* Study of the edge effects in the breakdown process of p^+−on-n bulk silicon diodes. Proceedings of the 8^{th} European Symposium of Semiconductor Detectors, Schloss Elmau, June 1998., 1998.

[99] D. Pitzl *et al. Nuclear Physics B*, 3A:340, 1991.

[100] R.Della Marina *et al.* Diagnostic tools and characterization techniques for large area integrated capacitors on high resistivity microstrip detectors. Boston, September 1998.

[101] R. Wheadon. Radiation tolerance studies of silicon microstrip detectors for the LHC. *Nucl.Instr. & Meth. A*, 342:126, 1994.

[102] O.Adriani *et al.* Beam test results for single and double-sided silicon detector prototypes of the CMS Central Detector. *Nucl.Instr. & Meth. A*, A396:76, 1997.

[103] E.Babucci *et al.* Temperature dependence of the behaviour of a single-sided irradiated silicon detector. CMS Note 98-011, 1998.

[104] R.Della Marina *et al.* Test results of heavily irradiated Si detectors. *Nucl.Instr. & Meth. A*, 422:238, 1999.

[105] B. Borgia S. Baccaro, A. Festinesi. Gamma and neutron irradiation facilities at ENEA - Casaccia Center (Roma). CMS TN/95 - 192, 1995.

[106] S. Dimopoulos and H. Georgi. *Nucl. Phys.*, B193:150, 1981.

[107] S.P. Martin. *A supersymmetry primer*, volume Perspectives on Supersymmetry of *Advanced Series on Directions in High Energy Physics*. World Scientific, Singapore, 1998.

[108] H. P. Nilles. *Phys. Reports*, 110:1, 1984.

[109] S. Wolfram. *Phys. Lett.*, B82:65, 1979.

[110] E. Norman. *Phys. Rev. Lett.*, 58:1403, 1987.

[111] S.Dimiopulos *et al.*. *Nucl.Phys.*, B452:496, 1995.

[112] B. Campbell. *Phys. Rev.*, D28:209, 1983.

[113] H. Baer and X. Tata. *Phys. Rev.*, D47:2739, 1993.

[114] H. Goldberg. *Phys. Rev. Lett*, 50:1419, 1983.

[115] F. Paige *et al.*. ISAJET 7.40. Technical report, Brookhaven National Laboratory, 1998. BNL-HET-98-39.

[116] S.Mrenna. SPYTHIA, a supersymmetric extension of PYTHIA. *Comput.Phys.Commun.*, (101):232, 1997.

[117] F. Pauss and M. Dittmar. Experimental challenges at LHC. In *Proceedings of 1998 ASI Summer School, St.Croix*, February 1999. CMS Note 99-8.

[118] The ALEPH Collaboration. Search for neutral Higgs boson in e^+e^- collisions at $\sqrt{s}$ up to 196 GeV. In *European Physical Society Conference HEP'99*, Tampere, Finland, July 1999. From Talk given by N.Kostantinidis.

[119] R. Kinnunen and D. Denegri. CMS Note 97-57, 1997.

[120] S.Abdullin. CMS Technical Note 1996/095, 1996.

[121] S.Abdullin *et al.*. CMSJET. CMS Technical Note 94-180, 1994.

[122] S.Abdullin and D.Denegri. CMS Note 97-70, 1997.

[123] L. Rurua. Possibilities for $h \rightarrow b\bar{b}$ with CMS at LHC. *Nucl.Instr. & Meth. A*, 408:128, 1998.

[124] ALEPH Collaboration. *Phys. Lett.*, B259:377, 1991.

[125] CDF Collaboration. Measurement of $b^0 - \overline{B}^0$ flavor oscillations using jet-charge and lepton flavor tagging in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV. Preprint FERMILAB-Pub-99-019-E (hep-ex/9903011), 1999.

[126] G. Ridolfi J. Ellis and F. Zwirner. *Phys. Lett.*, B237:423, 1990.

[127] D. Treille. In *Proceedings of the 29th International Conference on High Energy Physics ICHEP 98*, Vancouver, British Columbia, Canada, July 1998. World Scientific, Singapore.

[128] CDF Collaboration. *Phys. Rev.*, D56:1357, 1997.

[129] The ALEPH Collaboration. *Europ. Phys. J.*, C7:383, 1999.

[130] The L3 Collaboration. *Europ. Phys. J.*, C4:207, 1998.

[131] The OPAL Collaboration. *Europ. Phys. J.*, C8:255, 1999.

[132] *European Physical Society Conference HEP'99*, Tampere, Finland, July 1999. From Plenary Talk given by E. Gross.

[133] The D0 Collaboration. *Phys. Rev. Lett.*, 75:618, 1995.

[134] The CDF Collaboration. *Phys. Rev.*, D56:R1357, 1997.

[135] The CDF Collaboration. *Phys. Rev. Lett.*, 77:438, 1996.

[136] H.Baer *et al.*. Signals from minimal supergravity at the CERN LHC: multi-jet plus missing energy channel. *Phys.Rev.*, D52:2746, 1995.

[137] D. Nanopoulos J. Lopez and H. Pois. *Phys. Rev.*, D47:2468, 1993.

[138] B. Grinstein M. Dugan and L. Hall. *Nucl. Phys.*, B255:413, 1985.

[139] T.Skwarnicki. In *Proceedings of the 29th International Conference on High Energy Physics ICHEP 98*, Vancouver, British Columbia, Canada, July 1998. World Scientific, Singapore.

[140] M.Dittmar. Searching for the higgs and other exotic objects. In *30eme Ecole d'Ete de Physique des particules, Marseille, 7-11 septembre 1998*, February 1999. CMS Note 99-9.

[141] J. van der Spiegel *et al.* Theoretical and practical investigation of the thermal generation in gate controlled diodes. *Solid-St. Electr.*, 24:869, 1981.

[142] D.E.Sawyer and R.H.Rediker. Narrow base germanium photodiodes. *Proc. IRE*, 46:1122, 1958.

[143] J.W.Slotboom *et al.* Measurement of bandgap narrowing in silicon bipolar transistors. *Solid-St. Electr.*, 19:857, 1976.

[144] R. Della Marina et al. Production and test of silicon particle detectors optimized for optical transparency. Presented at the 7th Pisa Meeting on Advanced Detectors, La Biodola, Isola d'Elba, Italy, May 25-31, 1997.

[145] R. Battiston. The alpha magnetic spectrometer AMS. *Nucl.Instr. & Meth. A*, 409:458, 1998.

[146] R. Della Marina. AMS Silicon Detector Chip Design Report. Technical report, CSEM, 1996.

[147] F. and P. Pedrotti. *Introduction to Optics*. Prentice Hall Inc., 1993.