Optimisation of the ILC vertex detector and study of the Higgs couplings

Georgios Voutsinas

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DE PHYSIQUE ET CHIMIE PHYSIQUE

Institute Pluridisciplinaire Hubert Curien

THÈSE présentée par :

Georgios Voutsinas

soutenue le : 28 **Juin 2012**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Physique de Particules

Développement d'un Détecteur de Vertex de Nouvelle Génération pour le collisionneur ILC – Impact sur la Détermination des Rapports d'Embranchment du Boson de Higgs Standard

THÈSE dirigée par :

Dr Winter.Marc	Université de Strasbourg

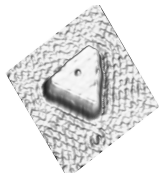
RAPPORTEURS :

Dr Wingerter Isabelle	LAPP
Dr Behnke Ties	DESY

AUTRES MEMBRES DU JURY :

Prof Le Diberder Francois	Université Paris Sud 11
Dr Dracos Marcos	IPHC
Prof Tsipolitis Yorgos	National Technical University of Athens

Georgios Voutsinas



Développement d'un Détecteur de Vertex de Nouvelle Génération pour le collisionneur ILC – Impact sur la Détermination des Rapports d'Embranchment du Boson de Higgs Standard

Résumé

Cette thèse est une contribution au document intitulé "Detector Baseline Document (DBD)" décrivant le concept de détecteur ILD envisagé auprès du collisionneur linéaire international électron-positon ILC (acronyme de l'anglais International Linear Collider).
Les objectifs de physique de l'ILD nécessitent un détecteur de vertex (VXD) particulièrement léger, rapide et très granulaire permettant d'atteindre une résolution sans précédent sur le paramètre d'impact des trajectoires reconstruites des particules produites dans les interactions étudiées. Le principal objectif de cette thèse est de montrer comment optimiser les paramètres du VXD dans le cas ou il est composé de Capteurs à Pixels Actifs fabriqués en technologie industrielle CMOS (CAPS). Ce travail a été réalisé en étudiant la sensibilité des performances d'étiquetage des saveurs lourdes et de la précision sur les rapports d'embranchement hadronique du boson de Higgs aux différents paramètres du VXD.
Le cahier des charges du VXD, particulièrement ambitieux, a nécessité le développement d'une nouvelle technologie de capteurs de pixels de silicium, les CAPS, dont le groupe PICSEL de l'IPHC est à l'origine. La vitesse de lecture et l'influence des paramètres qui régissent la fabrication des capteurs en fonderie ont été étudiées dans cette thèse, et des prototypes
de CAPS ont été caractérisés sur faisceau de particules. Enfin, les performances de trajectométrie d'un VXD composé de CAPS a été évalué avec des études de simulation.

Mots clés: ILC, ILD, Détecteur de Vertex, Étiquetage des
Saveurs Lourdes, Rapports d'embranchement du Higgs, Capteurs
CMOS à Pixels Actifs

Résumé en anglais

This thesis is a contribution to the " Detector Baseline Document ", describing the ILD detector which is intended for the International Linear Collider (ILC).

The physics goals of the ILD call for a vertex detector (VXD) particularly light, rapid and very granular allowing to reach an unprecedented resolution on the impact parameter of the tracks that reconstruct the particles produced in the studied interactions. The principle goal of this thesis is to show how to optimise the parameters of the VXD in the case that is composed of Active Pixel Sensors manufactured in industrial CMOS technology (CAPS). This work has been realised by studying the sensitivity of the performance of the heavy flavour tagging and the precision on the hadronic branching fractions of the Higgs boson as a function of different sets of VXD parameters.

The specifications of the VXD, particularly ambitious, call for the development of a novel silicon pixel sensors technology, the CAPS, which was pioneered by the PICSEL group of IPHC. The readout speed and the influence of the fabrication parameters have been studied in this thesis, and CAPS prototypes have been characterised in test beams. Finally, the tracking performance of a CAPS based VXD has been evaluated with simulation studies.

Key words : ILC, ILD, Vertex Detector, Flavour Tagging, Higgs Branching Fractions, CMOS Active Pixel Sensors

Contents

List of Tables

List of Figures

Preface

The Standard Model (SM) describes the elementary particles of Nature and the fundamental interactions between them. It has been extremely successful, being consistent with the experimental results at the per mill level. The only part missing to be complete is the experimental verification of the higgs sector. The higgs sector offers an explanation, inside the framework of the SM, for the ElectroWeak Symmetry Breaking (EWSB) and the generation of the fermion masses. Despite its successes, the SM cannot be the ultimate theory. The reason is that it cannot describe one of the four known fundamental interactions, the gravity. Moreover, there is an experimental observation that contradicts the SM. That is the discovery of the Dark Matter, which cannot be explained by the SM.

Therefore, the two dominant questions in today's particle physics are to understand the EWSB and to search for signs of new, beyond the SM, physics. The International Linear Collider (ILC) aims to address these two questions. The EWSB will be explored through detailed study of the higgs sector. The ILC physics case will be significantly enhanced if a light higgs boson will be discovered at the LHC*. Its mass, couplings, width and quantum numbers can be measured with high precision in a model independent way using the higgsstrahlung process, which is the main benchmark reaction proposed in the ILD Letter of Intent (LoI). The ILD (International Large Detector) is one of the two detector concepts approved by the International Detector Advisory Group (IDAG).

This thesis was a contribution to the ILD Detector Baseline Document (DBD). It concerned the optimisation of the ILD Vertex Detector (VXD) and its sensors, focusing on the CMOS sensors technology. The physics goals of the experiment require a VXD featuring unprecedented impact parameter resolution, capable to provide excellent performances in the reconstruction of displaced vertices and therefore in heavy flavour tagging. These requirements call for an innovative pixel technology. This thesis

*A particle consistent with the higgs boson, featuring a mass of 126 GeV has been discovered lately (July 2012) at the LHC.

mainly addressed the three following subjects:

- Test beam data analysis of CMOS sensors.
- Optimisation of the VXD via simulation studies of the higgsstrahlung process.
- Tracking performances of the proposed CMOS based VXD and Silicon Inner Tracker (SIT) detector configuration.

It consists of seven chapters. The first three of them are the introductory ones. In the first chapter the SM is introduced, giving more emphasis to the higgs sector. It is explained why SM cannot be the ultimate theory. A non exhaustive list of the beyond the SM theories is briefly summarised. Then we pass to the physics motivations of the ILC, and give a description of its expected physics program, assuming different scenarios of the LHC results.

The second chapter is an introduction to the ILC and the ILD. We give a description of the ILD subdetectors, obviously focusing on the VXD. A discussion of the reconstruction algorithms also takes place, where we focus mainly on the standalone silicon tracking.

The third chapter starts by reporting the VXD sensor requirements. Then the CMOS sensors are introduced, and it is explained why they are considered as a promising candidate for the ILC VXD. Finally, we propose a CMOS based VXD, that can satisfy the ILC physics goals while being compliant with the experimental running constraints.

The author had no contribution to the aforementioned first three chapters. The first part of the work conducted for this thesis is reported in the fourth chapter. There, after first describing the test setup and the tools used for the laboratory and the beam tests, we pass to the test beam results of the MIMOSA 22bis and MIMOSA 24 sensors, which were actually the author's contribution to the CMOS sensors R&D. MIMOSA 22bis was addressing the upstream part (pixel array) of MIMOSA 26, which offers the baseline architecture for numerous applications. The main purposes of MIMOSA 22bis test beam was to extract the optimal pixel design, featuring in–pixel amplification and pedestal subtraction, and to validate the column parallel readout. The whole R&D line that led us to MIMOSA 26 was based on the AMS 0.35 μm CMOS process. The scope of MIMOSA 24 was to enlarge the group of the validated CMOS processes for charged particle tracking by adding a new one, namely the XFAB 0.35 μm.

The fifth chapter addressed the optimisation of the VXD. In order to do that, we performed full simulation studies using the higgsstrahlung process. Two main alternative VXD geometries exist, one equipped with five single layers (VXD03),and one equipped with three double layers (VXD05). The performances of the two geometries were compared in terms of heavy flavour tagging efficiency and purity and in their sensitivities on the extraction of the higgs hadronic branching ratios. The study was

based on the LoI software tools, and assumed a generic sensor technology. The flavour tagging was provided by a neural network architecture being developed for the LoI. The neural nets were subjected to a dedicated training for each geometry. In the second part of this chapter, the parameter space of the sensor requirements was explored. The motivations were multifolded: to refer to some of them, was to see in what extent we can play with the sensors material budget and the distance from the interaction point without severely penalising the physics potential of the detector, to find the best compromise between the power consumption and the single point resolution of the outer layer's sensors, or to further probe differences in the behaviour of the VXD03 and VXD05 geometries. Finally we tried to evaluate the robustness of the VXD versus the beamstrahlung. The results obtained by this study were not reliable due to the non mature software tools.

In the sixth chapter we proceeded in tracking performance studies of the CMOS based VXD, in order to probe its robustness against the beam background. The obtained results are expected to be useful in the guidance of the CMOS sensors further development, especially concerning the spatial–time resolution interplay, the optimisation of each VXD layer separately, but as well as to extract the optimal overall detector configuration. The SIT is included in this study. Two options are examined and compared: the standard one, equipped with strip sensors, and an alternative equipped with CMOS pixel sensors.

The thesis is closing with a synthetic and comprehensive conclusion.

CHAPTER 1

Theoretical Background

The Standard Model (SM) is a single consistent theory, providing a joint description of the electroweak and strong interactions between the elementary particles. The only interaction that is not included is the gravity. Its development was one of the major scientific achievements of the last century. It has been very successful in the explanation of a wide variety of experimental observations. All the predicted particles but one have been already found. The one left is the higgs boson. The higgs field manifests itself through this particle. Therefore it has profound importance since in the SM, the particles acquire their mass via the interaction with the higgs field. With the discovery of the higgs boson, the SM would be completed.

But is this the end of the story? Obviously not. As we already mentioned, the SM does not account for one of the four fundamental interactions, the gravity. Until recently, there was no experimental result to contradict the SM. But now, cosmological observations from independent sources agree that only $\sim 5\,\%$ of the Universe is made of ordinary matter. For the remaining $95\,\%$, the SM has not any explanation. So it is obvious that it is not the ultimate theory. New theories, either extensions of the SM or beyond the SM, wait for experimental verification. Finally, even if the SM is in very good agreement with the experimental data, some of its sectors are quite obscure. We do not really understand why the electroweak symmetry breaking happens. Some other sectors, like the higgs mechanism, have been added ad hoc, while parameters like the fermion masses have been added "by hand". It is widely believed that the SM is a very good approximation of the truth at the electroweak energy scale. New physics may wait for us at the tera–scale.

So the time is ideal for new collider experiments which are going to explore the tera–scale. The Large Hadron Collider (LHC) at CERN has already started. It is a

powerful machine, with a collision energy of 14 TeV, that features a very high potential for discoveries. However, the extremely high data rate, the very high level of QCD background and the poorly defined parameters governing elementary interactions (a problem implicit to hadron colliders) complicate the realisation of high precision measurements. Moreover, the level of the precision achieved is not as high as to assertively establish or invalidate the SM. For example, if a higgs boson is found, one has to measure precisely its properties in order to find out to whether this is the SM higgs boson or something else. Finally, the extremely high data rate imposes the triggered selection of physics processes, based on their signature. A lepton collider, offering clean, well defined and tunable interaction conditions, would be ideal for precise measurements. Here we should stress the advantages stemming from precise measurements. Measuring precisely small quantum corrections, they can give us an insight for physics at much higher energy scales than the direct reach of the machine. Moreover, the sensitivity to new phenomena or new particles is enhanced because of the tunable, well defined and almost background free interaction conditions.

The International Linear Collider (ILC), which is an e^+e^- collision machine, is a strong candidate. The precision added to LHC preliminary assessments would be highly valuable. Its untriggered operation maximises its sensitivity to non–SM physics. This is another example of the fruitful LHC–ILC interplay, especially in the case where both machines are running concurrently. Feedback from the ILC could help the LHC experiments to optimise their triggers in order to maximise their efficiency. Last but not least they can establish a very interesting synergy with astrophysics and neutrino experiments, in order to address the cosmological questions.

In this chapter, we are going first to give a brief overview of the SM. Then we will refer to the main challenges that it has to confront, and the proposed solutions which are naturally leading us to plan experiments like those at the ILC. The physics potential of the ILC will be described. The main topic of this thesis is the ILC vertex detector. We are going to evaluate this detector by studying its performances in tracking and heavy flavour tagging, as well as with the accuracy on the determination of the branching ratios of the SM higgs. For the purposes of these studies, we are going to use mainly the higgsstrahlung process. Thus, the SM higgs production and decays will be described in the last part of the chapter. The bibliography of the theoretical part of the chapter is based on textbooks [Halzen 84], lectures and summer school lectures [School 09]. The part concerning the SM higgs physics at the ILC is based on [RDR b].

1.1 The Standard Model

1.1.1 Structure of the Standard Model

The SM is a gauge theory described by the gauge group

$$SU_C(3) \otimes SU_L(2) \otimes U_Y(1) \tag{1.1}$$

A gauge theory is invariant under a continuous set of local transformations. The particles of the SM are divided in elementary fermions, which are the constituents of matter and have spin $\frac{1}{2}$ and intermediate gauge bosons, which are the mediators of the interactions and have spin 1. This framework is complemented by the higgs boson, which is the only particle predicted by the SM yet to be found, and which is a product of the higgs mechanism which is assumed to be responsible for the generation of the masses and of the electroweak symmetry breaking.

There are 12 elementary fermions. Each of them has a corresponding antiparticle, featuring identical mass but the opposite electrical charge and magnetic dipole moment. The fermions appear in two categories, the quarks and the leptons. The leptons are moving freely in nature. On the contrary, the quarks are always found in bound states called hadrons. This happens due to the colour quantum number which the quarks are carrying. There are three colour (red, blue, green) and three anti–colour (anti–red, anti–blue, anti–green) quantum numbers. Only colourless particles are observed in nature. Thus the quarks can only be found in colourless combinations, the hadrons, of two (mesons) or three (baryons) quarks *. This effect is called colour confinement. The six quarks, carrying colour, electric charge and weak isospin, can interact through both electroweak and strong interactions. Three out of the six leptons, the electron, the muon and the tau, are carrying electric charge thus interact through the electromagnetic and the weak interactions. The three neutrinos, as their name implies, are neutral. They interact only via the weak interaction.

The fermions are further classified in three generations. However, the existence of more generations cannot be excluded. The measurement of the Z boson width, realised at the Large Electron Positron collider (LEP) at CERN, demonstrated that only three species of light left handed neutrinos exist. We should point here that if sterile neutrinos do exist (these are neutrinos that do not interact even through the weak interaction), they could allow for the existence of more than three generations of light neutrinos, despite the LEP results. However it should be stressed that up to date, there is no evidence for a fourth generation of fermions.

Each generation contains two quarks, one charged lepton and one neutrino. The total lepton number, defined by the generation number of the charged leptons, is always conserved in interactions. This is not true for the neutrinos, since neutrinos flavour oscillations have been observed. Concerning the quarks, mixing between their generations is allowed via weak charged currents, but it is suppressed with respect to interactions inside the same generation. The strength of the charged flavour changing currents is expressed by the Cabibbo–Kobayashi–Maskawa (CKM) matrix [PDG a]. The fermions of each generation have different masses, the ones of the first generation

*4 and 6 quarks bound states are predicted by SU(3) quark model.

being the lightest. The fermions belonging to the second or third generations decay to the ones belonging to a lower generation, therefore we rapidly end up only with fermions of the first generation. The ordinary matter in the whole Universe is made of fermions belonging to the first generation. Those are the electron and the u and d quarks whose combinations form the proton and the neutron. Fermions of the second or third generation can be found only in high energy collisions. For example, muons are produced when cosmic rays collide with the particles of the earth's atmosphere. A summary of the already observed SM particles can be found in table 1.1.

The matter comes with two chiralities, left–handed and right–handed. The left–handed fermions are found in isodoublets of the weak isospin, while the right–handed in singlets and vice versa for the antifermions (see table 1.2). In the SM, the neutrinos are assumed massless and only left–handed neutrinos and right–handed antineutrinos are included. However, numerous experiments have given evidence for neutrino oscillations, which imply the existence of non–zero masses for the neutrinos [Ahn 06]. Thus, reference frames exist where a neutrino has inversed helicity. Nevertheless, the neutrino's masses can be included in the SM without major modifications in its structure.

The interactions between the fermions are realised by exchanging gauge bosons. The strong interactions are mediated by the gluons. The gluons are massless bosons that carry colour charge, thus they also interact with themselves. This property leads to colour confinement. There are eight independent states of gluons, constructed from combinations of the three colour and anti–colour quantum numbers. They are described by the $\mathrm{SU}_c(3)$ sector of the formula 1.1. The electromagnetic interactions, described by the $\mathrm{U}_Y(1)$ factor of the formula 1.1, are mediated by the massless photon. Photon does not carry any electromagnetic charge itself: so, even if it feels the charge, there is no photon self–interaction. The weak force is described by the $\mathrm{SU}_L(2)$ factor of the formula 1.1. The force carriers are the massive charged bosons W^+ and W^-, and the neutral Z boson. The weak interaction strongly violates the parity. The charged weak currents (W^+W^- bosons) maximally violate parity since they couple only with left–handed fermions (or right handed antifermions). The neutral weak current (Z boson) also violate the parity, however they interact with fermions of both chiralities. The gauge bosons of the SM are summarised at the second part of the table Table 1.1.

The electromagnetic and weak interactions are combined into the electroweak interaction, being realised within the gauge group $\mathrm{SU}_L(2) \otimes \mathrm{U}_Y(1)$. Its carriers are the W^+, W^- and W^0 bosons of weak isospin with universal coupling g, and the neutral B^0 boson of weak hypercharge with universal coupling g', from the sectors $\mathrm{SU}_L(2)$ and $\mathrm{U}_Y(1)$ respectively. The photon and the Z boson are derived from the mixing of the W^0 and B^0 states:

$$\begin{pmatrix} Z^0 \\ \gamma \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^0 \\ B^0 \end{pmatrix} \tag{1.2}$$

where θ_W is the weak mixing angle. The electromagnetic constant e and the g are

PARTICLE FAMILY	SYMBOL	MASS $[\mathrm{MeV/c^2}]$	ELECTRIC CHARGE $[\lvert q_e \rvert]$
		FIRST GENERATION	
Quarks	u	1.5–3.3	$+2/3$
	d	3.5–6.0	$-1/3$
Leptons	e^-	0.511	-1
	$\bar{\nu}_e$	$\leq 2.2 \cdot 10^{-6}$	0
		SECOND GENERATION	
Quarks	c	1270^{+70}_{-110}	$+2/3$
	s	104^{+26}_{-34}	$-1/3$
Leptons	μ^-	105.7	-1
	$\bar{\nu}_\mu$	≤ 0.17	0
		THIRD GENERATION	
Quarks	t	170900 ± 1800	$+2/3$
	b	4200^{+170}_{-70}	$-1/3$
Leptons	τ^-	1784.1	-1
	$\bar{\nu}_\tau$	≤ 15.5	0

FORCE	GAUGE BOSONS	MASS $[\mathrm{GeV/c^2}]$	ELECTRIC CHARGE $[\lvert q_e \rvert]$
Strong	g (8 gluons)	0	0
Electromagnetic	γ (photon)	0	0
Weak	$W^\pm$	80.3980 ± 0.0250	± 1
	Z^0	91.1876 ± 0.0021	0

Tab. 1.1: Overview of the three generations of quarks and leptons families, along with the mediators of all forces, except the gravitational. The q_e stands for the value of the electric charge of the electron ($q_e = 1.6 \cdot 10^{-19}$ C). [PDG b]

$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} b \\ t \end{pmatrix}_L$
$u_R\ d_R$	$c_R\ s_R$	$b_R\ t_R$
$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$
e_R	μ_R	τ_R

Tab. 1.2: Weak isospin states of the fermions.

related with the θ_W via:

$$e = g \sin\theta_W \tag{1.3}$$

where $\sin\theta_W \sim 1/2$. The weak mixing angle has been measured at the LEP and the Stanford Linear Collider (SLC) with an accuracy of 0.1 %.

This very elegant formalism assumes massless bosons. The presence of a mass term for the gauge bosons breaks the gauge invariance of the electroweak Lagrangian. If one just introduces mass terms in the electroweak Lagrangian by hand, without giving any attention to the induced breaking of the symmetry, non renormalisable divergences will occur. Which means that the infinities arising in higher order calculations, would not be reabsorbed into the Lagrangian. Hence, the theory would lose its predictive power and become meaningless. This, however, is contradictory with what we observe in the Universe. The weak force has a very short range, thus the mediating bosons should be massive. Indeed, the experimentally observed physical particles W and Z bosons are massive. Another evident fact from our everyday experience is that the fermions have also mass. The mass is given to the particles through spontaneous symmetry breaking, even though fermions and bosons acquire their masses with a different way. We will give an overview of this process, as assumed in the SM, in the following section.

1.1.2 The Higgs Field

Through this section, we will discuss the Electroweak symmetry breaking and the subsequent generation of the masses of the gauge bosons, the mechanism responsible for the generation of the fermion masses and the calculation of the vacuum expectation

value of the higgs field.

Electroweak Symmetry Breaking and Generation of the Z, W Masses

The goal is to include the masses of the weak vector bosons $W^{\pm}$ and Z into a renormalisable electroweak theory. Obviously, the photon should remain massless. A solution to this problem is provided by the spontaneous electroweak symmetry breaking. The most simple way to do so is to introduce a new complex scalar SU(2) doublet field Φ, the higgs field. It breaks the electroweak symmetry by acquiring a non–zero vacuum expectation value [Higgs 64].

The following Lagrangian, invariant under local gauge transformations, describes the scalar field and its potential $V(\Phi)$:

$$\mathcal{L} = (D^{\mu}\Phi)^{\dagger}(D_{\mu}\Phi) - V(\Phi) \tag{1.4}$$

where $\lambda > 0$ and $V(\Phi)$ is:

$$V(\Phi) = \mu^2\Phi^{\dagger}\Phi + \lambda(\Phi^{\dagger}\Phi)^2 \tag{1.5}$$

The potential can take two forms, depending whether μ^2 is positive or negative. In the case $\mu^2 > 0$, the potential has a minimum at $\Phi = 0$ and describes a scalar field with mass μ. The interesting part comes with the case that $\mu^2 < 0$. Then an infinite number of minima is found, laying in a circle of radius υ, where $\upsilon^2 = -\frac{\mu^2}{2\lambda}$. These minima have equal probability. However, when one performs a measurement to the system, he must find one of them. So, we select this particular minimum without any loss of generality. The system remains symmetric. This effect is called spontaneous symmetry breaking.

So, the field Φ is expanded around this particular minimum, in terms of the field h(x) that represents the quantum fluctuations around the minimum:

$$\Phi_o = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \upsilon + h(x) \end{pmatrix} \tag{1.6}$$

By substituting equation 1.6 to the Lagrangian 1.4, we can find the masses generated by spontaneously breaking the electroweak symmetry for the gauge bosons, which are described from the following term:

$$(\frac{1}{2}\upsilon g)^2 W_{\mu}^{+}W_{\mu}^{-} + \frac{1}{8}\upsilon^2(W_{\mu}^{3}, B_{\mu})\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix}\begin{pmatrix} W^{3\mu} \\ B^{\mu} \end{pmatrix} \tag{1.7}$$

thus the masses of the W^{+}, W^{-} bosons are:

$$M_W = \frac{1}{2}\upsilon g \tag{1.8}$$

of the Z boson

$$M_Z = \frac{1}{2}\upsilon\sqrt{g^2 + g'^2} \tag{1.9}$$

and $M_A = 0$ for the photon. From equation 1.3, we get the ratio of the masses of the Z and W bosons:

$$\frac{M_W}{M_Z} = \cos\theta_W \tag{1.10}$$

where θ_W is the weak mixing angle. If one expands the potential $V(\Phi)$, one ends up with the following term in the Lagrangian:

$$-\lambda\upsilon^2 h^2 - \lambda\upsilon h^3 - \frac{1}{4}\lambda h^4 \tag{1.11}$$

where the first term gives the mass of the higgs boson, $M_H^2 = 2\lambda\upsilon^2$, while the second is the trilinear and the third the quartic higgs self–interactions. The higgs boson mass is given as a function of the parameter λ, which is one of the free parameters of SM. Thus, the SM cannot predict the higgs boson mass. This is one of the main reasons that makes the searches for the higgs boson so challenging. Its mass is one of the most sought goals of the high energy physics experiments of our days. Once its value is measured, all the other parameters of the higgs boson are automatically fixed inside the SM. So when (or if) the higgs boson is found and its mass is measured, the precise measurement of its parameters will reveal its nature, whether is pure SM nature or not.

Theoretical and experimental constraints exist on the higgs boson mass. They depend on the energy scale Λ, up to which the SM is expected to be valid. The theoretical boundaries on the expected higgs mass as a function of the scale Λ is shown at the figure 1.1. The upper limit is set by the validity of the perturbative expansion. Its value is a few hundred GeV. The unitarity in scattering amplitudes set also an upper limit at few hundred GeV. The lower limit follows from the intrinsic properties of the higgs potential [Branchina 05].

Direct searches at LEP have excluded the region $\leq 114.4\,\mathrm{GeV/c^2}$ at 95 % confidence level. A particle consistent with the Higgs boson, having a mass of $126.5\,\mathrm{GeV/c^2}$, has been found lately (July 2012) by the ATLAS and CMS experiments at the LHC with $5.0\,\sigma$ significance. The region up to $523\,\mathrm{GeV/c^2}$, except of a mass window at $121.8 < M_H < 130.7\,\mathrm{GeV/c^2}$, has been excluded at 99 % confidence level. There is still a possibility for a heavy higgs with a mass $\geq 600\,\mathrm{GeV/c^2}$ to be found. The latest higgs search results from ATLAS experiment ([ATLAS]) are illustrated in figure 1.2.

Very important informations concerning the higgs sector are also stemming from precision electroweak observables. The higgs mass contributes via loop corrections to

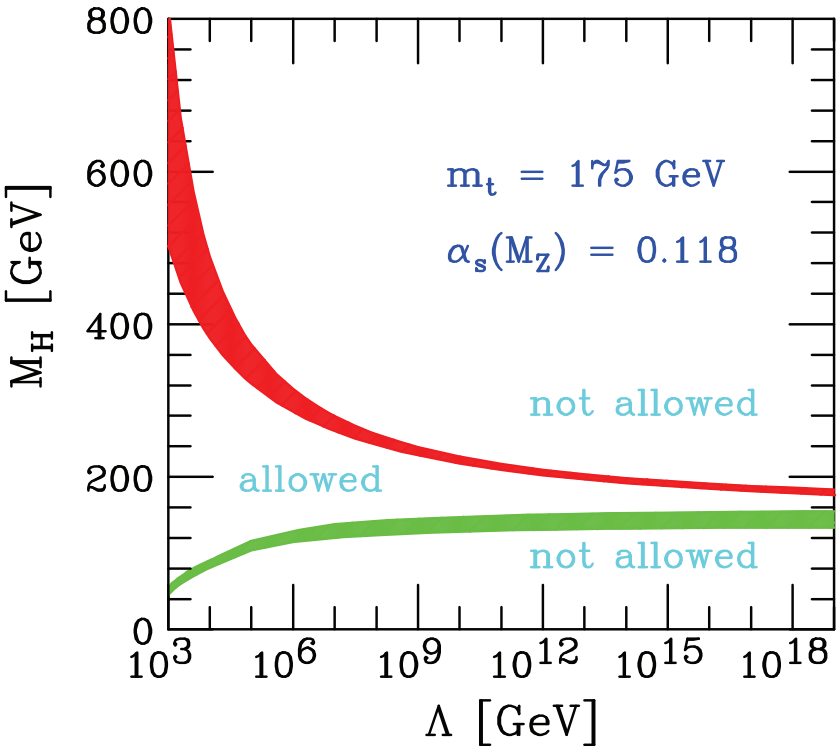


Fig. 1.1: Theoretical constraints on the higgs boson mass as a function of the energy scale Λ, with the assumption that the SM is valid up to that scale. The theoretical uncertainties on the limits are reflected on the width of the constraint bands.

several observables, like the W boson mass or several leptonic and hadronic asymmetries. Precise measurements of these observables at LEP, SLC and Tevatron allow us to constrain the higgs mass. The global electroweak fit to all these observables yields $M_H = 91^{+45}_{-32}\,\mathrm{GeV/c^2}$, with an upper limit of $186\,\mathrm{GeV/c^2}$ at 95 % confidence level. This fit can characterise the predictive power of the SM. The existence of a higgs boson with substantially larger mass requires some new physics, that will compensate for its effect on the loop corrections of the electroweak observables.

Fermion Masses Generation

The same higgs field is also assumed to be responsible for the fermion masses generation. The fermions acquire their masses through a Yukawa interaction between the scalar higgs field and the fermion field. Concerning the electron for example, the gauge invariant sector:

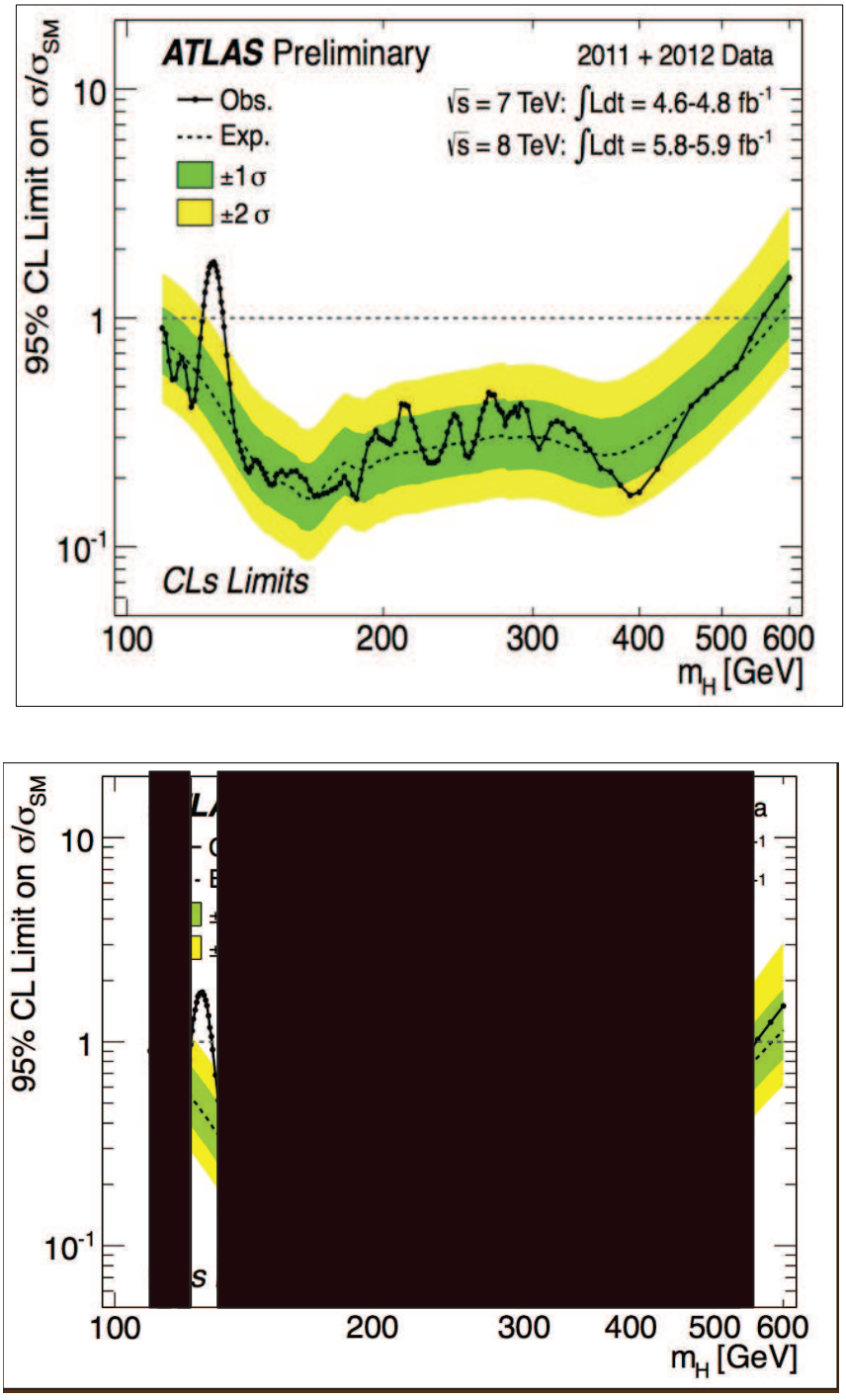


Fig. 1.2: The latest experimental results on SM higgs boson searches from ATLAS experiment. The black solid line corresponds to the experimental results. In the mass ranges where this line is below the horizontal dashed line, the existence of a higgs boson featuring a mass inside that range is excluded at 95 % confidence level. The dotted line shows the expected limit in the absence of a higgs boson. The green band shows the expected region in the absence of a higgs at 68 %, while the yellow band at 95 % confidence level. A significant excess is clearly observed at $\sim 126\,\mathrm{GeV/c^2}$. The shadowed areas at the bottom plot show the excluded mass regions at 99 % confidence level.

$$\mathcal{L} = -G_e \left[(\overline{\nu}_e, \overline{e})_L \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} e_R + \overline{e}_R (\Phi^-, \overline{\Phi}^0) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right] \tag{1.12}$$

is added to the Lagrangian. G_e is the Yukawa coupling between the electron and the higgs field. The mass of the electron will be generated by spontaneous symmetry breaking: we expand around the vacuum expectation value in terms of the higgs field h(x). After breaking spontaneously the symmetry, the Lagrangian takes the following form:

$$\mathcal{L} = -\frac{G_e}{\sqrt{2}} \upsilon (\overline{e}_L e_R + \overline{e}_R e_L) - \frac{G_e}{\sqrt{2}} (\overline{e}_L e_R + \overline{e}_R e_L) h \tag{1.13}$$

The first term of the equation 1.13 is the electron mass term and the second one is an interaction term with the higgs field. We should stress here that the SM does not predict a mass value for the fermions since G_e is not defined. We choose the appropriate values for the Yukawa couplings G_x, where x indicates a fermion, in order to be consistent with the observed fermion masses. Thus the validity of the SM cannot be checked directly versus the fermion masses. On the other hand, according to the second term of equation 1.13, the higgs couplings to the fermions are proportional to their masses. Thus the study of the higgs couplings at the ILC can be a crucial validity test of the SM. Another thing that should be stressed here is the wide span of the fermion masses over five orders of magnitude. The SM does not offer an explanation for this fact.

Higgs Vacuum Expectation Value

The Fermi's coupling constant G_F has been calculated very precisely through measurements of the muon lifetime. G_F is given by the following equation:

$$G_F = \frac{\sqrt{2} g^2}{8 M_W^2} \tag{1.14}$$

where g is the weak coupling constant and M_W the mass of the W boson. Taking M_W from equation 1.8 and substituting to equation 1.14 we get for υ, the higgs vacuum expectation value, that $\upsilon \approx 246\,\text{GeV}$.

1.1.3 SM Free Parameters

The SM contains a big number (26) of free parameters. Their values are not predicted by the SM but, as already mentioned for the fermion masses, need to be measured experimentally. For many years, these free parameters have been the fermion masses (9), the electromagnetic and strong coupling constants, the quark mixing parameters of the CKM matrix (4), the higgs boson mass, the Z boson mass and the weak coupling

constant. The θ_{CP}^{QCD} is a specific parameter belonging to the QCD Lagrangian. If this parameter was non zero, it would allow CP violation to occur in QCD interactions. This comes to a conflict with the experimental results obtained up to date, thus the θ_{CP}^{QCD} should be 0 or extremely small. We lack of an implicit explanation why this happens, and this is the strong CP problem [Dine]. The above parameters count to 19, and with the exceptions of the higgs mass, all the other parameters have been already experimentally measured. Recently, 7 more parameters were added after discovering that neutrinos do have masses. Out of this 7 additional parameters, three of them are for the neutrinos masses and four for the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix [PDG c], which is the equivalent of the CKM matrix for the neutrinos. This brings the total number of the SM free parameters to 26. Out of the 26 aforementioned parameters, 15 of them, which are the particle masses and the mixing parameters of the CKM matrix depend on the higgs field. This number may increase to 22, if also the neutrinos acquire their masses by interacting with the higgs field. Finally, cosmological observations suggest that we need to add one more parameter in order to describe nature: this is the cosmological constant, which expresses the energy density of the vacuum, which is a possible "explanation" for the Dark Energy (DE).

1.1.4 The Challenges of the Standard Model

All the particles predicted from the SM, except the higgs boson, are already found. Observables, precisely measured in several experiments, are consistent with the SM predictions, sometimes at the per mil level. With the exception of the neutrinos masses, which can be accommodated with some rather minor modifications of the SM, there are not other experimental results contradictory to the SM.

Despite its numerous experimental successes, the SM is not the ultimate theory of everything. Its insufficiency is underlined from cosmological observations and theoretical incompleteness. We will give an overview for some of the challenges that particle physics currently faces, in the following, non–exhaustive list.

Higgs boson: The higgs boson has still to be found. Its existence and the precise measurement of its parameters consist an ultimate test for the SM. A definite answer concerning its existence is expected to be derived in the near future by the LHC experiments. The higgs boson parameters may be precisely measured via an LHC–ILC interplay.

SM free parameters: Their existence do not consist a per se problem for the SM. But we lack a deeper understanding for their values. For example, as already mentioned, why the values of the fermion masses spread over five orders of magnitude?

Gravity The SM does not include one of the four fundamental interactions, the gravity.

Cosmological observations: Several independent cosmological observations converge that only $\sim 5\,\%$ of the Universe is explained by ordinary matter. $25\,\%$ of the energy density in the Universe is constituted by Dark Matter (DM) and the remaining $70\,\%$ by DE. The origin and nature of the DM and the DE are unknown, and they cannot be explained by the SM [Spergel 03].

Hierarchy problem: The SM does not predict the higgs mass. It however sets some theoretical bounds (see figure 1.1) with respect to Λ that limits the higgs mass to be $\leq 1\,\text{TeV}$. Taking into account the electroweak observables, the SM higgs boson mass has to be less than 200 GeV. On the other hand, quantum corrections to its mass via top quark, W, Z bosons and the higgs boson itself loops, are Λ^2 divergent, resulting in a huge mass. In order to avoid that, some very delicate cancellations between the quantum corrections should occur, a topic known as "fine tuning". For example, if Λ is at Planck scale, the first 33 decimal places must cancel each other. Even though fine tuning is not a showstopper for the SM, it seems quite unnatural, and we lack an implicit explanation why it occurs [Veltman 77].

CP violation: One would expect that at the Big Bang, equal amounts of matter and anti–matter have been created. In our days, we observe a baryon asymmetry in Universe. Therefore Universe, for some reason, favours matter at the expense of anti–matter. An explanation of this fact may be given via the CP violation. CP violation has not been observed in QCD. On the contrary, it is observed in weak interactions, but it cannot explain the observed amount of matter / anti–matter imbalance in the Universe. The explanation should lie in physics beyond the SM. [Sakharov 67]

GUT scale: The SM fails to find a common point where the coupling constants of the three fundamental interactions have equal strength. This point is called the Grand Unification Theories (GUT) scale.

1.1.5 Beyond the Standard Model

The aforementioned considerations are among the main motivations driving the current and future particle physics experiments, like the LHC and the ILC. Several theories that have the ambition to answer these questions, await for feedback from the experiments. These theories include Supersymmetry, Technicolour, String theory, Extra dimensions, Supergravity etc. Some of them try to extent appropriately the SM, striving to address

its weaknesses. To make an exhaustive list or give a detailed description is outside of the scope of this section.

Among the most popular is Supersymmetry (SUSY) [Singer], which is the only renormalisable Quantum Field Theory (QFT) among all alternatives of the SM considered up to now. SUSY is an additional symmetry that relates each known elementary fermion with a corresponding boson, and vice versa. The two related particles are called superpartners. They are a priori expected to have exactly the same mass and quantum numbers, except that their spin will differ by 1/2. Such a symmetry has not yet been found in Nature. There is no experimental evidence of any SUSY relating known elementary fermions and bosons. So, if supersymmetry exists, it will be a broken symmetry. This will allow the supersymmetric particle to acquire very high masses, explaining this way why they were not experimentally observed until now.

SUSY extension of the SM addresses a variety of the currently open issues of particle physics. It provides a natural solution to the hierarchy problem. The loop contributions on the higgs mass arising from the bosonic and fermionic superpartners will be mutually cancelled. Thus the hierarchy problem will be solved without the necessity to perform any fine tuning. The SUSY also allows for the unification of the three gauge interactions at a common GUT scale. Moreover, the lightest supersymmetric particle, will be a promising candidate for (at least a part of) the dark matter. Therefore, if SUSY manifestations are found, numerous fundamental questions should be answered. But a new one will also arise: why is SUSY broken?

A large fraction of the allowed parameter space of SUSY has been excluded by the LEP and Tevatron experiments. Within the LHC perspectives, it is expected that the SUSY energy scale is typically around 1 TeV. If this scale will be found to be well above 1 TeV would actually contradict the benefit of SUSY for controlling mass divergences. Thus if it exists, supersymmetric particles should be discovered at the LHC.

Alternatively, the Technicolour [Dimopoulos 79] theories explain the mass generation by introducing a new strong gauge interaction. This interaction substitutes the role of the higgs field for the spontaneous EW symmetry breaking, thus for the generation of the masses. Without assuming any higgs boson, it avoids the fine–tuning problem. A number of new resonant particles, likely to be accessible at the LHC energy, are predicted. These resonances are mostly decaying to top quark pairs.

Another active area of research for theories beyond the SM is the field of extra dimensions. They were introduced inside the framework of the efforts to unify quantum mechanics with gravity. Extra dimensions are predicted by String theories, that foresee 10 or 11 space–time dimensions. They also consider as a possible explanation for the hierarchy problem between the weak and the Planck scale (large extra dimensions, [Arkani-Hamed 98]). Their signature in collider experiments could be missing energy or modification of the differential cross sections for bosons or fermions pair production.

1.2 Physics at the ILC

The necessity for precision measurements was already underlined in several cases in the previous sections, where a brief overview of the landscape of the current physics issues took place. The ILC, as we will explain in the second chapter, offers an ideal environment for precision physics. The challenges on the particle physics listed in the section 1.1.4 consist the main motivations to proceed to the ILC.

To understand the nature of the Higgs boson found at LHC, detailed measurements of its properties are required, in particular of the Higgs couplings to the EW bosons and to the elementary fermions. Such precision measurements are a major priority for the upcoming LHC physics programme and have triggered an effort to achieve a ten-fold increase of the machine luminosity in the next decade.

The expected ultimate precision on the Higgs couplings [Hig 13] tends however to be limited by systematic uncertainties, resulting in accuracies which may not allow to unravel the exact nature of the Higgs boson.

Moreover, the precision on the invisible width, which provides a unique window on potential new physics (e.g. DM particles), will be determined with relatively modest precision. Finally, it is worth noticing that the LHC does only allow measuring rates expressing products of cross-sections with branching fractions and ratios of branching fractions; it does not allow to measure directly couplings, which would require determining the total width. The couplings are therefore derived from the measurements by assuming the SM value of the total width.

The ILC, on the other hand, allows for a direct determination of the total width using the Higgsstrahlung and the WW-fusion processes studied at 250 and 500 GeV. The couplings are measured in a model-independent way without assumption on the total width, and with a precision which can be about ten times better than at LHC for certain final states [Battaglia 01]. Finally, measurements at the ILC will be limited by statistics, meaning that any luminosity upgrade of the machine will benefit directly to the experimental precision.

The top quark physics studies that could be performed at the ILC are also of extreme interest. The accuracy on the measured top mass is one of the dominant sources of uncertainty for the electroweak precision tests. The ILC aims to measure the top mass with an uncertainty of $\sim 100\,\text{MeV}$. A discussion on what such an accuracy will bring, can be found at [Heinemeyer 03]. Furthermore, the top Yukawa couplings measurement will be an important test of the fermions masses generation via the higgs mechanism. Following the same line of thought, the precise measurement of other electroweak observables at the ILC, like the Z, W^+, W^- bosons masses and the weak mixing angle, will allow to further constrain the SM higgs boson mass, and to perform this way crucial tests on the validity of the SM. The high peak luminosity of the ILC, as well as its ability to vary its collision energy, will allow to accumulate very high statistics in the production either of top quark pairs, or gauge boson, leading this way

to very precise measurements of their parameters. Additionally, the foreseen beam polarisation will allow to effectively suppress background processes.

If SUSY exists, some supersymmetric particles ($\tilde{g}$, $\tilde{q}$) are likely to be discovered at the LHC. The mass spectrum of those particles will be partially directly accessible at the ILC($\tilde{l}$) . If finally SUSY particle candidates are discovered, a number of crucial questions will arise; the most straightforward among them will be whether it is really SUSY. To answer this question, we have to measure the quantum numbers of the new particles. The ILC offers an ideal ground to establish SUSY, with precise Yukawa and gauge couplings measurements and comparisons, spin verification and particle masses measurements (e.g. through threshold scans). In certain new particles searches, heavy flavour tagging and maybe, also vertex charge identification are crucial. The clean environment of the ILC, allows for the vertex detector to be placed closed to the Interaction Point, providing this way these informations with the required accuracy. Other very important questions that the ILC can contribute are how SUSY is realised in Nature, and why and how SUSY is broken.

A very interesting interplay can be established between the ILC and the cosmology. The nature of the DM is still unknown, but it is clear that it cannot be entirely explained with SM matter. The supersymmetric extensions of the SM, as well as the theories beyond the SM, provide candidates for DM. These candidates can be detected in collider experiments (LHC and ILC). Their expected signature can be cascade decays or elementary interactions with large amounts of missing energy. If DM particles will be discovered in collider experiments, the precise measurement of their mass and other properties is crucial. This precision measurements could allow us to understand which new physics model predict the discovered DM candidate particles. Using the model parameters, the relic abundance can be estimated. Then this estimation should be cross–checked with the cosmological observations, to find out whether they are consistent. Therefore, we will be able to know whether the anticipated model can explain the whole DM observed in the Universe. This roadmap could reveal the true nature of the DM. A second cosmological aspect that the ILC intents to search is the baryon asymmetry of the Universe. Several supersymmetric extensions of the SM try to give an explanation of this subject. Thus, searches for new particles and precise determination of their properties at the ILC may shed light on this mystery also.

1.2.1 Standard Model Higgs Physics at the ILC

One of the studies performed within the framework of the thesis, is the evaluation and optimisation of an ILC vertex detector performance on flavour tagging and the extraction of the higgs hadronic branching ratios. These studies are presented in chapter 5. They were based on a SM higgs boson, produced via the higgsstrahlung process. For this reason, we will finish the introduction with an overview of the production and decay processes of a SM higgs at the ILC.

Production of the SM Higgs at the ILC

The dominant production mechanisms for the SM higgs boson at the ILC are the higgsstrahlung $e^+e^- \to ZH \to f\overline{f}X$, where X are the decay products of the higgs and the WW fusion $e^+e^- \to \nu\overline{\nu}H$. The figure 1.3 shows the Feynman diagrams of these two processes.

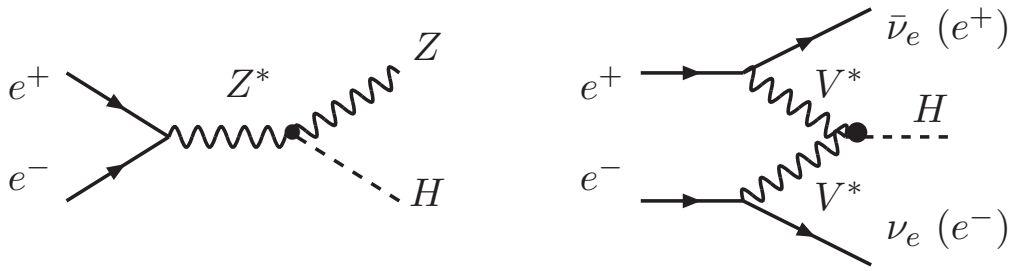


Fig. 1.3: Leading order Feynman diagrams of the dominant mechanisms for higgs production at the ILC. The higgsstrahlung process is illustrated at the left part of the figure, while at the right part shows the higgs production through the fusion of two virtual vector bosons. In case of W bosons we have two neutrinos in the final state, whereas in case of Z bobosons, an e^+e^- pair. WW fusion dominates by an order of magnitude the ZZ fusion.

The two dominant production mechanisms have similar cross section, O(50fb) for a higgs mass in the range of $115GeV \leq M_H \leq 200GeV$ at a collision energy of $\sqrt{s} = 500\,\text{GeV}$. The cross sections as a function of the higgs boson mass M_H for $\sqrt{s} = 500GeV$ and $\sqrt{s} = 1TeV$ are shown at the figure 1.4. At lower $\sqrt{s}$, the higgsstrahlung process dominates the higgs production. It sharply increases after the production threshold, reaching a maximum at $\sqrt{s} \sim M_Z + \sqrt{2}M_H$. After the peak, the cross section scales as $\sim 1/s$. The WW fusion dominates at the highest values of $\sqrt{s}$. It increases essentially like $\log s/M_H^2$.

Detection of the Higgs at the ILC

In the higgsstrahlung process, the Z boson is monoenergetic. At the ILC, the initial state of the event is accurately known. Thus the higgs boson can be reconstructed out of the recoiling system of the Z, without making any assumption concerning the higgs decay. This way, the higgs mass and the higgsstrahlung cross section can be defined in a model–independent way. The higgs mass M_H will be calculated from the following equation, which expresses energy conservation:

$$M_H^2 = s + M_Z^2 - 2\sqrt{s}(E_1 + E_2) \tag{1.15}$$

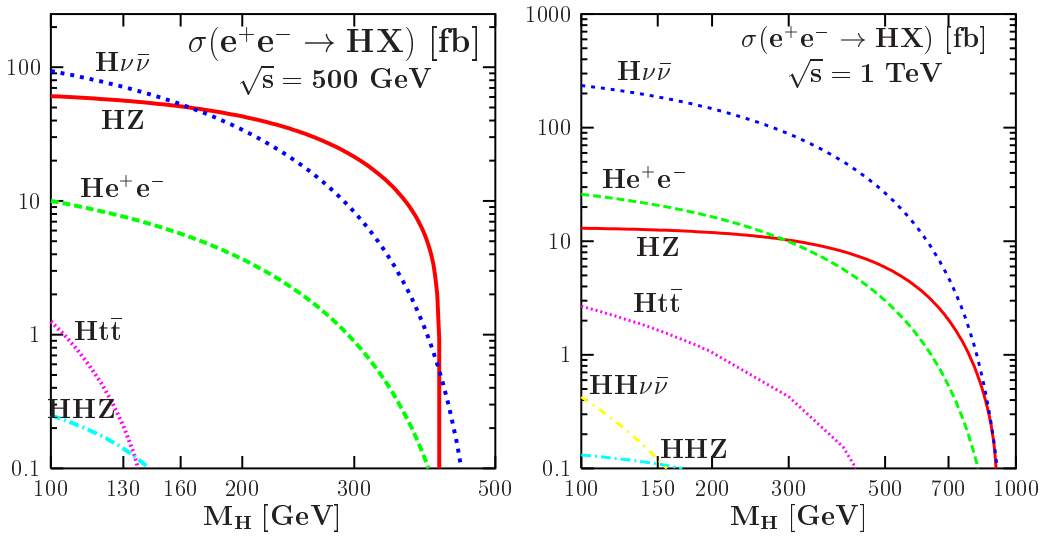


Fig. 1.4: Cross sections for the main production mechanisms of higgs at the ILC as a function of M_H for $\sqrt{s} = 500GeV$ (left plot) and $\sqrt{s} = 1TeV$ (right plot), obtained with the HPROD ([Djouadi a]) software. Radiative correction are not taken into account. ([Djouadi b])

where M_Z is the mass of the Z boson, reconstructed out of its decay products that feature energies E_1 and E_2. The Z can decay either to a pair of charged leptons, either to a quark pair or invisibly to neutrinos. All of these decay channels can be exploited in order to increase the statistics for the higgs mass calculation. The clearest signal, thus the best signal to background separation, is offered by the decay channels $Z \to \mu^-\mu^+$ and $Z \to e^-e^+$. In order to maximise the accuracy, the radiative effects due to the beamstrahlung (see section 2.2.1), the Initial State Radiation (ISR) and the Final State Radiation (FSR) should be effectively suppressed. Otherwise they are going to introduce high tails in the recoil mass distribution. The optimal $\sqrt{s}$ for this analysis is at 250 GeV, assuming $M_H = 120\,\text{GeV}$ where the cross section of the higgsstrahlung process is highest and the reconstruction of the Z boson is the easiest [Ito 09].

Higgs Decay at the ILC

According to the SM, once the mass of the higgs boson is determined, its properties are uniquely defined. The branching ratios of a SM higgs as a function of its mass are illustrated at the figure 1.5, being obtained by the HDECAY software ([Djouadi 98]). All the kinematically allowed higgs decays with branching ratios $> 10^{-4}$, including the loop mediated ones and the decays to off–shell vector bosons, are presented in the figure. It also includes the QCD and electroweak radiative corrections. The higgs couplings to the fermions are proportional to their mass. Thus, the scale of the fermionic higgs

branching ratios are defined by the fermion's mass. As a consequence, the precise measurement of the fermionic higgs branching ratios consists a crucial test that can reveal whether the detected higgs has a SM nature or not.

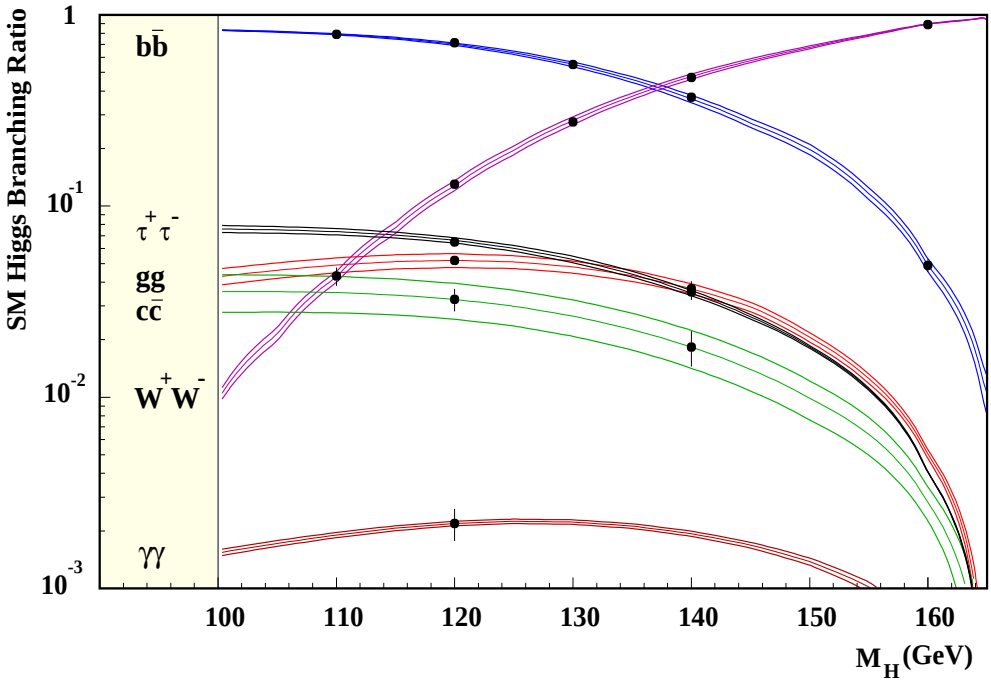


FIG. 1.5: The theoretical predictions for the branching ratios of the SM higgs boson as a function of its mass. The bands express the introduced uncertainties due to the uncertainties on the measurements of the quark masses and the strong coupling constants. The error bars of the points indicate the expected sensitivity of their measurement at the ILC, assuming $\sqrt{s} = 350\,\text{GeV}$ with $500\,\text{fb}^{-1}$. [Aguilar-Saavedra]

A light SM higgs, having a mass $M_H \leq 140\,\text{GeV}$, will mostly decay into $b\bar{b}$ pairs with a branching ratio of O(80 %). Other important decay channels are the $c\bar{c}$, $\tau^+\tau^-$ and pairs of gluons, featuring branching ratios of O(5 %). Since, as we already mentioned, the higgs couplings to the fermions are proportional to their mass, the branching ratios to light quarks are expected to be negligible. Therefore, the study of the SM nature of the higgs via its couplings will be performed by extracting the branching ratios to $b\bar{b}$ and $c\bar{c}$ pairs. Here we can stress the importance of a vertex detector that will provide excellent flavour tagging. For masses $M_H \geq 140\,\text{GeV}$, the higgs decays predominantly to gauge bosons pairs.

1.3 Conclusion

Through this introductory chapter, we have tried to briefly summarise the main challenges that fundamental physics has currently to confront with. On the one hand, these are the establishment or rejection of the SM, through the unveiling of the mechanism of the EW symmetry breaking and the understanding of how particles acquire their masses. On the other hand, SM is not the ultimate theory of Nature. A number of questions rise from its shortcomings, like to give a profound solution to the hierarchy problem, shed light to the nature of Dark Matter and Dark Energy, explain the baryon asymmetry in Universe and finally embede the gravity and unify all interactions in a more general theory.

Some of the above points can be adressed via collider experiments. For example, the LHC is expected to find the higgs boson(s), if exist(s). The task that comes directly after the potential discovery of the higgs boson is to understand its nature, meaning whether it is the higgs boson described by the SM or something else. In order to address this question, we need to measure very precisely the higgs properties. One of the most important among them, is the higgs hadronic branching ratios. These kind of studies need detectors featuring unprecedented sensitivity, and a machine that offering an environment that allows for high precision measurements. This machine can be an electron–positron linear collider. A proposed one is the ILC. In the next chapter, we are going to make an introduction to the ILC and one of its proposed detectors, the ILD. The requirements and optimisation of the detector, in order to maximise its sensitivity in the extraction of higgs hadronic branching ratios, will be the main subject of this thesis.

CHAPTER 2

ILC and ILD

A large fraction of the High Energy Physics (HEP) community has converged on the necessity of an electron–positron linear collider as the next large world wide machine. Currently, LHC has started running, and the world of physics is expecting ground-breaking results. So why already foresee another large multi–billion project in HEP? Especially in a period of global financial crisis? And why an electron–positron linear collider?

The answer lies in the complementary nature of the lepton and the hadron particle colliders. A hadron collider, having a high collision energy at the center of mass (hereafter $\sqrt{s}$), is a very powerful tool for direct discovery. The LHC for example, featuring $\sqrt{s} = 14\,\mathrm{TeV}$, has a limit for direct discovery of a few TeV, in case of single particle production. Here we should stress that large mass objects may be discovered at hadron colliders provided their cross section is large enough, and their final states can be reconstructed. The lepton colliders feature generally lower $\sqrt{s}$. Their value lies mainly in the fact that the provided experimental conditions allow for high sensitivity investigations. The complementarity of these two different collider types can be also underlined by the fact that they cover different parts of the mass scale sensitivity spectrum with respect of their $\sqrt{s}$. For example, sleptons of $\sim 250\,\mathrm{GeV/c^2}$ can be detected at the ILC but not at the LHC, even though they are kinematically accessible. On the other hand, LHC can directly probe SUSY particles of higher masses, which are not kinematically accessible at the ILC. It should be also mentioned that important discoveries have been made also at lepton colliders: the gluon at the e^+e^- storage ring PETRA and the τ lepton and the charm quark at the e^+e^- SPEAR collider.

The main advantages of lepton colliders can be summarised as:

- The leptons are structureless objects, at least up to the energies achieved at particle colliders until now. Therefore, during each collision the full $\sqrt{s}$ is available for the elementary process involved. This fact allows us to know precisely the initial four–vector of the momentum, thus we can fully reconstruct the event. On the contrary, the hadron colliders are actually parton assembly colliders. Only a part of the $\sqrt{s}$ is available for particle creation during each collision (typically 0.1 $\sqrt{s}$), so the initial four–momentum of an event is not known.

- Moreover, at the hadron colliders, the partons that do not participate in the elementary process contribute to parasitic (soft) interactions (QCD background). These mask the elementary process of interest, in particular at small polar angles, and exhaust the radiation tolerance of the detectors. It also imposes a selective trigger implementation. The lepton colliders, due to the fact that the leptons are elementary particles, do not suffer from this kind of background. Additionally, the event rate is low, compared to a hadron collider of similar energy. However the cross sections of the interesting electroweak physics processes are only 1–2 orders of magnitudes lower than the total cross section, see figure 2.1. This provides a relatively clean environment, ideal for precise measurements. The detectors, due to the low event rate, do not have to cope with extreme data rates. Thus, they can function in triggerless mode, being this way more sensitive to any possible new physics. For comparison, at the LHC, the total cross section of the interesting processes can be 8–11 orders of magnitude smaller than the one of the QCD background.

- The beam energy (and thus $\sqrt{s}$) is tunable, and both electron and positron beams can be polarised. Therefore the spin of the initial state is accurately known and tunable. Depending on the reaction under study, the selection of an appropriate polarisation can allow for the enhancement of the signal and the suppression of the background cross sections. The advantages on the ILC physics expected from the polarisation of both beams are well explained in [Moortgat-Pick 08].

The complementarity of hadronic and leptonic colliders can be easily illustrated with the following examples. The Z boson was discovered at the Super Proton Synchrotron (SPS), a proton–antiproton collider at CERN. The detailed study of Z boson, which was crucial in order to establish the SM, took place at a lepton collider, the Large Electron Positron collider (LEP), located also at CERN. The sensitivity of a lepton collider can indirectly probe for new physics much further than its $\sqrt{s}$. Small quantum corrections that can be precisely measured at lepton colliders could be a hint for new physics since they are influenced by the whole structure of an anticipated model. For example the mass of the top quark was predicted at the LEP, and the particle was finally discovered at the proton–antiproton collider Tevatron, at Fermilab. Moreover, LEP was the first machine able to search for the SM Higgs boson in the mass range

$\lesssim 100\,\mathrm{GeV/c^2}$, and was the first machine fixing upper bounds on its mass. A very detailed study for the LHC–Linear Collider interplay can be found at [Weiglein 06]. Last but not least, a particle having a very challenging decay for the LHC experiments (for example fully hadronic), could be missed at the LHC due to the vast amount of QCD background and the detector's triggers. The triggerless function of the lepton collider detectors allows to be potentially unbiased to any possible signature of new physics.

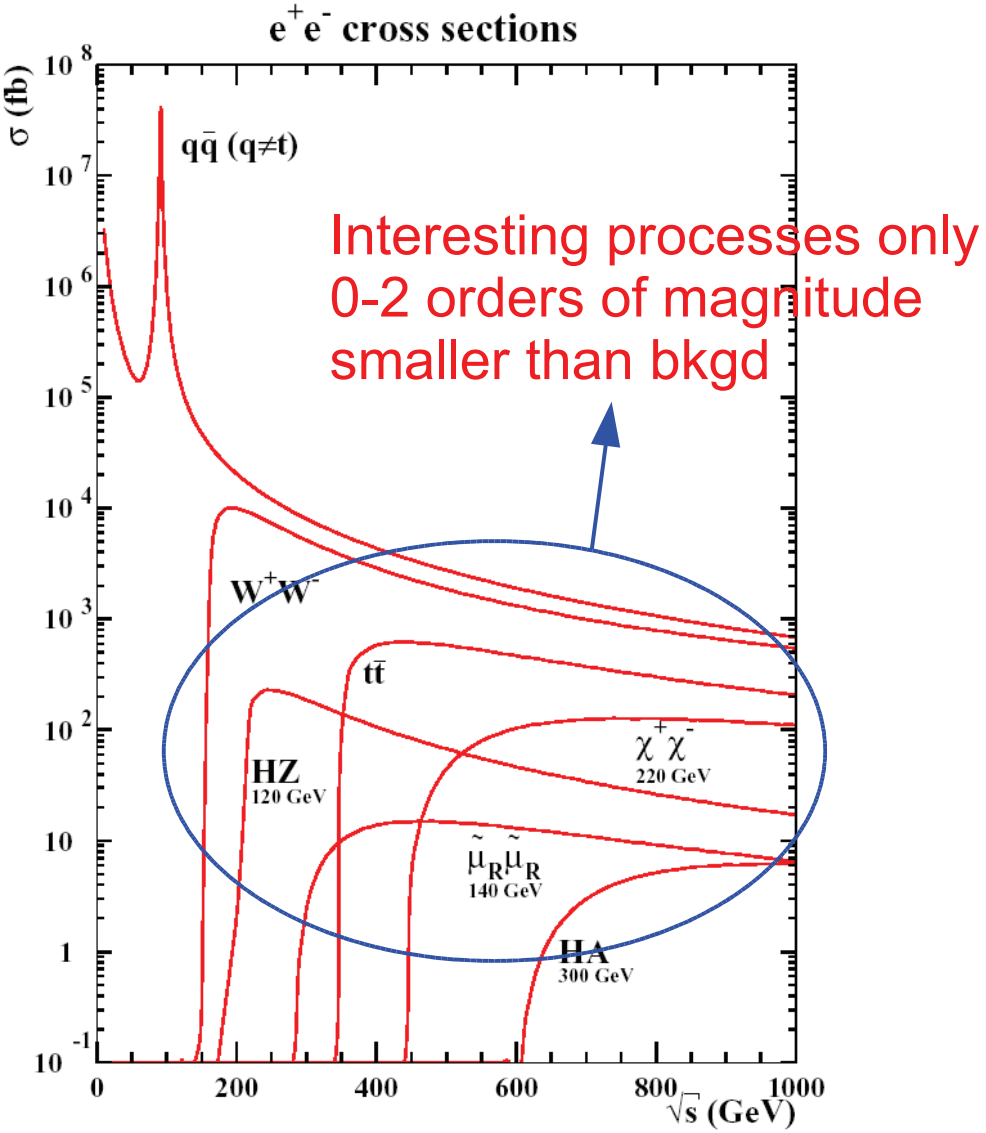


Fig. 2.1: Cross sections of several physics processes at the ILC as a function of the collision energy [Wienemann].

But why a linear and not a circular collider? Circular colliders offer a number of advantages. The particle bunches circle many times, being accelerated over an extended period of time. At a linear collider, the particles should reach the desired $\sqrt{s}$ in only one accelerating path. Therefore a linear collider would require a bigger number of accelerating cavities and would be, in general, bigger than a circular collider providing the same $\sqrt{s}$. The particles bunch repetition rate at the linear colliders is constrained by the damping ring (see sub–section 2.1.1) store time. Therefore it is 1–2 orders of magnitude lower than the one at the circular colliders. So in order for the linear colliders to reach a high luminosity (see equation 2.4), very small beam spot sizes should be achieved.

The choise for a linear collider is imposed by Bremmstrahlung. The trajectories of the particles at a circular accelerator are of course bent. This means that they lose energy through synchrotron radiation. Equation 2.1, which describes the synchrotron radiation emitted by a particle following a circular trajectory:

$$\Delta E_{syn} \propto \frac{E^4}{Rm^4} \tag{2.1}$$

where E is the energy of the particle, R the radius of its trajectory and m its mass. Due to the fact that the radiative energy loss is proportional to the particle's energy at the fourth power, a circular electron–positron accelerator with energy well beyond the one of the LEP is not plausible. It should have an extremely big radius. Thus, the energy loss via synchrotron radiation imposes that the next high energy lepton collider should be a linear one. Concerning hadron colliders like the LHC, equation 2.1 does not consist a problem due to the presence of the mass term at the fourth power at the denominator.

Currently, two projects for electron–positron linear colliders are under study, the ILC [ILC] and the Compact LInear Collider (CLIC) [CLIC]. A different approach, pursued by a part of the HEP community, is to address the same questions that the ILC and the CLIC try to answer, with a muon collider [MuonCollider]. Here it should be noted that equation 2.1 does not pose a problem for a muon collider, due to the large muon mass, thus it could adopt a circular design. Additionally the higher mass of the muon allows a higher $\sqrt{s}$. Nevertheless the complications imposed by the fact that the muons are unstable particles have not yet being solved.

The aforementioned experiments are merely motivated by the same fundamental questions. At this point one may raise the question, if both the ILC and the CLIC should be constructed and if not, which one of them. The answer of the first question is quite obscure. It depends not only on scientific, but also on political and financial criteria. The second question is also difficult. Each of the projects has its own advantages. The ILC features a more mature technology, and could run in parallel with the LHC, or the super LHC. That would give the invaluable opportunity of crosscheck and interplay between the LHC and the ILC. On the other hand, the CLIC, has a

more challenging technology which can make accessible a significantly higher $\sqrt{s}$. Yet it has to demonstrate its proof of principle. However, the physics needs cannot be absolutely predictable in advance. Therefore the early LHC results may show the way. Concerning the detectors, they share the same basic philosophy. Nevertheless, since the running conditions of each accelerator will impose substantially different environmental conditions, the detector optimisation will follow different roads. This thesis is studying a vertex detector for the ILC. From now on, we will mainly focus on this collider.

2.1 The ILC

The ILC is a proposed linear electron–positron collider. It will have a tunable $\sqrt{s}$ of 200-500 GeV, with an option of upgrade up to 1 TeV. Its length stretches to 31 km, and will reach 51 km in the case of the upgrade. Its design should also allow for electron–electron, photon–photon and electron–photon collisions.

The two main characteristic parameters of a collider are the $\sqrt{s}$ and the luminosity. Obviously, the higher the $\sqrt{s}$, the higher the mass spectrum available during the collisions for the creation of particles, as predicted by the equation relating the energy with the mass:

$$E = mc^2 \tag{2.2}$$

The second main reason that calls for high energy particle collisions is explained by De Broglie equation which relates the particle's momentum (p) with their wavelength (λ):

$$\lambda = \frac{h}{p} \tag{2.3}$$

so the higher the momentum, the finer the details we can probe in the structure of the matter.

The luminosity, assuming symmetrical beams featuring a Gaussian cross section, is expressed by the following equation:

$$L = \frac{n_b N^2 f_{rep}}{4\pi\sigma_x\sigma_y} \cdot H_D \tag{2.4}$$

where n_b is the number of particle bunches, N is the number of particles per bunch, f_{rep} is the repetition rate, σ_x and σ_y are the dimensions of the beam transverse profile and H_D is a multiplicative factor. The latter depends on the beam–beam interaction and enhances the luminosity by a factor ~ 2 (see section 2.2 about the pinch effect). The luminosity dictates the statistics that we expect by a certain process. If the cross section of the process is σ, then the event rate will be:

$$\frac{dN}{dt} = L\sigma \tag{2.5}$$

The ILC is deigned for a peak luminosity of $2 \times 10^{34} cm^{-2} s^{-1}$ at $\sqrt{s} = 500\,\text{GeV}$. Luminosity and energy are correlated in a certain extent, which calls for a compromise allowing for a minimal sensitivity at a given collision energy. This compromise is found at the ILC by relying on the computable cross sections of various fundamental processes of interest.

2.1.1 ILC Baseline Design

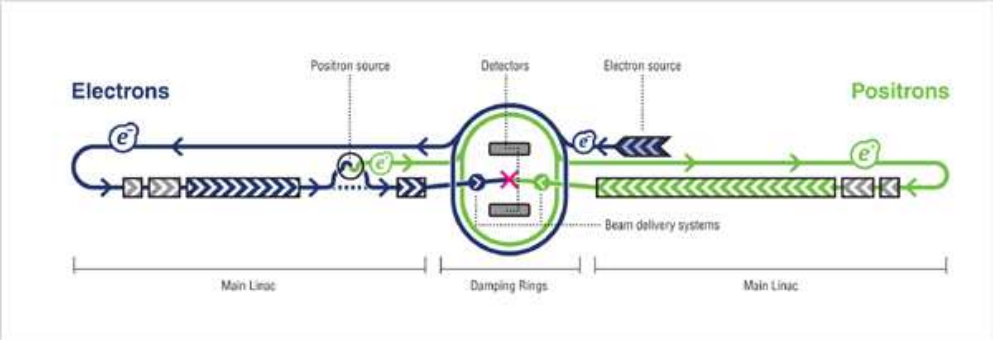


FIG. 2.2: Schematic layout of the ILC for the 500GeV option.

Figure 2.2 illustrates the baseline design of the ILC. Its mains components are the electron and positrons sources, the damping rings, the main linacs, the Beam Delivery System (BDS) and the Interaction Region (IR). It is designed to host two general purpose detectors that will share the same IR. They will be alternatively positioned on the IR using a push–pull scheme. In the following subsections we will briefly describe the main components of the ILC. A detailed description of the ILC can be found at the accelerator's Reference Design Report (RDR) [RDR a].

Electron Source

The electrons are produced by two independent laser–DC guns systems. The laser will fire high intensity, 2 ns pulses of circularly polarised photons on a photocathode

in order to produce polarised electrons. Right after their production, the electrons will be preaccelerated up to 76 MeV with the help of non–superconducting accelerating structures. After that, they will be injected in a 250 m long superconducting linac where they will reach the energy of 5 GeV.

Positron Source

The design of the positron source is more challenging than the electron one. The electron beam, when reaching an energy of 150 GeV, goes through a 150 m helical undulator, before returning to the main linac. When passing through the undulator, the electrons produce ~ 10 MeV photons that are directed to a Ti–alloy target. There, they produce electron–positron pairs. The positrons are separated from the electrons and the remaining photons which are dumped. The positron beam is first preaccelerated to 125 MeV at a normal conducting linac, transferred over 5 km and finally accelerated to 5 GeV at the superconducting Boost Linac before being injected into the dumping rings. The electron beam, after passing through the undulator where it undergoes an average energy loss of ~ 3 GeV, is transferred back to the main linac.

Damping Rings

The electron and positron beams, as produced from their sources, they feature a large emittance. In other words the dimensions and density of the particle bunches, as well as their momentum phase space, are quite outspread. In order to reach the desired luminosity, the emittance should be reduced. This takes place in the damping rings. There, magnets and wigglers force the particles to follow a bent track. They emit radiation in both transverse and longitudinal planes, but their energy is restored only on the longitudinal one. After 25 ms of damping time, the bunches obtain the desired emittance to be introduced into the main linacs. The radius of the damping rings is 6.7 km.

Main Linacs

After the extraction from the damping rings, the bunches follow a transfer line to the Ring To the Main Linac (RTML) structure. The RTML compresses the bunch length by a factor of 30–45 and accelerates them up to ~ 15 GeV. It also rotates the spin of the particles according to the desired polarisation at the collision point and finally delivers them to the main linacs. There, the particles will be accelerated to their final energy of 250 GeV. The main linacs use 1.3 GHz superconducting cavities, made of niobium, offering an average accelerating gradient of 31.5 MeV/m. Their working temperature is 2 Kelvin. Each linac hosts ~ 8000 cavities. Since the particles should reach their maximum energy in only one passing through the linacs, the total length of each main linac will be 11 km.

The accelerating cavities are organised in RF units. One RF unit is constituted from three SuperConducting RF (SCRF) cryomodules, containing in total 26 cavities. Each RF unit has its own RF source that includes a 120 kV modulator, a 10 MW multi-beam klystron and a waveguide which distributes the power to the cavities. The RF source, together with the power supplies and other support equipment, is hosted in a service tunnel, that runs parallel to the main linac tunnel. The main reasons for the service tunnel are to permit easy accessibility and reduce the radiation exposure of the RF source [RDR a]. This twin-tunnel design, which is the baseline design of the ILC, is currently under consideration, merely for cost reasons (see section 2.1.3).

Beam Delivery System

The scope of the BDS is to transfer the beams from the main linacs to the IR, and appropriately focus them in order to satisfy the luminosity goals of the ILC. It brings the beams to collide with a 14 mrad crossing angle. This allows for separated extraction lines for the 2 beams. Due to the elongated shape of the bunches, this crossing angle will lead to poor mutual penetration of the 2 bunches. To restore the luminosity, the bunches are rotated in the horizontal plane. This technique, which is illustrated by figure 2.3, is called crab crossing. The BDS serves other purposes as well. It should protect the detectors from miss–steered beams. It should also offer a screening from the beam halo muons. The task of the deflection of the muons is fulfilled by magnetised iron shielding.

Interaction Region

The ILC may host two general purpose detectors, the International Large Detector (ILD) and the Silicon Detector (SiD). Its initial design had foreseen separate interaction regions for each detector. This would have required two separate BDS, each of them being ~ 4 km long. Cost considerations finally led to the decision of a common IR, that will be shared by both detectors. The detectors will function alternatively according to a push–pull scheme. The push–pull scenario is depicted in the figure 2.4.

The ILD is planned to be placed on a $15 \times 20 \times 2\,m$ thick moving concrete platform that will transport the detector from the beam line to the garage position. This way of transportation has been chosen in order to minimise possible damages due to the transportation. It is inevitable that some substantial amount of time will be lost for mounting–dismounting and transporting the detectors. It is obvious that this time has to be as little as possible in order to exploit the beam time. The incoming detector needs also some time for alignment and calibration. The alignment requirements for the axis of the detector are ± 1 mm and $\pm 100\,\mu$rad, while for the final quadrapole magnet they are $\pm 200\,\mu$m and $\pm 5\,\mu$rad. Concerning the ILD, the calibration is expected to be realised by taking about $1\,\mathrm{pb}^{-1}$ at $\sqrt{s} = 91.2$ GeV, where the cross section for the Z

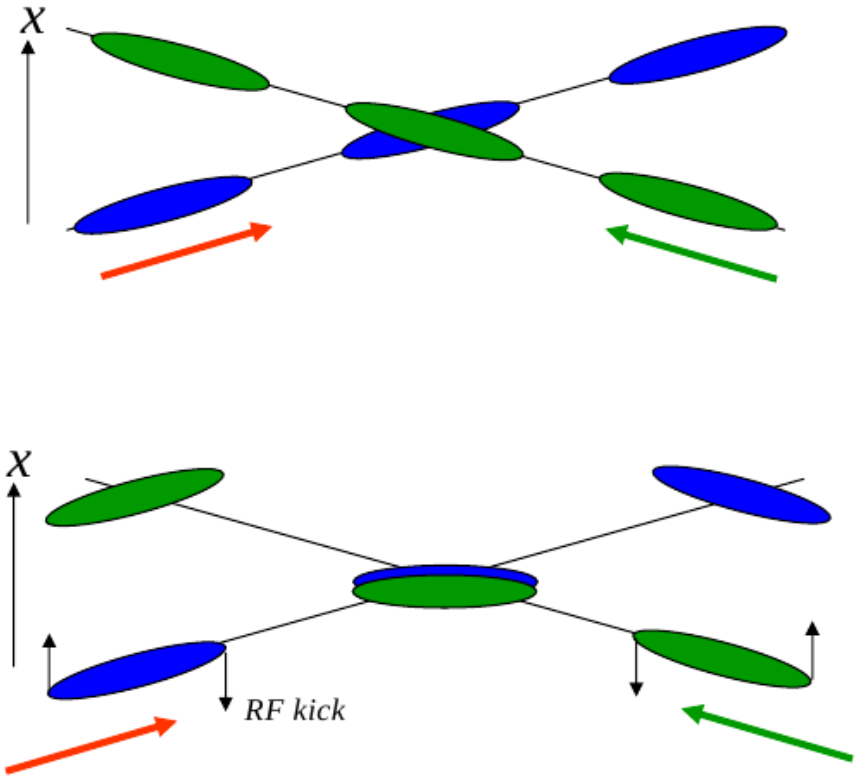


Fig. 2.3: Illustration of the crab crossing technique. The green and blue oval surfaces represent the bunch transverse section of the colliding beams. At the top of the figure one can observe the small overlapping area of the two beams, which leads to a loss of luminosity. Using the crab crossing technique (bottom of the figure), the overlapping area is maximised. From [Zimmerman]

boson production is maximised. Of course this assumes that the switch between the center of mass energies can be performed easily.

2.1.2 Beam Parameters

The beam parameters are designed in order to achieve the desired luminosity of $2 \times 10^{34} cm^{-2} s^{-1}$, while keeping as low as possible the beamsstralung effect (see section 2.2). The table 2.1 summarises the main beam and Interaction Point (IP) parameters for $\sqrt{s} = 500\,\text{GeV}$. The selection of this set of beam parameters is a result of the optimisation of the whole accelerating structure.

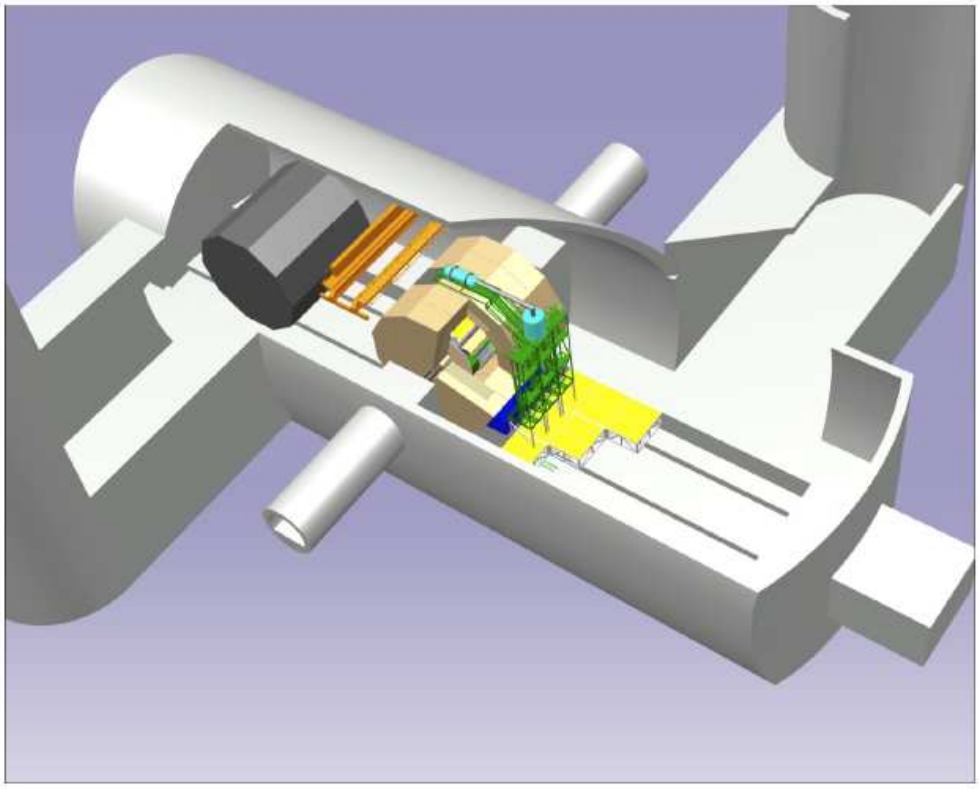


Fig. 2.4: Possible design of the foreseen detector hall for the ILC.

The design of the vertex detector is very sensitive to the environmental constraints imposed by each set of beam parameters. The frequency of the bunch trains is 5 Hz. Each bunch trains lasts about 1 ms, which can be translated into a machine duty cycle of $\sim \frac{1}{200}$. The time structure of the beam may allow to switch off the sensors of the vertex detector during the ~ 200 ms long train interval, thus reducing the dissipated power. We will resume the power cycling operational mode of the vertex detector in the section 3.1.1. Each bunch train is constituted by 2625 bunches, with a time interval of 369.2 ns between them. The hit density per unit time on the vertex detector (therefore its occupancy, see sub–section 3.1.2) depends on these numbers.

In order to cope with potential operational difficulties, there exist a number of different beam parameter options, that are designed to confront possible problems or limitations that may arise during the function of the ILC. They can all provide the luminosity of $2 \times 10^{34} cm^{-2} s^{-1}$ at $\sqrt{s} = 500$ GeV. The design of the ILC allows access

to any set of these beam parameters without any modification.

Low N "Low N" means that a lower number of particles per bunch (N) will be delivered. The big bunch populations may lead to disruption effects at the IP, microwave instabilities in the damping rings, or wakefield effects at the linacs. The Low N option, having half the number of particles that the nominal has, therefore the aforementioned effects can be suppressed. The consecutive loss in luminosity can be compensated via a better focus on the IP. The low space charge is also expected to suppress the beamstrahlung effect.

Large Y Tuning difficulties may not allow us to reach the desired emittance. The "Large Y" option focus more tightly in the horizontal plane in order to restore the luminosity, and elongates the bunches to reduce the beamstrahlung effects. A problem that could arise is the higher disruption of the beams at the IP.

Low P In the low power beam parameters the beam current is reduced by a factor of $\sim 30\,\%$ and the beam power by a factor of 2. The reason for that is to comply with possible limitations arising from the beam current or the beam power [RDR a], but also for run–cost reduction reasons. The luminosity is restored via a better focus on the horizontal plane. The bunches cannot be elongated due to the focusing on the vertical plane. Thus a significant enhancement of the beamstrahlung is expected.

Parameter	Units	Nominal	Low N	Large Y	Low P
Repetition rate	Hz	5	5	5	5
Particles/bunch	$\times 10^{10}$	2	1	2	2
Bunches/pulse		2625	5120	2625	1320
Bunch interval	ns	369.2	189.2	369.2	480.0
Energy loss by beamstrahlung	%	2.4	1.7	2.7	5.5
Luminosity	$10^{34}/cm^2/s$	2	2	2	2

Tab. 2.1: Main beam and IP parameters for $\sqrt{s} = 500GeV$ for various ILC optics options. (from [RDR a])

2.1.3 2009 Strawman Baseline

The 2009 Strawman Baseline option (SB2009) is proposing some major modifications at the design of the ILC, aiming to reduce its cost [SB2009]. The goal is the highest possible cost reduction, without compromising significantly the physics potential of the

ILC experiments. A strong argument of the SB2009 proposal is that the subsequent cost reduction can act as a buffer for future unexpected expenses that often occur in large projects. The most important modifications comparing to the nominal design are:

- The civil engineering is among the most important cost–driving factors of the ILC project. A significant reduction of the cost could be achieved if the twin tunnel design could be abandoned. Then the main linacs and the RTML would be integrated in a single tunnel. The single tunnel solution poses the challenge to place the linacs together with the sources and the distribution system of the RF, inside the same tunnel. Studies have proven the feasibility of such a concept. This is expected to reduce by $\sim 26\,\text{km}$ the length of the tunnels.

- The reduction of the beam power by 50 %. The beam power is proportional to $N \times n_b$, where N is the number of the particles per bunch and n_b the number of bunches per time unit. So the beam power could be reduced to 50 % by cutting by half either the number of bunches or the number of particles per bunch. In order to minimise the decrease of the luminosity, which is proportional to $N^2 \times n_b$ (see equation 2.4), it is the number of bunches that should be decreased. However, a reduction of 50 % of the luminosity is inevitable. The reduction of the beam power will allow for the decrease of the damping rings radius down to 3.2 km, as well as the decrease of the number of klystrons and modulators by 50 %.

- The modification of the design of the central region will lead to further suppression of the required tunnelling, reducing therefore the civil engineering costs.

- The undulator based positron source will be moved from the 150 GeV point to the end of the electrons main linac. Thus the length of the low energy positron transfer line will be decreased by several km. Moreover, the changing of the positron source will isolate most of the radiation–hazard structures at the central region.

The impact on the physics potential of the ILC experiments should be evaluated very carefully. A very important aspect of the SB2009 configuration is the reduction of the luminosity. It can be restored for the $\sqrt{s} = 500\,\text{GeV}$, but not fully for the $\sqrt{s} = 250\,\text{GeV}$, see table 2.1. Moreover, the SB2009 beam parameters will result in a higher beam induced background. Finally, it should be studied whether the SB2009 design allows for possible upgrades of the ILC for higher luminosity or $\sqrt{s}$.

2.1.4 Other Beam Options

According to its baseline design, the ILC should be able to conduct photon–photon or electron–photon collisions. The photons will be produced from the scattering of the

electrons with high intense pulse lasers. The scattering will take place a few millimeters in front of the IP. A photon energy $\sim 80\,\%$ of the electron energy may be reached [Bechtel 06].

2.2 Beam Induced Background

2.2.1 Beam-Beam Interactions

The main source of background in the ILD arises from the interaction between the very intense ILC beams. This effect has not been observed in such an extent at the previous lepton colliders (SLC, LEP). It is a result of the very high space charge density of the ILC bunches. Very detailed studies can be found at [Vogel 08]. Here we will give a brief description of the beam interactions effect and then focus on the pair background creation which is severely affecting the vertex detector and consists a major challenge for its design.

When the two opposite bunches approach each other, they exert mutually an electromagnetic force. The trajectories of the electrons (or positrons) are bent towards the center of the opposite bunch. This is called the pinch effect and it is both beneficial and detrimental. The further focusing of the beams, thus the decrease of the beam spot area, results in a luminosity enhancement by a factor of ~ 2. On the other hand, the deflection of the particles leads to the emission of hard bremstrahlung photons which degrade the energy spectrum of the beams. This is a very important effect because the precise knowledge of the initial state is crucial for a number of ILD analyses, like the model independent Higgs boson mass and production cross section measurement [Ito 09]. The average energy loss per particle is $\sim 1GeV$. The beamstrahlung photons travel along with the spent beams towards the beam dumps, therefore they are not a major background source. On the other hand, the scattering between two of these photons can create electron–positron pairs that are the largest source of background in the vertex detector.

2.2.2 The Pair Background

The e^+e^- pairs are created from bremstrahlung photons through two categories of processes, the coherent and the incoherent pair creation. The coherent pair creation is an effect of the same nature of the usual pair production where a photon interacts with an atomic nucleus field. Here the role of the atomic nucleus field is played by the collective field of the bunch. The contribution of this process to the total number of e^+e^- pairs in the ILC environment is estimated to be negligible. The vast majority is originates from the incoherent pair creation. In this case, the e^+e^- pairs are generated via the scattering of two photons coming from the two approaching beams. The incoherent

pair creation can be subdivided in the three following processes, depending on the nature of the photons that participate in the scattering. If both photons are virtual then we have the Landau–Lifshitz process which contributes to $\sim \frac{1}{3}$ of the pairs. When one photon is real and the other is virtual, the process is called the Bethe–Heitler. Its contribution is $\sim \frac{2}{3}$. In the Breit–Wheeler process, both participating photons are real. Its contribution is estimated to be at the per cent level. Figure 2.5 shows the Feynman diagrams of the aforementioned processes.

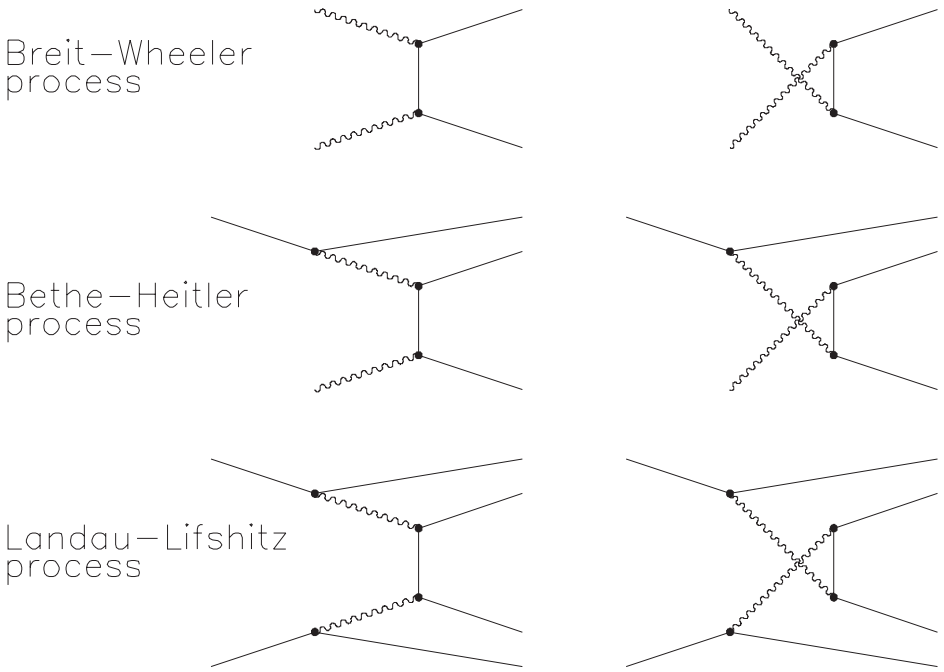


FIG. 2.5: Incoherent e^+e^- pair creation. From [Schulte 96]

The e^+e^- pairs have in principle relatively low P_T and are emitted in the very forward direction. The track density is illustrated in figure 2.6 as a function of the radius, the z coordinate and the strength of the magnetic field. This distribution has a conical shape whose radius is confined by the magnetic field. The shape of this cone dictates the design of the beam pipe and the vertex detector which both should be sufficiently far from the dense core of the track densities distribution. Thus it is used to define the minimal inner layer radius as well as the length in the z direction (along beam axis) of the ladders of this layer. However there will be still a small fraction of particles with higher P_T and/or polar angle that will directly reach and create hits in the vertex detector.

A second very important contribution arises from secondary backscattered parti-

cles. Even though the vast majority of the beam particles, after the collision, and the beamstrahlung are directed towards the beam dumps, a small fraction of them will be lost. This is happening because the two beams collide with a crossing angle of 14 mrad. These lost photons and electron/positrons may hit the calorimeters which instrument the very forward region of the detector, especially the BeamCal (see sub–section 2.3.4), and generate secondary particles. A fraction of these particles can be backscattered and reach the vertex detector, increasing the hit density mainly at small values of the azimuthal angle ϕ and, in a lesser extent, for large z. The latter contribution can be suppressed effectively with the application of the anti–DID (Detector Integrated Dipole) field ([Seryi 06]). The anti–DID is a magnetic field whose goal is to compensate for the spin misalignment of the beam particles, caused by the main field of the ILD, but also to guide the majority of the pair background particles towards the beam extraction holes (see also section 2.3.5). Concluding, the pair background* actually dominates the hit density at the vertex detector (VXD). At the innermost layer ~ 5 hits per cm^2 per bunch crossing are expected (meaning per 370 ns).

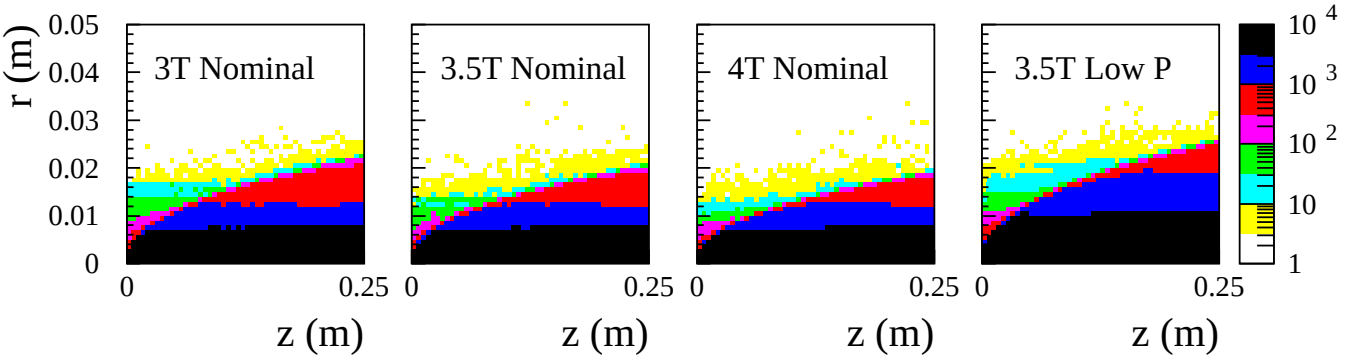


Fig. 2.6: Beamstrahlung e^+e^- pair particles track density as a function of the radius (vertical axis) and the distance in z (horizontal axis) for several magnetic fields and beam options. The magnetic field value and set of beam parameter are written on each plot. The colour scale at the right part of the figure, decodes the track density. From [LoI 10]

The Strasbourg group has performed detailed simulation studies in order to estimate the hit densities at the ILD vertex detector (VXD) caused by the e^+e^- pairs [De Masi 09]. The beamstrahlung was generated with the Guinea Pig software package [GuineaPig]. The produced files passed through the full simulation of the detector. Only the incoherent pair production has been taken into account. The figures 2.7

*From now on the terms pair background, beamstrahlung background and beam induced background refer to the beam background due to the incoherent pair creation.

and 2.8 illustrate the number of hits/mm^2/BX for each layer as a function of the z coordinate and the ϕ (azimuth) angle respectively, for the single layers (VXD03 - see section 2.3.2) model. For the layers 2-5, the hits densities have been multiplied by a factor of 5 for better visibility. The anti–DID field was not included in this study. In figure 2.8, the hit density in the region $|\phi| \leq 50^0$ is substantially higher. This effect is caused by the backscattered particles, since they are not produced symmetrically around the z axis. The reason for that is that the configuration of the very forward region, where those particles are created, it is not symmetric around the z axis due to the beams crossing angle. The peaks appearing for several values of the ϕ angle is an artifact; at these ϕ angles the two neighbouring ladders overlap (see section 2.3.2), thus a particle crossing this area will create two hits, one on each ladder.

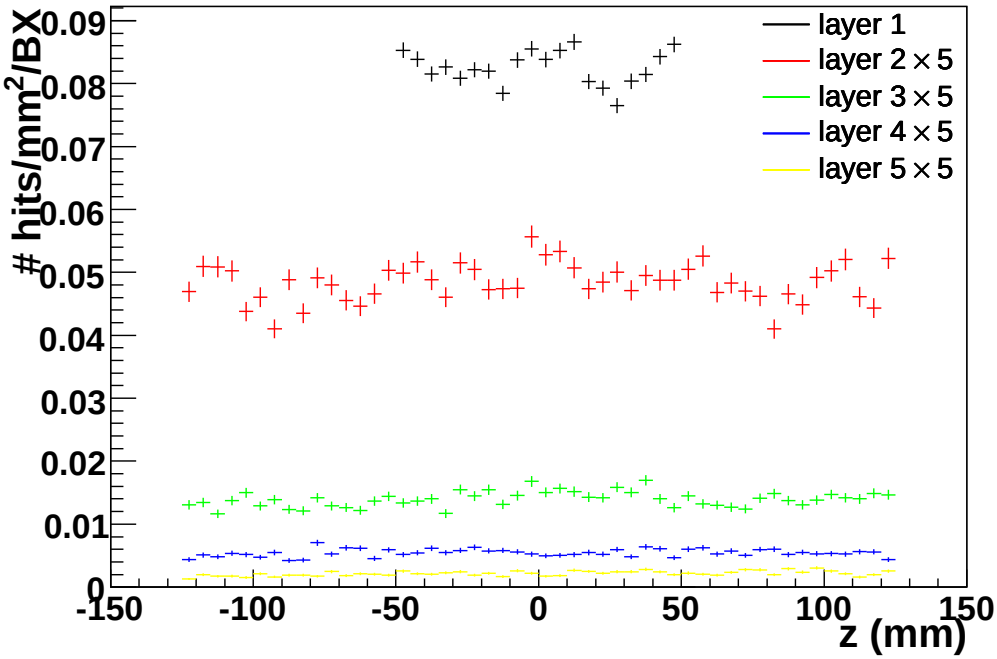


FIG. 2.7: Number of hits per mm^2 per BX for all 5 layers as function of the distance to the IP on the z-axis, assuming the nominal beam parameters. This number has been multiplied by a factor of 5 for layers 2-5. From [De Masi 09]

The anti–DID can effectively suppress the contribution arising from those backscattered particles. This is shown in the figures 2.10 and 2.9. They illustrate the hit densities arising from the pair background particles hitting the innermost layer of the VXD directly or after being backscattered, without and with the presence of the anti–DID field correspondingly. The hits labelled "large hit time" are due to the backscattered particles. The label "large hit time" is justified by the fact that the backscattered

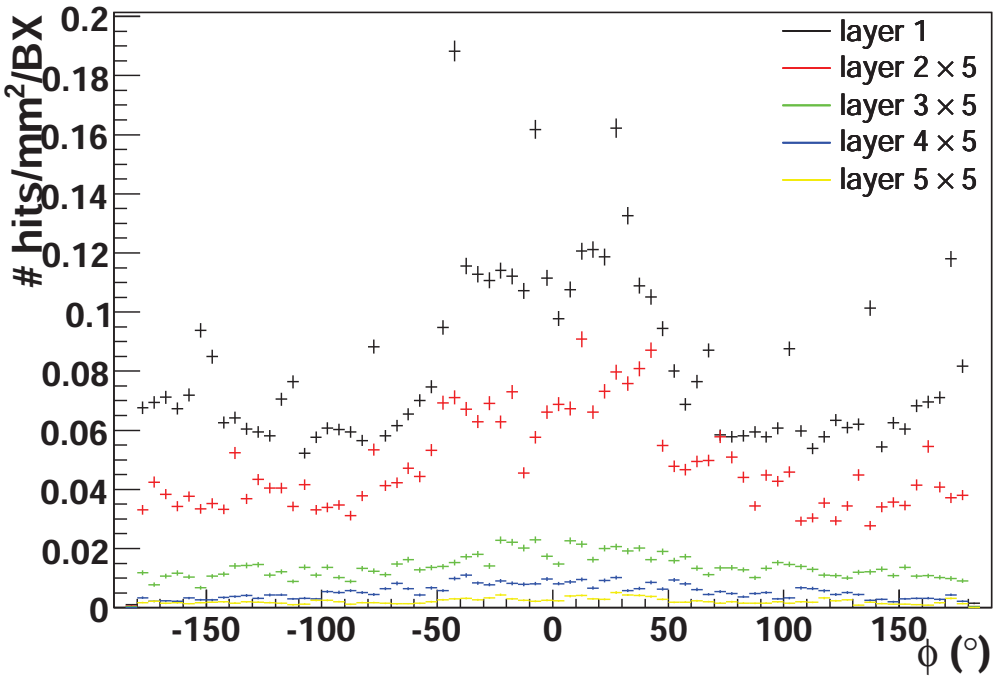


Fig. 2.8: Number of hits per mm^2 per BX for all 5 layers as function of the ϕ (azimuth) angle, assuming the nominal beam parameters. This number has been multiplied by a factor of 5 for layers 2-5. From [De Masi 09]

particles need to travel a distance of $\sim 4\,\mathrm{m}$ from their origin in order to create a hit at the vertex detector. The anti–DID field decreases their density by a factor of ~ 3, depending on the z and ϕ coordinates, which of course is directly reflected into the overall hit density. Moreover, it eliminates the peaks for large z, at the ends of the ladders, resulting in a much smoother hit density distribution. Thus its implementation is considered as highly beneficial for the background levels at the VXD. The table 2.2 summarises the hit densities for each layer for both geometries after implementing the anti–DID field.

2.2.3 The CLIC

The $\sqrt{s}$ that the ILC can reach is determined by the operational field gradient for $31.5\,\mathrm{MeV/m}$ delivered by its superconducting accelerating cavities. It has been demonstrated that the latter cannot exceed a gradient of $50\,\mathrm{MeV/m}$ (recently, superconducting accelerating cavities featuring a gradient that slightly exceeding $50\,\mathrm{MeV/m}$ have been constructed), due to thermal breakdown effects. Hence, if one wants to build a multi–TeV electron–positron collider, without making it extremely long, a novel lepton accelerating method has to be used. The CERN pursues the CLIC study [CLIC],

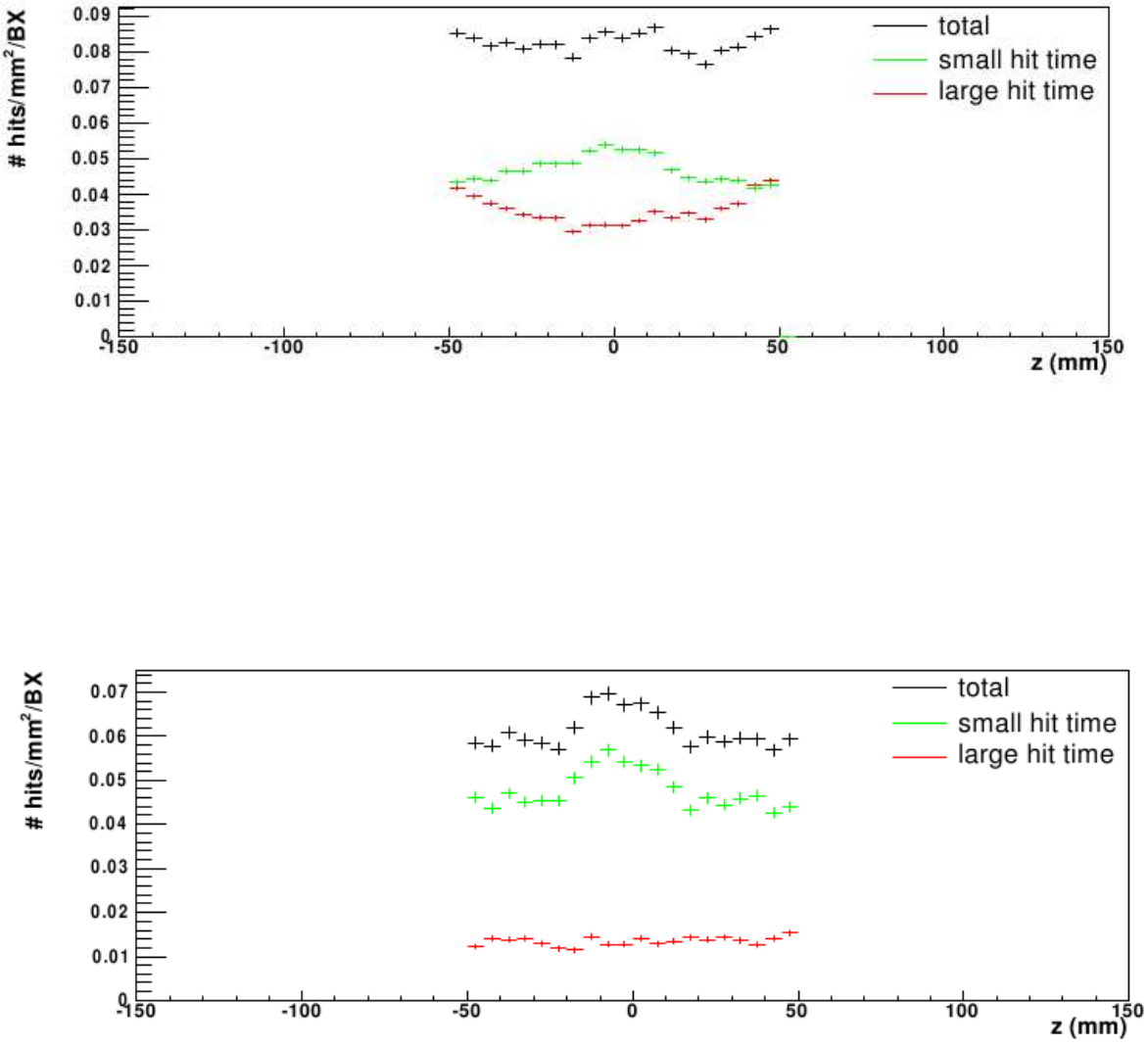


FIG. 2.9: Hit densities due to beamstrahlung in the innermost layer of the VXD as a function of the z coordinate. Top: without the anti–DID field. Bottom: with the anti–DID. The black crosses show the overall hit density, the green (small hit time) the one originating from particles impinging directly the VXD ladders, while the red ones (large hit time) originates from the backscattered particles. From [De Masi 09]

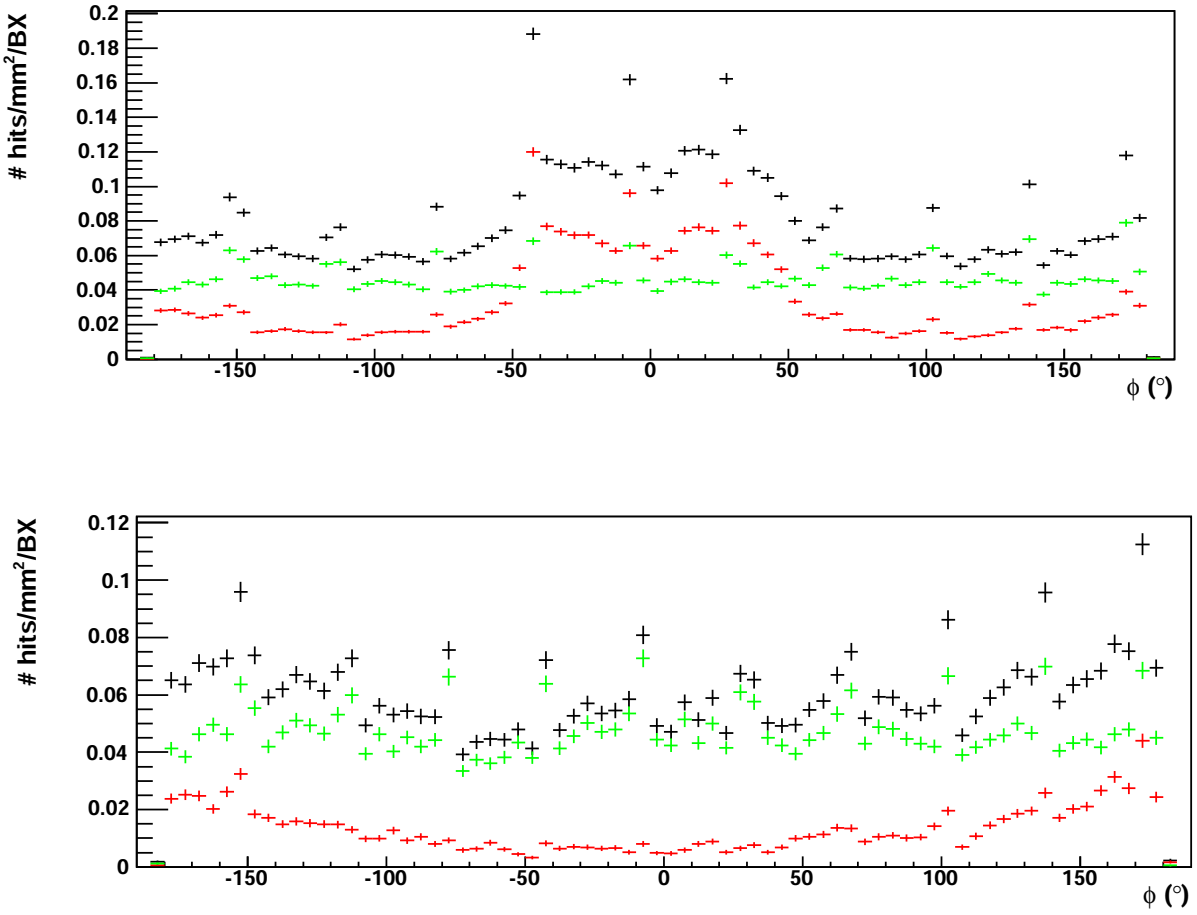


Fig. 2.10: Hit densities due to beamstrahlung in the innermost layer of the VXD as a function of the azimuthal ϕ angle. Top: without the anti–DID field. Bottom: with the anti–DID. The black crosses show the overall hit density, the green (small hit time) the one originating from particles impinging directly the VXD ladders, while the red ones (large hit time) originates from the backscattered particles. From [De Masi 09]

	VXD05		VXD03	
layer	$Hits/cm^2 \times BXs$	$r.o(\mu s) - (BXs)$	$Hits/cm^2 \times BXs$	$r.o(\mu s) - (BXs)$
1	4.4 ± 0.5	25(68)	5.3 ± 0.5	25(68)
2	2.9 ± 0.4	25(68)	$6.0 \pm 0.5 \times 10^{-1}$	50(136)
3	$1.54 \pm 0.14 \times 10^{-1}$	100(272)	$1.9 \pm 0.13 \times 10^{-1}$	100(272)
4	$1.34 \pm 0.11 \times 10^{-1}$	100(272)	$6.9 \pm 0.6 \times 10^{-2}$	100(272)
5	$3.2 \pm 0.7 \times 10^{-2}$	100(272)	$3.1 \pm 0.4 \times 10^{-2}$	100(272)
6	$2.7 \pm 0.5 \times 10^{-2}$	100(272)		

TAB. 2.2: Hits densities per cm^2 per BX, readout time and subsequent number of BXs superimposed per layer for the two main geometries, expected for the nominal beam options . The anti–DID field is included [De Masi 09].

which aims for an electron–positron collider that features an $\sqrt{s}$ of 0.5 to $\geq 3\,TeV$. The goal for the accelerating gradient is 100 MeV/m. In order to achieve this goal, the CLIC will use an innovative two beam acceleration scheme. The necessary RF power will be transferred to the main beam from the deceleration of a low energy - high current electron beam, called the drive beam. The proof of principle for the CLIC technology is still under way. Thus, the CLIC timescale goes well beyond the one of the ILC.

2.3 The ILD

Since the ILC running conditions are much less demanding than the LHC ones, one may consider that since detectors adapted to the LHC environment being at hand, the $R\&D$ for the ILC detectors does not pose any important challenges. This is a total misconception. The worth of the ILC physics program lays in the ability to perform high precision measurements. This pushes the detectors to target for unprecedented momentum, jet energy and impact parameter resolutions. These requirements make the $R\&D$ for the sensor technologies, as well as the reconstruction algorithms and the software, to be very challenging tasks. Additionally, even though it is true that most of the subdetectors are effectively almost background free, the beam induced background sets rather aggressive constraints to the vertex detector and to the forward trackers and calorimeters. Finally the design and integration of the ILC detectors should be compatible with the push–pull operational scheme, meaning they have to be easily movable and able to be aligned and calibrated accurately in a relatively short time.

Currently two general purpose detectors are being developed, the Silicon Detector (SiD) [SiD 09] and the International Large Detector (ILD) [LoI 10], which were subject of a Letter Of Intent (LoI) submitted to the relevant evaluation bodies [IDAG] in

2008/2009. The ILD was the product of the merge of two previous detector concepts, the Large Detector Concept (LDC) and the Global Large Detector (GLD), that share the same philosophy. The studies discussed in this thesis have been conducted inside the ILD framework, thus from now on we are going to focus on the ILD. The figure 2.11 presents an artistic view of the ILD. The time at which this thesis was written, the ILD concept was still rapidly evolving. Which means that the description of the detector and its reconstruction algorithms that will follow, will not be valid in all details by the time this document will be completed. So the goal of the following subsections is to introduce the structure and the general philosophy of the detector, but also to outline the framework in which the physics analysis object of this thesis were performed. This way the reader can evaluate whether the presented studies are up to date, or not. The following detector and algorithms descriptions are consistent with the ILD Letter of Intent.

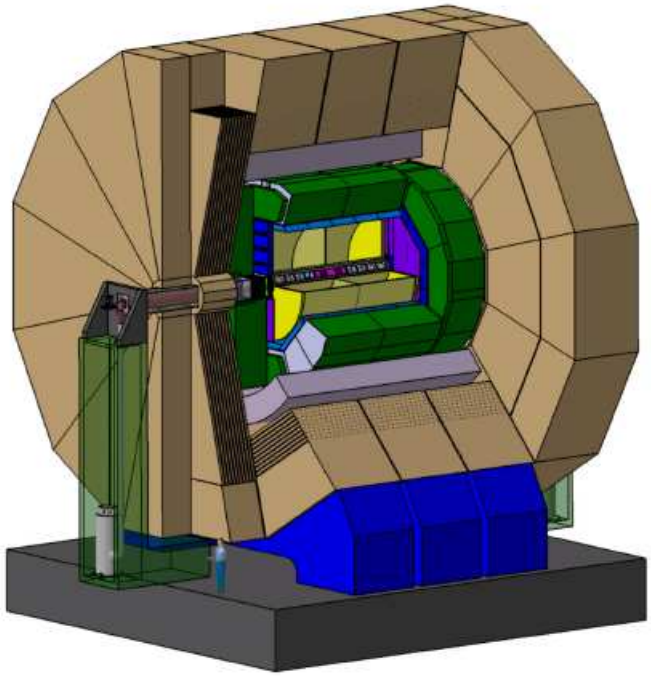


Fig. 2.11: An artistic view of the ILD.

2.3.1 ILD General Overview

The ILD follows the basic design scheme of the modern HEP detectors. Its design parameters are merely driven by the Particle Flow approach. A transverse sketch of one quadrant of the ILD is shown in figure 2.12. Going outwards from the IP, we will meet the following subdetectors: the Vertex Detector, located just outside the beam pipe; it is complemented by a silicon tracking system which is constituted by the Silicon Inner Tracker in the barrel and the Forward Tracking Disks at the endcaps. The main tracker is a Time Projection Chamber. A second auxiliary silicon tracking system is located between the main tracker and the calorimeters. It is constituted by the Silicon External Tracker in the barrel and the Endcap Tracking Detector at the endcaps. The Vertex Detector, the Time Projection Chamber and the complementary silicon detectors consist the ILD tracking system. Within this system, the Time Projection Chamber plays the central role for the track reconstruction, while the Vertex detector is mainly optimised for the reconstruction of displaced vertices. The calorimeters, both the electromagnetic and the hadronic, are located inside the coil, in order to optimise the jet energy resolution. The strength of the magnetic field is 3.5 T. The interplay between the strength of the magnetic field and the radius of the detector dictates its overall design. The specific values for the ILD were chosen after a cost–performance optimisation procedure [LoI 10]. An iron yoke surrounds the coil, being instrumented with muon detectors. The forward region is instrumented with the BeamCal, the LumiCal and the LHCAL. In the following sections, we will briefly describe the ILD subdetectors, giving of course emphasis to the Vertex Detector which is the main subject of this thesis.

We are going to use the following systems of coordinates: the cartesian, where x and y are the horizontal and vertical coordinates respectively, in the plane transverse to the beam line, while z is the coordinate along the beam line; the spherical, where r is the distance from the beam line, θ is the polar angle and ϕ the azimuthal one; and finally the cylindrical, where r is the distance from the beam line, ϕ the azimuthal angle and z is the coordinate along the beam line.

2.3.2 The Vertex Detector

The ILC physics program requires a Vertex Detector (called VXD hereafter) that exhibits excellent performance in terms of flavour tagging, meaning b- and c-jet identification, and track reconstruction. Flavour tagging is essential for numerous studies, the most prominent being the precise measurement of the Higgs boson couplings to the fermions. The key for the flavour tagging is the efficient reconstruction of the displaced vertices. To do so, one has to provide a detector that exhibits high spatial resolution, while featuring very low material budget, and located as close as possible to the IP. The minimum distance is dictated by the beamstrahlung induced background, as discussed

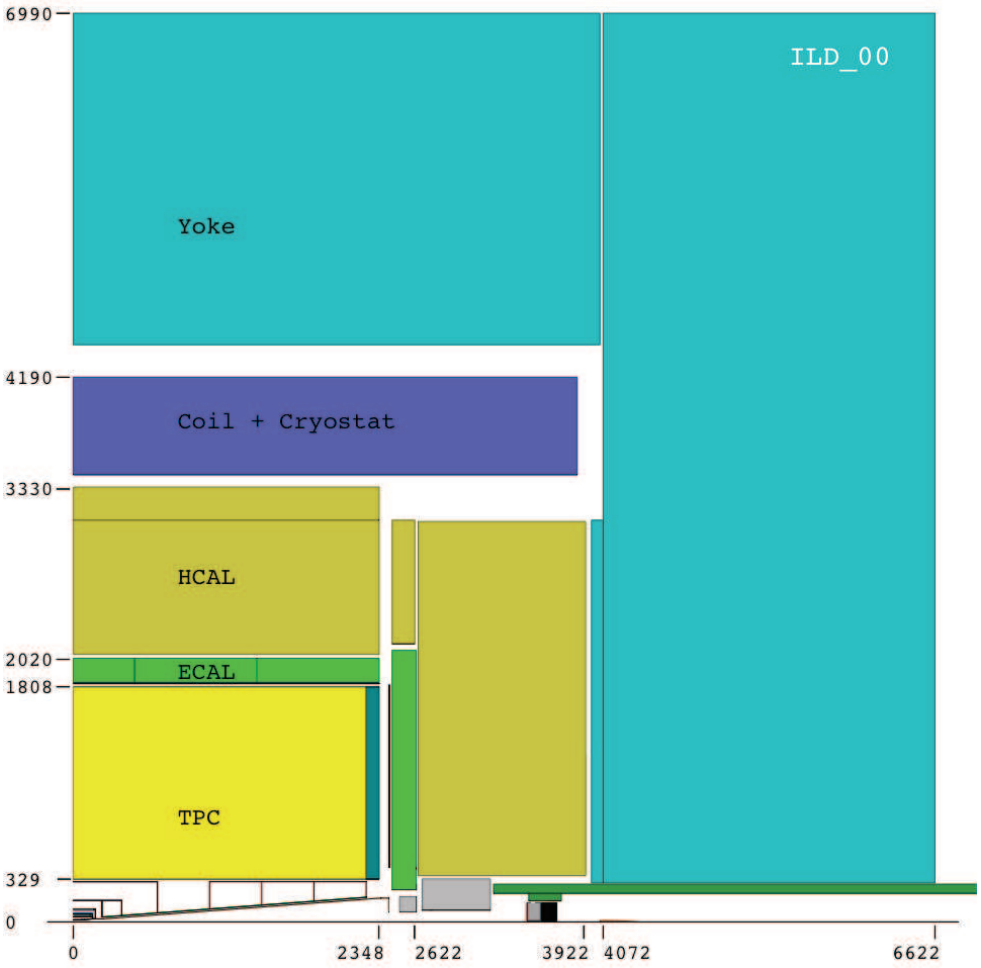


Fig. 2.12: Transverse view of one quadrant of the ILD ([LoI 10]).

in the section 2.2.2. The VXD, together with the forward detectors, are the subsystems of the ILC mostly affected by the beamstrahlung background. The imposed running constraints, as well as the trade offs between the various VXD requirements, will be explained in the third chapter.

The VXD also plays a central role in the track reconstruction. Providing 5 or 6 very high precision spatial points for each track, it offers a standalone track reconstruction ability. This is crucial for the track reconstruction of low momentum particles that do not reach the main tracker. Also, offering a polar angle coverage down to $\sim 13.5^0$, combined with the Forward Tracking Detector, it may help in the reconstruction of the shallow angle tracks. The ability to reconstruct low momentum and/or very forward tracks could be proven indispensable for the studies that require vertex charge

identification. Vertex charge has a binary behaviour, meaning that a single missed or misreconstructed track is sufficient to alter completely the measurement. An example of a study that requires vertex charge identification is the Top quark forward–backward asymmetry ([Ikematsu 09]).

The VXD is composed of ladders, arranged cylindrically in concentric layers to form barrels. The radius of the innermost layer, as well as the length of its ladders is limited by the conically shaped dense core of the pair background tracks, see figure 2.6. So the standard LoI design of the VXD is composed of long barrels. The beam background depends on the beam parameters. Therefore the minimum distance from the IP will be defined from the selected beam parameters.

There are two main VXD candidate geometries, one with 5 single layers (VXD03) and one with 3 layers (VXD05), as illustrated by figure 2.13. The VXD03 geometry features single ladders, which means one layer of silicon pixel sensors on a light mechanical support, that has to provide all the necessary services. The VXD05 geometry features double sided ladders, meaning that silicon pixel sensors will equip both sides of the support layer, with a 2 mm distance between them, and providing this way two very high precision measurements per layer. The main features for each geometry are summarised in table 2.3. For both designs, the material budget has to be limited to a few per mill of radiation length (X_0). The material budget stated in the LoI is 0.11 % and 0.16 % of X_0, for single and double sided ladders respectively. The high single point resolution and very light material budget required call for a pixelated sensor technology different from the one used previously, in particular at the LHC. A world wide effort was initiated to develop new pixel technologies for this purpose. Various are under development in perspective of an optimal at the time of the detector construction. The location of the support electronics is technology dependent. It could be put either at the ends, or at the sides of the ladder. An overlapping area of $\sim$ 500 mm between neighbour ladders has been foreseen, in order to allow for accurate ladder to ladder alignment.

	radius(mm)		ladder length (mm)		readout time (μs)		N_o of ladders	
geometry	VXD03	VXD05	VXD03	VXD05	VXD03	VXD05	VXD03	VXD05
layer1	15	16/18	125	125	25-50	25-50	10	10
layer2	26	37/39	250	250	50-100	100-200	11	11
layer3	37	58/60	250	250	100-200	100-200	11	17
layer4	48		250		100-200		14	
layer5	60		250		100-200		17	

TAB. 2.3: Geometrical and technological parameters of the two main vertex detector options.

The VXD05 geometry, as it is described inside the GEANT4 [Geant4] simulation

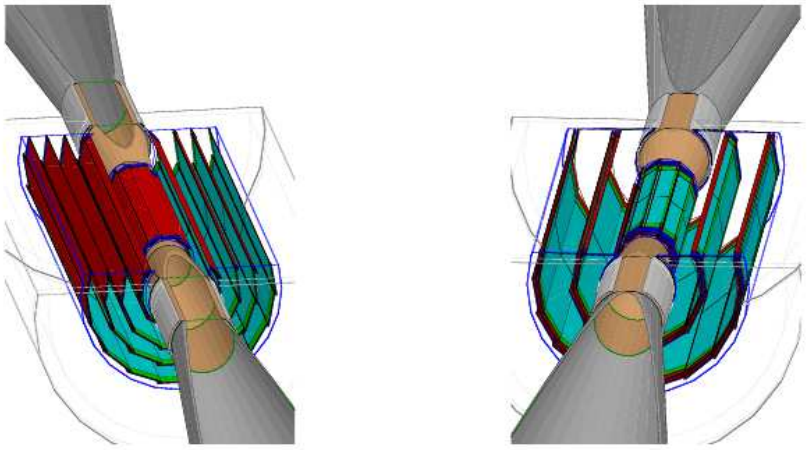


Fig. 2.13: VXD geometries: 5 single layers (VXD03, left) and 3 double layers (VXD05, right).

for the LoI studies, is illustrated in figure 2.14. The innermost layer is mounted on the beam pipe, while the outer layers are supported by a 500 μm thick, 65 mm radius cylindrical beryllium support. The whole structure is enclosed inside a 10 mm thick (0.05 %X_0) Styrofoam cryostat with an inner radius of 90 mm. As a shield against the Electromagnetic Interference (EMI) coming from the beam (see section 3.1.2), a Faraday cage made of a 0.5 mm thick aluminium foil (0.56 %X_0) surrounds the cryostat.

The simulation of cabling and services in general, inside and outside the VXD, is quite coarse, merely because these aspects depend strongly on the sensor technology. Concerning the LoI studies, the cabling inside the cryostat was simulated with flat kapton cables of 0.094 mm thickness. The services that will transfer the power, the data and the cooling from and to the outer world were missing. Recently, going towards the Detector Baseline Document (DBD), efforts have been made to estimate some generic values for the amount of the cabling needed. One would expect one cable per ladder. Hence, depending on the chosen geometry, the number of cables for power and data will typically amount from 60 to 80. Inside the simulation framework, they will be described as flat, 50 μm thick kapton cables with 0.018 % radiation length. This value may rise to $0.02 - -0.03$ % when the metal traces will be included. Further aspects which may severely affect the cables material budget cannot be examined at the current level of the VXD studies. One example is whether or not coaxial cables with metal screening will be required in order to avoid the EMI. Finally, the specifications for the cooling pipes are also not described yet, mainly because they are technology dependent.

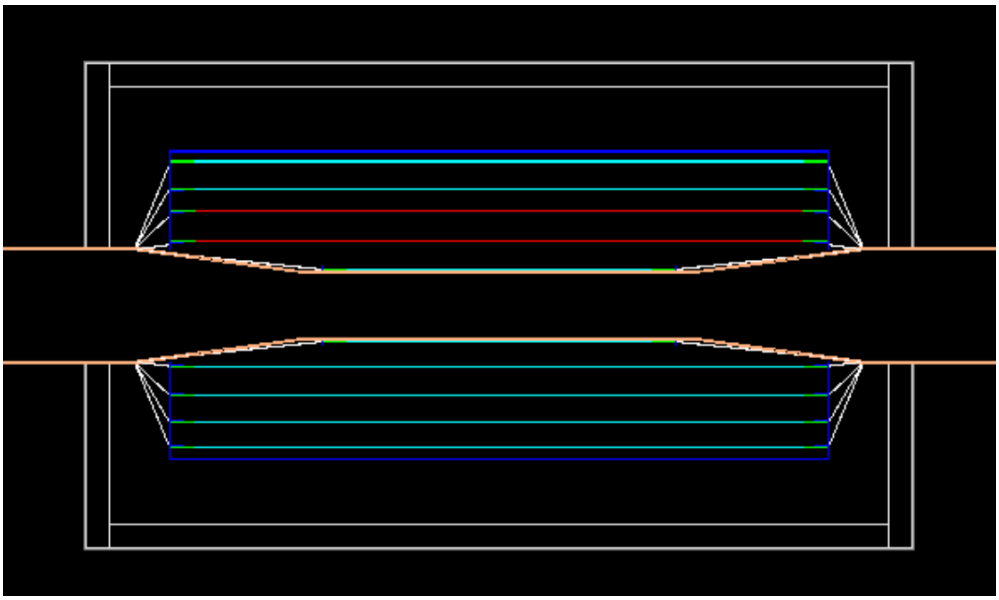


Fig. 2.14: Visualisation of a section of the VXD05 design as described inside the simulation framework (Mokka, see section 5.4) used for the LoI studies. We can see the three double sided ladders (the support is visualised with red, the sensitive with cyan and the supporting electronics with green), the Be support shell (blue), the kapton cables (white) and the cryostat (grey), while the beam pipe is visualised with orange colour.

A detailed description of the simulation status of the VXD that is going to be used for the DBD studies will be given in chapter 6.

2.3.3 The Tracking Detectors

At the ILD, the task of the track reconstruction is fulfilled by the combination of three subdetectors, that they feature standalone tracking abilities. The central tracker is a Time Projection Chamber (TPC), being complemented by the VXD, which is crucial especially for the reconstruction of the low momentum tracks, and the Forward Tracking Detector (FTD) that provide shallow angle coverage. Additional high precision spatial measurements for matching track segments reconstructed in the above mentioned subdetectors are provided by the Silicon Inner Tracker (SIT), the Silicon External Tracker (SET) and the Endcap Tracking Detector (ETD). The development of the tracking system was driven by the requirement to offer an overall momentum resolution for high momentum particles of $\sigma_{1/P_T} = 2 \times 10^{-5}\text{GeV/c}$.

The TPC

The TPC is a large volume gaseous detector that serves as main tracker of the ILD. The momentum resolution depends mostly on its performance. It is constituted of a central barrel part, with an inner radius of $\sim 33\,\mathrm{cm}$ and an outer radius of $\sim 180\,\mathrm{cm}$, and two endcap parts. Each TPC endcap has an area of $10\,\mathrm{m}^2$. It provides solid angle coverage up to $\cos\theta \sim 0.98$. The ILD physics program requires a TPC featuring a momentum resolution one order of magnitude better than the one achieved at the LEP experiments. This goal can be achieved with the use of micro–pattern gas detectors. The corresponding $R\&D$ issues are mainly investigated by the LCTPC collaboration [LCTPC].

Compare to silicon detectors, the TPC provides moderate spatial resolution, featuring $\leq 100\,\mu\mathrm{m}$ in $r\phi$ and $\sim 500\,\mu\mathrm{m}$ in z. On the other hand, it provides a huge number of spatial points in three dimensions (≥ 200), that compensate for the moderate resolution. The selection of a TPC as a main tracker offers several advantages to the overall detector concept. It offers a low material budget, being compliant with the particle flow analysis specifications. It can cope with the high magnetic field required in order to achieve high momentum resolution. Additionally, by measuring the energy loss of the particles, it can provide valuable particle identification information.

Silicon Tracker

The track reconstruction provided by the TPC and the VXD will be substantially improved with the aid of an auxiliary silicon tracking system developed by the SiLC collaboration [SILC]. It is constituted by four subsystems, the SIT and SET in the barrel region, and the FTD and ETD on the forward region. We will give a brief description of each subsystem in the following list.

SIT The SIT is constituted by two cylindrical false double layers equipped with silicon strip sensors, located in the space between the VXD and the TPC. They are called false double layers because the two silicon strip sensors are glued together, in a way that the strips on the two layers form a stereo angle between them. Following this approach, the SIT can provide two three–dimensional measurement points. The featured spatial resolution is $\sim 7\,\mu\mathrm{m}$ in $r\phi$ and $\sim 50\,\mu\mathrm{m}$ in the z coordinate. A material budget of $0.65\,\%\ X_0$ per layer is within reach, while the final goal is $0.5\,\%\ X_0$ per layer. Its main purpose is to bridge the gap between the TPC and the VXD, and aid in the linking of the two track segments reconstructed from the two aforementioned subdetectors. It will be also very helpful for the reconstruction of the low momentum tracks. Finally, due to the high time resolution of the strips, it can provide single bunch crossing timestamping.

SET A particle going from the TPC to the Electromagnetic CALorimeter (ECAL),

has to cross the TPC external wall that features a radiation length of 3 %. The multiple scattering effects will introduce some uncertainties on the track propagation. The SET is formed by a cylindrical layer of false double sided silicon strip sensors. It is located in the space between these two subdetectors, and will give a spatial point measurement in order to compensate for these uncertainties, benefiting the track–cluster matching (see section 2.4.4). Additionally, it can also provide single bunch crossing timestamping.

FTD The FTD is constituted by 7 disks of silicon sensors, complementing the TPC by covering the very forward region down to 0.15 rad. The first two or three disks will be equipped with silicon pixel sensors and the rest with silicon strips. It features its own standalone tracking capabilities. Given the fact that it is positioned in the very forward region, it is expected that will severely suffer from the beam induced background.

ETD The ETD is constituted by a cylindrical layer of false double sided silicon strip sensors. The motivations for the ETD are similar to the ones of the SET. It will provide a link between the TPC endcap and the calorimeter endcap. It will help to reduce the effect of the TPC endcap material budget, that it is expected to reach 15 %. The ETD measurements will help to significantly extent the lever arm of the track reconstruction, up to 2430 mm in the z coordinate.

2.3.4 Calorimeters

The physics program of the ILD requires typically a jet energy resolution of $\Delta E/E \leq 3.8\,\%$ for jet energies $\leq 100\,\mathrm{GeV}$, as it is stated at the [LoI 10]. This target can be expressed equivalently as $\Delta E/E = 30\,\%/\sqrt{E/\mathrm{GeV}}$. Such a fine resolution is out of reach of the conventional calorimetric systems. The ILD plans to achieve it with the help of the Particle Flow Analysis (PFA), see section 2.4.4. Thus the design of the calorimeters is merely driven by the PFA requirements, which means that they are optimised for high transverse and longitudinal segmentation. The energy resolution of the calorimeter itself remains of course an important consideration, nevertheless it is a secondary priority compared to the segmentation. This is because the desired energy resolution is going to be achieved through a combination of information from the tracking system and the calorimeters. There is a number of alternative technological options for the calorimeters. The choice will be mostly based on their pattern recognition performance. The $R\,\&\,D$ takes place inside the framework of the CALICE collaboration [CALICE]. Both hadronic and electromagnetic calorimeters will be positioned inside the superconducting solenoid in order to minimise the material located in front of them. This specification calls for a compact design, driven by cost concerns. They will be sampling calorimeters with layers of absorber material interleaved with

sensing layers. They will consist a barrel and two endcap parts, following the general ILD structure.

Electromagnetic Calorimeter The ECAL will be a sampling calorimeter using tungsten as an absorber. There are three options for the sensitive material. In one option, called the SiW ECAL, the sensitive material will be square silicon pin–diodes featuring a pitch of 5 mm. The second option, the ScECAL, foresees scintillator strips of $1 \times 4\,\mathrm{cm}^2$. In order to achieve the desired segmentation, they will be oriented orthogonally with respect to the strips of the neighbouring layers. This way an effective granularity of $1 \times 1\,\mathrm{cm}^2$ can be achieved. The third option is a variation of the SiW option, where the sensing and the readout are realised with pixelated CMOS sensors (see section 3.2). Featuring a pixel pitch of $O(50\,\mu\mathrm{m})$, a very high granularity can be achieved even with a digital readout. The use of standard industrial CMOS processes could lead to a reduced cost compared to the analogue silicon pin–diodes.

Hadronic Calorimeter The Hadronic CALorimeter (HCAL) will be a sampling calorimeter, using stainless steel as an absorber. Two different options for the sensitive material exist. The first option is scintillator tiles (Analogue HCAL), being read out by silicon photomultipliers. In the second option (Digital HCAL), the active layers will be equipped with gaseous devices, being read out in semi–digital mode (2 bits). A strong candidate is the Glass Resistive Plate Chamber (GRPC), that can provide very fine segmentation combined with high detection efficiency while maintain the cost at low levels.

Forward Calorimeters The very forward region of the ILD, very close to the beam pipe, will be equipped with luminosity–monitoring calorimeters. These are called LumiCal and BeamCal. The scope of the LumiCal is to provide precise measurements of the luminosity by measuring the Bhabha scattering process $e^+e^- \to e^+e^-(\gamma)$. Additionally, it will extend the solid angle coverage of the ECAL down to 10 mrad. The BeamCal will be used for a fast on–line estimation of the luminosity. The estimation will be based on the energy deposit due to the beamstrahlung pair background particles. The information of the BeamCal can be fed back to the BDS (see section 2.1.1). Then the BDS can apply the necessary beam optics corrections to maximise as much as possible the luminosity. The radiation tolerance requirements here are quite aggressive. The BeamCal has to withstand a dose of several MGy per year. Another reason to instrument the very forward region is the calorimeter hermeticity. The LHCal is expected to extend the solid angle coverage of the HCAL.

2.3.5 The Magnetic Field and the Yoke

The ILD is designed to operate in a nominal magnetic field of 3.5 T. Its main requirements are a high homogeneity inside the volume of the TPC and reduced fringing fields (the magnetic fields that are spread beyond the coil). They should be kept small in order not to magnetically perturbate the second detector, which is kept at a garage place at the IR. The main magnetic field is generated by a superconducting coil surrounding the whole tracking and calorimetric system. The size and geometry of the coil determines the overall structure of the ILD detector.

The coil is surrounded by an iron yoke, whose main functionality is to return the magnetic flux. It is constituted by a barrel and two endcap parts and its total weight amounts to 13400 t. It will feature a sandwich like structure, where iron plates will be interleaved with muon detectors. The muon detectors are also going to serve as tail catchers, improving the energy resolution for high energy jets not fully contained in the calorimeters.

A secondary magnetic field in the ILD detector is the anti–Detector Integrated Dipole (anti–DID) field. At the ILC is foreseen that the beam collision will happen with a crossing angle, most likely 14 mrad. In such a case, the main field will introduce a small deviation ($O(100\,\mu\text{m})$) on the vertical trajectory of the particles at the IP. This angle will cause a rotation of the particle's spin, resulting in a spin misalignment. Initially, the proposed solution to compensate for this effect was to superimpose a so–called DID field, that would align the field lines with the incoming beams. However, the DID field would have a large impact on the effect of the beamstrahlung on the detector. It is expected to increase the number of pair background particles that will backscatter on the BeamCal and hit the VXD. On the other hand, if the polarity of the DID is reversed (anti–DID), the field can guide the majority of the beamstrahlung particles to the beam extraction holes, thus minimising the hits due to backscattered beamstrahlung particles on the detector. The studies exposed in [Seryi 06], demonstrate that both the effects of the spin misalignment and the minimisation of the backscattered beamstrahlung particles can be realised with the anti–DID field.

2.4 Reconstruction

The physics reach of an experiment depends strongly on both the detector performances and the reconstruction algorithms used. In order to better understand the results of the detector optimisation, we should evaluate the detector inside the framework of the used tools, and understand the maturity of those tools. In the following sections, a brief description of the reconstruction algorithms used by the ILD collaboration will be given. We are going to emphasise those which strongly affect or involve the VXD.

2.4.1 Digitisation

The hits digitisation procedure, inside a simulation study, is the reproduction of the real detector's response to a particle crossing its sensitive volume. Different kind of detectors, like calorimeters or a TPC, but as well as the same kind of detectors that use different technologies, require their own specific digitiser. Concerning the ILD VXD, as has been repeatedly mentioned, there is not yet a specific pixel technology chosen. This fact imposes the utilisation of a very generic hit digitiser. This algorithm takes the hits as they are simulated by GEANT4, and reproduces the response of the sensor by applying a simple Gaussian smearing on the VXD ladders surface. The size of the smearing depends on the assumed sensors single point resolution.

2.4.2 Vertexing

The vertexing is the procedure to reconstruct the primary and the displaced vertices of a final state. The ILD uses a topological vertex finder, called ZVTOP [Jackson 97]. The jet flavour tagging is strongly based on the vertexing. The flavour tagging could be regarded as the most important task of the ILD VXD. Thus, the VXD performance on this aspect is crucial for its optimisation. For this reason, the parameters on which the vertex reconstruction and the flavour tagging are based, will be examined in chapter 5, where the VXD optimisation is discussed (see section 5.7).

2.4.3 Tracking

The tracking can be defined as the reconstruction of the charged particles trajectories. The track reconstruction procedure inside the ILD collaboration is still evolving. In this section, we are going to describe the track reconstruction algorithms that have been used for the purposes of the LoI studies. Two separate track segments are reconstructed, one at the silicon detectors and the other at the TPC. The two segments are then combined in a unique track. The TPC track reconstruction is based on the LEP tracking algorithms. It take into account the energy loss and the multiple scattering effects. A Kalman filter technique is used for the track fitting.

The silicon detectors tracking imposes certain requirements on the VXD performances. Thus we are going to briefly describe its pattern recognition algorithm. It starts by forming track seed candidates, in the VXD and SIT layers, which then follows throughout the detector. Inside a magnetic field, one needs three 3–dimensional measurements in order to initialise track reconstruction. Therefore, the track seeds are hit triplets that belong to different layers. In order to accelerate the hit search procedure, the VXD and the SIT are initially divided in stereo angle sectors. The seed formation is restricted each time inside a sector. The produced triplets are tested whether they satisfy the helix hypothesis. If they do, they are accepted as track seeds. Then the

algorithm searches the other layers of the same sector in order to attach additional hits to the track seed. The assignment is based on the distance between the position of the hit and the extrapolation of the track helix on this layer. A χ^2 criterion decides whether the hit should be kept. Finally all the reconstructed tracks are sorted according to their χ^2 over degrees of freedom value. The track with the smallest one is by definition accepted. Each hit will be assigned to only one track, meaning that those tracks candidates that share hits with an already accepted track, are rejected. The same procedure is applied to the forward tracking detectors. Finally, pairs are created, of all the possible combinations between the track segments reconstructed at the VXD–SIT and those from the FTD. The segments are examined whether they can be merged or not. The crucial merging criterion is the angle difference of the two track segments. In the case the criteria are fullfiled, a single track is created.

The approach of track pattern recognition used by the silicon tracking algorithm imposes some limitations to the evaluation of the VXD performances. The brute force seed triplet formation and the assignment of hits to a track based on predefined distance criteria, will introduce a huge combinatorial background in the case of high hit densities. This is the case of the ILD VXD, when beamstrahlung hits are considered. The maturity of the silicon tracking tool, does not allow it to cope well with the beam background. We will come back to this point when we will try to evaluate the VXD performance.

2.4.4 Particle Flow Analysis

As mentioned in the section 2.3.4, the goal for the jet energy resolution for the ILD reaches $\Delta E/E \leq 3.8\,\%$ (or equivalently $0.3\sqrt{E}$). This value is imposed by the requirement for a jet energy resolution comparable to the W and Z bosons width. Such a fine resolution is not achievable with conventional calorimetry. In order to achieve its goals, the ILD will follow the Particle Flow Analysis (PFA) paradigm [Thomson 09]. The PFA, rather than being a calorimetric technique, defines the overall detector optimisation procedure. The principle of the PFA is to individually measure the energy of each particle inside a jet. The energy of the charged hadrons and the leptons will be measured by the tracking system, the energy of the photons with the ECAL and that of the neutral hadrons with the HCAL. The jet fragmentation studies at LEP demonstrated that $\sim 60\,\%$ of the jet energy is carried by charged particles (mainly hadrons), $\sim 30\,\%$ by photons and only $\sim 10\,\%$ by neutral hadrons. Thus the aforementioned approach for the energy measurement can compensate for the large statistical fluctuations of the hadronic showers. To measure the jet energy using the PFA, one has to assign each reconstructed track to the correct calorimeter cluster to avoid double counting or neglecting the particle's energy. For example, if a neutral particle cluster in the calorimeters cannot be resolved from a charge hadron shower associated with a track, the energy of the neutral particle will be disregarded. This "confusion" term

dominates the jet energy resolution. Therefore, the granularity of the calorimeters is more important than the energy resolution. This makes the calorimeter design and optimisation for the ILC experiments to differ from the traditional calorimetry.

As already said, the design of the ILD, as well as its cost, are merely driven by the PFA requirements. Both calorimeters should be located inside the solenoid. So the compactness is another specification defining the calorimeters design. Large detector radius favours the PFA, since the physical separation of the hits in the calorimeters will be larger. It is also favoured from a strong magnetic field. The overall dimensions and magnetic field of the ILD, taking of course cost into account, have been optimised through detailed simulation studies. Other important factors are the tracking efficiency, both in barrel and in the forward region, the detector aspect ratio (mainly for the energy resolution of the forward jets, $0.80 \leq |\cos\theta| \leq 0.95$) and the VXD material budget. Within the ILD collaboration, the Pandora PFA algorithm is used [Thomson 09].

2.5 SiD

Two multipurpose complementary detectors are foreseen for the ILC, the ILD and the Silicon Detector (SiD). This will provide an essential tool for the validation and cross–check of the obtained results. The same approach has been followed by the LHC with its two multipurpose and complementary detectors, ATLAS and CMS. Another advantage of this approach are the benefits that will arise from their competition.

Compared to the ILD, the most substantial difference lays on the main tracking system. The SiD features a pure silicon tracker. It is made of five layers of silicon microstrip sensors, featuring a material budget of $0.8\,\%X_0$. Concerning track reconstruction, the high spatial resolution compensates for the reduced hits redundancy compared to a TPC. Since the readout of the strips takes place during one bunch crossing, it offers robustness against the beam induced background. Regarding the overall design of the SiD, it is also driven by the PFA requirements. So apart of the muon chambers, the main detector components are located inside the solenoid, which provides a 5 T magnetic field. The SiD follows the strong magnetic field - small detector radius approach, mainly for cost–conscious reasons, since the cost is mainly driven by the volume of the calorimeters and the solenoid. The vertex detector is formed from five layers in the barrel region, equipped with silicon pixel sensors, and four inner and three outer pixel disks at the endcaps. The electromagnetic calorimeter uses tungsten as the absorber and silicon as the sensing material. The hadronic calorimeter has an iron–RPC structure. An iron yoke which surrounds the solenoid returns the magnetic flux. The yoke is equipped with muon chambers. The detailed description of the detector can be found at the Letter of Intent of the SiD [SiD 09]. Figure 2.15 shows the illustration of one quadrant of the SiD.

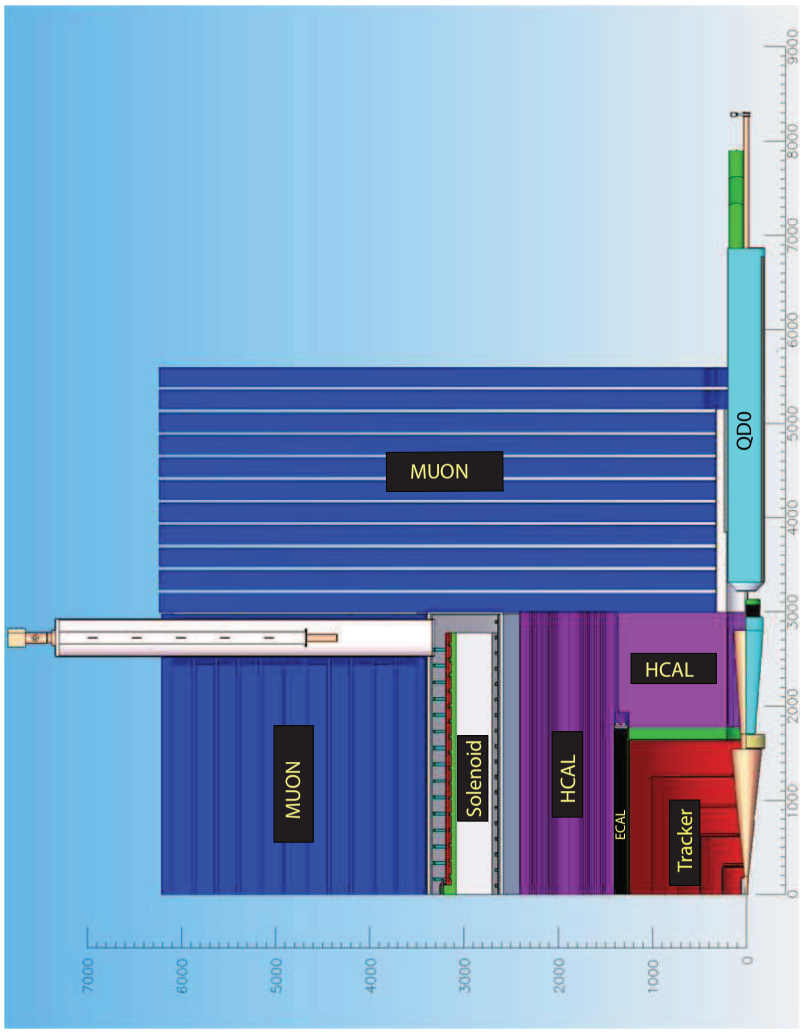


Fig. 2.15: Transverse view of one quadrant of the SiD ([SiD 09]).

2.6 Conclusion

Through this chapter, we have introduced the future electron–positron linear collider ILC, and one of its proposed detectors, the ILD. A brief description of the reconstruction tools was given as well. The main focus was on the VXD description, whose optimisation studies is the main topic of this thesis. The ILC offers a clean environment, ideal for high precision measurements, but is not completely background free: its most important source is the beam induced background. This effect constraint the VXD geometrical design and its sensors specifications. In the following chapter we will introduce the CMOS sensors technology, which is a promising candidate for the VXD sensors. We will describe a VXD based on CMOS sensors, and try to find its optimal design, that will satisfy its physics goals while be compliant with the constraints imposed by the beam background.

CHAPTER 3

A Proposed CMOS Based Vertex Detector

The physics motivations driving the design of the ILD vertex detector have been presented in the second chapter. Here we will discuss what these mean in terms of its performance. It is required that the vertex detector will be able to identify charm and beauty hadrons. The decay length of the charm hadrons is $O(100\,\mu\text{m})$, thus a precision of $O(10\,\mu\text{m})$ for the reconstruction of displaced vertices is required. The figure of merit for a vertex detector, that indicates its precision in the reconstruction of displaced vertices, is the impact parameter resolution. The impact parameter resolution (σ_{IP}) can be expressed by the following parametric equation:

$$\sigma_{IP} = \text{a} \oplus b/p \cdot \sin^{\frac{3}{2}}(\theta) \tag{3.1}$$

where p is the particle's momentum and θ the polar angle. In order to achieve the goals of the experiment, the required values for the parameter a is below $5\,\mu\text{m}$ and for b is $\leq 10\,\mu\text{m} \cdot GeV/c$. In this chapter we will present the goals on the vertex detector design and discuss the sensor technology required to achieve them. Of course, the experiment's running conditions should be considered. These are mainly arising from the beam-beam interaction and determine the tolerance on the ionising and non - ionising radiations as well as the readout time of the sensors.

The ILD requirements for an unprecedented impact parameter resolution (see table 3.1), which imposes very segmented ultra light sensors, together with the constraints from the ILC running conditions, that demand a moderate radiation tolerance and swift readout, call for an innovative silicon pixel sensor technology. A number of candidate technologies exist (they are summarised in section 3.1.3). In this chapter we will focus

on CMOS* Pixel Sensors (CPS). The latter are a novel technology in the High Energy Physics area, even though it is widely used by the optical industry. The principle of operation and the general performance of CMOS sensors will be presented. The architecture and the performance of the MIMOSA 26 sensor, the first full scale digital sensor of the MIMOSA† series, will be described. Finally, using MIMOSA 26 as a baseline sensor architecture and assuming the double layers VXD design (see sub–section 2.3.2), a CMOS based vertex detector will be discussed.

3.1 The ILD Vertex Detector Specifications

To satisfy the requirements imposed both from the physics goals of the experiment and its running conditions consists a major challenge, since these requirements are often in conflict. We can highlight this statement with the following example. The very high single point resolution implies a highly segmented sensor. This comes in contradiction with the rather fast readout speed, imposed by the beam induced background. Moreover, the big number of columns of a highly segmented sensor increases the power consumption. That might necessitate a complex cooling system, which could add on significantly on the overall material budget of the detector. In the following sections we are going to explain these trade–offs, and describe the optimisation procedure for the sensors and the detector.

3.1.1 Physics Driven Requirements

Impact Parameter Resolution

In order to reconstruct displaced vertices, or to measure the vertex charge, each track has to be assigned correctly to its corresponding vertex. To achieve that, the impact parameter of the tracks has to be measured with the highest possible accuracy. The impact parameter is defined as the distance of the closest approach of the track helix to the Interaction Point (IP) (see equation 3.1). The parameters a and b can be estimated analytically, and are approximately given by the two following expressions:

$$\mathrm{a} = \sigma_{s.p.} \frac{R_{int} \oplus R_{ext}}{R_{ext} - R_{int}} \tag{3.2}$$

$$b = R_{int} \cdot 13.6\, MeV/c \cdot z \cdot \sqrt{\frac{x}{X_0 \sin\theta}} \left[1 + 0.038 \cdot \ln\left(\frac{x}{X_0 \sin\theta}\right)\right] \tag{3.3}$$

*standing for Complementary Metal Oxide Semiconductor
†standing for Minimum Ionising MOS Active pixel sensor

where $\sigma_{s.p.}$ is the spatial resolution of the sensors, z is charge of the impinging particle, $\frac{x}{X_0 \sin\theta}$ the material crossed by the particle given in units of radiation length and R_{int} and R_{ext} are respectively the inner and outer layers radii.

Thus the parameter a depends on the single point resolution $\sigma_{s.p.}$ and the lever arm. The lever arm is defined as the distance $R_{ext} - R_{int}$. The parameter b depends on the distance of the innermost layer to the IP and the material budget (x/X_0). It becomes the dominant factor for either low momentum particles or particles crossing the VXD layers at a rather shallow angle, while a dominates for high momentum particles. This is clearly deduced from equation 3.1.

To optimise the VXD performance, one has to target for sensors with fine granularity, the lightest possible detector's ladders, placed at a minimal distance from the IP and a large enough lever arm, while keeping the power consumption of the sensors at low levels. Power consumption generates heat, which needs to be swept away by an appropriate cooling system. The higher the heat level, the more complex, therefore massive, the cooling system. Thus the discussion for power consumption is justified inside the framework of the material budget minimisation. During the procedure of the VXD performance optimisation, we have to take into account the running constraints imposed by the experiment. Moreover, it often happens that the simultaneous optimisation of all those aspects is not feasible, because it ends up with conflicting requirements. These efforts will be described in the following sections.

The ILD collaboration has set as target values for the parameters a $\leq 5\,\mu$m and for b $\leq 10\,\mu\text{m} \cdot GeV/c$. Such a high impact parameter resolution has never been achieved by collider experiments in the past. The table 3.1 summarises the values of a and b parameters achieved or foreseen by several collider experiments. The ILD indicative sensor's and ladder's design specification required to achieve the desired performances are the following:

- sensor single point resolution $\lesssim 3\,\mu$m
- radius of the first layer ~ 15 mm
- material budget of the first layer restricted to a few per mill of radiation length

The limiting factor for the radius of the first layer is the fact that the beam tube should be sufficiently far from the dense core of tracks from pair background particles, see section 2.2. The foreseen beam pipe radius is $\sim 14\,$mm, which means that the minimum innermost layer radius could not be less than $\sim 15\,$mm. This innermost layer radius, combined with a sensor's single point resolution of $\sim 3\,\mu$m can provide the required pointing accuracy. Concerning the material budget for each layer, the target values are $\leq 0.11\%X_0$ for the single layers and $\leq 0.16\%X_0$ for the double layers geometry. These values are quite aggressive and are fairly well inside the upper limit that b $\leq 10\,\mu\text{m} \cdot GeV/c$ sets. They are considered as target values, which are not yet

collider	a (μm)	b ($\mu\mathrm{m} \cdot GeV/c$)
LEP	25	70
SLC	8	33
LHC	12	70
RHIC	13	19
ILC	≤ 5	≤ 10

Tab. 3.1: Achieved (or ambitioned) impact parameter resolution for different collider experiments.

proven to be achievable. Presently, there is not any available technology that can offer such light ladders, while preserving the overall performance. Total material budgets of about twice the target values seem however already within reach. The above values concerning granularity, material budget and distance from the IP are the reference values as stated in the LoI of the ILD, thus they are going to guide our efforts.

Cooling System

The VXD cooling system is a crucial element of the design since it introduces additional material. Being located outside the VXD fiducial volume, it does not affect its performance, but it may impact (locally) the outer tracking performance and the particle flow analysis. As already mentioned in section 2.3.2, the material budget of the cryostat is bigger than the one of all VXD layers together. A more detailed analysis of the subject will be given in the section 3.4. In any case it is very promising that the time structure of the beam (see sub–section 2.1.2), with a machine duty cycle of 0.5 %, may allow to switch off the sensors between the bunch trains (power cycling). This can reduce substantially the sensor power consumption, thus the complexity of the cooling system.

Efficiency

A very important characteristic for a detector is the efficiency. The efficiency is a quite wide concept. Here we are going to stress the importance of the detection efficiency and the tracking efficiency for the VXD. The detection efficiency can be defined as the ratio of the number of detected particles over the total number of particles traversing the sensitive volume of the detector. On a multi-layer vertex detector missing a hit in one layer results in a less accurate track reconstruction. This is even more important if the missing signal is in the innermost layer. Then the pointing accuracy will be significantly degraded.

Correspondingly, the tracking efficiency can be defined as the ratio of the recon-

structed tracks, versus the total number of tracks of an event within the geometrical acceptance of the detector. It is very difficult for the VXD to cover the very forward area while keeping the material budget and multiple scattering effects at an acceptable level. Additionally, the vast majority of the pair background particles are directed to the very forward region. Thus this region cannot be instrumented. Finally, there are purely practical limitations, like the beam pipe.

Summarising, full hermeticity cannot be achieved, and it is in particular inevitable that tracks with very shallow angle will be lost. Regarding the design of the forward part of the vertex detector, two options are envisaged. The first is a long barrel design, and the second is short barrel complemented by forward endcap disks. The advantage of having forward endcap disks is that a particle crossing them at a shallow angle sees less material budget compared to the other option, thus suffers less from multiple scattering. On the other hand, the cabling and services of the vertex detector have to pass in front of the disks, resulting in a degradation of the overall performance. Taking into account the aforementioned considerations, the VXD adopts a long [‡] barrel approach. The polar angle coverage reaches down to 13.5^0, when the FTD hits are also included.

3.1.2 Running Constraints

The running constraints of the ILC environment play a major role on the VXD design, comparable to the one of the physics goals themselves. They are imposed by the function of the collider itself. Concerning the VXD, these are mainly the beamstrahlung, and (possibly) the electromagnetic interference with the RF power emitted by the beam.

Pixel Occupancy

As mentioned in section 2.2.2, the hits from pair background dominate the VXD occupancy. The occupancy can be calculated from the hit density provided in [De Masi 09] and requires accounting for technology–dependent parameters. The latter include the pixel pitch, the thickness of the sensitive part of the sensor, the cluster multiplicity and the detector time resolution. Thus the occupancy can be calculated only within a specific technology. The results are summarised in table 3.2, assuming average CMOS parameters. For the cluster multiplicity, the thermal diffusion of the charge was taken into account. Concerning the innermost layer, assuming the existence of the anti-DID field, and a 25 μs integration time (68 bunch crossings), the estimated occupancy amounts to $\sim 1.5\,\%$.

The effects of a high occupancy on the physics studies are mainly the following: the formation of a combinatorial background on the track reconstruction and the degrada-

[‡] In general, a “short” barrel design is complemented with endcaps, while a “long” one not.

technology	Integration Time (μs)	Pitch (μm)	Sensitive Thickness(μm)	Occupancy (%)
CMOS	25	20	15	1.5

Tab. 3.2: The sets of parameters that were used to calculate the occupancy for the CMOS technology option. The occupancy was calculated assuming the presence of the anti–DID field.

tion of the impact parameter resolution (when a beamstrahlung hit is superimposed on a physics hit). The smaller the number of bunch crossings over which one integrates within a readout frame, the smaller the combinatorial background. A small time resolution is thus desirable in order to keep the occupancy at an acceptable level. Its achievement is the purpose of intense $R\,\&\,D$.

There are two alternative philosophies on how to read out a sensor, both of them following from the particular time structure of the beam. Either the readout is performed continuously, or the hit information is stored until the end of the bunch train and the sensor is read between two consecutive bunch trains. In the latter case, a time stamping strategy is essential. In the former, providing the readout is fast enough, time stamping within a frame is not necessary. Hence, depending on the approach that a technology follows, time resolution is provided either by the integration time of the sensor or by the timestamp. The choice between both options is influenced by the degree of complexity of the signal processing logic that can be integrated in the sensor. These aspects are strongly dependent on the sensor's technology. A more detailed discussion of the time resolution, concerning MAPS, will follow in section 3.4.

Double Hit Separation

Directly related to the occupancy due to the beamstrahlung, is the necessity for the VXD sensors to feature a fine double hit separation. It can be expressed as the ability to discriminate between two different hit clusters that are very close to each other. Hit merging will result in a biased reconstruction of the hit position which will affect the track extrapolation. This will lead to a degradation of the impact parameter resolution. Taking into account the expected event rate of the ILC originating from hard electroweak interactions, this will happen very rarely. On the other hand, it could be important concerning the beamstrahlung hits, especially for the inner layers. A fine double hit separation will prevent the superposition of the physics hit clusters with the pair background clusters, preserving this way the sensors single point resolution.

Double hit separation has a great importance for the track reconstruction efficiency as well. In the framework of particle flow analysis, where each particle inside a jet should be reconstructed, a fine double hit separation distance is crucial in order to extrapolate efficiently the tracks towards the inner layers of the VXD.

Radiation Tolerance

The main source of ionising radiation is the pair background. Obviously, the innermost layer of the VXD will be subject to bigger rate, thus will suffer from more severe radiation damage. Taking into account the calculated hit density which amounts to 5 hits per cm^2 per BX, the estimated dose will be $\sim$ 150 kRad per year. This value includes a safety factor of 3 to account for the uncertainties of the simulations.

The sources of the non–ionising radiation are again the pair background particles, which dominate over the neutron gas that circulate inside the detector. The neutron gas is produced through photo-nuclear reactions from photons, electrons or positrons that hit the beam pipe, the mask or the beam dump. The estimated flux coming from neutrons amounts to $\sim 10^9\, n_{eq}/\mathrm{cm}^2$. Another potential neutron source comes from the calorimeters of the experiment. Their fluence may reach $\sim 10^9\, n_{eq}/\mathrm{cm}^2$ per year as well.

The annual flux that corresponds to the pair background is $\sim 10^{11}\, n_{eq}/\mathrm{cm}^2$, one or two order of magnitudes higher than the one coming from neutrons. A safety factor of 3 is also included in this estimation and the assumed NIEL [§] factor of the electrons is 1/10. Summarising, a sensor for the VXD should withstand a dose of $\sim 300\, kRad$ and a flux of $\sim 2 \times 10^{11}\, n_{eq}/\mathrm{cm}^2$, assuming at least 2 years of operation before replacement [Winter 09].

Electromagnetic Interference

The beam bunches, while traversing the Interaction Region (IR), emit vast amounts of RF power. The RF power comes from the beam wake fields, and reflects the complexity of the instrumentation around the IR seen by the beam particles. If the beampipe is made of a continuous metal enclosure, it should in principle keep the RF confined in the beam pipe.

However, in order to control the luminosity, a number of monitoring equipment should penetrate the beam pipe. These could create apertures from where the RF could escape and saturate the VXD electronics. Such a case was confronted at the Stanford Linear Detector (SLD), the only pixelated vertex detector existing so far on a linear collider, where after each bunch crossing a δ function shaped RF pulse was observed. The source of this RF leakage at SLD is still unknown. It caused a saturation on the analogue front end vertex detector electronics. The solution of this problem was to delay by 10 μs the readout [Damerell 05]. Unfortunately a similar solution will not be applicable at the ILC.

A possible approach could be the use of an aluminium Faraday Cage (FC). Currently, the FC is a part of the official design of the ILD. The main drawback of this

[§]Standing for Non–Ionising Energy Loss

concept is the additional material budget (which is however rather modest), see section 2.3.2. However, they do exist some considerations whether it will be effective or not. Special care should be taken about cabling and other services (for example cooling pipes) that have to penetrate the FC to enter in the vertex detector, in order not to provide apertures for the RF power.

Finally, whether the effect of the Electromagnetic Interference (EMI) will be a problem or not, is strongly depending on the sensor's technology. Generally, sensors that will be read out between the trains, are considered more robust against EMI. Concluding, we must stress that the danger of EMI is far from being established.

3.1.3 Sensor Technologies

To summarise on the VXD specifications, we can say that the physics goals and running conditions call for sensor technologies that offer a high granularity and a low material budget, while minimising occupancy on the inner layers. On the other hand, the constraints on radiation tolerance are moderate. A number of candidate technologies exists. A non exhaustive list includes the planar CMOS sensors [Turchetta 01], the 3D integrated CMOS [3DIC], the deep n–well CMOS [Traversi 08], the DEPFETs [Feld 09], the FPCCDs [Sugimoto 09] and the ISIS [ISIS]. Currently, no present technology can satisfy all the above requirements. However, intensive R&D is going on and it is expected that in the near future some of them will reach the needed level of maturity. The most innovative approaches, i.e. 3D integrated sensors, offer the highest potential but its development still needs many years. In the following sections, we are going to focus on the planar CMOS sensor technology.

3.2 CMOS sensors

Active Pixel Sensors (APS) are widely used since the mid-nineties in the imaging industry as an alternative solution to Charge Coupled Devices (CCDs). They are called active because each pixel contains a photodiode and its own active amplifier. In our days, they are widely used for web or mobile phone cameras. They are manufactured via standard CMOS processes (thus they are also called CMOS sensors). CMOS is the dominant technology for integrated circuits construction mostly because offering submicron feature size, it allows for a very high density of logic functions. It exhibits very good noise performance (even though CCDs remain better in this aspect). Moreover it is characterised by a low power consumption, since the circuitry in each pixel remains active only during readout.

CMOS refers to a certain type of Field Effect Transistor (FET). MOSFETs can be manufactured in extremely small sizes and their fabrication procedure is relatively simple. Moreover, digital functions and memories can be realised in circuits that use

only MOSFETs, without including any resistors or diodes, thus they keep a low power consumption. These are some of the reasons that lead CMOS to be the dominant technology for integrated circuits fabrication nowadays. Figure 3.1 illustrates the structure of a MOSFET. On a p-substrate, 2 highly doped n-diff implantations are created, the source and the drain. On the surface of the substrate, between source and drain, an extremely thin layer of silicon oxide (SiO_2) serves as an insulator. On top of it a layer of metal is deposited in order to serve as the gate electrode. Currently, polysilicon has substituted metal as the preferred gate material. Metal lines (usually aluminium made) are used as ohmic contacts between the structures of the transistor. The number of metal lines depends on the feature size of the fabrication process. The bigger the number, the more complex logic can be integrated at the chip.

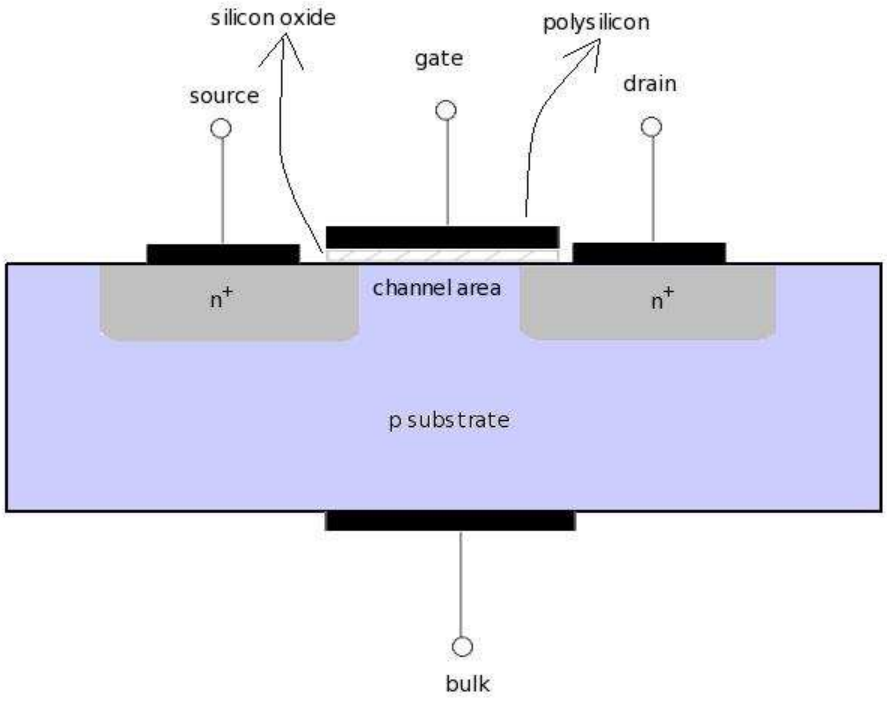


Fig. 3.1: Section of an NMOS FET.

The current in a MOSFET is constituted from only one type of charge carrier. The one illustrated in figure 3.1 is an n-channel MOSFET or NMOS, where the current

is constituted only from electrons. If p-doping is replaced by n-doping and vice versa we end up with a PMOS transistor. Here the current is constituted from holes. In the CMOS acronyms, C stands for complementary. This implies that both kind of transistors are used in CMOS processes, which allows the fabrication of microcircuits with excellent features and complex logic.

3.2.1 CMOS Sensors as Charged Particles Detectors

CMOS sensors suitable for charged particle tracking detection are manufactured with the twin–tub fabrication process. The NMOS transistors are built in p–wells and vice versa. The n– and p–well doping concentrations can be tuned independently. Industry uses this process because this way the NMOS and PMOS transistors can be separately optimised, therefore the performance of the chip can be maximised. These wells are grown on a thin lightly doped epitaxial layer. Its presence makes the fabrication process a bit more complex, but it is justified because the relatively high resistivity of the epitaxial layer (due to its low doping) significantly reduces the possibility of latch–up. As we will see, the epitaxial layer plays the role of the sensitive volume of the CMOS sensors. The twin–tube process allows to make the whole surface of the sensor sensitive to impinging particle, which is mandatory for charged particle tracking applications.

The special features that make CMOS sensors especially attractive for charged particle tracking are high granularity, low material budget, and cost effectiveness. The most original feature is the fact that signal sensing and processing share the same substrate. For this reason they are also called MAPS (stands for Monolithic Active Pixel Sensors). We will further discuss these aspects later.

The structure and the operation principle of CMOS sensors is illustrated in figure 3.2. The basis is a moderate quality highly doped p^+ substrate. Moderate quality implies the presence of a relatively big amount of crystal defects. On top of it, a p^- epitaxial layer is grown. The epitaxial layer is the sensitive part of the detector, hence silicon of higher quality is required in order to avoid the recombination of the charge carriers (electron–holes). The charge collection is realised by n–wells, which form pn junctions with the epitaxial layer. The area of the collection diode is just a small fraction of the total area of the pixel. Around the collection (or sensing) diode there are highly doped p–wells whose boundary surfaces with the epitaxial layer create potential barriers that guide the charge to the sensing diode. They also serve as bulks for the NMOS transistors implementation.

Industrial CMOS processes provide low resistivity epitaxial layers (at least until recently - see section 3.4.2). Typical values for the doping concentrations are 10^{15}at/cm^3 for the epitaxial layer, 10^{19}at/cm^3 for the substrate and 10^{17}at/cm^3 for the p–wells. The size of the depletion region is defined through the doping concentration. This leads to the fact that the sensor's sensitive volume is mostly undepleted, except for a very small region around the pn junction. The depth of this region is limited to a fraction

of a micrometer.

The charge carriers generated by the traversing particle diffuse thermally inside the sensitive volume. Across the boundaries of the epitaxial layer with the p–wells and the substrate, a built-in voltage is produced due to the different doping levels, defined by the following equation:

$$V_b = \frac{kT}{q} \ln(\frac{N_{p+}}{N_{p-}}) \tag{3.4}$$

where k is the Boltzmann constant, T the temperature, q the elementary charge and N_{p+} and N_{p-} the different doping concentrations on both sides of the interface. This potential acts as a barrier for the charge carriers (electrons) between the boundary surfaces of the epitaxial layer, the substrate and the p-wells. So the charge carriers are restricted to diffuse inside the epitaxial layer. This way they are guided towards to an n–well diode, where they are finally collected (except of them being recombined within the medium traversed).

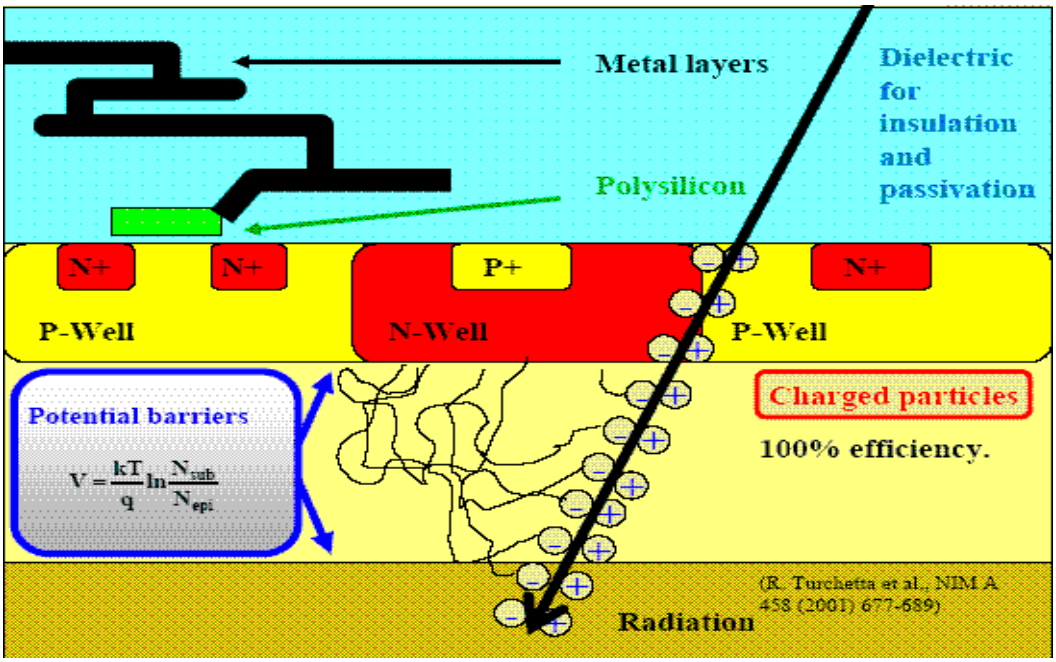


FIG. 3.2: Structure and principle of operation of CMOS sensors. The sensitive part is the epitaxial layer, which is illustrated with light yellow colour. The epitaxial layer is located on top of the substrate (bottom part of the picture, with brown–yellow colour). Over the epitaxial layer is the signal sensing and processing part.

The charge collection through thermal diffusion impacts the charge collection efficiency (CCE), which is relatively different from the one of a fully depleted sensor. This

is directly related to the sizeable average path of the charge carriers before being collected. The probability of the electron–hole recombination obviously increases with the path, which leads to a lower CCE. Moreover the charge is shared between more pixels, compared to fully depleted sensors. Thus the average charge collected by one pixel will be lower. The sizeable average path of the charge carriers results also in relatively long collection times ($\sim 100\,ns$). On the other hand, spreading the charge between several pixels, can substantially improve the spatial resolution if center of gravity positioning algorithms are employed. Due to the high quality of the epitaxial layer, the lifetime of the charge carriers happens to be large enough to avoid losing a significant part of the charge through recombination.

Main Features of CMOS Sensors

As we already mentioned, the MAPS technology offers some special features that makes it a very attractive option for realising low mass, high precision tracking and vertexing devices for HEP applications. These aspects are developed below.

Monolithic approach: The main feature of the MAPS is that signal processing can take place on the same substrate as the charge sensing element. The complexity of the logic that can be integrated in pixel is however limited by the fact that only NMOS transistors can be used, and of course by the pixel dimensions. As mentioned in section 3.2, a PMOS transistor needs a n–well as a bulk. If this n–well is formed inside the sensitive area it will behave as a parasitic charge collecting diode. Thus the charge collection efficiency of the sensor would be substantially decreased. For this reason, PMOS transistors are integrated only at the periphery of the chip.

Granularity: CMOS processes can offer a very small feature size ($< 100\,nm$). The minimal dimension of a pixel can reach $\sim 10\times$ the feature size, so CMOS sensors may obtain a pixel pitch of a few micrometers. Taking advantage of the charge sharing between the neighbouring pixels, a very fine (submicron) spatial resolution can be achieved.

Material budget: Only the epitaxial layer, whose thickness amounts to $O(10\,\mu\text{m})$, is used for signal generation. The substrate provides mechanical stability to the chip and reflects the charge carriers back into the epitaxial layer. The charge reflection is actually served only by the interface between the 2 layers. Hence, the bulk of the substrate can be partially removed. The whole sensor can be thinned down to $\sim 50\,\mu\text{m}$, while still maintain a good production yield and a comfortable handling during the detector assembly. A drawback of the very thin sensitive volume is the small magnitude of the generated signal. Since MIPs ¶

¶MIP stands for Minimum Ionising Particle

produce typically 80 electron - hole pairs per micrometer, the generated signal will be $O(1000e^-)$. Therefore, very low noise electronics should be used for in-pixel signal processing. Finally, the total material budget benefits from the fact that the sensitive volume and the signal processing circuitry are integrated on the same chip. Here we should stress that a peripheral circuitry of necessary supporting electronics, non sensitive to impinging particles, adds–on the overall material budget of the sensor. The reader could refer to section 3.3.3 for more informations about the dimensions and the functions of the peripheral circuitry of a real size CMOS sensor, that is used for HEP applications.

Radiation tolerance: Ionising radiation merely induces surface damage to MAPS, by charge accumulation at the interface between zones of different doping (e.g. gate oxide and polysilicon). These charges can increase the diode's leakage current. The thinner the diode, the smaller the amount of accumulated charges, thus the higher the immunity of the sensor to ionising radiation. Transistors featuring very thin gate oxide can be manufactured using submicron fabrication processes. Non–ionising radiation induces bulk damages by displacing atoms of the crystal lattice, increasing this way the crystal defects. These constitute traps for the charge carriers that diffuse inside the epitaxial layer. This way, their average lifetime is reduced. Therefore, the CCE decreases. One solution to compensate for the loss of CCE due to non–ionising radiation is to try to decrease the charge collection time. This can be realised with two ways: either going for smaller pixel sizes or for high resistivity epitaxial layers, where the depletion region occupies a significant fraction of the pixel volume (see sub–section 3.4.2).

Cost effectiveness: Last but not least, one of their prominent features is the cost effectiveness, since industrial processes are used. This fact has the drawback that HEP applications, which represent moderate income potential for this industry, cannot drive the development of the technology in favour of the charged particle tracking performance. One needs to adapt to the process parameters as offered in the market. On the other hand, the vast amounts of money invested by the industry into CMOS imagers R & D, led to such a fast development of the CMOS processes which is undoubtedly beneficial also for the HEP applications. The low cost of industrial CMOS processes is of primordial importance mostly for the R & D, where numerous of prototype sensors should be fabricated in order to optimise the technology. It also comes out to be a determining factor for the achievement of large area detectors.

3.2.2 Pixel Architecture

As aforementioned, the amount of charge generated by an impinging MIP for typical epitaxial thicknesses is $\sim 1000\, e^-$. Since the charge collection is realised with thermal

diffusion, the charge is expected to be shared between several neighbouring pixels. Hence, the amount of charge collected from a sensing diode in a pixel will usually not exceed a few hundreds of electrons. It is crucial to obtain a high Signal to Noise Ratio (SNR). The SNR is the figure of merit for a sensor since it dictates its basic performance; the efficiency–fake hit rate interplay and the spatial resolution (see sub–section 3.2.3). The SNR follows a Landau distribution. This Landau shape is imposed by the the shape of the distribution of the energy loss (thus signal charge generated) of MIPS in thin silicon layers; an asymmetrical distribution with high tails towards large values of energy loss [Bischel 88]. The SNR of a sensor is defined as the Most Probable Value (MPV), (e.g. the peak), of the Landau distribution.

Obviously, the signal is generated inside the sensitive volume of the sensor, while noise contributions could added from everywhere. Therefore, in order to achieve a high SNR value, it is recommended to amplify the signal the closest possible to its source, e.g. in the pixel itself. Thus a preamplification stage is built inside the pixel (see figure 3.3). This is accomplished from a source follower which is constructed by a NMOS transistor (M2) connected to a current source outside of the chip. The collected charge is converted into voltage through the collecting diode's parasitic capacitance. The passage of a particle appears as a discharge of the capacitor C_d, thus a voltage drop at the gate of the M2 transistor (node K). The amplification is performed through the modulation of the transistor's current cause of the voltage variations at the node K. Here we have to stress that the C_d is subjected to a continuously discharging, with a slow pace, due to the leakage current. The signal is detected as a more abrupt discharging, thus a sudden voltage drop at the node K. The discharging due to the diode's leakage current need to be compensated for, otherwise it would lead the pixel into saturation. In the following sections we are going to describe the 2 basic designs of the built-in preamplifiers. Their main difference lies on the way they compensate for the diode's leakage current. In one design (3T) a reset transistor is used for the leakage current compensation, while in the other (SB) a biasing diode.

3 Transistors Pixel Design

The function of the standard 3 transistors (3T) pixel design relies on 2 operational phases. During the first one, the readout phase, the pixel is addressed 2 times for readout, separated of course by a time interval equal to the integration time of the sensor. During this phase the pixel is sensitive. The transistor M1 (see left part of figure 3.3) is closed and the charge of the C_d is slowly decreasing due to the diode's leakage current. In the second phase, initiated when a reset signal arrives at the pixel, M1 is open in order to allow the recharge current to restore C_d to its initial voltage. During this time, the pixel is not sensitive. The dead time is usually $\sim 50\,\%$, but it maybe much shorter.

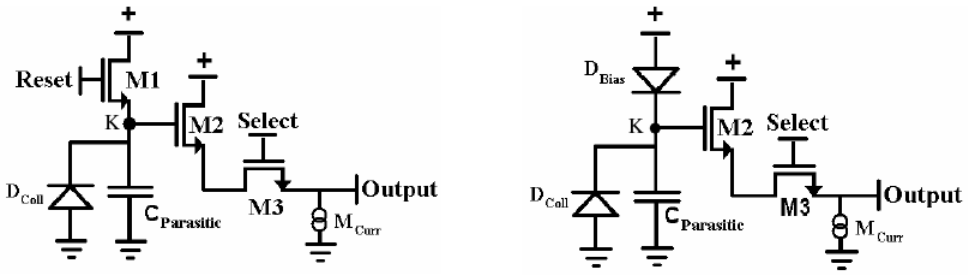


FIG. 3.3: 3T (left) and SB pixel (right) architectures.

Self-Biased Pixel Design

The self-biased (SB) design is an alternative way to compensate the leakage current that, unlikely the 3T, features no dead time. Here the recharge current is provided by a forward biased p^+-n junction placed inside the n–well. Until a signal occurres, the leakage current of the diode is in equilibrium with the recharge current. The signal created by a particle crossing the sensor, causes a sudden discharging of C_d, which is followed by a recharge current. In order to detect the physics signal, the recharge procedure should be relatively slow compared with the integration time. If the charge of C_d is restored too fast, the voltage drop that occurred at the node K will not be detectable at the time when the pixel will be addressed for readout, because the voltage would have been already restored. Therefore, we will never be notified for the passage of the particle. A slow recharging allows for a detectable difference in the value of voltage between two consecutive readout, thus the detection of the particle. However, a loss of a small fraction of the signal cannot be avoided. Several considerations could be arise regarding the behaviour of the SB pixels against a high particle flux. Concerning the VXD, this is crucial mainly for the inner layer(s) due to the high occupancy generated from the pair background. It was demonstrated analytically ([Deveaux 08b]) that provided a fast readout and an accordingly slow recharging, SB pixels can cope with a high hit rate. However they do have an upper limit for their occupancy.

3.2.3 Sources of Noise in MAPS

The performance of a sensor can be characterised by its delivered SNR. The applied threshold on the SNR defines the trade–off between efficiency and fake hit rate. A high SNR threshold effectively suppresses the reconstruction of fake hits, caused by large noise fluctuations. On the other hand, it may cut out also hits due to particles, degrading this way the efficiency. A low SNR threshold allows for 100 % efficiency, but may also allow noise fluctuations to be considered as particle hits. The same logic applies for the resolution. The spatial resolution of MAPS benefits from the charge sharing between the adjacent pixels, and the consecutive application of a positioning algorithm (section 4.2.3). A tight cut on SNR will cut out pixels with a lower charge sharing, reducing the accuracy of the positioning algorithms. On the other hand, a loose SNR cut may allow noise fluctuations to enter into the calculation, thus spoil the hit's position reconstruction. Therefore, the SNR thresholds should be selected carefully, in accordance with the specific requirements of each experiment. A high SNR value, can yield simultaneously excellent overall performance on efficiency, fake hit rate and spatial resolution.

The origins of the noise of a CMOS sensor can be generally divided in 2 main categories, the Fixed Pattern Noise (FPN) and the Temporal Noise (TN). The FPN originates from non-uniformities of pixel response inside the pixels submatrices. It can be regarded as the offset (or pedestal), which needs to be subtracted from the pixel's response in order to measure the signal.

The TN may arise from a variety of factors. Its main sources are the shot noise, the thermal noise or the 1/f noise [Deptuch 02]. Depending on the operation phase (reset, integration or readout), difference sources may contribute on the TN. This implies that a separate noise analysis should be done for each operation phase.

Noise during reset: This noise arises when the transistor M1 is open in order to restore the C_d charge. Thus it appears in the 3T-like pixel designs. It can be described by the equation that governs thermodynamic fluctuations:

$$\overline{V_n^2} = \frac{kT}{C_d} \tag{3.5}$$

where V_n is the potential of the noise, k is the Boltzmann constant, T the temperature and C_d the diode's capacitance. It can be the dominant source of TN.

Noise during integration: This is mainly the shot noise caused from the statistical fluctuations of the leakage current of the diode. Provided the readout is fast, the shot noise contribution gets marginal. It can be described by the equation:

$$\overline{V_n^2(t_{int})} = \frac{q i_{leak}}{C_d^2} t_{int} \tag{3.6}$$

where i_{leak} is the leakage current and t_{int} is the integration time.

Noise during readout: During readout, the noise originates from the transistors M2 and M3, the source follower current source and the column switch, which has a column capacitance C_1. The noise contribution of each transistor is proportional to $\frac{kT}{C_1}$ and is a function of their transconductance, as well as of the output's conductance.

Correlated Double Sampling

A method to suppress the noise contribution arising from some of the noise sources is the Correlated Double Sampling (CDS). During the readout of the sensor, two samples are taken. The first sample is subtracted from the second in order to search for possible signals. It is a very efficient noise suppression process. Concerning MAPS, it eliminates the FPN and the reset noise, which has a large contribution to TN. It also reduces the 1/f noise to some extent. In MAPS, the CDS can either be performed at the sensor periphery or inside the pixels.

3.2.4 MAPS Readout

The typical time needed for the charge collection by diffusion in a pixel is $\sim 100\,\text{ns}$ for an undepleted epitaxial layer. This sets an ultimate value for the time resolution of these detectors which, concerning MAPS, is actually dominated by other aspects. These are the time needed to obtain the information from all the pixels of the sensor, as well as the pixel occupancy.

Addressing sequentially the pixels and multiplexing their output in one single bus was the first approach for the readout of the MAPS. The discrimination thresholds for hit identification were set offline and could be selected specifically for each pixel. This is a very efficient way to compensate for the non uniformity of the pixel response. The integration time is a function of the operational frequency of the bus, which is typically $\sim 50\,\text{MHz}$, and of the number of the pixels connected to the bus line. This method is very time consuming for a real size sensor of $\sim 10^5 - 10^6$ pixels. It is in the order of magnitude of a millisecond. For the needs of the VXD, this time should be decreased by two orders of magnitude. A second major concern is the expected high data flow ($O(100)\,Gbits/s/cm^2$) that may set severe requirements on the data acquisition system (DAQ). The data sparsification should take place as close as possible to the sensing pixels arrays.

An approach could be to timestamp the hit and integrate the whole data sparsification functionalities inside the pixel. This approach is followed with Hybrid Pixels Sensors (HPS) used by LHC experiments but it cannot be followed by a CMOS based VXD for the ILC. Such a complex signal processing requires both NMOS and PMOS transistors. As already mentioned, the n–wells that will be serve as the bulk for the PMOS transistors will become parasitic charge collection diodes. Moreover, the space

needed for the implementation of these circuits will lead to the increase of the pixel's pitch, thus degrading the single point resolution. Finally, large pixel density, which follows from their smallness, typically 50 times smaller than HPS used at LHC, would generate a large power dissipation.

A viable solution could be to increase the readout speed of the pixels arrays and integrate the data sparsification logic on the periphery of the sensor. A way to improve the time resolution is to group the pixels in columns that will be read in parallel. This method is usually called column parallel readout or rolling shutter mode and its principle is depicted by figure 3.4. The lines are addressed to be read sequentially by a shift register. The integration time is independent of the number of columns, since each column has its own bus line, and is a function only of the time to read a pixel and the number of the pixels per column. The number of readout buses is equal to the number of columns of the sensor. That in a real size sensor ($\sim 10^6$ pixels, ~ 1000 rows and ~ 1000 columns) can be translated in a decrease of $O(10^3)$ in the number of the pixels connected to one bus line. This corresponds to a decrease factor on the readout time of $O(10^3)$. The sparsification will be provided by discriminators, located at the bottom of each column and being common for all the pixels of a column. However, the on chip data sparsification requires more in-pixel functionalities, which increase the time needed to read a pixel. The pixels feature differences in the level of the FPN in the same order of magnitude as the generated signal. So in order to apply a common discriminator threshold for all the pixels in one column, an in-pixel CDS is mandatory. More information concerning architectures featuring in-pixel CDS can be found in [Dorokhov 06], [Degerli 05]. For this purpose, more than one clock cycle is required to read a pixel. Finally, an additional advantage of the column parallel readout is the power effectiveness. Only the line that is actually being read needs to be active. This fact reduces the power dissipated by the pixel array by two or three orders of magnitude compared to alternative approaches (e.g. in pixel sparsification) requiring each pixel to be continuously activated.

3.2.5 General Performance

More than 25 prototypes of the MIMOSA series have been produced since 10 years. The proof of principle of the MIMOSA sensors for charged particle tracking has been fully demonstrated. Several prototypes have been manufactured with different fabrication processes, featuring different epitaxial layer thicknesses and/or feature size. Moreover, they had various pixel pitches and designs. They have been extensively assessed in lab and beam tests. The tests included sensors irradiated at different level of radiation doses or neutron fluxes, carried out usually at room (up to 40^0C) or slightly negative temperatures.

With sensors delivering analog output, the collected signal and the pixel noise can be measured directly. Taking into account the very thin epitaxial layer and the charge

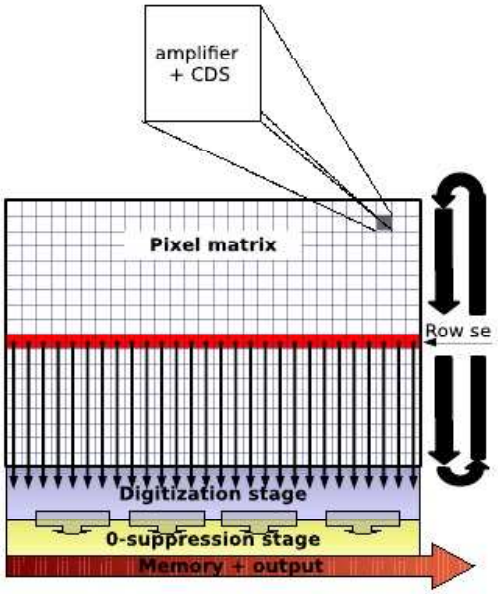


FIG. 3.4: The principle of column parallel readout [Baudot 10].

sharing, the signal collected by the seed pixel (see sub–section 4.2.3) amounts to few hundreds of electrons (usually 200–300). On the other hand, the observed pixel noise is $\sim 10e^-$ at 20^0C. That means that a value for the SNR of the seed pixel of ∼20–30 can be achieved. Such an SNR value can be translated to $\sim 100\,\%$ efficiency, with an average fake hit rate per pixel staying below 10^{-4}.

The spatial resolution depends on the pixel pitch and of course on the data output mode, analog or digital. For analog sensors it increases, roughly linearly, from $\leq 1\,\mu$m for a $10\,\mu$m pitch to $\sim 3\,\mu$m for a $40\,\mu$m pitch. These values are achieved using the non-linear η positioning algorithm (see section 4.2.3), which takes advantage from distribution of the signal charge over several pixels. It allows for a finer spatial resolution than the one expected from the binary positioning algorithm (see sub–section 4.2.3). For a binary output, a resolution of $\lesssim 3\,\mu$m for a pitch of $\sim 15\,\mu$m can be achieved.

Ionising radiation leads to an increase of the diode's leakage current, which may deteriorate significantly the noise performance of the sensor. A radiation tolerance

for doses up to 1MRad has been achieved, assuming a moderate cooling. It could be substantially improved by moving to a smaller feature size process, because the latter provides a thinner gate oxide.

The tolerance to non–ionising radiation depends strongly on the pixel's pitch. Sensors with a pitch of $10\,\mu$m can stand a flux of $10^{13}\,n_{eq}/\mathrm{cm}^2$. It decreases down to $2 \times 10^{12}\,n_{eq}/\mathrm{cm}^2$ for $20\,\mu$m pitch, see [Deveaux 08a]. Sensors featuring a small pixel pitch, due to the smaller average path of the charge carriers before being collected, lose less charge by trapping. They are thus more tolerant to non–ionising radiation. Regarding the radiation environment of the ILC experiments, present MAPS already exhibit a satisfactory performance. The tolerance to non–ionising radiation will be dramatically improved for sensors featuring a high resistivity epitaxial layer. A very detailed study on the radiation damage on MAPS can be found in [Deveaux 08b].

3.3 MIMOSA 26

Reticle size analog output sensors, with a pixel pitch of $O(10\,\mu\mathrm{m})$, are limited to a readout speed of $\sim 1\,$ms. This value is definitely inadequate for the ILD VXD readout time needs. The speed should be improved by 2 orders of magnitude to comply with the occupancy expected on the VXD inner layers. Such a dramatic improvement can be obtained by passing from analog to digital sensors where the pixels are grouped in columns and being read in parallel.

3.3.1 R & D Framework

The R & D for MIMOSA 26 was performed inside the framework of the European Union's EUDET program [EUDET], which aimed to provide infrastructure concerning the R & D for the ILC detectors. MIMOSA 26 was designed to equip the reference planes of the EUDET beam telescope, depicted in figure 3.5. The telescope consists of two arms with three planes of two MIMOSA 26 sensors each. MIMOSA 26 provides also the baseline architecture for the sensor foreseen for the STAR ([Szeleniak 08]) Heavy Flavour Tracker (HFT) upgrade. IN STAR HFT, it will be the first time that a MIMOSA sensor will be used in a collider experiment. First physics results are expected inside 2012. The ILD running constraints, especially of the VXD inner layers, call for a faster sensor. In the remaining part of the chapter, having MIMOSA 26 as a baseline, we will discuss the plans for the development of CMOS sensors suitable for the vertex detector of the ILD.

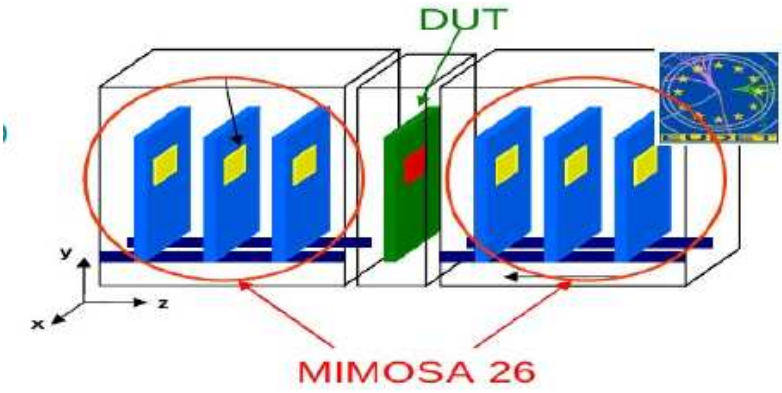


Fig. 3.5: The EUDET beam telescope. It consists of two arms each made of three planes of two MIMOSA 26 sensors. The Detector Under Test (DUT) is placed between the two reference planes (see sub–section 4.1.2 for beam telescopes general description).

3.3.2 Earlier prototypes

The first prototype of this readout scheme is MIMOSA-8, fabricated in the TSMC 0.25 μm technology featuring $\lesssim 7\,\mu$m epitaxial layer thickness. It consists of 32 columns. Out of them, 24 end with a discriminator in order to provide digital output. The remaining 8 columns have analog output for test purposes. It features in-pixel preamplification and CDS. The sensor was tested at the DESY beam test facility, that offers a 5 GeV e^- beam. The Equivalent Noise Charge (ENC) after CDS was found below 15 e^-. The measured detection efficiency was $99.3 \pm 0.1\,\%$ for an average fake hit rate per pixel of $\sim 10^{-3}$. The resolution was measured at the 120 GeV pion beam at the CERN-SPS, and was found $\sim 7 - 8\,\mu$m, compatible with the binary resolution (see sub–section 4.2.3) expected from the pitch (25 μm). The outcome of the tests was the validation of the pixel design, but also the necessity to move to a process that offers thicker epitaxial layer, mainly in order to obtain better detection efficiency and spatial resolution. The exploration of fabrication processes with several MIMOSA prototypes (for example MIMOSA 9, 11, 14) concluded that AMS 0.35 μm was the most suitable technology (by that time) for charged particle tracking. The MIMOSA 16 sensor is the full translation of MIMOSA 8 to AMS 0.35 μm. The test beam of the sensor took place with a 180 GeV pion beam at the CERN–SPS. The measured detection efficiency was $\sim 99.8\,\%$ for an average fake hit rate per pixel below 10^{-5}. The spatial resolution was $\sim 5 - 6\,\mu$m. MIMOSA 22 and its more radiation tolerant variant, 22bis, are the final prototypes of this R & D line. A detailed description of MIMOSA 22 / 22bis and full test beam data's analysis of MIMOSA 22bis will follow in the next chapter.

3.3.3 Sensor Description

MIMOSA 26 is the first full scale digital sensor of the MIMOSA series with integrated signal processing. It adopts the optimal pixel design of MIMOSA 22 / 22bis combined with the zero suppression chip SUZE-02, that performs on line data sparsification. Fabricated in the AMS 0.35 μm technology, it consists of 576 rows and 1152 columns resulting in a sensitive area of $\sim$ 2 cm^2 (figure 3.6). The pixel pitch is 18.4 μm. Column parallel readout and a nominal operation frequency of 80 MHz ensure an integration time of 115.2 μs. The maximum operational frequency reached is 115 MHz, meaning that the integration time can shortened to $\sim$ 80 μs. The signal is amplified by a preamplification stage and then undergoes the CDS stage. Both are performed in–pixel. This allows to digitise the signal using column level discriminators. A second double sampling takes place at column level in order to compensate for the FPN (see sub–section 3.2.3). Finally, the output is pipelined to the zero suppression circuit integrated on the chip. For test purposes the chip incorporates also an analogue output offering access to the pixel response at reduced readout speed. The pixel signals are therefore influenced by a larger integrated leakage current in this case. The inactive part of the chip is constituted by the following areas: the 200 μm wide top band that serves only for test purposes, and could be removed in view of the material budget minimisation for a vertex detector; at the chip bottom there is a 3 mm zone for the discriminators, the zero suppression circuit and the slow control; finally the 300 μm wide side band that performs the row steering, which could also be moved at the chip bottom in a fabrication process that offers more metal layers than the AMS 0.35 μm technology.

3.3.4 On Chip Data Sparsification

Fast readout vertex detectors have typically high data flow. In order to better handle the data, and relax the requirements of the DAQ, it is desirable to sparsify the data as close as possible to the signal sensing. The best option would be to integrate the zero suppression circuit on the sensor itself. CMOS technology allows that. The zero suppression circuit of MIMOSA 26, SUZE 02, is fabricated with the same process as the sensing part, hence integrated on the same chip. It takes as input the output of the discriminators and searches for fired pixels row by row, using a sparse data scan algorithm [Himmi 09]. It can handle up to 9 hits per row. This number is consistent with the estimated hit density foreseen for EUDET and STAR applications. The information is stored in a 96 kbit memory splitted in two buffers, allowing for continuous readout. The data suppression level is between 10 and 1000, depending on the occupancy.

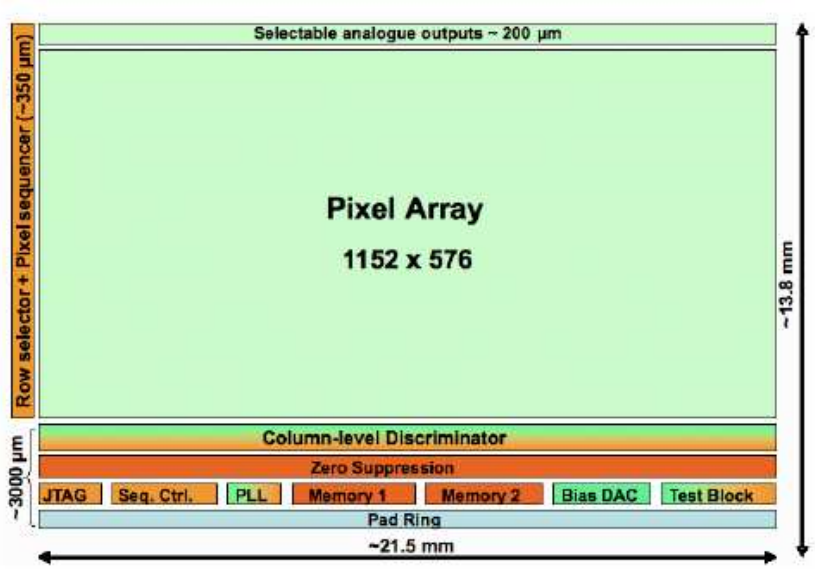


Fig. 3.6: Layout of the MIMOSA 26 sensor.

3.3.5 Performance

The sensor was extensively tested in the laboratory and in beam tests. In laboratory, using a ^{55}Fe source for calibration, it was possible to measure the ENC. It was found $\sim 12 - 13e^-$ when the sensor was operated at 80 MHz clock frequency. MIMOSA 26 was tested in beam at CERN–SPS, exposed to 120 GeV π^-. It exhibited a detection efficiency of $\sim$ 99.5% for an average fake hit rate $\leq 10^{-4}$ per pixel and a spatial resolution of $\sim 3.5\mu m$ (figure 3.7). The exhibited performance are considered adequate for the EUDET and STAR applications.

3.4 A Vertex Detector for ILD based on MAPS

In this section we will propose a possible configuration for the VXD based on MAPS. We are going to assume a double layers geometry and MIMOSA 26 as the baseline sensor architecture. Special care should be given to the integration of the MAPS on the VXD system. This is crucial in order to preserve the sensor's performance. Moreover, exploiting innovative integration techniques can further improve the detector's performance itself. Since some of the requirements for the inner layer are different from those for the outer layers, separate approaches will be followed.

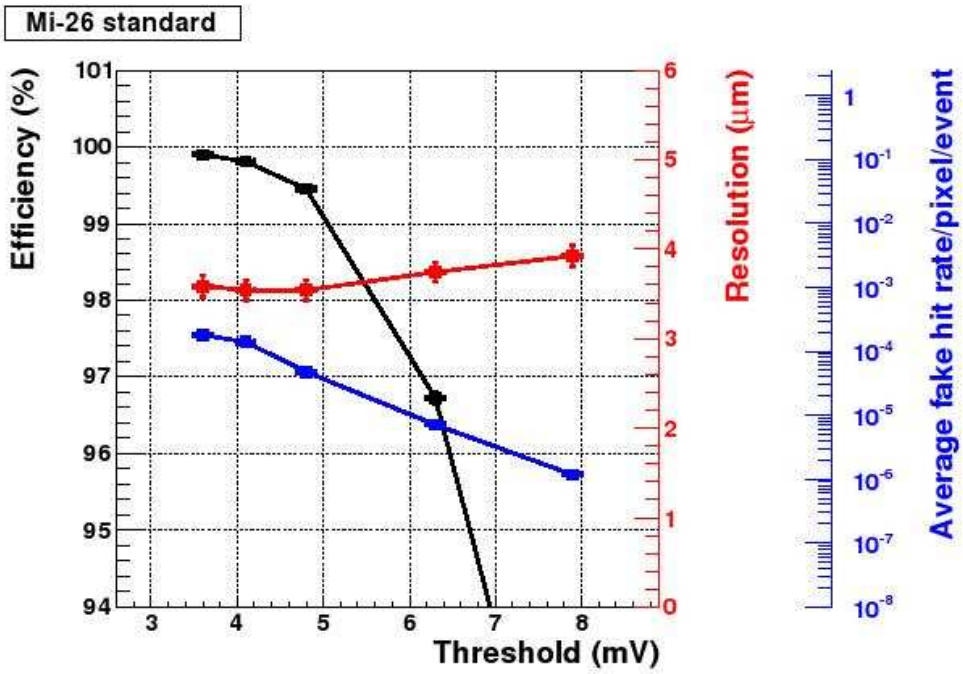


Fig. 3.7: MIMOSA 26 sensor detection efficiency (black line), spatial resolution (red) and fake hit rate (blue) as a function of the discriminator thresholds in units of mV.

3.4.1 Integration Issues

The sensors have to be integrated in a detector system in order to form a functional vertex detector. The integration framework covers all the necessary services: the mechanical support structure of the sensors, the cabling that will carry the data and provide the necessary electric power and steering signals, the cooling issues, etc. The challenge is to construct a functional detector, that will provide the required mechanical stability with the minimum additional material. This is crucial in order to preserve the excellent sensor's performance. In the VXD the sensors will be mounted on ladders, which will be arranged following cylindrical symmetry. The integration techniques that will exploit the double sided layers of the VXD05 geometry in order to improve the detector's performance will be presented in the following sections.

PLUME

The PLUME collaboration (stands for Pixelated Ladder using Ultra–light Material Embedding) has been set up with the objective to study the feasibility and the benefits of light ladders equipped on both sides with sensors, according to the double layers VXD geometry (VXD05 design). The layout of the PLUME ladder is illustrated in figure 3.8. The mechanical support of the ladder is made of 2 mm silicon carbide foam featuring a density of 4–8 %. Two kapton–metal flex cables are glued on each side of the support. The sensors are bonded on top of the flex cable. The flex cables transfer the output data from the sensors and provide all the necessary services to them. PLUME ladders are designed to be compatible with thinned sensors of different technologies. The first complete prototype, equipped with 6 thinned MIMOSA 26 sensors on each side was fabricated in 2011, featuring a cross–sectional material budget of $\sim 0.6\,\% X_0$. The target value of the total material budget is $\sim 0.3\% X_0$. It motivates the realisation of the next prototype, to come out in 2012. The material budget estimation for each component is summarised in table 3.3. The two different values stated for each component correspond to a rather aggressive, or a more conservative estimation.

Layer	thickness (μm)	budget (%)X_0
SiC foam (4–8 %)	2×10^3	0.092–0.184
sensor	35–50	0.037–0.053
metal	10–20	0.007–0.014
kapton	20–50	0.007-0.018
total		0.232–0.392

Tab. 3.3: Thickness and material budget estimation for each component of the PLUME ladder.

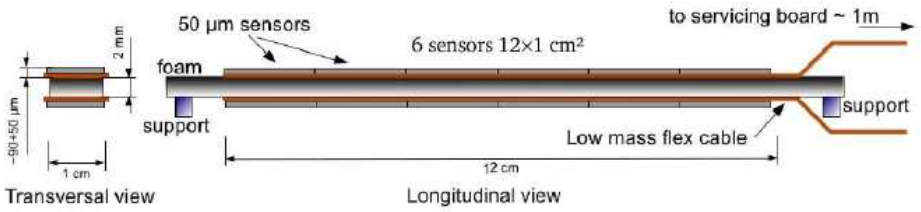


Fig. 3.8: The layout of the PLUME double sided ladder.

The PLUME project intends to investigate added values of the double sided ladders. The hits generated on the two sides by a traversing particle could be correlated to construct a minivector. This is expected to improve the hit-track matching efficiency, resulting in an improved spatial resolution. The improvement is expected to be more pronounced for shallow angle tracks. As it will be described in the section 3.4.3, the hit correlation can provide a significantly better time resolution.

The collaboration intends to investigate the following issues concerning the integration and functionality of the double sided light ladders for the ILD.

- The operation of the power pulsing in the 3.5 T magnetic field and the resulting Lorentz forces in order to preserve the impact parameter resolution.
- The heat dissipation on the ladder and the compliance with air flow cooling.
- The effect on resolution of possible vibrations due to the air flow cooling.
- The position accuracy of the sensors on the ladder in view of an effective alignment.

Alignment Studies

In order to maintain the high impact parameter resolution required by the ILD physics goals a very efficient alignment of the vertex detector should be performed. The positioning of the ladders should be accurate enough to fully exploit the intrinsic resolution of the sensors. For this purpose, an infrastructure that will allow the study of the alignment of multiple ladders is under construction within the EU FP7 program AIDA ([AIDA]). It is composed of a Large Area beam Telescope (LAT) and of a

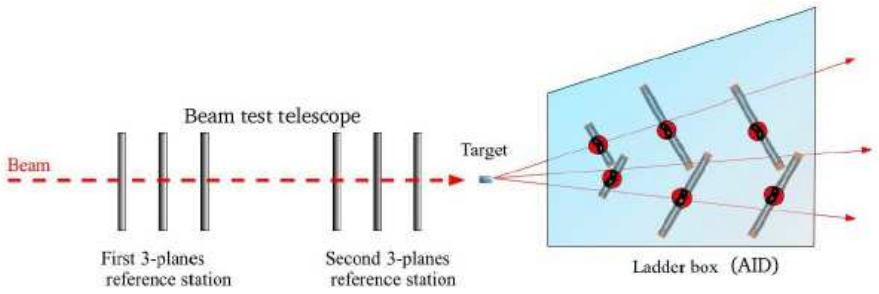


Fig. 3.9: The beam test infrastructure provided within the AIDA framework aiming to investigate the alignment issues on vertex detectors.

ladder box called AID (standing for Alignment Investigation Device). It is illustrated in figure 3.9. The AID box will contain 3 consecutive pairs of ladders (which could be PLUME ladders or other concepts), being like a ϕ sector of a vertex detector. A target could be placed in front of the box, allowing to study the vertex reconstruction capabilities.

Stitching

In the modern industrial CMOS processes the photomask is projected on a small area of the silicon wafer, called reticle. Thus the maximum available size of a CMOS sensor is restricted to the reticle size, which amounts to a few cm^2. The length of a VXD ladder reaches 12.5 cm, which means that several sensors should be mounted on the same ladder. Individual mounting of sensors implies inevitably dead areas between them. A way to overcome this problem is offered by stitching. The stitching is the replication of the same sensor on a single wafer without insensitive area in between them. Postprocessing allows to connect the replicas in order to form a single integrated circuit.

Stitching offers a number of advantages; the dead zones between the sensors can be eliminated; the mechanical rigidity of a large single object is higher than separate sensors put together, eventually resulting in the minimisation of the stiffening material - the silicon carbide foam in the case of PLUME. Furthermore, if sensors with the size of the VXD ladder are produced, the material budget of the flex cable will decrease significantly. Last but not least, stitched sensors are already aligned with respect to each other. So the whole ladder can be considered as one module for the alignment procedure. Alternatively, individually mounted sensors on a ladder could introduce a

degree of misalignment. Each sensor has to be considered as an independent module. The increased number of modules makes the alignment task more challenging and complex.

Despite the aforementioned benefits, the decision for the use of the industrial stitching will depend on the fabrication yield, which should be well above 90 %. Currently the yeld ranges from 60 to 90 %, values that constitute a showstopper for its use. So the current status favors the individual mounting of the sensors on the ladder. Investigation on the dicing techniques is going on, aiming at reducing the dead zones around the sensors. Depending on the outcome of these investigations, sets of 2 or 3 stitched sensors may form the optimal solution.

3.4.2 Exploration of High Resistivity Epitaxial Layers

Recently, CMOS chips featuring a high resistivity ($\gtrsim 400\Omega \cdot \mathrm{cm}$) epitaxial layer became available. By applying standard CMOS voltages to the sensing pn junction, a few μm deep depletion area is created below the diode, instead of a fraction of micron in case of low resistivity. Figure 3.10 illustrates the simulation's results of [Dorokhov 10] showing the amplitude of the depleted region created beneath the diode in a high resistivity ($1\mathrm{k}\Omega \cdot \mathrm{cm}$) epitaxial layer MAPS. The average path length of the charge carriers is shorter, resulting in a faster charge collection and enhanced CCE, as well as in a substantial improvement of the non-ionising radiation tolerance. Table 3.4 demonstrates the enhancement in CCE for MIMOSA 26 sensors fabricated in high resistivity epitaxial layers of different thicknesses compared to the standard one. This technology is presently under study. It seems very promising for improving the performance of CMOS sensors, eventually paving the way to new applications.

Epi-layer resistivity $\Omega \cdot \mathrm{cm}$	thickness μm	seed (%)	2×2 (%)	3×3 (%)
Standard ~ 10	14	21	54	71
400	10	36	85	95
400	15	31	78	91
400	20	22	57	76

Tab. 3.4: CCE for MIMOSA 26 sensors with standard and high resistivity epitaxial layers of different thicknesses for the seed pixel, 2×2 and 3×3 pixel clusters [Hu-Guo 10].

3.4.3 Design of the Innermost Layer

The major challenge for the inner double layer is the high occupancy induced by pair background hits. Occupancy higher than $1 - 2$ % is not acceptable. To cope with the

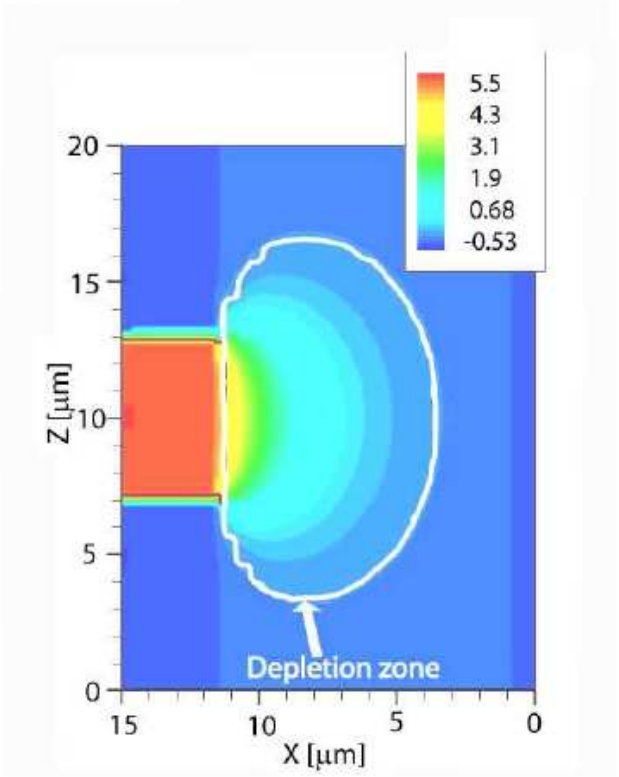


Fig. 3.10: Simulation results from [Dorokhov 10] showing the depleted region created around the diode in a high resistivity ($1\text{k}\Omega \cdot \text{cm}$) epitaxial layer MAPS.

occupancy, one has to perform a fast sensor readout ($\sim 25\,\mu$s). Of course, it is crucial to preserve the very high spatial resolution. These are somehow conflicting requirements since fine spatial resolution requires a highly segmented sensor. On the other hand, assuming a column parallel readout, high segmentation means a big number of pixels per column thus slower readout. One could suggest as a solution the replacement of the discriminators with few bit Analog to Digital Converters (ADCs), which can preserve spatial resolution while allowing for a larger pixel pitch. This approach is not viable: lets consider the two following cases. In the first case, one increases the pixel pitch, which will naturally lead the signal charge to be shared between fewer pixels. This will end up to a degradation of the spatial resolution. This will happen because in order to exploit the ADCs, charge sharing between few pixels is required. The reduced charge

sharing between pixels, will lead to reduced accuracy of the position reconstruction algorithms. And this is something we do not want to happen, especially at the inner layers of the VXD. In the second case, we assume that even though the pixel pitch is increased, the cluster multiplicity remains the same, thus the spatial resolution is preserved. In this case, we have failed in our initial motivation, the decrease of the pixel occupancy. Occupancy means the percentage of pixels firing in one event. Increasing the pixel pitch means that the total number of pixels is decreased. Then, if the cluster multiplicity remains unchanged, the percentage of pixels firing in one event increases. Therefore, even if the time resolution is improved this way, this approach will finally lead us to the same occupancy.

Moreover, we should consider other trade–offs of this approach. In general, an ADC has increased power consumption compared to a discriminator. However the increased power consumption could be compensated by the smaller number of columns per sensor that can offer an ADC - large pixels approach. Another important aspect is the material budget minimisation. The ADCs would lead to a bigger insensitive area, thus additional material budget. The innermost layer of the VXD is the part of the detector that is closest to the IP, hence the minimisation of the material is crucial for the overall detector performance. Anyway, deep submicron CMOS technology can improve significantly the speed - power dissipation - material budget performance of ADCs. Last but not least, it should be stressed that the design featuring discriminators is much less complex to implement than the one with ADCs. ADCs restricted, at one dimension, to the pixel pitch size is an innovative technique, therefore their realisation is a very challenging task from the technical point of view.

Having in mind the aforementioned considerations, a viable approach could be the following. The innermost layer could be equipped with sensors adopting the MIMOSA 26 readout architecture, providing binary output. Starting from the minimum integration time of MIMOSA 26 which is $\sim 80\,\mu s$ we can go down to $\sim 40\,\mu s$ by performing the readout from both sides of the columns. This means that the layout illustrated in figure 3.6 should be modified. An insensitive band occupied by the column level discriminators, zero suppression etc, should be added at the top of the chip. This way the length of the columns will be divided by a factor of 2. With the column parallel readout, the integration time is proportional to the number of pixels in the column, so it will be also reduced by a factor of 2. This approach will however lead to an increase of the insensitive area occupied by the readout and data sparsification circuits on the bottom and top of the sensor. Moving to deep submicron processes ($\leq 18\,\mu$m) will allow us to suppress the insensitive area by a factor of 25 to 50 %. Moreover, the smaller capacitance of the metal lines can allow us to increase the operational frequency which will also benefit the readout speed.

The time resolution performance shall be improved substantially by taking advantage of the aforementioned double sided structure of the ladder. One side will

be equipped with highly segmented square pixels aiming for high spatial resolution, while on the other side the pixels will be elongated in the column direction aiming for high time resolution (figure 3.11). Equipping the innermost side with square pixels of $\sim 15\mu$m pitch and binary readout can provide the desired $\lesssim 3\mu$m spatial resolution. A crucial step here is to implement the MIMOSA 26 pixel architecture from a pitch of $18.4\,\mu$m to $\sim 15\mu$m. Deep submicron processes can make that feasible.

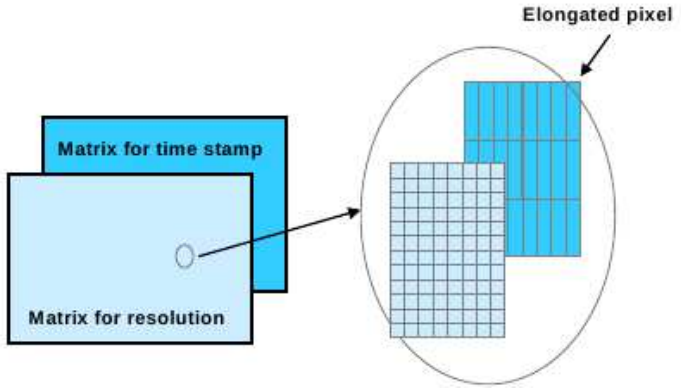


Fig. 3.11: Structure of the double sided ladder for the innermost layer.

On the other side, the sensors will have pixels elongated in the column direction. A pitch increase by a factor of 4 in this direction results in a subsequent reduction of the number of pixels per column, and thus the integration time is also reduced by the same factor. The key point of this design is that hits generated on both, $\sim 2mm$ distant, sides of a ladder by traversing particles can be correlated, as was discussed in the section 3.4.1. Thus one could timestamp the tracks. This can be crucial during reconstruction, since it is expected to suppress the combinatorial background. The sensors equipped with elongated pixels, aiming for high time resolution, are called AROM[‖].

A critical aspect of the elongated pixels design, is the CCE. In the column direction, the spacing between collecting diodes will be 4 times bigger than in the row direction. The danger is that a significant part of the charge will be lost due to recombination. This point has started to be attentively studied in beam tests, with prototype sensors featuring elongated pixels. We are going to quote results obtained with the MIMOSA 22–AHR sensor, which has a submatrix equipped with elongated pixels with pitch dimensions of $18.4\,\mu\text{m} \times 73.2\,\mu$m. This sensor was manufactured with the AMS $0.35\,\mu$m process, featuring a relatively high resistivity ($400\,\Omega \cdot$ cm) epitaxial layer of $15\,\mu$m thickness. The obtained MPV of the Landau distribution of the SNR was 29.6 ± 0.3.

[‖]Stands for Accelerated Read–Out Mimosa sensors

This SNR value can ensure us that the potential loss due to recombination of the charge carriers generated by an impinging particle is not significant. For the next beam tests campaigns, it is planned to study sensors with elongated pixels having more than one collecting diode, as well as sensors where the diodes are placed in a staggered geometry.

The MIMOSA 30, fabricated during the last quarter of 2011, is a prototype sensor aiming to demonstrate the feasibility of the realisation of the VXD inner layer sensors. It was manufactured in the AMS 0.35 μm process, featuring a high resistivity epitaxial layer. It consists of two parts; the first, the high spatial resolution one, is equipped with square pixels of 16 μm pitch; the second part, the timestamping one, is equipped with elongated pixels of $16 \times 64\,\mu\text{m}^2$. The CDS takes place in–pixel. The pixel matrices are read with a double sided column parallel readout. Each column ends with a discriminator, providing this way digitised output. The expected performance for the first part of MIMOSA 30 is a spatial resolution of $< 3\,\mu$m and a readout time of $\lesssim 50\,\mu$s; while for the second part, the expected spatial and time resolution are $\sim 6\,\mu$m and $\sim 10\,\mu$s respectively. The sensor is going to be tested in beam (CERN–SPS) during the summer of 2012.

3.4.4 Design of the Outer Layers

The pair background decreases exponentially as a function of the distance from the IP. The occupancy of the outer layers of the vertex detector is expected to be lower, compared to the innermost one, by 1 and 2 orders of magnitude for the second and the third layer respectively. This means that the requirements on readout time are less severe. A sensor integration time of 100 μs is considered adequate. Here, the main goal of the optimisation is to decrease the power dissipation. This is motivated by the relatively large surface to equip. Power dissipation is proportional to the number of columns of the sensor. It is not affected by the number of pixels in each column. Hence a possibility would be to decrease the number of columns by increasing the pixel pitch. Of course, the high spatial resolution should be maintained. A pixel pitch of 35 μm equipped with a 3–4 bits ADC per column can provide a spatial resolution $\sim 3 - 4\,\mu$m which satisfies the vertex detector requirements. This value for the resolution is an assumption, stemming from extrapolated results from tests of different sensors, see table 3.5. The fabrication and test of a sensor with such features is under way.

The power consumption of one column of MIMOSA 26 is 520 μW. If the column ends with a 3–4 bits ADC the estimated consumption is expected to stay below 1 mW. Considering the size of the pixel matrices of the sensors for both the inner and the outer layers, and taking into account their peripheral circuitry, we can calculate the instantaneous power dissipation of the whole VXD. It would amount to $\lesssim$ 900 W. These calculations are based on sensors fabricated in 0.35 μm process. Moving to a smaller feature size technology ($\leq 0.18\,\mu$m), is expected to further decrease the power consumption due to the lower voltages used. Table 3.6 summarises the estimated

Pitch (μm)	20	20	30	35	40
Number of bits	12	4	12	4	12
epi–layer	low–res	low–res	low–res	high–res	low–res
$\sigma_{s.p}(\mu\text{m})$	1.5	1.7	2.1	≤ 4 (extrapolated)	3

Tab. 3.5: Measured and calculated single point resolution as a function of the pixel pitch, the ADC encoding and the resistivity of the epitaxial layer.

instantaneous power dissipation of the whole VXD and each layer separately, assuming it is equipped with sensors fabricated in 0.18 μm technology. The full VXD power dissipation is estimated to be $\lesssim 700$ W.

The ILC beam structure (see section 2.1.2) allows to adopt the power cycling concept in order to reduce the average power consumption of the VXD. The pixels can be switched off at the time interval between two bunch trains. Concerning the elements of the peripheral circuitry, all of them can be switched off, with the exceptions of the discriminators/ADCs and the bias DAC (Digital to Analog Converter). Assuming a duty cycle of $\leq 2\,\%$, the estimated average power consumption of the full VXD will be $\lesssim 15$ W.

The estimated values for the power dissipation allow for an air flow cooling with low circulation speed. The benefits for the VXD design arising from this fact are mostly the two following: first, the complexity and additional material budget of a liquid cooling system is avoided. Second, the air flow could cause vibrations of the ladders, which are suspectible to degrade the spatial resolution. The lower the intensity of the air flow, the smaller the vibrations that the ladders have to withstand.

A prototype sensor, aiming to equip the outer layers of the VXD, is MIMOSA 31. It was fabricated during the last quarter of 2011. It consists of 48 columns of 64 pixels, having a pitch of $35 \times 35\,\mu\text{m}^2$. Each column ends with a 4–bits ADC. The expected spatial resolution is $\sim 3.5\,\mu$m. The size of the prototype chip is the 1/10 of the size of the sensor foreseen (maybe) to equip the VXD. The sensor is planned to be exposed to the DESY e^- beam during the first quarter of 2013.

3.4.5 Summary

In this chapter we have addressed the possibility to realise a vertex detector for the ILD based on planar CMOS sensors. The physics goals of the experiment call for a VXD equipped with ultra light and highly segmented sensors, while its running conditions require a swift readout for the inner layers and a moderate radiation tolerance. CMOS sensors, with their genuinely very small feature size and very low material budget, naturally become a very promising candidate. The main challenge they are facing is to realise a fast readout for the inner layers, while preserving their other performance.

The optimisation procedure is not an easy task, since the aforementioned requirements are often conflicting: sensors with small pixel pitch offer optimal performance in spatial resolution and radiation tolerance, while by enlarging the pitch, one can improve the time resolution and reduce the power dissipation.

The above statement, combined with the substantially different running constraints on the inner and the outer layers of the VXD, led us to follow a separate optimisation procedure for the inner and the outer layers. For the inner layers, the main goal is to reach a time resolution of $O(10\,\mu\text{s})$, while the spatial resolution should by any means remain $\lesssim 3\,\mu$m. The proposed solution is taking profit of the double sided structure of the ladders; the first side is optimised for high spatial resolution; on the second side, the spatial resolution in one coordinate is sacrificed, in order to time stamp the hit. An intensive R & D effort to study this approach has already started, and even with the present status, a spatial resolution of $\lesssim 3\,\mu$m and a time resolution of $\lesssim 10\,\mu$s seem within reach. Concerning the outer layers, the target is the minimisation of the power dissipation. This can be accomplished using sensors with a large, $\sim 35\,\mu$m, pixel pitch. Replacing the discriminators with few bits ADCs, a spatial resolution $\sim 3.5\,\mu$m can be achieved. The feedback from the test beam campaigns of MIMOSA 30 and 31 sensors will be valuable in our efforts to design a VXD for the ILC, based on CMOS technology.

The table 3.6 summarises the required sensor performance for the CMOS based VXD. The sensor performance may be improved substantially in the future, benefiting from the progress of the industrial CMOS processes (smaller feature size). Finally, the importance of the integration techniques in preserving, or even improving the sensor performance, should be underlined. As a concluding remark, we should stress that moving from planar CMOS technologies to 3D integration techniques may bring striking improvements to the overall sensor performance.

Layer	Radius (mm)	spatial resolution (μm)	Output	time resolution (μs)	P_{inst} (W)
DL1	16-18	≤ 3	binary	10	~ 300
DL2	37-39	$\sim 3-4$	3–4 bits ADC	100	~ 150
DL3	58-60	$\sim 3-4$	3–4 bits ADC	100	~ 250
Total					$\lesssim 700$

Tab. 3.6: Sensor performance requirements of a CMOS based VXD adopting the double layers geometry. Sensors assumed to be fabricated in a $0.18\,\mu$m process.

CHAPTER 4

Test Beam Analyses of the MIMOSA 22bis and MIMOSA 24 Prototype Sensors

Since 1998, when the idea to use CMOS sensors for high precision tracking detectors emerged at IPHC Strasbourg, numerous prototypes of MIMOSA pixel sensors have been fabricated and tested. The main goal of the first MIMOSA sensors was the establishment of the proof of principle that the CMOS technology could be used for HEP applications. After this primary goal was achieved, the optimisation procedure was initiated. Several other prototypes have been realised and tested, exploring different CMOS fabrication processes and sensor architectures.

A series of tests are conducted with each prototype, in order to evaluate its performance. The testing of the prototypes includes tests with light sources, radioactive sources and high energy particle beams. Since the goal is to develop high precision trackers, the measurement of the tracking performance of the prototype is a crucial step. In order to make these measurements, the sensor is exposed to a high energy particle beam. This procedure is called test beam, and usually takes place at an accelerator facility which provides the particle beam.

The procedure of a typical CMOS test beam is the following: the prototype, which is called the DUT (Device Under Test), is mounted on a beam telescope, which reconstructs the beam particle tracks. The beam telescope is a detector formed by several planes of well tested silicon sensors. It provides a very accurate track interpolation on the DUT surface, and then matched with the hits on the DUT. The matching procedure allows for the extraction of the DUT spatial resolution, the tracking efficiency and the fake hit rate. Other important goals of a test beam is to measure the charge collection on the seed pixel and the cluster, the SNR (see sub–section 3.2.2) and the cluster

multiplicity. The above concepts will be introduced in the following sections. In this chapter, we are going to analyse the test beam data of the MIMOSA 22bis (section 4.3) and MIMOSA 24 (section 4.4) sensors. MIMOSA 22bis is a real scale ($\sim 79 \times 10^3$ pixels) swift readout digital sensor. Its main purpose was the development of a proper amplification scheme working together with in pixel CDS (see sub–section 3.2.3), exhibiting low noise and being radiation tolerant, but as well to validate the column parallel readout architecture with integrated discriminators. MIMOSA 24 is an analog sensor. Its main goal was the evaluation of a new fabrication process, XFAB 0.35 μm. But first we will proceed to a brief description of the experimental setup and the data analysis framework.

4.1 Experimental Setup

4.1.1 SPS Test Beam Facility

The test beam campaigns examined in this thesis, took place at the North Area test beam facility, at CERN. The SPS proton beam of 450 GeV is directed towards the three targets of the North Area facility. The produced secondary beams can be composed of electrons, muons or pions, and their momentum range is $10 - 400$ GeV/c. Our test beams were performed with a 120 GeV/c π^- beam. The spill duration was 8 s and its period was either 33.6 s or 48 s.

4.1.2 Beam Telescopes

In the beam tests described in this thesis two different beam telescopes have been used:

Strip Telescope

The strip telescope [Colledani 97] is formed of eight reference silicon strip planes, arranged in two arms of four planes each. The DUT is placed between these arms. The area of the planes is $12.8 \times 12.8\,\mathrm{mm}^2$ and the thickness is 300 μm of silicon. Each plane has 256 strips of 50 μm pitch. The four planes of each arm form two modules, each one made of two neighbouring planes placed 2 mm apart. The strips of the two planes belonging to the same module are oriented orthogonally with respect to each other, see figure 4.1. The provided spatial resolution on the DUT surface is $\lesssim 2\,\mu$m. The way to calculate the spatial resolution of a beam telescope, with respect to the position and single point resolution of each plane, is well explained in [Meier 99]. Two plastic scintillators are placed perpendicular to the direction of the beam, one in front of the telescope and the other behind, covering areas of $2 \times 4\,\mathrm{mm}^2$ and $7 \times 7\,\mathrm{mm}^2$ respectively. The trigger is provided each time a coincidence signal between the two scintillators is

recorded. The location of the scintillators should be such, that their "shadow" covers the part of the DUT we want to test.

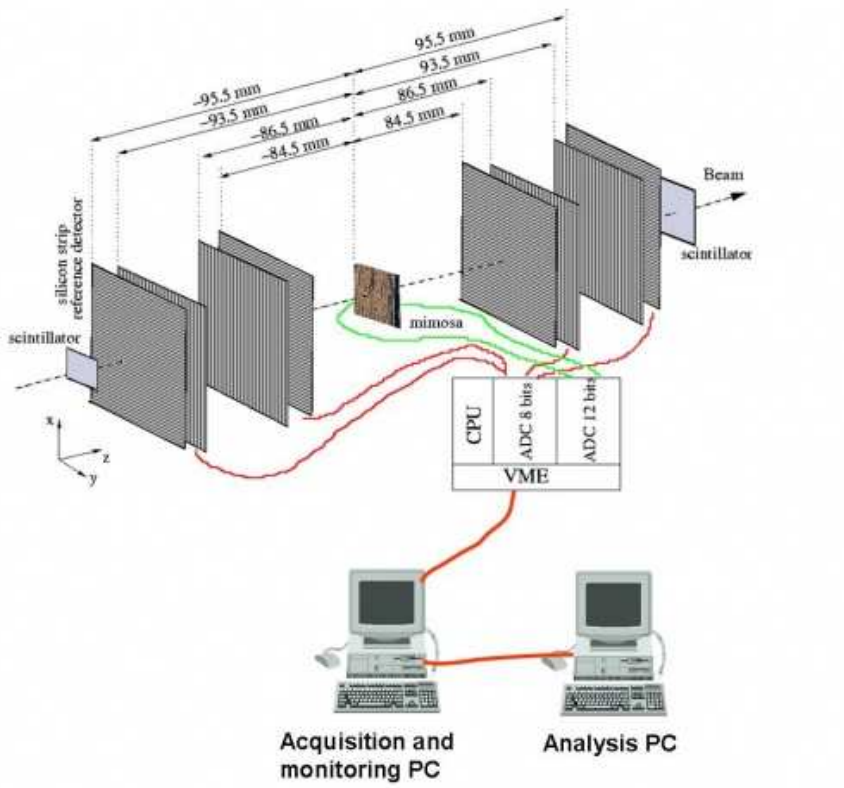


Fig. 4.1: Schematic representation of the silicon strip telescope and its DAQ system.

Pixel Telescope

The TAPI (Télescope à Pixel de l' IPHC) is a general purpose, portable pixel beam telescope, designed and constructed at IPHC, Strasbourg [Dulinski 07]. It is formed from four reference planes, arranged in two arms of two planes, equipped with the MIMOSA 18 pixel sensor. The DUT is placed between the two arms. The MIMOSA 18 was fabricated using the AMS 0.35 OPTO process. It has four submatrices with 256×256 pixels of $10\,\mu$m pitch, thus it provides an active area of $5 \times 5\,\text{mm}^2$. The reference planes are contained inside aluminium boxes. An external cooling system keeps their temperature stable, at 16^0 C. The signal coincidence of two scintillators, placed along the beam axis before and after the telescope, provides the trigger. The interpolated spatial resolution on the DUT surface is $\leq 1\,\mu$m.

4.1.3 Electronics

In order to facilitate and systemise the testing procedure of the various MIMOSA prototypes, a modular and flexible test setup has been developed at IPHC. It is composed of three modules: the proximity board, the interface card and the data acquisition card.

The proximity board is the PCB (Printed Circuit Board) where the prototype sensor is wire bonded. It hosts the first stage external amplifiers and generates the clock signal that sequences the chip operation. For each prototype, a specific proximity board is required. The rest of the test setup can be used with any prototype sensor under test. It may change according to which beam telescope is used (see sub–section 4.1.2). The interface card, as its name suggests, is the card used as the interface between the sensor and the Data Acquisition System (DAQ). It transfers the data from the sensor to the DAQ, and provides bi–directional transmissions of the control signals between them. It also provides the power supply and the reference voltages for the amplifiers on the sensor. The interface card allows for the integration of the minimal required logic on the proximity card. This introduce a high degree of flexibility in the testing procedure; one can take the prototype sensor wire bonded on its PCB, and test it with his own setup. A second motivation to minimise the several functionalities integrated on the proximity board is to reduce its power consumption, thus the heat on its surface.

Two different DAQ systems have been used for the tests described in this thesis: one based on the VME standard [Claus 04] for the test beam of the MIMOSA 22bis, and one based on a home made board, called TNT [Arnold 05] for the MIMOSA 24. They will be described in the following section.

Data Acquisition System

The MIMOSA 22/22bis (see section 4.3.2) provides both analog and digital output. The analog data provided by each of the 8 columns ending without a discriminator, are fed to an ADC on the DAQ (or imager) card. Each card has four ADCs and four buffers, used to perform the CDS, thus two cards are required. Concerning the digital part, the output of the 128 columns ending with a discriminator is multiplexed in 16 lines. The imager card features 16 inputs for digital signals, thus one card is enough. The CDS of the digital part is taking place in–pixel, so there is no need for such a functionality on the imager card. Therefore, three imager cards are required in total for the DAQ of MIMOSA 22/22bis. The VME protocol is used for power supply. The data are sent to the DAQ PC, where they are merged via a USB connection. The user often wants to access the DAQ board. For this reason it is not convenient to be placed in the beam area. Hence it is located in the control room, $\sim 40\,\mathrm{m}$ from the beam area.

MIMOSA 24 was tested with the TAPI–TNT2 system. The TNT2 board makes a loop over all the pixels of the sensor, and stores the analog signal of each one at a corresponding memory cell. When it proceeds to the next iteration, it first compares the previous value of the signal of each pixel with the new one. In other words, it performs an on–line CDS. If the CDS gives a value higher than a specified threshold, the information is addressed for readout. The data sparsification is performed this way. In the next step, TNT substitutes the old with the new analog signal of the pixel.

The readout takes place only inside the time window after the arrival of a trigger signal. This time window is equal to the integration time of the sensor. For example, if the trigger arrives when the TNT addresses the pixel i, and for the time interval until the TNT makes a full loop on the pixels and reach again the pixel i, if there is any pixel whose CDS result exceeds a specified threshold, the information goes to the readout. If a new trigger arrives during this time interval, the time window is extended accordingly. The data are transferred to the PC via a USB connection.

4.2 Data Analysis

4.2.1 Software

The main tasks of the test beam data analysis software is to reconstruct the hit clusters and the tracks, to generate the η functions (see sub–section 4.2.3) and to perform the alignment, both of the telescope planes and the DUT. Then, it provides the framework to perform the offline analysis and to characterise the DUT. Two different software packages, both developed at IPHC, have been used for the data analysis. The MIMOSA 22bis sensor was analysed with the MIMOSA Analysis Framework (MAF) software, while MIMOSA 24 with the TAPI Analysis Framework (TAF). Both are using the C++ language and are based on the ROOT analysis software [ROOT]. MAF has

been designed specifically to analyse data coming from prototypes tested with the silicon strip telescope. It is very well tested, being used for more than ten years. The TAF software, designed for the TAPI telescope, is based on MAF, sharing in a large extent the same code. It is a very flexible software, that can be applied in any configuration.

4.2.2 Estimation of Pedestal, Noise and Common Mode Shift

The first step of the data processing is to perform the CDS by subtracting two consecutive frames. After the CDS, we obtain the image of the sensor, which contains the raw signal for each pixel (or strip if we refer to the reference planes of the strip beam telescope. In this section, the term pixel will be used referring to both). The value of the raw signal, $r_k(n)$, at the pixel k in the event n, is a combination of several contributions, described by the following equation:

$$r_k(n) = s_k^{phy}(n) + q_k^{ns}(n) + p_k(n) + c(n) \tag{4.1}$$

where $s_k^{phy}(n)$ is the signal generated by a possible traversing particle, $q_k^{ns}(n)$ is the random noise, $p_k(n)$ is the pedestal and c(n) the common mode noise. The value we seek after is the physics signal, $s_k^{phy}(n)$. In order to extract it from the raw data, all the other contributions should be calculated. When there is no physics signal $s_k^{phy}(n)$, the value of the pedestal is calculated by

$$p_k(n) = \frac{1}{N}\sum_{n=1}^{N}(p_k(n) + c(n) + q_k^{ns}(n)) \tag{4.2}$$

where the random and common mode noise are averaging to 0. In the presence of a signal, the true value of the pedestal is:

$$p_k(n) = \frac{1}{N}\sum_{n=1}^{N}(r_k(n) - s_k^{phy}(n)) \tag{4.3}$$

Practically, the equation 4.3 cannot be used because the value of $s_k^{phy}(n)$ is not known (it is actually what we search for). If we simply average over the raw signals, the measurement will be biased due to the presence of the $s_k^{phy}(n)$. A possible solution is to measure the pedestal when no particles are present, thus no physics signals occur. This approach was not always possible during test beam campaigns, since for certain test setups a trigger signal (thus a traversing particle) was required in order to record the data from the sensor. Using another configuration that does not require the passage of a particle to take the raw data, does not yield reliable results. The reason is that the pedestal is sensitive to aspects like the configuration, the DAQ system and also to environmental factors. It should be measured with the same configuration and

environmental conditions used for data taking. Thus, an approach that will offer a good approximation to the equation 4.3 should be used instead. Due to some substantial differences between the test setups used for MIMOSA 22bis and MIMOSA 24, different approaches have been used for the pedestal estimation.

The test setup used for the MIMOSA 22bis test beam, required a trigger arising from the passage of a real particle. Thus, the pedestal has to be measured with the beam on. The first step is to estimate an initial value for the pedestal for each pixel, using the first few hundred events, through the summation over the raw signal. A sampling method is used to suppress the contribution on the raw signal from the $s_k^{phy}(n)$ term: the raw signal values of the first 200 events are grouped in samples of 5 events, for each pixel. Out of these samples, the events featuring the extreme values of the raw signal are rejected. The reason is that if the signal coming from a particle contributes to the raw signal in an event, it is expected to acquire a higher value. Thus the rejection of the maximum value of the 5 event group, ensures that the events where the $r_k(n)$ value contains also the $s_k^{phy}(n)$ term, will not be accounted for the pedestal estimation. Now, one may ask what if a second particle hits the same pixel, during this 5 event time interval. The particle intensity during the test beam was rather low, making such a case very improbable. The next obvious question would be, since the beam intensity is so low, how can we be so sure that the maximum raw signal value of the 5 event group contains the contribution of a physics signal. Actually, we cannot be sure. So, in order not to bias our measurement by cutting out the highest values of the true pedestal for each pixel, the minimum raw signal of the group is also rejected. The initial estimation of the pedestal is calculated through the summation over the leftover values of the raw signals. The pedestal value is continuously updated during the data taking, after the 200 initial events, according to the following recursive formula:

$$p_k(n) = \frac{1}{A}[(A-1)p_k(n-1) + r_k^{3\sigma}(n) - c(n)] \tag{4.4}$$

where A is a weighting value equal to 10, selected in order to reduce the sensitivity to the fluctuations and $r_k^{3\sigma}(n)$ is the physics signal suppressed raw value. The latter maens that raw signal values being $> 3\,\sigma$ far from the previous pedestal value are not accounted for the pedestal update. Obviously because such a difference in the raw value of one pixel between two consecutive events is very probable to originate by the signal of a crossing particle.

The TNT DAQ system provides sparsified data, thus we cannot apply the same method for MIMOSA 24. On the other hand, the TNT can accept a fake trigger, therefore the pedestal can be measured from raw data when the beam is off. Small data taking runs, with the beam switched off, were performed with MIMOSA 24 in order to measure the pedestal and the noise. Due to the data sparsification, a recursive formula that will update the noise and pedestal values during the data taking, cannot be applied here.

After the pedestal is measured, the noise is extracted from the standard deviation of the pedestal:

$$q_k^{ns}(N) = \sqrt{\frac{1}{N-1}[\sum_{n=1}^{N} r_k(n)^2 - N\overline{p}(N)^2]} \tag{4.5}$$

The noise is updated during the data taking, according to the following recursive formula:

$$q_k^{ns}(n) = \sqrt{\frac{1}{B}[(B-1)(q_k^{ns}(n-1))^2 + (r_k^{3\sigma}(n) - p_k(n) - c(n))^2} \tag{4.6}$$

where B is a weighting value equal to 10.

Some variations of the response of all the pixels of a sensor can be observed from frame to frame. This is called the common mode shift. It is estimated according to the formula:

$$c(n) = \frac{1}{K}\sum_{k=1}^{K}(r_k^{3\sigma}(n) - p_k(n)) \tag{4.7}$$

This calculation is very time consuming. To speed up the procedure, it is restricted to a region of the sensor that contains K pixels. After the estimation of the pedestal and the common mode shift, the physics signal can be calculated from the equation 4.1.

$$s_k^{phy}(n) = r_k(n) - p_k(n) - c(n) \tag{4.8}$$

The physics signal contains the contribution coming from the traversing particle, plus the random noise.

4.2.3 Hit Reconstruction

The impact of a particle traversing a sensor is called a hit. Due to the charge sharing between neighbouring pixels, the traversing particle generates a cluster of firing pixels. The pixel of the cluster with the highest signal is called the seed pixel. The hit identification and reconstruction procedures differ in some details, depending on whether we have analog, digital or sparsified data (TNT). In general, the first step of the hit reconstruction is to identify the seed pixel. The seed candidate is accepted, if its SNR is higher than the defined threshold. The threshold is used in order to suppress the fake hits due to noise fluctuations. There is not a general optimal threshold; one has to find the suitable one, according to the sensor specifications and the running conditions. The noise value used in the SNR computation, is the one the recursive formula 4.6 gives, for the current event. In case of sparsified analog data, we cannot apply the equation 4.6.

Thus, the threshold for the seed selection is applied to the signal value of the candidate pixel.

If the candidate seed is accepted, we try to reconstruct a cluster around it. The requirement for the cluster suppresses effectively fake hits, caused by large noise fluctuations of a pixel. Thus a second cut, on the total SNR of the neighbouring pixels that form the cluster is applied. On the other hand, we have to mention that with the cluster requirement all hits having a single pixel cluster are going to be rejected. Hence a few per–cent loss in detection efficiency is expected. The decision whether to apply the cluster selection criteria or not should be dependent on the cluster's pixel multiplicity. The cluster reconstruction can be realised with two alternative ways: we either accept all the neighbouring pixels that satisfy the applied selection criteria, or require the cluster to be consisted of a predefined number of pixels. After the first hit cluster is reconstructed, the algorithm searches for other seed candidates. Obviously, pixels that are already part of a cluster, cannot belong to another one.

In case of a digital output sensor, like the main part of MIMOSA 22bis, each pixel with signal value that pass the discriminator threshold becomes automatically a seed candidate. Obviously the pixels cannot be sorted out through their signal value, because it is equal to 1 for all of them. The seed is defined as the pixel with the highest number of direct neighbour firing pixels. As direct neighbours are defined pixels that share one border.

Once the hit cluster has been reconstructed, we want to find the hit position, with coordinates (u,v). To do so, we use the following positioning algorithms. The accuracy of each method, is quantified from the residual $u_h - u_t$ ($\sigma_{residual}$), where u_h is the reconstructed position of the hit given by the algorithm, and u_t is the interpolated position of the track on the surface of the detector plane. To measure the resolution of the detector, we need to take into account the track interpolation accuracy $\sigma_{interpolation}$, apart from the residual. In order to disentangle the track interpolation accuracy of the telescope from the spatial resolution of the sensor, the latter is calculated using the following equation:

$$\sigma_{sensor} = \sqrt{\sigma^2_{residual} - \sigma^2_{interpolation}} \tag{4.9}$$

Obviously, the spatial resolution depends both on the sensor single point resolution and the positioning algorithm. In the following sections, we will describe briefly the hit position reconstruction algorithms. For more information one can refer to [Meier 99].

Digital Method

The digital positioning algorithm reconstructs the hit position at the center of the seed pixel. It does not use any information coming from the cluster. The provided resolution is $\frac{P}{\sqrt{12}}$, where P is the pixel pitch.

Center of Gravity

The Center of Gravity (CoG) algorithm takes advantage of the charge spread among the cluster pixels, in order to refine the calculation of the hit position. Assuming a reference axis where the center of the seed pixel is at 0, the pixels to the left have negative coordinates, and the ones to the right positive, the hit position is calculated as:

$$u_h = \frac{1}{Q_{tot}} \sum_{n \epsilon k} u_k \times q_k \tag{4.10}$$

where the summation goes over the pixels belonging to the cluster. Q_{tot} is the total charge of the cluster, u_k the coordinate of the center of the k^{th} pixel and q_k its charge. The CoG algorithm assumes that the charge is distributed with equal probability among the pixels. This assumption is not true, thus it introduces a systematic error in the calculation of the hit position.

The η Method

Even if the beam illuminates uniformly a detector's surface, the charge is not distributed linearly between neighbouring pixels or strips. This can become clear if we plot the η distribution. The definition of the η is given by the following equation;

$$\eta = \frac{s_L}{s_L + s_R} \tag{4.11}$$

where $s_{L(R)}$ is the signal collected on the left (right) pixel. If the charge was distributing linearly, η would have a flat distribution. On the contrary, the η distribution has the shape shown on the left side of figure 4.2, built from data coming from one plane of the strip beam telescope, which is far from being linear. An explanation for this effect can be found in [Turchetta 93]. The η distribution can be built from the data, using the first few hundreds (or thousands) events of a series of frames recorded (i.e. run).

The right part of the figure 4.2 shows the integral of the η distribution over N events, for the same plane. This histogram is used to remove the bias of the hit's position reconstruction. The correction procedure is the following: one takes the difference of the center of the seed pixel with the position given from the CoG algorithm, $U_{CoG} - U_{DIG}$, normalised to the pixel's pitch. One then can take take this number, which will be between 0 and 1, go to the histogram of the normalised integrated η function (right part of the figure 4.2), and find the bin which corresponds to it. Finally, corrects the digital hit position with the value of integrated η function for this specific bin. An issue of this approach is that the U_{CoG} can be reconstructed outside the seed pixel, thus the aforementioned procedure does not make sense. However, this happens very rare (probability $\sim 10^{-3}$).

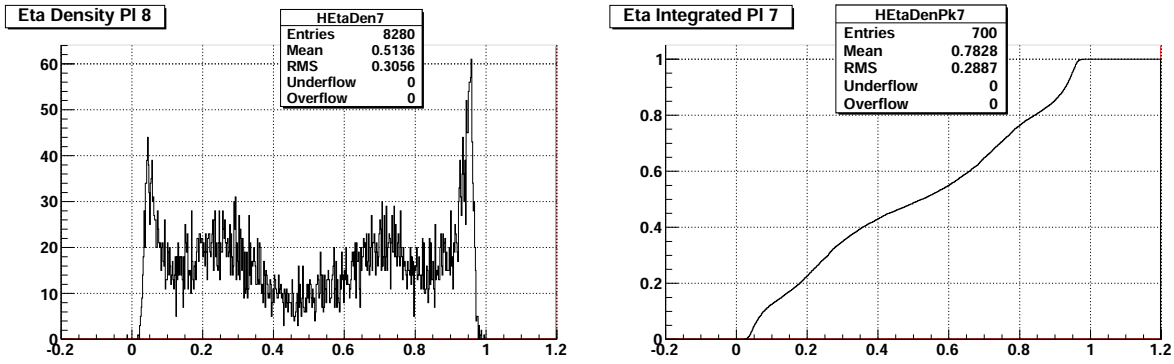


Fig. 4.2: Left part: the η distribution built on a reference plane of the silicon strip beam telescope. Right part: the normalised integral of the η distribution over 700 events for the same plane.

The η method provides the most accurate hit position reconstruction. This can be demonstrated by comparing the resolution provided by the different positioning algorithms. Figure 4.3 illustrates the hit–track residuals for all the above positioning algorithms, derived from the test beam data analysis of the analog part of the submatrix S5 of the MIMOSA 22bis sensor (see sub–section 4.3.4). The smallest residuals, $\sim 2\,\mu$m, are provided by the η method.

4.2.4 Alignment

The purpose of the alignment procedure is to identify and correct the possible relative translations or rotations between the telescope planes and/or the DUT. The procedure proceeds through two separate steps. First the telescope is aligned, and next, the DUT is aligned with respect to the beam telescope. Both telescopes follow the same strategy for their alignment.

Telescope Alignment

Regarding the strip telescope, one module in each arm is considered as fixed. The same principle is applied for TAPI, with the only difference that here is one plane at each arm. These planes, called the primary planes, define a reference system. The rest of the telescope planes, called secondaries, are aligned with respect to that reference system. The track is reconstructed from hits in the primary planes. Then it is interpolated to the secondary planes. The possible rotations and translations are identified by plotting the residuals $r_u = u_h - u_t$, where u_h is the reconstructed hit position on the secondary plane and u_t is the interpolated track position, versus the u_t. The MAF software can

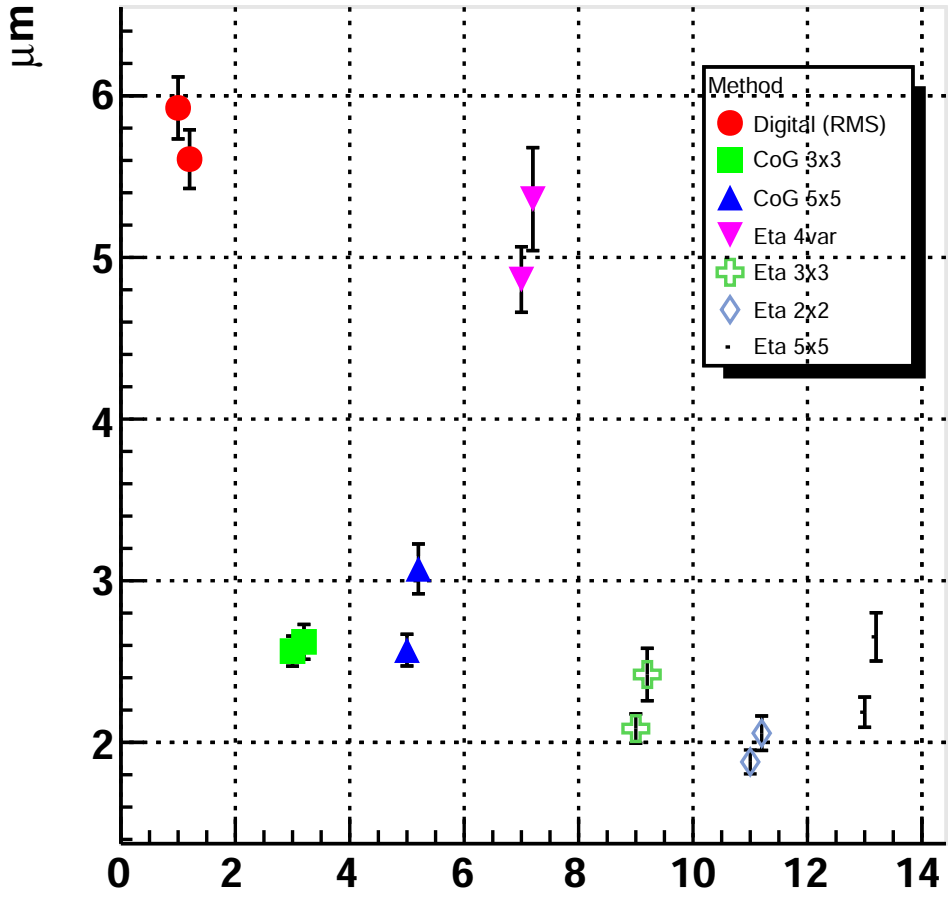


Fig. 4.3: Track hit residuals, in u and v coordinates, obtained with several positioning algorithms. The algorithms used are reported in the legend. The CoG stands for the Center of Gravity while the Eta for the η algorithm. Next to the algorithm type is mentioned the size of the used cluster. Results are derived from the test beam data analysis of the analog part of the submatrix S5 of the MIMOSA 22bis sensor (see sub–section 4.3.2). The horizontal axis, stands for the code number of each algorithm.

handle two translations perpendicular to the beam axis, and one rotation with respect to the beam axis. It was used for the MIMOSA 22bis analysis. The TAF software, used for the MIMOSA 24 analysis, can cope with any of the three translations and three rotations.

One can perform a linear fit to the residuals r versus the v_t coordinate of the track interpolation on the plane's surface:

$$r = u_{off} + m \cdot v_t \tag{4.12}$$

where m is the slope, being equal to the tangent of the rotation angle with respect to the beam axis, and u_{off} is the offset. For a pixelated telescope, like TAPI, the alignment of a plane could be performed on both directions. Thus, the linear fit is performed also for the residuals $r_v = v_h - v_t$ versus the u_t coordinate. The fit parameters are computed with a least squares minimisation. The alignment proceeds in an semi–automatic way. In the beginning, we do not have any idea about the degree of the misalignment of the plane. Thus the boundaries on the track–hit distances that enter in the calculation of the residual r for the equation 4.12 include all the surface of the plane. We proceed to the initial estimation of the alignment parameters, and align the telescope according to them. The procedure is repeated. At each iteration the track–hit boundaries are decreasing. The iterations stop when the shift in the computation of the translation parameters become smaller than a defined limit, equal to 0.5 μm for our study. The procedure of telescope alignment was repeated for each new run.

DUT Alignment

The alignment of the DUT follows the alignment of the beam telescope. The strategy remains the same, the minimisation of the track–hit residuals. Usually, prototype sensors have various pixel submatrices featuring different designs, dimensions etc. The DUT alignment should be performed separately for each submatrix during the offline analysis. The three possible rotations and translations of the DUT are taken into account, so six alignment parameters have to be calculated in total. They are calculated numerically, using the MINUIT software package [MINUIT].

4.3 MIMOSA 22/22bis

4.3.1 Objectives of the MIMOSA 22/22bis Prototype Sensors

The MIMOSA 22 and MIMOSA 22bis sensors were the final prototypes of the development program for the upstream section of the signal processing architecture of MIMOSA 26. It encompasses the signal collection, pre–amplification and discrimination. The importance of MIMOSA 26 as a milestone of the CMOS sensors development, being the first real size digital sensor used in HEP applications, is underlined in section 3.3. MIMOSA 22 was designed and fabricated in the second half of 2007. Its main goal was to test the column parallel readout mode, with digital output.

This readout scheme was first tested with MIMOSA 16 [Degerli 09]. MIMOSA 16 is a small sensor, which features $\sim$ 3000 pixels. The main question that MIMOSA 22 had to address, was whether the performance of MIMOSA 16 could be obtained when this readout scheme is implemented in a real scale sensor, having $\sim 10^5$ pixels. A crucial aspect here is to check the uniformity of the discriminators response. The second main goal of MIMOSA 22 was the optimisation of the pixel design. The signal

processing of MIMOSA 16, integrated in a pixel pitch of 25 μm, had to be adapted in a pitch of 18.4 μm. The main motivation to decrease the pixel dimensions is to obtain the required spatial resolution of the EUDET telescope. Of similar importance was the optimisation of the sensing diode dimensions. MIMOSA 22/22bis sensors were also designed to offer improved chip testability, compliant with the EUDET beam telescope requirements. A JTAG steering logic has been implemented in the chip. It is used to define the functional modes of the chip, as well as to set the common threshold for the discriminators.

MIMOSA 22bis was designed and fabricated in the first half of 2008. The main aim was the development of a proper amplification scheme working together with in pixel CDS, exhibiting low noise and tolerance to radiation. A second motivation was to fully validate the sensor design, which means to demonstrate the ability that the same architecture can be fabricated in two submissions and exhibit similar performance. It also offered the possibility to conduct further studies concerning the pixel architecture optimisation, and to implement and test radiation tolerant structures.

4.3.2 Chip Description - Architecture

MIMOSA 22/ 22bis has been fabricated in the AMS 0.35 μm OPTO process, which features a 14 μm thick epitaxial layer. It has 576 rows and 136 columns. 128 of the columns end with a discriminator, featuring a common adjustable threshold, that provide binary output. The remaining 8 columns provide analog output for pixel characterisation. There are 17 and 14 pixel submatrices on MIMOSA 22 and MIMOSA 22bis respectively.

Several aspects of the pixel design are tested in each submatrix. These are the sensing diode dimensions, the amplification architecture and the implementation of radiation tolerant structures. The pixel pitch is 18.4 μm. The nominal clock frequency is 100 MHz. Each pixel readout needs 16 clock cycles, which gives an integration time of 92.5 μs. The layout of the sensor is illustrated in figure 4.4, where the analog and digital parts are pointed out.

Three different amplification schemes have been implemented in the sensor. The first features a reset diode for leakage current compensation (3T like design, see sub–section 3.2.2), with common source amplifier (see figure 4.5, a). A variation of this design exists, (figure 4.5, b), where a capacitive load based on two NMOS transistors provides about 10 % higher gain. The second amplification scheme is a self–biased diode (see sub–section 3.2.2) with a Common Source (CS) amplifier (see figure 4.5, d). It also uses a load based on two NMOS transistors, in order to maximise the amplifiers gain (Enhanced Common Source–ECS). This is the reference pixel design. The third one, (figure 4.5, c), tries to combine the reset with the self–biased diode designs. All pixel designs include in–pixel CDS. The above amplification schemes are combined with various diode dimensions.

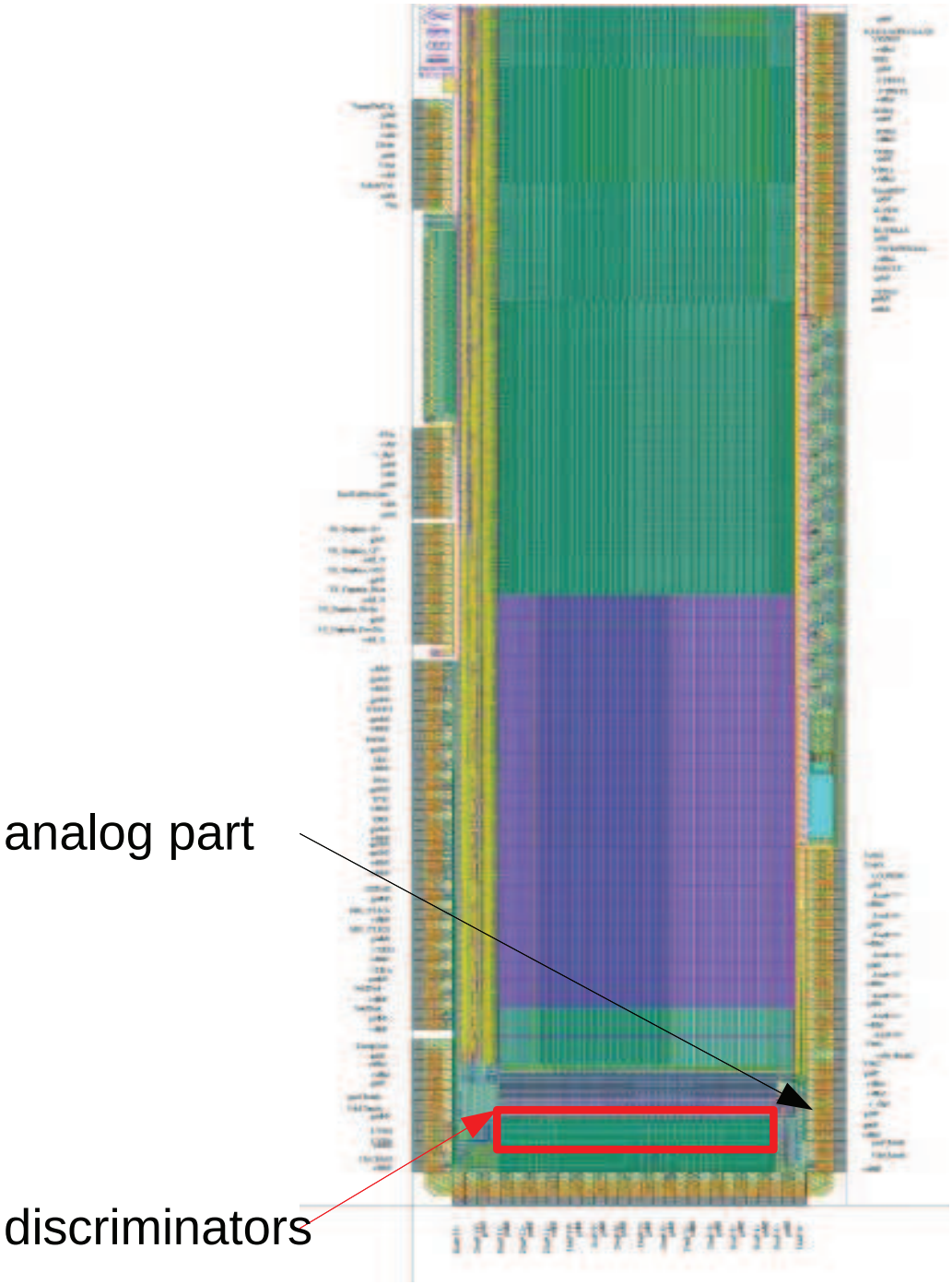


FIG. 4.4: Layout of the MIMOSA 22 sensor. The rolling shutter is performed from the top to the bottom of the pixel array.

Both MIMOSA 22 and MIMOSA 22bis share a reference pixel architecture, implemented at S6 and S5 submatrices respectively, that has been very well tested and exhibits excellent performance. From now on, it will be referred as the reference pixel design. It features a self–biased diode with radiation tolerant design of the sensing diode. A polysilicon ring surrounds the diode, aiming to break the path of the leakage current. Its amplification architecture is an enhanced common source amplifier with static feedback (figure 4.5, d). Some radiation tolerant variations of the reference pixel design have been added in MIMOSA 22bis. In the first variation, implemented in the submatrix S4, the low pass filter feedback transistor is substituted by an enclosed layout transistor (FELT), which is expected to improve the radiation tolerance. In the second variation, an extra voltage shifting transistor has been added (VST). Submatrices S1 and S3 feature a combination of the two variations (FELT + VST). Submatrix S13 amplification scheme is based on a common source amplifier (figure 4.5, a), while its diode's biasing scheme is based on a reset circuit. During the MIMOSA 22bis test beam, we focused mainly on the characterisation of the reference submatrix (S5), and on submatrices S4 and S13. Their main features are summarised in table 4.1.

Submatrix	diode's size (μm^2)	diode's biasing	ampl. scheme	diode's rad. hardness
S4	14.62	SB	ECS+FELT (d)	yes
S5	14.62	SB	ECS (d)	yes
S13	15.21	3T	CS (a)	yes

Tab. 4.1: Main features of the pixel submatrices of MIMOSA 22bis sensor, that were tested during September–October 2008. The diode's biasing refers to the leakage current compensation (see sub–sections 3.2.2, 3.2.2). In the amplification scheme description, the letter which is inside the parenthesis refers to figure 4.5. The diode's radiation hard design refers to if a polysilicon ring has been implemented around the diode.

4.3.3 Laboratory Test Results

Before exposing the sensor to the beam, several tests have been conducted in the laboratory. The sensor was exposed to a ^{55}Fe source, emitting in particular 5.9 keV X–rays, in order to calibrate its Charge to Voltage Conversion factor (CVC). Each submatrix has been calibrated separately since it has its own settings and gain. The analog output derived from the 8 columns that terminate without discriminator, has been used for the calibration and the evaluation of the noise performance of the sensor.

The CVC was measured to be between 40 and 60 μV per electron. After the calibration, the ENC can be measured. The TN (see sub–section 3.2.3) of the pixels has been measured between 10–14 e^-, at a temperature of $\sim 20^0 C$. The submatrices with

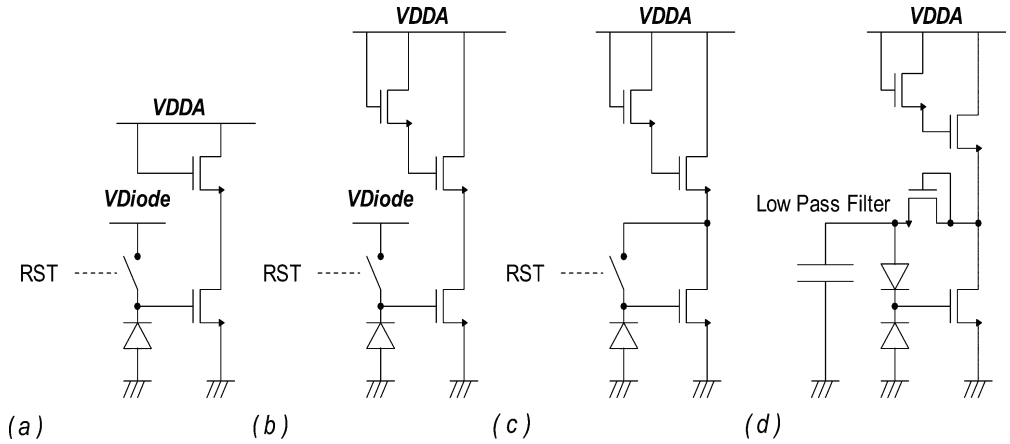


FIG. 4.5: MIMOSA22 amplification schemes. See the text for details.

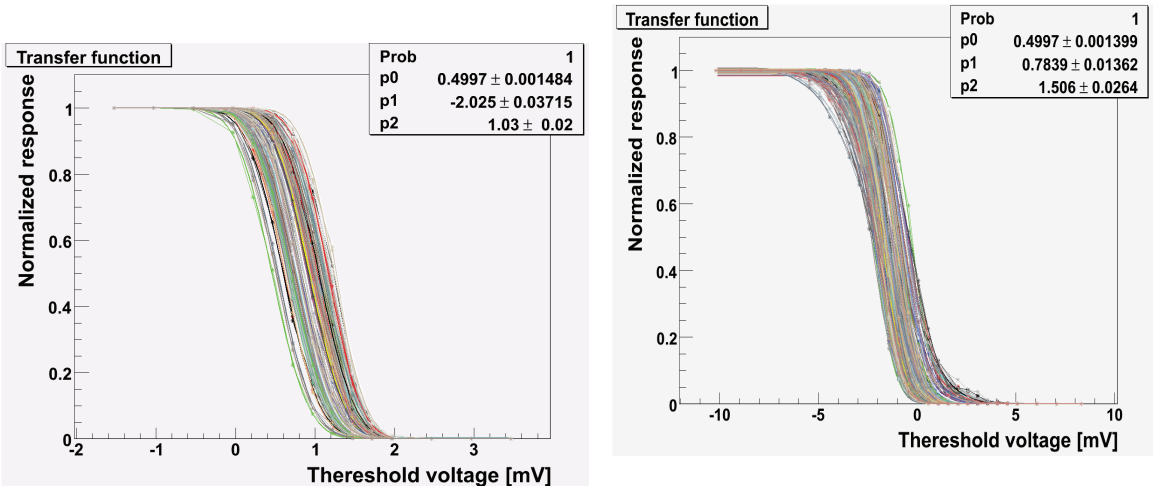


FIG. 4.6: Transfer functions of the MIMOSA 22bis discriminators when applying a threshold scan: Left, standalone; right, connected with the pixel submatrices. The tests were conducted at a temperature of $20^0\,C$.

radiation tolerant diode exhibited a TN increased by $\sim 1e^-$. The explanation of this effect is that the polysilicon ring around the diode increases its capacitance. The FPN was measured to be about $5\,e^-$.

The ^{55}Fe source has been also used to measure the cluster Charge Collection Efficiency (CCE), which expresses the percentage of the charge generated by the impinging X–ray photon that is collected by a cluster of a certain size. It was observed to be 70–80 % for a 3×3 and 80–90 % for a 5×5 pixels cluster. The CCE has been calculated with the following procedure. The distribution of the charge collected in the cluster

is plotted. Then, the CCE is derived from the ratio of the peak of the cluster charge distribution with the calibration peak.

Concerning the digital part of the sensor, the discriminators have been characterised, standalone and connected with the pixel matrices. In order to characterise standalone the discriminators, we connect them to a fixed voltage source. Then the output of the discriminators is scanned by adjusting the common discriminator thresholds through the JTAG. This way the transfer curves for each discriminator (see figure 4.6) have been obtained. These curves show how the response of the discriminators change from 0 to 1, depending on the applied threshold. Excellent performance have been obtained in both cases.

By fitting the transfer curves, one can obtain the values of the TN and the FPN. The TN is estimated from the slope of the curve, while the FPN from its dispersion. The TN and the FPN are displayed in figure 4.7, for the submatrix 5 of the MIMOSA 22bis sensor, after being irradiated with an integrated dose of 300 kRad. The observed values for TN and FPN, in units of ENC, are $\sim 17\,e^-$ and $\sim 5\,e^-$ respectively. Thus, even for irradiated sensors, the TN stays below $20\,e^-$, while the FPN which mainly arises from the dispersion of the discriminator's response, is at least twice lower than the TN. Its quadratic contribution to the total noise is thus $\lesssim 10\,\%$.

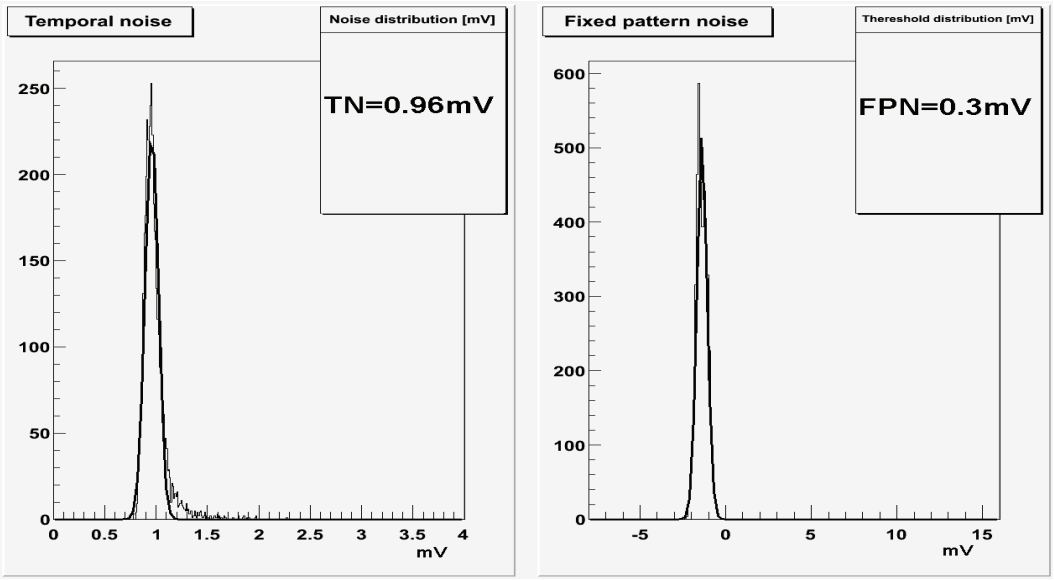


Fig. 4.7: Noise performance of the submatrix S5 of MIMOSA 22bis combined with the discriminators, after being irradiated with an integrated dose of 300 kRad, at a temperature of $20^0\,C$. Left, temporal noise; it corresponds to $\sim 17\,e^-$ ENC. Right, fixed pattern noise dispersion; it corresponds to $\sim 5\,e^-$ ENC.

4.3.4 Test Beam Data Analysis

Four different chips were mounted on the silicon strip beam telescope and tested in the CERN–SPS test beam facility, during a two week period at September–October 2008. Table 4.2 lists the tested chips, their irradiation level, the test temperature and the number of events acquired. A separate analysis has been performed for the analog and the digital part of each sensor.

Concerning MIMOSA 22bis, the submatrices S1, S2, S3, S4, S5 and S13 have been analysed. Each one of them consists of 32×136 pixels. We are going to focus mainly on the results obtained for the reference pixel design (submatrices S6 and S5 for MIMOSA 22 and MIMOSA 22bis respectively).

Sensor	Irradiation	Temperature(^0C)	Events($\times 10^6$)
MIMOSA 22bis	0	20	2.48
MIMOSA 22bis	150 kRad	20 & 35	2.17
MIMOSA 22bis	300 kRad	20	0.66
MIMOSA 22	$10^{12} n_{eq}/\mathrm{cm}^2$	20	1

TAB. 4.2: List of sensors tested during Sep/Oct 2008 test beam. For each sensor are reported its irradiation level, the temperature in which the tests were conducted and the number of obtained events.

The concepts of the major performance, that can characterise a sensor, should be defined. The temporal noise (TN) and the pedestal dispersion (FPN) have been already defined in sub–section 3.2.3, and the way to extract them via the data has been described in sub–section 4.2.2. The concept of the detection efficiency is explained in sub–section 3.1.1. We accept that the DUT has detected the passage of a MIP, if a hit is reconstructed "close" to the interpolation point of a good track on its surface. As good tracks, are defined the ones that have a reconstructed hit in each of the reference planes of the telescope, and have a $\chi^2/DoF \leq 1$. For the purposes of our study, "close" means ≤ 3 times the pixel pitch.

The reported uncertainty of the efficiency is purely statistical. Its estimation is based on the assumption that the efficiency follows a binomial distribution. In case of 100 % efficiency, the binomial approach lead us to an uncertainty equal to 0. This is obviously not true. To avoid this, the uncertainty is computed assuming that one of the tracks is not matched.

The resolution is calculated from the equation 4.9. The residual, unless stated otherwise, is calculated with the η positioning algorithm, on a 3×3 pixel cluster. The telescope resolution is considered to be $1.4\,\mu$m. The telescope is considered perfectly aligned with the DUT. Figure 4.8 shows the distribution of the track–hit residuals on the U and V coordinates for the reference submatrix of a non–irradiated chip.

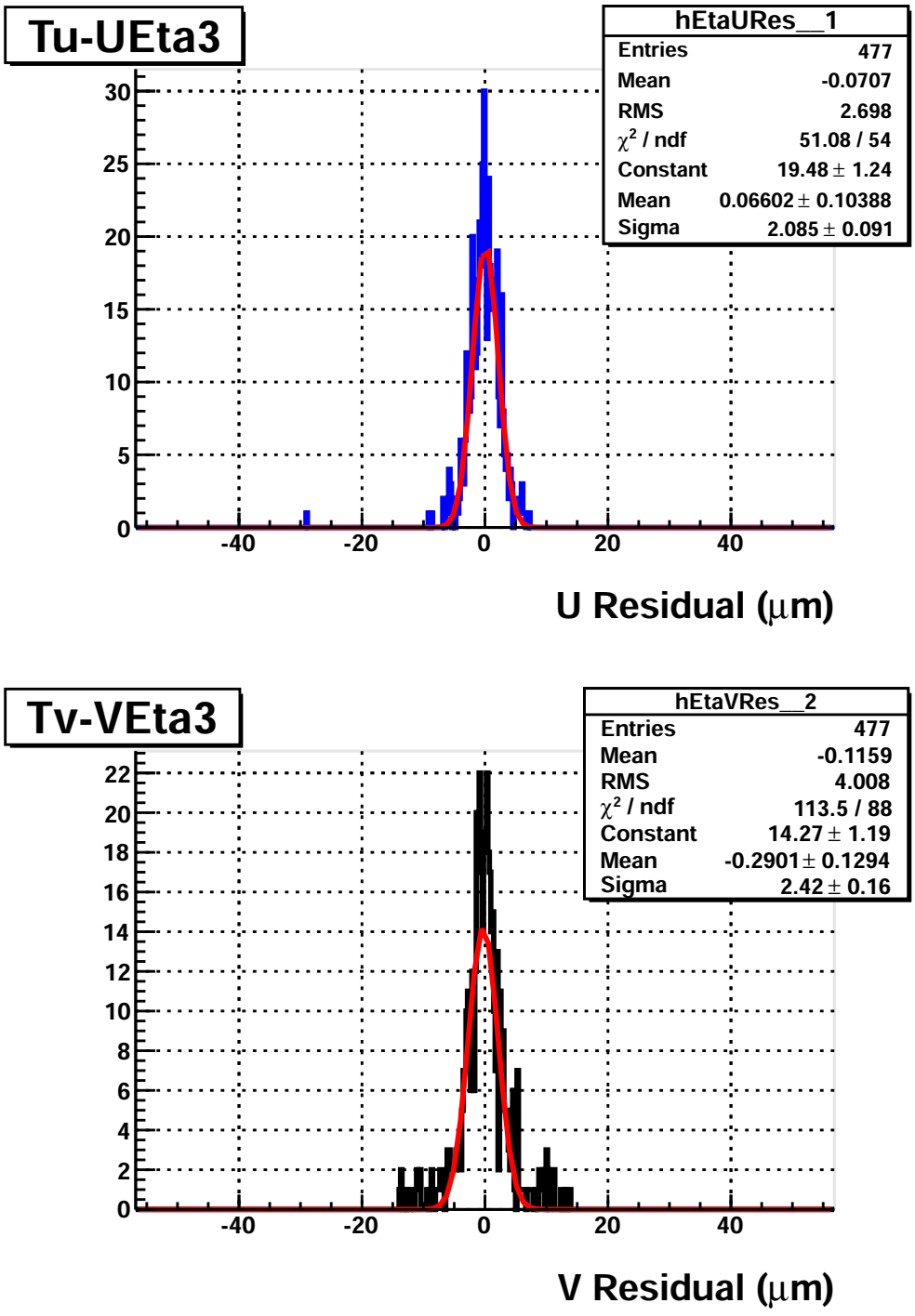


Fig. 4.8: Track–hit residuals of submatrix 5 of a non–irradiated chip. The hit position was recostructed with the η algorithm. The residuals on the U coordinate are illustrated at the top of the figure, while at the bottom are the ones on the V coordinate.

The average fake hit rate per pixel is measured with the following procedure: the MAF software reconstructs only one track per event. In order to measure the fake hit rate, we consider events where the track crosses the DUT far away from the submatrix under study. This implies that if there are any hits on the submatrix surface, they are definitely not caused by a real traversing particle. The statistical uncertainty of the fake hit rate is estimated assuming that it follows a Poisson distribution. Finally, it should be stressed that all the reported uncertainties are purely statistical, and no chip to chip dispersion is taken into account.

Analog Part

The eight columns that provide analog output allow to perform the analysis described in sub–section 4.2.2, and to estimate the pedestal and the noise during data taking, while the beam is on. As seed candidates are considered the pixels with signal to noise ratio (SNR) value higher than the applied threshold, which in this case is equal to 4. The cluster SNR threshold is $SNR_{neig} \geq 0$. The seed pixel's noise is shown at the left part of the figure 4.9, for the reference pixel design (S5) of the non–irradiated MIMOSA 22bis. Its mean value is measured to be $11.61 \pm 0.03 e^-$, when the temperature is stabilised at $20^0\,C$.

The seed pixel SNR is shown in figure 4.9, at the right part. Being fitted with a Landau distribution, the estimated most probable value (MPV) is 18.6 ± 0.2. Such an SNR value promises an excellent, $\sim 100\,\%$, detection efficiency. Actually, the detection efficiency measured with the analog part, reaches $99.95 \pm 0.04\,\%$.

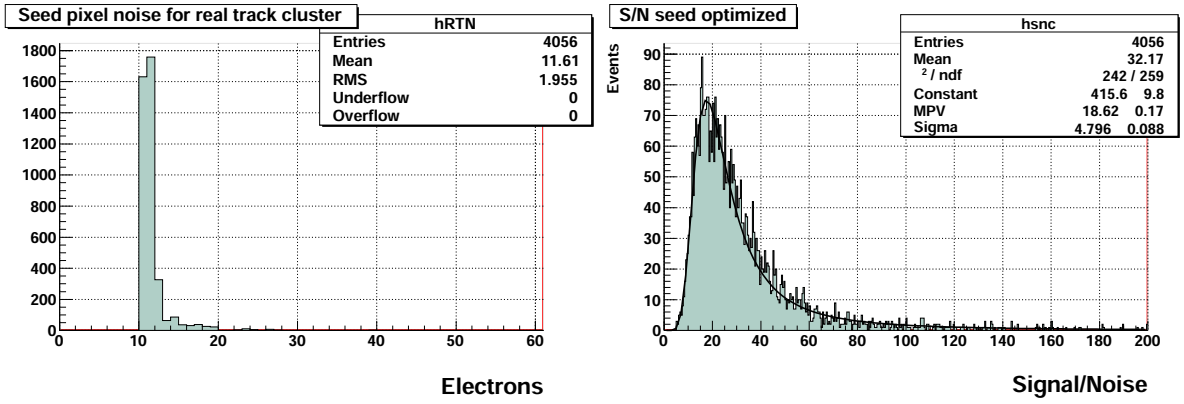


FIG. 4.9: MIMOSA 22bis reference pixel design: left side, seed pixel noise; right side, seed pixel SNR. Both are measured at $20^0\,C$.

Digital Part

The digital output of the sensor is provided by the 128 columns ending with a discriminator. We mainly focus on the detection efficiency, the average fake hit rate and the spatial resolution. Excellent detection efficiency and fake hit rate performance are somewhat conflicting requirements, since both depend on the discriminators SNR threshold in opposite ways (see sub–section 3.2.3). The discriminator threshold is set in units of voltage, but each voltage value corresponds to a specific SNR value.

A scan over the thresholds has been realised, in order to find the optimal one, offering an efficiency of $\sim 100\%$, while keeping the fake hit rate to an acceptable level. The results for the submatrix S5 of MIMOSA 22bis are displayed in figure 4.10. Both efficiency and fake hit rate decrease when the threshold increases. The spatial resolution appears to be nearly independent from it, having a value of $\sim 3.5\,\mu$m. Its fluctuations are not really understood. A working point providing an efficiency of $\sim 99.8\,\%$ for an average fake hit rate per pixel of $O(10^{-5})$ and a resolution of $\sim 3.5\,\mu$m can be selected. It corresponds to a discriminator threshold of 3.0 mV, which can be translated to a SNR threshold of 6.2.

Radiation Tolerance Studies

The radiation tolerance requirements of the EUDET beam telescope are rather moderate. The telescope is going to be operated at the DESY test beam facility. There, it will be exposed to an annual flux of $\sim 10^{11} e^-$ of a few GeV. This means that it should withstand an annual dose of ≤ 3 kRad and a fluence of $\leq 10^{10}\, n_{eq}/\mathrm{cm}^2$. Several users would prefer to operate the telescope at the CERN–SPS test beam facility, where more severe radiation damage is expected. The annual flux at the CERN–SPS would at most be several 10^{12} pions having an energy of O(10^2) GeV. Then the radiation tolerance requirements would increase to an annual dose of O(10) kRad and a fluence up to $O(10^{12})\, n_{eq}/\mathrm{cm}^2$.

To test the radiation tolerance of the different pixel designs, sensors irradiated with various integrated ionising doses and fluences were mounted on the telescope and exposed to the beam. Two MIMOSA 22bis chips were irradiated with 10 keV X–rays up to an integrated dose of 150 and 300 kRad respectively. One MIMOSA 22 chip was irradiated with an integrated flux of $O(10^{12})\, n_{eq}/\mathrm{cm}^2$ in a nuclear reactor.

Figures 4.11, 4.12 and 4.13 summarise the observed performance in terms of detection efficiency, resolution and average fake hit rate per pixel for the submatrix S5 of MIMOSA 22bis for several integrated radiation doses. These performance should be presented as a function of the discriminator threshold in units of SNR, since for different chips, identical voltage thresholds correspond to different SNR thresholds.

As it was expected, the detection efficiency and spatial resolution performance are degrading when the radiation levels increase. On the other hand, a quite strange result

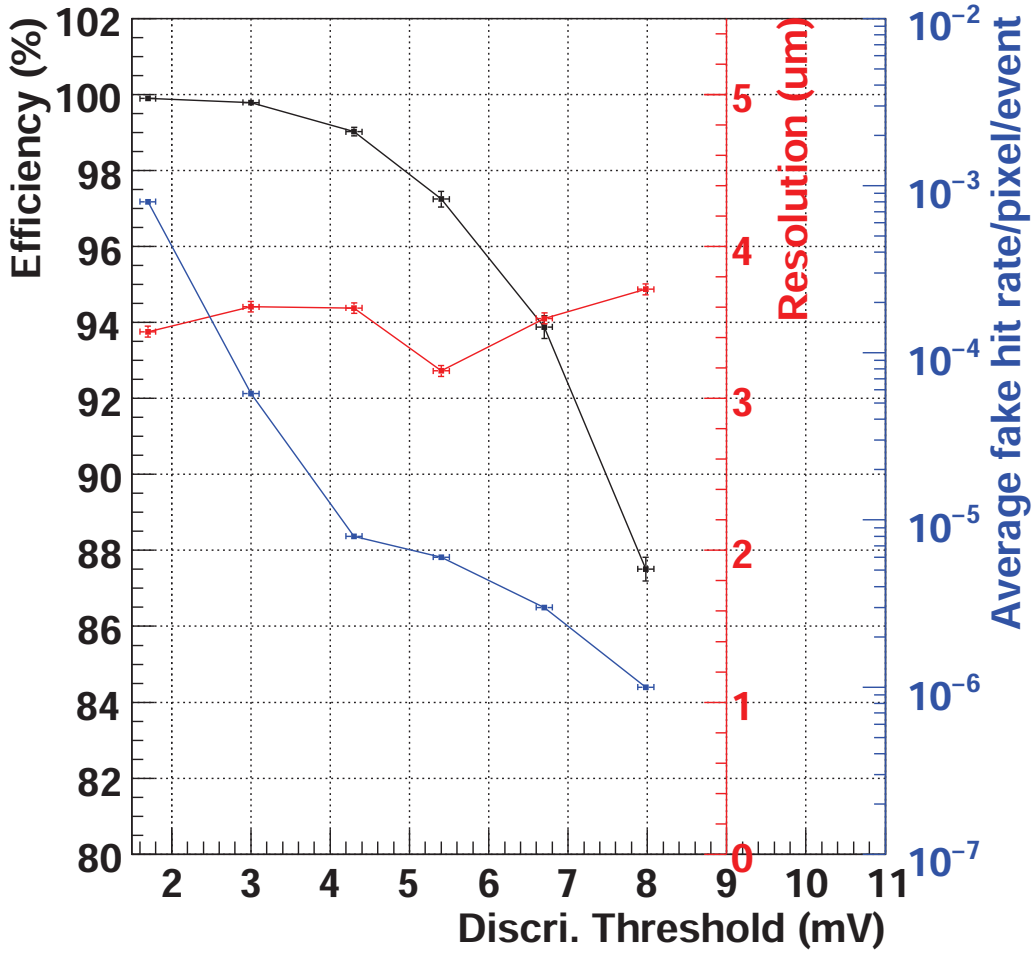


FIG. 4.10: MIMOSA 22bis detection efficiency (black), spatial resolution (red) and fake hit rate (blue) as a function of the discriminator thresholds, in units of mV, for the submatrix S5. Unless otherwise stated, the tests were conducted in a temperature of $20^0\,C$.

is that the fake hit rate seems also to decrease as the radiation dose raises. A second strange result is that the degradation of the spatial resolution performance with the irradiation level seems to mitigate when heating the sensor. In figure 4.12, we see that for the same sensor, being irradiated with an integrated dose of 150 krad, when the testing temperature increases from 20^0 C to 35^0 C, the spatial resolution improves by $\sim 0.4\,\mu$m. But as was mentioned, the measured spatial resolution appears to fluctuate at $\sim 0.5\,\mu$m for the same sensor, and the reason for that is not really understood. Therefore, the results of figure 4.12 should not lead us to the conclusion that spatial resolution improves when the temperature increases. Table 4.3 presents the observed performance, obtained at the optimal* discriminator threshold, for sensors irradiated with 0, 150 kRad and 300 kRad at room temperature ($20^0\,C$) for the submatrix S5.

Radiation dose (kRad)	Efficiency (%)	Resolution (μm)	Fake hit rate per pixel
0	99.79 ± 0.06	3.60 ± 0.04	5×10^{-5}
150	99.66 ± 0.07	4.06 ± 0.04	8×10^{-5}
300	99.42 ± 0.09	4.30 ± 0.04	$O(10^{-4})$

Tab. 4.3: Performance of the submatrix S5 for several integrated ionising radiation doses, obtained for their optimal discriminator threshold. Uncertainties are statistical only.

Finally, we have to select the pixel architecture that is most suitable for the EUDET beam telescope sensor. Most of the submatrices exhibit a satisfactory performance. The most promising candidates are the S5 of MIMOSA 22bis (same as the S6 of MIMOSA 22), and the S4. The S4 features the same diode design and amplification scheme as the S5. Their difference is that the low pass filter feedback transistor of S5 is substituted by an enclosed layout transistor (FELT) in S4, which is expected to provide a more radiation tolerant design.

The detection efficiency versus the average fake hit rate per pixel for various radiation doses and temperatures is shown in figure 4.14, for these two pixel architectures. As one can clearly see, their performance are similar. Finally, the S5 pixel architecture was selected, because it has been validated in two sensor submissions, i.e. those of MIMOSA 22 and MIMOSA 22bis.

The main outcome of this test beam is that the column parallel readout architecture, with integrated discriminators, has been fully validated in two sensor's submissions. The requirements of the EUDET telescope sensor are satisfied. The EUDET beam telescope aims for an extrapolated resolution on the DUT surface of $\sim 2\,\mu$m. To achieve that, the reference planes sensors should feature a spatial resolution of $\leq 4\,\mu$m.

*The optimal threshold is the one offering the best compromise between the detection efficiency and the fake hit rate, while preserving the spatial resolution.

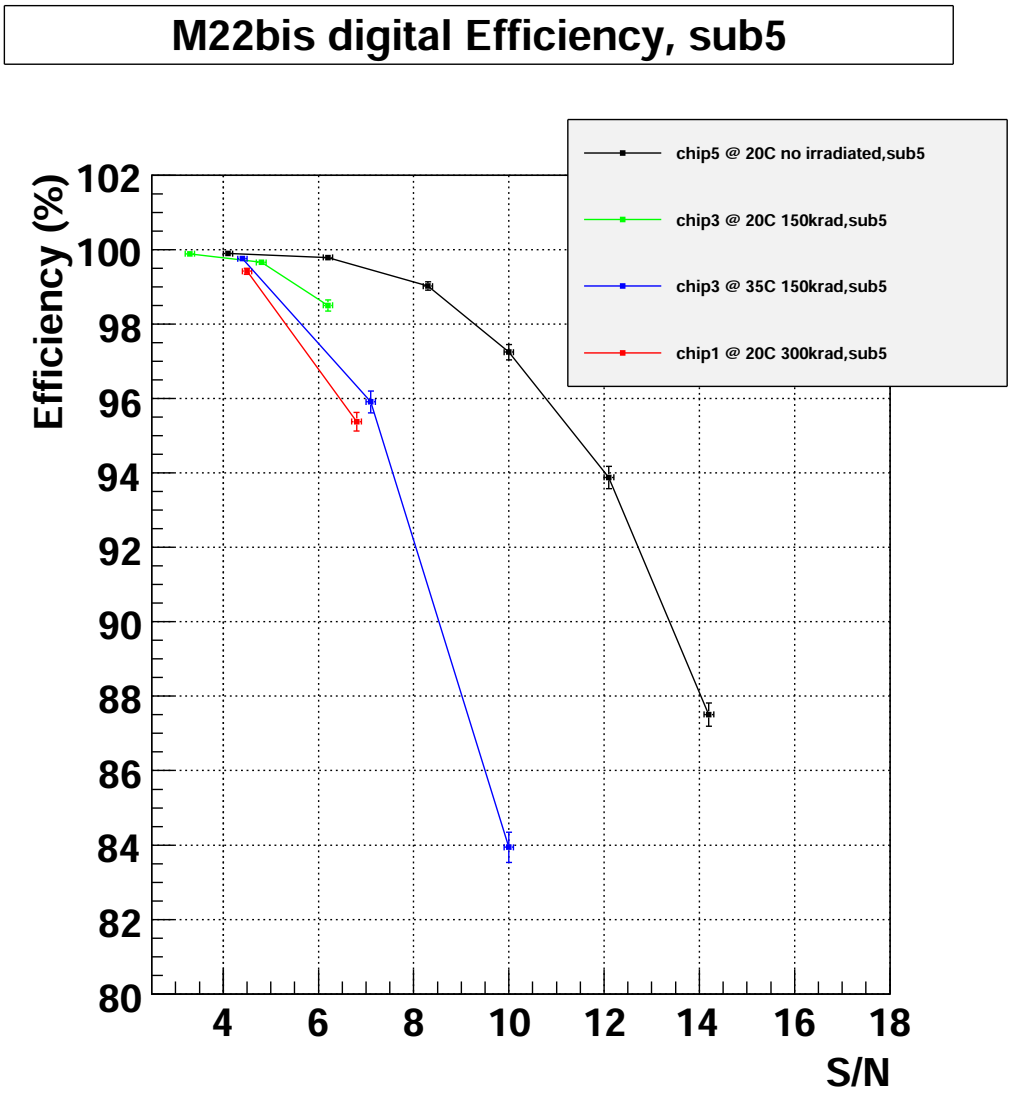


FIG. 4.11: Detection efficiency of submatrix 5 for several integrated radiation doses of the chip and temperatures. Unless otherwise stated, the tests were conducted in a temperature of $20^0\,C$. The black line corresponds to 0 radiation, the green to an integrated dose of 150 kRad, the blue to an integrated dose of 150 kRad at a temperature of $35^0 C$ and the red to an integrated dose of 300 kRad. the horizontal axis is the discriminators threshold in units of SNR.

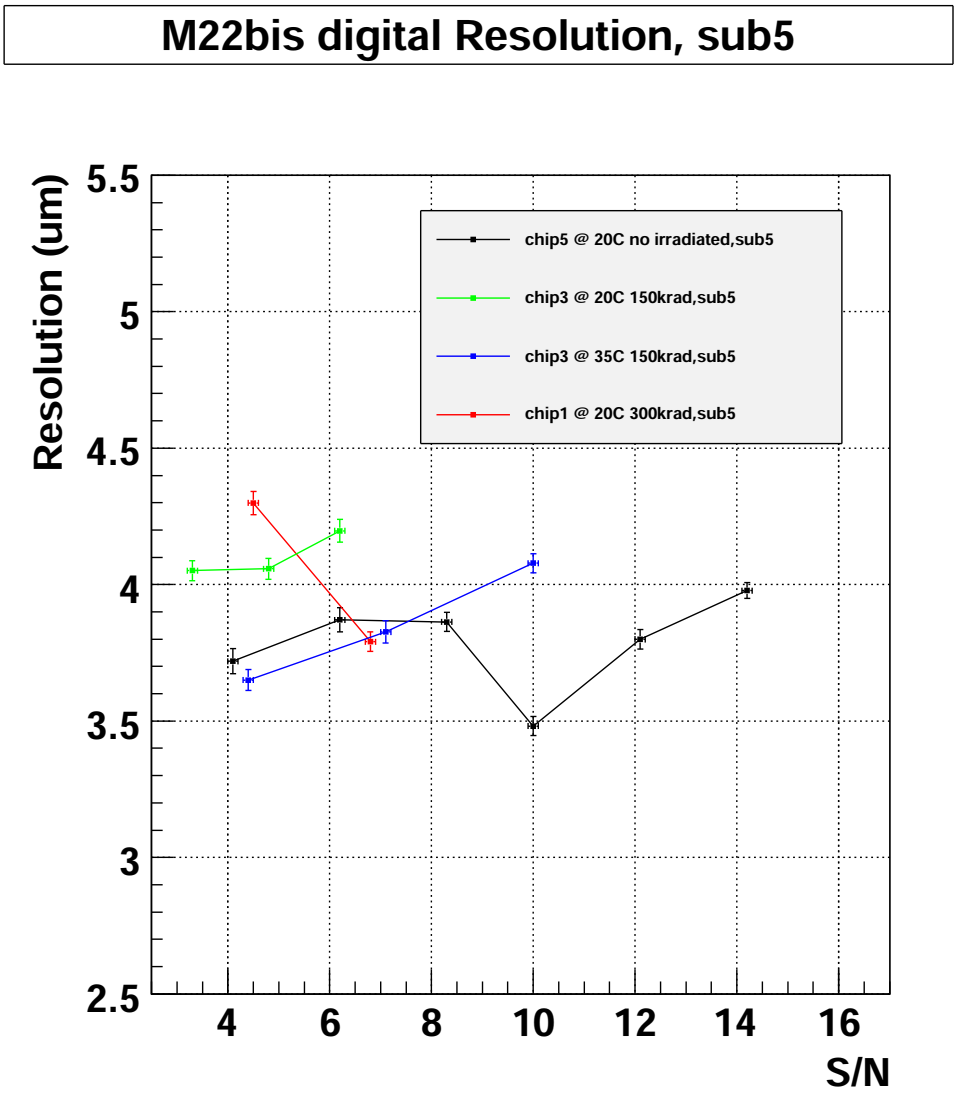


Fig. 4.12: Spatial resolution of submatrix 5 for several integrated radiation doses of the chip and temperatures. Unless otherwise stated, the tests were conducted in a temperature of $20^0\,C$. The black line corresponds to 0 radiation, the green to an integrated dose of 150 kRad, the blue to an integrated dose of 150 kRad at a temperature of 35^0C and the red to an integrated dose of 300 kRad. the horizontal axis is the discriminators threshold in units of SNR.

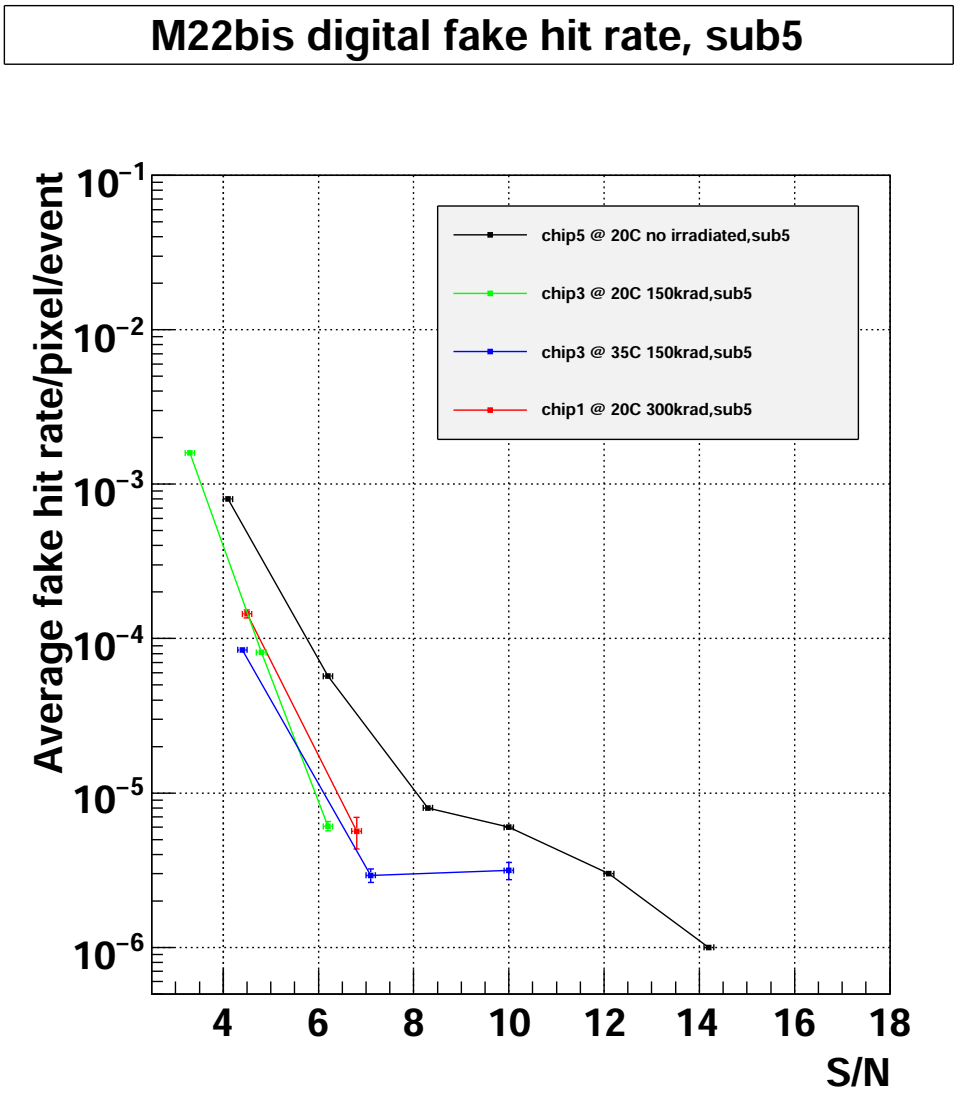


FIG. 4.13: Average fake hit rate per pixel of submatrix 5 for several integrated radiation doses of the chip and temperatures. Unless otherwise stated, the tests were conducted in a temperature of $20^0\,C$. The black line corresponds to 0 radiation, the green to an integrated dose of 150 kRad, the blue to an integrated dose of 150 kRad at a temperature of $35^0 C$ and the red to an integrated dose of 300 kRad. the horizontal axis is the discriminators threshold in units of SNR.

The performance exhibited from the reference pixel design are compliant with this goal. The radiation tolerance requirements are also satisfied, as it was demonstrated with the study of the irradiated sensors. Thus, the reference pixel design will be implemented in the final sensor, MIMOSA 26.

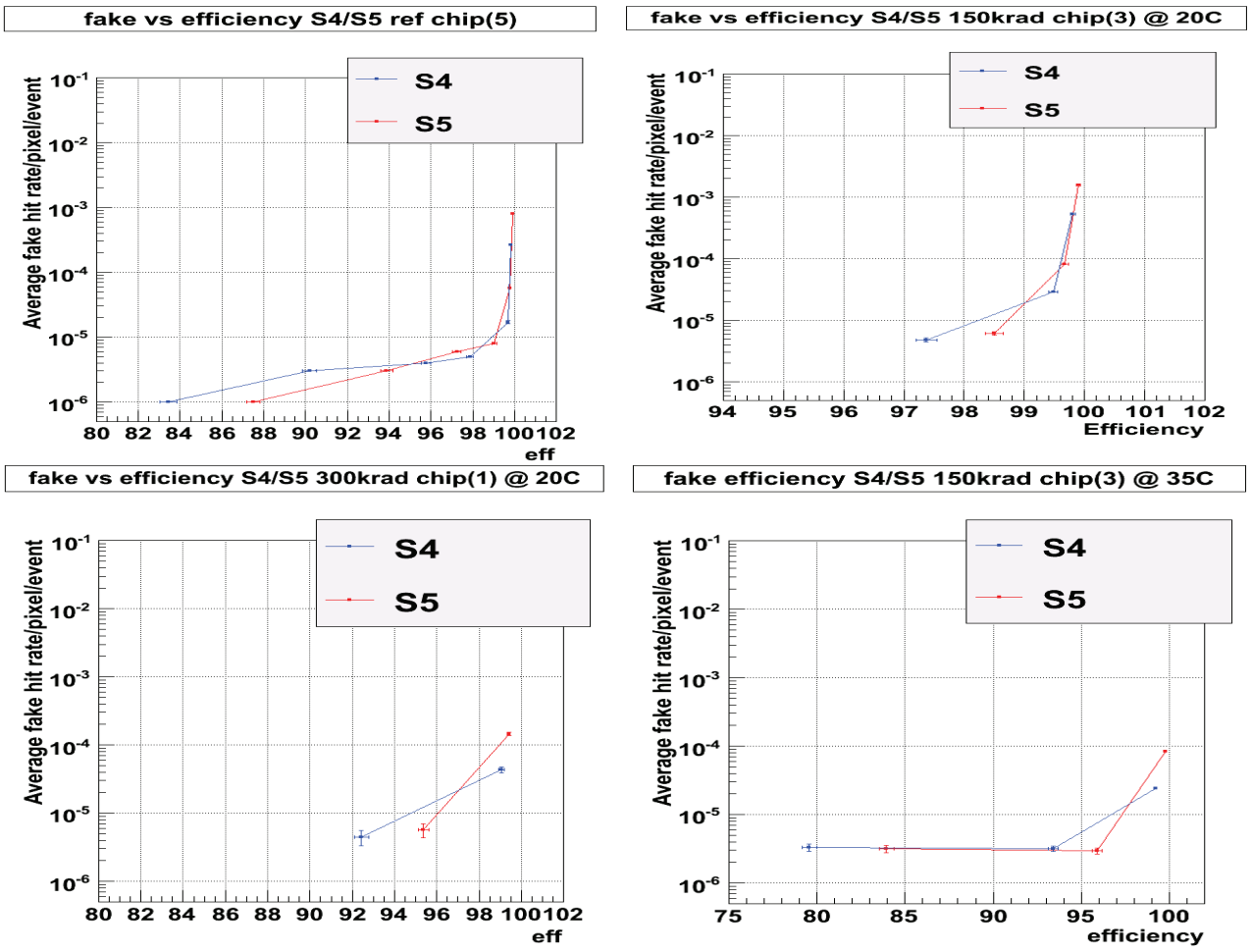


Fig. 4.14: Detection Efficiency–average fake hit rate per pixel trade off for the submatrices S4 (blue line) and S5 (red line) of MIMOSA 22bis. Unless otherwise stated, the tests were conducted in a temperature of $20^0\,C$. Top: on the left is for the non irradiated chip, on the right for an integrated dose of 150 kRad. Bottom: on the left for an integrated dose of 150 kRad at $35^0 C$, on the right for an integrated dose of 300 kRad.

4.4 MIMOSA 24

The detection performance of CMOS sensors depend strongly on various manufacturing parameters, which are specific to each fabrication process. A characteristic example is the thickness of the epitaxial layer, which determines the magnitude of the collected signal. Several CMOS fabrication processes exist, and the industry invests in an intensive $R\&D$ effort for their optimisation for optical imaging applications. As already mentioned, charged particle tracking applications have a marginal impact on this $R\&D$ effort. Therefore, an important task of the MIMOSA sensors development consists in exploring various fabrication processes in order to find the optimal one for charged particle tracking.

The design of the MIMOSA 24 sensor can be included inside this framework of the exploration of fabrication processes. In this case, the XFAB 0.35 μm process was assesed. The main motivation to investigate this process was the announced intention of XFAB to launch an option featuring a high resistivity epitaxial layer. The advantages of a high resistivity epitaxial layer are described in sub–section 3.4.2. The first sensor featuring a high resistivity epitaxial layer, MIMOSA 25, had demonstrated very encouraging results [Dorokhov 10]. It was fabricated with the XFAB process, but it had a rather large feature size of 0.60 μm. It is thus very interesting to test the XFAB process, moving to the smaller feature size of 0.35 μm. Therefore MIMOSA 24 emerged, featuring yet a low resistivity epitaxial layer though.

The scope of the MIMOSA 24 tests are to evaluate the basic performance of this process, and compare them with the ones obtained with the AMS 0.35 μm OPTO, which was the one providing the best performance for the MIMOSA, at least by that time. To perform these comparative studies, we are going to use the results obtained with MIMOSA 9 [Dulinski 04], a sensor having similar architecture as MIMOSA 24 but fabricated in the AMS 0.35 μm OPTO process. It should be stressed that currently, more manufacturers provide high resistivity epitaxial layers, thus moving to XFAB is not an one–way road.

4.4.1 Architecture

MIMOSA 24 is an analog output sensor with an integration time of 160 μs. Its layout is displayed in figure 4.15. Its pixel array is organised in four parts, each one with its own analog output. Each of these parts is further subdivided in two submatrices, thus the sensor has eight pixel submatrices in total. The main design features of each submatrix are summarised in table 4.4. The number of pixels in each of them is 32×32 for the first three parts and 16×32 for the fourth one.

The study of the ionising radiation hardness of the XFAB 0.35 μm process is of special interest. Two radiation tolerant diode designs have been implemented in the third part of the sensor. The first one reproduces the design used in the sensors manu-

factured in the AMS process, thus it is noted with the suffix AMS. It was however not possible to implement exactly the same design with the XFAB process. The second one is the enclosed layout transistor design, noted as ELT. In both cases, the implementation of the thin oxide used to interrupt the leakage current path could not reach up to the n–well–epitaxial layer boundary surface. For more information on these radiation tolerant structures, one can refer to [Koziel 09].

Submatrix	Pixel pitch (μm)	Diode dimensions (μm^2)	Design	Rad. Tol design
0	20	4.3×3.4	SB	
1	20	6×6	SB	
2	20	4.3×3.4	3T	
3	20	6×6	3T	
4	20	4.3×3.4	3T	ELT
5	20	4.3×3.4	3T	AMS
6	30	4.3×3.4	3T	
7	30	5×5	3T	

Tab. 4.4: Pixel characteristics of each submatrix of the MIMOSA 24 sensor.

4.4.2 Test Beam Data Analysis

The test beam took place in August 2009 at the CERN–SPS. The tests were conducted with the TAPI–TNT configuration (see sub–section 4.1.2) and the data were analysed using the TAF software (see sub–section 4.2.1). Four chips, exposed to various integrated radiation doses and fluences were tested. These are the non irradiated reference chip 9, the chip 7 that was exposed to a radiation dose of 500 kRad, and two chips, numbered 4 and 5, with an integrated fluence of $3 \times 10^{12} n_{eq}/\text{cm}^2$ and $1.3 \times 10^{13} n_{eq}/\text{cm}^2$ respectively.

To evaluate better the results, we should stress some consequences arising from the data sparsification provided by the TNT DAQ board. Concerning the noise and pedestal extractions, the recursive formulae 4.6 and 4.4 could not be used. Due to the threshold applied to the DAQ, only pixels with raw value higher than this threshold are read out. This affects also the cluster reconstruction. The number of pixels in a cluster depends on the TNT threshold. It is not always possible to reconstruct an $n \times n$ cluster around the hit. Additionally, the charge collected in an $n \times n$ cluster when the data are sparsified will be in average higher than the one collected in a cluster with the same multiplicity, in the case of non sparsified data. The reason is that in the non sparsified data case, pixels with relatively low charge enter in the cluster reconstruction.

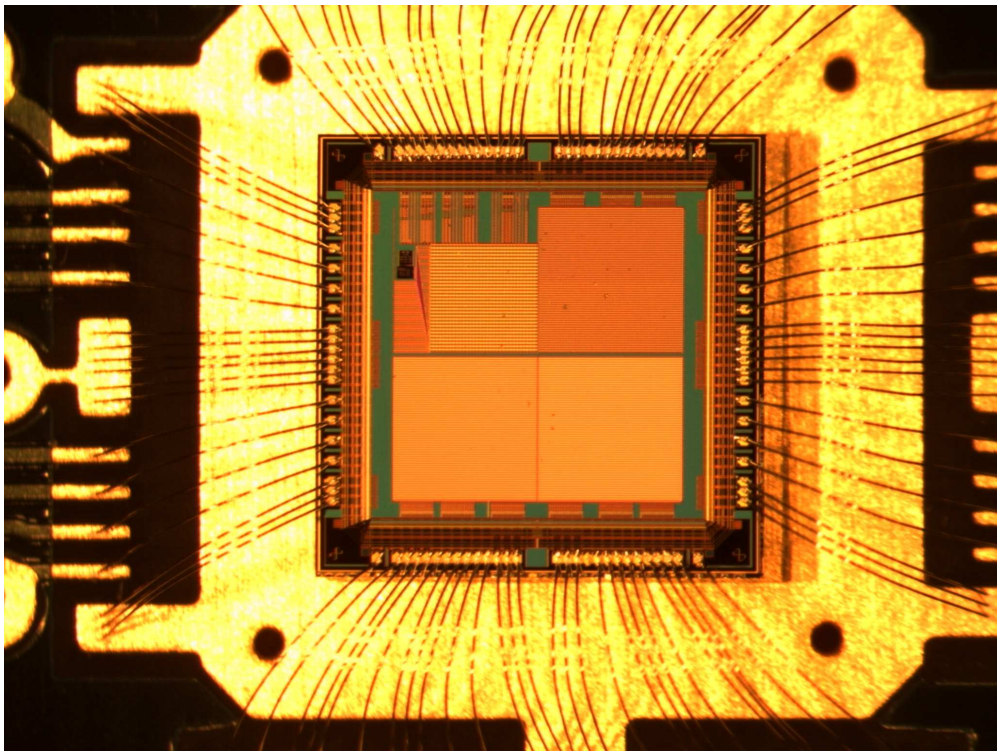


Fig. 4.15: Layout of MIMOSA 24 sensor.

These pixels will be cut out if a threshold is applied on the DAQ. Thus one cannot compare cluster charge or multiplicities between sensors tested with and without data sparsification.

Finally, another point is that during this test beam, for most of the chips we have obtained enough statistics to perform a reliable analysis only for the submatrices 0 and 1. Thus we are going to focus on the analysis of these two submatrices.

Basic Performance

The noise estimation is performed using short dedicated noise runs, when the beam is off and the TNT threshold set to 0. A mean noise value was obtained for each pixel. Thus it was mandatory to estimate the seed SNR using an indirect way: the charge collected in the seed pixel during a data taking run is divided by the noise extracted for the same pixel during the noise run. Figures 4.16 and 4.17 show the charge collected, the noise and the SNR for the seed pixel for the submatrices 0 and 1. The applied thresholds for the hit reconstruction are ≥ 5 for the seed candidates and ≥ 2 for the cluster candidates, in units of SNR. Independently of these thresholds, the pixels that do not exceed the TNT DAQ threshold are not taken into account.

For the submatrix 0, featuring a small diode of $4.3 \times 3.4\,\mu\text{m}^2$, the charge collected in the seed pixel, being fitted with a Landau distribution, has an MPV value of $248.8 \pm 0.7\,e^-$. The obtained noise in the seed pixel is $11.63 \pm 0.10\,e^-$, leading to an SNR of 21.2 ± 0.4. The submatrix 1, with diode dimensions equal to $6 \times 6\,\mu\text{m}^2$, features a seed charge of $327.2 \pm 0.5\,e^-$, a seed noise of $15.64 \pm 0.11\,e^-$, thus an SNR equal to 20.3 ± 0.3.

The statistics obtained for the submatrices 4 and 5, especially for the irradiated chips, were not sufficient to perform a reliable analysis. Thus it was not possible to test whether the AMS radiation tolerant design can be successfully implemented in the XFAB process, or to evaluate the ELT design.

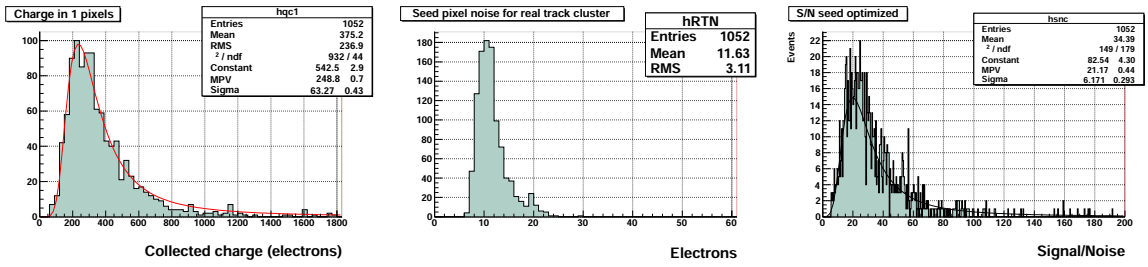


Fig. 4.16: MIMOSA 24 submatrix 0 (diode size equal to $4.3 \times 3.4\,\mu\text{m}^2$) test beam results: left part, charge collected in the seed pixel; middle part, seed pixel ENC; right part, seed pixel SNR.

In order to measure the resolution of the MIMOSA 24, it is first required to estimate the track extrapolation accuracy of the telescope. The individual planes of the TAPI

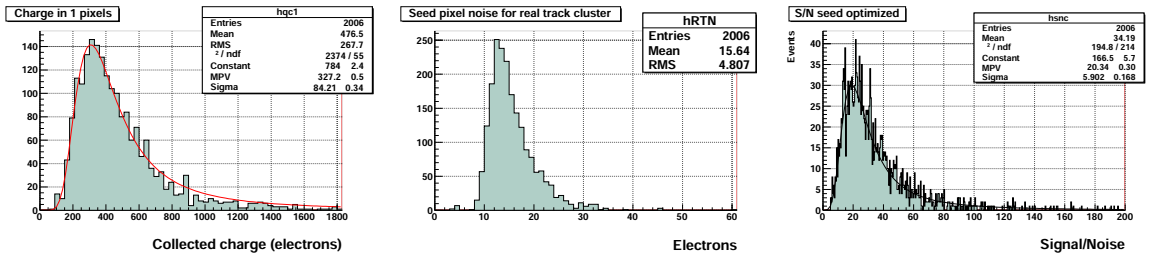


Fig. 4.17: MIMOSA 24 submatrix 1 (diode size equal to $6 \times 6\,\mu\text{m}^2$) test beam results: left part, charge collected at the seed pixel; middle part, seed pixel ENC; right part, seed pixel SNR.

telescope were aligned with an accuracy of $\sim 1.5 - 2\,\mu$m. This can be translated in a track interpolation accuracy on the DUT surface of $\leq 1\,\mu$m. It was demonstrated, as illustrated by figure 4.3.4, that the best resolution is obtained when the hit position is reconstructed with the η algorithm. However, it is not possible to apply the η correction, in the way it is currently implemented in the software, to the data obtained with the TAPI–TNT configuration. The reason is illustrated by figure 4.18. The DAQ threshold does not always allow to reconstruct a cluster around the hit. In the case of one pixel clusters, the hit position given by the CoG algorithm is identical to the digital one. Thus the difference $U_{CoG} - U_{DIG}$ is equal to 0. This results in a very sharp peak, δ–function like, of the η distribution at the position corresponding to the center of the pixel (left part of figure 4.18). When one tries to integrate the η function, this sharp peak leads to a discontinuity at the center of the pixel (right part of figure 4.18), that biases the reconstruction of the hit position at the borders of the pixel. This effect is illustrated in figure 4.19, where we see the interpolated position of the track in the vertical axis while on the horizontal one is the reconstructed hit position. We observe that, independently of the interpolated position of the track impact, the hit is always reconstructed inside a narrow band, close to the pixel borders.

For this reason, the CoG algorithm will be used. An implementation of the η algorithm that can cope with sparsified data is under study. Certain hit selection criteria, concerning the cluster multiplicity and thresholds on the seed and cluster charge should be applied. The resolution obtained with the CoG algorithm, for both directions, of the submatrices 0 and 1 of MIMOSA 24 is presented in table 4.5. The results are compared with the measured resolution of MIMOSA 9, with the same method. We observe slightly better resolution performance with MIMOSA 24. Moreover, the bigger diode seems to yield finer resolution than the small one.

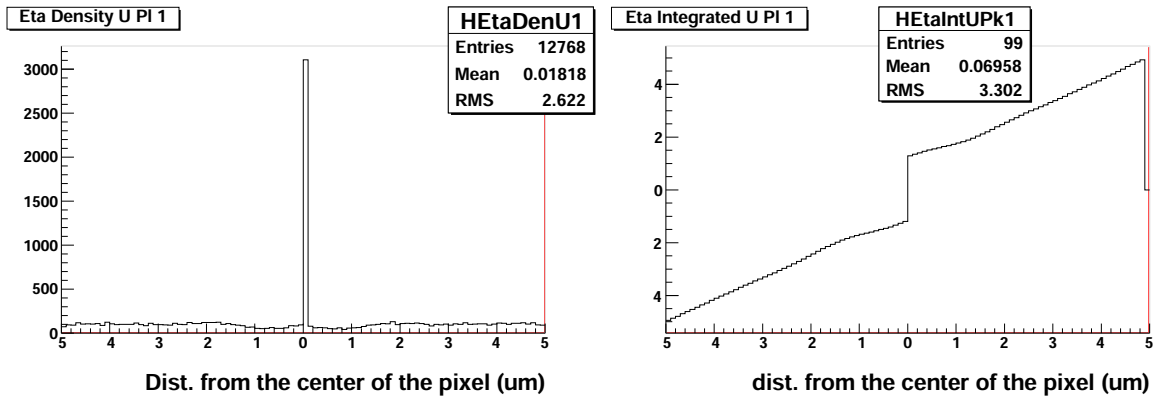


Fig. 4.18: η distributions for the first plane of the TAPI telescope (MIMOSA 18 sensor).Left part: η density, right part: η integrated density. At the horizontal axis is the distance from the center of the pixel.

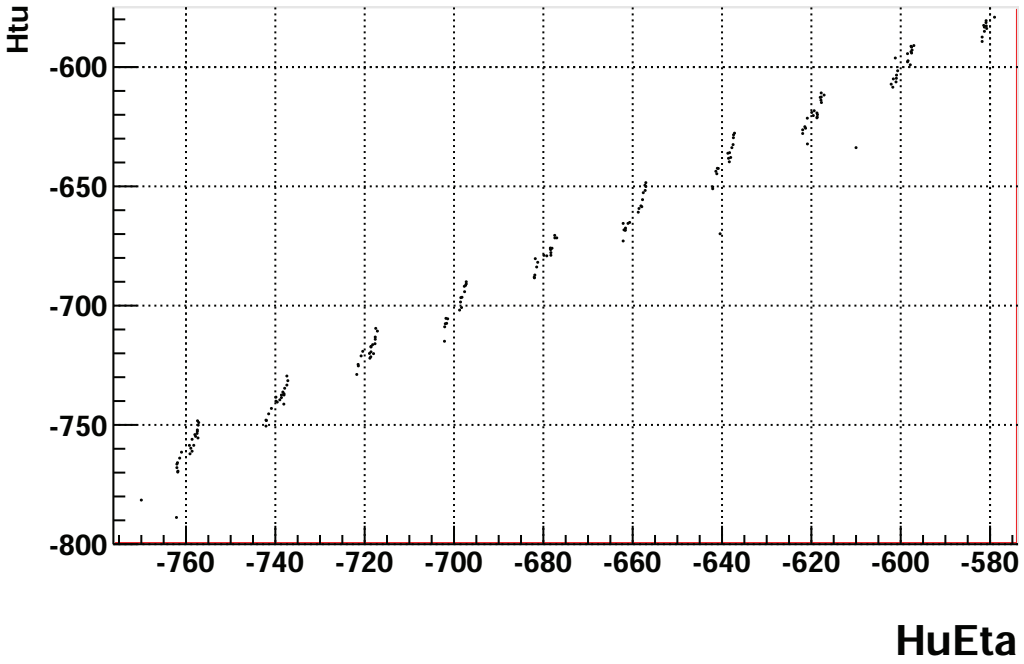


Fig. 4.19: Correlations of the extrapolated position of the track (vertical axis) with the reconstructed hit position for the MIMOSA 24 sensor (horizontal axis), when the hit position is reconstructed with the η algorithm.

sensor	diode (μm^2)	U resolution (μm)	V resolution (μm)
M24	3.4×4.3	2.30 ± 0.03	2.06 ± 0.03
M9	3.4×4.3	2.30	2.03
M24	6×6	1.89 ± 0.03	1.55 ± 0.04
M9	6×6	2.18	1.83

Tab. 4.5: Spatial resolution of MIMOSA 24 and MIMOSA 9. Uncertainties are statistical only.

Comparison with MIMOSA 9

The main motivation for the fabrication of the MIMOSA 9 sensor, by 2004, was to explore a new (then) fabrication process, the AMS 0.35 μm OPTO. The description of the sensor can be found at [Dulinski 04]. Two MIMOSA 9 submatrices are similar to the submatrices 0 and 1 of MIMOSA 24. They feature a pixel pitch of 20 μm, self biased diode design, and diode dimensions of $4.3 \times 3.4\,\mu$m^2 and $6 \times 6\,\mu$m^2 respectively. Thus, by comparing the basic performance of the corresponding MIMOSA 9 and MIMOSA 24 submatrices, the new process XFAB 0.35 μm can be characterised with respect to the very well tested AMS 0.35 μm OPTO.

We should stress at this point that the two sensors have been tested with different configurations, MIMOSA 9 with the silicon strip telescope while MIMOSA 24 with the TAPI–TNT system. Due to this fact, as was explained in sub–section 4.4.2, information coming from the cluster of the hit cannot be compared. Therefore, the study will be restricted to the information coming from the seed pixel, which means charge collection, noise and SNR. The results are presented in table 4.6.

submatrix	seed charge (e)	seed noise (e)	seed S/N
M9 3.4×4.3 diode	250.4 ± 0.3	9.234 ± 0.02	26.3 ± 0.3
M24 3.4×4.3 diode	248.8 ± 0.7	11.63 ± 0.10	21.2 ± 0.4
M9 6×6 diode	338.8 ± 0.4	12.32 ± 0.02	26.8 ± 0.3
M24 6×6 diode	327.2 ± 0.5	15.64 ± 0.11	20.3 ± 0.3

Tab. 4.6: Comparison between MIMOSA 24 and MIMOSA 9 seed pixel information. The reported uncertainty is the statistical one.

Charge collection on seed yields similar values for both sensors. Roughly 20% higher noise was measured with MIMOSA 24, which obviously reflects in the SNR value. However, MIMOSA 24 exhibits a satisfactory SNR, which is ≥ 20 for both submatrices.

Concluding, both sensors exhibit similar performance. As already mentioned, due to the different test setups used, as well as to the low statistics acquired for the irradiated chips, it was not possible to fully validate the XFAB process with respect to the AMS one. Nevertheless, we did not observe any striking results that could discourage us from fabricating sensors in the XFAB 0.35 μm process. Thus, we can conclude that XFAB offers a promising candidate technology for fabricating CMOS sensors for HEP applications.

4.5 Summary and Perspectives

In this chapter we went through the test results of the MIMOSA 22/22bis and MIMOSA 24 sensors. The tests of these sensors have addressed crucial aspects concerning the realisation of a CMOS based VXD. As we have discussed in the previous chapter, the CMOS technology genuinely offers high granularity and low material budget. The main challenge towards a VXD for the ILC is to provide a sensor having the above features combined with a swift readout. The challenge lies in the strong anti–correlation among those requirements. A prominent example is that high spatial resolution requires a highly segmented sensor with a large number of pixels, while high time resolution is favoured by a smaller number of pixels.

Full (reticle) size sensors complying with the VXD spatial resolution requirements should be segmented into $O(10^6)$ pixels. A sequential readout of such a pixel matrix would lead to readout times in the order of a few ms, which is too slow. The organisation of the pixels into columns, and the application of a column parallel readout can improve the time resolution by two order of magnitudes, bringing it close to the VXD requirements. This idea was fully validated throughout the tests of the MIMOSA 22/22bis sensors. Implementing more than one rolling shutter in the sensor will further improve the readout time, going down to few μs, an adequate performance even for the VXD innermost layer.

A second requirement arising from the relatively high hit rate on the inner layers of the VXD is the online data sparsification. In–pixel data sparfication would require digital logic (a discriminator) to be implemented inside the pixel. However, digital logic needs both PMOS and NMOS transistors integrated inside the pixel sensitive area, which would a–priori result in parasitic charge collection (see section 3.2.1). Thus the data sparsification should take place on the sensor periphery. The discriminators are located at the end of each column and they feature a common threshold. The application of a common threshold became feasible due to a double CDS strategy that has been invoked, which compensates for the FPN of the pixel array. The obtained results are quite encouraging. It can provide an efficiency of $\sim 99.8\,\%$ with a corresponding average fake hit rate per pixel of $O(10^{-5})$ for the radiation level expected at the ILC environment. The pixel matrix of the MIMOSA 22/22bis featuring the optimal

pixel design (S5), combined with a data sparsification logic constitute the MIMOSA 26 sensor (see sub–section 3.3), being developed for the EUDET beam telescope.

Obviously, MIMOSA 22/22bis does not cover all of the VXD requirements. However, implementing the S5 architecture in a smaller pixel, by moving to fabrication processes with a smaller feature size, the required spatial resolution can be achieved. Having in mind the required tolerance on ionising radiation (see sub–section 3.1.2), we can see that the reference pixel design of MIMOSA 22/22bis could withstand the expected annual dose. Smaller feature size, combined with a smaller pixel pitch, will improve also the tolerance on both ionising and non–ionising radiation.

As we have discussed, the CMOS based VXD will rely on a double optimisation strategy, which uses both small and large pixel sizes depending on the layer in order to achieve the best compromise between time and spatial resolution, and power dissipation. When one goes to large or elongated pixels, a major concern is the charge collection efficiency (as well as the non–ionising radiation hardness, but this is not a major issue for the ILC). One can envisage to improve the charge collection efficiency by increasing the size of the depletion area of the pixel. A way to do this is to use high resistivity epitaxial layers, which was something non standard to the CMOS processes used by the optical industry until a few years ago. XFAB offered sensors with epitaxial layers featuring a resistivity of a few hundreds $\Omega \cdot \mathrm{cm}$. But since crucial parameters change from one fabrication process to another, before moving to a new one it is desirable to assess its performance against a well–known process. That was the purpose fulfilled from the MIMOSA 24 tests. Even though a complete validation could not be extracted from the tests of this sensor, it was at least demonstrated that using XFAB $0.35\,\mu$m is an open possibility.

In chapters 3 and 4 we have addressed the CMOS sensors technology, and presented results obtained in test beam campaigns. The sensors tested are considered as intermediate steps towards the long term goal, i.e. the development of the final ILD sensor(s). In the next chapter we are going to change our approach. We will assume a sensor technology that satisfies the ILD requirements, and proceed to simulation studies of physics final states, in order to optimise the VXD geometrical configuration.

CHAPTER 5

ILD Vertex Detector Optimisation Studies

5.1 Introduction

The purpose of the ILD vertex detector (VXD) is to provide excellent heavy flavour tagging and to participate in the track reconstruction, where it plays a crucial role especially for the low momentum tracks. The term heavy flavour tagging means the identification of b– and c–jets, and of τ leptons. The performance of the VXD on these aspects is crucial for the extraction of the Higgs couplings to fermions, as well as for the reconstruction of the vertex charge. The feasibility of the aforementioned analyses have been already demonstrated through detailed studies in the framework of the ILD Letter of Intent [LoI 10].

The LoI physics analyses have assumed one of the alternative VXD geometries (see sub–section 2.3.2), the VXD05. We should stress that we do not intent to repeat these analyses. Our main goal in this chapter is to use these studies as benchmarks in order to compare the performance of the alternative VXD geometries, focusing on the two main ones, called VXD03 and VXD05. We seek to find whether there is an argument, concerning the physics performance of the two geometries, that pushes us to choose one of them. To do so, we will examine their heavy flavour tagging performance, and their sensitivity in the extraction of the Higgs hadronic branching ratios.

Apart from studying the two main VXD geometries, we will try also to explore a small part of the available parameter space of their geometrical and technological features. Our motivations are merely the two following: first to see whether we can further optimise the VXD performance; second, some of the LoI–targeted values, mainly the innermost layer radius and the ladder's material budget, may be proven not to be

feasible. In that case, we are interested to find the margin allowed for relaxing the requirements on these parameters, without degrading severely the physics potential of the detector.

5.2 Physics Process–Event Generation

The selected physics channel is the Higgs boson production via the Higgsstrahlung process at $\sqrt{s} = 250\,\text{GeV}$ (left part of the figure 1.3), which is a benchmark process for the ILC physics program. The Higgsstrahlung cross–section maximises near production threshold, therefore for an assumed mass M_H for the Higgs boson of $120\,\text{GeV/c}^2$, the optimal $\sqrt{s}$ is in the range of $\sim 220-250\,\text{GeV}$. Moreover, operating at $\sqrt{s} = 250\,\text{GeV}$ offers a better mass resolution on the $Z \to \mu^+\mu^-$ decays [Bambade]. The SM Higgs branching ratios for this M_H value are summarised in table 5.1.

Decay mode	BR (%)
$b\bar{b}$	68.16
$c\bar{c}$	3.04
gg	6.71
W^+W^-	13.45
ZZ	1.48
$\tau^+\tau^-$	6.74

Tab. 5.1: Branching ratios for a SM Higgs boson of $120\,\text{GeV/c}^2$ mass [Djouadi 98].

The Monte–Carlo generator files produced for the LoI studies have been used. For the Higgsstrahlung process we are using the files produced at SLAC with the WHIZARD event generator [WHIZARD]. The effect of beamstrahlung on the energy spectrum of the beams, thus the resulting effects on $\sqrt{s}$ and the cross–sections, are reproduced with the CIRCE software [Ohl 97]. The effect on $\sqrt{s}$ can be seen in figure 5.1. It shows the distribution of the total Invariant Mass (IM) of all Monte–Carlo particles (which corresponds to the effective $\sqrt{s}$), produced at a Higgsstrahlung event.

The main background processes ZZ and WW (see section 5.3), have been generated at KEK with PYTHIA version 6.409 [PYTHIA]. The assumed energy spread is 0.28 % for the electron and 0.18 % for the positron beam. These values do not account for the low IM distribution tail.

We will focus on the events where the Z boson decays to a pair of muons. The expected branching ratio of $Z \to \mu^-\mu^+$ is $\sim 3.5\%$. This specific channel selection limits severely the statistics. On the other hand, it also limits the potential background sources, thus can provide a relatively easy signal–background separation. The cross–section of the $e^+e^- \to ZH \to \mu^+\mu^- X$ process, where X stands for the Higgs decay

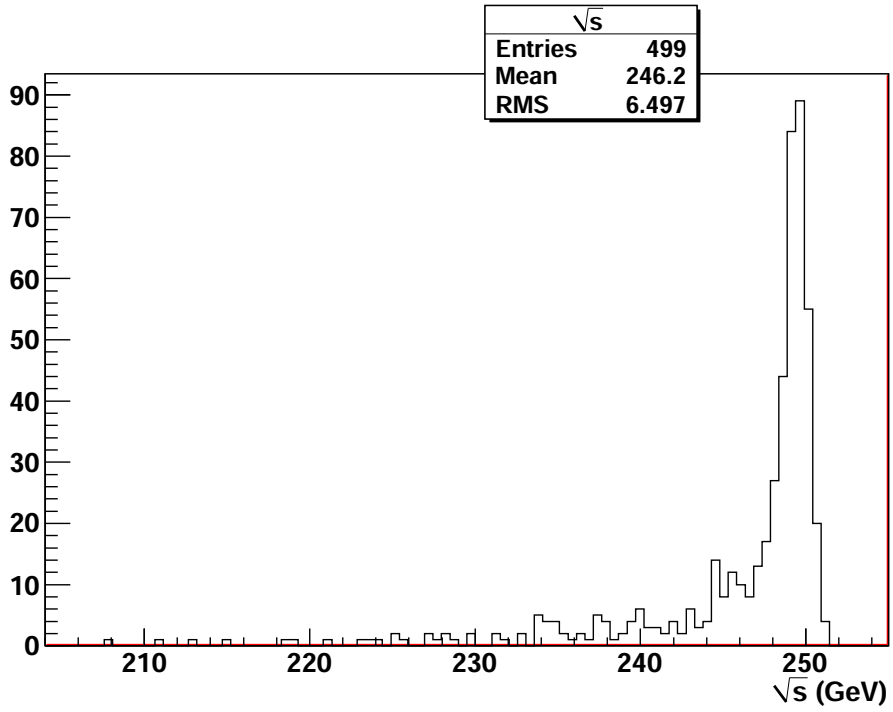


Fig. 5.1: The total invariant mass of all Monte–Carlo particles (which corresponds to the effective $\sqrt{s}$), produced in Higgsstrahlung events.

products, assuming unpolarised beams, is 7.78 fb for $\sqrt{s} = 250$ GeV. For the beam polarisation option of 80 % for e^-_R and 30 % for e^+_L, the cross–section is 7.02 fb. The beam polarisation option of 80 % for e^-_L and 30 % for e^+_R leads to a cross–section of 10.40 fb.

5.3 Physics Background

Each process that could mimic the signal's final state topology, which is constituted by two highly energetic muons originating from the Z-boson decay and two jets from the Higgs boson decay, should be considered as a possible background source. The studies exposed in [Kuhl 07], showed that the requirement for a muon pair, where each muon has at least 15 GeV/c of momentum, suppress almost completely the majority of the potential background. Significant background contributions could arise from the following processes:

qq: The requirement for two high momentum muons will suppress very effectively the background contamination coming from this channel. Some contribution may nevertheless arise from the missidentification of the final state pions as muons.

The excellent purity of the Pandora particle flow algorithm (see table 5.3) should eliminate this effect.

WW: The requirement for 2 high momentum muons should marginalise the contribution that arises from the fully hadronic channel. Additionally, it is expected to suppress in a large extent the semi–leptonic channel. However, the semi–leptonic channel $WW \rightarrow q\overline{q}\,\mu\,\nu_\mu$ can produce a final state whose topology resembles to the signal one, having two jets and a high energetic muon in the final state (plus missing energy from the neutrino). A pion misstagged as a muon (even though this is not very probable to happen - see table 5.3), or a semi–leptonic heavy flavour hadron decay could add a second energetic muon to the final state. This will result in a final state similar to the signal one. Since this channel has a relatively large cross–section, it is a potential source of background. Its cross–section can be suppressed by choosing the appropriate electron and positron polarisation, 80 % for e^-_R and 30 % for e^+_L, see table 5.2. Even though this selection reduces the signal cross–section by $\sim 30\,\%$, it reduces by at least one order of magnitude the WW cross–section. Hence this is a strong argument to choose this set of polarisation. To estimate the signal contamination that could arise from this channel, we have fully reconstructed 10^4 events, corresponding to an integrated luminosity of $31.3\,\mathrm{fb}^{-1}$. Only one event was found to pass the selection criteria, therefore we may assume the signal contamination introduced by this channel as marginal. Of course, in this analysis of 10^4 events, we have to stress the small size of the sample which introduces large statistical uncertainties regarding the background contribution of this process. But our assumption can be verified by the ones obtained for the LoI studies. Therefore, the effect of this process will be assumed marginal.

ZZ: The fully hadronic channel, as well as the semi–leptonic channel $ZZ \rightarrow q\overline{q}\nu\nu$, will be eliminated by requiring a high momentum muon pair. On the other hand, the semi–leptonic channel where $ZZ \rightarrow \mu^+\mu^- q\overline{q}$ has exactly the same topology as the signal. Thus it is expected to provide the largest background contribution. This background should be fully understood and controled. A sample corresponding to $250\,\mathrm{fb}^{-1}$ has therefore been passed through full detector simulation and reconstruction.

single Z: The single Z production is the second most important background source, even though its contamination is estimated to be at least one order of magnitude smaller than the one of the double Z process ([Kuhl 07]).

single W: The background contribution of this channel is expected to be negligible.

The simulation and reconstruction of all background processes for each candidate VXD geometry, and the relevant event selection optimisation, would require vast com-

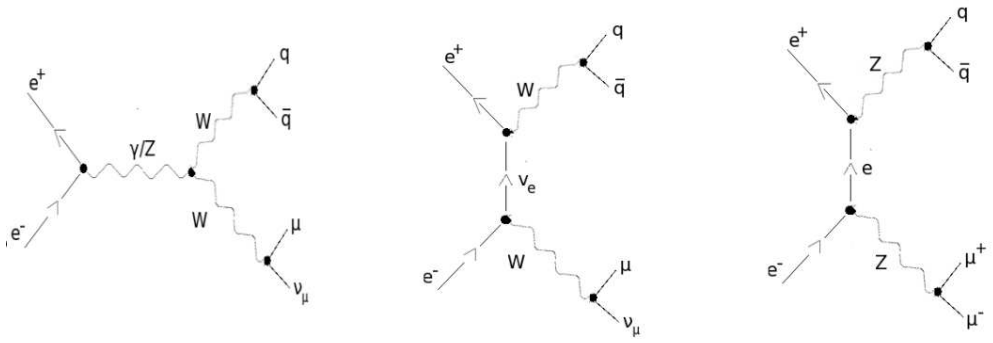


Fig. 5.2: First order Feynman diagrams of the main background processes. The left shows the s–channel of the WW production; the middle shows the t–channel of the WW production; the right shows the ZZ production.

puter resources. As already mentioned in the introduction, we should stress that we do not intent to repeat the physics analyses done for the LoI. That task would encounter very big practical difficulties, without adding an important contribution. Thus, we will be guided by the results obtained for the LoI, focus on the dominant background channels, and be inspired by the suggested event selection variables. The reason is, apart from the technical ones already mentioned, that we do not expect that the different VXD geometries will have a sizeable effect on the signal–background separation of the minor background processes. Therefore, for the purpose of the studies exposed in this chapter, only the contamination due to the double Z process will be examined, since it is by far the most important. We consider that the signal contamination from other processes will be effectively suppressed to marginal levels, by requiring the event selection that was used in [Kuhl 07].

Table 5.2 summarises the signal and background processes that will be studied, together with their corresponding cross–sections for the assumed polarisation, as well as the number of events that have been fully simulated and reconstructed for each of them. The Feynman diagrams for these processes are illustrated in figure 5.2.

5.4 Detector Simulation

The detector has been fully simulated using the GEANT4 [Geant4] based software Mokka [Mokka]. Mokka uses the geometry information stored in a set of MySQL [MySQL] databases, which is translated by a set of C++ drivers into GEANT4 objects.

The overall detector simulation model is being built from a combination of a number

process	Signal	ZZ	WW
final state	$\mu^+\mu^- q\overline{q}$	$\mu^+\mu^- q\overline{q}$	$\mu\nu_\mu q\overline{q}$
polarisation	$e_L^-80\%, e_R^+30\%$		
cross–section (fb)	10.33	124.03	5338.87
polarisation	$e_R^-80\%, e_L^+30\%$		
cross–section (fb)	7.02	71.89	319.42
Events	1750	17975	10000

Tab. 5.2: Cross–sections and number of events fully simulated and reconstructed for the signal and the main background processes. The number of reconstructed events corresponds to an integrated luminosity of 250 fb^{-1} for the signal and ZZ processes, and to 31.3fb^{-1} for the WW production for the assumed beam polarisation of $e_L^-80\%, e_R^+30\%$.

of subdetectors, like for example the TPC, the ECAL, the magnetic field or the beam pipe. In order to build a subdetector, the relevant database should be queried by the relevant driver. This scheme introduces a great deal of flexibility into the optimisation procedure, since a driver can query various databases. For example, a change in the dimensions of a module could be tested just by modifying the relevant parameter in the database, without touching the code. This also facilitates the optimisation process, since it is quite straightforward to test several overall detector options by adding, removing or substituting subdetectors during run time through steering files.

On the other hand, the increased amount of flexibility might become slightly dangerous, and result in inconsistencies like overlapping volumes. A workaround is offered by the superdrivers concept; these are self–scaling drivers that can scale the whole detector in length and radius in order to avoid possible overlaps.

Mokka takes as an input a generator file, usually in *stdhep* format, and simulates the interactions with the full detector. The output file include the hit collections created at the several subdetectors during the simulation. They have the LCIO [LCIO] persistency format. The LCIO is the event data model used for the linear collider studies. If Mokka is compiled with the GEAR [GEAR] package, it provides an *xml* file that holds geometrical parameters, essential for the reconstruction procedure. Figure 5.3 illustrates the software organisation. The analyses presented in this chapter have been realised with the ILD_00 detector model, which is the model used for the LoI studies. Then, depending on the desired VXD geometry, the relevant VXD subdetector variant was substituted.

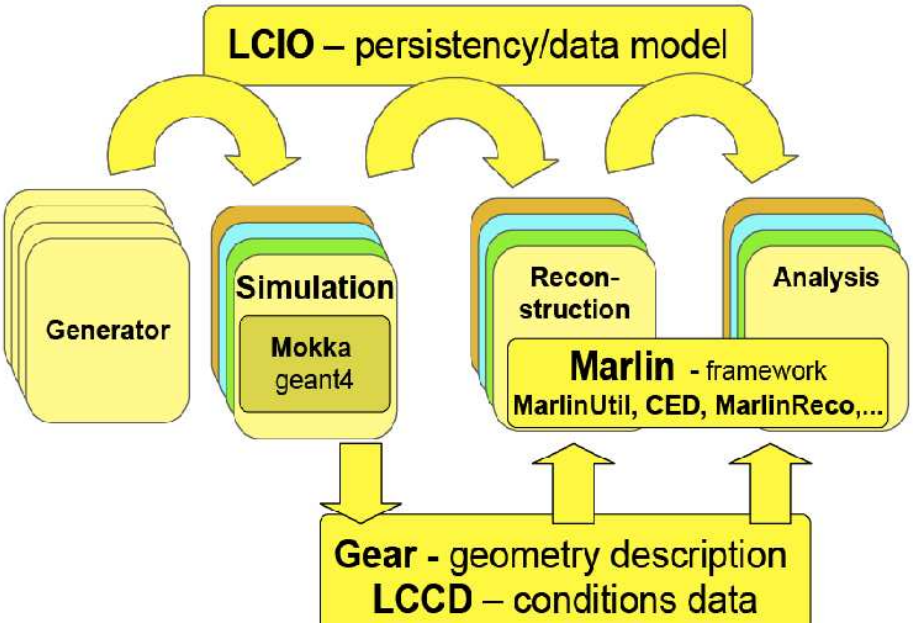


Fig. 5.3: Schematics illustrating the ILD software organisation [Gaede 10].

5.5 Event Reconstruction

The reconstruction chain is handled by Marlin [Marlin], which is a modular C++ application framework. Each task during the reconstruction procedure is fulfilled by a dedicated module, called processor. An *xml* steering file defines which processors will be called during reconstruction, and in which order, as well as the values of their parameters. Marlin uses the LCIO format for its input and output data, but also as a transient data model. The main part of the reconstruction, meaning the digitisation, the tracking and the clustering (see section 2.4) is realised via MarlinReco, which is a Marlin based package that provides the relevant algorithms. The resulting tracks and clusters are fed to the particle flow analysis, provided by the PandoraPFA software [Thomson 09].

The reconstruction of our data samples will proceed accordingly through the following steps. At the first stage, the events are scanned for $\mu^+\mu^-$ pairs. The identification of the muons is provided by the PandoraPFA. The PandoraPFA performance on muon identification can be quantified by calculating the relevant efficiency and purity. The

efficiency is defined as the ratio of the true muons that have been identified versus the total number of true muons. The purity is the ratio of the number of the true muons that have been identified versus the number of reconstructed particles that have been identified as muons. We have calculated these ratios for our signal sample, and the results are summarised in table 5.3. Very high purity is provided. On the contrary the efficiency is more moderate, and is going to introduce a statistical loss: since two high energy and opposite charged muons are needed for the Z reconstruction, with the provided efficiency of $\sim 90\%$, about 20% of the signal events are expected to be rejected.

PandoraPFA performance	
efficiency	90.6%
purity	99.9%

Tab. 5.3: Performance of the PandoraPFA algorithm on the muon identification for the signal sample.

Out of all the possible $\mu^+\mu^-$ combinations, the one most susceptible to be a Z boson is selected for the Z reconstruction. This *best* candidate is defined as the pair having its invariant mass (IM) closest to the Z boson mass, as it is reported in [PDG b]. A first set of selection criteria (see section 5.6) is applied to this candidate. If they are satisfied, we proceed to the jet clustering. The selected muon pair is then removed from the particle's collection. The rest of the reconstructed particles of the event are forced to jet reconstruction using the Durham k_T algorithm [Catani 92], requiring a fixed value of 2 jets. If the jet clustering succeeds, the jet–wise flavour tagging provided by the LCFI collaboration (see section 5.7) is applied on them.

5.6 Event Selection

A pure sequential cut based event selection is used. The cut variables and tuning are inspired from the LoI analyses that used the Higgsstrahlung process in order to measure the Higgs boson mass, the production cross section and the branching ratios [Kuhl 07], [Ito 09]. The cuts are applied at two stages. The first stage is after the Z boson reconstruction where an event selection takes place. The selected events undergo the jet clustering. Then a second set of cuts is applied to the reconstructed jets. The selection criteria are the following:

1. **Two high momentum muons:** As already mentioned in section 5.3, the requirement for two high momentum muons effectively suppresses the contamination

coming from the $q\overline{q}$ production, and from the fully hadronic decays of the ZZ and the WW final states.

2. **$70 \leq Z_{IM} \leq 110$ GeV :** The cut applied here demands that the invariant mass of the muon pair should be consistent with the invariant mass of the Z boson. That is expected to eliminate the contamination arising from the semiloptenic channel of the WW process. After the first two cuts, the leftover background contribution mostly comes from the double Z boson production, which is obviously not affected by these cuts. So, from now on, we will focus on the double Z boson background. The upper left plot of figure 5.4 shows the Z boson invariant mass distribution for the signal and the ZZ background. There are a large number of entries towards the low mass tails of these distributions. Their source is the misefficiency in the muon identification where one or both of the muons coming from the Z decay are not identified. In this case the Z boson is reconstructed from a muon pair where either one or both muons are not really the decay products of the Z. The vast majority of those events are rejected by this cut.

3. **$117 \leq$Recoil Mass≤ 150 GeV :** The aim of the last three cuts is mainly to discriminate the Higgsstrahlung from the double Z boson production process. The variable with the highest separating power between these two processes is the recoil mass of the Z system. Its distribution is illustrated in the upper right part of the figure 5.4. The recoil mass is calculated using the following equation:

$$\text{Recoil Mass} = \sqrt{s + Z_{IM}^2 - 2\sqrt{s}(E_1 + E_2)} \tag{5.1}$$

where $\sqrt{s}$ is the center–of–mass energy, Z_{IM} is the reconstructed invariant mass of the Z boson and E_1, E_2 the energies of the two muons used for the Z reconstruction. Concerning the signal, a very sharp peak at $120\,\text{GeV/c}^2$ is clearly visible, which is consistent with the generated mass of the Higgs. The distribution of the ZZ process has a peak around $91\,\text{GeV/c}^2$, the Z invariant mass.

Both histograms have a large number of entries in the area of large recoil mass values. These high tails of the distributions are due to radiative effects. The beamstrahlung degrades the center–of–mass energy, thus increasing the recoil mass. This is explained from the fact that in order to calculate the recoil mass, using the equation 5.1, we assume a $\sqrt{s}$ to be equal to 250 GeV. If its real value is smaller, equation 5.1 will lead to an overestimation of the recoil mass. The Initial State Radiation (ISR), having a similar effect as the beamstrahlung, will also contribute to the high recoil mass values. Additionally, it may reduce the production cross–section if a very high energy photon ($P \geq 40\,\text{GeV/c}$) is created (therefore the Higgsstrahlung process will not be kinematically accessible). Moreover, the Final State Radiation (FSR) will add some entries to the high

tails of the recoil mass distribution. The reason is that the final state muons are expected to radiate bremstrahlung photons, loosing energy this way. But given the relatively large mass of the muon, bremstrahlung is expected to be highly suppressed. Thus the contribution of the FSR will only play a marginal role. Finally, we observe that there is a significant decrease of the events at the high tails after 159 GeV/c^2. This is the value of the recoil mass at the mass threshold of the Higgsstrahlung process.

4. **cos $(\theta) \leq$ 0.9 :** The Higgsstrahlung being a pure s–channel process, the reconstructed Z boson is expected to have a uniform polar angle distribution. The ZZ production being essentially a t–channel process, the angular distribution is expected to peak at small polar angles. The down left plot of figure 5.4 is illustrating the above, providing another discriminating variable for the two processes.

5. **100 $\leq$ di–jet IM $\leq$ 140 GeV :** The final cut is applied to the reconstructed di–jet invariant mass. For the signal, the distribution has a peak near the generated Higgs boson mass, while for the ZZ background it peaks near the Z boson invariant mass.

Table 5.4 summarises the numbers of events for the signal and the background that survive after the application of each cut. The presented results were obtained with the VXD03 geometry after the analysis of an integrated luminosity of 250 fb^{-1}, both for the signal and the background. Almost identical results were found with the VXD05 geometry. Finally we end up with a significance of the signal of $\frac{S}{\sqrt{S+B}} = 24.0 \pm 0.9$.

Cut	signal	ZZ
Initial event sample	1750	17975
1	1607 (91.8%)	15883 (88.4%)
2	1349 (77.1%)	10064 (56.0%)
3	1258 (71.9%)	1152 (6.4%)
4	1205 (68.9%)	1061 (5.9%)
5	924 (52.8%)	556 (3.2%)

Tab. 5.4: Cut flow for the signal and the ZZ background.

5.7 Flavour Tagging

The physics reach of an experiment is based on the detector performance as well as on the maturity of the reconstruction algorithms. Thus, before proceeding to the

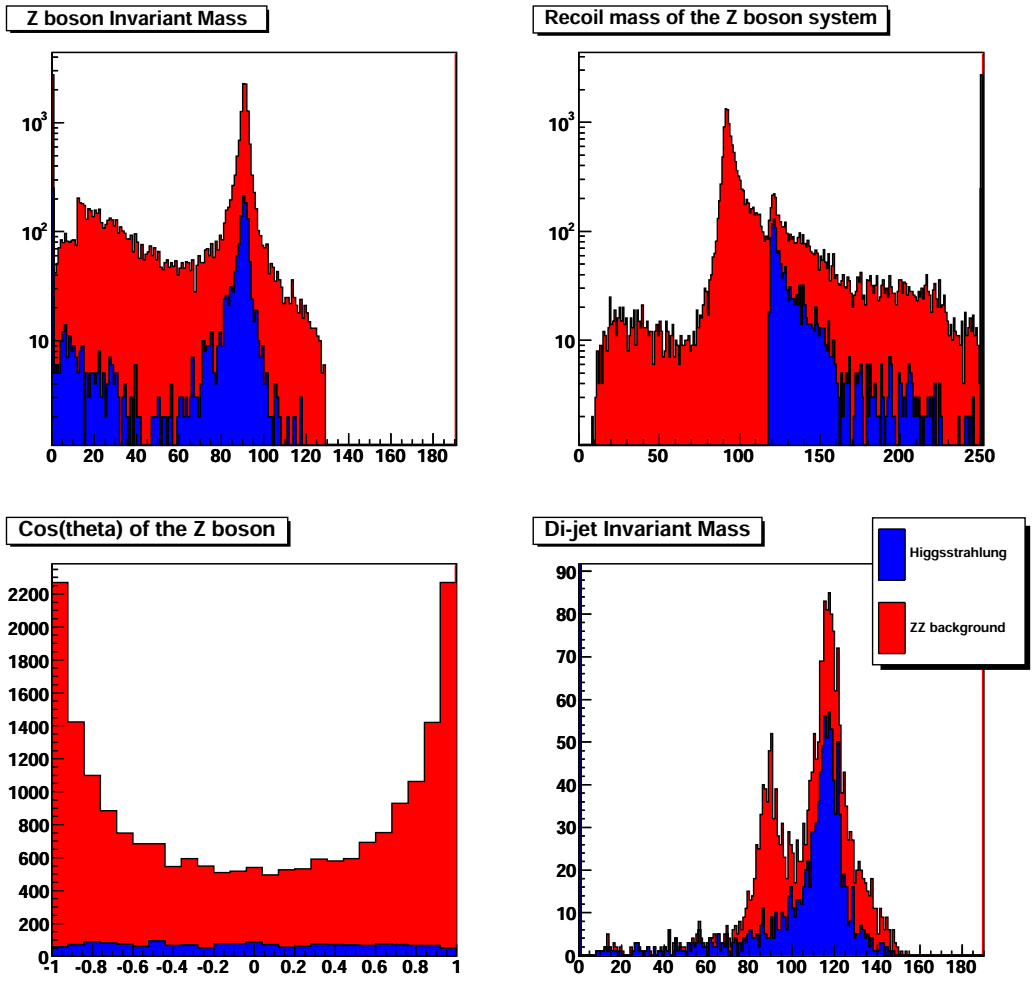


Fig. 5.4: Main discriminating variables. The plots filled with blue colour represent the signal and the ones filled with red the ZZ background distributions. The upper left displays the Z boson invariant mass reconstructed from the best candidate muon pair. The upper right shows the recoil mass of the Z boson system. Down left is the cosine of the Z polar angle. Down right is the di–jet invariant mass, being calculated after applying the cuts on the three other variables (invariant mass, recoil mass, polar angle).

study of the VXD performance, we will introduce the algorithms, used by the ILD collaboration, that provide the heavy flavour tagging. The scope of section 5.7, more precisely sub–sections–5.7.1 and 5.7.2, is not to fully describe these tools, but rather to make an introduction, in order to facilitate the understanding of the optimisation procedure. The identification of the jet's flavour is derived from the neural net based

flavour tagging developed by the LCFI collaboration [Bailey 09]. We should stress here that the secondary vertex reconstruction is jet–based, and not event–based. The task of vertex reconstruction is fulfilled with the topological vertex finder ZVTOP [Jackson 97], an algorithm originally developed for the Stanford Linear Detector (SLD) and currently adapted to the ILC by LCFI.

Two cases are distinguished, which lead to two different sets of input variables to the neural nets. In case *a*, only the primary vertex has been found, while in case *b* displaced vertices have also been found inside the jets. For this latter case, the information derived from the additional vertices is crucial. It is thus very important to suppress the reconstruction of fake vertices that could stem from wrong track associations. Vertices coming from photon conversions or K_S and Λ decays should also be removed, since they may confuse the flavour tagging. Those unwanted vertices can be effectively suppressed by an appropriate track selection. Details about the selection criteria and their performance in terms of photon conversion reconstruction, and K_S and Λ identification can be found in [Bailey 09].

5.7.1 The ZVTOP Vertex Finder

The ZVTOP software package contains two complementary algorithms for topological vertex reconstruction, called ZVRES and ZVKIN. ZVRES is used to reconstruct multiprong vertices while ZVKIN is specialised for the case of one prong vertices. Here we are going to give a brief description of the ZVRES algorithm. For further information one can refer to [Jackson 97]. The main idea of the algorithm is to construct a function $V(\vec{r})$, which describes the vertex probability at the position $\vec{r}$. In order to construct this function, each track is ascribed a probability tube in space, obtained from the track helix parameters. This Gaussian probability tube is defined from the following function:

$$f_i(r) = exp\{-\frac{1}{2}(\vec{r}-\vec{p})V_i^{-1}(\vec{r}-\vec{p})^T\} \tag{5.2}$$

where $\vec{p}$ defines the point of closest approach of the track to the point defined by $\vec{r}$, and V_i is the position covariance matrix of the track at $\vec{p}$. The basic vertex density function is derived by combining the track probability functions, and is described by the following equation:

$$V(\vec{r}) = \sum_{i=1}^{N} f_i(r) - \frac{\sum_{i=1}^{N} f_i^2(r)}{\sum_{i=1}^{N} f_i(r)} \tag{5.3}$$

where N is the number of tracks. The function $V(\vec{r})$ will acquire high values in space regions where several tracks overlap, thus $f_i(r)$ has also a high value. On the other hand, in regions where the function $f_i(r)$ has an important value for ≤ 1 track no

vertex should be found, and $V(\vec{r})$ will be close to 0. Out of the vertex density a number of secondary and tertiary vertices can be reconstructed. Additional information, like the position of the Interaction Point (IP), could be added to the expression of $V(\vec{r})$ in order to suppress the probability of the reconstruction of fake vertices.

The same function is used in order to decide whether to merge or to resolve two reconstructed vertices. The resolution of the vertices at two points, defined by $\vec{r_1}$ and $\vec{r_2}$, is based on the following criterion:

$$\frac{min\{V(\vec{r})\} : \vec{r} \in \vec{r_1} + \alpha(\vec{r_2} - \vec{r_1}), 0 \leq \alpha \leq 1}{min\{V(\vec{r_1}), V(\vec{r_2})\}} \leq R_0 \tag{5.4}$$

where the denominator is the lowest of the two values $V(\vec{r_1})$, $V(\vec{r_2})$. The numerator is the minimum value of the vertex function along the straight line that connects the space points defined by $\vec{r_1}$ and $\vec{r_2}$. R_0 is a tunable parameter which defines the number of reconstructed vertices.

5.7.2 Neural Nets Flavour Tagging

The ILC experiments use the neural nets based heavy flavour tagging developed by the LCFI collaboration. One of its main features is that it uses a different set of input variables, depending on the jet's vertex multiplicity.

If only the primary vertex has been found, then the following set of 8 input variables is used:

- The impact parameter significance in $r-\phi$ of the first and second most significant tracks of the jet
- The impact parameter significance in z of the first and second most significant tracks of the jet
- The momentum of the first and second most significant tracks of the jet
- The joint probability in $r - \phi$ and z that all the tracks in the jet belong to the primary vertex

The term significant means the impact parameter significance of the track. It is defined as b/σ_b, where b is the distance of the closest approach of the track to the IP, and σ_b its uncertainty.

Heavy flavour hadrons usually travel a short distance of a few tenths or hundreds (depending on their flavour) of micrometers before decaying. Therefore, tracks originating from them have a positive impact parameter. The sign of the impact parameter is defined by the intersection of the track with the jet axis. If it is in front of the

primary vertex it is defined as positive, otherwise it is negative. The impact parameters of the tracks that originate from the IP feature a symmetric distribution, centered at zero. Tracks with negative impact parameter reflect an artefact, expressing the finite resolution of the detector. The probability P_i that the track i, having an impact parameter significance b/σ_b, originates from the IP is:

$$P_i = \frac{\int_{b/\sigma_b}^{\infty} f(x)dx}{\int_0^{\infty} f(x)dx} \tag{5.5}$$

where $f(x)$ is the distribution of impact parameter significances for IP tracks. The function $f(x)$ can be defined from the data. The negative part of the impact parameter significances corresponds almost uniquely to IP originated tracks. Thus, by fitting it, we can extract the shape of the $f(x)$ for the IP originated tracks. Using the joint probability definition, the probability that N tracks are coming from the IP is given by:

$$P_J = y \sum_{m=0}^{N-1} \frac{(-\ln y)^m}{m!} \tag{5.6}$$

where $y = \prod_i P_i$. The value of P_J for light jets is close to 1, while for b and c jets it is close to 0. Joint probability is the observable with the highest discriminating power for flavour tagging when the jet contains no displaced vertices.

The impact parameter significances of the most significant track of the jet are a crucial input for charm jet identification. About 40% of the D^+ hadrons are expected to decay to only one charged track. In this case no secondary vertex will be formed, but a single charged track with high impact parameter. On the other hand, when an event contains a B hadron decaying in one charged track track followed by an one prong decay of the subsequent D hadron, the expected signature is 2 high impact parameter tracks. Thus the impact parameter significance of the second most significant track of the event could be used to discriminate b– from c–jets.

If at least 1 displaced vertex is found inside the jet, we try to exploit the displaced vertex information. The following set of input variable is used:

- The decay length and the decay length significance of the secondary vertex
- The sum of the momenta of all the tracks belonging to this secondary vertex
- The P_T corrected mass of the secondary vertex
- The track multiplicity of the secondary vertex
- The secondary vertex probability

- The joint probability in $r-\phi$ and z that all the tracks in the jet belong to the primary vertex

The highest discriminating power is provided by the P_T corrected mass of the secondary vertex. The calculated invariant mass of the secondary and potential tertiary vertices, provides a very powerful discrimination between the b– and the c–jets. The applied correction takes into account the missing mass due to neutral particles. This is achieved by using the difference in the direction of the vector of the resultant of the vertex momenta and the direction of the vertex axis. The latter is defined as the vector which points from the primary to the secondary vertex.

5.7.3 Neural nets training

The b(c)–jet identification will be based on a dedicated set of b(c) tagging neural nets. The b–tagging neural nets will consider as signal the b–jets and as background the light and the c–jets. Alternatively, the c–tagging neural nets will consider as signal the c–jets and as background the light and the b–jets. As it was mentioned before, the set of input variables depends on the outcome of the search for displaced vertices inside the jet. If such vertices are found, the distributions of the input variables are quite different according to the following two cases: a. only one secondary vertex has been reconstructed and b. more than one secondary vertices have been reconstructed. So finally, three different sets of b(c) tagging neural nets are trained, one for each of the following jet subsamples:

- Only the primary vertex has been reconstructed
- The primary and only one secondary vertex has been reconstructed
- The primary and more than one secondary vertices have been reconstructed

The neural nets used for the ILD LoI studies have been trained with the $e^+e^- \rightarrow Z \rightarrow q\,\overline{q}$ sample at $\sqrt{s} = 91.2\,\text{GeV}$. We are going to use the same channel for our study. This process provides jets with an average energy of 45 GeV. This value is quite close to the average jet energy expected from the Higgsstrahlung process at $\sqrt{s} = 250\,\text{GeV}$. A dedicated training will be performed for each of the two main VXD geometries, the VXD03 and the VXD05 options. Consequently, a separate fit of the negative impact parameter significances of the tracks will be performed for each of them. Generally speaking, each time we make a modification on the VXD geometry or its technological parameters, it would be preferable to use a dedicated, tuned set of neural nets. Therefore, inside the framework of the VXD optimisation studies, we will perform a dedicated training for each modification under study. For the purpose of these studies, 10^3 events of $Z \rightarrow b\overline{b}$, 10^3 events for $Z \rightarrow c\overline{c}$ and 10^3 events for $Z \rightarrow$

light jets have been fully simulated and reconstructed, for the purpose of the neural nets training.

The architecture of the neural nets is a multilayer perceptron. It has 8 inputs, a hidden layer with 14 nodes and one output. A possible pitfall of the neural nets training procedure is the overtraining. After too much training, the neural nets become too specific to the training sample, including its statistical features. So they rather memorise the training data sample, than learn from it. This leads to poor predictive performance towards an independent sample, meaning that these neural nets are rather useless for any application. So after the training, the possibility of overtraining has been checked. The figure 5.5 illustrates the output of the b-tagging nets for a pure b sample, independent from the one that has been used for the training. A very sharp peak is observed at 1. On the other hand there is a small accumulation of events near 0. However, the overall very good performance towards an independent sample is a clear evidence that no overtraining took place, at least at a significnt level.

A second aspect that should be checked are the training uncertainties. It is not guaranteed that a neural net will yield the maximum performance after one training. It is possible that if we train again our neural nets with the same data sample, they may yield a slightly different performance [Walsh 09]. Hence the training uncertainties should be estimated, and compared with the statistical uncertainties of the data sample. In order to check this, we trained five times the same set of neural nets and compared the dispersion of the training outcome with the statistical uncertainties of the sample on c–tagging. The dispersion of the neural nets training has been estimated from the rms of the purities obtained from each independent training. In figure 5.6, the relative dispersion is compared with the relative statistical uncertainties, for c–nets thresholds that vary from 1 % to 100 %, in steps of 1 %. It is clear that the uncertainties due to the finite statistics of the data dominate over the neural nets training uncertainties, except a very narrow zone near 0.

5.8 Flavour Tagging Performance

Now we can proceed to the analysis part. The goal of this section is to evaluate the performance of the VXD in flavour tagging, as a function of the geometry and various technological aspects. The test sample is the Higgsstrahlung process. The physics background does not play any role in this study, thus will be neglected. However, in order to obtain a similar jet sample with the one we are going to use for the Higgs branching ratios extraction, the event selection described in the section 5.6 will be applied also here. The performance can be characterised from the efficiency and the purity on b– and c–tagging. The efficiency is standing for

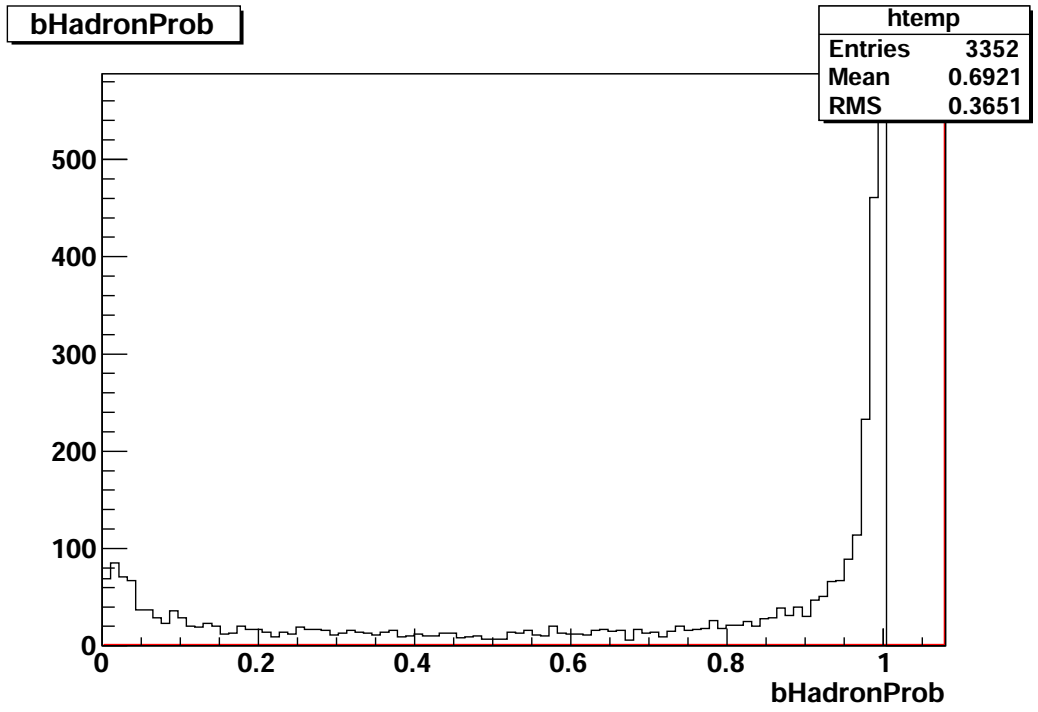


Fig. 5.5: Output of the b-tagging algorithms for an independent pure $h \to b\overline{b}$ sample. A sharp peak is observed near 1, while there is a small accumulation of events near 0. This is a strong indication that no overtraining occurred.

$$\text{Efficiency} = \frac{\text{b(c)-tagged jets} \cap \text{true b(c)-jets}}{\text{true b(c)-jets}} \tag{5.7}$$

and the purity is defined as follows

$$\text{Purity} = \frac{\text{b(c)-tagged jets} \cap \text{true b(c)-jets}}{\text{b(c)-tagged jets}} \tag{5.8}$$

A jet is tagged as b(c) if the b(c)–tagging probability, given by the output of the b(c)–nets, is higher than the applied cut. The value of this cut depends on the requirements of each specific study and defines the efficiency and purity of the selected b(c) sample. In general it goes like that: the stricter the cut, the higher the purity, the lower the efficiency. To plot the efficiency–purity curves, a scan is applied on the cut value, going from 0 to 1 in 100 steps. The term true b(c)–jet standing obviously for a jet initiated by a b(c)–quark. The determination of the true jet flavour is provided by the TrueAngularJetProcessor of the LCFI vertex package. This processor has access to the Monte–Carlo (MC) generator information. It determines the true jet flavour by angular matching of heavy quarks to the jets. The jet is assigned the flavour of the

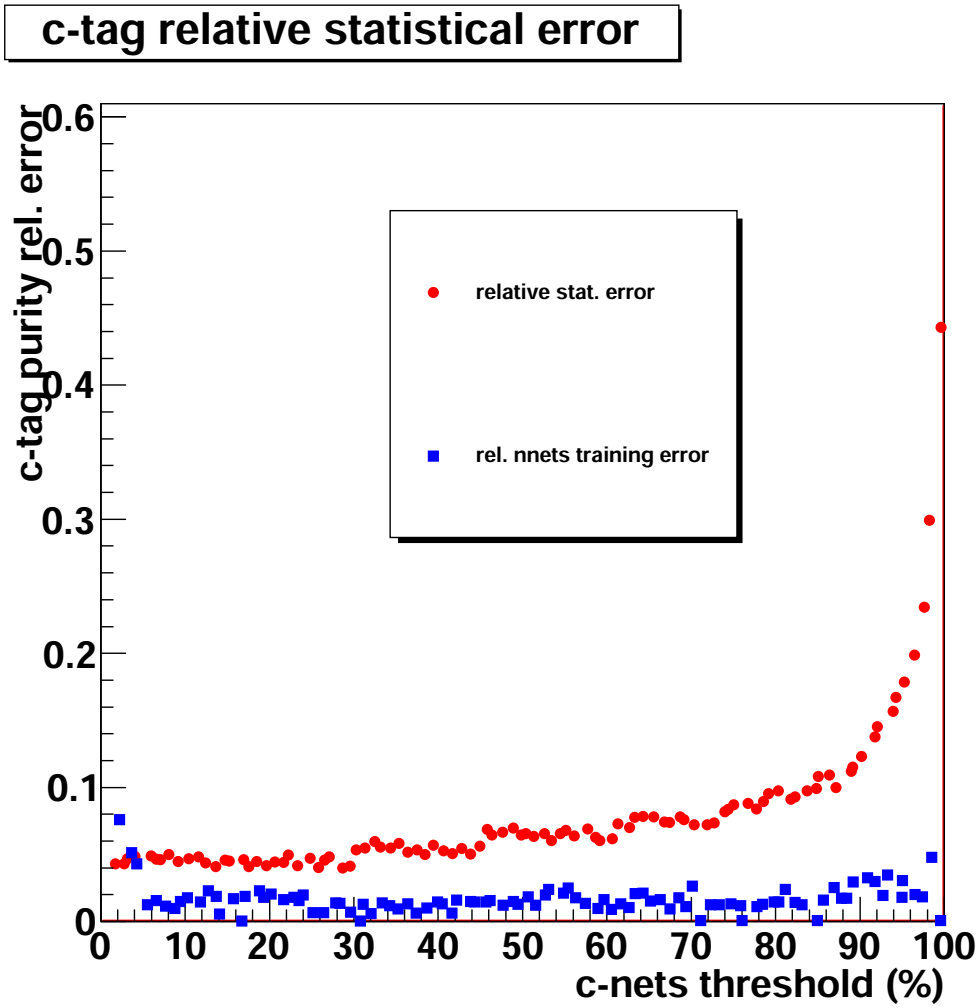


Fig. 5.6: C–tagging purity uncertainties: comparison between relative statistical uncertainties due to the finite statistics of the data sample (red full circles) and the relative variance of the neural nets performance after 5 independent trainings (blue full squares).

heaviest quark of its hadron that appears first in the MC decay chain. In the case of two jets events, like in our study, it is quite clear which hadrons belong to which jet. For multijet events, some ambiguities may arise.

We are going to focus mainly on the performance comparison of the two main candidate VXD geometries, VXD03 and VXD05. Inside the framework of the VXD optimisation, modifications of these two basic models will be also studied. The study will be based on the reconstruction tools provided by the MarlinReco package that were described in the section 2.4. It was already mentioned that there is not yet

any specific technology chosen for the VXD sensors. For the purposes of the study, a very generic one will be assumed. The digitisation of the hits will be based on the default digitisation algorithms of MarlinReco, meaning that there is no pixelisation. The assumed sensor's single point resolution, unless otherwise stated, is 2.8 μm for every layer of every model. The hit finding efficiency is assumed to be 100 % for every layer.

Using the LoI reconstruction framework, which nevertheless was the optimal one by the time those studies were conducted, imposes certain constraints on the VXD performance. This is true in particular for the silicon tracking algorithm. When one discusses the VXD performance and optimisation, the pair background effect is an essential input. However the standalone silicon tracking algorithm, as was explained in sub–section 2.4.3, it is not mature enough to cope with it. For this reason, during the first stage of our studies, we merely prompt into the geometrical and technological features of the VXD, without taking into account the beam background. At a second stage, the beam background hits are included using the VTXNoiseClusters processor (see sub–section 5.8.3). It is expected that the silicon tracking will be optimised in the near future, in order to cope with the beam induced background.

5.8.1 Comparison of the Main Alternative Geometries

The analysis has been performed with a Higgsstrahlung event sample equivalent to 857 fb^{-1} of integrated luminosity, $e^+e^- \rightarrow ZH \rightarrow \mu^+\mu^- X$. This integrated luminosity is translated to 6000 events. We search for any possible striking performance differences, that could indicate which VXD geometry is best suited to the physics objectives.

The purpose of this study is to serve as the first step in the comparison of the performance of the VXD alternative geometries. Effects like the beam background, or possible ladder or sensors misalignments are not included at this stage. In its next steps, it should be complemented by adding more realism.

The figure 5.7 presents the efficiency - purity plots for b and c–tagging of the VXD03 and the VXD05 models. Both geometries demonstrate excellent b–tagging, and satisfactory c–tagging performance, that meet the ILD requirements. The observed differences between the performance of the two geometries are well inside the limits of the statistical uncertainties, except of a narrow region in the range of high efficiency–moderate purity of c–tagging. There, the difference in purity reaches a value ~ 0.03 for an efficiency of 0.65 - 0.70. This corresponds to ~ 15 % higher purity for VXD05 which appears only in this narrow efficiency range. However, the difference remains marginal, and both single and double layers geometries can be claimed having similar potential in terms of heavy flavour tagging.

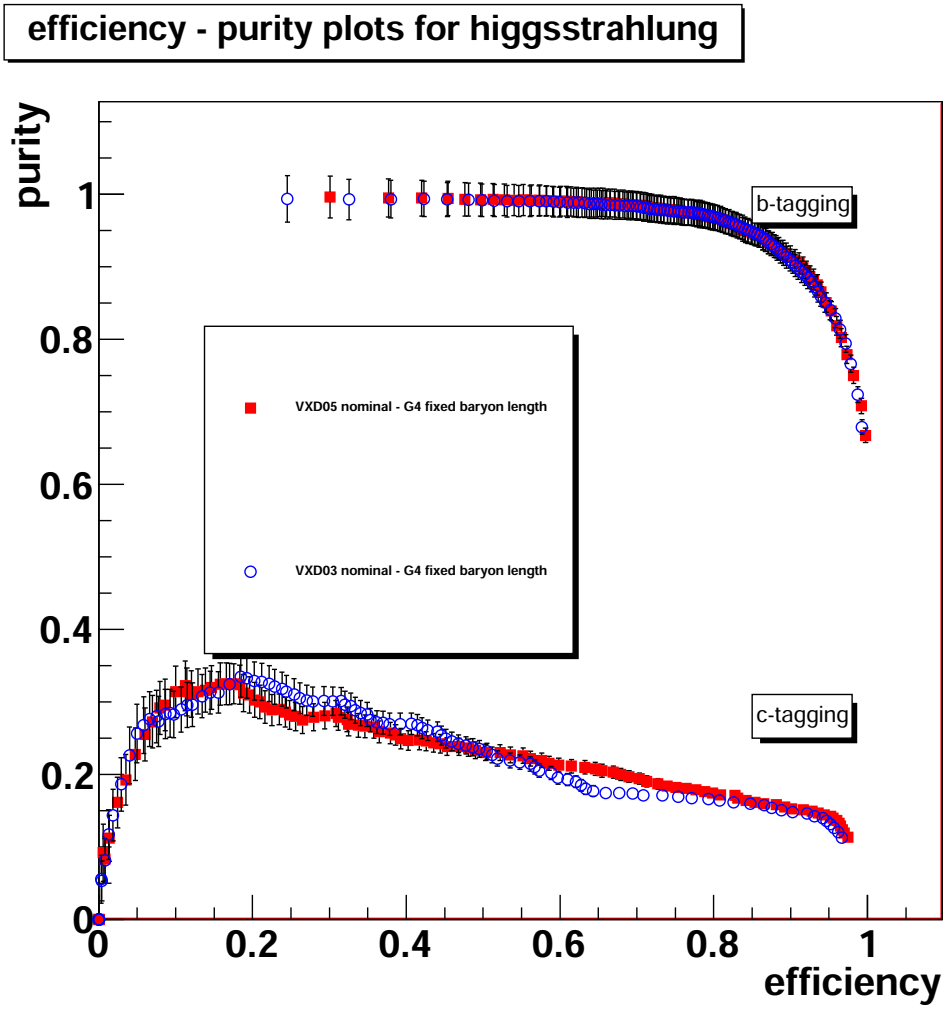


Fig. 5.7: Purity versus efficiency for b- and c-tagged jets for each geometry. The red full squares stand for the VXD05, the blue circles for the VXD03. Only statistical uncertainties are shown.

5.8.2 Vertex Detector Optimisation Studies

The flavour tagging performance of the VXD is quite sensitive to aspects such as its geometry and its technological features, the running conditions of the experiment and the reconstruction algorithms. In this section, we will focus on the geometry and the technological features, and will try to study the performance of the VXD as a function of them. There are several motivations to scan partially the parameter space of the geometry and technology features. The most important are summarised in the following list:

- To see whether the performance of the VXD can be further optimised.
- To complement the study presented in the previous section by adding more realism. During this stage, the added realism is relevant to the sensors specifications: the emergence of a sensor technology that will satisfy all the requirements stated in sub–section 2.3.2 in the near future is not certain. Thus the assumed parameters, on which the results presented in figure 5.7 were based, may change. It is therefore important to find out in which margin these parameters may change, without degrading significantly the performance.
- To test specific aspects of the CMOS based VXD.

For practical reasons (time, computer resources), it is unrealistic to scan the whole parameter space. We are therefore going to address the following aspects:

1. The distance from the IP of the innermost layer
2. The sensor's single point resolution
3. The material budget
4. The detection efficiency of the sensors

Study of the Inner Layer Radius

The distance of the innermost layer from the IP (or inner layer radius–R_{int}) is one of the most important parameters of the VXD design. The smaller R_{int} the better, since the extrapolation error of the track to the IP increases with the distance. Both a and b parameters of the impact parameter resolution, see equation 3.1, depend on R_{int}. The effect is expected to be bigger on the parameter b, since it increases linearly with the distance from the IP (see equation 3.3). The lower limit on R_{int} is set by the beam pipe radius, see figure 2.6, which is 14 mm. How close the inner layer of the VXD can go to the beam pipe will be finally determined by the acceptable level of the beam background induced occupancy (see sub–section 3.1.2), and by system integration aspects.

The two VXD options VXD03 and VXD05 aim to go the closest possible, featuring innermost layer radii of 15 mm and 16 mm respectively. It is not certain that the occupancy will finaly allow us to reach these values. They cannot be accurately predicted, since they result from a convolution of several factors. To remind some them, these are the beam background hit density (which is estimated through simulation studies), parameters of the sensor technology (which is not yet determined) like the time resolution or the dimension of the pixels, and overall detector parameters, like the magnetic field configuration. Here we are going to focus on the latter, and more specifically on

the presence of the anti–DID field. The implementation of the anti–DID field (see subsection 2.2.2) is still under consideration. If it will not be implemented, the hit density will be substantially increased, especially for small values of the azimuthal angle ϕ and - in a small extent - at the ends of the ladders of the innermost layer (see figures 2.10 and 2.9). This will may lead us to move the innermost layer away from the IP and place it at an inner radius of $\sim 20\,\mathrm{mm}$.

It is thus important to study the effect of the innermost layer radius on the VXD performance. It can be used as an argument, whether the anti–DID field should be implemented or not. For the purposes of the study, we will create a modified version of the VXD05 geometry (called VXD05_rint20 hereafter), where the inner layer radius will be moved from 16 mm to 20 mm.

First we will compare the impact parameter resolution of the VXD05 and the VXD05_rint20 models. From equation 3.1, the effect of the increased R_{int} can be calculated analytically. It is expected to degrade the parameters a by $\sim 12\%$ and b by $\sim 25\%$. In order to compare with the nominal VXD05 geometry, we will proceed to a simulation study. We will estimate the impact parameter resolution of muons for varius values of P_T, emitted at a polar angle of $\theta = 85^0$. The P_T spectrum will cover the area between 300 MeV/c to 200 GeV/c. The results are displayed in figure 5.8.

The degradation of the b parameter by 25 % is more pronounced for low P_T values. Out of these results, it is difficult to conclude whether the displacement of the innermost layer radius to 20 mm will really impact the physics performance of the VXD. The tracks inside the jets of our Higgsstrahlung events have a wide momentum spectrum, see figure 5.9. It is thus necessary to proceed to studies addressing physics final states in order to evaluate the inner layer radius effect. For the purpose of the study, a new neural net set was trained for the VXD05_rint20 geometry. The purity and efficiency curves for b– and c–tagging, compared to the ones of the nominal VXD05 model, are demonstrated in figure 5.10.

Both options have almost identical b–tagging performance. Concerning the c–tagging, the nominal VXD05 geometry appears to have better performance, as expected. In the range of low efficiency (~ 0.1 to 0.2) the relative difference in purity between the two options reaches $\sim 20\,\%$. This may be important for some specific studies, since in this efficiency range the detector provides its highest purity. For values of efficiency above 0.25, the relative difference in purity is $\lesssim 15\,\%$.

Study of the Material Budget

The LoI goals for the material budget are $0.11\,\%\,X_0$ for single sided and $0.16\,\%\,X_0$ for double sided ladders. These values are considered rather aggressive, and are currently out of reach of all the candidate VXD sensor technologies. Without of course excluding the possibility of a breakthrough in the future, it is interesting to test the performance of the VXD equipped with ladders featuring more realistic material budget. The PLUME

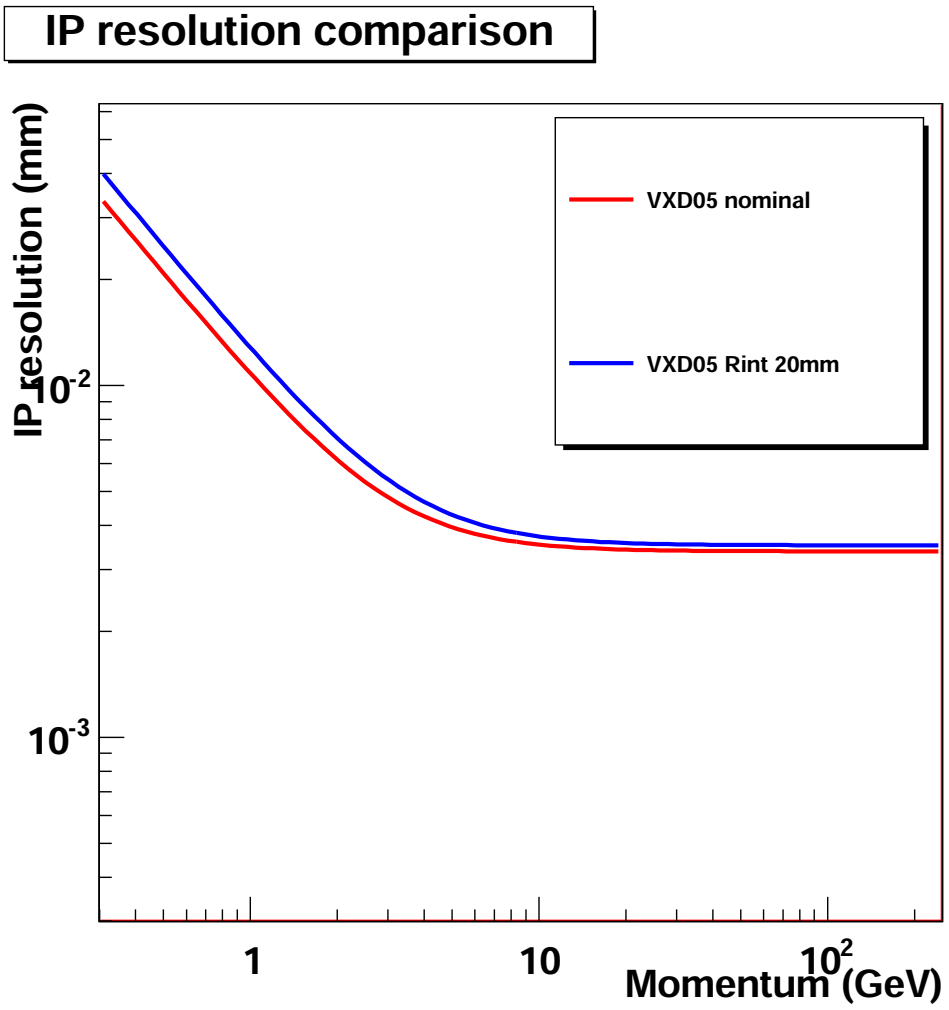


Fig. 5.8: Impact parameter resolution. The blue line is standing for the nominal VXD05 geometry and the red for the VXD05_rint20.

project (see sub–section 3.4.1), is expected to provide a double sided ladder of $0.232 - 0.392\,\%\,X_0$ by 2012. Therefore, the flavour tagging performance of the VXD05 equipped with double sided ladders featuring a material budget of $0.235\,\%\,X_0$ and $0.39\,\%\,X_0$ will be studied and compared with the LoI configuration. The flavour tagging neural nets undergo a dedicated training for each modification of the material budget. The results are compared in figure 5.11.

In case the LoI material budget estimations will be proven unrealistic, then as figure 5.11 indicates, degraded VXD performance is expected. The b–tagging remains almost unaffected, but the c–tagging suffers from an important degradation. In the region of high purity (efficiency between 0.1 to 0.25) the nominal VXD05 option has $\sim$

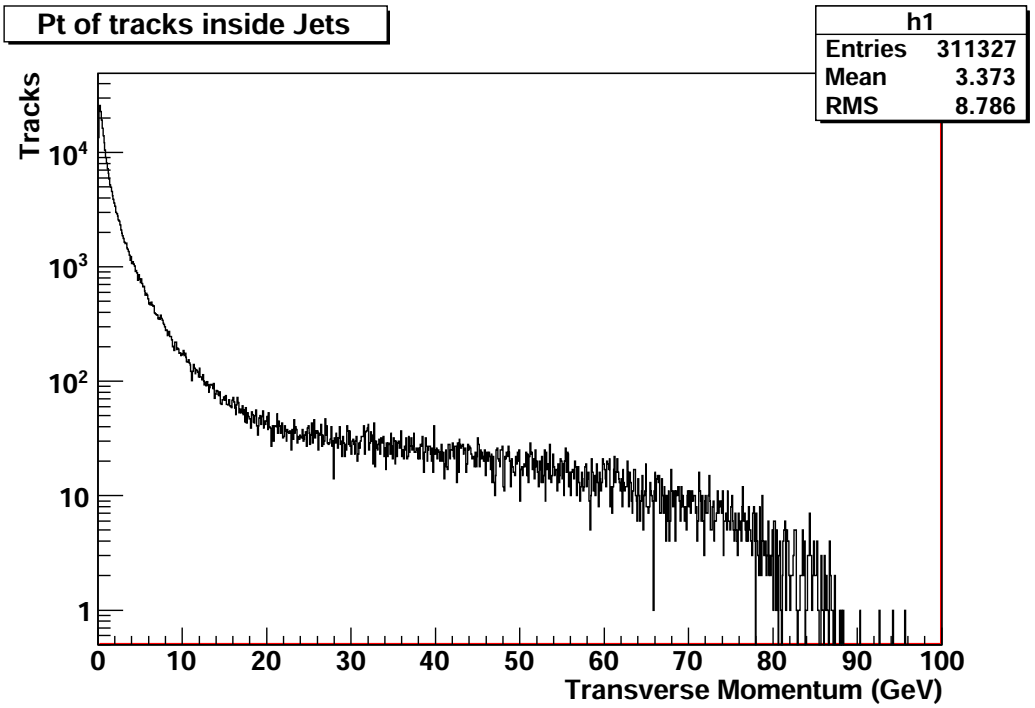


Fig. 5.9: Transverse momentum of the reconstructed tracks inside the jets.

30 % better performance on c–tagging than the VXD05 geometry featuring a radiation length of 0.235 %, and $\sim$ 50 % better than the one featuring a radiation length of 0.39 %. For values of efficiency ≥ 0.25, the nominal VXD05 option exhibits ≤ 10 % better performance than the option with 0.235 % radiation length, and $\sim$ 30 % than the one with 0.39 % radiation length. It is obvious that the more pessimistic approach of 0.39 % X_0 will severely affect the potential of the ILD to perform physics studies that require high performance flavour tagging, like the study of the Higgs couplings to fermions. The more optimistic scenario, induces also a degradation on the c–tagging performance, which is less pronounced. It could be tolerated, with the condition that no other compromises on the VXD design will take place (for example concerning the innermost layer radius).

One may think that the effect of an increased material budget would be less pronounced in the case of a single layer geometry. The reason is that single sided ladders are lighter than the double sided ones; thus the innermost layer introduces less material budget so close to the IP. This thought motivated us to perform a similar study with the VXD03 geometry. The estimation of the thickness of a single ladder is not straightforward, since there is not a corresponding project like PLUME for them. Obviously, the thickness of the sensor remains the same. The same should hold true for the flex cable: therefore, the thicknesses of the sensor, kapton and metal layers will

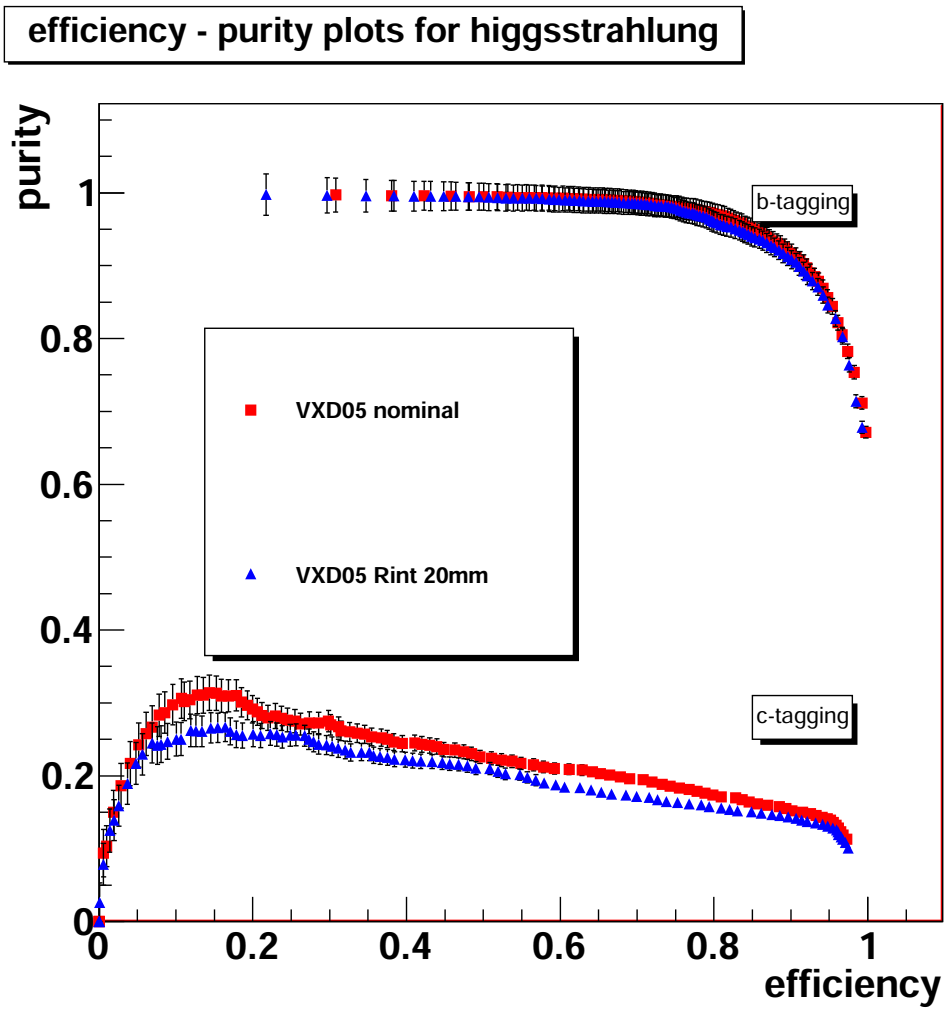


Fig. 5.10: Purity versus efficiency for b- and c-tagged jets in case of VXD05, for two values of the innermost layer radius. The red full squares stand for for 16 mm and the blue full triangles for 20 mm (so–called VXD05_rint20 variant).

be given by table 3.3. The component of the ladder whose thickness is expected to change significantly is the support structure. A conservative estimation for its material budget is $0.1\,\% X_0$. Therefore, for the whole single sided ladder, a material budget of $0.18\,\% X_0$ can be considered as a quite realistic approach.

Figure 5.12 compares the flavour tagging performance of the nominal VXD03 option and the one featuring a material budget of $0.18\,\% X_0$. The performance degradation of c–tagging is much less striking than the one observed for the double sided geometry. The c–tagging of the nominal option is better by $\leq 10\,\%$.

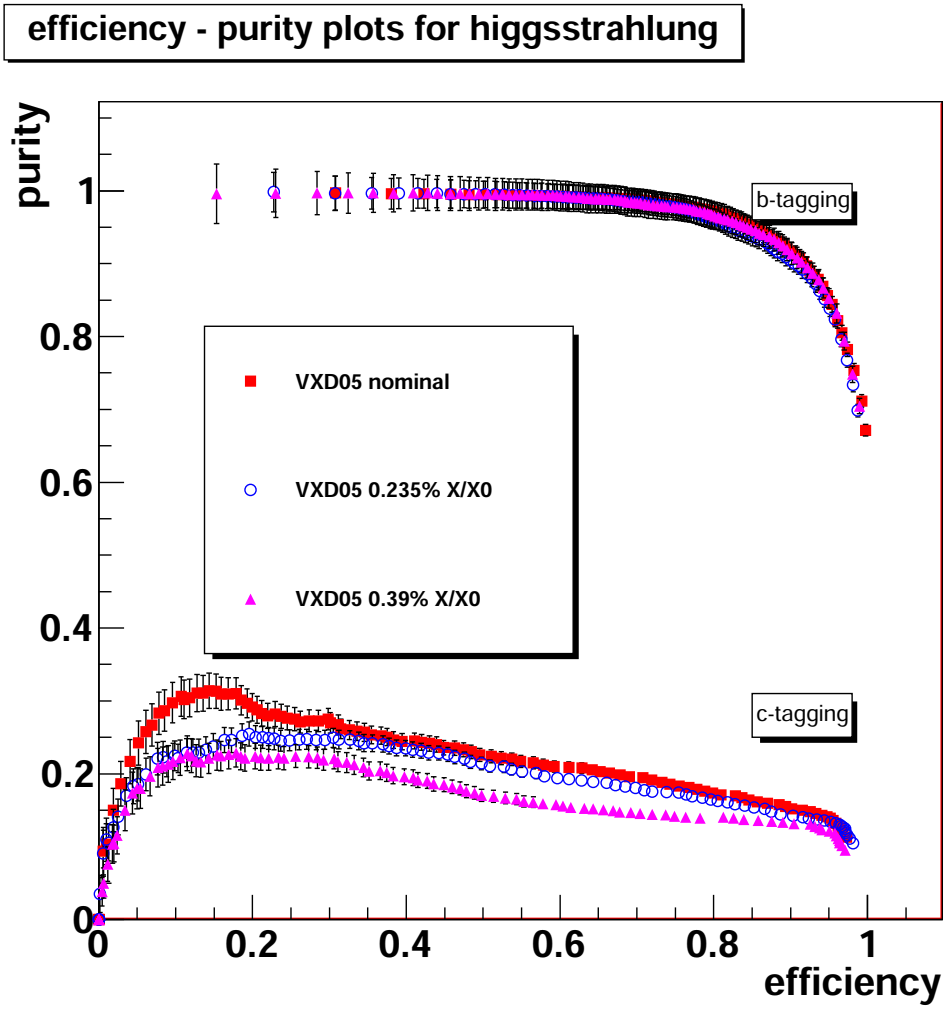


Fig. 5.11: Purity versus efficiency for b- and c-tagged jets as a function of the material budget of the ladder for the VXD05 geometry. The red full squares stand for the nominal material budget of $0.16\,\%\,X_0$, the blue circles for $0.235\,\%\,X_0$ and the pink full triangles for $0.39\,\%\,X_0$

Effect of the Single Point Resolution of the Outer Layers

The outer layers design of the proposed CMOS based VXD, as explained in section 3.4.4, is optimised for reduced power consumption. A way to achieve this goal is to reduce the number of sensor columns, i.e. by increasing the pixel pitch. Thus, a slightly worse single point resolution is expected, see table 3.5. This fact motivated us to study the flavour tagging performance as a function of the single point resolution of the outer layers sensors.

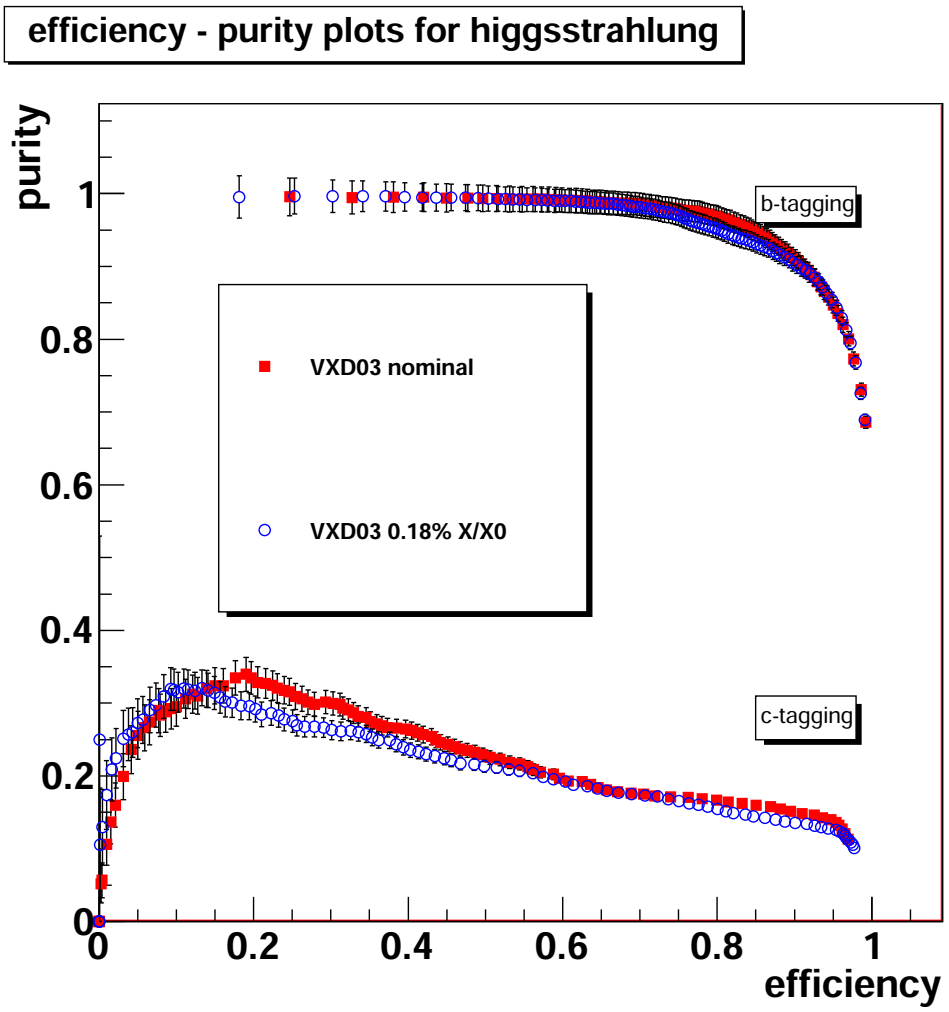


Fig. 5.12: Purity versus efficiency for b- and c-tagged jets as a function of the material budget of the ladder for the VXD03 geometry. The red full squares stand for the nominal radiation length of 0.11 % while the blue circles stand for 0.18 %.

Figure 5.13 compares the performance of the VXD05 geometry featuring the default value of the resolution (2.8 μm) in all layers, with the performance of the same model when the resolution is increased to 5 μm for the 2 outer double layers. The observed effect on the flavour tagging is negligible. This is quite encouraging for the efforts to minimise the power dissipation.

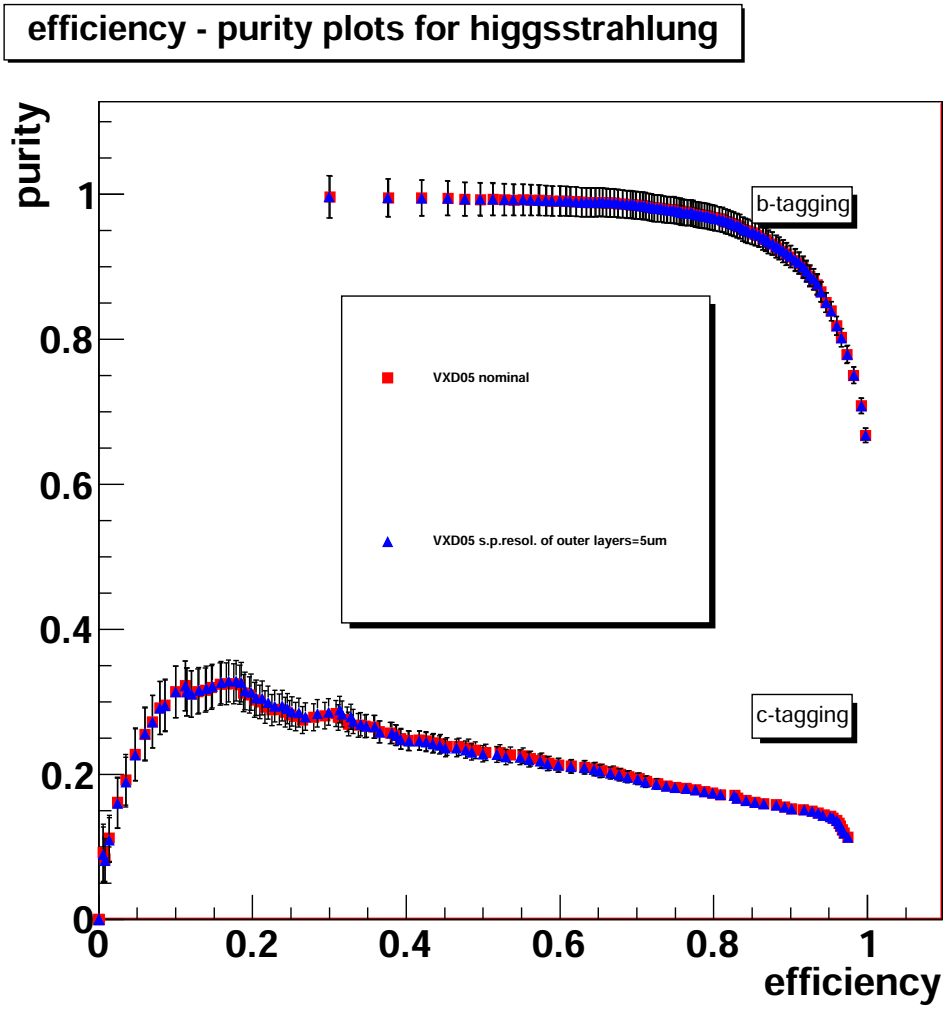


Fig. 5.13: Purity versus efficiency for b- and c-tagged jets using the VXD05 model. The red full squares stand for an assumed sensor's single point resolution for the 2 outer double layers of 2.8 μm, while the blue full triangles stand for a 5 μm resolution instead.

Study of the Hit Detection Efficiency

As already mentioned, the assumed hit finding efficiency is 100 % for every layer. This is an ideal case. It is interesting to study the performance of the two main candidate VXD geometries, VXD03 and VXD05, when inefficiency is being introduced. For a double sided ladder, assuming a detection efficiency of ϵ, the probability to get ≥ 1 measurement at a superlayer is $(1-(1-\epsilon)^2)$ [Baudot]. For a single sided ladder, the probability is obviously ϵ.

The motivation of the study is to see whether the double sided structure of the ladder can compensate for the loss of one hit due to inefficiency. In order to test that, we will compare the flavour tagging performance of the VXD03 and VXD05 designs, assuming an efficiency of 99 % during the digitisation level. An efficiency of 99 % is considered a rather poor performance. It is expected that the VXD will be equipped with sensors featuring an efficiency of $\geq$ 99.9 %. The value of 99 % is chosen for this study in order to indicate whether one of the geometries has an intrinsic advantage compared to the other, in case of inefficiency.

The results of the study are presented in figure 5.14. For both geometries, the effect of a rather low detection efficiency on the flavour tagging performance is marginal. Thus the detection efficiency will not be a strong argument for the decision for the selection of one of the two geometries.

5.8.3 Flavour Tagging Performance Including Beam Pair Background

The optimisation of the VXD without taking into account the beam background effect is certainly biased. On the other hand, as mentioned in the introduction, the present version of the silicon tracking cannot cope with high hit density. Thus if we will take it into account for our physics studies, it will not be possible to obtain reliable results.

The studies of the beamstrahlung effect at the ILC have been carried out with the Guinea Pig generator [GuineaPig]. Overlaying the simulated beam background hits (expected from the number of bunch crossings corresponding to the VXD integration time) to the ones of the physics events would require vast computer resources. Therefore a random hits generator will be used, the VTXNoiseClustersProcessor of the MarlinReco package, which generates hits with uniform distribution over the ladder surface. The produced hit collection is merged with the physics hit collection before the digitisation level. The density of the hits is according to the estimated hit densities for each layer, as given in table 2.2, multiplied by the expected number of bunch crossings over which each layer is being read out. The use of a random hits generator may result into an underestimation of the track combinatorial background. This is because the beamstrahlung hits are created by real particles that follow trajectories inside the detector's magnetic field, which means they belong to real tracks. These tracks should be reconstructed by the tracking tools. On the contrary, using a random hit generator, the hits of course do not belong to any track. Therefore, only ghost* tracks will be reconstructed.

The standard ILD track reconstruction scheme is used also here. That means that standalone tracking is performed independently in the TPC and the silicon detectors, followed by an attempt to merge the two track segments. The number of physics hits in

*Tracks that do not correspond to any particle, but are artefacts of the tracking algorithm.

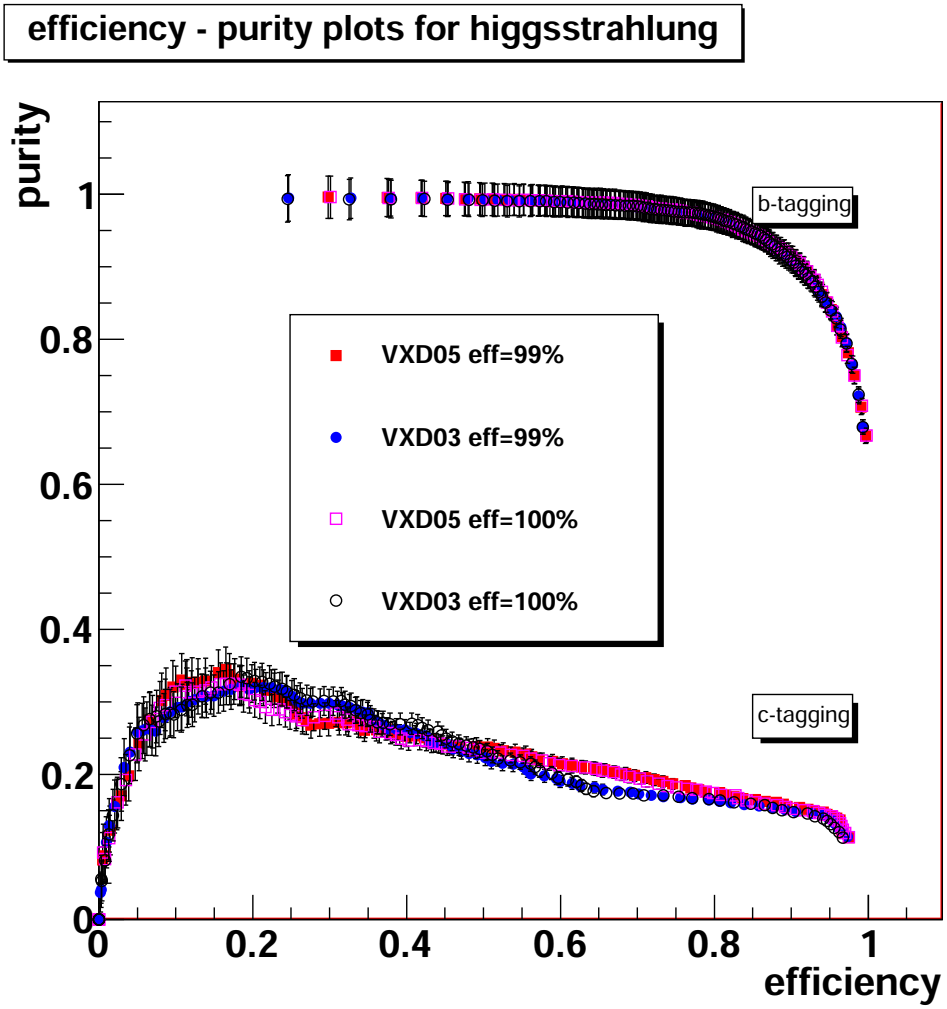


Fig. 5.14: Purity versus efficiency for b- and c-tagged jets for each geometry. The red full squares stand for the VXD05 and detection efficiency of 99 %, the pink open squares stand for the VXD05 and detection efficiency of 100 %, the blue full circles for the VXD03 and detection efficiency of 99 % and the black open circles for the VXD03 and detection efficiency of 100 %

one event ((created in one electroweak interaction) on the VXD are $O(10^2)$. When we superimpose the expected beam background hits (according to each layer's integration time), the total number of hits on the VXD increases by three orders of magnitude and reaches the number of $O(10^5)$. This results in a huge increase of the hits triplets used as track seed candidates (see section 2.4.3), leading to an extremely large computation time. In order to avoid this, the hit triplets formation is restricted to the SIT and the outer layers of the VXD, where the hit density is relatively low.

The average numbers of reconstructed tracks per event, without and with beam background, are compared in table 5.5 for the two main VXD geometries. An increase of the number of reconstructed tracks by two orders of magnitude is observed. These are obviously ghost tracks, since they occure from the combination of random noise hits. Therefore the tracking purity is very low.

One may argue that the tracking purity can be restored (at least partially) by imposing stricter selection criteria during track reconstruction. This approach makes sense only if we have a very satisfactory tracking efficiency. Thus we proceed to a tracking efficiency study, in the presence of beam background hits. For the purpose of the study, a track is considered as found if at least 75 % of its hits are coming from the same particle. The figure 5.15 demonstrates that the tracking efficiency stays below 30 % in the presence of the beam background hits, a rather discouraging result.

These tracking results strongly suggest that a separate neural nets training that will include the random noise hits should be performed, in order to pass to flavour tagging studies. This is because the number and the features of the track sample used as input to the neural nets are altered radically. The computer resources required for this task go beyond the ones we have available. Therefore, the evaluation of the effect of the beam background on the VXD flavour tagging performance, using the current tools, could not be conducted.

geometry	beam background	standalone silicon tracks	combined tracks
VXD05	no	25 ± 10	29 ± 12
VXD05	yes	1069 ± 54	697 ± 37
VXD03	no	25 ± 10	29 ± 12
VXD03	yes	5018 ± 76	1834 ± 50

Tab. 5.5: Average number of tracks per event reconstructed in the silicon tracker and the overall ILD tracking system. Comparison for the two main VXD geometries with and without beam background.

Concluding, table 5.5 implies that we should rather give up the standalone VXD tracking. At least as far as the current silicon tracking algorithm is used. The important question that should be addressed is whether a standalone VXD tracking is crucial for the ILD physics program. If not, then a way should be found to improve the VXD performance in the presence of the beam background. An idea is to use only tracks that reach the TPC. The track reconstruction will be performed at the TPC, which is much less affected by the pair background. The TPC tracks will be used as seeds that will be extrapolated to the silicon detectors. Obviously, this idea imposes that all the tracks with $P_T \leq 300\,\mathrm{MeV/c}$, that do not reach the TPC, will not be accounted for. In that case, the two following points should be addressed. First, if this is compliant

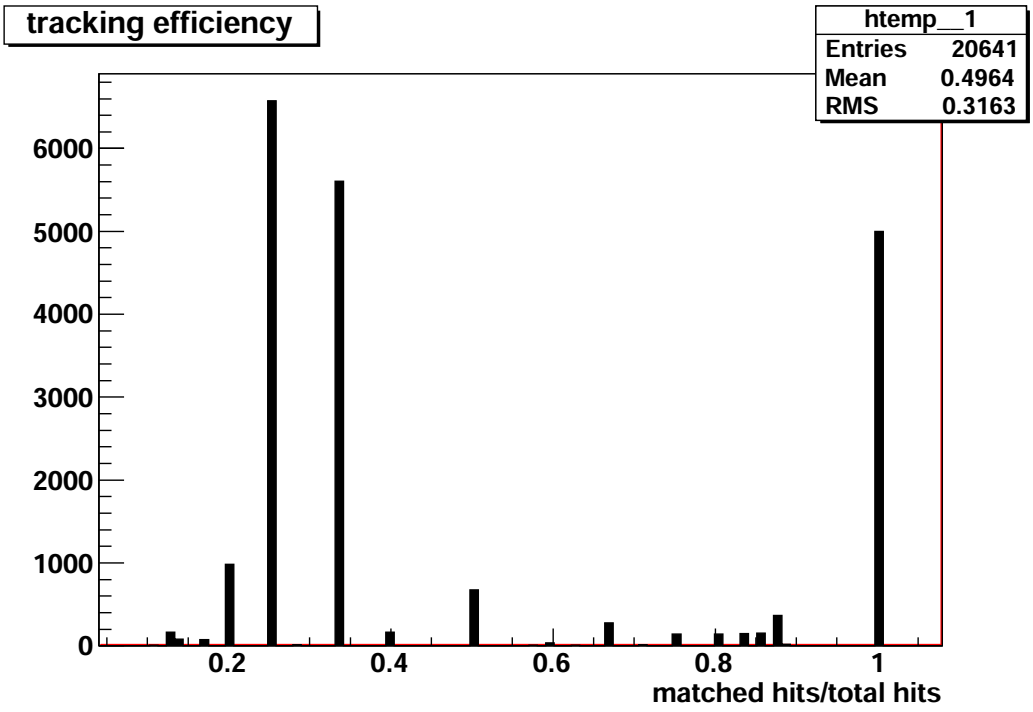


Fig. 5.15: Tracking efficiency of the standalone silicon tracking algorithm, when the beam background hits are included. The horizontal axis represents the ratio of the hits associated to a track that have been generated from the same true particle, versus the total number of hits associated to this track.

with the ILD physics requirements. Second, to study the tracking performance when the track is extrapolated from the TPC to the silicon detectors, in the presence of the beam background. This particular topic will be examined in the next chapter.

5.9 Higgs Branching Ratios Extraction

The feasibility of the Higgs hadronic branching ratios analysis, using the Higgsstrahlung process, has been demonstrated for the LoI. The sensitivity of the branching ratios extraction has been estimated for each of the three decay channels of the Z boson ($Z \rightarrow q\overline{q}, Z \rightarrow l^+l^-, Z \rightarrow \nu\nu$). By combining the three channels, an improved overall sensitivity is obtained, as compared to each channel alone (see table 5.6). The ILD_00 simulation model has been used, which corresponds by default to the VXD05 vertex detector geometry. The specific analysis of the $ZH \rightarrow l^+l^-q\overline{q}$ channel, where the leptons pair are either muons or electrons, has been performed by the LCFI group [Goldstein]. The obtained sensitivities for the $H \rightarrow b\overline{b}$, $H \rightarrow c\overline{c}$ and $H \rightarrow gg$ decay

channels are summarised in table 5.7.

Channel	$H \to b\bar{b}$ (%)	$H \to c\bar{c}$ (%)	$H \to gg$ (%)
$ZH \to l^+l^- q\bar{q}$	$2.7 \oplus 2.5$	$28 \oplus 2.5$	$29 \oplus 2.5$
$ZH \to \nu\bar{\nu}h$	$1.1 \oplus 2.5$	$13.8 \oplus 2.5$	
$ZH \to q\bar{q}c\bar{c}$		$30 \oplus 2.5$	
Combined	2.7	12	29

Tab. 5.6: Relative uncertainties on the Higgs hadronic branching ratios extraction from the LoI. The added 2.5% on the uncertainty corresponds to the uncertainty on total Higgs production cross–section. The symbol l^+l^- stands either for muons or for electrons [LoI 10].

Channel	$H \to b\bar{b}$ (%)	$H \to c\bar{c}$ (%)	$H \to gg$ (%)
$Z \to \mu^+\mu^-$	4	50	50
$Z \to e^+e^-$	6	60	50

Tab. 5.7: Relative uncertainties on the Higgs hadronic branching ratios extraction. Using the Higgsstrahlung process, where the Z decays either to a pair of electrons, or a pair of muons [Goldstein].

The scope of this section is to compare the sensitivities of the two main candidate VXD geometries, VXD03 and VXD05, on the Higgs hadronic branching ratios extraction. Their performance will be a crucial factor for the VXD geometry selection procedure. For the purpose of the study, the $e^+e^- \to ZH \to \mu^+\mu^- X$ process will be used, the same one as for the flavour tagging studies performed earlier in this chapter. The events are reconstructed according to section 5.5, and the cut based selection described in section 5.6 is applied. For the Higgs branching ratios extraction, we are going to follow a strategy similar the one used in [Kuhl 07] and for the LoI studies. The event wise tagging variable, the X–likeness, where X can be either b or c, is defined as follows:

$$X - likeness = \frac{X_1 X_2}{X_1 X_2 + (1 - X_1)(1 - X_2)} \tag{5.9}$$

The X_1 and X_2 variables are the output of the flavour tagging neural nets for the first and the second jet of the event respectively (to remind, the recoil products of the Z boson are forced to 2–jet clustering). The output of the b(c) neural nets expresses actually the probability that this jet is a b(c) jet. Obviously, the b–likeness is constructed from the b–tagging nets, and the c–likeness from the c–tagging nets.

The X–likeness is constructed in a way that requires both the jets to have a X–tag probability close to 1, in order to be also close to 1. In any other case, the X–likeness value is close to 0. Hence, the likeness variable can characterise the Higgs decay. A value of b(c)–likeness close to 1, implies a high probability that the Higgs decayed to a pair of b(c)–quarks. Using the event wise tagging variable, we can directly proceed to the measurement of the Higgs branching ratios.

5.9.1 Methode

Based on results obtained from previous studies of the Higgs hadronic branching ratios at electron–positron linear colliders, it is proven that a pure cut based analysis does not provide the highest sensitivity. Indeed, in the cut based analysis presented in [Brau], the obtained sensitivity for $H \to b\overline{b}$ reached $\sim 3\,\%$ and for $H \to c\overline{c} \sim 45\%$. Compared to the sensitivities obtained in [Kuhl 07] with a template fitting technique, of $\sim 1\,\%$ and $\sim 12\,\%$ respectively, they are significantly lower. Since our goal here is to compare the performance of two VXD geometries, one has to use the method that yields the higher sensitivity.

A Monte–Carlo (MC) template fitting technique is going to be used. We are going to give a brief description of this technique. For a complete description, one can refer to [Barlow 93]. We start from a sample of Higgs decay events, of all the possible decay channels, inevitably mixed with physics background events. This is actually the situation we suppose we will face with the real data. Out of this sample, the Higgs branching ratios to b– and c–quark pairs have to be measured. The Higgs decay of each event can be characterised by a number (N) of event wise tagging variables. In our case, we will use two variables of this kind, the b–likeness and the c–likeness. We build the 2–dimensional distribution of the b–likeness and the c–likeness out of our "data". The goal is to extract the number of b– and c–tagged events from this distribution, i.e. the Higgs hadronic branching ratios.

If we could reproduce theoretically the same distribution for each Higgs decay channel, the composition of our data could be resolved by fitting those theoretical templates to the data. An analytical form of this distribution does not exist. However, it can be constructed out of MC simulated samples. 2–dimensional templates of the b– and c–likeness will be constructed out of MC samples of exclusive Higgs decays to $H \to b\overline{b}$, $H \to c\overline{c}$, $H \to gg$ and for the background.

To do that, the initial sample is splitted into two subsamples. $250\,\mathrm{fb}^{-1}$ will be used for the “data” and $500\,\mathrm{fb}^{-1}$ for the MC templates. Obviously, the two subsamples should be independent. Both subsamples undergo the same event reconstruction and selection. As backgrounds are considered the physics background, coming from the $e^-e^+ \to ZZ$ process, and the non-hadronic Higgs decays.

Next, the data are binned into the 2–dimensional template of b– and c–likeness. The number of data events in bin i is called d_i. The expected number of events f in bin

i, coming from all the Higgs decays channels, will be given by the following equation:

$$f_i = N_D \sum_{j=1}^{m} P_j \alpha_{ij} / N_j \tag{5.10}$$

where the sum runs over the m possible Higgs decay channels, N_D is the number of events in the "data" sample, N_j the number of events of the j MC exclusive Higgs decay channel and α_{ij} the number of events in bin i, coming from the j decay channel. The P_j is going to give us the ratio of the events coming from the j channel. We want to determine through our data the normalisation factor P_j. In order to do that, we are going to fit the MC templates of the exclusive Higgs decay channels to the data, letting the normalisation factors P_j to vary. For a SM Higgs, since the template fitting is performed with SM MC samples, the P_j are expected to be equal to 1. With real data, a deviation from this value would be an indication of a non–SM Higgs boson. Having estimated the P_j factors, the Higgs branching ratios can be determined by the following equation:

$$BR(H \to j) = P_j \times BR(H \to j)_{SM} \times \frac{\sigma(e^+e^- \to HZ)_{SM}}{\sigma(e^+e^- \to HZ)} \tag{5.11}$$

Where P_j are the parameters obtained from the fit, $BR(H \to j)_{SM}$ the SM branching ratios, $\sigma(e^+e^- \to HZ)_{SM}$ the SM cross–section for the Higgsstrahlung process and the $\sigma(e^+e^- \to HZ)$ the measured cross–section. In the SM, for a given mass of the Higgs boson, the branching ratios are uniquely defined. At the ILC, the Higgs mass and cross–section can be measured with a model independent analysis [Ito 09]. A systematic uncertainty is going to be introduced on the Higgs branching ratio analysis from this measurement, whose value is currently estimated to 2.5 %.

The normalisation factors P_j for the $H \to b\bar{b}$ (P_{bb}), $H \to c\bar{c}$ (P_{cc}), $H \to gg$ (P_{gg}) decay channels, will be estimated from the fit of the MC templates to the data. The 2–dimensional templates are illustrated by figure 5.16. The normalisation factor of the background processes, P_{bkg}, will be fixed to 1. This is easily explained for the SM physics background, the main background contribution coming from the ZZ process. The cross–section of this process, thus the contamination of our data sample, is expected to be accurately known. On the other hand, the non–hadronic Higgs decays are considered also as background processes. So fixing the P_{bkg} can be tricky. We do not know the non–hadronic Higgs branching ratios. Consequently, using in the fitting procedure a fixed normalisation factor it is in principle not correct. However it can be justified, if the non–hadronic Higgs branching ratios can be measured with negligible contamination from the Higgs hadronic branching ratios. There are studies that demonstrated this [Brient 04]. For the purpose of our study, we will assume that the non–hadronic Higgs branching ratios are already measured, thus we can fix P_{bkg} to 1.

5.9.2 Fitting Technique

Concerning the fitting method itself, a number of different approaches exist, each one with its own potential and limitations. A straightforward way is to construct the χ^2 of the data d_i and the expected values f_i:

$$\chi^2 = \sum_i \left(\frac{d_i - f_i}{\sigma_{d_i}} \right)^2 \tag{5.12}$$

where the sum runs over all the bins i of the 2–dimensional distributions of b– and c–likeness and σ_{d_i} is the statistical uncertainty of d_i. By minimising the χ^2, the values of P_j can be estimated. In order to use it, the number of events d_i in each bin should follow a Gaussian distribution. Actually, since there are products of a particle decay, they follow a Poisson distribution. For a relatively large number of events in each bin, the Gaussian distribution can be a very good approximation of the Poissonian. This is a necessary condition in order to use equation 5.12. By examining figure 5.16, it is clear that this condition is not true for all the bins of the templates.

A second issue is the statistical fluctuations of the MC samples. The uncertainty on the denominator of the equation 5.12, $\sigma_{d_i} = \sqrt{d_i}$, takes into account only the statistical fluctuations of the data. This is correct, if the statistical fluctuations of the MC samples are negligible. This is true, for the ideal case that MC samples of infinite size are used, or more realistically, if their size is $\gtrsim 1$ order of magnitude bigger than the one of the data sample. In our study the MC samples have only twice the size of the data samples. So the MC statistical fluctuations should be accounted for. Thus, equation 5.12 should be modified to:

$$\chi^2 = \sum_i \frac{(d_i - f_i)^2}{d_i + (\frac{N_D}{N_{MC}})^2 \sum_j a_{ij}} \tag{5.13}$$

where N_{MC} is the sum of the events of the MC samples. Finally, the equation 5.13 is going to be used for the fit. In order to justify the Gaussian approximation, all the bins with less than five entries will not be used for the fit. This workaround was inspired from the study exposed in [Goldstein]. The number of bins of the 2–dimensional templates were restricted to 10×10. The motivation was not to split the events into a too large number of bins, to mitigate the number of bins with 0 or very few events.

A way to overcome the problem induced by the bins with very few events, it is to use the binned maximum likelihood method. The likelihood function will be constructed assuming that the number of events in each bin follows the Poisson distribution. The normalisation factors P_j will be estimated by maximising the logarithm of the likelihood:

$$\ln L = \sum_{i=1}^{n} d_i \ln f_i - f_i \tag{5.14}$$

The equation 5.14 is the appropriate way to estimate the P_j, if high statistics MC samples are available. This is because this equation supposes that the MC samples undergo negligible statistical fluctuations. Thus, given the available statistics in our study, it is not applicable here. A method that can cope both with the statistical fluctuations of the MC samples and the bins with very few events, it is presented in [Barlow 93]. It is implemented in the TFractionFitter function of the ROOT software [ROOT]. Nevertheless, in order to use that method, the total number of events should be significantly bigger than the number of events of any bin. Examining the histograms of figure 5.16, we can see that this requirement is not always satisfied. This is more pronounced for the template of the $H \rightarrow b\bar{b}$ exclusive decay. Due to the excellent b–tagging performance of the detector, the vast majority of these events are in the bin with b–likeness equal to 1 and c–likeness equal to 0. Hence, this approach is also not appropriate for this study. Equation 5.13 was therefore prefered for the estimation of the P_j.

5.9.3 Results

The 2D templates of b– and c–likeness, which are going to be used for the fit, of the "data" and the MC samples of the exclusive Higgs hadronic decays and the background (ZZ process and non–hadronic Higgs decays) are shown in figure 5.16 for the VXD05 geometry. The integrated luminosity is $250\,\mathrm{fb}^{-1}$ for the "data", and $500\,\mathrm{fb}^{-1}$ for the MC samples. We are going to fit the $H \rightarrow b\bar{b}$, $H \rightarrow c\bar{c}$, $H \rightarrow gg$ and background MC templates to the "data" (Higgs decays + ZZ background), using the equation 5.13.

The estimated normalisation factors P_j, for the VXD03 and VXD05 geometries, are summarised in table 5.8. After estimating the P_j, the Higgs hadronic branching ratios with their associated errors can be directly extracted from the equation 5.11. The last factor of this equation, the ratio $\frac{\sigma(e^+e^-\rightarrow HZ)_{SM}}{\sigma(e^+e^-\rightarrow HZ)}$ of the SM cross–section to the measured one, in the case of a simulation study using SM data samples, is obviously by definition 1. Of course, for a study with real data, the cross–section of the Higgsstrahlung process should be measured with a model independent analysis. Hence, in our case, we are going to use only the systematic uncertainty that is going to be introduced by the measurement of the model independent Higgsstrahlung cross–section, which amounts to 2.5 %. The extracted Higgs hadronic branching ratios are summarised in table 5.9.

The obtained results are in accordance with the branching ratios expected for a SM Higgs boson, having a mass of 120 GeV (see table 5.1). The estimated sensitivities for the $H \rightarrow c\bar{c}$ and $H \rightarrow gg$ decays are rather low: the relative statistical uncertainties are $\sim 50\,\%$. But here we should stress that the low sensitivity to these decays comes

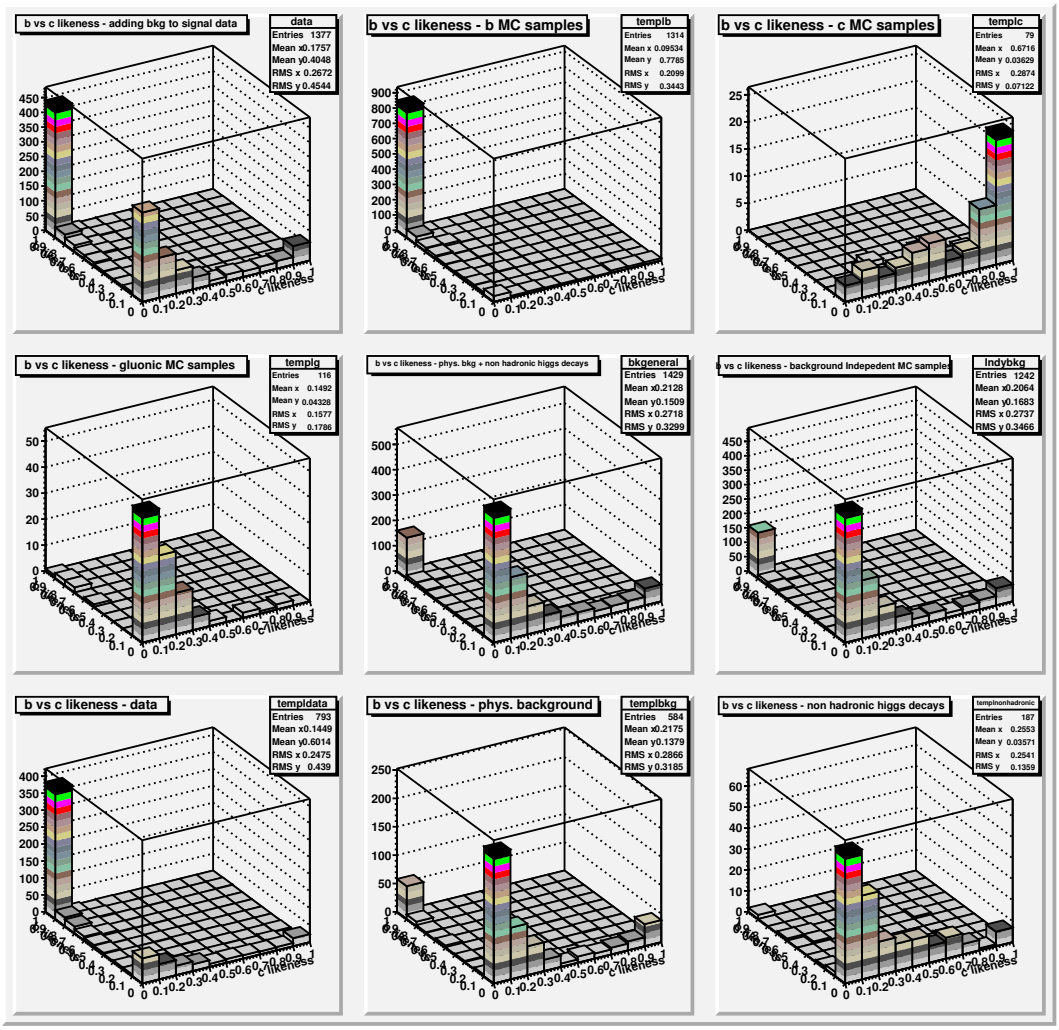


Fig. 5.16: 2–dimensional templates of b– and c–likeness. Starting from the top left plot: data (only Higgs decays), MC sample exclusive $H \to b\overline{b}$, $H \to c\overline{c}$, $H \to gg$, MC background (non–hadronic Higgs decays + ZZ background), MC sample ZZ background, MC sample non–hadronic Higgs decays, data (Higgs decays + ZZ background), data (only ZZ background).

as a consequence of the low statistics of our sample, since we restrict our analysis of Higgsstrahlung process only to the events where the Z boson decays to a pair of muons (3.5 % branching ratio). If we consider also its leptonic decays to an electron–positron pair, as well as its hadronic and invisible decays, the obtained combined sensitivity will reach the results stated in table 5.6. But in our study, we are interested in the direct comparison of the sensitivities obtained with each one of the two main candidate

	VXD05	VXD03
P_{bb}	0.93 ± 0.06	0.99 ± 0.06
P_{cc}	0.93 ± 0.59	0.86 ± 0.54
P_{gg}	1.68 ± 0.58	0.88 ± 0.61

Tab. 5.8: The estimated normalisation factors P_j for the Higgs hadronic branching ratios for the main VXD geometries, VXD03 and VXD05.

	VXD05	VXD03
$H \to b\bar{b}$ (%)	$63.39 \pm 4.09 \oplus 2.5$	$67.48 \pm 4.09 \oplus 2.5$
$H \to c\bar{c}$ (%)	$2.83 \pm 1.79 \oplus 2.5$	$2.61 \pm 1.64 \oplus 2.5$
$H \to gg$ (%)	$11.27 \pm 3.89 \oplus 2.5$	$5.90 \pm 4.09 \oplus 2.5$

Tab. 5.9: The Higgs hadronic branching ratios for the main VXD geometries, VXD03 and VXD05.

geometries.

On the Higgs hadronic branching ratios extraction, the VXD geometry affects only the estimation of the factors P_j. Thus we can compare the VXD03 and VXD05 sensitivity, by comparing the uncertainties on the estimation of the P_j. There is a contribution on this uncertainty coming from the finite size of the MC samples. In principle, it should be identical for both geometries. The statistical uncertainties reported in table 5.8 indicate that both geometries appear to have similar sensitivity. Therefore, the conclusion of this study is that the development efforts should be continued for both the VXD03 and the VXD05 candidate geometries.

5.10 Conclusions on the Vertex Detector Optimisation

In this chapter, the first point we have aimed to address, was the comparison between the single (VXD03), and the double layers (VXD05) geometries for the ILD vertex detector. The standard ILC software and reconstruction algorithms were used. We have tried to probe into the performance of the two VXD models, by studying their performance on heavy flavour tagging, and their sensitivities on the extraction of the Higgs hadronic branching ratios. The selected physics process was the Higgsstrahlung, $e^+e^- \to ZH \to \mu^+\mu^- X$.

The results presented in figure 5.7 and in table 5.8, show that both geometries demonstrate excellent b–tagging, and very good c–tagging performance, that satisfy

the ILD physics goals. We did not observe any striking performance differences between the two geometries, that could suggest to favour one of them. Therefore, the outcome of this first step is that it is suggested to continue pursuing the development efforts for both the main candidate geometries.

The studies were based on some indicative characteristics on the overall geometry and the ladder design, which are crucial in order to achieve the desired performance. However, it is not certain that all of these target specifications are achievable. To be more specific, this fear mostly concerns the minimum distance of the innermost layer of the IP and the material budget of the ladders. So, the second point of this analysis, has adopted the philosophy of a worst case scenario study.

The effect of the beamstrahlung on the VXD may be significantly higher than the one anticipated from the simulation studies. Even if the simulations are correct, the hit density on the VXD may be significantly higher than the one presented in table 2.2, if finally the anti–DID field will not be implemented. In any of the two aforementioned cases, the innermost layer radius should be moved away from the IP, approximately at $\sim 20\,\text{mm}$. This will degrade the VXD c–tagging performance by a factor of $\sim 15-20\,\%$, as it is shown at in figure 5.10.

In parallel, it is not certain if the aimed material budget of the VXD ladders can be realised. The figure 5.11 demonstrates the flavour tagging performance of a VXD featuring less challenging material budget values. Concerning c–tagging, it appears to be important to stay below $0.3\,\%\ X_0$. Here we have to underline that in both the aforementioned tests, the b–tagging remains excellent and unaffected. Thus, our optimisation studies focus on c–tagging.

Concluding, if one of the assumed VXD features proven to be not feasible, the overall degradation of the VXD c–tagging performance is expected to be tolerable. But if we fail to realise both of them, the VXD performance will be severely penalised.

A third point of the study was to test various technological parameters of the VXD. We studied the performance of both VXD03 and VXD05 geometries as a function of the detection efficiency. Selecting a rather low efficiency, 99 %, the observed effect on the flavour tagging performance was negligible. Thus, it is not expected that the detection efficiency will play an important role on the selection between the two main geometry options.

The other technology parameter tested was the single point resolution of the sensors equipping the outer layers of the VXD. This study was inspired by the design of the outer layers of the CMOS based VXD (see subsection 3.4.4). One can envisage to reduce the number of columns of the sensors at the outer layers, by increasing their pixel pitch, in order to reduce the power consumption. The question is how much this can affect the VXD performance. The studies presented in figure 5.13, show that the expected performance degradation by moving to pixels featuring a pitch of $\sim 35\,\mu\text{m}$, being readout with 3–4 bits ADCs, will be negligible.

An essential aspect for the VXD optimisation design, which has not been taken into account for these studies, is the alignment. All the studies were performed assuming a perfect alignment. Of course, this is an ideal case. An extremely accurate alignment is needed, $\leq 1\,\mu$m, in order for the VXD to fulfill the ILD physics program requirements. It is suspected that both geometries do not allow for the same alignment accuracy. The overlapping areas of the ladders of the VXD03 geometry can offer a region with two hits in the same layer, while the VXD05 offers a region with four hits. The increased number of hits is likely to be beneficial for the alignment, either of the overall VXD, or ladder to ladder. This could be an argument in favour of the double layers geometry option. However the alignment accuracy has not been quantified yet. Hence, no missalignment was assumed for these studies.

Finally, it is not possible to conclude on the evaluation of the ILD VXD, without considering the beam background effect. This effect was neglected in the physics analysis studies presented in this chapter. The reason is exposed in the section 5.8.3. The current silicon tracking is not mature enough to cope with it. As has been already indicated, the tracking in the silicon detector is still an open issue. Therefore, it does not make sense to repeat the physics studies including the beam background effect in the current reconstruction framework.

The development of a tracking algorithm is outside the scope of this thesis. Thus, in order to finalise the ILD VXD studies, we will try to examine the effect of beam background on the tracking itself. In other words, to test which configuration of silicon detectors can provide the optimal tracking, in terms of tracking performance. To do so, since we cannot count on the standalone silicon tracking, we will use tracks reconstructed in the TPC and propagate them to the silicon detectors. It is clear that the SIT will be play a crucial role in this process, since it bridges the gap between the VXD and the TPC. This will be the subject of the next chapter.

CHAPTER 6

Silicon Tracking Studies

6.1 Introduction

The VXD optimisation studies inherited us with an open issue: the optimisation of the VXD without considering the beam background is quite incomplete. As already mentioned, the standalone silicon tracking used for the LoI studies is not appropriate for tracking in environments with high hit density, like the innermost VXD layer. As a consequence, physics analyses including beam background, conducted in the current reconstruction framework, cannot provide reliable results. Hence, we will try to approach the robustness of the VXD versus the beam background from a different point of view. In order to do so, we will examine the tracking performance that can be provided by a given VXD design. The methode used is discussed in the following section. We will focus on a CMOS based VXD, described in sub–section 6.3.1, which is an optimised version of the basic design described in section 3.4.

This study is going to include also the SIT. Our motivation is to combine the VXD with the SIT in a silicon detector configuration, capable to provide highly performant standalone silicon tracking. Moreover, the role of the SIT is crucial in the overall track reconstruction scheme, since it provides the link between the main tracker, which is the TPC, and the VXD.

The standard option for the ILD SIT, as described in the LoI [LoI 10] (see also sub–section 2.3.3), is equipped with silicon strip sensors providing single bunch crossing time resolution, but a rather modest spatial resolution, especially along the z coordinate. We will discuss whether an SIT equipped with CMOS pixel sensors can provide better tracking performance in the presence of the beam background. The excellent spatial

resolution of the highly granular pixel sensors may compensate for their lower time resolution. Of course, a question we also need to address, is the specifications of those pixel sensors.

A pixel SIT would use the same pixel sensor technology as the VXD. Therefore, the complexity of the $R\&D$ effort of the final system could be significantly reduced. That argument also, motivated the studies presented in this chapter.

6.2 The Methodology

For the purposes of the study, we are going to develop a pattern recognition based on the naive track following algorithm. This methode is based mainly on three elements: the provider of the track seeds, a way to propagate the track parameters and the associated error matrix from layer to layer and associate candidate hits, and finally, a quality criterio which will be used in order to sort the candidate tracks, select the "best" ones and reject the incompatible and "bad" tracks.

The trajectory of a charged particle inside a homogeneous magnetic field is a helix. In order to initiate a helix, one needs three spacepoints. Therefore, the track seed could be provided by a triplet of hits in the outer layers of the silicon detectors, where the beam background effect is less pronounced. However, developing a standalone tracking algorithm is outside of the scope of this thesis. Thus, we are going to use as seeds, tracks that have been reconstructed in another detector, namely the TPC. The initial track parameters and their associated errors will be taken from the TPC track.

The seed track will be extrapolated inwards to the silicon detector layers, which will be populated by both the "physics" and the beam background hits. There, we will search for hits that are close (we will come again on the subject of the track extrapolation area later) enough to the track extrapolation point, trying to find the "best" candidate hit and add it to the track. As best hit, we consider the one that is closest to the point where the track crosses the detector surface. After adding the hit (if exists), the track is refitted. The updated track parameters are extrapolated to the next detector layer, where this procedure is repeated.

The IMarlinTrack tool [Aplin 11] will be used for the propagation of the track parameters and its error matrix. The tool takes into account material effects. The propagation of the track error matrix, and more specific of the uncertainties of the impact parameters on the $R-\phi$ plane (δd_0) and the parallel to beam axis plane (δz_0), can provide a first indication concerning the robustness versus the beam background. These two uncertainties define the area where the track is expected to cross a given detector layer. Obviously, the smaller this area, the less background hits will be considered during the track reconstruction. The probability to pick the correct "physics" hit increases accordingly, and the number of the reconstructed ghost tracks decreases. The term ghost (or fake) tracks refers to tracks that do not correspond to a Monte–Carlo

particle, but are artefacts of the tracking algorithm.

The search area σ_{eff} is given by the following equation:

$$\sigma_{eff}^2 = \sigma_{ext}^2 + \sigma_{s.p.}^2 \tag{6.1}$$

where σ_{ext} reflects the track extrapolation area and $\sigma_{s.p.}$ stands for the sensor's single point resolution. For the purpose of the study we decided to consider as candidates, hits being inside a radius of $5 \times \sigma_{eff}$ from the track extrapolation point.

A number of different criteria exist in order to decide if a hit is suitable to be attached to the track. According to the naive track following approach, we accept as the proper hit the one that is closest to the track extrapolation point. Of course, an association of a "wrong" hit to the track, instead of the proper one, may happen. It is always possible that a wrong hit, which does not originate from the particle that has generated the track, is closer to the track extrapolation point. This hit may stem from another particle, from the detector noise, or from the machine induced backgrounds. The latter case is considered to be by far the biggest challenge for the track–hit association at the silicon detectors of the ILD.

The probability to match with the correct hit depends on the track extrapolation area, the hit density on the layer, the accuracy on the reconstruction of the hits positions (a convolution of the sensor single point resolution and the position reconstruction algorithm) and of course the hit selection criteria. In general, the probability for a wrong association increases when already wrong hits have been attached to the track, either during the seed formation, or during the previous steps of the track following. As aforementioned, the hit density mostly depends on the distance of the layer to the IP and the time resolution, while the track extrapolation area reflects the sensor's single point resolution, the layer material budget and the geometrical configuration of the detector. We are going to focus on these aspects.

We will examine the tracks coming from a benchmark physics channel. This will be the Higgsstrahlung process, $e^+e^- \rightarrow ZH \rightarrow \mu^+\mu^- X$ at $\sqrt{s} = 250\,\text{GeV}$, exactly the same one used for the VXD optimisation studies of the previous chapter. Then the tracking efficiency can be calculated from the ratio of the track hits originating from the same particle to the total number of hits that the particle created at the Silicon detectors.

6.3 Silicon Detectors Configurations

6.3.1 VXD Description

The VXD model used for the tracking studies is based on the one described in section 3.4. The values of spatial and time resolution for each layer are summarised in table 6.1. The value of the innermost layer's time resolution has decreased from $\sim 10\,\mu$s

(table 3.6) to $\sim 2\,\mu$s. More details on the justification of these values are provided in sub–section 6.3.3. Moreover, the strategy of optimising one face of a double sided ladder for spatial and the other for time resolution, is not restricted anymore to the first layer.

The main motivation for generalising this strategy to the whole VXD (maybe to the SIT as well) is that we want the silicon detectors to be able to support a standalone tracking algorithm. Thus, the combinatorics during the seed formation (see sub–section 2.4.3) should be minimised as much as possible.

The goal of the improvement of the innermost layer's time resolution is mainly the robustness against the beam background hit density. The overall ILD design is not yet finalised. It may happen that the anti–DID field will not be implemented, or that a new machine optics will generate higher energy electrons which the anti–DID will hardly sweep away. In that case, due to the increased beam background hit density, the innermost VXD layer should either be moved away from the IP (which is not desirable), or being read faster. Additionally, we should have in mind the increased beam background in the case that the ILC will run at $\sqrt{s} = 1\,\text{TeV}$.

The GEANT4 implementation of the VXD is updated from the one described in section 2.3.2, used for the LoI studies, to the one that is going to be used for the Detector Baseline Document (DBD) (see figure 6.1). For the purpose of the DBD studies, a more realistic description is required, especially in terms of material budget. The description of the ladders is based on already existing ones, realised inside the framework of the PLUME project (see section 3.4.1). The total material budget of each double sided ladder is $\sim 0.3\,\% X_0$. The sensitive part (i.e. the two sensors) amounts to $\sim 0.1\,\% X_0$. The support consists of 1.88 mm of silicon carbide foam of 6 % density, serving as a spacer and a supporting structure, featuring a material budget of $\sim 0.13\,\% X_0$. On both sides of the SiC foam are the flex cables. They are described as being made of $10\,\mu$m of Al and $50\,\mu$m of kapton, which gives us a material budget of $\sim 0.03\,\% X_0$ for each flex cable.

The ladders are enclosed in an hermetic Beryllium support. Three spiralling cooling tubes made of Titanium are located on the support endplates. Each double layer is assigned a dedicated cooling tube. The cables are described by three cones, consisting of kapton and Copper, of thickness varying with the z coordinate. Each cone starts from the place where the ladders are supported by the Be support (the Be annulus blocks), and going outside the VXD through an aperture in the cryostat.

6.3.2 SIT Description

The standard option of the SIT, featuring microstrip sensors, is assumed to provide single bunch crossing timing information and a spatial resolution of $7\,\mu$m in $R\phi$ and $50\,\mu$m in z direction. The corresponding features of the pixelated SIT are summarised in table 6.1.

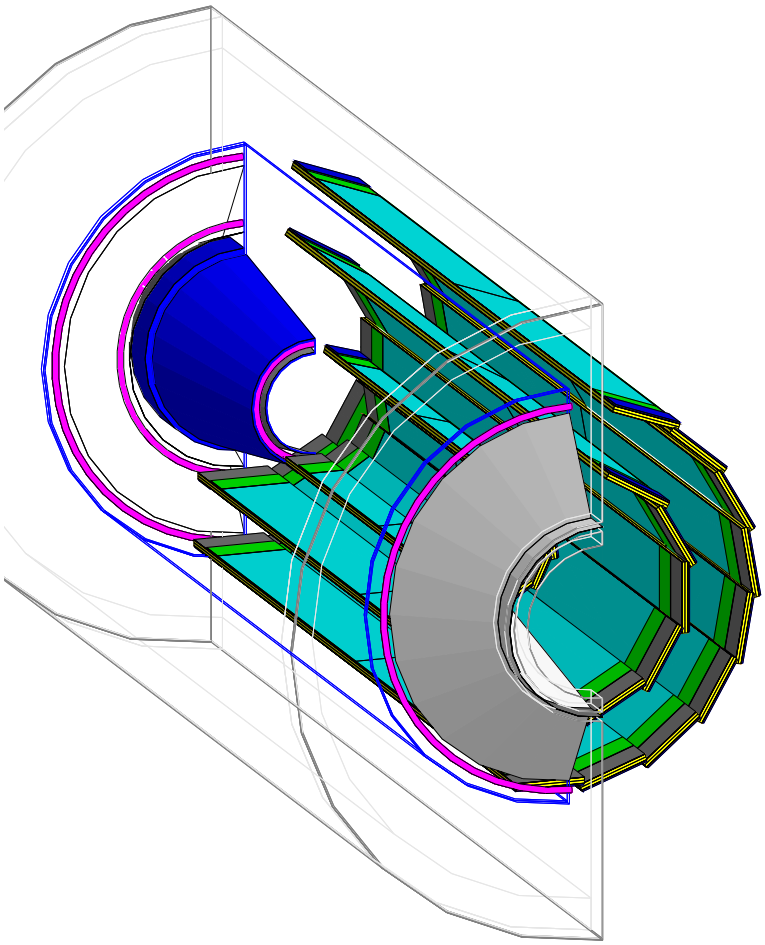


Fig. 6.1: Visualisation of the VXD, as described inside the simulation framework used for the DBD studies. The image is separated in two hemispheres for better visibility. The double sided ladders structure is illustrated in the right hemisphere: with cyan colour is the sensitive part of the ladder, with yellow the support, while with green is the support electronics, located at the end of each ladder. In the left hemisphere we can see with grey colour the cabling, with pink the Titanium cooling tubes, spiralling on the transparent Beryllium support endplate and with blue the Beryllium support cone ensuring the hermeticity of the VXD. The barrel part of the Beryllium support and the cryostat are transparent.

Layer	$\sigma_{s.p.}$ (μm) MIMOSA/AROM	t_{int} (μs) MIMOSA/AROM
VXD1	3/6	50/2
VXD2	4/10	100/7
VXD3	4/10	100/7
SIT1	4/15	100/7
SIT2	4	100

Tab. 6.1: Values of spatial and time resolution per layer, for the VXD and the SIT.

The SIT geometry, as it is implemented inside Mokka for the DBD studies, is illustrated in figure 6.2. It consists of two superlayers, featuring double sided ladders. The distance between the two sensitive layers of each superlayer is 2 mm. The first superlayer is located at a radius of 154 mm, having 10 ladders, while the second one is at 301 mm and features 19 ladders. The layer's thicknesses are expected to be different for the two alternative SIT options. Concerning the strip SIT, the sensitive layer is made of 200 μm thick Silicon, while the support is made of 2 mm of graphite. Therefore, the total material budget of each superlayer is 1.36 % X_0. The pixel sensors that are going to equip the pixelated SIT feature a thickness of 50 μm of Silicon. The material budget of the support is considered to be the same as for the strip SIT. Therefore, the total material budget of each superlayer of the pixelated SIT is expected to be 1.05 % X_0.

3D Versus 2D Measurements

Silicon strip detectors measure only in a projected space, providing 2D measurement points. In order to obtain 3D spatial information from them, measurements from two layers, featuring a stereo angle between them, should be combined. Combining information from two measurement layers can lead to ambiguities in the position of the reconstructed hits. This will result in the reconstruction of a number of ghost hits. Figure 6.3 illustrates why this effect happens. The number of ghost hits is a function of the strips stereo angle (see [Baudot 11]).

As was mentioned, in the simulation, neither strips nor pixels are really described. There is an active silicon layer segmented into ladders and sensors. Therefore, the number of ghost hits should be calculated during the reconstruction stage. This task is accomplished by a dedicated Marlin processor.

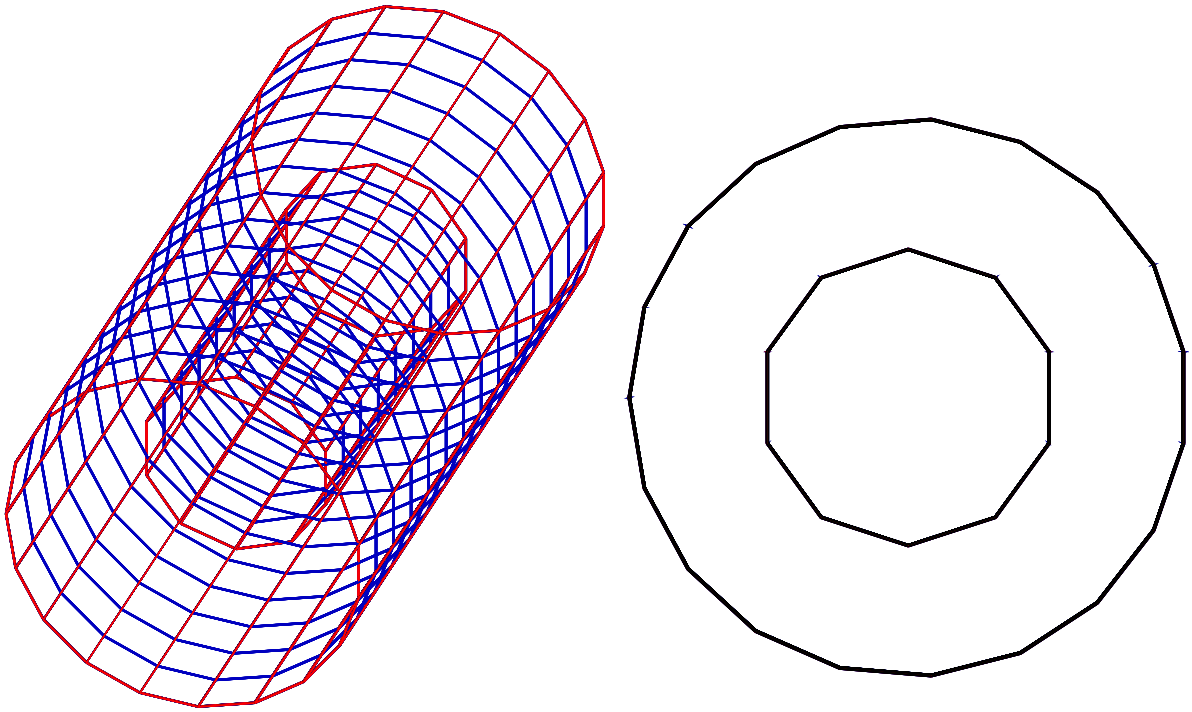


Fig. 6.2: Visualisation of the SIT, as described inside the simulation framework used for the DBD studies. Left: 3D view of the SIT. We can observe the ladder and sensor structure. Right: view of the SIT from the $R\phi$ plane.

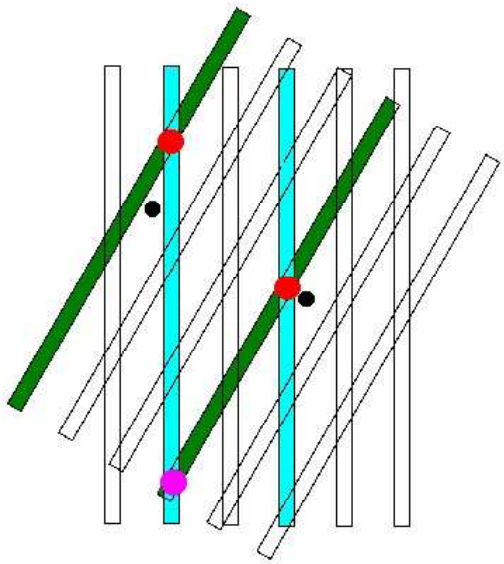


Fig. 6.3: Hit position reconstruction ambiguities with silicon strip detectors. The figure displays a false double layer equipped with silicon strips placed with a stereo angle. Coloured are the strips that give a hit signal. The two black dots represent the real hits, while their reconstructed positions are shown in red. A ghost hit, marked with magenta colour, will be reconstructed.

6.3.3 Pixel Sensors for the VXD–SIT Configuration

In table 6.1 we state some values for the time and spatial resolution of the VXD–SIT configuration. Their origin is explained hereafter. In section 3.4 we have described how a time resolution of $\sim 10\,\mu$s can be obtained for the innermost layer, and how the instantaneous power consumption can stay at $\lesssim 700\,W$ for the whole VXD. From this starting point, here we should explain how we can refine the innermost's layer time resolution to $\sim 2\,\mu$s. We also need to recalculate the power consumption expected, when the strategy of optimising one face of a double sided ladder for spatial and the other for time resolution is applied to the outer VXD layers as well. Finally, the specifications of the pixelated SIT sensors raise their own challenges, mainly due to the large area that has to be covered.

A conservative approach to improve the time resolution, which may be envisaged with the $0.35\,\mu m$ fabrication process, is to implement two discriminators per column. This way, two rows can be read out simultaneously, reducing the integration time by a factor of 2, down to $\sim 5\,\mu s$. The columns should however be enlarged well beyond $20\,\mu m$ (at the expense of the spatial resolution), and the limited number of the metalisation layers (i.e. 4 at most) hampers the possibility of achieving a compact data compression circuitry integrated in the chip.

Moreover, in order to further improve the time resolution down to the desired value of $\lesssim 2\,\mu s$, the discriminator should be implemented inside the pixel itself. This is only feasible with a smaller, $\leq 0.18\,\mu m$, feature size technology. When the discriminator is inside the pixel, a factor of two is gained directly, since there is no need to drive the analog signal to the end of the column. Additionally, more than one row can be read out simultaneously. For example, if one reads 4 rows simultaneously, the total gain in readout time is 2×4, thus from $10\,\mu s$ readout time one can go down to $\lesssim 2\,\mu s$.

Care should be given to the power consumption: more rows switched on means more power dissipated. On the other hand, that will be at least partially compensated by moving to smaller feature size processes, since they offer intrinsically reduced power consumption. The average power consumption of the whole VXD is therefore expected to be $\sim 12\,W$, a value which complies with a lightweight cooling system.

An alternative challenging approach is the following: the standard approach for the rolling shutter is to perform the latter from both the top and the bottom of the sensor, which means two rolling shutters per pixel matrix. One could envisage to further subdivide the pixel matrix into submatrices, that will be read out in parallel. According to the number of the submatrices, we can gain in direct proportion in terms of time resolution, however again at the expense of an increased power consumption.

Turning to the SIT pixel sensors, the main concern is to keep the power consumption at an acceptable level, something challenging due to the large ($\sim 4\,\mathrm{m}^2$) surface needs to be covered. A value for the instantaneous power consumption $\lesssim 4\,$kW for the whole SIT seems affordable, which may be compatible with a relating discrete cooling system

assuming a detection duty cycle of $\lesssim 2\,\%$.

6.4 Background Hit Densities

For simplicity, the beam background hits will be generated via a random hit generator. The argumentation justifying this decision was developed in sub–section 5.8.3, and is also valid here. The considered hit densities are the ones expected from the nominal beam parameters. They are summarised in table 6.2, as they were calculated in [De Masi 09] (for the VXD), [Vogel 08] (for the SIT).

Layer Layer	Radius mm	nominal $Hits/cm^2 \times BXs$
SIT		
L2	301	10^{-3}
L1	154	2.5×10^{-3}
VXD		
L5	60	2.7×10^{-2}
L4	58	3.2×10^{-2}
L3	39	1.34×10^{-1}
L2	37	1.54×10^{-1}
L1	18	2.9
L0	16	4.4

Tab. 6.2: Background hit densities on the VXD and the SIT, estimated for the nominal beam parameters.

6.5 Studies on Tracking Performance

Our goal is to study the tracking performance of the silicon detector system, which consists of the CMOS based VXD combined with both options for the SIT. The evaluation of the various detector configurations will be based on the following criteria:

Track finding efficiency: If the majority ($\geq 75\,\%$ for the purpose of this study) of the hits associated to the track have been generated from the same particle, the particle trajectory is considered to be correctly reconstructed, thus found. Obviously, in order to perform this analysis, the Monte–Carlo information should be mapped very carefully during the reconstruction stage. The track finding efficiency should be always stated together with the rate of the ghost track reconstruction.

Track extrapolation area: As already mentioned, the uncertainties on the impact parameters δd_0 and δz_0 define the size of the track extrapolation area. The robustness versus the beam background depends strongly on the size of this area. Hereafter, when we refer to the search area or the track extrapolation area, we mean this area around the track extrapolation point whose dimensions are defined by the δd_0 and the δz_0 in the $R-\phi$ and z coordinates respectively.

The table 6.3 summarises the tracking efficiencies calculated for each detector configuration, as well as for the combined VXD plus SIT system. The features of each detector system are described in table 6.1, while the physics process used is the Higgsstrahlung (see section 6.2). 2600 Higgsstrahlung final states have been analysed for the strip SIT and 3000 states for the pixelated one. The considered beam background hit densities are the ones estimated for the nominal beam parameters.

Here we should clarify the composition of each track subsample, i.e. the total and the found tracks. The total tracks subsample consists of tracks reconstructed in the TPC, having a $\chi^2/ndf < 20$. These tracks are extrapolated to the silicon detectors, where a track–hit association procedure takes place. If at least one hit is associated, we end up with a track segment reconstructed in the silicon detectors. The calculation of the tracking efficiency is based only on the hits belonging to that segment. The track segment is considered as found if $\geq 75\,\%$ of its hits belong to the same Monte–Carlo particle. Obviously, the reported efficiency values have been calculated from the ratio of the found track segments, to the total tracks extrapolated to the silicon detectors. The reported uncertainty is purely statistical, and has been calculated using the assumption that the efficiency follows a binomial distribution.

Detector	# total tracks	# found tracks	Efficiency (%)
CMOS VXD	50285	40494	$80.53 \pm 0.18\,\%$
Strip SIT	45841	45640	$99.56 \pm 0.03\,\%$
Pixel SIT	55692	53599	$96.24 \pm 0.08\,\%$
CMOS VXD + Strip SIT	52450	43928	$83.75 \pm 0.14\,\%$
CMOS VXD + Pixel SIT	61561	51729	$84.03 \pm 0.15\,\%$

Tab. 6.3: Results on tracking efficiencies for various silicon detector configurations.

Here we should stress the two following important aspects, that are not incorporated in the studies presented in this chapter. First, as already mentioned, results concerning the tracking efficiency should always be stated together with the rate of fake tracks. But since our track sample consists of tracks reconstructed in the TPC, being extrapolated to the silicon detectors, the number of fake tracks cannot be estimated. A calculation of the rate of fake tracks would make sense only if we were using a standalone silicon tracking.

Second, the fake hits produced due to the detector noise are not taken into account. The fake hit rate depends on the sensor technology. It should in general be well below the rate of background hits. One could expect that in the SIT (or even the outermost layer of the VXD), the fake hit density would be comparable or larger than the one due to the beam background.

6.5.1 From the TPC to the Silicon Detectors

When propagating the track from the TPC to the Silicon detectors, we first meet the SIT outermost layer. It is very important to pick the correct hit there: if a wrong hit is associated to the track, the whole reconstruction will follow a wrong path. Therefore, we will first estimate the pointing accuracy from the TPC to the silicon detectors, as a function of the track's momentum. In order to do so, we will use single muon tracks having momenta of 300 MeV, 500 MeV, 1 GeV, 2 GeV and 10 GeV. The impact parameter uncertainties δd_0 and δz_0 are displayed in figure 6.4.

From the impact parameter uncertainties and the beam background hit density, one can roughly estimate the probability to find a beam background hit inside the track's extrapolation area. We can consider this area as an ellipse, which half–axes are the Most Probable Values (MPV) of the δd_0 and δz_0 distributions. The area of this ellipse, for example for 500 MeV muons, would be $0.0018\,\mathrm{cm}^2$. The contribution of the sensor's single point resolution term of the equation 6.1 is marginal compared to the one of the track error matrix propagation. We can consider the track extrapolation area to be roughly an order of magnitude larger than the ellipse area: this is in order to include the tails of the δd_0 and δz_0 distributions. Taking into account the beam background hit density on the outermost SIT layer, the probability to find a beam background hit inside the track–hit matching search area is 1.8×10^{-5}. For a pixelated SIT, where the layers's integration time expands over 200 bunch crossings, this probability becomes 3.6×10^{-3}. We should underline that this number expresses the probability that one beam background hit exists inside the track extrapolation area, and not the probability of a false track–hit matching.

The aforementioned probability is inconsistent with the results we have obtained through our simulation studies. For the same muon sample ($P_T = 500\,MeV$), when we extrapolate the TPC track segment to the outermost SIT layer, in 3.7 % of the cases we associate a beam background hit to the track. This number is one order of magnitude larger than the probability of having one beam background hit inside the track extrapolation area. This impressive contradiction between our initial estimation and the simulation results may be explained by the following fact: for each 500 MeV muon, $\sim$ 14 track segments are reconstructed in the TPC. Obviously, all these track segments cannot be matched with hits in the silicon detectors generated by the initial muon. We will come back in the topic of the proper selection of the tracks that will consist our track sample, when discussing the combined VXD+SIT systems tracking

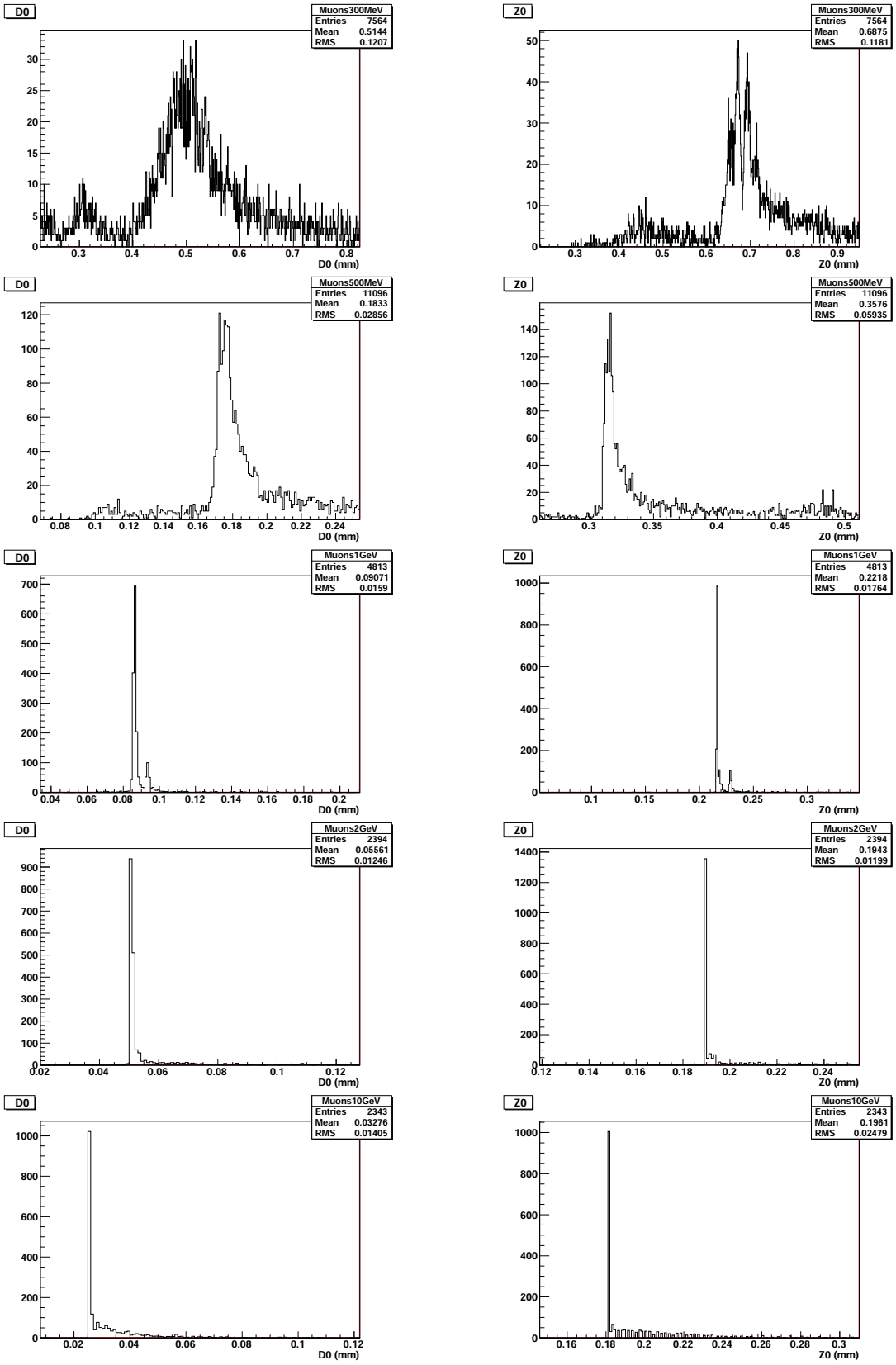


Fig. 6.4: Uncertainties on the extrapolation from the TPC to the outermost SIT layer of the impact parameters δd_0 and δz_0, for muon tracks featuring various momenta. Left: δd_0. Right δz_0. Going from top to bottom the muons momenta are 300 MeV, 500 MeV, 1 GeV, 2 GeV and 10 GeV.

performance.

6.5.2 Strip–Pixel SIT Comparison

A first overview of the results summarised in table 6.3, pushes us to conclude that the strip SIT features an excellent tracking efficiency, significantly higher than the one of the pixelated SIT. As mentioned earlier, the wrong track–hit association may originate either from a beam background hit, either from a hit generated by a second particle, or from a hit reconstructed due to the combinatorial background (ghost hit). The latter is an effect specific to the strip sensors. There is a different way to deal with each of the above cases, in order to reduce the probability of a wrong track–hit matching. Concerning the separation of tracks originating from different Monte–Carlo particles, a fine spatial resolution is required. The time resolution is not so relevant, considering the low ILC event rate (rate for electroweak interactions $\lesssim 1\,\text{Hz}$). On the other hand, for the beam background hits, this probability could be reduced either by optimising the time, or the spatial resolution.

The vast majority of the wrong track–hit associations is due to the beam background hits. The effect of the associated hits that belong to another physics particle is rather marginal. The same holds true for the ghost hits; even if the sensors equipping a double sided ladder having their strips oriented perpendicular to each other, which is the worst case scenario, the combinatorial background effect is marginal compared to the beam background. Obviously the above statement holds true if the hit density stays below a certain limit; we could expect that if it exceeds that limit, the subsequent increase of the hits per area will lead us to a dramatic increase of the ghost hits as well.

Therefore, our main concern is to deal with the beam background hits. It is thus worthwhile entering into the time–spatial resolution interplay between the two SIT options. The strip sensors provide an excellent time resolution (single bunch crossing timestamping), while the time resolution of the pixelated SIT is 2 orders of magnitude worse. Which means that if we consider the same search area for track–hit association, the pixelated SIT would have 2 orders of magnitude more background hits to deal with. On the other hand, the potential of the pixelated SIT lies on its spatial resolution, which results into smaller track extrapolation areas. The resolutions on the parameters d_0, z_0, of tracks belonging to the Higgsstrahlung sample, are illustrated in figures 6.5, 6.6 for both options. It is not easy to directly say how much smaller are the track extrapolation areas for the pixel SIT, because they strongly depend on the track's momentum; additionally, they change from layer to layer, since they depend on the traversed material budget and the extrapolation distance. This has some importance, since it is more important to restrict the search area for track–hit associations on the layers which have a high hit density.

For the reasons above, it will be more illuminating to compare the track extrapolation accuracy, for the two SIT options, on each layer. To better follow this discussion,

we should name the SIT layers. The left part of figure 6.2 illustrates the $R\phi$ view of the SIT. It consists of two double layers (or superlayers). Going inwards, the outer superlayer consists of the layers 3 and 2, while the inner one of the layers 1 and 0. Thus, the track segment reconstructed in the TPC is extrapolated to the layer 3 of the SIT. The strip SIT provides one 3D measurement per superlayer, while the pixelated one provides two.

Concerning the relatively high momentum tracks (the ones to which the MPV of the distributions shown in figures 6.5, 6.6 correspond), the δd_0, going from the third to the second layer, is $\sim 23\,\mu$m for the strip SIT and $\sim 5\,\mu$m for the pixelated one. This difference is explained by: first the better spatial resolution that is provided by the pixel sensors ($4\,\mu$m versus $7\,\mu$m for the strips); second by the fact that for the strip SIT, since the superlayer consisting by the layers 3 and 2 provides only one 3D measurement, there is no update of the track parameters going from the third to the second layer. Extrapolating the track from the outer superlayer to the inner one (distance $\sim 150\,$mm), the δd_0 is $\sim 12\,\mu$m and $\sim 7\,\mu$m for the strip and pixel SIT respectively. Finally, going from layer 1 to layer 0, the track extrapolation uncertainty remains almost the same for the strip SIT. The reason is that, despite that the track parameters are not updated, the extrapolation distance is two small (2 mm). Concerning the pixel SIT, we observe a small degradation of the δd_0. This is because layer 1 (see table 6.1) is dedicated to timestamping: it therefore offers a rather moderate spatial resolution ($15\,\mu$m).

We expect a more striking difference for the δz_0, since the resolution of the strip sensors in the z direction is $50\,\mu$m, while stays $4\,\mu$m for the pixel sensors. Indeed, extrapolating the track to layer 2, the δz_0 values are $\sim 180\,\mu$m for the strip SIT and $\sim 5\,\mu$m for the pixel SIT. Going from the outer superlayer to the inner one, the values become $\sim 56\,\mu$m and $\sim 15\,\mu$m respectively. These values remain effectively unchanged when going from layer 1 to layer 0.

In figures 6.5, 6.6, we observe that the distribution of the extrapolated impact parameter uncertainties sometimes appear to have a multiple peak structure. This is an effect of the track fitting code. Sometimes, even though a candidate hit is found inside the search area, it is not associated to the track because the fitting is not successful. In that case, the track is just extrapolated to the next layer without its parameters being updated. On the contrary, the impact parameter uncertainties undergo a small degradation, caused by the extrapolation distance and the crossed material. For example, in the top plot of figure 6.6, which displays the δz_0 when extrapolating the track from the third to the second layer of the pixel SIT, we observe a secondary peak at $180\,\mu$m. This is almost equal to the value of δz_0 when extrapolating the track from the TPC to the third SIT layer (see bottom right plot of figure 6.4). If the track parameters are not updated, the δz_0 will remain more or less the same when going from the third to the second layer. Here we should stress that this effect has nothing to do with the detector itself, therefore we will not take into account in the evaluation of the silicon

detector performance. It is an effect relevant only to the track fitting code. By the time this thesis was written, the tracking tools were substantially improved with respect to the ones used for the LoI studies, but they were still in an evolving stage towards the DBD. Issues like that are thus quite reasonable to occur.

Going back to the strip–pixel SIT comparison, we observe that with the exception of the δz_0 when extrapolating from layer 3 to layer 2, the hit search area is a few times smaller for the pixel SIT than for the strip SIT. From both figures 6.5, 6.6, we observe that the strip SIT δd_0 and δz_0 distributions have larger high tails (which corresponds mostly to lower momentum tracks) than the pixelated one. Reminding that the time resolution of the strip SIT is 2 orders of magnitude better, while the track extrapolation is only a few times more precise for the pixelated one, the results of the tracking efficiency comparison between the two options can be understood. Even in the case that one of pixel SIT layers is equipped with sensors featuring elongated pixels, being dedicated to timestamp the track, (this is the case for the results reported in table 6.3 and in figures 6.5, 6.6) the tracking efficiency performance of the strip SIT cannot be reached.

But before stating that a SIT equipped with silicon strip sensors is the optimal solution, we should proceed to a more profound analysis. Reminding sub–section 6.3.2, the strip SIT provides 2D measurements that can be combined to up to 2 3D measurements, while the pixelated one can provide up to 4 3D measurements. Moreover, pixelated SIT offers more precise measurements. Therefore, the comparison of the two alternative SIT options should take place inside the framework of the whole tracking scheme. The abundance of the reconstructed hits is an essential issue for the performance of a tracking system. Therefore, we will not conclude which is the best SIT option before we study their performance combined with the VXD, in sub–section 6.5.3.

6.5.3 Silicon Detectors Combined System

Table 6.3 reports the tracking efficiency performance of the CMOS based VXD, alone or combined with the two alternative SIT options. Due to the much higher beam background hit densities, the wrong track–hit associations taking place in the VXD dominate the overall inefficiency. For both options, the tracking efficiency is $83-84\,\%$. These values indicate a rather low performance. But we have strong indications to believe that this low performance stems from the composi tion of the track sample. The track sample is composed from all track segments reconstructed in the TPC, that feature a $\chi^2/ndf < 20$. Therefore, inside that sample, there should be a significant number of tracks originating from photon conversions in the material of the detector, or π and K decays. This assumption is verified by the momenta spectrum of the tracks sample: many tracks do not have large enough transverse momentum in order to reach the TPC. Obviously, these tracks do not originate from the IP. Subsequently, they may not have generated hits in the silicon detectors (or at least in all of their layers). In

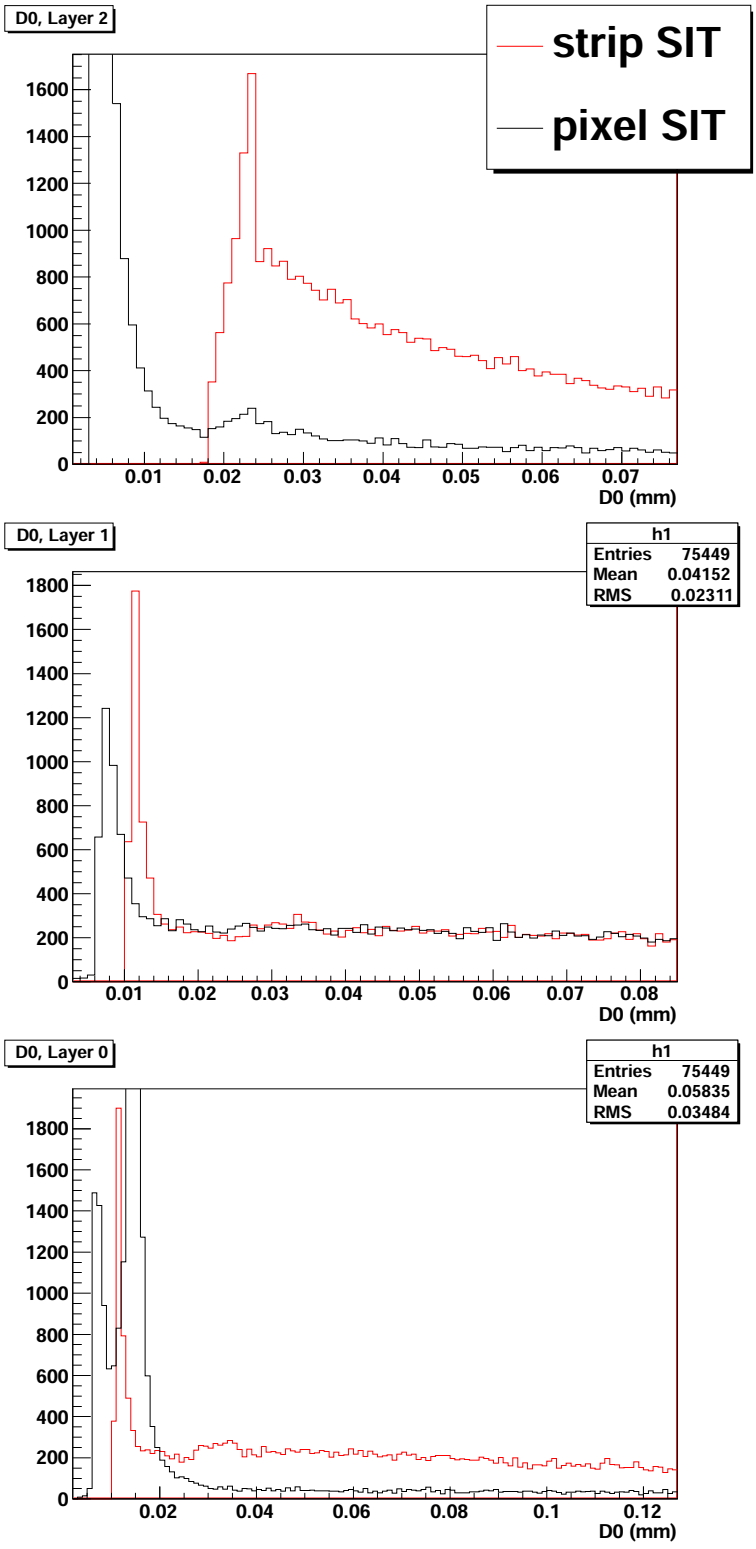


Fig. 6.5: Uncertainty δd_0 on the track extrapolation for the strip SIT (red) and the pixel SIT (black). Top: extrapolation from layer 3 to layer 2. Center: extrapolation from layer 2 to layer 1. Bottom: extrapolation from layer 1 to layer 0.

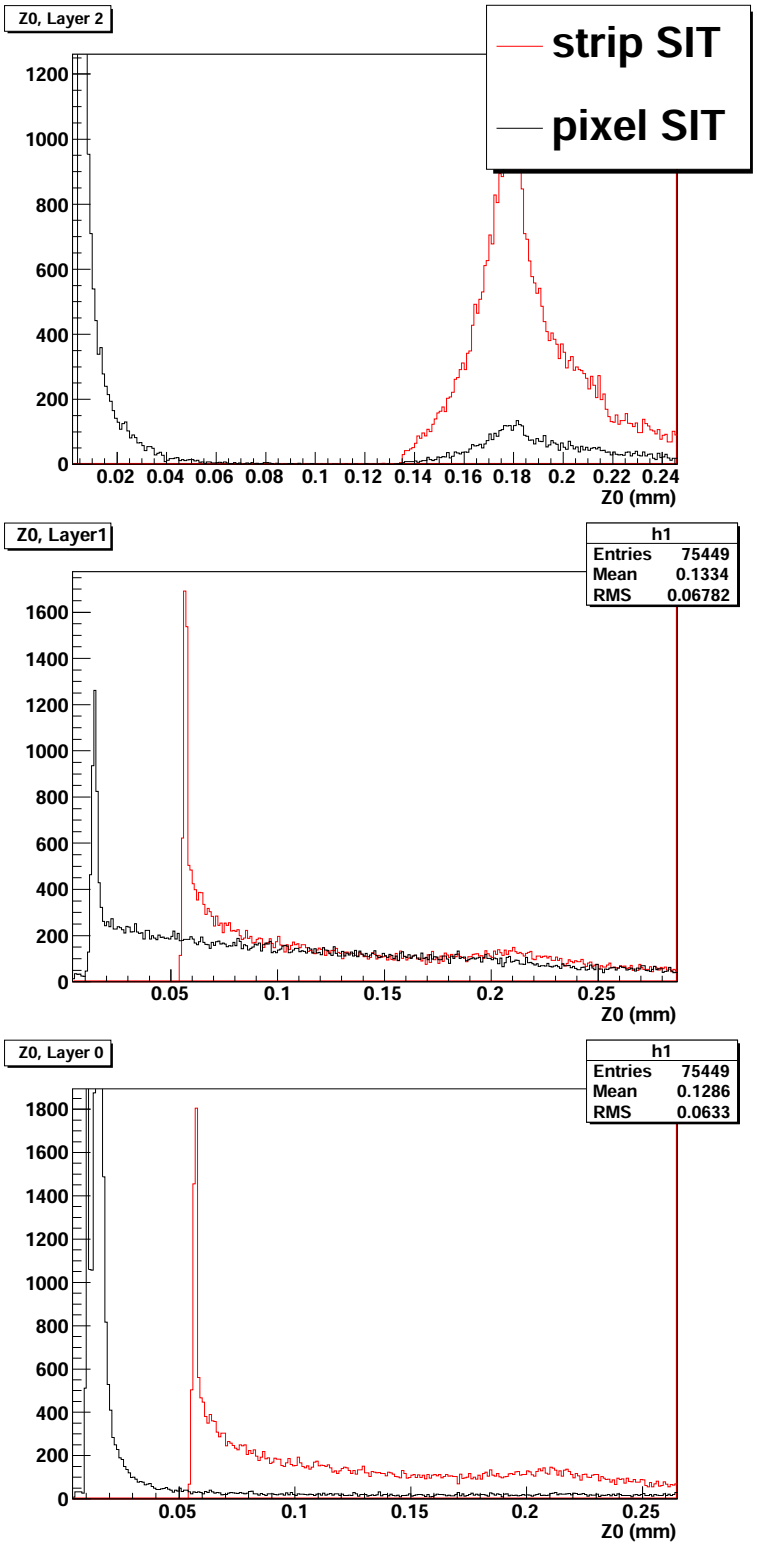


Fig. 6.6: Uncertainty δz_0 on the track extrapolation for the strip SIT (red) and the pixel SIT (black). Top: extrapolation from layer 3 to layer 2. Center: extrapolation from layer 2 to layer 1. Bottom: extrapolation from layer 1 to layer 0.

that case, when we extrapolate these track segments from the TPC inwards to the silicon detectors, we are bound to make wrong track–hit associations. An appropriate selection of the track sample needs to take place, in order to evaluate the tracking efficiency of the silicon detectors combined system reliably.

We have tried to investigate this assumption by applying some provisional, naive cuts to the parameters d_0 and z_0. As a first test, any tracks with $d_0 > 2\,\text{mm}$ and $z_0 > 5\,\text{mm}$ were rejected. We expected that this way the tracks that originate far from the IP, thus they may have not generated hits in the silicon detectors, are removed. This is admittedly a naive selection that will compromise the detector's physics performance, since a number of tracks originating from secondary vertices stemming from b and c hadron decays will be also removed. But the main point of this test is not to optimise the physics potential of the experiment, but rather to evaluate the VXD+SIT tracking efficiencies performance: if the lost tracks originate from particles that have indeed crossed the VXD+SIT system, then we should rather rethink the design of the VXD and its sensors specifications. After applying the aforementioned cuts, the VXD+SIT system tracking efficiency improves to 90–91 % (for both SIT alternatives). This is a quite encouraging result. However, a more refined track selection has to be applied.

Additionally, we may should rethink which is the lower limit on momentum for the track reconstruction scheme, when the tracks are reconstructed in the TPC and then being extrapolated and pick up hits from the silicon detectors. This can be highlighted by the results presented in sub–section 6.5.1, where on average ~ 14 TPC tracks were reconstructed for each 500 MeV muon. The issue of looping tracks in the TPC, is very difficult to be addressed. Roughly, particles with momentum less than 1 GeV/c, do not cross the TPC. Therefore, for this range of momenta, in order to obtain reliable results for the silicon detectors tracking efficiency, the looping tracks should be removed first.

A second important observation, is that the overall performance of the silicon tracking system is similar whether the SIT is equipped with strips or pixel sensors. This comes to verify that despite the fact that the strip SIT has significantly higher tracking efficiency, it is compensated by the higher number of precise 3D measurements offered by the pixel SIT. The above statement could be highlighted by figure 6.7. This figure shows the track extrapolation uncertainties δd_0 and δz_0 when we extrapolate the track from the SIT to the outermost VXD layer. The pixel SIT, especially for δz_0, provides a much more precise extrapolation. Thus it reduces significantly the probability for a wrong track–hit association in the outermost VXD layer. Concluding, we should underline that the performance of the combined strip SIT–VXD system remains slightly better.

6.5.4 Silicon Detectors Configuration Optimisation Studies

We are going to proceed to a number of studies that will explore a small part of the available parameter space for the optimisation of the combined VXD–SIT system.

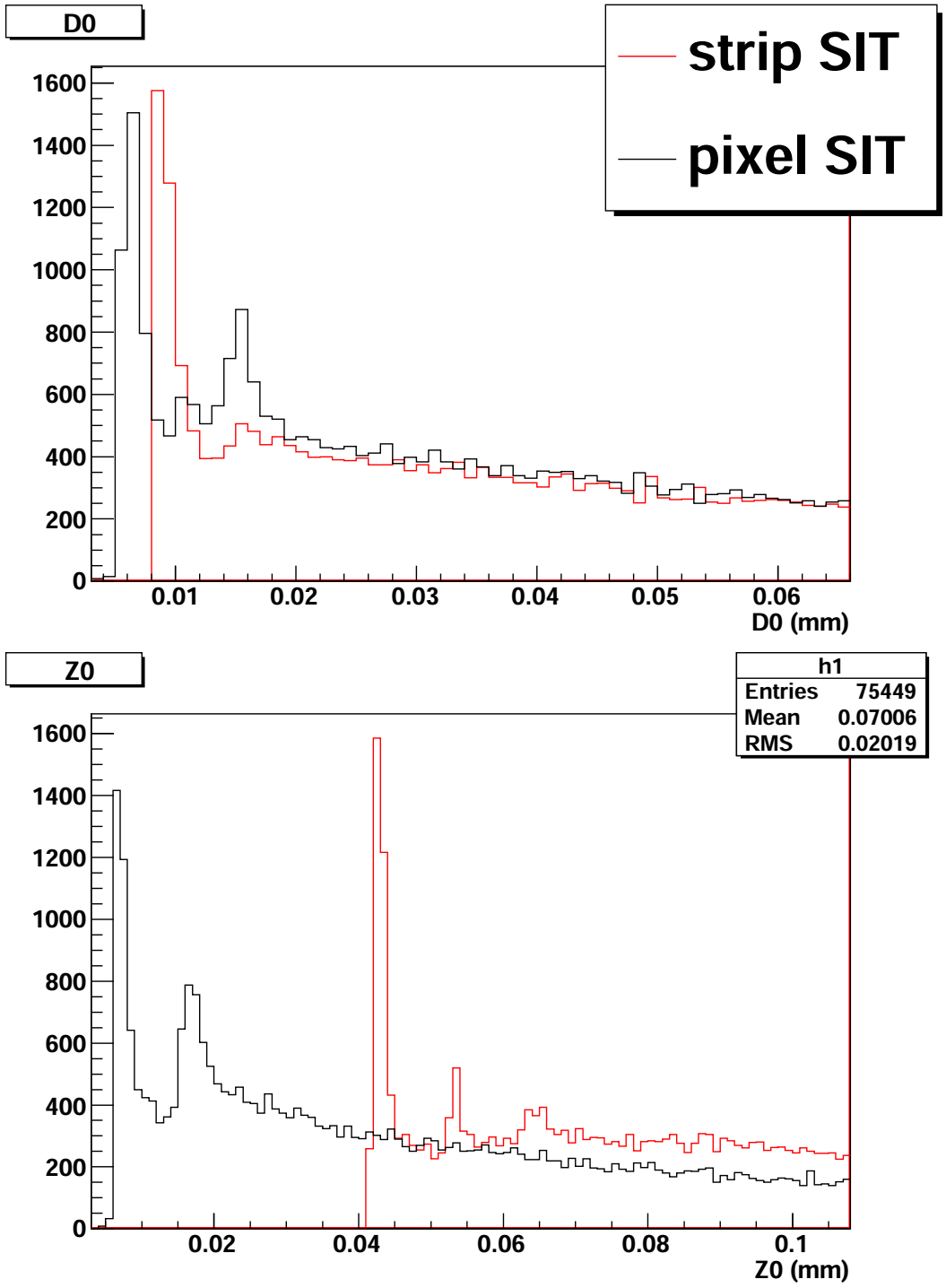


Fig. 6.7: Extrapolation of the δd_0 and δz_0 from the strip SIT (red) and the pixel SIT (black) to the outermost VXD layer. Top; δd_0 and bottom; δz_0

Of particular interest is a hybrid SIT design, where one of the SIT layers will be equipped with strips and the other with pixel sensors. For convenience, let us name as HybridSIT1 the design where the inner layer of the SIT is equipped with strip sensors, and call the one where the outer layer is equipped with strip sensors HybridSIT2. Finally, we will test the standalone performance of a pixel SIT design, where there is not any layer dedicated to timestamping. The motivation for this latter study is actually to estimate the added value of this timestamping layer. The results of the above tests are reported in table 6.4.

Detector	total tracks	found tracks	Efficiency (%)
Pixel SIT (no timestamping)	48933	46202	94.42 ± 0.10 %
HybridSIT1	47821	46725	97.71 ± 0.07 %
HybridSIT1+VXD	53337	44539	83.50 ± 0.16 %
HybridSIT2+VXD	53078	43981	82.86 ± 0.16 %

Tab. 6.4: Results on tracking efficiencies for two alternative hybrid SIT designs, and a pixelated SIT without a dedicated timestamping layer.

The pixelated SIT design without a timestamping layer demonstrates a probability of 94.42 ± 0.10 % to reconstruct correctly a track segment. This value should be compared with the 96.24 ± 0.08 % probability, exhibited by the pixelated SIT that includes a timestamping layer(see table 6.3). The above result establishes the added value of such a layer.

The hybrid SIT design, specifically the HybridSIT1, features a probability for a correct reconstruction of a track segment equal to 97.71 ± 0.07 %. This probability value places this design in a position of being slightly better than the pixelated SIT (96.24 ± 0.08 %), and slightly worse than the strip one (99.56 ± 0.03 %). So, we proceed to characterise it as a part of the combined VXD–SIT system. The tracking efficiency of the HybridSIT1+VXD configuration was found to be 83.50 ± 0.16 %, which is comparable with the strip SIT plus VXD and the pixel SIT plus VXD configurations. In other words, we did not observe any striking results by combining strip and pixel sensors. Despite the fact that the hybrid design exhibits a better track–hit matching efficiency concerning the reconstruction of the SIT track segment, compared to the pixelated design, its pointing accuracy to the VXD is degraded. This effect is clearly demonstrated in figure 6.8. This figure shows the extrapolation of the track impact parameter uncertainty δz_0 from the hybrid and the pixel SIT to the outermost VXD layer. Reminding the figure 6.7, it is exactly the extrapolation of the δz_0 that demonstrates the pointing accuracy of the pixelated SIT. As we can observe, this pointing accuracy is severely degraded if we substitute one pixelated layer with one equipped with strip sensors. The reason for that is the degraded spatial resolution of the strip

sensors, and the fact that the hybrid design can provide up to three 3D measurements, compared to the four offered by the pixelated one.

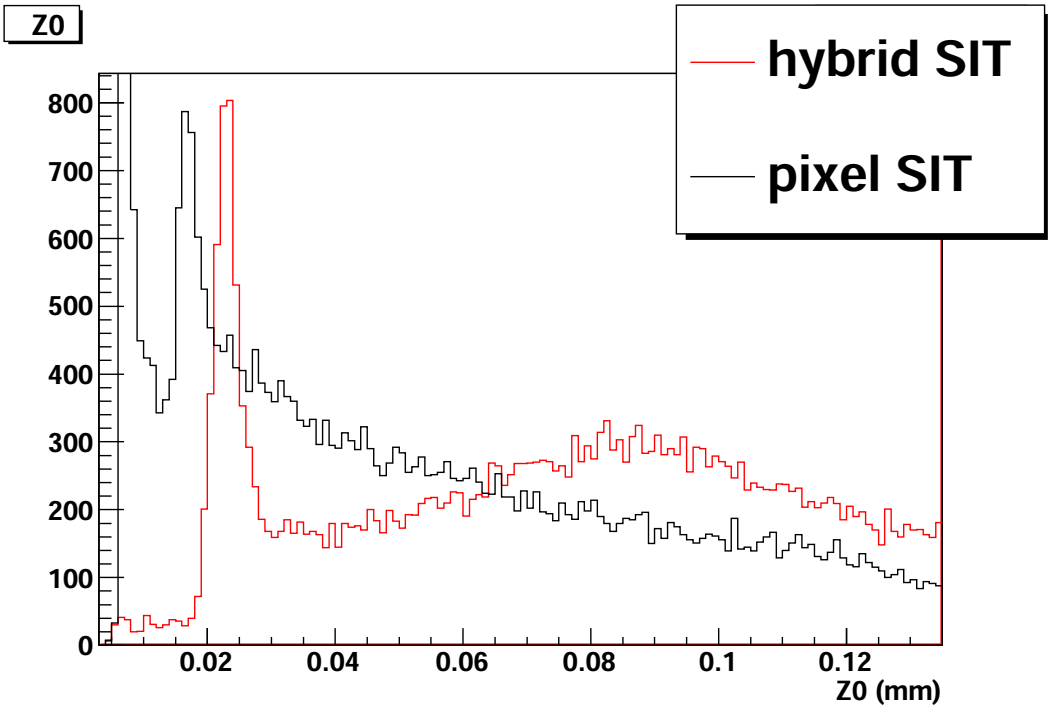


Fig. 6.8: Extrapolation of the δz_0 from the hybrid SIT (red) and the pixel SIT (black) to the outermost VXD layer.

6.6 Conclusion and Perspectives

In this chapter we have tried to optimise the VXD and SIT detectors based on their tracking performance. After conducting the VXD optimisation studies, based on the heavy flavour tagging and the Higgs hadronic branching ratios extraction, presented in the previous chapter, an element we were missing was the robustness of the VXD versus the beam background hits. Moreover, we were motivated to examine the combined VXD and SIT system, and probe the potential of a pixelated SIT. Here, we have tried to address these questions. The means used in order to characterise the above systems were the tracking efficiency and the extrapolation of the track impact parameter uncertainties δd_0 and δz_0. We have used a more realistic GEANT4 detector description, the one that is going to be used for the DBD studies. Also, we have focused on a VXD equipped with CMOS sensors, in contrast with the previous chapter studies that were considering sensors of a generic technology.

Comparing the two main SIT options, the standard one equipped with strip sensors versus the one equipped with pixel sensors, we observed that the strip SIT offers a much better track–hit association: $99.56 \pm 0.03\,\%$ of the track segments reconstructed in the SIT could be considered that they are reconstructed correctly, according to the criteria we have set. For the pixel SIT, the percentage of the correctly reconstructed track segments is $96.24 \pm 0.08\,\%$. These results are due to the excellent time resolution of the strip sensors; the higher spatial resolution of the pixels could not balance the two orders of magnitude worse time resolution.

On the other hand, the four high precision 3D measurements offered by the pixel SIT provide a better link between the track segments reconstructed in the TPC and the VXD. Overall, the tracking efficiencies of the combined VXD+SIT systems are therefore quite similar: $83.75 \pm 0.14\,\%$ in case of a strip SIT and $84.03 \pm 0.15\,\%$ for the pixel SIT case.

These values of tracking efficiencies for the combined VXD+SIT system are rather worrying. Our assumption is that the lost tracks are mostly coming from photon conversions and pions/kaons decays: in that cases, we can reconstruct a track segment in the TPC, but it is not possible to match it with any segment reconstructed in the silicon detectors. Thus, when we extrapolate the track from the TPC inwards to the silicon detectors, we are bound to associate it with wrong hits. This obviously degrades the tracking efficiency performance. A second argument supporting this assumption, is that tracks having momenta less than $1\,\mathrm{GeV/c}$ are looping inside the TPC. This issue is very difficult to be addressed. An important question, which has been set also in chapter 5, is whether a standalone silicon tracking is required in order to properly reconstruct the low momentum tracks.

Another argument in the strip–pixel SIT comparison could be drawn out of momentum resolution studies. Maybe one of the alternative silicon detector configuration has a higher potential in reconstructing the momenta of low momentum tracks. These studies are still on–going.

The studies conducted for this chapter showed that the strip SIT exhibited better standalone performance. The performance of the combined VXD+SIT system are similar for both SIT options. However, there is still some arguments that may favour the pixelated SIT option. The strongest of this arguments is the perspective to develop a local pattern recognition algorithm for a standalone tracking in the ILD silicon detectors. For the formation of a track seed inside a magnetic field, we need three 3D points that will initiate a track helix. In the case of the strip SIT, these three points can only be provided by the outer and the inner superlayer of the SIT, and the outermost layer of the VXD. These layers are located at radii of 301 mm, 154 mm and 60 mm respectively. On the one hand, the relatively large distance between the hits is something good, since it provides a big lever arm and thus a more accurate track extrapolation. On the other hand, the probability to include wrong hits in the track seed, thus biasing

from the beginning the track reconstruction could be quite high. The reason is that the stereo angle sector of the detector that is going to be used for the seed formation will be also large, therefore an elevated number of seeds that contain wrong hits will be formed.

The track seed on the pixel SIT could contain two, 2 mm distant hits, from one superlayer and one more hit from the other superlayer. The small distance between the pair of hits located on the same superlayer could ensure us that these two hits are indeed belonging to the same track (especially in the low hit density environment of the SIT). In the same time, a large lever arm is provided. Moreover, the pixelated SIT alone could provide a track seed, or even a track segment. Nevertheless, the above argumentation concerning the standalone silicon tracking should be supported with tracking studies, or even better with physics analyses results. A study on flavour tagging performance of the various silicon detector configurations, given the existence of a local pattern recognition algorithm for a standalone silicon tracking, may lead to final conclusions about the optimal configuration.

Other motivations for going to a pixelated SIT could be offered by the exploitation of the benefits provided by double sided ladders equipped with pixel sensors; the argumentation is similar to the one developed to support the double layers VXD geometry. To remind them briefly, these benefits could be:

- The development of a silicon tracking algorithm based on mini–vectors.
- More accurate sensor and ladder alignment.
- The possibility to correlate beam background hit clusters in the 2 mm distant layers and reject them offline.

The study of a hybrid SIT design, which combines strip and pixel sensors, did not provide us with any striking results. Thus such a design is not suggested, at least from the point of view of track finding efficiency. Nevertheless, the argumentation that a decisive conclusion can only be drawn through physics analyses that includes a standalone silicon tracking, is also valid here.

Finally, the study exposed in this chapter may then be revisited once the real potential of the new, fast, pixel architecture explored by these sensor prototypes will be estimated on the basis of real measurements.

CHAPTER 7

Summary and Conclusion

The ILC is expected to offer unique possibilities of investigating the EWSB, the origin of mass and the physics beyond the SM. This machine of unprecedented potential is however only relevant if high precision experiments, typically 10–100 times more sensitive than at LHC, are achievable, allowing to fully take advantage of the machine precision, flexibility and running conditions. This applies in particular to the vertex detector, the key sub–system for flavour tagging and a component essential for low momentum track reconstruction. A highly performant flavour tagging is in particular required to extract the Higgs hadronic branching ratios with high sensitivity, which is a key measurement to unravel the nature of the Higgs boson.

The central topic of this thesis was the optimisation of the vertex detector (VXD) of ILD, one of the two experimental concepts developed within the ILC project, and to contribute to the development of CMOS sensors that will satisfy the VXD requirements. This thesis was conducted during the period of the finalisation of the LoI document and the preparation for the DBD. It has contributed to the LoI, but the main idea was to use the LoI studies as a starting point, and continue them towards the DBD.

The LoI studies have proven that the VXD provides the necessary performance to obtain the desired precision in the extraction of the Higgs hadronic branching ratios, therefore in the measurement of the Higgs couplings. For the purposes of the study it was assumed a double sided layers geometry, and a generic and rather ideal sensor technology that can fulfil all the VXD requirements. The feasibility of the study was proven, but there are still open issues that motivated this phd. These issues are mainly the level of realism of the LoI studies, and the optimisation of the VXD geometrical configuration.

Concerning the geometry, the main point was the comparison between the two

alternative VXD geometries, VXD03 which features 5 single layers, and VXD05 which features 3 double sided layers and was the one used for the LoI. To do that, we have probed into the ability of the detector to measure the physics observables describing the final states, by performing full simulation studies. The selected physics channel was the Higgs boson production via the Higgsstrahlung process $e^+e^- \to ZH \to \mu^+\mu^- X$, at $\sqrt{s} = 250\,\mathrm{GeV}$, assuming a Higgs mass of $120\,\mathrm{GeV/c^2}$, which is a benchmark process of the ILC physics program.

Firstly, we evaluated the flavour tagging performance for the two main candidate geometries. A dedicated neural net training and parameter tuning was performed for each geometry. The obtained results showed that both geometries demonstrate excellent b–tagging and satisfactory c–tagging performance that meet the ILD physics goals. The two geometries did not show any significant performance differences. The second test was the evaluation of the sensitivity of each candidate geometry to the extraction of the Higgs hadronic branching ratios. The most important physics background that may contaminate our selected process, which arise from the double Z boson production process, was accounted for. The relevant cross sections were estimated by assuming beam polarisations of 80 % for e^-_R and 30 % for e^+_L. The Higgs branching ratios to beauty and charm quarks, and gluons, were calculated using a template fitting technique, which is assumed to provide the highest sensitivity. Both geometries exhibited satisfactory performance, without any striking differences between them.

Even though we have said that the CMOS sensors R & D and the simulations concerning the VXD geometry optimisation are not directly related, we can still extract from the simulations very useful indications that can guide the direction of the CMOS sensor development. To do that, we have substituted the LoI target values for the VXD specifications with more realistic ones. The purpose of this study was, on the one hand, to see whether the CMOS sensors performance and features are adequate to fulfil the ILD physics goals, and on the other hand, if not, to find out on which issues we have to focus our R & D efforts. The parameter space scanned included the material budget of the ladders, the distance of the innermost layer from the IP and the single point resolution of the outermost layers. Each time a parameter was modified, a new set of neural nets was trained and the flavour tagging performance reevaluated. The b–tagging performance were found to be extremely robust. The c–tagging was more sensitive to modifications. Its performance will be significantly degraded if:

- The ladder's material budget exceeds 0.3 % X_0. Thus, useful conclusions can be drawn concerning the margins inside which can one relax the ladder's material budget requirements without severely degrading the physics potential of the VXD.

- The innermost layer will be moved away from the IP, at a radius of 20 mm. Therefore, it is suggested to stay at a radius of 16 mm. This may impose either

a faster readout of the innermost layer's sensors ($\sim 10\,\mu$s) and/or the implementation of the anti–DID field which can substantially decrease the hit density due to backscattered beam induced background particles.

Concerning the outer layers single point resolution, scanning single point resolution values in the range of $2.8\,\mu\text{m} - 5\,\mu\text{m}$, the flavour tagging performance degradation was negligible. This is quite encouraging for the efforts to minimise the power dissipation.

Up to this point, the effect of the beam background hits was not taken into account directly in our studies. This consists an important missing point, since it dictates the running constraints of the VXD. Including the effect of the beam background is the safest way to extract conclusions concerning the required time resolution of the CMOS sensors. As was earlier mentioned, improving the time resolution is the main challenge that the CMOS sensors technology has to confront. Time resolution comes at the expense of spatial resolution, thus this study will give useful points to our effort to find the best trade–off among the sensors specifications. Since the maturity of the tracking tools does not allow us to perform studies of physics final states, we have tried a different approach.

Therefore the third part of this thesis was addressing the tracking performance of the combined VXD–SIT system, and tried to evaluate its robustness versus the beam background. We have implemented in the simulation a realistic CMOS based VXD design, with realistic parametrisation that follows close the sensors and integration techniques developments. Concerning the SIT, we have mainly focused on two alternative designs, the standard one equipped with strip sensors and an alternative one, equipped with pixel sensors. The beam background effect has been included in this study. Our goal was to evaluate the tracking performance, meaning the track pointing accuracy, the tracking efficiency of the track segment reconstructed at the silicon detectors and the momentum resolution, of the combined silicon detector system. On the other hand, the obtained performance and the robustness versus the beam background, could provide the main directions that the sensors and integration techniques R & D should follow.

The combined VXD–SIT configuration provides a tracking efficiency of $\sim 83-84\,\%$. These values are obtained whether a strip or a pixelated SIT is considered. They are rather low. We assume that this low performance is an outcome of the non–refined track selection of our track sample; most of the lost tracks should originate from photon conversions and pions/kaons decays. Studies to support this argument are on–going.

Concerning the pixel–strip SIT comparison, it was observed that the strip SIT provides an excellent track–hit association (99.56 % tracking efficiency on the silicon detectors track segment). The pixelated SIT features a more compromised track–hit association (96.24 % tracking efficiency), but on the other hand it provides four high precision 3D measurements, instead of two of the strip SIT, and also a better pointing accuracy towards the VXD. Thus, examining it as a part of the overall track recon-

struction scheme, could be a better option. Final conclusion concerning the CMOS based VXD requirements and the best SIT design will be drawn by a complete physics analysis, like for example heavy flavour tagging, performed with a standalone silicon tracking based on a local pattern recognition algorithm.

This thesis has been conducted inside the PICSEL group of Strasbourg, which is the pioneer in developing CMOS sensors for HEP experiments. Therefore, we have focused on this specific technology. The CMOS sensors technology, combining high granularity and low material budget with a swift readout and satisfactory for the ILC environment radiation tolerance, is a promising candidate for the VXD sensors. The contribution of this thesis to the CMOS sensors R & D consisted in the analysis of test beam data of two sensors, one (MIMOSA 22/22bis) addressing the question of read–out speed and one (MIMOSA 24) investigating a new CMOS process.

The main challenge that the CMOS sensor technology has to confront is to realise a fast readout for the inner VXD layers, while preserving their other performance. In order to cope with the hit flux expected in the ILC environment, and keep the pixel occupancy at acceptable levels ($\lesssim 1-2\,\%$), a column parallel readout has to be integrated in the sensors, together with a data sparsification microcircuit. An important step towards this goal was the realisation of the MIMOSA 22/22bis sensors, which among others, have addressed and validated the column parallel readout. These sensors were the Intermediate Digital Chips (IDCs) that led us to MIMOSA 26, a $1 \times 2\,\mathrm{cm}^2$ sensor equipping the beam telescope of the EU project EUDET, featuring integrated data sparsification.

The observed performance of MIMOSA 22bis were fully satisfactory for the EUDET telescope, and very promising for future applications; it exhibited an efficiency of $\sim 99.8\,\%$ for an average fake hit rate per pixel of $O(10^{-5})$ and a resolution of $\lesssim 3.5\,\mu\mathrm{m}$. This performance is preserved in acceptable levels after irradiation with an integrated dose of 150 kRad, which is a pessimistic estimate of the expected annual dose in the ILC environment. MIMOSA 26 could offer the baseline architecture for the VXD sensors: it already satisfies most of the requirements for the sensors that are going to equip the outer layers of the VXD. On the other hand, in order to fulfil the innermost's layer requirements, its readout time should be decreased by an order of magnitude, and its spatial resolution improved to $\lesssim 3\,\mu\mathrm{m}$.

The scope of MIMOSA 24 was to investigate a new fabrication process (XFAB $0.35\,\mu\mathrm{m}$). The main motivation to explore this process was the intention of XFAB to move to high resistivity epitaxial layers. MIMOSA 24 is not directly related with the effort for the development of the VXD final sensors. Moving to high resistivity epitaxial layers will substantially improve the tolerance to non-ionising radiation, which is not a severe issue in the ILC environment. However, the exploration of a new fabrication process is a very important part on the CMOS sensors technology R & D. Very interesting in particular was the comparison between MIMOSA 24 and MIMOSA

9, the latter being fabricated in the AMS 0.35 μm OPTO process, which is a very well tested process that exhibits satisfactory performance for charged particle tracking applications (e.g. MIMOSA 22). The main outcome was that no striking deviations in the performance of the two sensor were observed, thus moving to the XFAB 0.35 μm process was shown to be an open possibility.

Concluding, we can claim that the VXD is well suited to the ILC physics program. To evaluate this outcome, one should consider the level of realism of the simulation studies performed in this thesis. Adding more realism is required. The effect of beam induced background hits, a more realistic material budget description and the estimation of the spatial resolution degradation due to misalignment will degrade the VXD performance, thus they should be accounted for in future physics analysis. Nevertheless, these effects are expected to be compensated by the new, fast pixel sensors which are under development. Improving the sensors time resolution by an order of magnitude, while keeping the remaining sensors specifications at the same levels, is expected to preserve the VXD physics potential.

Bibliography

[3DIC] 3DIC. Consortium for developing 3d/vertically integrated readout electronics and sensors [Webpage]. Available from http://3dic.fnal.gov/.

[Aguilar-Saavedra] J.A Aguilar-Saavedra *et al. Tesla TDR, part III, Physics at an e^+e^- Linear Collider*. hep–ph 0106315.

[Ahn 06] M. H. Ahn *et al. Measurement of neutrino oscillation by the K2K experiment*. hep–ex/0606032v3, 2006.

[AIDA] AIDA. Advanced european infrastructure for detectors at accelerators (eu-fp7) [Webpage]. Available from http://aida.web.cern.ch/aida/index.html.

[Aplin 11] S. Aplin. *ILD Tracking, LCWS 2011*. Granada, Spain, 26-30 September 2011. Available from http://goo.gl/SU9CU.

[Arkani-Hamed 98] N. Arkani-Hamed *et al. The hierarchy problem and new dimensions at a millimeter*. Physics Letters B, vol. 429, pages 263–272, 1998.

[Arnold 05] L. Arnold *et al. TNT Digital Pulse Processor, Real Time Conference, 14th IEEE–NPSS*. 2005. Available from http://www.iphc.cnrs.fr/-tnt-.html.

[ATLAS] ATLAS. Latest results from atlas higgs search [Webpage]. Available from http://goo.gl/jTH3r.

[Bailey 09] D. Bailey *et al. The LCFIVertex package: Vertexing, flavour tagging and vertex charge reconstruction with an ILC vertex detector.* NIM A, vol. 610, page 573, 2009.

[Bambade] P. Bambade *et al. Strategy to measure the Higgs mass, width and invisible decays at ILC.* hep–ph 0703173v1.

[Barlow 93] R. Barlow *et al. Fitting using finite Monte Carlo samples*, 1993. MAN/HEP/93/1.

[Battaglia 01] M. Battaglia *et al. Determination of the Higgs Profile: HFITTER*, 2001. LC–PHSM–2001–053.

[Baudot] J. Baudot. *private communication.*

[Baudot 10] J. Baudot. *Future of Low Mass Pixel Systems with MAPS, 19th International Workshop on Vertex Detectors.* Loch Lomond, Scotland, UK, June 6–11 2010.

[Baudot 11] J. Baudot. *Développements de capteurs à pixels CMOS pour la physique de hautes énergies et measure de production de étrangeté dans les collisions d' ions lourds ultra–relativistes avec l' experiénce STAR auprès du RHIC.* HDR, Université de Strasbourg, 2011. Available from http://www.iphc.cnrs.fr/IMG/pdf/HDR_JBaudot_28octobre2011.pdf.

[Bechtel 06] F. Bechtel *et al. Studies for a Photon Collider at the ILC*, January 2006. DESY 06-007.

[Bischel 88] H. Bischel. *Straggling in Thin Silicon Detectors.* Rev. of Mod. Physics, vol. 60, page 663, 1988.

[Branchina 05] V. Branchina *et al. Effective potential (in)stability and lower bounds on the scalar (Higgs) mass.* Physical Review D, vol. 72, 2005.

[Brau] J.E. Brau *et al. Higgs Branching Ratios Measurements at a Future Linear Collider.*

[Brient 04] J.C. Brient. *Measurement of the Higgs decays into WW^* at future e^+e^- linear colliders*, 2004. LC-PHSM-2004-001.

[CALICE] CALICE. Calice [Webpage]. Available from https://twiki.cern.ch/twiki/bin/view/CALICE/WebHome.

[Catani 92] S. Catani *et al.* *The k_T reconstruction algorithm for jets in deep inelastic scattering and hadron collisions*. Phys. Lett. B, vol. 285, page 291, 1992.

[Claus 04] G. Claus *et al.* A portable system for monolithic active pixel sensors characterization [Webpage]. 2004. Available from http://www.iphc.cnrs.fr/IMG/poster_systeme_d_acquisition.pdf.

[CLIC] CLIC [Webpage]. Available from http://clic-study.org/.

[Colledani 97] C. Colledani *et al.* *A Submicron Precision Telescope for Beam Test Purposes*. NIM A, vol. 372, pages 379–384, 1997.

[Damerell 05] C. Damerell. *Vertex Detectors—How to Overcome Electromagnetic Interference*. 2005 ALCPG & ILC Workshops Snownmass, USA, 2005. Available from http://goo.gl/9ms4b.

[Degerli 05] Y. Degerli *et al.* *A Fast Monolithic Active Pixel Sensor With Pixel-Level Reset Noise Suppression and Binary Outputs for Charged Particle Detection*. IEEE Trans. Nucl. Sci., vol. 52, page 3186, 2005.

[Degerli 09] Y. Degerli *et al.* *Development of Binary Readout CMOS Monolithic Sensors for MIP Tracking*. IEEE Trans. Nucl. Sci., vol. 56, February 2009.

[De Masi 09] R. De Masi *et al.* *Estimation of the background on the vertex detector of ILD from beamstrahlung*, 2009. Available from http://goo.gl/56u7B.

[Deptuch 02] G. Deptuch. *New Generation of Monolithic Active Pixel Sensors for Charged Particle Detection*. Ph.D. Thesis, Institut Pluridisciplinaire Hubert Curien, Strasbourg and UMM Krakow, 2002. Available from http://www.iphc.cnrs.fr/Grzegorz-DEPTUCH.html.

[Deveaux 08a] M. Deveaux. *Design considerations for the Micro Vertex Detector of the Compressed Baryonic Matter experiment*. PoS (VERTEX 2008) 028, 2008.

[Deveaux 08b] M. Deveaux. *Development of fast and radiation hard Monolithic Active Pixel Sensors (MAPS) optimized for open charm*

meson detection with the CBM - vertex detector. Ph.D. Thesis, Goëthe Universität, Frankfurt am Main and Institut Pluridisciplinaire Hubert Curien, Strasbourg, 2008. Available from http://www.iphc.cnrs.fr/IMG/Deveaux_michael_these.pdf.

[Dimopoulos 79] S. Dimopoulos *et al.* *Mass without scalars*. Nuclear Physics B, vol. 155, page 237, 1979.

[Dine] M. Dine. *TASI Lectures on the Strong CP Problem*. hep–ph/0011376v2.

[Djouadi a] A. Djouadi [Webpage]. Available from http://www.cern.ch/djouadi/programs.html.

[Djouadi b] A. Djouadi. *The Anatomy of Electro–Weak Symmetry Breaking*. hep–ph 0503172v2.

[Djouadi 98] A. Djouadi *et al.* Comput. Phys. Commun., vol. 108, page 56, 1998.

[Dorokhov 06] A. Dorokhov. *NMOS-based high gain amplifier for MAPS, VIth International Meeting on Front End Electronics*. Perugia, Italy, May 17–21 2006. Available from http://goo.gl/se0sm.

[Dorokhov 10] A. Dorokhov *et al.* *Improved radiation tolerance of MAPS using a depleted epitaxial layer*. NIM A, vol. 624, page 432, 2010. Available from http://goo.gl/6pkPm.

[Dulinski 04] W. Dulinski. Mimosa 9 tracker test chip submission [Webpage]. 2004. Available from http://www.iphc.cnrs.fr/IMG/MIMOSA9_info.ppt.

[Dulinski 07] W. Dulinski. *Beam telescope for medium energy particles based on thin submicron precision MAPS*. Hawai, USA, October 27, November 3 2007. Available from http://www.iphc.cnrs.fr/IMG/wd_NSS07_hawai.doc.

[EUDET] EUDET. Eudet coll. (eu-fp6) [Webpage]. Available from http://www.eudet.org.

[Feld 09] L. Feld *et al.* *DEPFET Pixel Vertex Detector for the ILC*, 2009. Available from http://www.hll.mpg.de/~lca/ilc/DEPFET-VTX-ILC.pdf.

[Gaede 10] F. Gaede. *Towards the Software Baseline for ILD, ECFA*. Geneva, 18-22 October 2010. Available from http://goo.gl/MSHVJ.

[Geant4] Geant4 [Webpage]. Available from http://geant4.web.cern.ch/geant4/.

[GEAR] GEAR. A geometry description toolkit for ilc reconstruction software [Webpage]. Available from http://ilcsoft.desy.de/portal/software_packages/gear/.

[Goldstein] J. Goldstein *et al.* *Analysis of the Higgs Hadronic Branching Ratios in the ZH to llqq channel.* Available from http://goo.gl/JuDhQ.

[GuineaPig] GuineaPig [Webpage]. Available from Guinea-Pig:https://trac.lal.in2p3.fr/GuineaPig/.

[Halzen 84] F. Halzen & D. Martin. Quarks & leptons: An introductory course in modern particle physics. John Wiley & sons, 1984.

[Heinemeyer 03] S. Heinemeyer *et al.* *Physics Impact of a Precise Determination of the Top Quark Mass at an e^+e^- Linear Collider*, 2003. CERN–TH/2003–103.

[Hig 13] Higgs hunting workshop [Webpage]. 2013. Available from http://events.lal.in2p3.fr/conferences/higgshunting2013/index.html.

[Higgs 64] P.W. Higgs. *Broken symmetries, massless particles and gauge fields.* Physics Letters, vol. 12, page 132, 1964.

[Himmi 09] A. Himmi. *A Zero Suppression Micro-Circuit for Binary Readout CMOS Pixel Sensors, TWEPP 2009.* Paris, France, 21–25 September 2009. Available from http://www.iphc.cnrs.fr/IMG/ah_suze_Twepp2009_proc.pdf.

[Hu-Guo 10] C. Hu-Guo. *Development of Fast and High Precision CMOS Pixel Sensors for an ILC Vertex Detector, International Linear Collider Workshop 2010.* Beijing, China, 26–30 March 2010.

[IDAG] IDAG. International detector advisory group [Webpage]. Available from http://www.linearcollider.org/physics-detectors/Detectors/IDAG.

[Ikematsu 09] K. Ikematsu *et al.* *A study of the forward-backward asymmetry of $e^+e^- \to t\bar{t}$ processes in all-hadronic decay channel at $\sqrt{s} =$ 500 gev with the ILD detector*, 2009.

[ILC] ILC [Webpage]. Available from http://www.linearcollider.org/.

[ISIS] ISIS. In-situ storage image sensor [Webpage]. Available from http://www.physics.ox.ac.uk/lcfi/ISIS2.htm.

[Ito 09] K. Ito *et al.* *HZ Recoil Mass and Cross Section at ILD*, 2009. Available from http://goo.gl/9jDyI.

[Jackson 97] D.J. Jackson. *A Topological Vertex Reconstruction Algorithm for Hadronic Jets*. NIM A, vol. 388, page 247, 1997.

[Koziel 09] M. Koziel. *MIMOSA 24 sensor description*, 2009. Available from http://www.iphc.cnrs.fr/IMG/mimosa24.pdf.

[Kuhl 07] T. Kuhl *et al.* *Simulation of the measurement of the hadronic branching ratios for a light Higgs boson at the ILC*, 2007. LC–PHSM–2007–001.

[LCIO] LCIO [Webpage]. Available from http://lcio.desy.de/.

[LCTPC] LCTPC. A time projection chamber for a future linear collider [Webpage]. Available from http://www.lctpc.org/.

[LoI 10] LoI. *International Large Detector Letter of Intent*, February 2010. DESY 2009–87. Available from http://goo.gl/FHQTp.

[Marlin] Marlin. Modular analysis and reconstruction for the linear collider [Webpage]. Available from http://ilcsoft.desy.de/portal/software_packages/marlin/.

[Meier 99] D. Meier. *CVD Diamond Sensors for Particle Detection and Tracking*. Ph.D. Thesis, CERN, Geneva, 1999.

[MINUIT] MINUIT. Function minimisation and error analysis [Webpage]. Available from http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html.

[Mokka] Mokka. A detailed geant4 simulation for the international linear collider detectors [Webpage]. Available from http://polzope.in2p3.fr:8081/MOKKA.

[Moortgat-Pick 08] G. Moortgat-Pick *et al.* *The role of polarized positrons and electrons in revealing fundamental interactions at the Linear Collider.* Phys. Rept., vol. 460, page 131, May 2008.

[MuonCollider] MuonCollider [Webpage]. Available from http://www.fnal.gov/pub/muon_collider/.

[MySQL] MySQL [Webpage]. Available from http://www.mysql.com/.

[Ohl 97] T. Ohl. *CIRCE version 1.0: Beam spectra for simulating linear collider physics.* Comp. Phys. Comm., vol. 101, page 269, 1997.

[PDG a] PDG. The ckm quark–mixing matrix [Webpage]. Available from http://pdg.lbl.gov/2011/reviews/rpp2011-rev-ckm-matrix.pdf.

[PDG b] PDG. Pdg summary tables [Webpage]. Available from http://pdg.lbl.gov/2011/tables/contents_tables.html.

[PDG c] PDG. The pmns neutrino–mixing matrix [Webpage]. Available from http://pdg.lbl.gov/2011/reviews/rpp2011-rev-neutrino-mixing.pdf.

[PYTHIA] PYTHIA [Webpage]. Available from http://home.thep.lu.se/~torbjorn/Pythia.html.

[RDR a] RDR. *ILC accelerators Reference Design Report.* Available from http://goo.gl/NYGll.

[RDR b] RDR. *Reference Design Report, Physics at the ILC.* Available from http://goo.gl/UuJHf.

[ROOT] ROOT [Webpage]. Available from http://root.cern.ch/drupal/.

[Sakharov 67] A.D. Sakharov. Pisma Zh. Eksp. Teor. Fiz., vol. 5, page 32, 1967.

[SB2009] SB2009. *SB 2009 proposal.* Available from http://ilc.kek.jp/SB2009/SB20091216A.pdf.

[School 09] Third Linear Collider Physics School [Webpage]. 2009. Available from http://www.hep.lancs.ac.uk/lc2009school/Programme.htm.

[Schulte 96] D. Schulte. *Study of Electromagnetic and Hadronic Background in the Interaction Region of the TESLA Collider.* Ph.D. Thesis, Universität Hamburg, 1996.

[Seryi 06] A. Seryi *et al. IR Optimization, DID and anti–DID*, 2006. SLAC–PUB–11662.

[SiD 09] SiD. *Silicon Detector Letter of Intent*, March 2009. Available from http://goo.gl/1BpBy.

[SILC] SILC. The silc collaboration [Webpage]. Available from http://lpnhe-lc.in2p3.fr/.

[Singer] A. Singer. *ABC of SUSY*. hep–ph/0905.4630.

[Spergel 03] D. N. Spergel *et al. First Year Wilkinson Microwave Anisotropy Probe Observations: Determination of Cosmological Parameters*. Astrophys. J. Suppl., vol. 148, page 175, 2003.

[Sugimoto 09] Y. Sugimoto. *FPCCD vertex detector R&D for ILC*, 2009. Available from http://www-jlc.kek.jp/subg/vtx/pub/FPCCD_Rev.pdf.

[Szeleniak 08] M. Szeleniak *et al. CMOS pixel vertex detector at STAR*. PoS (VERTEX 2008) 032, 2008.

[Thomson 09] M.A. Thomson. *Particle Flow Calorimetry and the PandoraPFA Algorithm*. Nuclear Physics B, July 2009.

[Traversi 08] G. Traversi *et al. Performance of a DNW CMOS active pixel sensor designed for the ILC Vertex Detector Development of deep N-well MAPS in a 130 nm CMOS*. IEEE NSS, page 1361, 2008.

[Turchetta 93] R. Turchetta. *Spatial resolution of silicon microstrip detectors*. NIM A, vol. 335, page 44, 1993.

[Turchetta 01] R. Turchetta. *A monolithic active pixel sensor for charged particle tracking and imaging using standard VLSI CMOS technology*. NIM A, vol. 458, page 677, 2001.

[Veltman 77] M. Veltman. *Second Threshold in Weak Interactions*. Acta Physica Polonica B, vol. 8, page 475, 1977.

[Vogel 08] A. Vogel. *Beam–Induced Backgrounds in Detectors at the ILC*. Ph.D. Thesis, Departments Physik der Universität Hamburg, 2008. Available from http://goo.gl/PKCVS.

[Walsh 09] R. Walsh. *Flavour tagging optimisation studies at ILD*, 2009. Available from http://goo.gl/hEZid.

[Weiglein 06] G. Weiglein *et al. Physics Interplay of the LHC and the ILC*. Phys. Rept., vol. 426, page 47, 2006.

[WHIZARD] WHIZARD. The whizard event generator [Webpage]. Available from http://projects.hepforge.org/whizard/.

[Wienemann] P. Wienemann. Physics at the ilc [Webpage]. Available from http://goo.gl/VzuPl.

[Winter 09] M. Winter. *Development of Swift, High Resolution, Pixel Sensor Systems for a High Precision Vertex Detector suited to the ILC Running Conditions*. DESY PRC R&D Nr 01/04, November 2009.

[Zimmerman] F. Zimmerman [Webpage]. Available from http://goo.gl/vpBZb.

Acknowledgements

It has been a long trip, and I owe many thanks to many people. First, I would like to express my gratitude to my supervisor Dr. Marc Winter, for being a mentor and for giving me the opportunity to work with PICSEL group. Also, I would like in particular to thank Dr. Rita De Masi for her inspiring guidance and the help she offered me during my first steps in Strasbourg.

This work has been conducted inside the PICSEL group of IPHC Strasbourg, and I would like to thank all the members of this group. They have created an excellent working environment, and provided me with invaluable support. Especially I want to thank Dr. Jerome Baudot and Dr. August Besson for the fruitfull discussions we had and for sharing their knowledge in CAPS with me. But also for their guidance through the mysteries of the MAF/TAF software. Mathieu Goffe has been a colleauge and friend for me, and the experience I gained working in test beams with him was invaluable.

I would like to thank Dr. Steve Aplin from DESY, for his help to understand the philisophy of track reconstruction and to properly use the ILC software.

The almost four years I spent in Strasbourg have been a wonderful time thanks to the friends I met there. Christianna who grew up 100m away from my place, and I had to go to Strasbourg to discover her sitting in the office across the corridor. I am grateful to the Gods of luck for that. In Leonidas and Mario I found two true friends, and a sense of nostalgy is flooding me when I remember our adventures in Strasbourg, especially our evenings in Taverne Francais or even having fun in CNRS cantine. Michael, with whom I had the pleasure to have "heated" discussions on a broad spectrum of subjects, giving me a different perspective towards them, and finally always cool down over a glass of cold blanche. Artemis (who by the way unraveled many mysteries of Latex and SVN for me), Nikos, I learned many things discussing together. Sofia and Katerina, showed me Strasbourg as a party-city. I met Daria few months before I finished my PhD; she transformed these months, which usually are the most difficult in the life of a PhD student, to the happiest ones. All these people made me feel Strasbourg being

my home, almost as much as Pireas is.

Finally, to my parents and sister, and my childhood's friends: they are a lighthouse for me.