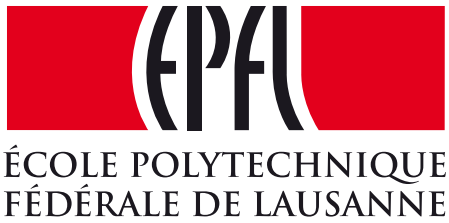


ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

ELECTRICAL AND ELECTRONIC ENGINEERING SECTION

MASTER THESIS

Performance-Optimization Studies for the CLIC Vertex Detector

Author:
Niloufar ALIPOUR TEHRANI

Supervisors:
Prof. Dr. Pascal FROSSARD (EPFL)
Dr. Philipp ROLOFF (CERN)

This project has been carried out at CERN.

August 29, 2013

Abstract

The Compact Linear Collider (CLIC) is a mutli-TeV linear e^+e^- collider currently under development at CERN. In the post-LHC era, CLIC will allow to explore a great number of searches for New Physics such as the precise measurements of the Higgs boson. In this master thesis, we mainly focus on the development and the improvement of the vertex detector. The vertex detector requires excellent spatial resolution, low mass, geometrical coverage down to low polar angles, high rate readout for the sensors and new cooling technologies for heat removal. Considering such requirements, the CLIC vertex detector technology is far more advanced in comparison to the technologies currently used in particle physics.
This project consists of two main parts. In the first part, we study the vertex detector and optimize its geometry for the use of airflow cooling techniques and also for flavor tagging. In the second part, we implement a decoder which can respect the timing constraints for the CLICpix chip, a silicon pixel detector developed for the CLIC vertex detector.

Contents

1 Introduction

Elementary particle physics searches to answer the fundamental question on what matter is made of. This quest was initiated a long time ago and still has a long way to go. Many theories have been put in place to explain the nature of matter and various kinds of the elementary particles. The Standard Model (SM) is the most successful theory in modern physics which predicts many phenomena like the existence of the heavier quarks and the massive gauge bosons of the electroweak interaction [24]. Many of the SM predictions have also been proven experimentally. The SM also predicts the existence of the Higgs boson, which was experimentally observed in 2012 by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) at CERN. Even though the Higgs has been found, its true nature still remains unclear and its couplings to the other particles in the SM have to be observed in order to fully validate this model.

The LHC, performs collisions of protons in a circular trajectory. In the high-energy physics (HEP) community, it is agreed that the next collider which can complement the LHC experiments would be a lepton (e^+e^-) collider. Indeed, a lepton collider would allow more precise measurements as the electron and the positron are elementary particles and their collisions generate less backgrounds than proton colliders.

Two concepts of future lepton colliders being considered are the International Linear Collider (ILC) [15] with a center-of-mass energy up to 500 GeV and the Compact Linear Collider (CLIC) [29] with a center-of-mass energy up to 3 TeV. Two detector concepts have been developed for linear colliders: the International Large Detector (ILD) [9] and the Silicon Detector (SiD) [11]. In this report, we mainly focus on the SiD concept for CLIC called CLIC_SiD.

This master project focuses on the vertex detector designed for the CLIC_SiD. The latter is the innermost detector in CLIC. The vertex detector is part of the so-called tracking system. It contributes to the measurement of the curved tracks of high-energetic charged particles in the magnetic field of the experiment. One of its primary goals is to reconstruct the vertex of heavy quarks (beauty and charm) and to distinguish them from lighter quarks (up, down and strange). The vertex detector is designed for the tagging of tau-leptons. The sensors in the vertex detector are made of silicon pixel detectors and need to have very small pixels in order to increase the spatial resolution. As the vertex detector is very close to the interaction point, it

is highly exposed to beam-induced backgrounds which lead to high pixel occupancies. Precise time-tagging capabilities are needed to separate the physics hits from the background hits. Heat removal from the sensors is also a big challenge. The material budget constraints suggest the use of air cooling in the detector. During this project, we study new geometries for the vertex detector in order to make it compatible with the use of airflow cooling techniques. Also the use of more detection layers will help to have precise track reconstruction and to reduce the material budget. For this reason, we study the effect of using double-layered silicon sensors on the particle recognition.
As a second step, we study a decoding algorithm for the CLICpix chip, a silicon pixel detector prototype developed for the CLIC vertex detector which uses 65 nm CMOS technology.
First of all, we give an overview on CLIC and its sub-detectors in Chapter 2. Secondly, we study the SM and the physics motivation for CLIC and introduce the silicon pixel detectors used in the tracking system. In Chapter 3, the flavor tagging technique is introduced which is used to identify heavy quarks.
Chapter 4 describes the software packages used for the implementation of the geometry of the CLIC detector. In addition, the packages used for the track reconstruction and the flavor tagging are explained. Chapters 5 and 6 describe the developments made during this project. First the implementation of the new vertex detector geometries and then the study of their flavor-tag performance for quark identification are discussed. Chapter 7 presents the decoder algorithms implemented for the CLICpix chip prototype and finally Chapter 8 recapitulates the conclusions of the studies and gives an outlook on possible future improvements.

2 The Compact Linear Collider

The Compact Linear Collider (CLIC) is a future linear particle collider for electrons (e^-) and positrons (e^+) which is being developed at CERN [29]. As opposed to the Large Hadron Collider (LHC) where the particles (protons) are accelerated in a circular accelerator, CLIC uses a linear beamline. When the trajectory of a high-energy charged particle is deviated radially in a magnetic field, it radiates and loses energy. This effect is called *synchrotron radiation*. At a given energy it affects electrons much more than protons. This is the reason why a linear collider is used for electron beams.
Today, the LHC is the largest particle accelerator being able to collide two opposing particle beams of protons with a center-of-mass energy up to 8 TeV. So far, the main result of the LHC is the observation of the Higgs boson and the determination of its mass in 2012. However, the experiments at the LHC can not fully answer the questions on the nature of this particle.
In the post-LHC era, CLIC will allow to determine the properties of the Higgs boson with a very high precision. CLIC can measure collisions with center-of-mass energies from a few hundred GeV up to 3 TeV.
In this chapter, we will briefly discuss the standard model of particle physics and attempt to understand how CLIC can determine more precisely the properties of the Higgs boson. Then the CLIC detector and its components are described. We also give an introduction to silicon detectors which are used in the CLIC tracking system.

2.1 The Standard Model

In particle physics, the Standard Model (SM) is a theoretical framework that describes how the interaction between elementary particles is governed by four fundamental forces. This theory, developed in the early 1970s, explains most of the experimental results.
According to the Standard Model, matter is made of elementary particles. They can be regrouped into three basic kinds: *quarks, leptons* and *gauge bosons*.
There are six leptons and six quarks classified into *generations* based on different properties and described in Tables 2.1 and 2.2.

	l	Q	L_e	L_μ	L_τ
First generation ⇒	e	-1	1	0	0
	ν_e	0	1	0	0
Second generation ⇒	μ	-1	0	1	0
	ν_μ	0	0	1	0
Third generation ⇒	τ	-1	0	0	1
	ν_τ	0	0	0	1

Table 2.1 – Lepton classification. From [24].

Leptons, are arranged in three generations according to their charge (Q), electron number (L_e), muon number (L_μ) and tau number (L_τ) [1]: the *electron* and the *electron neutrino*, the *muon* and the *muon neutrino* and finally the *tau* and the *tau neutrino* corresponding respectively to the first, second and third generation. The electron, the muon and the tau have an electric charge and a sizable mass but the neutrinos are electrically neutral and have a very small mass. There are six antileptons with the reversed signs for the lepton properties. In total, there are 12 leptons.

	q	Q	D	U	S	C	B	T
First generation ⇒	d	-1/3	-1	0	0	0	0	0
	u	2/3	0	1	0	0	0	0
Second generation ⇒	s	-1/3	0	0	-1	0	0	0
	c	2/3	0	0	0	1	0	0
Third generation ⇒	b	-1/3	0	0	0	0	-1	0
	t	2/3	0	0	0	0	0	1

Table 2.2 – Quark classification. From [24].

Similarly to the leptons, there are six types of quarks also called *flavors* of quarks (Table 2.2). They are classified by their charge (Q), strangeness (S), charm (C), beauty(B), top (T), upness (U) and downness (D)[2].
The lighter particles belong to the first generation and the heavier belong to the second and third generations. The six quarks are: *up* and *down* (the first generation), *charm* and *strange* (the second generation) and finally *top* and *bottom* (or beauty). The quarks mix together to form, for example, protons (uud) and neutrons (udd). Protons and neutrons, together form atoms and molecules. As said before, the Standard Model, describes not only the elementary particles but also the interactions between them. There are four fundamental forces responsible for these interactions: the *strong* force, the *weak* force, the *electromagnetic* force and the *gravitational* force. The differences between them are their effective ranges

[1] L_e, L_μ and L_τ are leptonic family numbers and assigned to +1 for particles of corresponding family, -1 for antiparticles and 0 for leptons of other families.

[2] S, C, B, T, U and D are quantum numbers like the leptonic numbers and are conserved in a quantum system.

and strengths. The electromagnetic and the gravitational forces have infinite ranges but the electromagnetic force is stronger than gravity. For example, the electromagnetic force ties the electron to the atomic nucleus to form atoms. The weak and the strong forces are dominant only at the subatomic level. The weak force is stronger than gravity but it is the weakest among the three other interactions and is responsible for the radioactive decay of heavy particles into lighter ones. The strong force is the strongest of all forces as its name suggests and is responsible for gathering the quarks to form hadrons such as protons and neutrons. Hadrons are indeed made of quarks held together by the strong force. The theory of the strong interaction is called *Quantum Chromodynamics* (QCD).
Every interaction has its *mediator*. The *gauge bosons* are the force-carrier particles responsible for exchanging three fundamental forces. The *gluon* responsible for exchanging the strong force and bending the quarks together, the *photon* for the electromagnetic force and the *W and Z bosons* for the weak force. The gravitational force is believed to be exchanged by the force-carrying particle *graviton* but the theory of gravity at the microscopic scale has not yet been established.
The fundamental constituents of matter identified are shown in Figure 2.1.

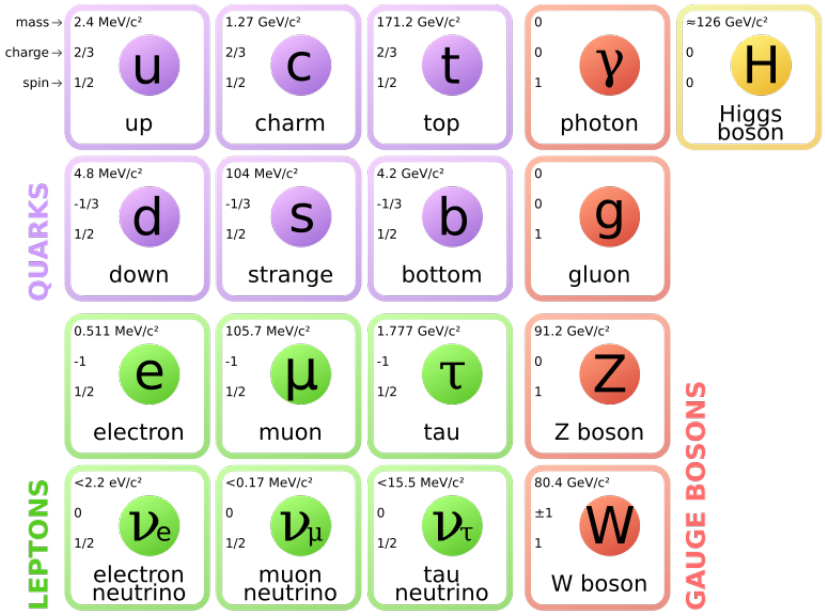


Figure 2.1 – The building blocks of matter. From [7].

Today, the Standard Model is the best theory describing the subatomic world. However, it does not answer questions like the nature of dark matter. This theory also predicts the existence of the Higgs boson which gives the mass to all particles. It was experimentally observed in 2012 by the LHC experiments.
The weak and the electromagnetic forces are closely related to each other and can be unified as the *electroweak* force and the equations describing the unification predict the force-carrying particles (the photon, the W and Z bosons). The only problem is that all the force-carrying

particles are described as being massless which is true for the photon, but the W and Z bosons have a mass about 100 times larger than that of the proton. To solve this problem, the Brout-Englert-Higgs mechanism was introduced which suggests that the Higgs boson gives the mass to the W and Z boson by interaction with a *Higgs field*.
The Higgs particle can be produced in a particle collider by accelerating particles to high energies and speed (close to the speed of light) and colliding them together. Heavy particles, like the Higgs boson, are occasionally produced and then detected by a particle detector. The Standard Model predicts different mechanisms to produce the Higgs boson and the probability to produce it is very small. For example, in LHC only 1 Higgs boson is produced per 10 billion collisions.
CLIC studies different mechanisms to produce the Higgs boson as shown in Figure 2.2.

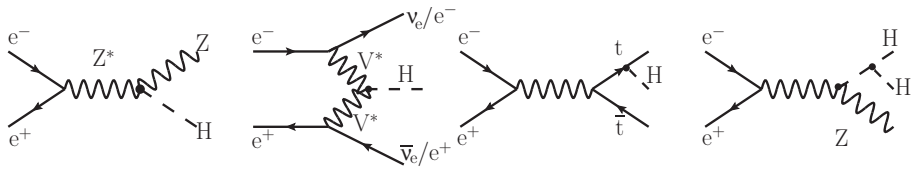


Figure 2.2 – Standard Model Higgs boson production mechanisms at CLIC. From [29].

The cross sections (the probability for a specific process to occur in a given interaction) to produce a Higgs with a mass of $M_H = 125$ GeV as a function of the center-of-mass energy[3] $\sqrt{s}$ is given in Figure 2.3. For lower $\sqrt{s}$, the HZ mechanism is dominant. For higher energies, the $H\nu_e\bar{\nu}_e$ mechanism gets dominant.

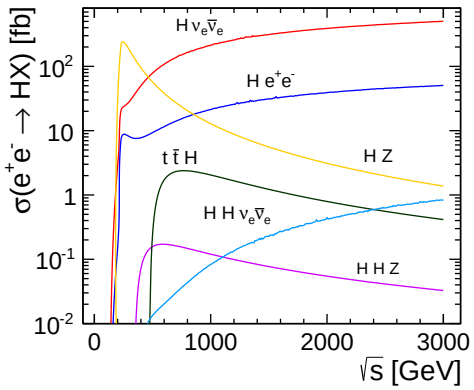


Figure 2.3 – Cross sections for different production mechanisms for a $M_H = 125$ GeV Higgs boson as a function of the e^+e^- center-of-mass energy. From [27].

Like many other particle, the Higgs boson decays quickly into a set of lighter particles. It can

[3] The center-of-mass energy is given by $\sqrt{s} = \sqrt{(\sum_i E_i{}^2 - \sum_i p_i{}^2)}$.

decay through many different processes and each has its own probability. As its lifetime is very small, it can not be detected directly by the detectors. It can be recognized by the reconstruction of its decay products in the detector. Each part of a particle detector is optimized for detecting specific particle properties.
An electron-positron collider allows to perform precision measurements because the colliding beams are made of elementary particles. In fact, with elementary particles, the center-of-mass energy and the polarization of the colliding particles can be selected precisely. And unlike proton-proton collisions (used in the LHC experiments), there is no underlying event from proton remnants as shown in Figure 2.4. This is the reason why CLIC can do more precise measurements on the Higgs bosons and provide complementary information on the LHC results.

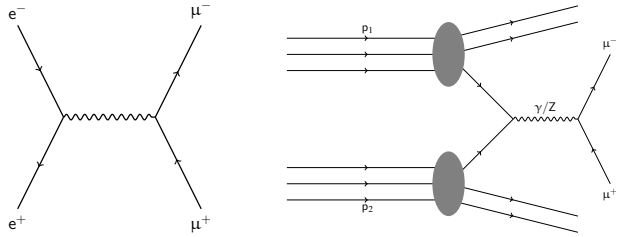


Figure 2.4 – Schematic view of electron-positron collisions at CLIC (left) and proton collisions at the LHC (right). From [32].

In the following Section, we describe different components of the CLIC detector.

2.2 The CLIC_SiD Detector Concept

Figure 2.5 shows the layout of the CLIC accelerator for a center-of-mass energy of 3 TeV. The electron and positron beams are accelerated on a linear trajectory and collide in the central region of the machine (in the interaction point), where the CLIC detector is placed.

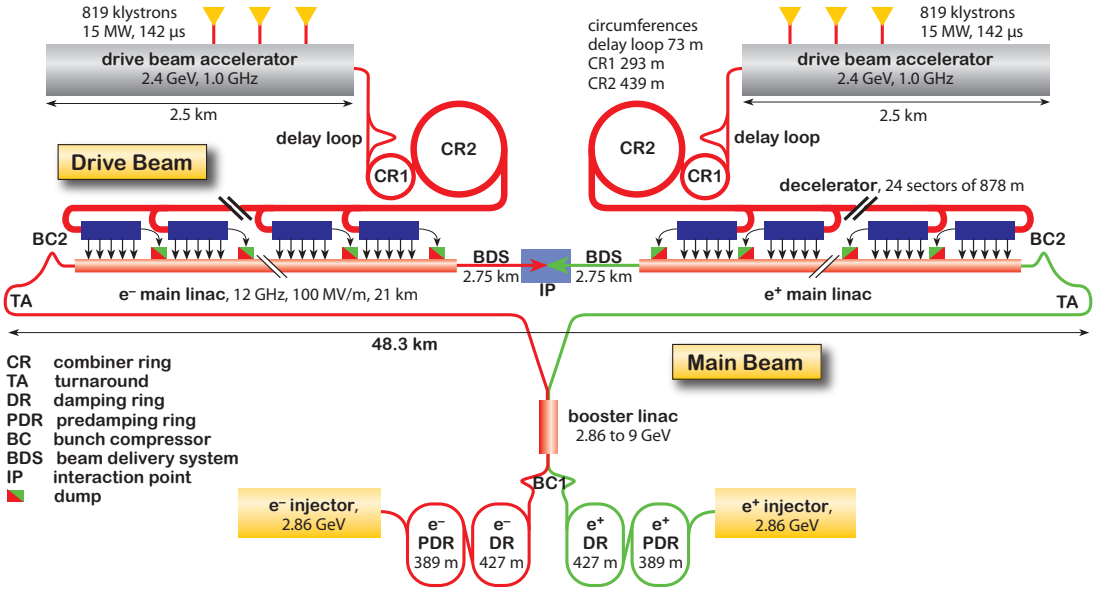


Figure 2.5 – The schematic view of the CLIC layout for a center-of-mass energy of 3 TeV. From [20].

The site studies have shown that CLIC could be placed near CERN underground as shown in Figure 2.6. It shows how CLIC could be built in stages. For each energy increase more accelerating modules will be needed, making the accelerator longer. The site length for 3 TeV will be 48 km.

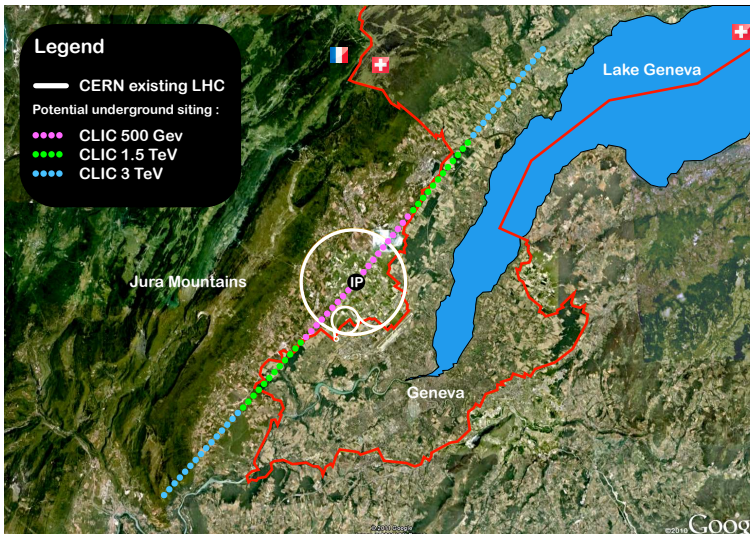


Figure 2.6 – CLIC footprint near CERN, for different implementation stages. From [20].

The CLIC detector model studied in this thesis is called CLIC_SiD which is based on the SiD detector concept [11] developed for the International Linear Collider (ILC) [15]. ILC is another linear e^+e^- collider under development. It is based on a different acceleration technique. ILC can reach 500 GeV center-of-mass energy for a site length of 32 km, while CLIC can reach 3 TeV for a site length of 48 km. The ILC may be implemented in Japan.
The SiD detector is considered as a compact detector with an overall length of 12.39 m and a total height of 12.50 m. It is composed of several sub detectors. Its main components are shown in Figures 2.7 and 2.8 and briefly described in the following sections.

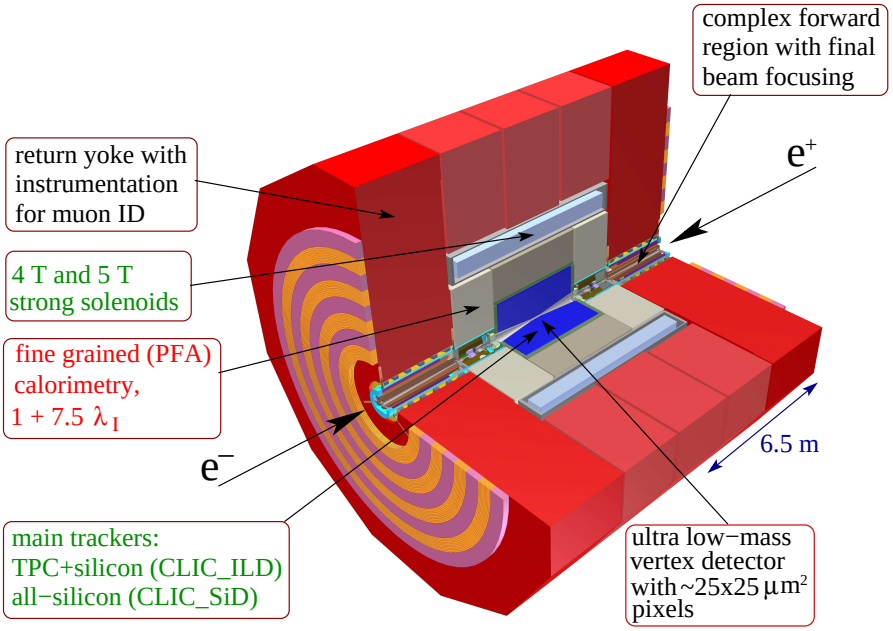


Figure 2.7 – The CLIC detector model. The beams come from the two sides of the detector and collide at its center in the Interaction Point (IP). From [27].

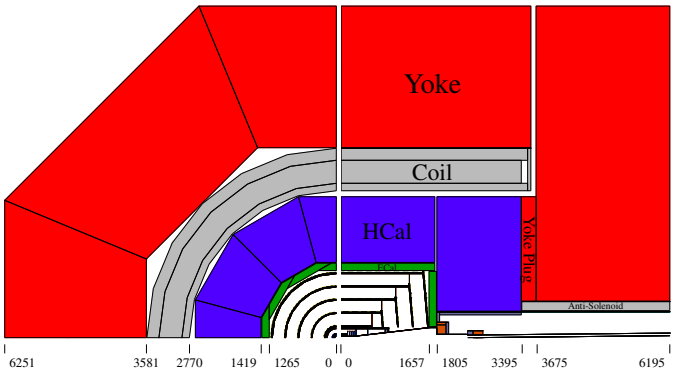


Figure 2.8 – One quadrant of the CLIC_SiD detector model in the xy-plane (left) and in the zx-plane (right). Distances are given in millimeters. From [23].

2.2.1 Tracking Detectors

The tracking system consists of silicon detectors with excellent point resolution and low material budget placed in the innermost part of the detector and very close to the interaction point. It covers a wide range of polar angles and the lowest angle measured is for $\theta = 8°$. The tracking system is made of two main components: the vertex detector and the main tracking system which are detailed in the sections below.
The tracker measures the position of the charged particles in several layers and allows for the precise measurement of particle trajectories. The measurement of the curvature of the track yields very precise information on the particle's momentum (cf. Equation 5.1).
The vertex detector uses silicon pixel detectors while the main tracking system uses silicon microstrip detectors. In the region close to the interaction point, the density of the particles is very high. To measure the position of the charged particles with a high precision, the pixels provide a high spatial resolution. But they need a complex readout system and they are very expensive. For these reasons, they are used only at the inner part of the detector.
The main tracking system consists of silicon microstrip detectors. The strips are placed along a specific direction. In order to have the coordinates of a particle hit, two strips with two different directions are used. The overlap between the strips provides the position of the hit in the detector. But the strips can not provide precise measurements when the density of the particles is high because many strips can be activated which corresponds to many overlaps between the strips. For this reason, they are placed in the regions further away from the interaction point. They need a simpler readout system compared to the pixel detectors and they are more cost-effective.
Also particles with higher momentum reach the region covered by the tracking system and their radius of curvature is lower in the magnetic field and they have a more linear trajectory. Silicon strips give enough precision on their track reconstruction. The pixel detectors improve the precision of the track reconstruction for low-momentum particles as they bend more in the magnetic field.

Vertex Barrel and Endcap

The vertex barrel is composed of five concentric layers each made of several modules. The vertex endcaps cover the forward region close to the interaction point of the detector and is composed of four disks each made of several trapezoidal modules. The vertex barrel and endcaps are described in detail in Section 5.2. Each module contains a silicon sensor with a thickness of 50 μm and with a pixel size of 25 µm×25 µm. The single-point accuracy obtained is 3 µm. The silicon sensor is the sensitive part of each module. In Section 2.3, an introduction on silicon sensors is given. Figure 2.9 illustrates the vertex barrel and endcaps for the CLIC_SiD as defined in [29].

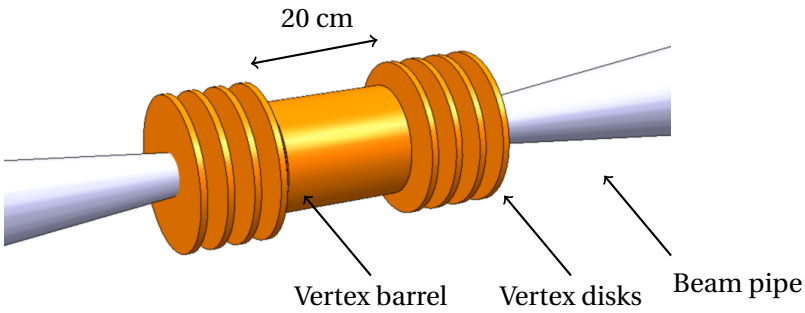


Figure 2.9 – Schematic picture of the vertex barrel and endcaps. For CLIC_SiD, the barrel contains five layers and the endcaps are made of four disks of silicon pixel detectors.

Tracker Barrel and Tracker Endcap

The tracker barrel surrounds the vertex detector. It contains five layers of silicon microstrip detectors. The tracker endcap covers the forward region with four concentric disks (after the vertex endcaps) as shown in Figure 2.10.
The silicon microstrip detectors have a typical strip length of 10 cm and a strip pitch of 50 µm. The typical measurement accuracy for single hits is approximately 7 µm in the bending plane.

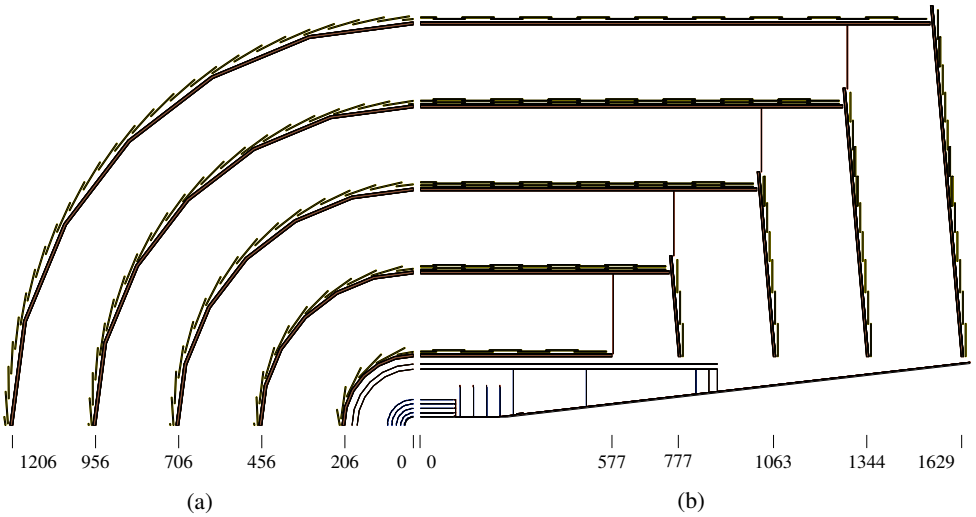


Figure 2.10 – The layout of the tracking system in the xy-plane (a) and in the zx-plane (b). The tracker modules are shown in yellow and the support structures in brown. All distances are given in millimeters. From [23].

2.2.2 Calorimeters

Calorimeters are placed after the tracking systems. Two types of them are used: the silicon-tungsten electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The calorimeters are used to measure the energy of the particles. The ECAL is designed to measure the energy of particles interacting via the electromagnetic interactions and the HCAL is aimed at measuring particles interacting via the strong nuclear force.

2.2.3 Superconducting Coil

The superconducting coil provides a solenoidal magnetic field of 5 T. It is placed outside of the calorimeters. The magnetic field deflects the trajectory of charged particles. Tracking detectors use the radius of curvature to measure the momenta of charged particles (cf. Section 5.3.2).

2.2.4 Iron Yoke

The iron yoke surrounds the whole detector. It forms an integral part of the magnet system and contains detection layers to identify the muons.

2.3 Introduction to Silicon Detectors for the Tracking System

The tracking detectors are made of silicon detectors. A silicon atom has four electrons in the valence band. In the fundamental state of a semi-conductor, the conduction band is empty and behaves like an insulator. When the energy of the electrons raises by thermal excitation or energy absorption, they can jump from the valence band to the conduction band. The *band gap* between the valence and the conduction band is 3.62 eV (at 300 K) [25]. When an electron moves from the valence to the conduction band, it leaves behind a hole representing a positive charge. This hole can be filled by an electron from a neighboring atom. The electric conduction is then generated by the electron-hole displacement.
By introducing impurities, like doping the silicon with atoms having five valence electrons, an n-type silicon is obtained. Indeed, the fifth electron from the impurity can jump easily in the conduction band. A p-type silicon is obtained by doping with silicon atoms having three valence electrons. By creating a junction between a p-type and an n-type silicon, the electrons from the n-type silicon move to the p-type side and holes move in the opposite direction. This charge movement continues until a balance is created and no charge moves. A *depletion* zone is created which stops the conduction between the two areas. By applying a voltage between the two sides of the junction where the n side is negative compared to the p side, the junction conducts the current. The depletion zone can be increased by applying a positive voltage to the n side compared to the p side. This configuration is shown in Figure 2.11.
Indeed, in such a system, when a high-energy charged particle goes through the depletion

zone, it produces electron-hole pairs by losing its energy. The Bethe-Bloch equation [13] gives the rate of ionizing loss of a charged particle in matter. For high-energy particles (at energies above several hundred MeV) having velocities approaching the speed of light, the energy loss approaches a constant minimum value. *Minimum ionizing particles* (MIP) refers to relativistic particles with high velocities and they have the same energy loss behavior in matter. But for a non relativistic charged particles ($v \ll c$), the energy loss varies inversely with particle energy as they spend greater time in matter.
In the existing tracking systems, silicon sensors with a thickness of 300 µm are used which provide around 20000 electron-hole pairs. CLIC requires very low material in the tracking system. Based on the studies done in [29], the silicon sensor has to have a thickness of 100 µm which provides around 7000 electron-hole pairs. In this case, the readout system has to be very precise with a high noise rejection.
As a voltage is applied to the both sided of the p-n junction, the created charge moves out of the depleted zone and a detectable electrical current is generated. Then the generated current is measured using readout pads. In the vertex barrel and endcaps areas close to the beam pipe, small pixels are used as readout pads. For a depleted detector, the speed of charge collection is a few ns which allows the use of a high rate readout system.

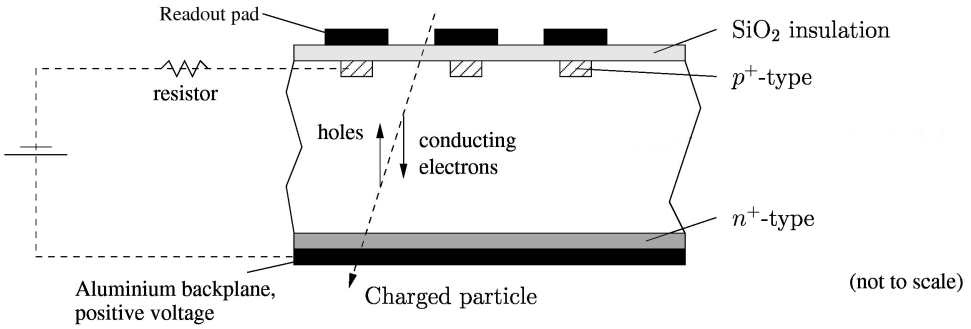


Figure 2.11 – Schematic view of a silicon sensor. The applied voltage, creates a depletion zone between the n-type and the p-type regions. When a charged particle passes the zone, it deposits some of its energy in the silicon and releases electrons from the silicon atoms and creates a detectable current. From [36].

3 Flavor Tagging for CLIC

Heavy quarks produced in high-energy e^+e^- collisions at CLIC play an important role in many physics processes. They can be distinguished from light-flavored quarks through a measurement of their displaced decay vertex in the detector, using high-precision vertex detectors and sophisticated pattern recognition software.
Quarks transform to jets of particles in the so-called hadronization process, governed by non-pertubative QCD, as shown schematically in Figure 3.1. The hadronization process takes place instantaneously after the primary interaction. As quarks can not live alone, they will associate with other quarks to form a hadron. The u, d and s quarks, hadronize at the interaction point and produce stable hadrons.
B hadron is produced by the association of b quark with a lighter quark (c, u, d or s). On average, it gets 90% of the momentum of the jet. C hadron is produced by the association of c quark with u, d or s quark. On average, it gets 50% of the momentum of the jet and has a shorter lifetime than the B hadron. Equation 3.1 defines the decay length l of a particle. It corresponds to the distance a particle travels before it decays to a lighter particle.

$$l = c\tau\beta\gamma \tag{3.1}$$

where c is the speed of light in vacuum, τ the particle lifetime, $\beta = \frac{v}{c}$ with v the velocity of the particle and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. For B hadrons, $c\tau$ is around 500 µm. For C hadrons, $c\tau$ is between 60 and 300 µm.
For example, for a B hadron at 250 GeV, l is around 25 mm. After this distance, the B hadron will decay to a lighter hadron.
The vertex detector can detect the B and C hadrons and reconstruct their displaced decay vertex.
The top quark plays a special role as it decays through the weak interaction into a W boson and a b quark, before the hadronization can occur.

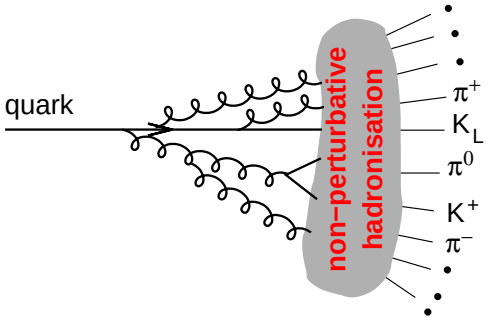


Figure 3.1 – Schematic view of a jet produced by the hadronization of a quark. From [33].

The Higgs boson has a very short lifetime. By studying its decay products in the detectors, one can probe the predictions of the SM and search for effects of new physics. An example for a physics process involving heavy quarks is the production of Higgs bosons with subsequent decay to pairs of b quarks in the process:
$e^+e^- \rightarrow H\nu\bar{\nu} \rightarrow b\bar{b}\nu\bar{\nu}$.
In this process, lots of background processes produce two jets other than b quarks.
Flavor tagging is the process of identifying the flavor of the quark a jet originates from. The identification of beauty quarks is called *b-tagging*. The beauty quark is heavier than c, u, d and s quarks and is the most easy to identify. The flavor tagging is only possible for bottom and charm quarks. For the lighter quarks, individual tagging is currently not considered for the CLIC detector, as the jets from u, d and s quarks have too similar properties.
The flavor tagging algoritms used for the study presented in this thesis is described here-below. It is based on the LCFIPlus software package [34](derived from LCFIVertex [19]) which is described later in Section 4.6.

3.1 Vertex Finding and Jet Classifying

The flavor tagging algorithms first tries to find the location of the *Interaction Point* (IP) which corresponds to the position of the collision of the particle beams in the detector. The reconstructed position of the IP is called the *primary vertex*. This is done by reconstructing the tracks passing the vertex detector which is the closest sub-detector to the IP. For CLIC, the beam bunch dimensions are very small (σ_x =45 nm, σ_y =1 nm and σ_z =44 µm). Therefore the IP position is used as a constraint to find the primary vertex.
After finding the primary vertex, an algorithm is used to look for secondary vertices. The number of secondary vertices found in a jet is a useful criterion to distinguish between b, c and uds jets. Based on the number of secondary vertices, the jets are grouped into four different classes as illustrated in Figure 3.2.
In the first class (Figure 3.2a), no secondary vertices are found. This class, most likely corresponds to the light quarks (u, d and s jets) which are very light and hadronize at the IP and their produced hadrons are stable.

In the second class (Figure 3.2b), in addition to the primary vertex, one secondary vertex is found. Charm jets are most likely reconstructed in this class.
For the third class (Figure 3.2c), in addition to the primary vertex, one secondary vertex and one track with high impact parameter[1] (called pseudovertex) is found. In this case, it is assumed that there are two secondary vertices but only one of them is reconstructed. It corresponds most likely to B hadrons which are heavy with longer lifetime and they decay to C hadrons and then lighter hadrons. This is the reason why two secondary vertices are produced.
Finally, in the fourth class (Figure 3.2d), two secondary vertices are reconstructed. Most jets in this class originate from b quarks.

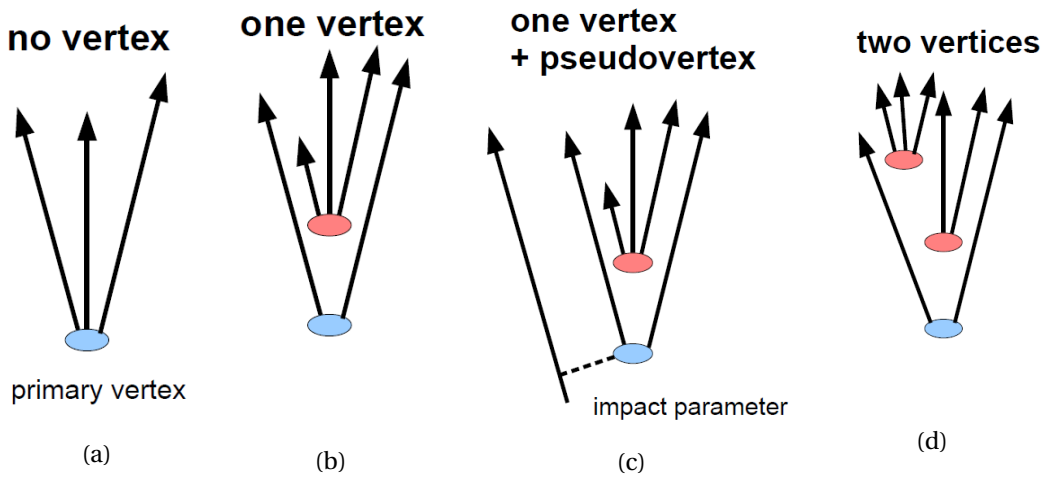


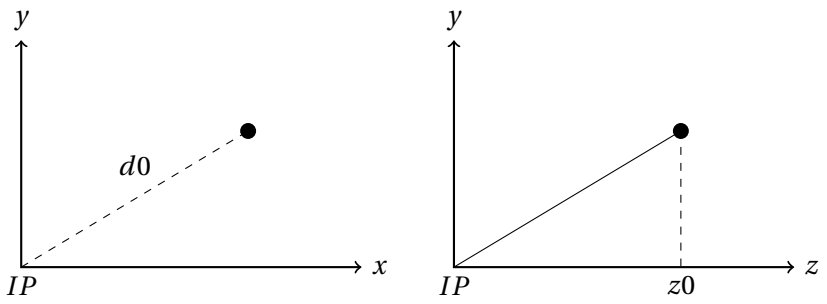
Figure 3.2 – Schematic of different jet classes, based on the number of secondary vertices. Figure from [31].

3.2 Flavor Tagging

As explained previously, flavor tagging consists of identifying the type of the quark a jet originates from. We want to be able to distinguish between b and c and light-flavor quarks.
In high energy physics (HEP) the events of the investigated process are called signal events (in our case, they contain jets from b or c quarks) and need to be differentiated from background events (events containing jets from light flavor quarks, b or c quarks depending on the definition of signal events).
In section 3.1, the classification of the jets into different classes is explained. The identification of the flavor of a quark inducing the jets follows the identification of the jet class. Besides the number of secondary vertices, other discriminating *input variables* enter a multivariate discrimination.
For the study reported in this document the LCFIPlus flavor tagging software was used, see

[1] The impact parameters are explained in Section 3.2.

also Section 4.6. A list of the LCFIPlus input variables is given in [34]. For each jet class (Figure 3.2), an optimized set of variables is used (see Table 3.1).
The impact parameters d_0 and z_0 are important parameters used for flavor tagging. The impact parameter of a track is the distance between the tracks's point of closest approach to the IP [19]. The impact parameter d_0 gives this distance in the xy-plane and z_0 to the z-axis. The *significance* of an impact parameter is the impact parameter divided by its uncertainty. Figure 3.3 illustrates the impact parameters d_0 and z_0.



(a) d_0 (b) z_0

Figure 3.3 – Impact parameters d_0 and z_0. The black dot corresponds to the closest point of approach of a track to the IP projected in the xy-plane (a) and zy-plane (b).

A brief explanation for each flavor tag input variable is give below (from [34]):

- trk1d0sig/trk2d0sig: d0 significance of track with highest/second highest d0 significance.
- trk1z0sig/trk2z0sig: z0 significance of track with highest/second highest d0 significance (ordering by d0, not z0).
- trk1pt/trk2pt: Transverse momentum of track with highest/second highest d0 significance.
- trk1pt_jete/trk2pt_jete: trk1pt/trk2pt divided by the jet energy.
- jprobr/jprobz: Joint probability in the r-phi plane/z projection using all tracks.
- vtxlen1_jete: Decay length of the first vertex in the jet (zero if no vertex is found) divided by the jet energy.
- vtxsig1_jete: Decay length significance of the first vertex in the jet (zero if no vertex is found) divided by the jet energy.
- vtxdirang1_jete: The angle between the momentum (computed as a vector sum of track momenta) and the displacement of the first/second vertex multiplied by the jet energy.

- vtxmom1_jete/ vtxmom2_jete: Number of tracks included in the first vertex (zero if no vertex is found) divided by the jet energy.
- vtxmass1/vtxmass2: Mass of the first/second vertex computed from the sum of track four-momenta.
- vtxmult1/vtxmult2: Number of tracks included in the first/second vertex (zero if no vertex is found/number of vertex is less than two).
- vtxmasspc: Mass of the vertex with minimum p_T correction allowed by the error matrices of the primary and secondary vertices.
- vtxprob: Vertex probability. For multiple vertices, the probability P is computed as $1-P=(1-P_1)(1-P_2)...(1-P_N)$.
- vtxlen2_jete: Decay length of the second vertex in the jet (zero if number of vertex is less than two) divided by the jet energy.
- vtxsig2_jete: Decay length significance of the second vertex in the jet (zero if number of vertex is less than two) divided by the jet energy.
- vtxdirang2_jete: The angle between the momentum (computed as a vector sum of track momenta) and the displacement of the second vertex multiplied by the jet energy.
- vtxlen12_jete: Distance between the first and second vertex (zero if number of vertex is less than two) divided by the jet energy.
- vtxsig12_jete: vtxlen12 divided by its error as computed from the sum of the covariance matrix of the first and second vertices, projected along the line connecting the two vertices divided by the jet energy.
- vtxdirang12_jete: The angle between the two vectors as defined as follows. The first vector is the displacement vector from vertex 1 to vertex 2. The second vector is the difference of the vertex momentum 1 and vertex momentum 2. The computed angle is then normalized by the jet energy.
- vtxmom_jete: The vertex momentum normalized by the jet energy, where the vertex momentum is computed by simply summing the momentum of tracks used to create vertices. If there are multiple vertices they are simply added.
- vtxmass: Vertex mass as computed from the sum of four momenta of all tracks forming secondary vertices.
- vtxmult: Number of tracks which are used to form secondary vertices (summed for all vertices).
- 1vtxprob: Vertex probability with all tracks associated in vertices combined.

- jprobr5sigma/jprobz5sigma: Joint probability in the r-ϕ plane/z projection using all tracks having impact parameter significance exceeding 5 sigma.
- d0bprob/d0cprob/d0qprob: Product of b/c/q-quark probabilities of d_0 values for all tracks, using b, c, q d_0 distributions.
- z0bprob/z0cprob/z0qprob: Product of b/c/q-quark probabilities of z_0 values for all tracks, using b, c, q z_0 distributions.
- trkmass: Mass of all tracks exceeding 5 sigma significance in d_0 / z_0 values.

3.3 Boosted Decision Trees for Multivariate Classification

Using multiple input variables for event classification is called *Multivariate classification*. In order to establish decision boundaries to discriminate the signal and the background, machine learning methods are used.
Boosted decision trees (BDT) are widely used for multivariate classification. The boosting algorithm is a very powerful learning technique which combines many classifiers (which can have a weak performance) to obtain a more powerful classifier [37]. BDTs are fast and robust with respect to the correlations between the input variables. They have a high performance and also for small training sample sizes they do not require extensive parameter tuning.
In this classifier, the selection of the signal and background events is based on the result of several decision trees. An example of a simple decision tree is given in Figure 3.4, where S and B mean respectively signal and background.
A decision tree contains successive decision nodes starting from the *root node* and ending into *leave nodes* (in Figure 3.4, they are shown in boxes). In the leaves the events are categorized as either signal or background. Each node uses only one discriminating variable to arrive at a conclusion if the event is signal-like or background-like.

The *cut criteria* between signal and background are defined during the training process using samples of simulated signal and background training events. The training starts with the root node and the selecting variable and cut value are chosen in a way to obtain the best separation between signal and background. By using this cut criterion, the training sample is divided into two training subsamples: a signal-like and a background-like sample. Two nodes are associated to each subsample and the same process used for the root node is then used to define new cut criteria. The stopping condition for the division is then defined either by a minimum number of events reached, or a minimum or maximum signal purity (Equation 3.2). Finally, the leave nodes are called signal or background, depending on the majority of events classified as the respective type in the node.
The boosting algorithm then assigns a larger weight to the signal events in the training sample which end up in a background node (and vice versa) than to events ending up in correct nodes. A new decision tree is built based on the re-weighted event sample. The boosting is done between 100 and 500 times to obtain a set of decision trees (called forest).

During the application phase, each event is successively evaluated with all decision trees from the forest. A likelihood estimator is assigned to it depending on how often it is classified as signal or background. A cut value on this estimator can be chosen depending on the required efficiency and purity of the event selection. The purity and efficiencies for the signal and the background (ϵ_{signal} and $\epsilon_{background}$) are defined as:

$$Purity = \frac{N_{signal,selected}}{N_{signal,selected} + N_{background,selected}} \tag{3.2}$$

$$\epsilon_{signal} = \frac{N_{signal,selected}}{N_{signal}} \tag{3.3}$$

$$\epsilon_{background} = \frac{N_{background,selected}}{N_{background}} \tag{3.4}$$

The achieved purity for a given efficiency characterizes the performance of the BDT classifier. It is obtained from applying the classifier to an independent test sample of signal and background events and varying the cut value on the estimator between minimum efficiency/-maximum purity and maximum efficiency/minimum purity.
LCFIPlus uses in total eight BDTs. For each jet category, one BDT is used for beauty and one for charm and in total there are four jet classes.

Category 1	Category 2	Category 3	Category 4
trk1d0sig	trk1d0sig	trk1d0sig	trk1d0sig
trk2d0sig	trk2d0sig	trk2d0sig	trk2d0sig
trk1z0sig	trk1z0sig	trk1z0sig	trk1z0sig
trk2z0sig	trk2z0sig	trk2z0sig	trk2z0sig
trk1pt_jete	trk1pt_jete	trk1pt_jete	trk1pt_jete
trk2pt_jete	trk2pt_jete	trk2pt_jete	trk2pt_jete
jprobr5sigma	jprobr	jprobr	jprobr
jprobz5sigma	jprobz	jprobz	jprobz
d0bprob	d0bprob	vtxlen1_jete	vtxlen1_jete
d0cprob	d0cprob	vtxsig1_jete	vtxsig1_jete
d0qprob	d0qprob	vtxdirang1_jete	vtxdirang1_jete
z0bprob	z0bprob	vtxmom1_jete	vtxmom1_jete
z0cprob	z0cprob	vtxmass1	vtxmass1
z0qprob	z0qprob	vtxmult1	vtxmult1
trkmass	vtxlen1_jete	vtxmasspc	vtxmasspc
	vtxsig1_jete	vtxprob	vtxprob
	vtxdirang1_jete	1vtxprob	vtxlen2_jete
	vtxmom1_jete	vtxlen12all_jete	vtxsig2_jete
	vtxmass1	vtxmassall	vtxdirang2_jete
	vtxmult1		vtxmom2_jete
	vtxmasspc		vtxmass2
	vtxprob		vtxmult2
	trkmass		vtxlen12_jete
			vtxsig12_jete
			vtxdirang12_jete
			vtxmom_jete
			vtxmass
			vtxmult
			1vtxprob

Table 3.1 – LCFIPlus input variables used for different jet classes. The variables in blue are added to the input variables during this project.

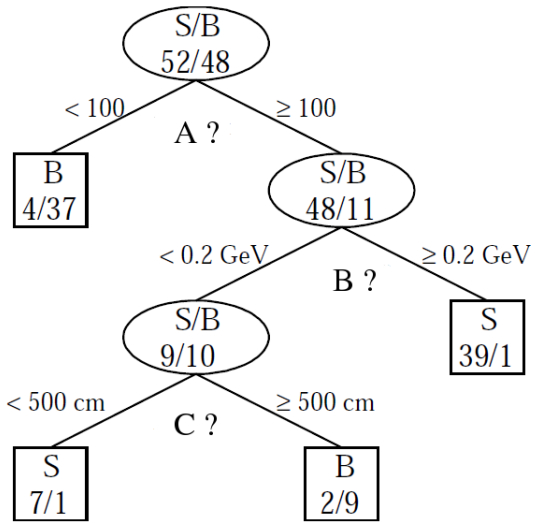


Figure 3.4 – Schematic of a decision tree. The numbers in the boxes refer to the number of signal and background events after the successive application of the selection cuts in each decision node. A, B and C are input variables used to perform cuts. From [37] with the use of general input variables A, B and C.

4 Simulation and Reconstruction Software Chain

For the simulations of the CLIC detectors, the same packages are used as the ones developed for the ILC. However they are modified in order to fulfill the requirements of the CLIC detectors. In this project, we mainly explore those related to the implementation of the detector's geometry and the flavor tagging packages. The default geometry considered is based on the CLIC_SiD concept. This geometry is the starting point of our study. By modifying its vertex detector, new geometries for CLIC are designed. Then, we perform the flavor tagging for these new geometries and the performance of each geometry is compared to the default one.
The software chain used for the geometry modification and flavor tagging is described in this chapter. First of all, the detector geometry is described using GeomConverter [2]. Then the SLIC software [6](a GEANT4 [21] based simulation software) is used to simulate the passage of particles through matter. And finally LCSim [4] is used to reconstruct and analyze the tracks of the particles.
After the reconstruction of particles tracks, the LCFIPlus package is used to perform vertex and jet finding and also flavor tagging. This software contains all the algorithms described in Section 3 and will be described further in Section 4.6.

4.1 GeomConverter

The geometry of the detector is first described in a compact way in an XML file. An example of this file can be found in [3]. This XML file is written in the form of HTML elements consisting of tags enclosed in brackets. The file is made of several sections each enclosed in tags starting for example with <h> and ending with <\h>. Each section describes different parts of the detector. First, the constants related to the radii of each part of the detector, the number of the modules and other parameters related to the position of each sub-detector are defined.
In $< material >$ section, the composite materials are described. For each of them the constituents and their fractions are given.
In $< display >$ section, the visualization parameters (color, visibility and ...) for each sub-detector are given.
The $< detectors >$ section, describes each sub-detector using the constants, the material and

also their visualization parameters as defined earlier. This part describes the envelope for each module and gives indications on how each module is placed in relation to the other modules in the simplest and the least redundant way possible. It also determines which materials are sensitive to the passing of the particles and perform measurements.
And finally some information about the readout electronics and the magnetic field in the detector is given in the sections $< readout >$ and $< fields >$.
By using the information about the relative position of each module and sub-detector in the XML files, GeomConverter creates new formats such as *lcdd*. The full information about the detector and all the modules and sub-detectors are explicitly written in this file. The XML file is readable by humans as the parameters are compact but the *lcdd* is a big file only readable by the LCSim software which uses the information to reconstruct and analyze the flow of particles in every module of the detector.

4.2 SLIC

SLIC is the GEANT4 based full simulation software used for the study of the SiD detector concept. SLIC has access to the default functions of GEANT4 and some extra commands suited for the SiD concept.
GEANT4, is a toolkit for simulating the passing of particles through matter. This software is used in wide areas of applications such as the high energy, nuclear and accelerator physics, medical and space sciences.
As the modern particle physics is going towards large-scale detectors, more accurate simulations are needed in order to understand the complex situations. GEANT4 provides software components which can be used to study basic phenomena, geometries and full-scale detector simulations for experiments at e.g. the Large Hadron Collider (LHC). The GEANT4 simulations take into account the geometry of the system, the materials involved, the fundamental particles, the generation of primary particles, the tracking of particles through materials and external electromagnetic fields, the physics processes involved in the possible interaction of the particles with the materials they are passing through, the response of sensitive components, etc.
GEANT4 also provides a visualization package, called DAWN. The latter can display the detector geometry and the trajectories of the particles through different parts of the detector.
In order to obtain the hits of the particles in the detector, SLIC uses as input the geometry description of the detector (an *lcdd* file) and the particles separately. The particles are generated using dedicated particle generators (based on Monte Carlo methods [14]). The output of SLIC is a file with *.slcio* extension. It contains the information about the amount of charge in a given volume in the detector.

4.3 LCSim

LCSim is the software used to analyze and to reconstruct the tracks of the particles registered in the events generated by SLIC. It digitizes, reconstructs and analyzes the signals in the detector. The digitization simulates the electronics used in the detector and the measurement of signals in the sensors. This will allow to compute the hits in the detector which correspond to the reconstructed signal of a particle measured in a detector component. LCSim can also be used with real measurements to reconstruct tracks.

4.4 Marlin

Marlin (Modular Analysis and Reconstruction for the LINear Collider) [5], is a C++ application framework for the analysis and the reconstruction based on LCIO [22] developed for the linear collider software. Linear Collider I/O (LCIO) is an event data model for linear collider detectors. All the data corresponding for example to one collision is contained in the LCIO event. The event data is stored in various collections specific to the type of information. The hits and the reconstructed tracks are stored in the LCIO event.
In Marlin, every task is implemented as a processor or module which can analyze LCIO events and create supplementary output collections and add them to the events. These processors parameters are defined in a *steering file* in the XML format. Figure 4.1 illustrates how Marlin processors work.

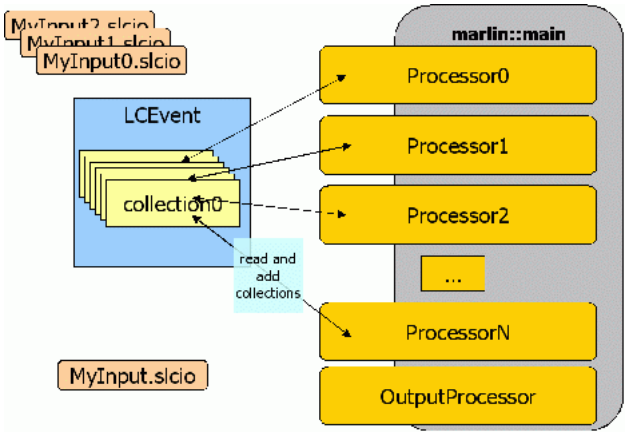


Figure 4.1 – A Schematic illustrating Marlin processors. From [5].

The flavor tagging software chain uses a Marlin processor to find the vertices and register them in the LCIO event.

4.5 ROOT

ROOT [16] is an object-oriented program and library widely used by the high-energy physics community for data analysis and is developed at CERN. It is written in C++ and provides packages for plotting, statistics and data processing applications. All the plots and data processing for this thesis are done by using ROOT. An important data container in ROOT is called a *tree*. With its *substructures* and *leaves* containing the raw data, it is optimized to reduce disk space and provides a high access speed.

4.6 LCFIPlus

Once the track reconstruction of the particles has been completed using the softwares described in sections 4.1, 4.2 and 4.3, the flavor tagging can be performed. LCFIPlus contains all the algorithms described in chapter 3. First, it performs the vertex and jet finding which are implemented as a Marlin processor.
Three major steps needed for the flavor tagging are: vertex finding, makentuple and training.

4.6.1 Vertex Finding

During the vertex finding, all primary and secondary vertices are identified. The LCFIPlus step *VertexFinder* performs this process. It is performed using an XML steering file. All the algorithms and the parameters to be used for vertex finding are specified in the steering file. The output of this step is written as an LCIO file and then processed by the next step.

4.6.2 MakeNtuple

During this step, all the input variables which serve as input for the flavor tagging are extracted from the event and written in the form of ROOT trees. The input variables used for our results are listed in Table 3.1. The LCFIPlus algorithm used in this step is called *MakeNtuple*.

4.6.3 TMVA

The Toolkit for Multivariate Analysis (TMVA) is a software designed for high-energy physics applications with a ROOT-integrated machine learning environment [8]. It contains different classifiers such as neural networks, boosted decision trees (BDT), etc. Each classifier is implemented in C++ and the framework provides training, testing and performance evaluation of the algorithms. LCFIPlus uses TMVA with BDTs for the jet classification. During the training phase, the BDT is trained using different input variables for different jet categories (cf. Chapter 3). Only half of the events are used for the training and the other half is used to evaluate the performance of the flavor tagging. We use this performance to compare different vertex detectors implemented for CLIC.

4.7 Computing Grid

The packages mentioned previously use a large amount of computing resources and can not be run locally on a single PC. For each simulation 10000 events are considered. This is the reason why we need to use the CERN computing Grid and the analysis is done in the data centers around the world.
DIRAC [1] (Distributed Infrastructure with Remote Agent Control) is used to have access to distributed computing resources.

5 Implementation of Different Geometries for the Vertex Detector

In Chapter 2, a brief overview on the CLIC_SiD detector is given. As seen in Section 2.2.1, the vertex detector is responsible for the tagging of the heavy quarks and need to have high spatial resolution, precise timing capabilities, full geometrical space coverage for low polar angles θ, low mass and sufficient heat removal from sensors and readout. These conditions, push the technology beyond its current limits. The size of the pixels for the CLIC vertex detector is much smaller than the pixels used at the hadron colliders, while complex on-chip readout and ultra-thin materials are needed. In addition, the vertex detectors are located very close to the interaction point where the beam-induced background rates are very high. Several R&D programs are addressing the various challenges such as sensors, readout, interconnects, power pulsing, thin supports and air cooling.
In this section, we are interested in the geometry of the detector and the placement of the sensors in order to have a better performance for the flavor tagging. This study will be used later to define the optimal geometry of the vertex detectors for CLIC. This study is part of a project to establish an improved full CLIC detector model by the end of 2014.
Right now, in the CLIC studies, the CLIC_SiD vertex geometry as described in [29] is used. This geometry, called the default geometry in this report from now on, is used as a starting point for the implementation of new geometries. The new suggested geometries are conceived by varying the default geometry. Two new geometries are studied during this project: the first one contains a spiral instead of disks in the vertex endcap and the second one contains double-sided sensors in the vertex barrel and in the spiral vertex endcap.

5.1 Coordinate System

The polar coordinate system in which the detector is defined is shown in Figure 5.1. The beam line is parallel to the z-axis. The polar angle θ is the angle between the z-axis and the radial coordinate r and ϕ the azimuthal angle.

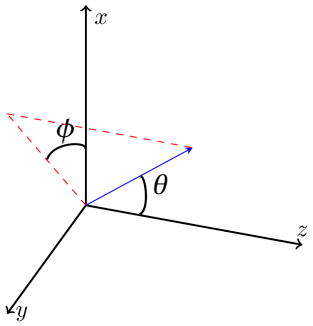


Figure 5.1 – The coordinate system in which the detector is defined. The beam line is parallel to the z-axis.

5.2 The Default Vertex Detector

The default vertex detector refers to CLIC_SiD. Its model is fully described in SLIC [6], a GEANT4 [10, 12] based package (cf. Chapter 4).
The vertex detector is placed in the innermost part of the detector and performs the flavor tagging. The default vertex detector is made of five layers in the barrel and four disks of silicon pixel detectors in the endcaps. Figure 5.2 illustrates the vertex barrel.

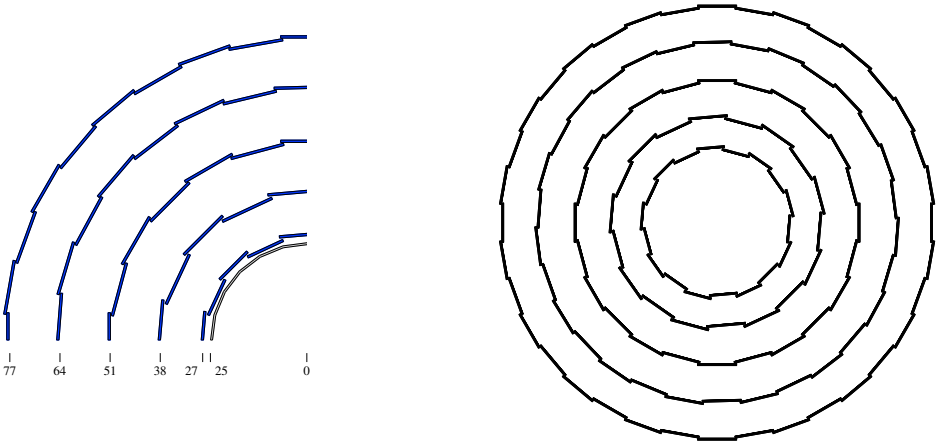


Figure 5.2 – Schematic layout of the vertex barrel detector in the xy-plane for the default detector. On the left, the distances are given in millimeters. From [23].

Each layer in the barrel is made of several modules as listed in Table 5.1. Each module contains a silicon sensor with a thickness of 50 μm. The silicon sensor is the sensitive part of each

module which detects the particles passing through it. In order to simulate the electronics used for the readout of the sensor and also the material used as support for each sensor, the silicon sensor is followed by a layer of carbon fiber with a thickness of 130 μm.

Layer	N	r[mm]	z[mm]	w[mm]
1	18	27.0	98.5	9.8
2	18	38.0	98.5	13.8
3	24	51.0	98.5	13.8
4	30	64.0	98.5	13.8
5	36	77.0	98.5	13.8

Table 5.1 – Parameters for the vertex detector barrel layers for the default geometry, where N represents the number of the modules in a layer, r the radius, z the distance from the interaction point and w the width of the module. Each module is made of 50 μm of silicon followed by 130 μm of carbon fiber in the GEANT4 simulations. From [23].

The vertex endcap contains 4 disks in the forward region made of silicon pixel detectors. Figure 5.3 shows a schematic layout of the vertex detector with highlighted vertex barrel and endcap. The whole vertex detector is then surrounded by the main tracking system made of silicon strip detectors. The parameters for the vertex endcaps are given in Table 5.2.

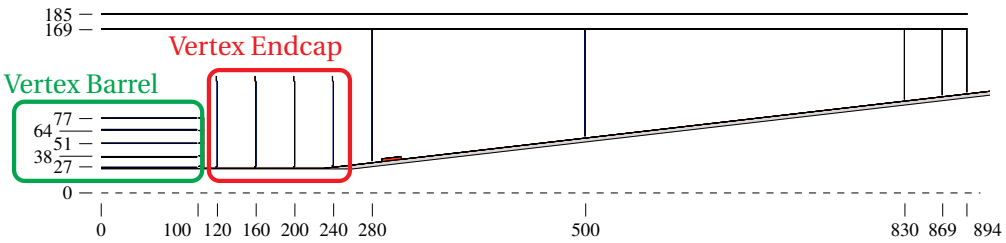


Figure 5.3 – Schematic layout of the vertex detector in the zx-plane. The distances are given in millimeters. From [23].

Disk	N	r_{in}[mm]	r_{out}[mm]	w_{in}[mm]	w_{out}[mm]	z[mm]
1	16	27.0	115.0	10.8	45.1	120.0
2	16	27.0	115.0	10.8	45.1	160.0
3	16	27.0	115.0	10.8	45.1	200.0
4	16	28.1	115.0	11.3	45.1	240.0

Table 5.2 – Parameters for the vertex endcaps for the default geometry, where N represents the number of the modules in a disk, r_{in} and r_{out} the inner and the outer radius for the disks, z the distance from the interaction point. Each module is made of trapezoidal sensors with 50 μm of silicon followed by 130 μm of carbon fiber in the GEANT4 simulations. For the trapezoidal modules, w_{in} and w_{out} represent the inner and the outer widths. From [23].

The vertex detector is designed to provide excellent point resolution with low material budget in order to minimize multiple scattering in the detector. Figure 5.4 shows the coverage of the vertex barrel and the vertex endcaps separately. The vertex detector can measure tracks down to a polar angle of about $\theta = 8°$. The number of measured points affects the performance of the particle track reconstruction. The track reconstruction performance increases by increasing the number of sensors in the detector.

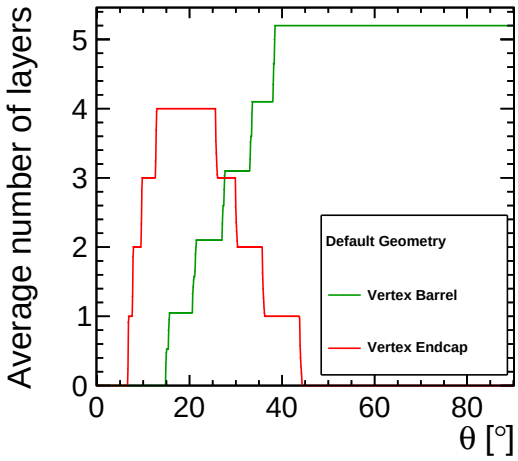


Figure 5.4 – The coverage of the vertex detector with respect to the polar angle θ. The number of layers is averaged over the azimuthal angle ϕ.

5.3 The Spiral Endcap Geometry

The material budget used in the detector affects the precision of the measurements, especially in the vertex detector. Several studies are done to reduce the material which does not serve to perform measurements such as the cables, the supports and also the cooling. In the vertex

detector, many silicon pixel detectors are used and they need an efficient heat removal system. The cooling strategy is very important to reduce the material used. For example the cooling solutions with pipes and liquids can increase significantly the material budget. The aim is therefore to use airflow cooling for the CLIC vertex detector.
However, the default geometry designed for the vertex detector is not optimal for the airflow cooling. The vertex endcaps, which are made of disks stop the air and do not allow the air to flow through the entire detector.
One solution is to use a spiral placement for the modules of the detector [30]. Figure 5.5 illustrates the cooling for the spiral vertex endcaps. The air can easily flow along the sensors used in the endcaps and remove the heat.

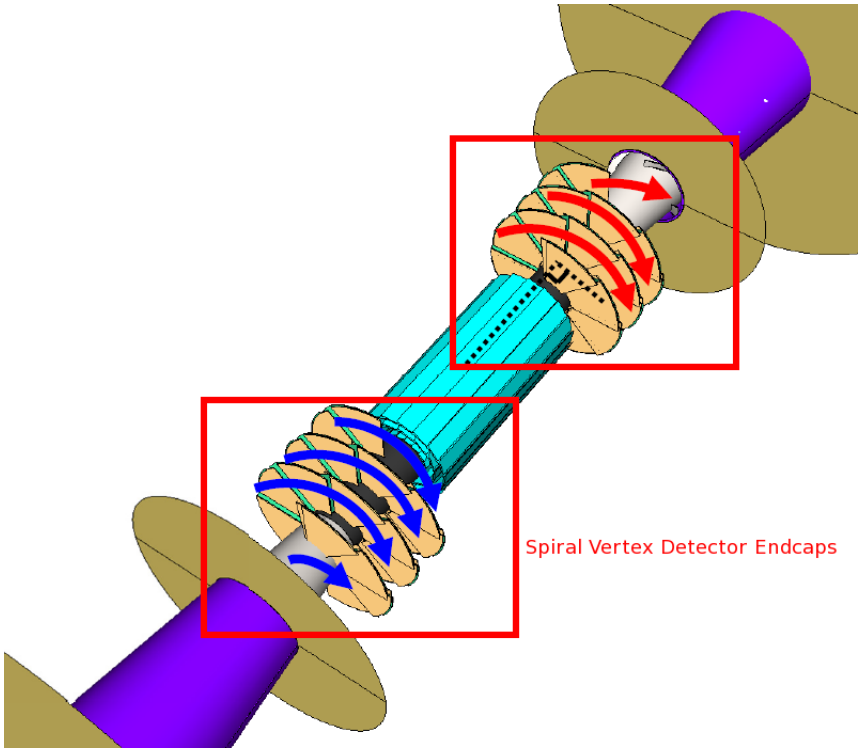


Figure 5.5 – Cooling the CLIC detector using the airflow. A solution is the use of spiral geometry instead of disks in the vertex endcaps. The airflow comes from one side of the detector and cools down the vertex barrel detectors and endcaps and exits at the other side. From [30].

Mechanical studies for the spiral endcaps are described in [18] and it is possible to build this geometry from the engineering point of view. However, for the physics point of view it is important that the spirals have a comparable performance to the disks and the placement of the spirals in the endcaps does not affect the flavor-tagging efficiency. In this chapter the material budget and also the impact parameter resolution of this new geometry are studied. The next chapter will focus on the flavor-tag performance.
The parameters for the spirals in the endcaps are given in Table 5.3. For this geometry the number of modules N in a layer is reduced to 8 compared to 16 for the default geometry. In

Table 5.3, z gives the distance of the first module of a spiral layer and the modules are spaced from each other by a distance of Δz=3.6 mm.
Figure 5.6 illustrates the spiral endcap geometry. The same barrel geometry as the default detector is used in the spiral endcap geometry.

Layer	N	r_{in}[mm]	r_{out}[mm]	w_{in}[mm]	w_{out}[mm]	z[mm]
1	8	27.0	115.0	22.7	96.6	120.0
2	8	27.0	115.0	22.7	96.6	150.0
3	8	27.0	115.0	22.7	96.6	180.0
4	8	28.1	115.0	23.6	96.6	210.0

Table 5.3 – Parameters for the spiral geometry used for the endcaps, where N represents the number of the modules in a layer, r_{in} and r_{out} the inner and the outer radius for the layers, z the distance from the interaction point for the first module of the layer. The other modules are placed at a distance of Δz=3.6 mm from the previous module in the z direction. Each module is made of trapezoidal sensors with 50 μm of silicon followed by 130 μm of carbon fiber in the GEANT4 simulations. For the trapezoidal modules, w_{in} and w_{out} represent the inner and the outer widths.

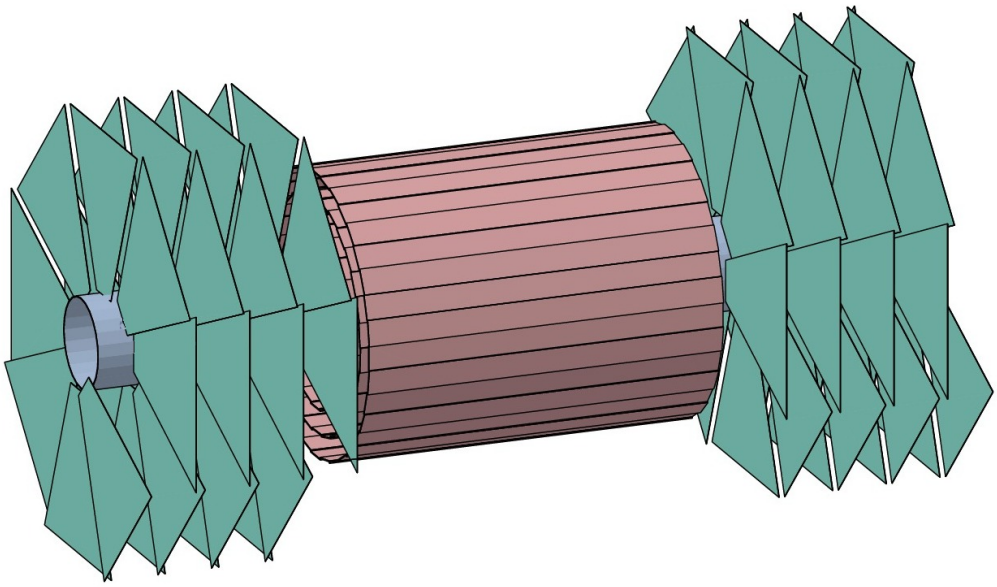


Figure 5.6 – Schematic view of the vertex detector of the spiral endcap geometry. The barrel is shown in red and is the same as the one used for the default geometry. The endcaps are shown in green.

5.3.1 Material Budget

It is important to verify that the material budget of the spiral endcap geometry does not differ much from the default one. Figure 5.7 plots the material budget for both, the default and the spiral endcap geometries.
The material budget is computed using the radiation length X_0. This parameter is a characteristic of the material and it is defined as «*the distance over which the electron energy is reduced by a factor* $\frac{1}{e}$ *due to radiation loss only*»[28].
In the computation of the material budget, we also take into account the inner parts of the detector like the beam pipe, cables and also the inner support material. For each polar angle θ, the material budget is averaged over the azimuthal angle ϕ.
As shown in Figure 5.7, the amount of material does not change much for the spiral endcaps and it remains close to that of the default geometry.

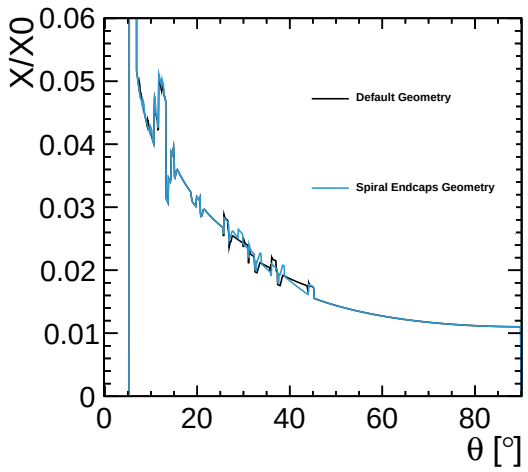


Figure 5.7 – The material budget for the simulated vertex detector for the default and the spiral endcap geometries averaged over the azimuthal angle ϕ.

The number of silicon layers as a function of the polar angles θ averaged over ϕ is given in Figure 5.8. The average number of layers for each polar angle is very similar to the default geometry. The main difference compared to the disks (see Figure 5.4) is that in the spiral endcaps, the number of layers varies with respect to the ϕ angle as each module is situated further away from the previous module along the z-axis.

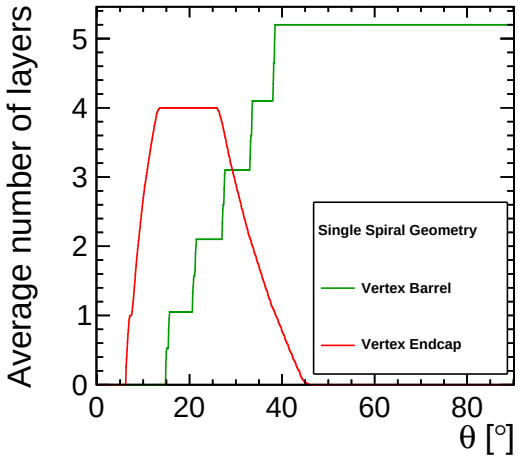


Figure 5.8 – The coverage of the vertex detector with spirals in the endcaps as a function of the polar angle θ. The number of layers is averaged over the azimuthal angle ϕ.

5.3.2 Resolution Checks

In particle detectors, the momentum is a main observable which can be computed by the Lorentz force. In a homogeneous magnetic field, charged particles are deflected by the Lorentz force and they follow a helix trajectory defined by the particle momentum $\vec{p}$, charge and the magnetic field $\vec{B}$. Assuming a magnetic field parallel to the z-axis, the radius of the circular path of the particle with a charge Q in the $r\phi$-plane is given by the Lorentz force:

$$\frac{m{v_T}^2}{\rho} = Qv_T B \tag{5.1}$$

where m is the mass of the particle, v_T is the absolute value of the particle velocity projected to the $r\phi$-plane and ρ the radius of curvature. Knowing

$$mv_T = p_T, \tag{5.2}$$

where $p_T = p\sin(\theta)$ is the transverse component of the momentum perpendicular to the magnetic field. The curvature κ is given by

$$\frac{1}{\kappa} = \rho = \frac{p_T}{QB}. \tag{5.3}$$

Charged particles which form a measurable track carry a charge of $\pm e$. The transverse momentum is then computed as

$$p_T = \frac{0.3B}{\kappa}\frac{GeV}{Tm}. \tag{5.4}$$

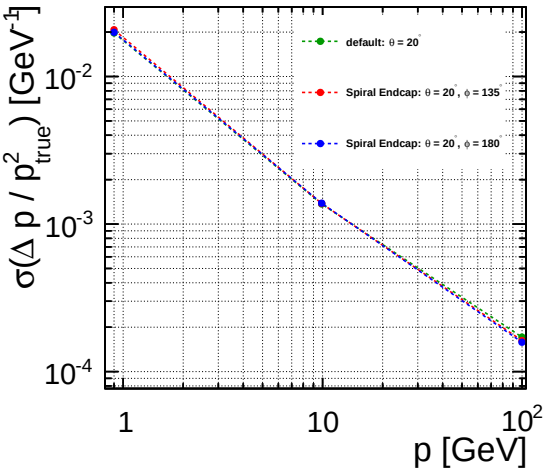
The units in the Equation 5.4 are the ones widely used in high-energy physics community. The *electron volt* (eV), is a unit of energy and corresponds to the energy acquired by an electron when accelerated through a potential difference of one volt. 1 eV is equal to 1.6×10^{-19} joules. In our detector, it is important to know how precise the momentum and the impact parameters z_0 and d_0 are measured as they are important parameters used for flavor tagging.
Figure 5.9 shows the momentum, d_0 and z_0 resolutions for the default and the spiral endcap geometries. The momentum, d_0 and z_0 resolutions are given by the following equations:

$$\sigma\left(\frac{1}{p}\right) = \sigma\left(\frac{p_{Track} - p_{Track,MC}}{{p_{True}}^2}\right) = \sigma\left(\frac{\Delta p}{{p_{True}}^2}\right), \tag{5.5}$$

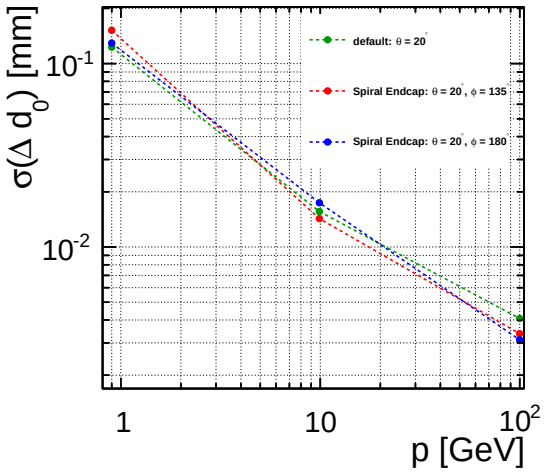
$$\sigma\left(d_{0,Track} - d_{0,Track,MC}\right) \tag{5.6}$$

$$\sigma\left(z_{0,Track} - z_{0,Track,MC}\right) \tag{5.7}$$

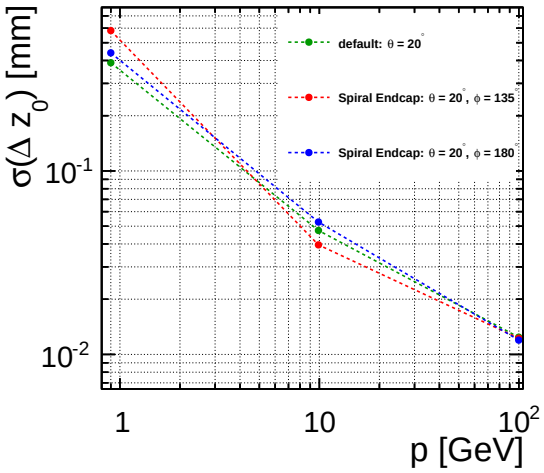
where p_{Track} is obtained by simulation and computed from the reconstructed track of the particles in the detector using Equation 5.4. $p_{Track,MC}$ and p_{True} correspond to the true value of the particle momentum before entering in the detector simulation. In an ideal detector, p_{Track} and $p_{Track,MC}$ would be the same. Indeed particles lose some of their momentum in the material in the detector and the number of the measurement points on the particles is limited.
$d_{0,Track}$ and $z_{0,Track}$ are the reconstructed impact parameters obtained by simulating the passing of the particles through the detector. $d_{0,Track,MC}$ and $z_{0,Track,MC}$ are the true impact parameters of the particles before entering the detector.
The above-mentioned resolutions are computed for the default and spiral endcap geometries using single muons with a polar angle of $\theta = 20°$ as shown in Figure 5.9 (three different momentum values are considered). The points are obtained from a Gaussian fit using 10000 simulated events and reconstructed tracks using SLIC version v3r0p3 and LCSim version 2.5. For the spiral endcap, we compute the resolutions using muons having different azimuthal angles of $\phi = 180°$ and $\phi = 135°$. This allows to compare the first and the last module of the first endcap layer as shown in Figure 5.10.
The momentum resolution for the first and the last module of the first endcap layer is very similar but the spiral endcap affects the d_0 and z_0 resolutions especially for low-momentum particles. In fact, the trajectories of these particles have smaller radius in the magnetic field and they follow a spiral trajectory and could not be measured by the modules in the spiral endcap which are placed further than the first module. In general, the higher the momentum of the particle, the better the resolution.



(a) p resolution



(b) d_0 resolution



(c) z_0 resolution

Figure 5.9 – p, d_0 and z_0 resolutions for the default and the spiral endcaps geometries for singles muons at $\theta = 20°$.

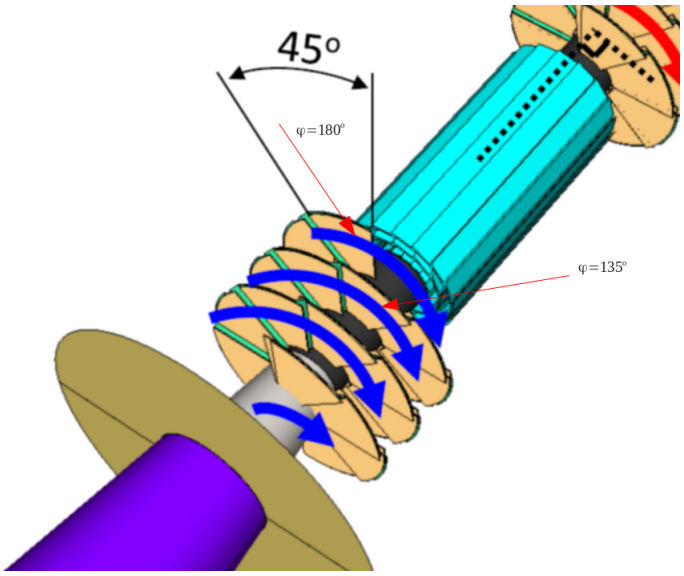


Figure 5.10 – $\phi = 180°$ and $\phi = 135°$ respectively correspond to the first module and the last module of the first endcap layer.

5.4 The Double Layer Geometry

As seen previously, in the detector it is favorable to use less material in order to minimize multiple scattering and achieve more precise measurements. In the vertex detector, we would like to have more silicon sensors and at the same time, minimize the amount of material used. The double layer geometry, uses two silicon sensors on a single support in the barrel and in the endcaps. And as the airflow is foreseen for the heat removal of the detector, spirals are used in the endcaps. Since the both sides of the sensor are close to each other, they provide more precise measurements for a more precise track reconstruction.
The vertex barrel detector of the default geometry is modified in a way such that instead of containing five layers of single-layered sensors, it contains three layers of double-layered sensors. In the vertex endcap, instead of four layers, there are three layers of double-layered sensor in a spiral arrangement.
Both sides of each module contain silicon sensors with a thickness of 50 μm. The overall thickness of the carbon fiber used to simulate the mechanical support and the electronics is 130 μm and the rest of the module is filled with air. The overall thickness of a double-layered sensor is 2 mm and is based on the CLIC_ILD study [29]. A schematic view of the double-layered sensor is shown in Figure 5.11. The thickness of the carbon is the same as the one used for the default geometry because we consider that the same amount of support structure and cables is used for the single and the double-layered sensors. The parameters of the double layer barrel and endcap are given in Tables 5.4 and 5.5.

Layer	N	r[mm]	z[mm]	w[mm]
1	18	27.0	98.5	9.8
2	24	51.0	98.5	13.8
3	36	77.0	98.5	13.8

Table 5.4 – Parameters of the vertex detector barrel layers for the double layer geometry, where N represents the number of the modules in a layer, r is the mean radius, z is the distance from the interaction point and w is the width of the module. Each module is made of two layers of silicon sensors with a thickness of 50 μm.

Layer	N	r_{in}[mm]	r_{out}[mm]	w_{in}[mm]	w_{out}[mm]	z[mm]
1	8	27.0	115.0	22.7	96.6	120.0
2	8	27.0	115.0	22.7	96.6	160.0
3	8	27.0	115.0	22.7	96.6	200.0

Table 5.5 – Parameters for the double layer geometry used for the endcaps, where N represents the number of the modules in a layer, r_{in} and r_{out} the inner and the outer radius for the disks, z the distance from the interaction point for the first module of the layer. The other modules are placed at a distance of Δz=5 mm from the previous module in the z direction. For the trapezoidal modules, w_{in} and w_{out} represent the inner and the outer widths. Each module is made of two layers of silicon sensors with a thickness of 50 μm.

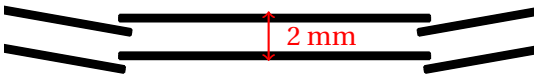


Figure 5.11 – In the GEANT4 simulations, each sensor in the double layer geometry is simulated as two silicon sensors on top of each other and the overall thickness of the sensor is 2 mm.

A schematic layout of the vertex barrel with double-layered sensors is shown in Figure 5.12. Figure 5.13, illustrates the spiral endcaps implementation with double-layered sensors.

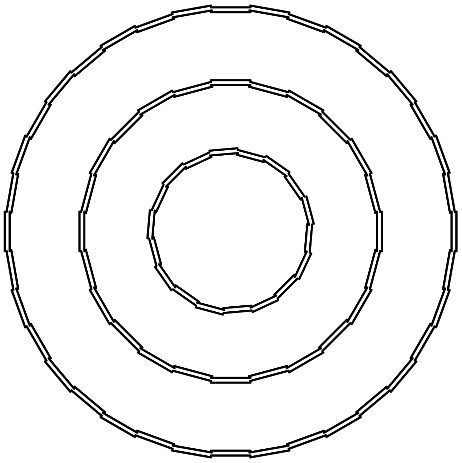


Figure 5.12 – Schematic layout of the vertex barrel detector with double-layered sensors in the xy-plane.

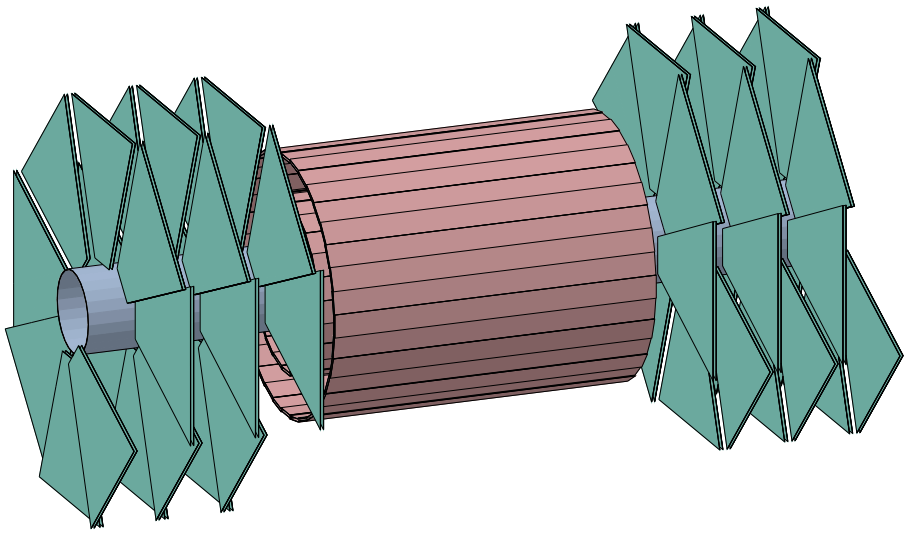


Figure 5.13 – Double-layered barrel (in red) and spiral endcaps (in green) for the vertex detector.

5.4.1 Material Budget

The material budget for the geometry using double-layered sensors is shown in Figure 5.14. It is computed as explained in section 5.3.1. The amount of material used for double layer geometry is very similar to the default geometry.

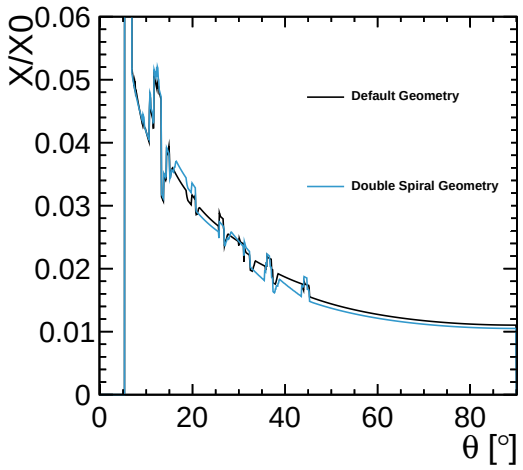


Figure 5.14 – The material budget for the simulated detector for both, the default and the double layer vertex detectors averaged over the azimuthal angle ϕ.

We have also calculated the material budget for both geometries at the polar angle of $\theta = 90°$. The theoretical and the simulated values are compared in Table 5.6. We can observe that both, theoretical and simulated values for the material budget are quite close, but the simulation gives higher values. This can be explained by the fact that for the simulation, the material budget is integrated over the ϕ angle and for some azimuthal angles, the modules overlap each other. In contrast, the theoretical computation does not consider these overlaps.

	Default	Double Layer
Theory	1.07%	1.00%
Simulation	1.10%	1.05%

Table 5.6 – Theoretical and simulated values of the material budget for the default vertex barrel and the double layer vertex barrel at $\theta = 90°$.

Figure 5.15 shows the coverage of the double layer geometry. The average number of layers in the vertex endcap is higher than for the default and the spiral endcap geometries with similar material budget (see Figures 5.4 and 5.8).

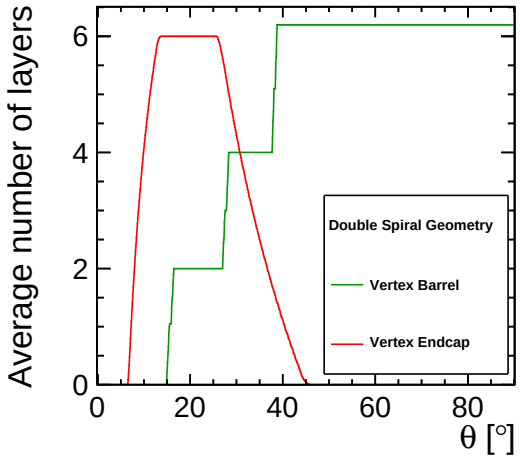
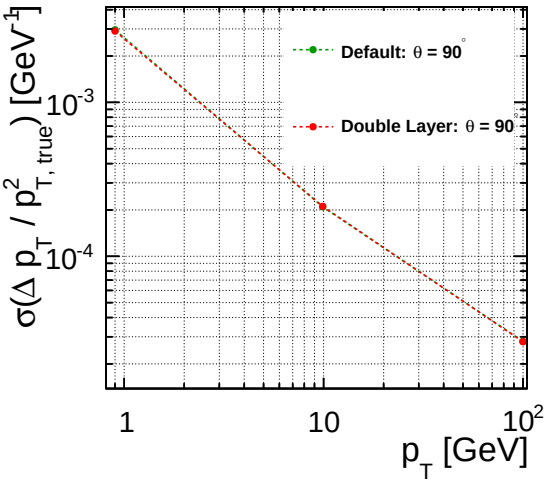


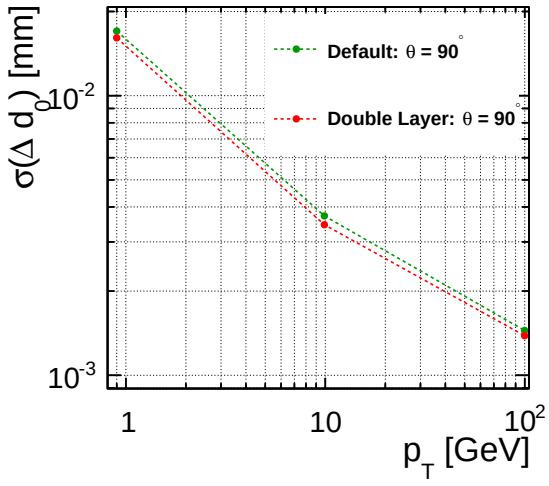
Figure 5.15 – The coverage of the vertex detector with double barrel and spirals in the endcaps with respect to the polar angle θ. The number of layers are averaged over the azimuthal angle ϕ.

5.4.2 Resolution Checks

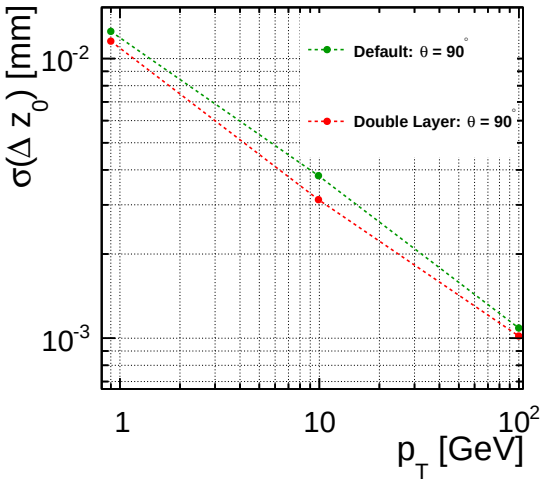
We compare the p_T, d_0 and z_0 resolutions for the default and the double layer geometries using the same setup used for the spiral endcap geometry. The resolutions are computed by using single muons with the polar angle of $\theta = 90°$ and with a momentum of 1 GeV, 10 GeV and 100 GeV (Figure 5.16).



(a) p_T resolution



(b) d_0 resolution



(c) z_0 resolution

Figure 5.16 – p_T, d_0 and z_0 resolutions for the default and the double layer geometries for singles muons at $\theta = 90°$.

We can conclude that the double layer barrel slightly improves the d_0 and z_0 resolutions. In fact, d_0 and z_0 are mostly measured by the first layer in the barrel and having two layers of sensors close to each other helps to get more measurements close to the interaction point. z_0 has a better resolution than d_0 because it consists of measuring a horizontal distance on the z-axis for $\theta = 90°$.
Figure 5.17 summarizes the coverage of the whole vertex detector (the vertex barrel and endcaps) for the above-mentioned geometries with respect to the polar angle θ (averaged over ϕ). In total, the double layer geometry has more sensitive layers in the barrel and the endcaps with similar material budget as the default and spiral endcap geometries.

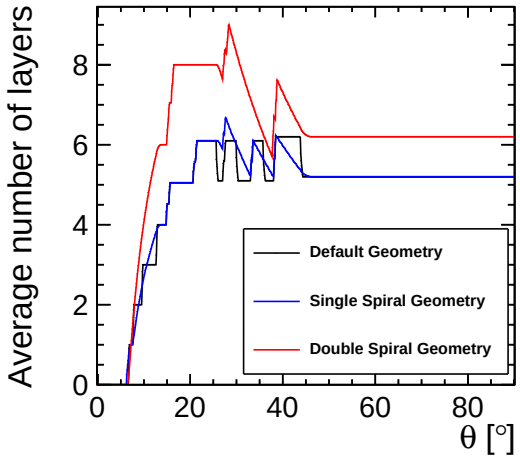


Figure 5.17 – The coverage of the whole vertex detector for 3 different geometries with respect to the polar angle θ. The number of sensitive layers are averaged over the azimuthal angle ϕ.

6 Comparison of the Flavor-Tagging Performance for Different Vertex Detector Geometries

In Chapter 5, different geometries for the vertex detector have been described and their material budget and resolutions have been compared to the default geometry. In this chapter, their performance for the flavor tagging of b and c quarks is evaluated.
The performance of flavor tagging is evaluated using dijets generated after the e^+e^- collisions. The events are generated using Monte Carlo methods [14]. Figure 6.1 sketches a dijet event in which two jets are oriented in opposite directions (by the conservation of the momentum): one from the quark and the other from the antiquark. The b quark hadronizes immediately after the collision to a B hadron. The B hadron has a rather long lifetime and decays later than a C hadron. Charm quarks hadronize immediately to a C hadron which has a shorter lifetime than a B hadron. Light quarks, shower immediately into jets after the interaction and the produced hadrons are stable. Secondary vertices are generated when heavier hadrons decay into lighter ones (cf. Chapter 3). For the flavor tagging, we want to distinguish between b, c and uds jets.

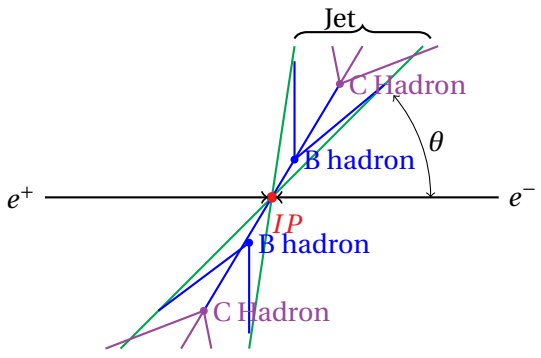
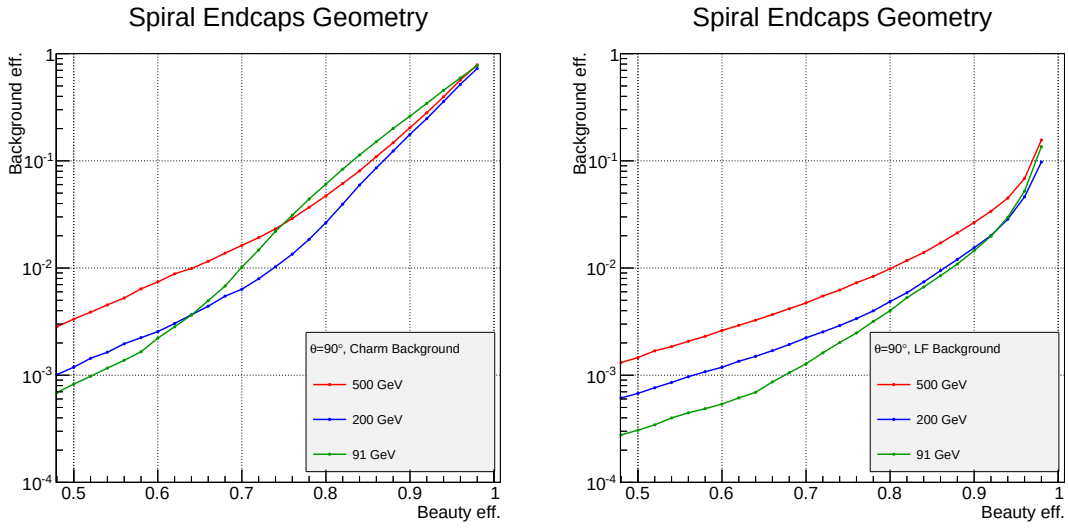


Figure 6.1 – Schematic of a dijet event used for the simulations. This example illustrates the process $e^+e^- \rightarrow b\bar{b}$.

Flavor tagging is performed using the LCFIPlus package (cf. Chapter 4.6). Its performance is dependent on the jet energy and polar angle. For this reason, the flavor tag is done for dijets with center-of-mass energies of 500 GeV, 200 GeV and 91 GeV having polar angles of $\theta = 10^\circ, 20^\circ, ..., 90^\circ$ with a flat distribution in ϕ angles. For electron-positron interactions at 1 and 3 TeV, the typical energy for the jets is 200 and 500 GeV, respectively. The center-of-mass energy in dijet events is twice the energy of a single jet.
First, the effect of spirals in the endcaps is studied by comparing the flavor-tag performance with the default geometry for different jet angles (Section 6.2). In this study, for each jet flavor and each angle, 80000 events are considered. Then the effect of double-layered sensors on the flavor tagging is evaluated in Section 6.3.
For the simulations, boosted decision trees are trained using 50% of the generated events and the other 50% is used for testing the performance of the flavor tagging. The input variables used are listed in Table 3.1.
As the flavor tagging is performed with known events, we can compute the efficiency of recognizing a particle by using Equation 3.3. We are also able to compute the background efficiency by using Equation 3.4 which corresponds to the fake rate of recognizing a background particle as a b or c quark.

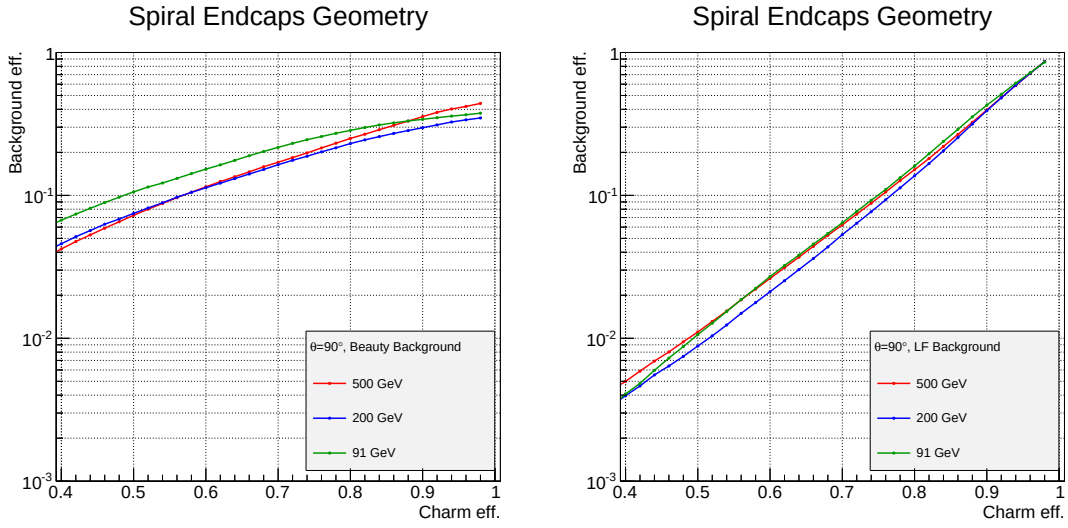
6.1 Flavor-Tag Performance Dependence on the Jet Energy and Angle

Flavor tagging is very dependent on the jet energies. Figures 6.2 and 6.3 show the flavor-tag performance in the barrel region of the vertex detector for three different dijet energies.
In Figure 6.2, the b-tag efficiency (on the x-axis) is plotted versus the fake rate of recognizing charm jets and light flavor jets as beauty jets. Figure 6.3 plots the c-tag efficiency versus the fake rate of recognizing beauty and light flavor jets as charm jets.
In general, the b-tag performance is better for jets with lower energies. This could be explained by the fact that the B hadron with lower energy has a shorter decay length and decays earlier (likely before the first barrel layer) while the B hadron with higher energy decays sometimes after the first layer. The first detecting layer can not perform any measurement on the high energy B hadron. This degrades the track reconstruction and thus the vertex finding and flavor tagging. In Chapter 3, the B and C hadron decay lengths are computed.



(a) (b)

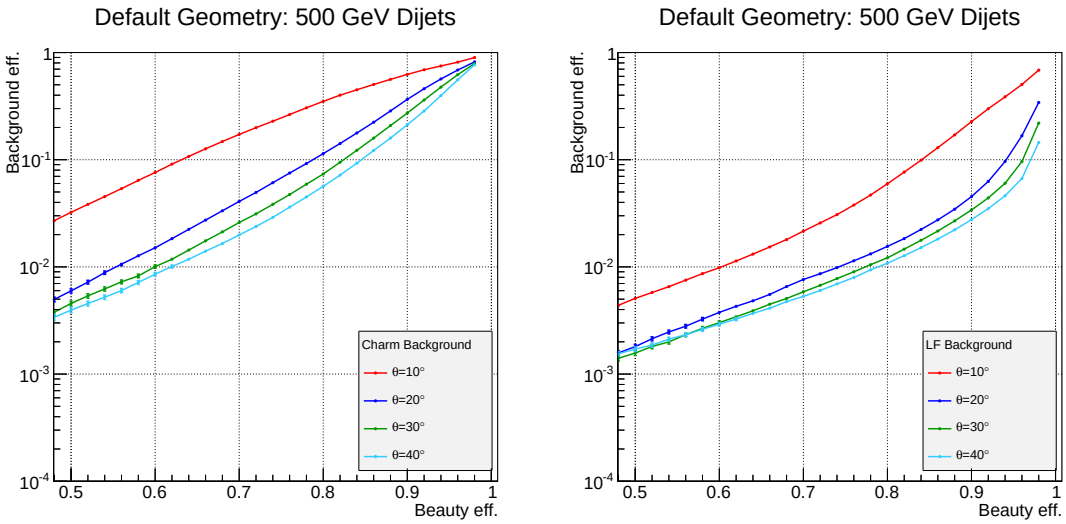
Figure 6.2 – b-tag efficiency for dijets at $\theta = 90°$ for the spiral endcaps geometry. (a) shows the fake rate for recognizing charm jets as beauty jets and (b) shows the fake rate for recognizing light flavor jets as beauty jets.



(a) (b)

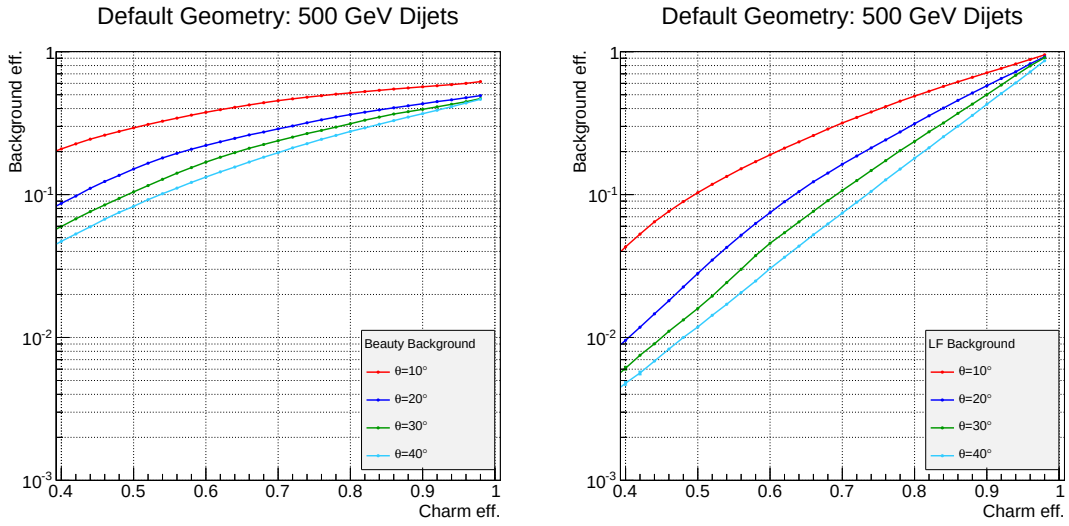
Figure 6.3 – c-tag efficiency for dijets at $\theta = 90°$ for the spiral endcaps geometry. (a) shows the fake rate for recognizing beauty jets as charm jets and (b) shows the fake rate for recognizing light flavor jets as charm jets.

Flavor-tag performance is also dependent on the jet angle. In the forward region, several factors are responsible for decreasing the performance. For low jet polar angles, some fraction of the jet is lost in the beam pipe as the endcaps do not cover the whole region and also the vertex detector resolution in the forward region is worse than in the other parts (see Figures 5.9 and 5.16). Also, the number of layers in the detector decreases with decreasing polar angles (see Figure 5.17).
Figures 6.4 and 6.5 show the dependency of the flavor-tag performance on the polar angle for dijets at 500 GeV using the default geometry (for the other jet energies and detector geometries see Appendix A.2 and A.3). It can be seen that by increasing the polar angle, the flavor tagging gets better.
In these Figures the errors on the efficiencies are shown to give an idea on the magnitude of the uncertainties. Indeed when we compute an efficiency, we select some events among all available events and the uncertainty on the selection is given by Binomial errors. The error is given by: $\sqrt{\frac{e\cdot(1-e)}{m}}$, where e is the efficiency and m the total number of jets. The errors on the computed efficiencies are very small (around 10^{-4}).



(a) (b)

Figure 6.4 – b-tag efficiency for dijets at 500 GeV with different polar angles using the default geometry.

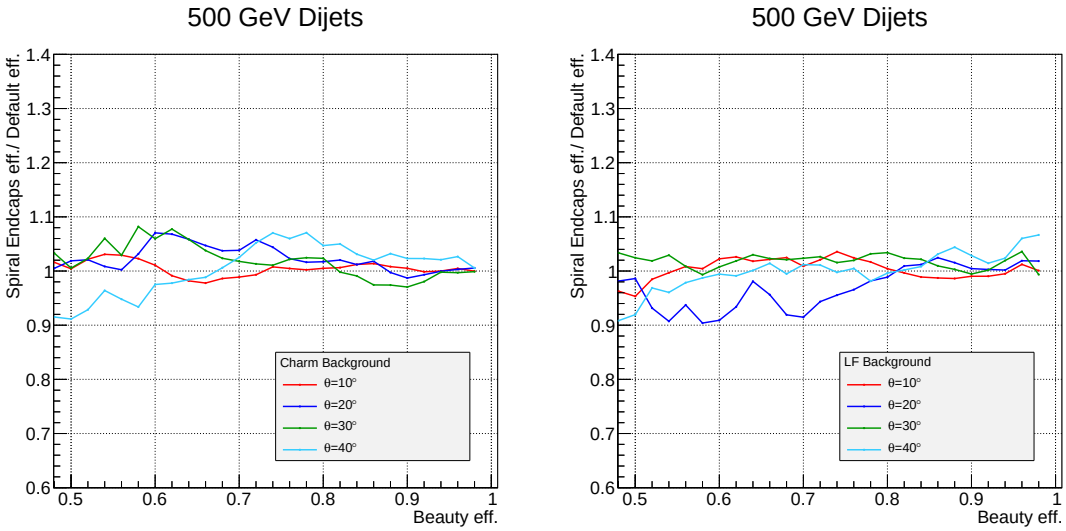


(a) (b)

Figure 6.5 – c-tag efficiency for dijets at 500 GeV with different polar angles using the default geometry.

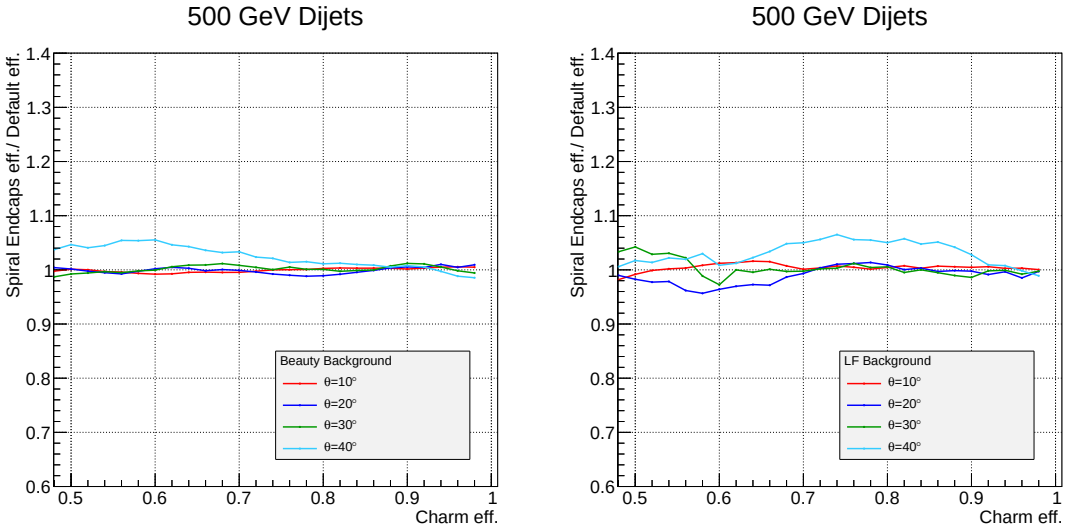
6.2 Spiral Endcap vs. Default Geometry

In order to compare the performance of the two geometries, we compute the ratio between the background efficiency for two different geometries. In this section, we are mostly interested in the forward region and we want to study the effect of the spirals endcaps on the flavor tagging. Note that the default and the spiral endcap detectors have identical barrel regions.
Figures 6.6 and 6.7 show the ratio between the background efficiency of the single spiral geometry and the default geometry for dijets at 500 GeV with polar angles of $\theta = 10°, 20°, ..., 40°$. If the ratio is smaller than one, then the spiral endcaps geometry has a better performance than the disks used in the default geometry. If the ratio is greater than one, the default geometry has a better performance.



(a) (b)

Figure 6.6 – b-tag efficiency for dijets at 500 GeV in the forward region. On the y-axis, the background efficiency of the spiral geometry is divided by the default one. (a) shows the fake rate for recognizing charm jets as beauty jets and (b) shows the fake rate for recognizing light flavor jets as beauty jets.



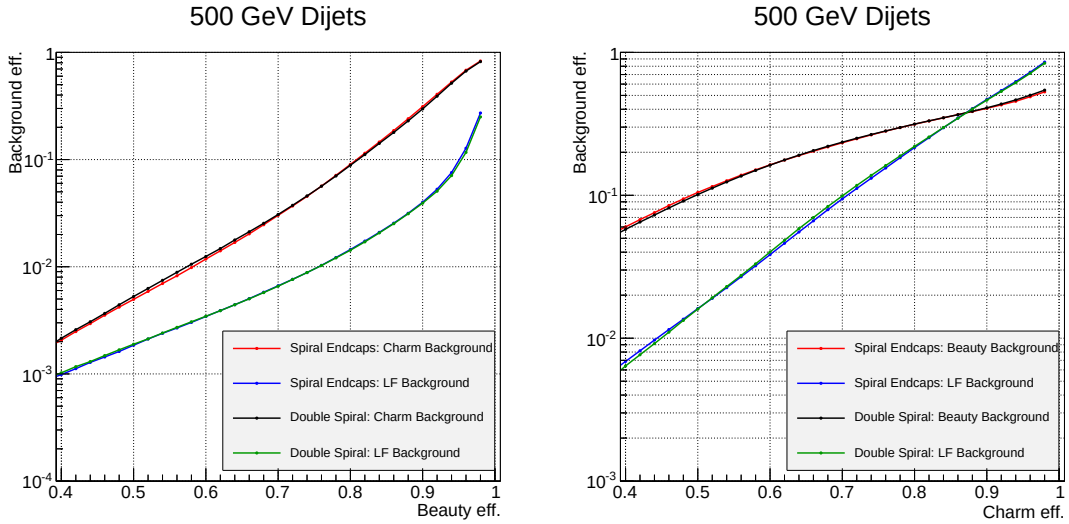
(a) (b)

Figure 6.7 – c-tag efficiency for dijets at 500 GeV in the forward region. On the y-axis, the background efficiency of the spiral geometry is divided by the default one. (a) shows the fake rate for recognizing beauty jets as charm jets and (b) shows the fake rate for recognizing light flavor jets as charm jets.

By comparing the ratios between the fake rates, we can see that spiral endcaps do not affect much the flavor tag in the forward region (for any dijet energy and angle considered). For 200 GeV and 91 GeV, the results are given in Appendix A.4.2 and A.4.3.
The changes are smaller than 10% except for dijets at lower energies at $\theta = 40°$, where the difference is around 20% for b-tagging (cf. A.4.2 and A.4.3).

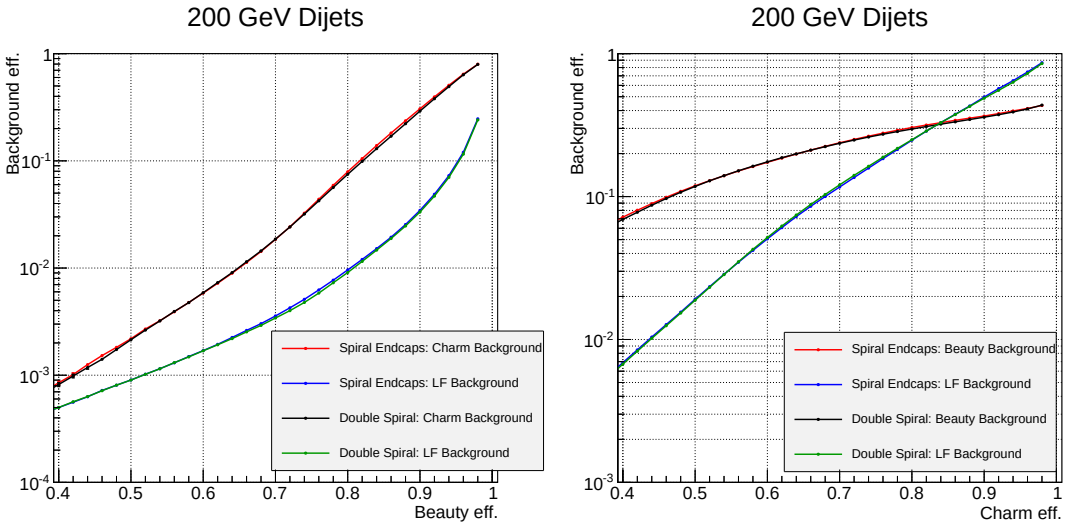
6.3 Double Layer Geometry vs. Spiral Endcap Geometry

In order to compare the double layer and the spiral endcap geometries, for each energy, we train the BDTs considering all the dijets with a mixture of polar angles. In total, for each jet flavor 720000 events are considered. Having large number of events helps to reduce the statistical fluctuations.
b-tag and c-tag efficiencies for the spiral endcaps and double layer geometries are given in Figures 6.8, 6.9 and 6.10.



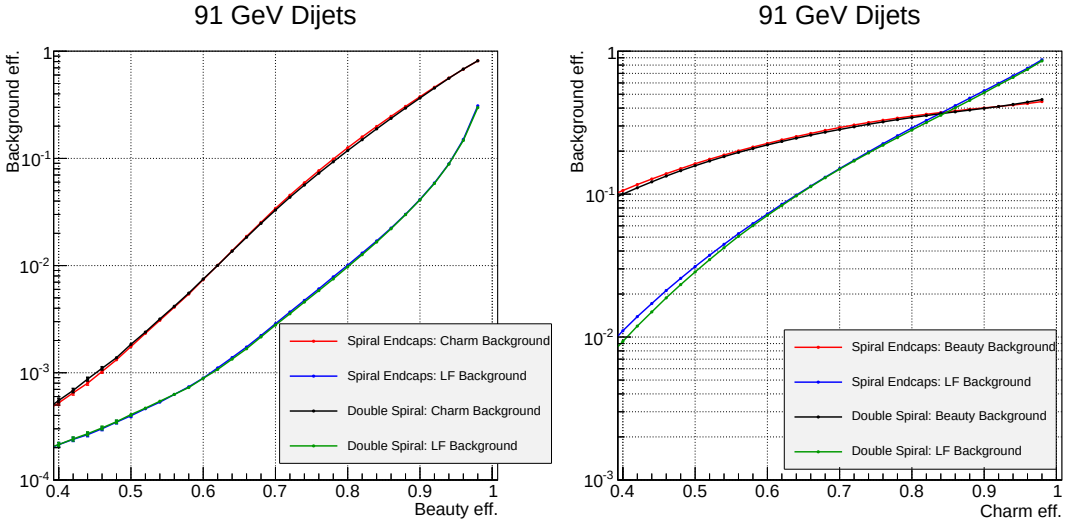
(a) (b)

Figure 6.8 – Dijets at 500 GeV (with a mixture of polar angles): b-tag and c-tag efficiencies for the spiral endcap and the double layer geometries versus the fake rate of recognizing background jets as b or c jets.



(a) (b)

Figure 6.9 – Dijets at 200 GeV (with a mixture of polar angles): b-tag and c-tag efficiencies for the spiral endcap and the double layer geometries versus the fake rate of recognizing background jets as b or c jets.



(a) (b)

Figure 6.10 – Dijets at 91 GeV (with a mixture of polar angles): b-tag and c-tag efficiencies for the spiral endcap and the double layer geometries versus the fake rate of recognizing background jets as b or c jets.

Figure 6.11 shows the ratio between the flavor-tag efficiency of the double layer and the spiral endcap for different backgrounds and jet energies. If the ratio is greater than 1, then the spiral endcap has a lower fake rate recognition over the double layer geometry. The performances

of these two geometries are quite similar, except for dijets at 91 GeV the c-tag performance shows better results for light-flavor rejection with the double layer geometry. This could be explained by the fact that double-sided sensors provide two measurements which are close to each other which improves the track reconstruction.
For the default and the spiral endcap geometries, if the decay of B hadrons occurs after the first layer (in the barrel or in the endcap), the first layer can not be used to perform measurements on the b jets. For the double layer geometry, 2 layers can not perform the measurements as double-layered sensors are used. But in total, the number of hits is identical if the decay of B hadrons occurs after the first layer for the three geometries.

We also compared the double layer and the spiral endcaps geometries using dijets having different polar angles. The results are shown in Appendix A.4.7, A.4.8 and A.4.9. The comparison is also done for the double layer and the default geometries (cf. Sections A.4.4, A.4.5 and A.4.6). Figures 6.12 and 6.13 illustrate the b-tag and c-tag performances for dijets at 500 GeV at different polar angles θ.
The results in Figure 6.12 are very dependent on the jet polar angles. We also notice that for dijets at around $\theta = 50^{\circ}$, the performance of the flavor tagging gets worse up to 40% for the double layer geometry compared to the spiral endcap geometry . Indeed, this polar angle falls in the transition region between the endcap and the barrel. Table 6.1 compares the number of layers in this transition region (only in the endcaps). With the spiral geometry (single or double-layered sensors), there are fewer layers than the default geometry and the number of layers becomes very dependent on ϕ. The number of layers for the double layer geometry is more affected. The number of layers in function of θ and ϕ angles are given in Figure 6.14.
But the c-tag performance is not affected by the double layer geometry (Figure 6.13). As said before, if the B hadrons decay after the first layer, fewer hits are observed in the detector and the track reconstruction is less precise.

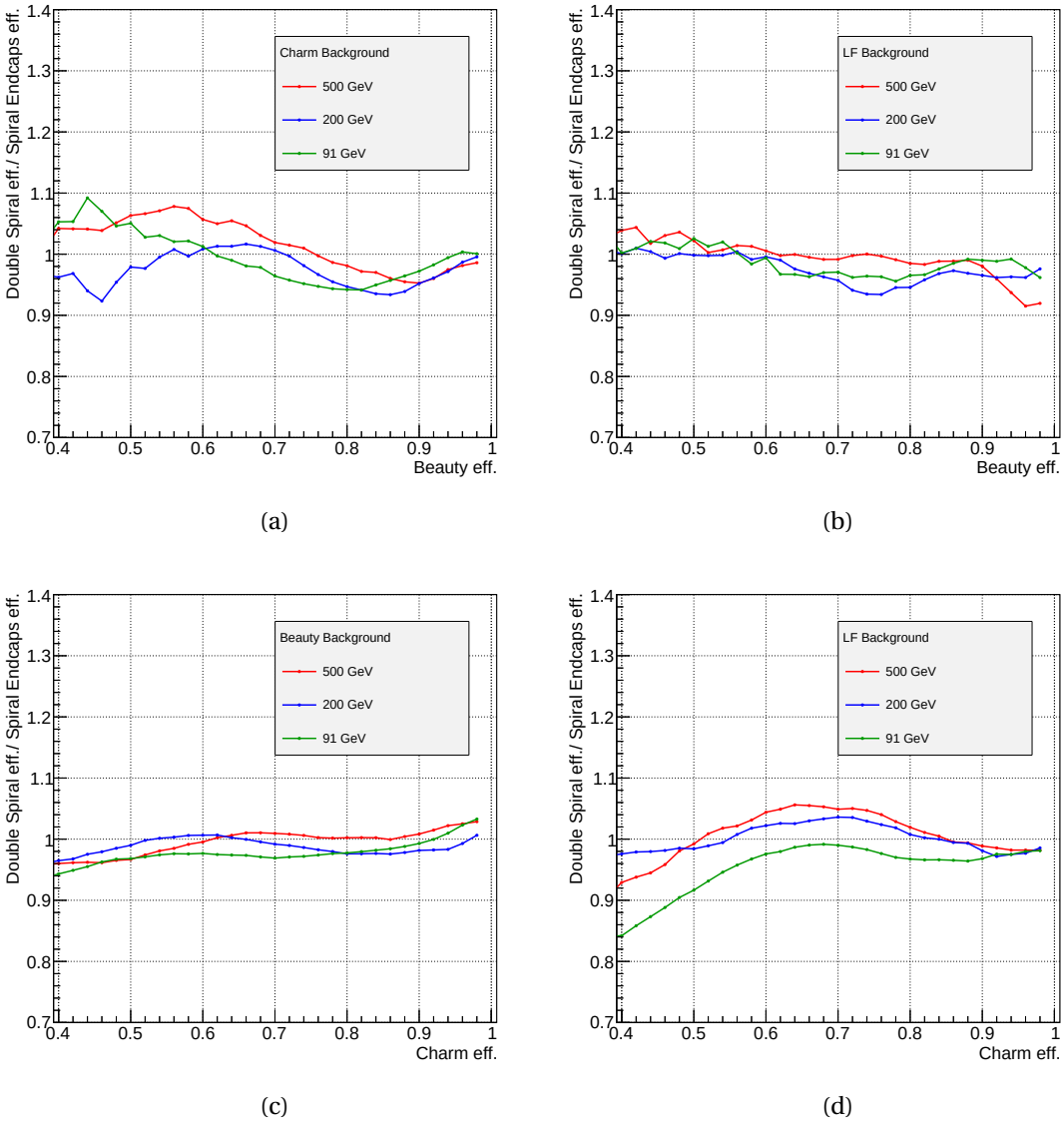
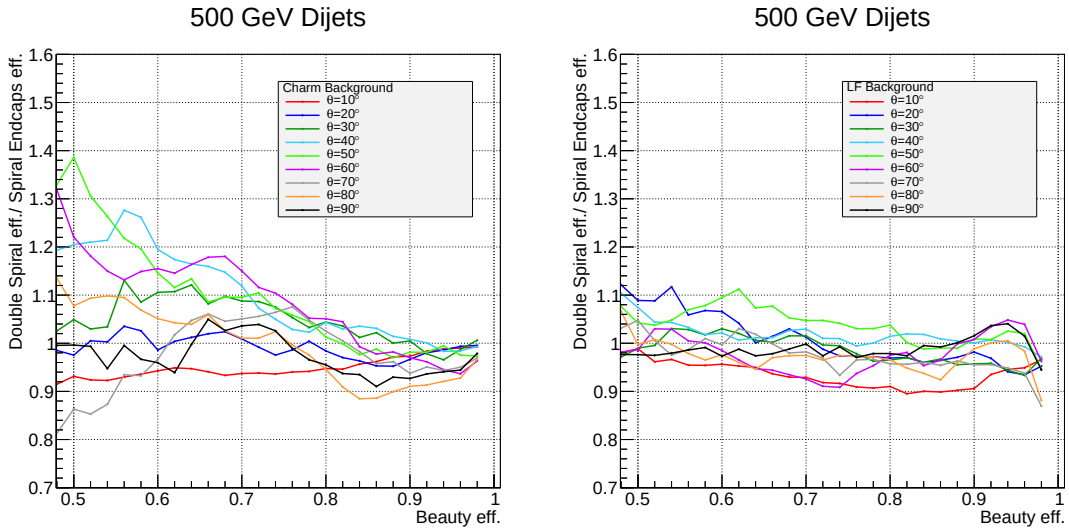
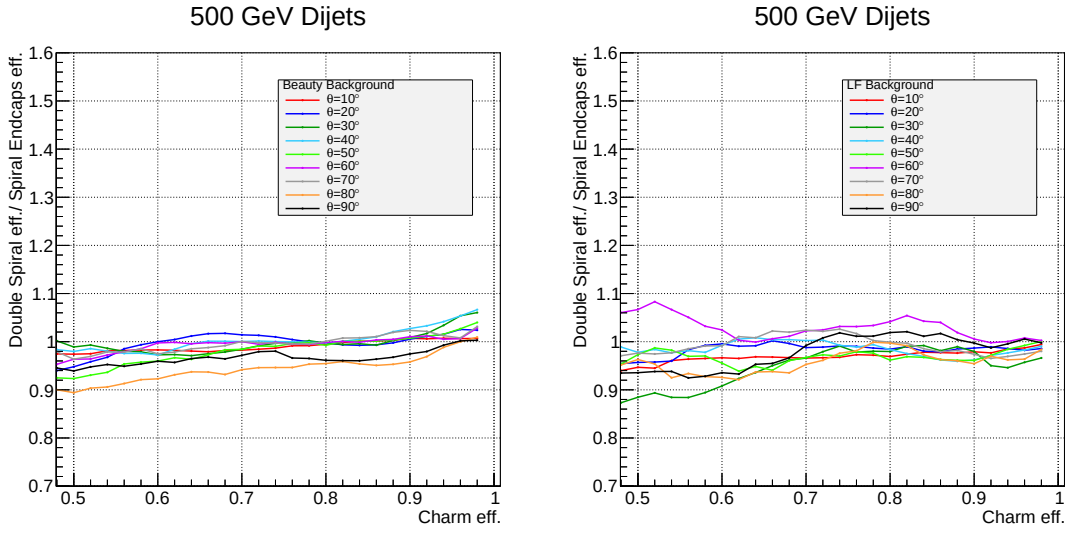


Figure 6.11 – The ratio between the background efficiency of the double layer geometry over the spiral endcaps geometry for jets at different energies.



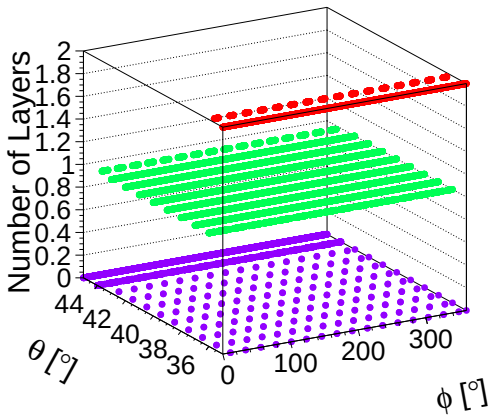
(a) (b)

Figure 6.12 – b-tag efficiency comparing the double layer and the spiral endcaps geometry for dijets at 500 GeV.

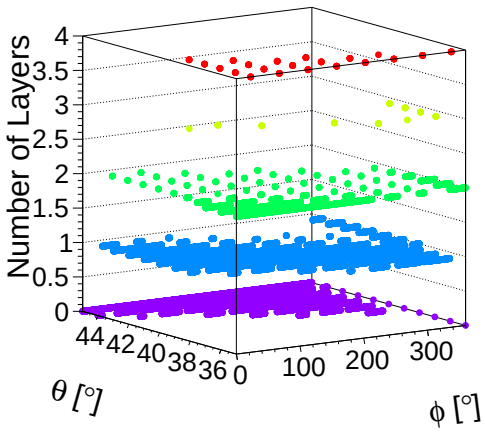


(a) (b)

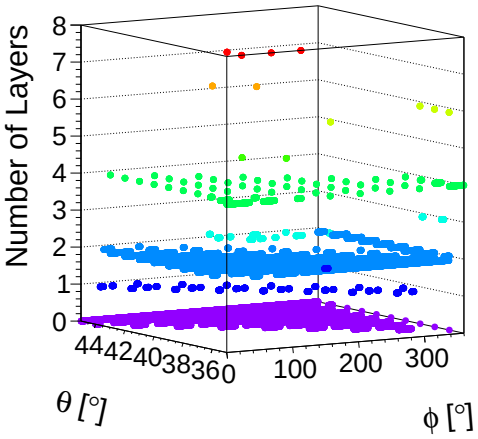
Figure 6.13 – c-tag efficiency comparing the double layer and the spiral endcaps geometry for dijets at 500 GeV.



(a) Default geometry



(b) Spiral endcaps geometry



(c) Double layer geometry

Figure 6.14 – Number of layers in the vertex endcap in function of θ and ϕ angles. The default geometry has a constant number of layers as it uses disks in the endcap region. The double layer geometry is very dependent on the ϕ angles.

θ[◦]	Default	Spiral Endcaps	Double Layer
35	1.98	1.74	2.54
36	1.25	1.52	2.23
37	0.99	1.31	1.94
38	0.99	1.12	1.67
39	0.99	0.97	1.40
40	0.99	0.79	1.15
41	0.99	0.62	0.89
42	0.99	0.47	0.65
43	0.99	0.30	0.41
44	0.39	0.14	0.19
45	0	0	0

Table 6.1 – Number of layers hit in the endcap for polar angles close to $\theta = 40°$, averaged over ϕ. In this region, the default geometry has more layers than the two other geometries. The spiral geometry loses many layers as the number of modules is dependent on ϕ.

6.4 Flavor-Tag Performance Effect on the Higgs Boson Analysis

For example, flavor tagging is used to measure the Higgs boson decay to $b\bar{b}$ and $c\bar{c}$ quark pairs. The Standard Model predicts that the production of the 125 GeV Higgs boson is dominated by the process: $e^+e^- \rightarrow H\nu\bar{\nu}$ at 3 TeV. The study of this process is described in [26] for the default geometry of the CLIC_SiD detector.
As seen previously, the spiral endcaps and the double layer geometries increase or decrease the performance of flavor tagging by up to 20%. We illustrate the effect of the variation of the fake rates on the precision of H→$b\bar{b}$ and H→$c\bar{c}$ measurements described in [26].
First, we assume that:

- For H→$b\bar{b}$, the backgrounds do not contain b-jets (they are mostly light jets).
- For H→$c\bar{c}$, the backgrounds do not contain c-jets (they are mostly beauty and light quark jets).
- The flavor tags are fully uncorrelated with the other selection variables.

Table 6.2 gives the numbers of events for the decays of the Higgs to $b\bar{b}$ and $c\bar{c}$ quark pairs after the selection performed in the analysis described in [26].

	$H\to b\bar{b}$	$H\to c\bar{c}$
Signal events	282×10^3	660×10^1
Background events	130×10^3	350×10^2

Table 6.2 – Number of signal and background events after selection for $H\to b\bar{b}$ and $H\to c\bar{c}$ decays. From [26].

If the fake rates increase or decrease by 20%, the number of background events scales by 1.2^2 and 0.8^2, respectively.
We are interested in the precisions on $\sigma(e^+e^-\to H\nu\bar{\nu})\times BR(H\to b\bar{b}, H\to c\bar{c})$, where $\sigma(e^+e^-\to H\nu\bar{\nu})$ is the cross section[1] and BR the *branching ratio*. The branching ratio is the probability for the Higgs to decay to a given final state. This precision is given by the inverse of the significance which is defined as: $S/\sqrt{S+B}$. S and B are the number of signal and background events, respectively.
Table 6.3 gives the uncertainties for the default case from [26] and when the fake rates are increased or decreased by 20%.
By comparing the results, the impact of fake rates on $H\to c\bar{c}$ is higher than $H\to b\bar{b}$. This can be explained by the fact that the purity for the $H\to c\bar{c}$ selection is much smaller.
In conclusion, a 20% change in the fake rate for light jets leads to 6-7% effect on the precision for $H\to b\bar{b}$. A change of 20% in the light quark and beauty fake rates leads to a 15% change on the precision of $H\to c\bar{c}$.

Precisions on:	$\sigma(e^+e^-\to H\nu\bar{\nu})\times BR(H\to b\bar{b})$	$\sigma(e^+e^-\to H\nu\bar{\nu})\times BR(H\to c\bar{c})$
Default	0.23%	3.1%
20% increased fake rates	0.24%	3.6%
20% decreased fake rates	0.21%	2.6%

Table 6.3 – Uncertainties for the default case (from [26]) and for the cases considering 20% increased and decreased fake rates.

[1] The probability for a specific process to occur in a given interaction.

7 Implementation of the decoding algorithm for CLICpix DAQ

To provide highly efficient flavor-tagging for heavy-flavor quark states and tau-leptons, the studies in [29] have shown that the pixel sensors need a single-point resolution of 3 µm. A material budget of $X < 0.2\% X_0$ for the beam pipe and for each of the detection layers in the vertex detector is required. This material budget corresponds to 200 µm of silicon per layer including the readout electronics and mechanical support. The single-point resolution of 3 µm can be achieved with 25 µm×25 µm pixels using an analog signal readout.
The CLICpix demonstrator is a prototype of the pixel detector readout chip designed at CERN for CLIC. It is implemented in 65 nm CMOS technology [35] and will later be bump-bonded to a silicon sensor. The small prototype ASIC is produced with an array of 64 × 64 pixels.
The time structure of the beam for CLIC and ILC is shown in Figure 7.1. As the machines accelerating the beams are quite different for the two colliders, the beam characteristics are also different. For CLIC, the beam consists of bunch trains of 312 bunches with a repetition rate of 50 Hz. Within a bunch train, the bunches are separated by 0.5 ns and for 3 TeV center-of-mass energy, there are $3.72 \cdot 10^9$ particles per bunch [29]. The data acquisition system has to work in high rates in order to be able to read the signal from the chip and to prepare the chip for the next bunch train.
CLICpix can provide a 4-bit Time of Arrival (TOA) and a 4-bit Time over Threshold (TOT). The TOA measurement helps to find the bunch to which each hit corresponded. TOT measures the energy of the incoming particle. For each hit in a pixel, a current pulse is generated. A preamplifier and a capacitor are used to convert the current to a voltage. The length of time the voltage stays above a certain threshold is used to measure the energy. The longer the TOT, the higher the energy deposited by the particle.

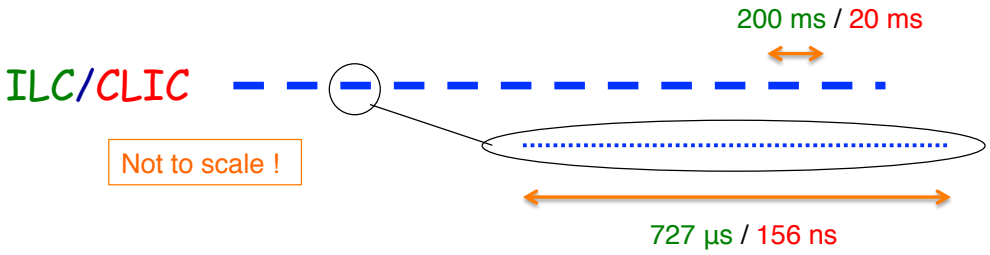


Figure 7.1 – The CLIC and ILC beams. From [17].

In this chapter, we study the data acquisition system designed for the CLICpix chip prototype and implement a decoder for the chip data.

7.1 The CLICpix Data Acquisition System

The data acquisition (DAQ) system for the CLICpix chip has to be optimized for high data rates. Figure 7.2 shows the CLICpix measurement setup. This setup is used to evaluate CLICpix features. The final DAQ system will be designed based on the measurements obtained with this setup. The CLICpix board contains the chip and the power supply and it is connected to the FPGA board. The FPGA board contains the Xilinx Spartan-6 LX45 FPGA with Xilinx MicroBlaze microprocessor and can be connected with an Ethernet connection to the PC. The CLICpix board is not part of the DAQ system.

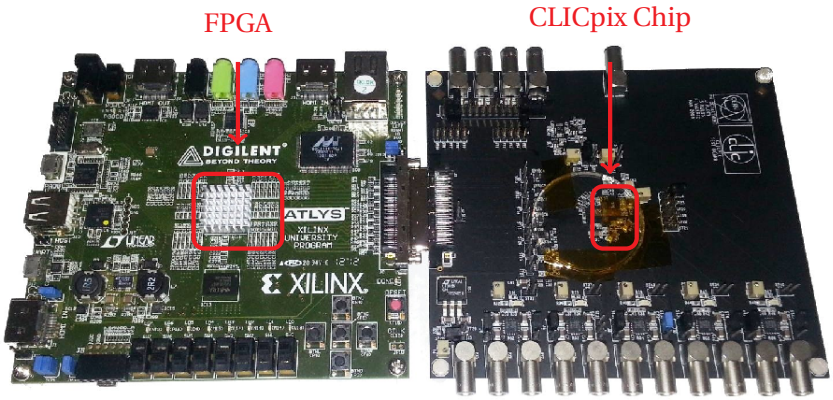


Figure 7.2 – Readout board (left) and CLICpix chip board (right).

The CLICpix chip is designed to operate at the frequency of 320 MHz. It sends its data as one bit per clock cycle and only indicates the start of the frame with a start signal. The data is stored in the registers available in the FPGA device and once the chip has finished sending the whole frame, the data is sent to the PC. This will minimize the Ethernet connection occupancy. The TCP protocol is used for sending the data to the PC. As a TCP packet has a big size, it is

more convenient to send more data for each packet.
Section 7.2 explains the procedure through which the chip sends its data. The chip can send its data using compressed or uncompressed modes. In compressed modes, the chip does not indicate when it has finished sending an entire frame of data. For this reason, it is important to understand the compression. We implement a decoder in the FPGA which decodes the data from the chip and indicates the end of a frame. Once the end of the frame is known, the chip data is sent to the PC and finally the TOT and TOA values are extracted offline.

7.2 CLICpix Compression Algorithm

For the first CLICpix prototype several data compression modes are designed. The user can configure the compression mode used for the chip. In this project we only focus on the full compression mode in which only non-zero pixel signals are transferred.
For any compression mode the CLICpix chip associates a bit flag to a group of pixels. The chip is a matrix of 64 × 64 pixels and its array is partitioned into blocks of pixels. A bit flag is associated to each block. For the fully compressed mode, if the flag is set to "0", no pixel in that block is hit and the chip does not send zero values for the pixels in that block. If the bit flag is set to "1", then at least one pixel in that block has a non-zero signal.
The compression levels are described below and are shown in Figure 7.3:

- A *double column* (DC) regroups two columns of the array and a bit flag is associated to it. If the flag is "0", then none of the 128 pixels in the DC contains non-zero data (no hit in the DC). One chip contains 32 DCs.

- A DC is divided into 8 *Super Pixels* (SP). A SP regroups an array of 2 × 8 pixels. Like the DC, a bit flag is associated with each SP. If the flag is "1", then at least one of the pixels in the SP is hit.

- Finally, a *Pixel* has its own flag. If it is hit, then the flag is "1" and the data for that pixel comprise 8 bits: 4 bits for the TOT and 4 bits for the TOA. If the pixel flag is "0", then the TOT and the TOA are not sent as both of them are null.

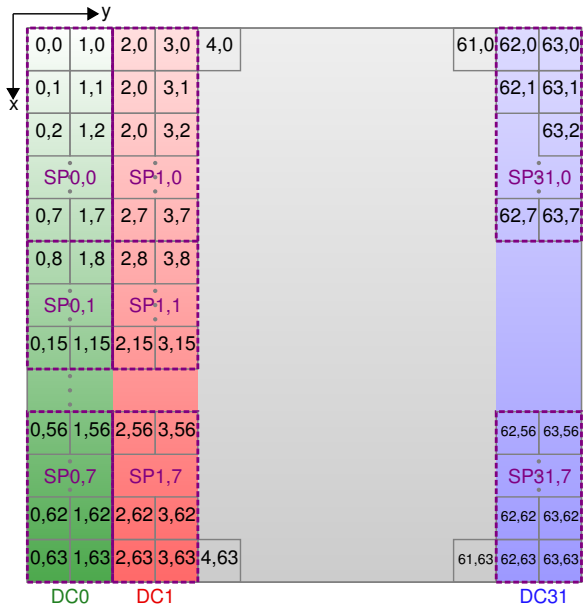


Figure 7.3 – CLICpix matric numbering for the compression.

The chip transmits the bits in series: one double column at a time. By using a fast readout clock (of around 320 MHz), each pixel shifts the data to the next one and the counters work as a long shift register. This procedure is shown in Figure 7.4.

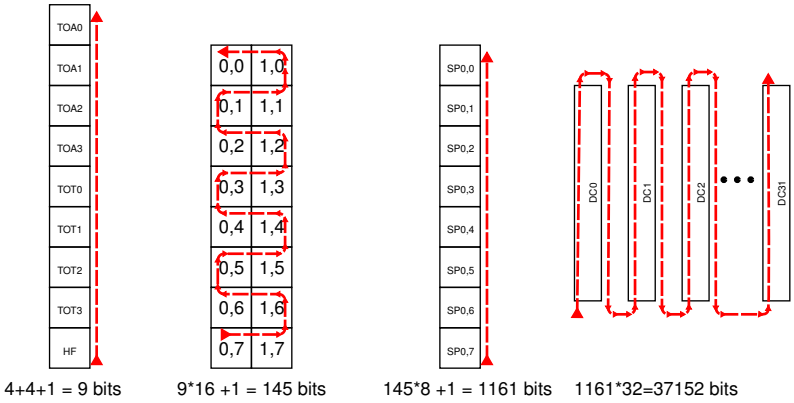


Figure 7.4 – CLICpix matrix readout: at pixel, super pixel, double column and chip level.

7.3 Data Deserialization

For Spartan-6 FPGA devices, serializing and deserializing circuits are embedded hardware in order to achieve higher operational data rates. The data coming from the chip can be

deserialized by a factor of 2, 4 or 8. The deserializing circuit can operate at higher frequencies than the other parts of the FPGA. This circuit can be used to acquire the data from CLICpix at a very high frequency. At the output of the deserializer, the data from the CLICpix is grouped in a vector and can be processed by a slower clock.
The Figure 7.5, shows the deserialization of the input data by a factor of 2. In the 2 bit input case, two data bits can be processed at the same time with a clock two times slower than the input clock to the deserializer.

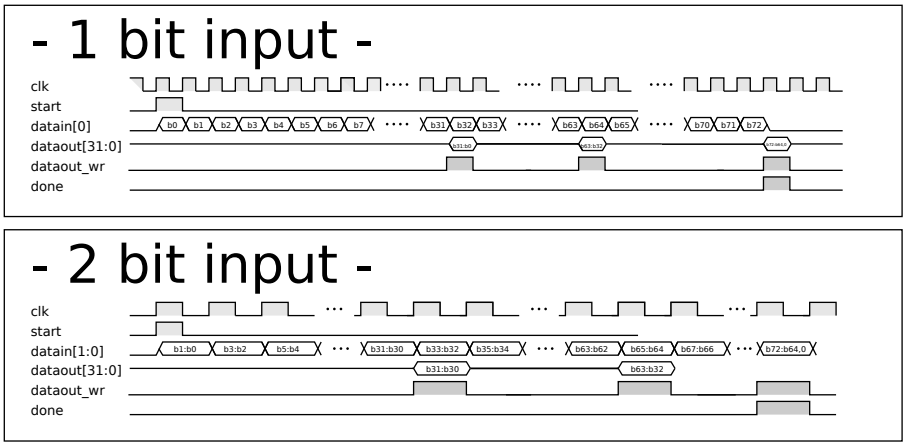


Figure 7.5 – Deserialization by a factor of 2.

7.4 Decoder Implementation

The state diagram shown in Figure 7.6 is used to implement the decoder in VHDL for the FPGA. VHDL (VHSIC Hardware Description Language) is used for the design of digital systems. The rectangles in Figure 7.6 represent the state of each bit. In this state diagram, we only have 5 states: Idle, Double Column, Super Pixel, Pixel and Data Bits.
The Idle state is the state before the chip starts sending data. When the signal *start* changes its value from "0" to "1", then the chip sends its very first bit. The Double Column, Super Pixel and Pixel states mean that the last bit received corresponded to one of these blocks' flag bit. When we are in one of the mentioned states, depending on the value sent by the chip and the values of the counters in the next rising edge of the clock we make a transition. In the diagram, the data coming from the chip is surrounded by a blue rectangle. The values of the counters are verified in the decision blocks in diamond shapes. The abbreviations used in the state diagram are listed in Table 7.1.
The Data Bits state corresponds to the TOT and the TOA value of a pixel with a non-zero bit flag. The diamonds are the decision blocks. They check if the counters have reached their maximum value or if the signal start has changed its value to "1". Finally, the output of the state diagram is surrounded by a red rectangle. When this output value is "1", then the chip has finished sending its data and we can reset the chip for a new measurement.

Counter		Counter Maximum Value	
BC	TOT and TOA bit counter	NB	7
PC	Pixel counter	NP	15
SPC	Super pixel counter	NSP	7
DCC	Double column counter	NDC	31

Table 7.1 – Abbreviations used in Figures 7.6 and 7.8.

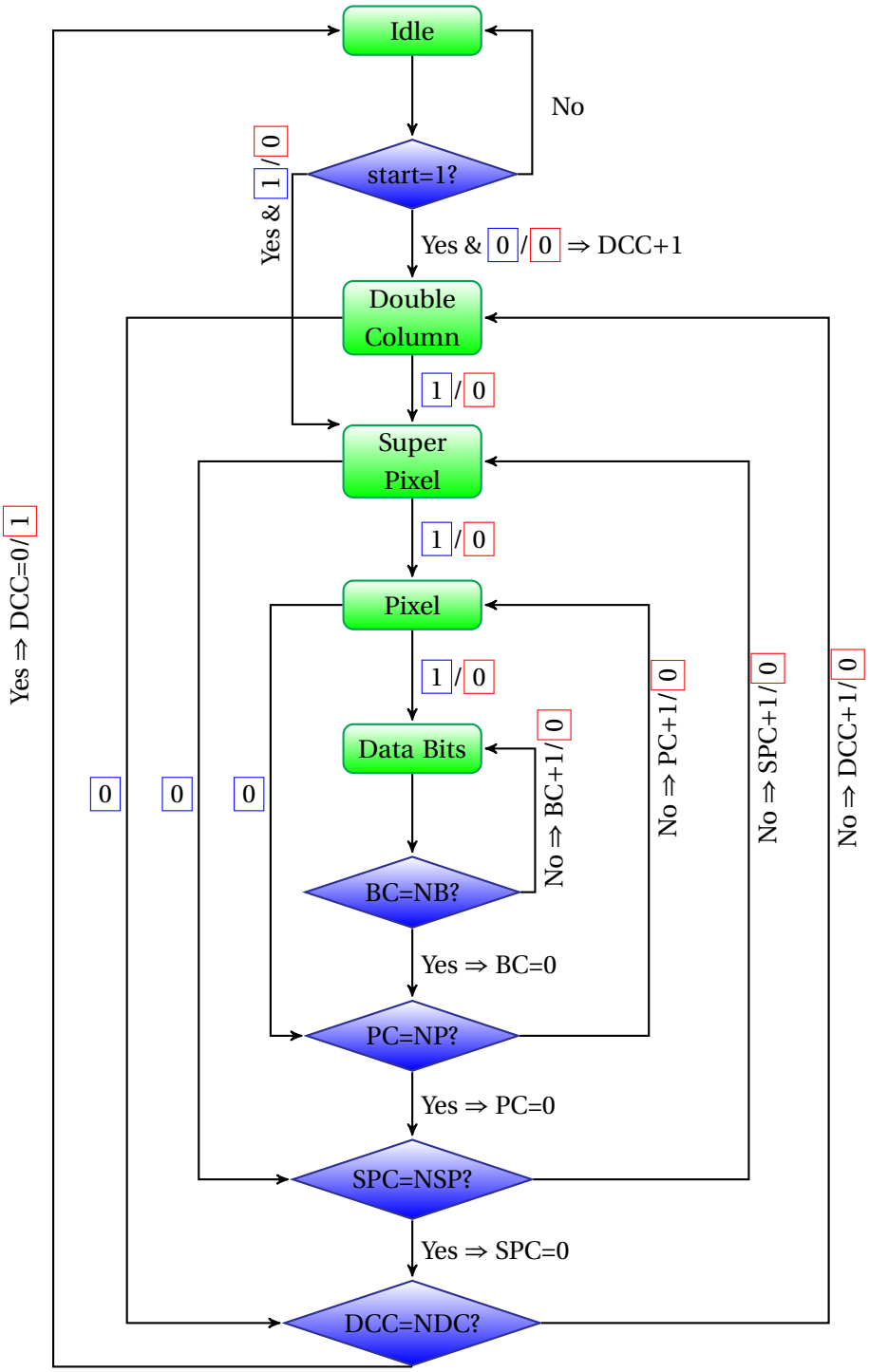
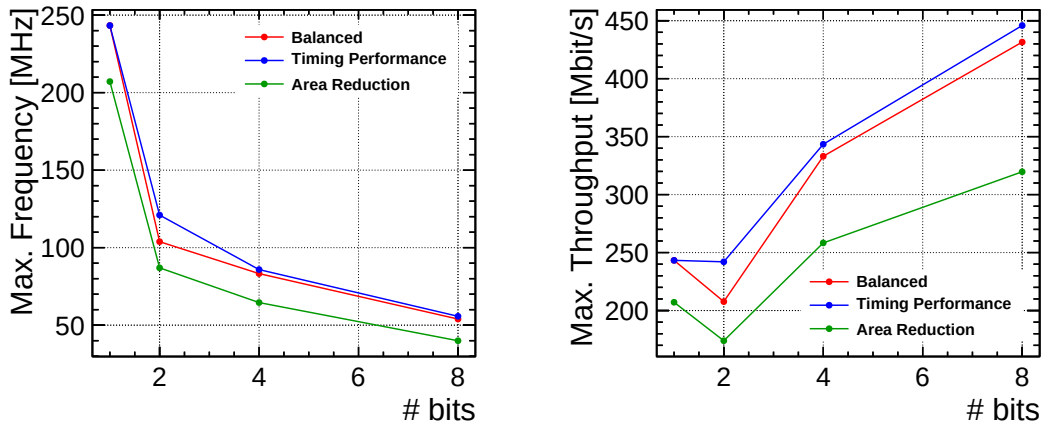


Figure 7.6 – The State Diagram for decoding the CLICpix data in full compression mode. The inputs (coming from the chip) are surrounded by a blue rectangle and the outputs (which inform if the chip has finished sending data) are surrounded by red rectangle. Table 7.1 lists the abbreviations used.

Figure 7.7 shows the maximum frequency and throughput at which the decoder can operate. On the x-axis, the total number of serialized bits are given, and on the y-axis the maximum throughput or frequency achieved by the decoder. These values are obtained by simulation and after synthesizing the decoder code in VHDL (without any hardware test) using the Xilinx ISE (Integrated Software Environment) 14.4. This tool is provided by the Xilinx board and is used for the synthesis and the analysis of the HDL (Hardware Description Language) designs. The synthesis can be done in several ways to achieve different goals. The *timing performance* improves the speed of the generated logic by reducing the levels of logic. The *area reduction* goal, optimizes the design for area by reducing the total amount of logic used for the implementation. And the *Balanced* option makes a trade-off between high speed and area reduction.



(a) The maximum frequency

(b) The maximum throughput

Figure 7.7 – The maximum frequency and the maximum throughput obtained in simulation by implementing the Diagram in Figure 7.6 on the Spartan-6 FPGA board using different design goals for the synthesis.

By increasing the number of deserialized bits sent by the CLICpix chip, it is possible improve the maximum frequency and throughput achieved by the decoder. The goal is that the decoder operates in real time at a frequency of at least 320 MHz.

As we want that the decoder follows the data rates of the CLICpix chip, we have increased the number of the states in the state machine. In Figure 7.8, the state Data Bit 7 is added which corresponds to the last bit of the pixel data corresponding to the TOA+TOT. The results for this new implementation of the decoding algorithm are given in Figure 7.9. More the number of states is increased in a state diagram, less tests are needed and hence less logic is needed for its hardware implementation.

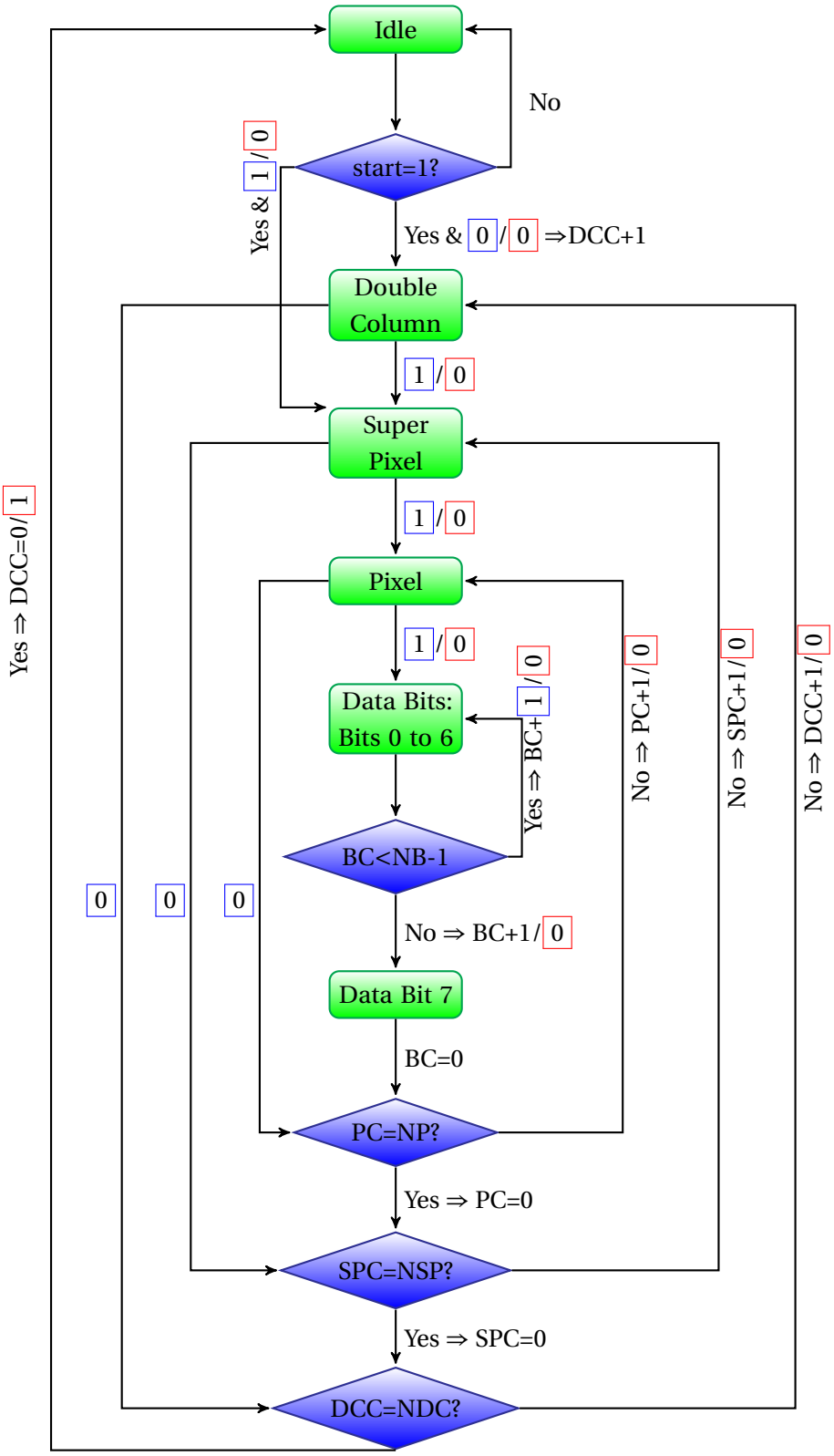
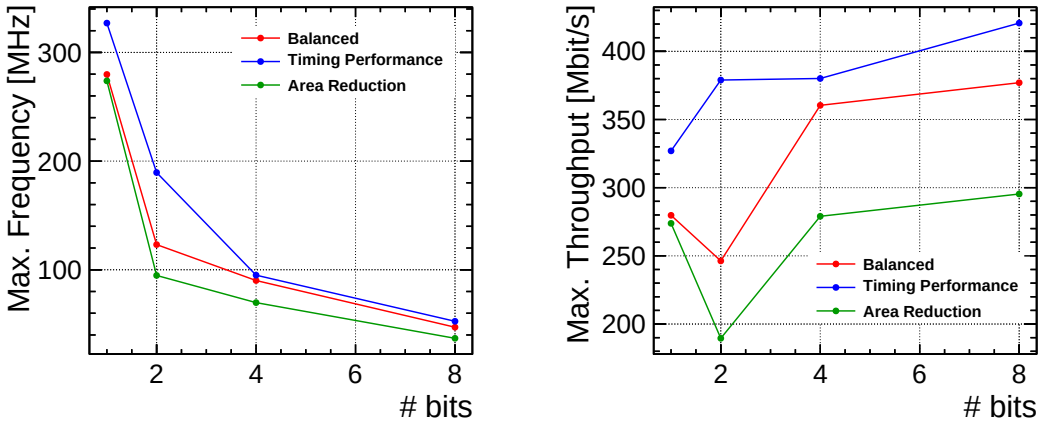


Figure 7.8 – The State Diagram for decoding the data from CLICpix. The inputs (coming from the chip) are surrounded by a blue rectangle and the outputs (which inform if the chip has finished sending data) are surrounded by red rectangle. Table 7.1 lists the abbreviations used.



(a) The maximum frequency

(b) The maximum throughput

Figure 7.9 – The maximum Frequency and the maximum throughput obtained in simulation by implementing the Diagram in Figure 7.8 on the Spartan-6 FPGA board.

As we can see in Figure 7.9, the decoding algorithm described in Figure 7.8 allows to operate at higher frequency than the algorithm described in Figure 7.6 as expected.

7.5 Decoder Tests on the CLICpix Demonstrator

The decoder is also tested on the CLICpix chip. For the CLICpix prototype, the silicon sensor is not yet bonded to the ASIC. For testing the chip, there is the possibility of applying a voltage to the preamplifier of each pixel. The charge is then collected by each pixel and CLICpix sends the data to the FPGA.
For the testing, the frequency of the clock with which the chip sends the data is set to 50 MHz (lower than the maximal data rates the chip can achieve). The data from the chip are deserialized by a factor of 8 in the FPGA which means that the decoder decodes 8 bits at a time with a frequency of 6.25 MHz.
When we want to read a frame, first a STROBE signal is sent to the chip. When it is received by CLICpix, the chip sends a READY signal to the FPGA which indicates that the chip is ready to send its data. The delay between the READY signal and the first data bit sent by the chip is fixed to 42 clock cycles by the chip designer. For a 50 MHz clock, this delay corresponds to 840 ns.
In Figures 7.10, 7.11 and 7.12 some readout examples are given. The READY signal is read by the channel 3 (in magenta) of the scope, the CLICpix data by the channel 2 (blue) and the end of the frame is read by the channel 1 (yellow).
Figure 7.10, shows the decoding time when reading an empty matrix. The delay between READY and the end of frame is Δt=1.830 µs which corresponds to 91.5 clock cycles. The signals are very noisy. This is the reason why the measurement is not very precise. The decoding took

around 49 clock cycles. We would have expected 32 clock cycles corresponding to 32 double columns bit flags. The difference could be due to the deserialization. When the data from the chip are deserialized by 8 bits, the READY signal is perceived by the decoder 8 clock cycles later. In addition, the READY signal can be anywhere in the 8-bit vector which can introduce a delay of maximum 8 clock cycles.
Figures 7.11 and 7.12 show the decoder output when the pixels with the coordinates (0, 0) and (62, 0) are set to one, respectively. In both cases the delay between the READY signal and the end of frame signal is Δt=2.48 µs. This corresponds to 82 clock cycles if we subtract the 42 clock cycles delay of the chip. With one pixel on, 64 clock cycles are needed to finish the decoding in theory. But again, the data takes more time to be decoded and this can again be explained by the delays due to the deserialization.
This study will be pursued in more detail in the near future.

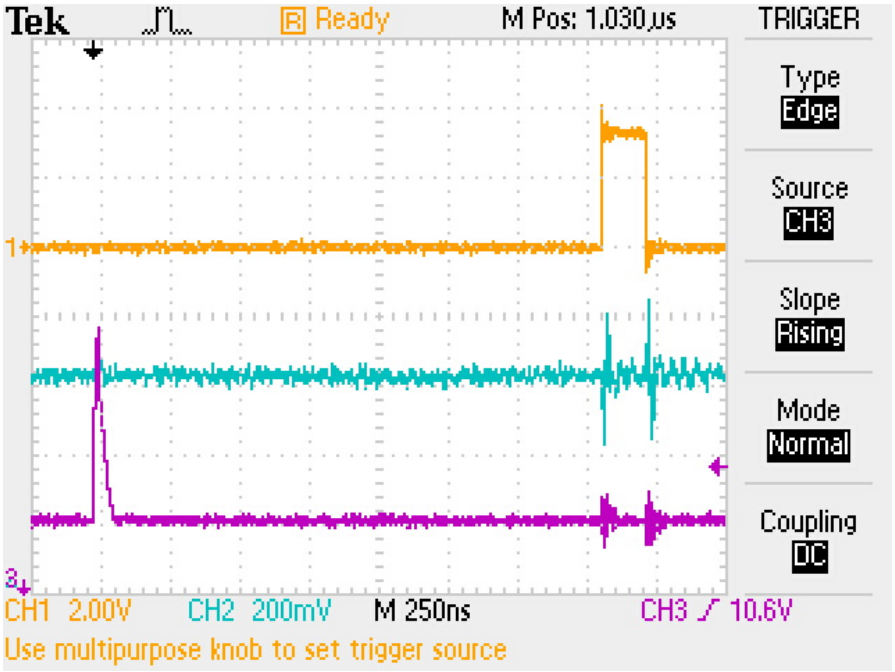


Figure 7.10 – CLICpix demonstrator tested with only zeros for the pixel values. The magenta signal corresponds to the READY signal sent by CLICpix to the FPGA. The blue signal corresponds to the data sent from the chip. The yellow signal corresponds to the end of the frame.

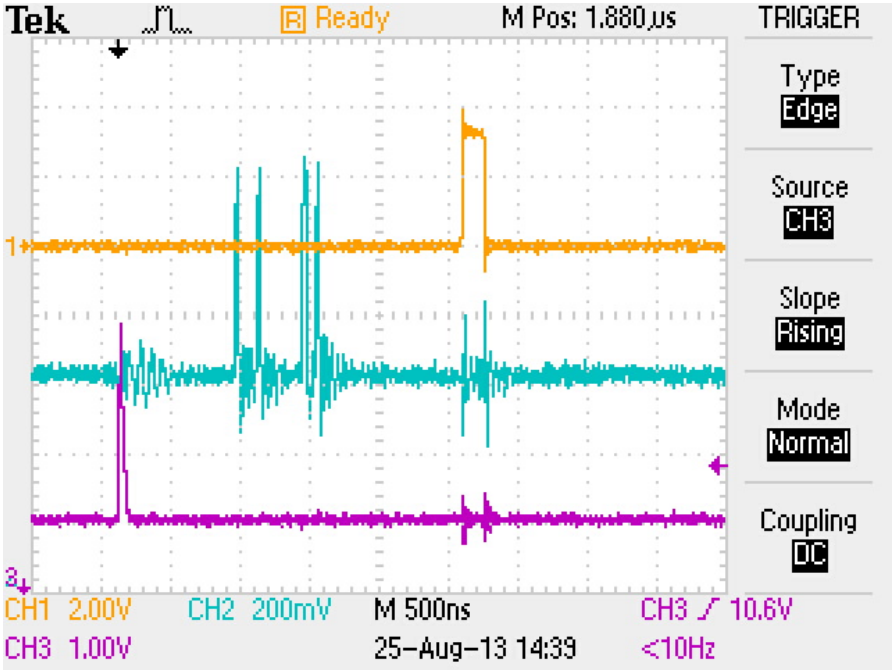


Figure 7.11 – CLICpix demonstrator tested with only one pixel having a non-zero value (with coordinates x=0 and y=0).

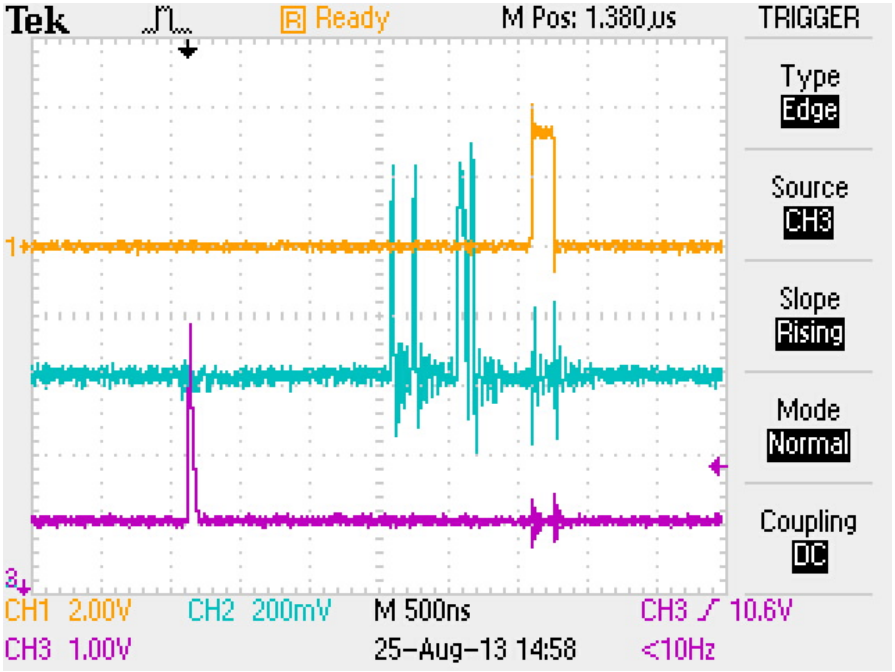


Figure 7.12 – CLICpix demonstrator tested with the only one pixel having a non-zero value (the last one read with coordinates x=62 and y=0).

8 Conclusion

The vertex detector for CLIC is very different from the vertex detector existing in today's particle detectors. The experimental conditions at the interaction region are very challenging for existing technologies at CLIC.

In the first part of this thesis, we focus on improving the geometry of the CLIC vertex detector model established in [29]. This study is part of a project which will define a complete new CLIC detector model fulfilling the physics requirements by the end of 2014.

For this study, first we implemented two new concepts: the spiral endcaps and the double layer geometries. The spirals in the endcaps allow to use airflow cooling which can reduce the material budget significantly. The double layer geometry, not only allows this cooling technique but also with the same amount of material budget as the default detector, provides more sensitive material for more precise measurements. These two geometries have similar impact parameter resolutions as the default geometry.

The flavor tagging is performed for simulated dijet events using the newly implemented and the default geometries. The b-tag or c-tag efficiency versus the fake rate for the identification of other jets as b or c jets is compared for each geometry. The overall results show that all the geometries are very similar in terms of the flavor-tag performance. While studying the endcap region, we observed that for dijets with a polar angle θ of around 40°, the flavor tagging degrades for the spiral and the double layer geometries compared to default geometry. For this polar angle, the number of sensors for the spiral geometries is very dependent on the azimuthal angle ϕ. This effect is worse for double-sided sensors and might be partially compensated using different optimizations of the tracking depending on the ϕ-angle.

The double-layered sensors provide more measurement points for charged particle tracks with the same material budget compared to the geometry using single-sided sensors.

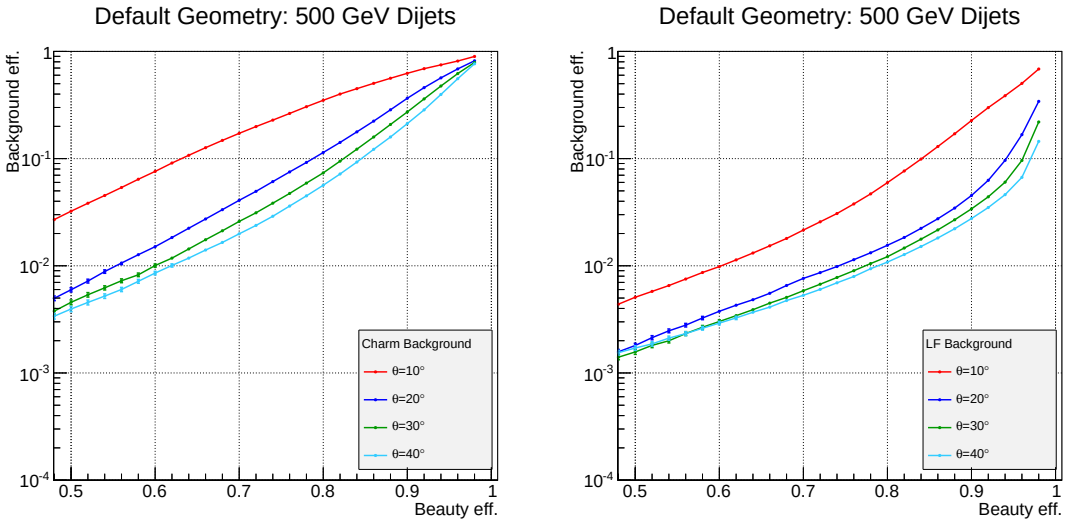
With these results we can conclude that the spiral endcap and double layer geometries for the vertex endcaps can be considered as solution for the improvement of the CLIC detector.

In the second part of this thesis, we worked on a decoder implementation for the CLICpix chip. The simulation results show that it is possible to decode the chip data in real time and decrease the time needed to read a frame. Further studies will be done for optimizing the data coding.

A Flavor-Tag Performance for All Energies and Jet Angles

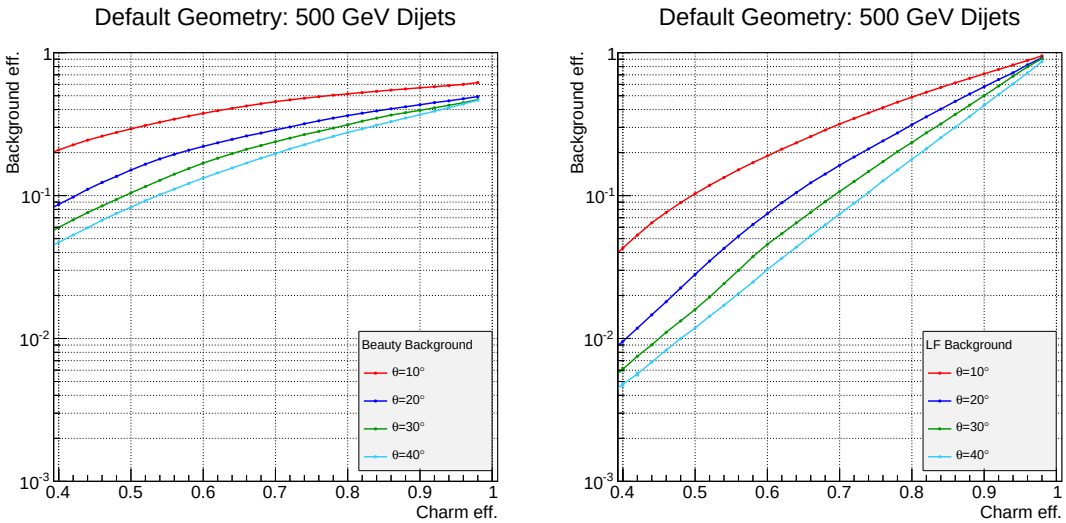
A.1 Dijets at 500 GeV

A.1.1 Default Geometry



(a) (b)

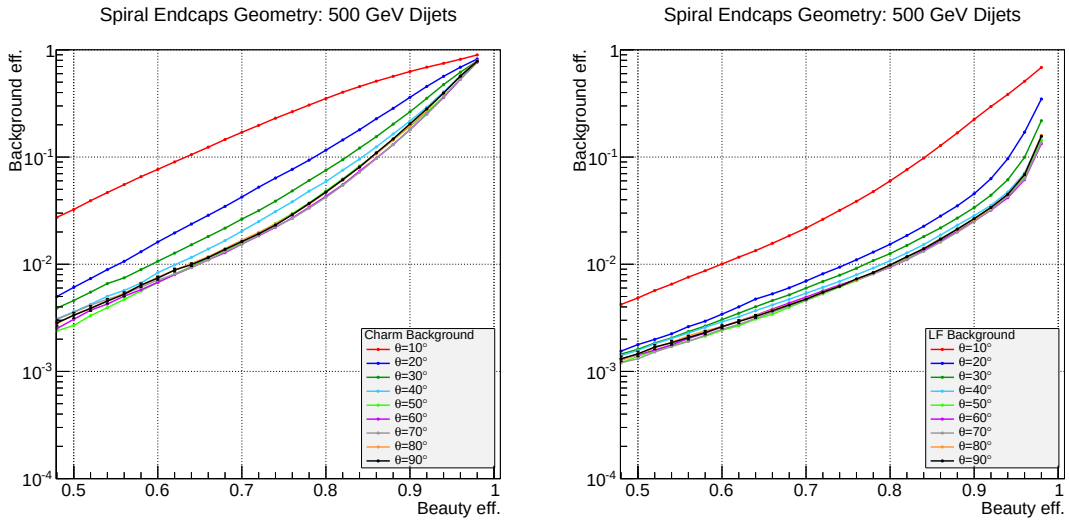
Figure A.1 – b-tag efficiency for dijets at 500 GeV using the default geometry.



(a) (b)

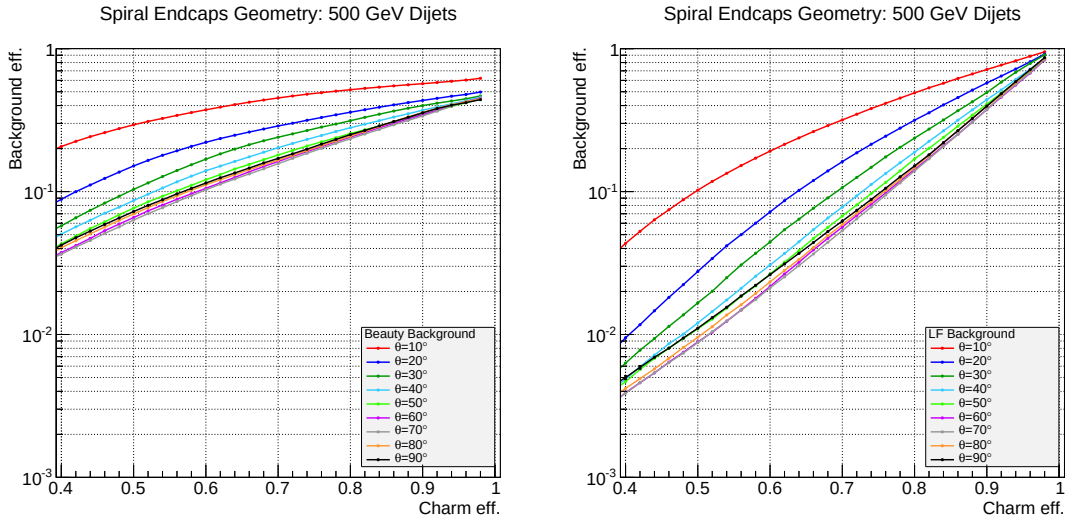
Figure A.2 – c-tag efficiency for dijets at 500 GeV using the default geometry.

A.1.2 Spiral Endcaps Geometry



(a) (b)

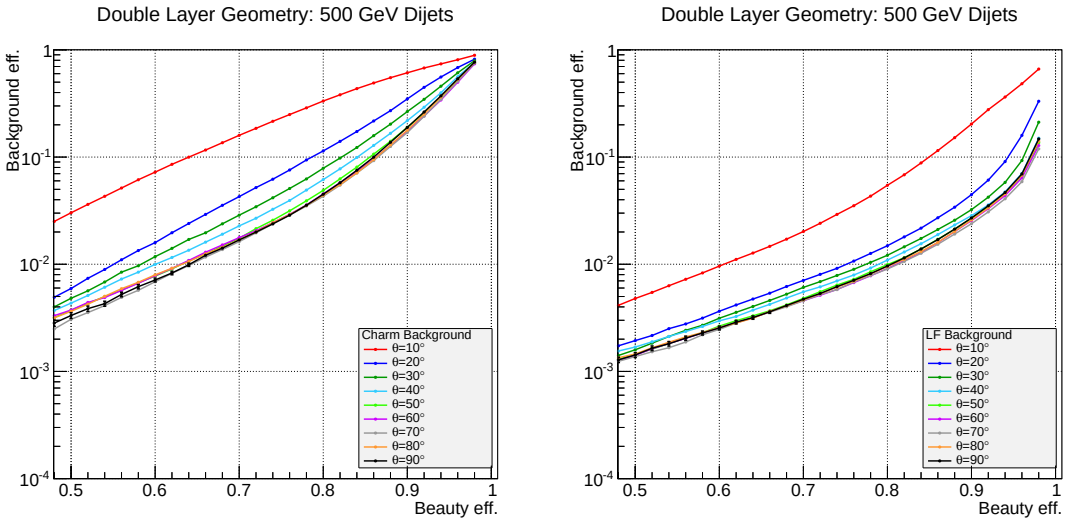
Figure A.3 – b-tag efficiency for dijets at 500 GeV using the spiral endcaps geometry.



(a) (b)

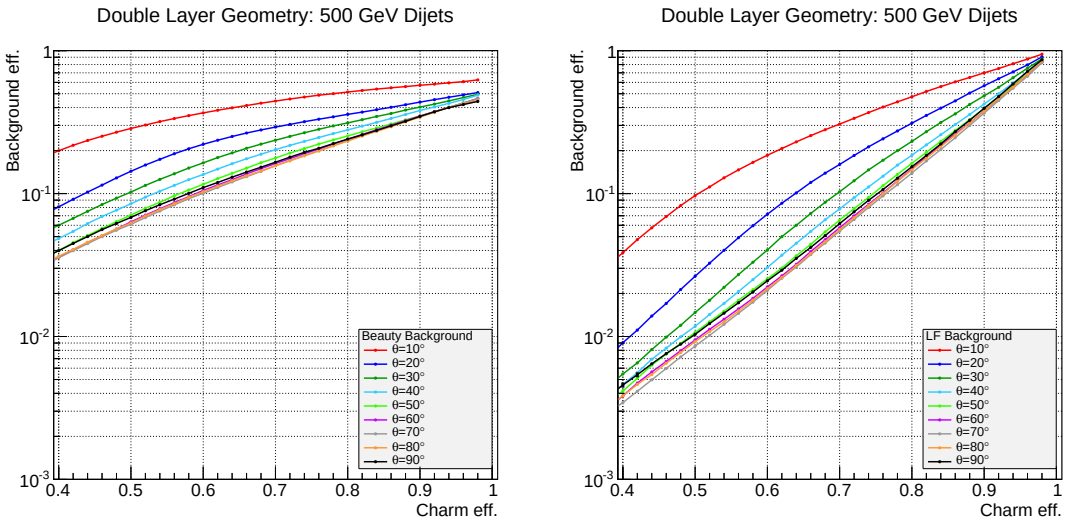
Figure A.4 – c-tag efficiency for dijets at 500 GeV using the spiral endcaps geometry.

A.1.3 Double Layer Geometry



(a) (b)

Figure A.5 – b-tag efficiency for dijets at 500 GeV using the double layer geometry.

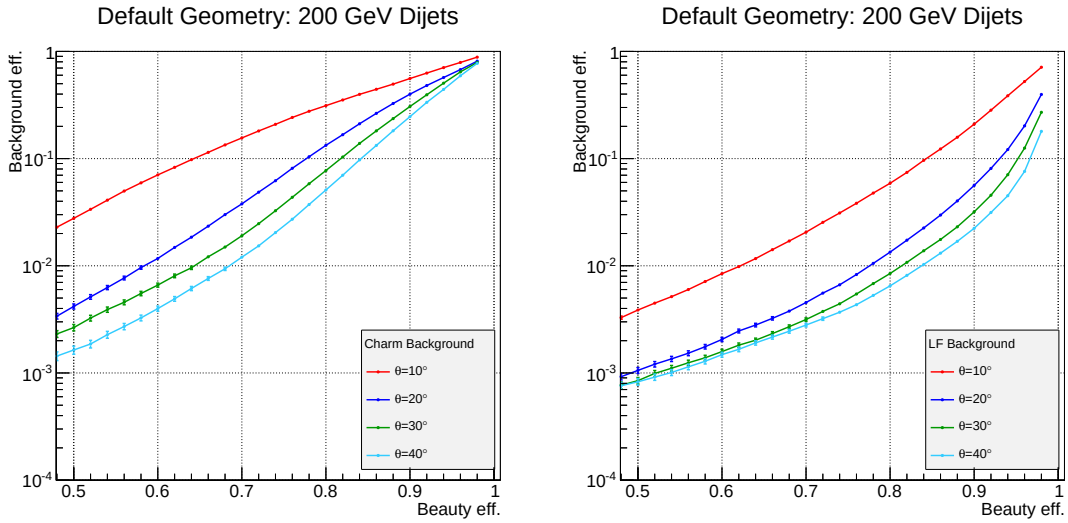


(a) (b)

Figure A.6 – c-tag efficiency for dijets at 500 GeV using the double layer geometry.

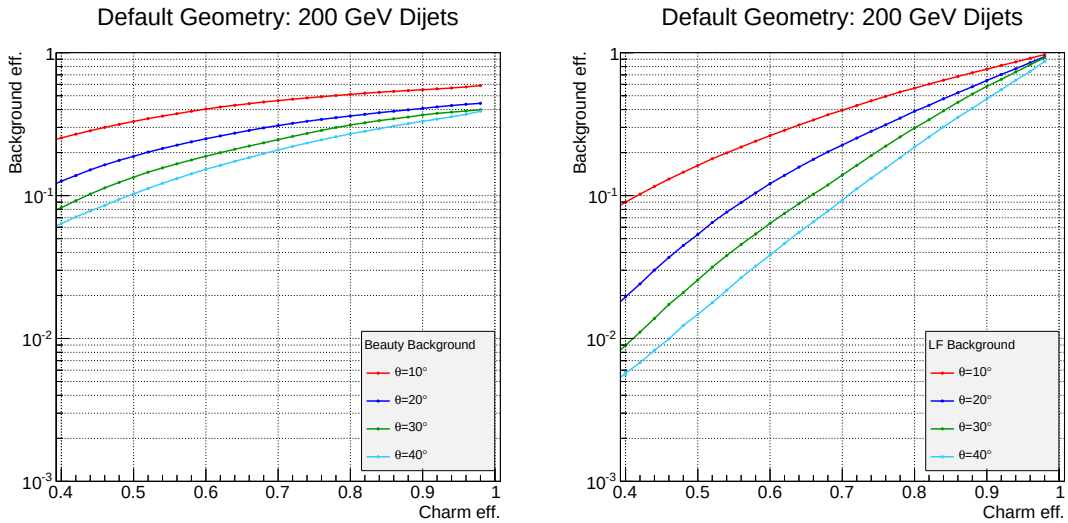
A.2 Dijets at 200 GeV

A.2.1 Default Geometry



(a) (b)

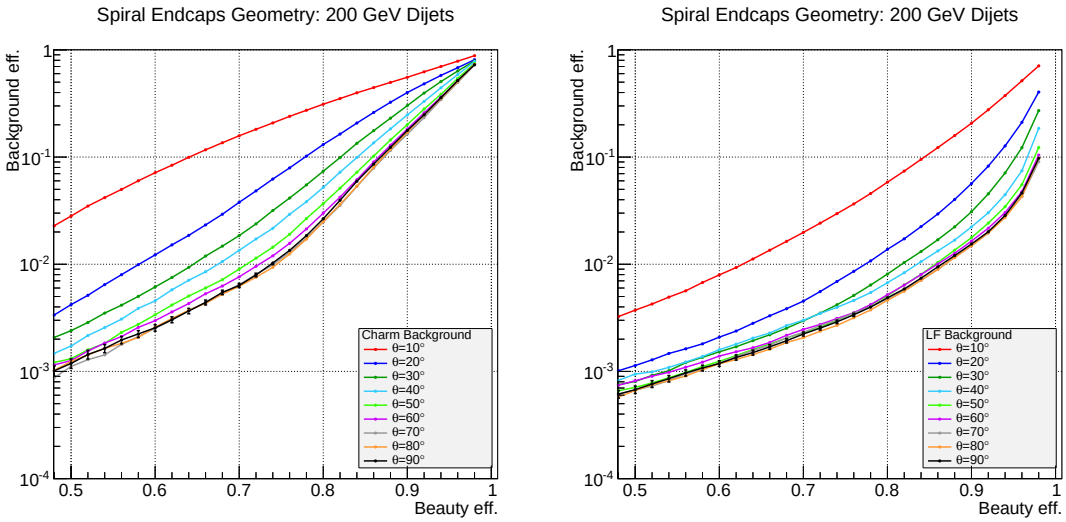
Figure A.7 – b-tag efficiency for dijets at 200 GeV using the default geometry.



(a) (b)

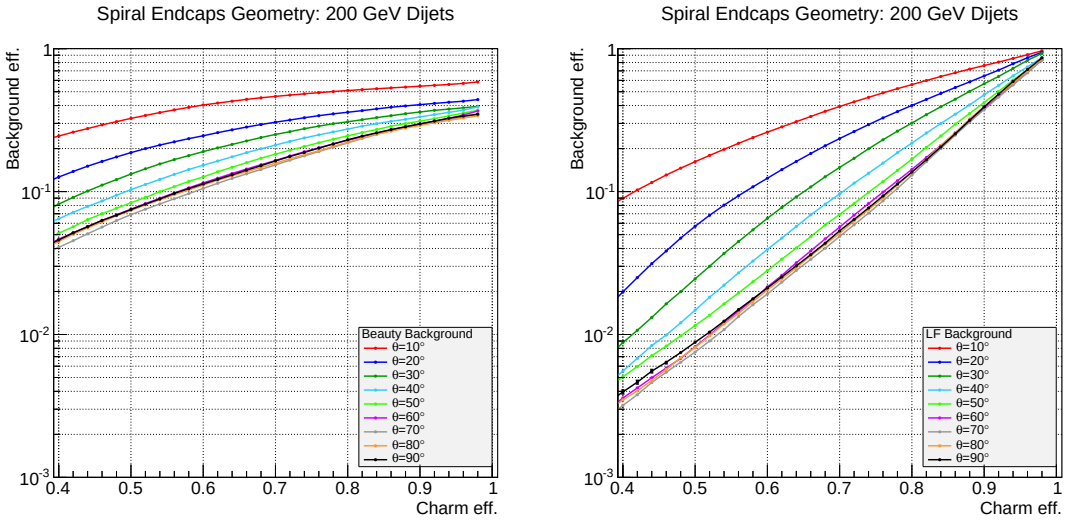
Figure A.8 – c-tag efficiency for dijets at 200 GeV using the default geometry.

A.2.2 Spiral Endcaps Geometry



(a) (b)

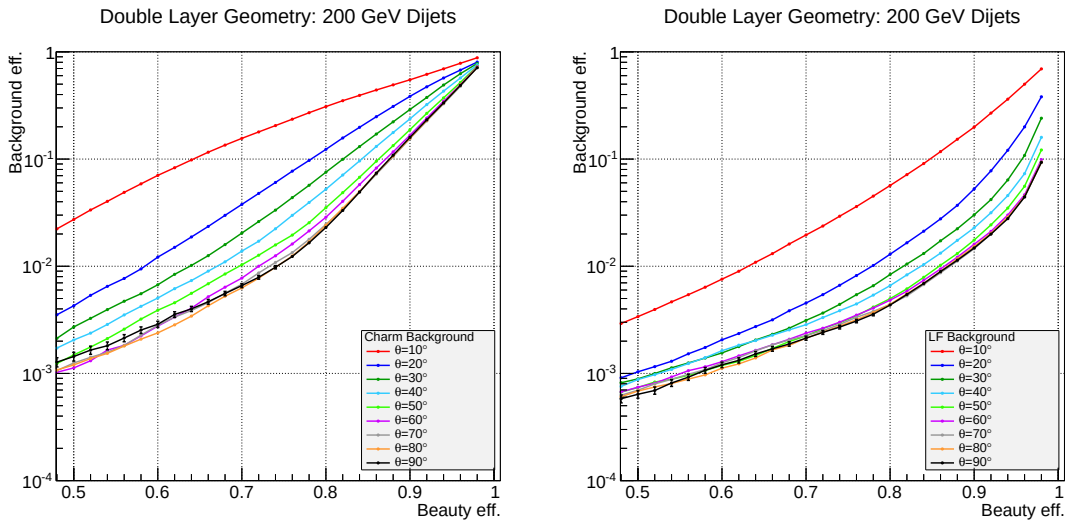
Figure A.9 – b-tag efficiency for dijets at 200 GeV using the spiral endcaps geometry.



(a) (b)

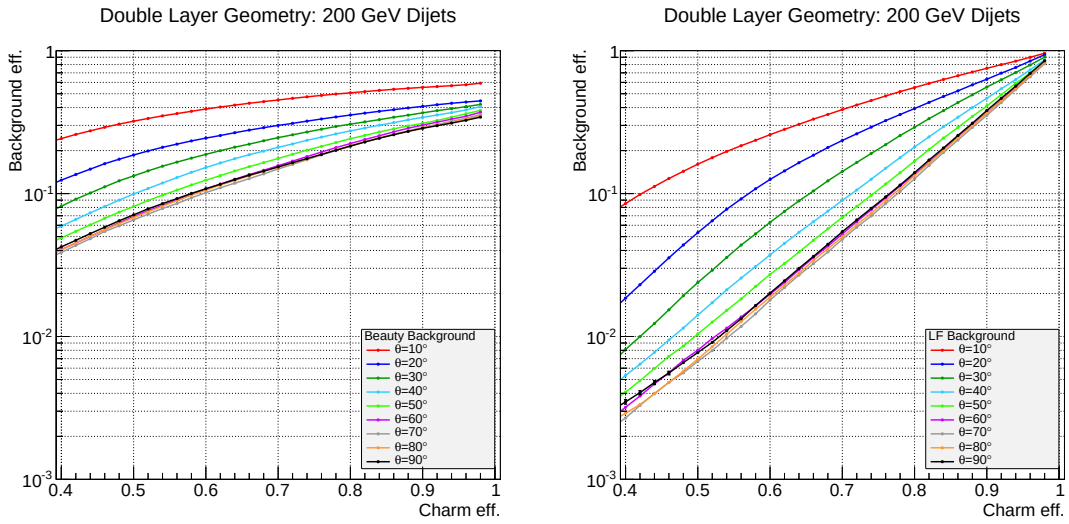
Figure A.10 – c-tag efficiency for dijets at 200 GeV using the spiral endcaps geometry.

A.2.3 Double Layer Geometry



(a) (b)

Figure A.11 – b-tag efficiency for dijets at 200 GeV using the double layer geometry.

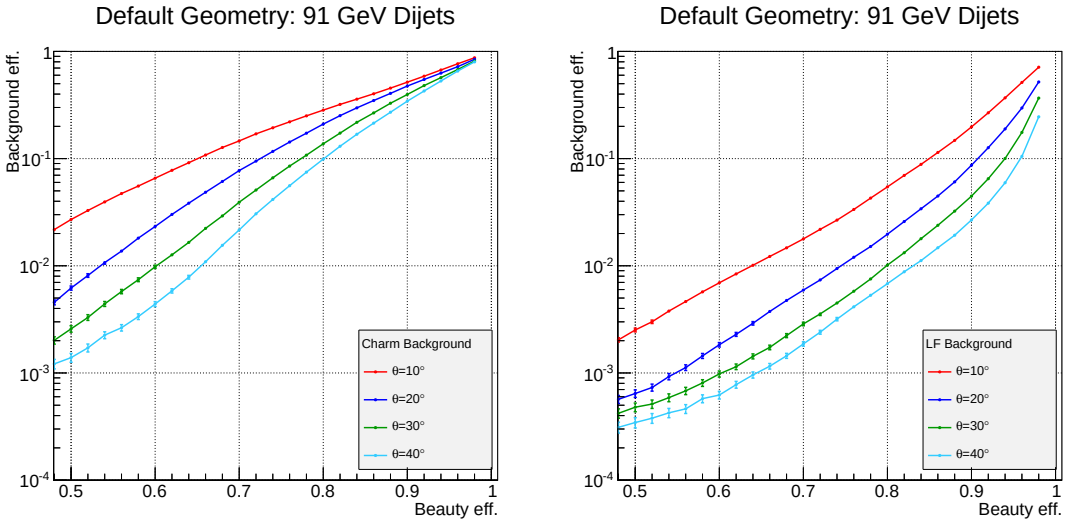


(a) (b)

Figure A.12 – c-tag efficiency for dijets at 200 GeV using the double layer geometry.

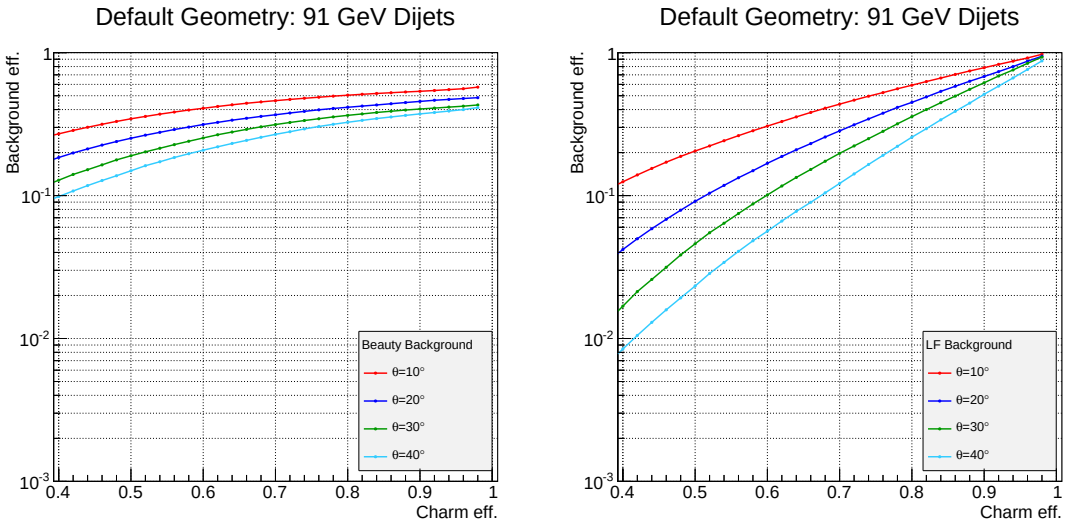
A.3 Dijets at 91 GeV

A.3.1 Default Geometry



(a) (b)

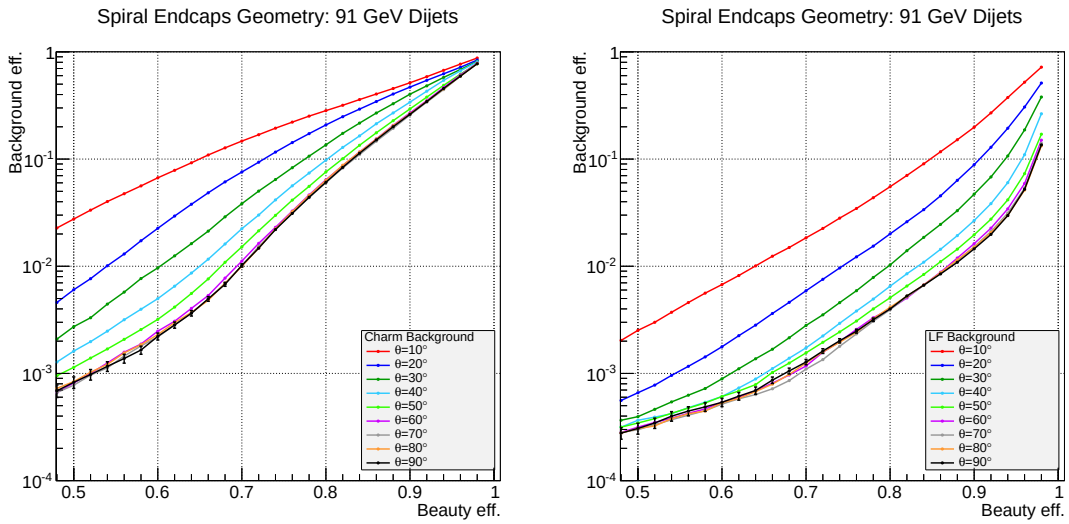
Figure A.13 – b-tag efficiency for dijets at 91 GeV using the default geometry.



(a) (b)

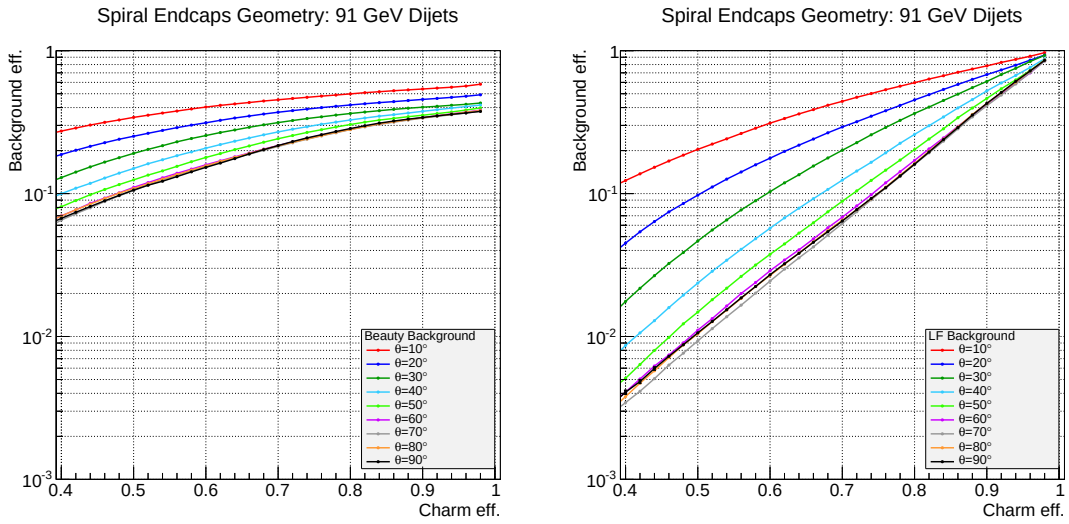
Figure A.14 – c-tag efficiency for dijets at 91 GeV using the default geometry.

A.3.2 Spiral Endcaps Geometry



(a) (b)

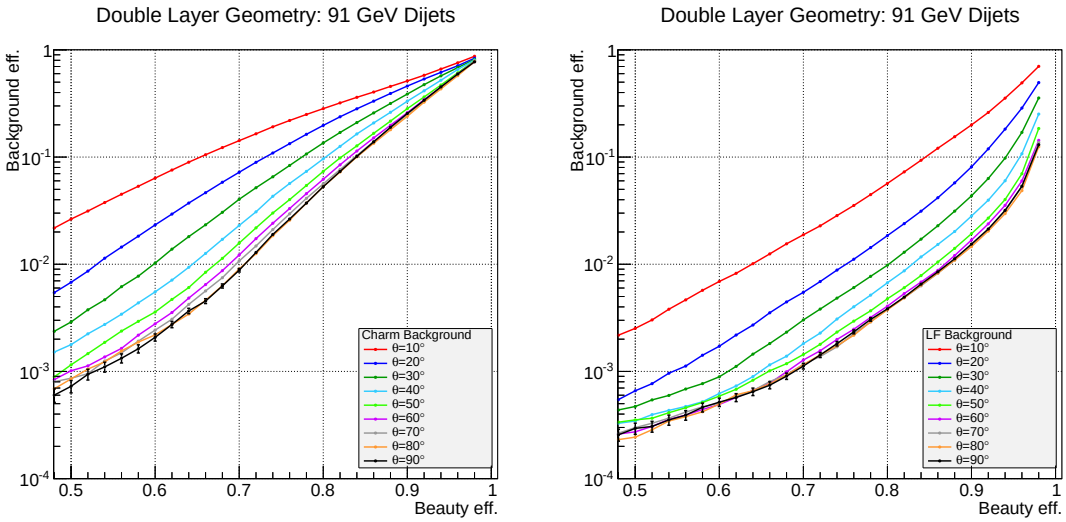
Figure A.15 – b-tag efficiency for dijets at 91 GeV using the spiral endcaps geometry.



(a) (b)

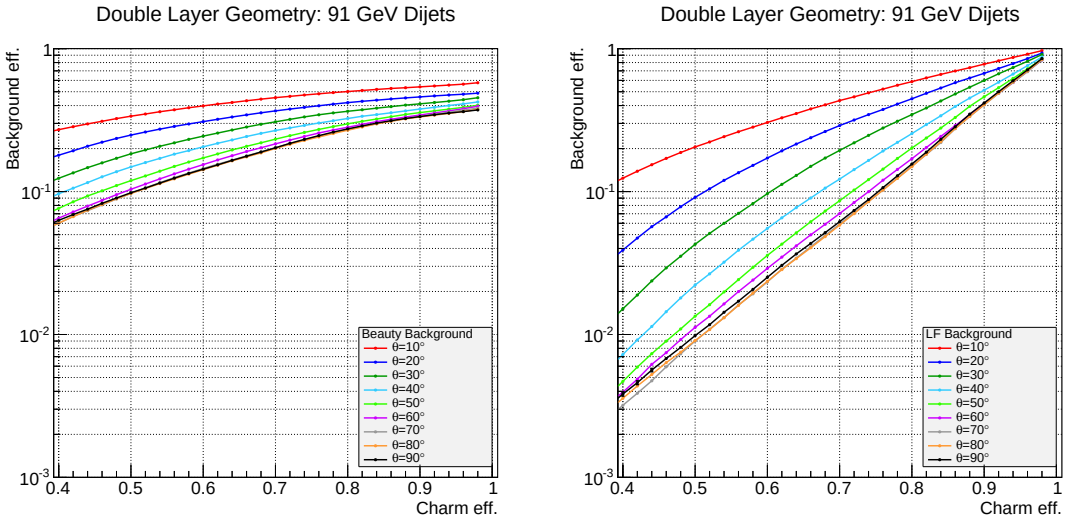
Figure A.16 – c-tag efficiency for dijets at 91 GeV using the spiral endcaps geometry.

A.3.3 Double Layer Geometry



(a) (b)

Figure A.17 – b-tag efficiency for dijets at 91 GeV using the double layer geometry.

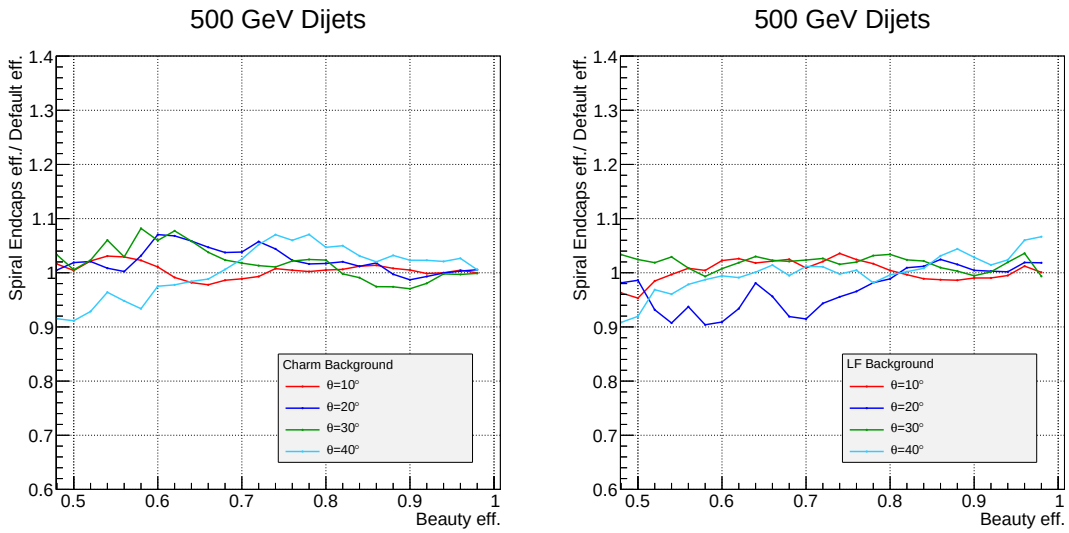


(a) (b)

Figure A.18 – c-tag efficiency for dijets at 91 GeV using the double layer geometry.

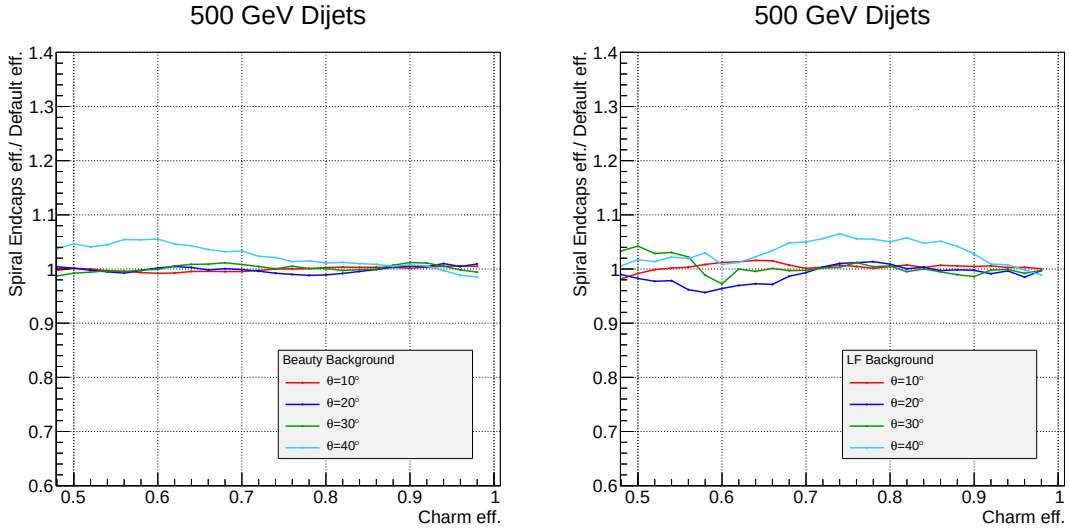
A.4 Background Efficiency Ratios for Different Geometries

A.4.1 Dijets at 500 GeV: Spiral Endcaps vs. Default Geometry



(a) (b)

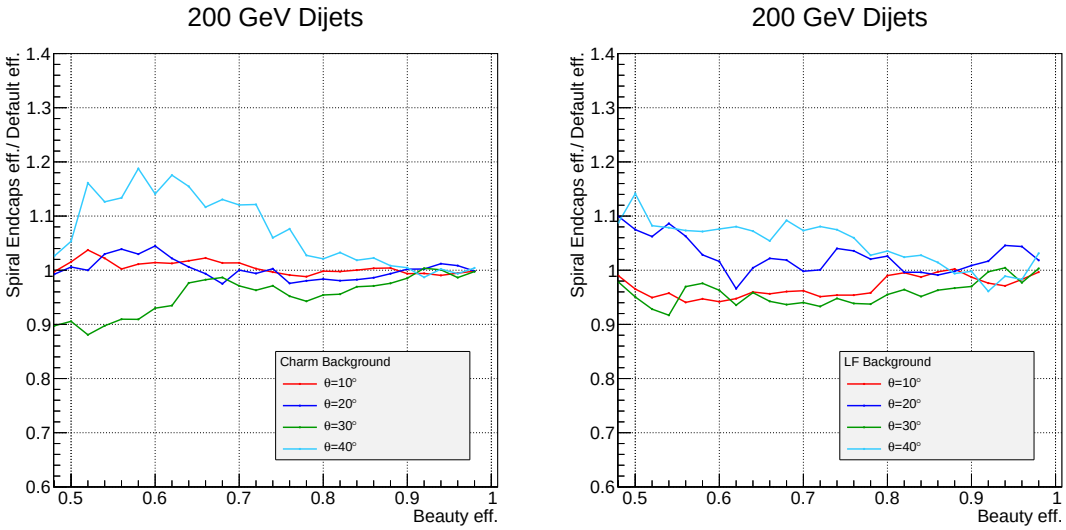
Figure A.19 – b-tag efficiency comparing the spiral endcaps and the default geometry in the forward region.



(a) (b)

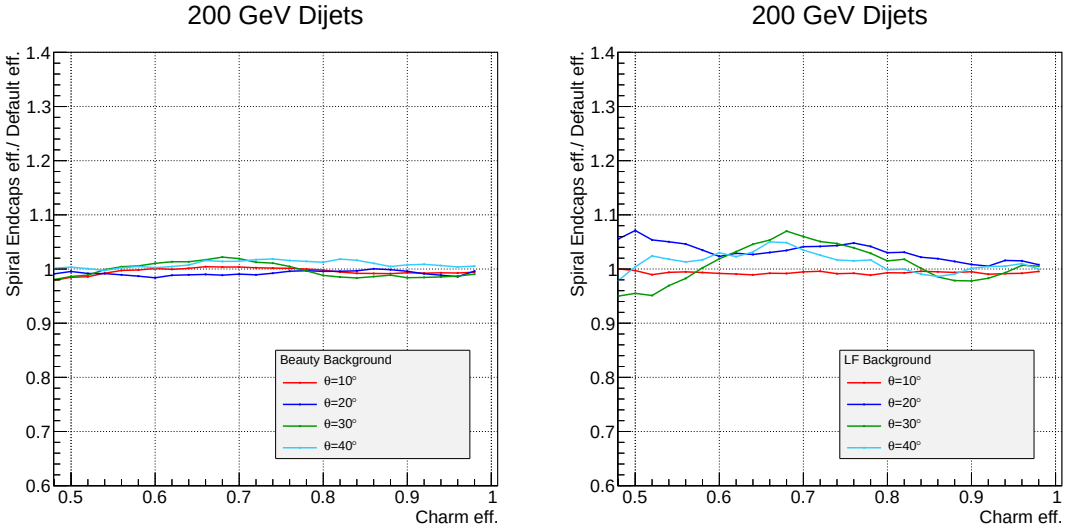
Figure A.20 – c-tag efficiency comparing the spiral endcaps and the default geometry in the forward region.

A.4.2 Dijets at 200 GeV: Spiral Endcaps vs. Default Geometry



(a) (b)

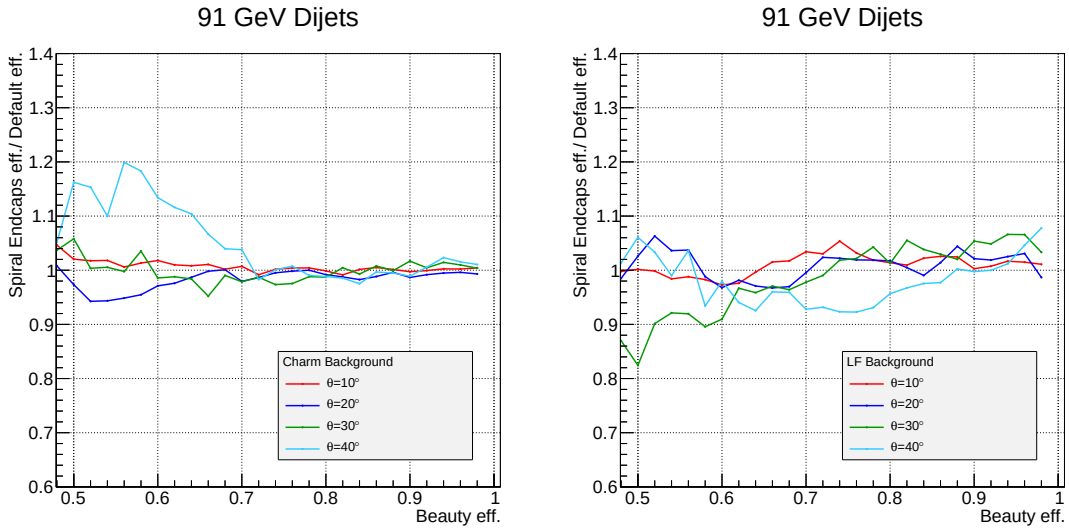
Figure A.21 – b-tag efficiency comparing the spiral endcaps and the default geometry in the forward region.



(a) (b)

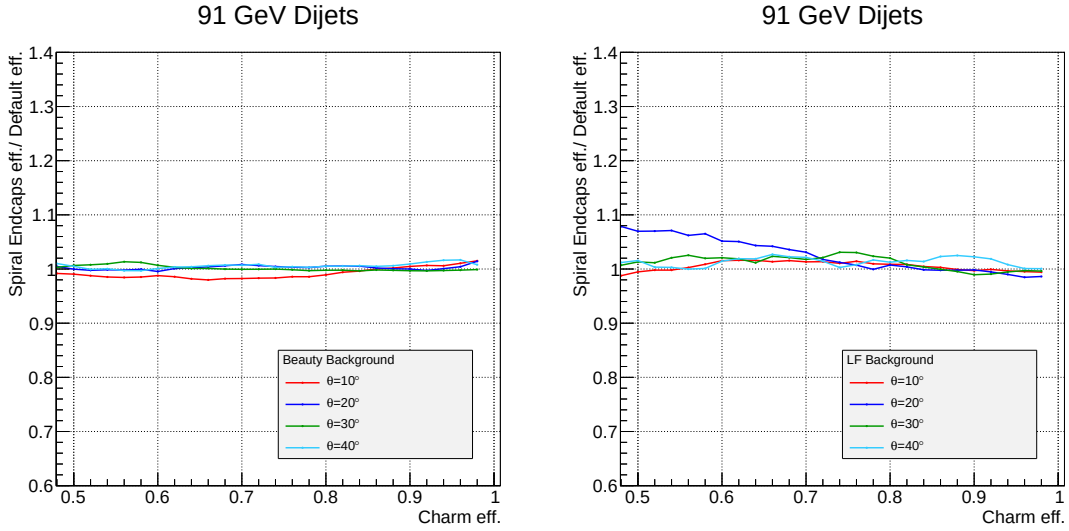
Figure A.22 – c-tag efficiency comparing the spiral endcaps and the default geometry in the forward region.

A.4.3 Dijets at 91 GeV: Spiral Endcaps vs. Default Geometry



(a) (b)

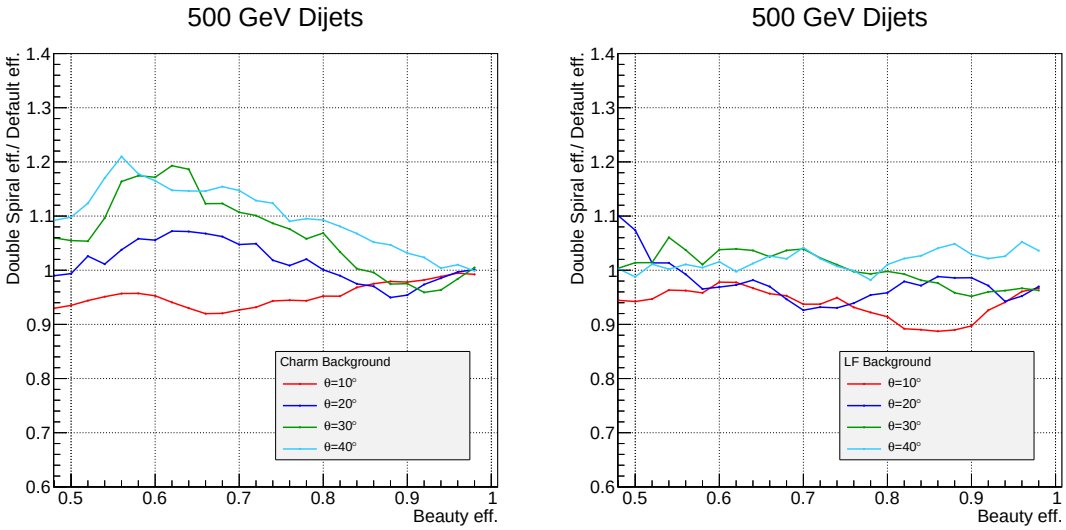
Figure A.23 – b-tag efficiency comparing the spiral endcaps and the default geometry in the forward region.



(a) (b)

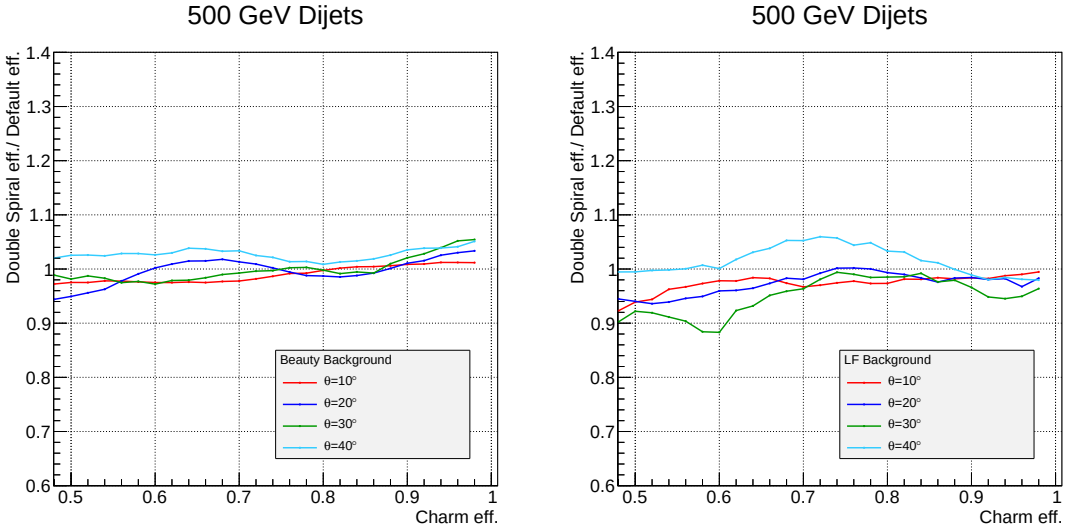
Figure A.24 – c-tag efficiency comparing the spiral endcaps and the default geometry in the forward region.

A.4.4 Dijets at 500 GeV: Double Layer vs. Default Geometry



(a) (b)

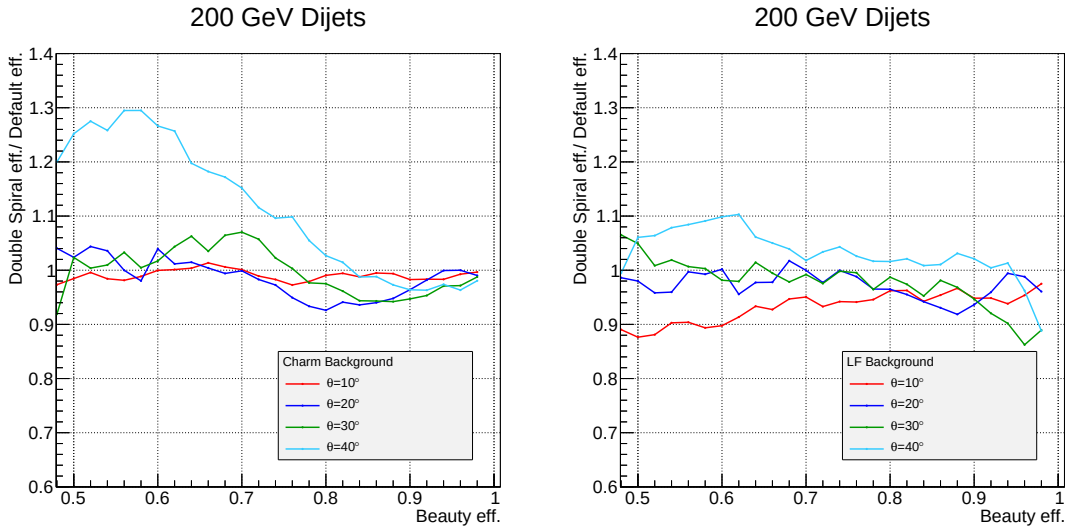
Figure A.25 – b-tag efficiency comparing the double layer and the default geometry in the forward region.



(a) (b)

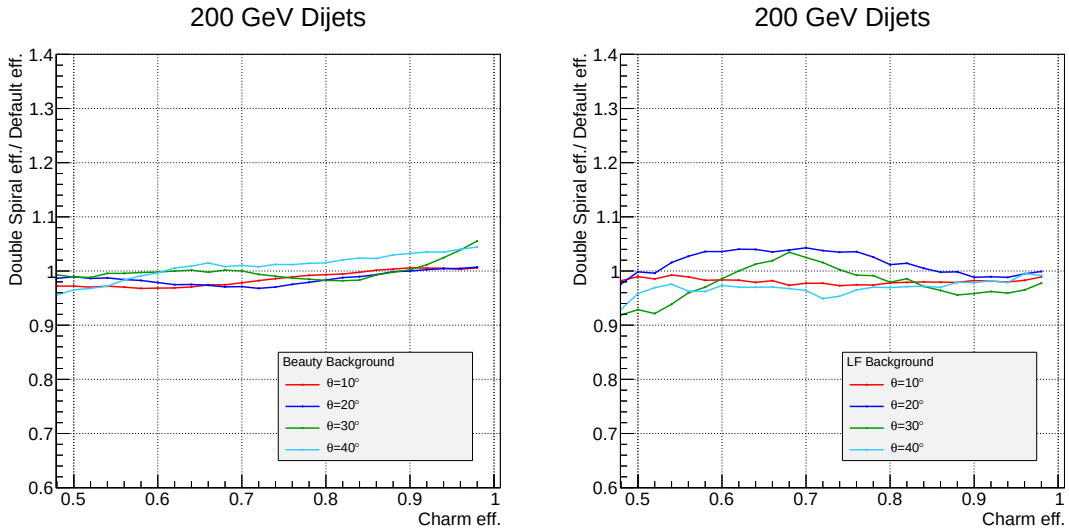
Figure A.26 – c-tag efficiency comparing the double layer and the default geometry in the forward region.

A.4.5 Dijets at 200 GeV: Double Layer vs. Default Geometry



(a) (b)

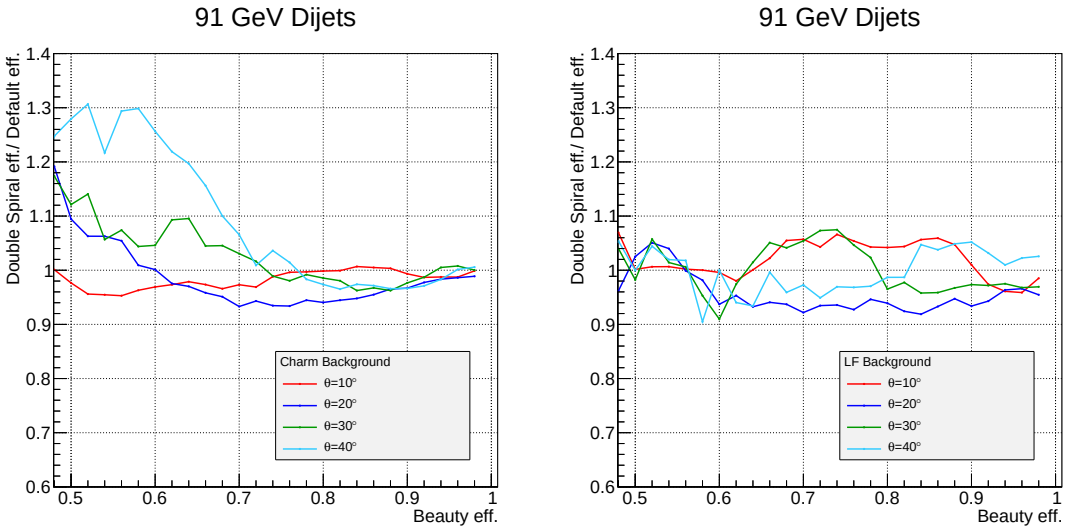
Figure A.27 – b-tag efficiency comparing the double layer and the default geometry in the forward region.



(a) (b)

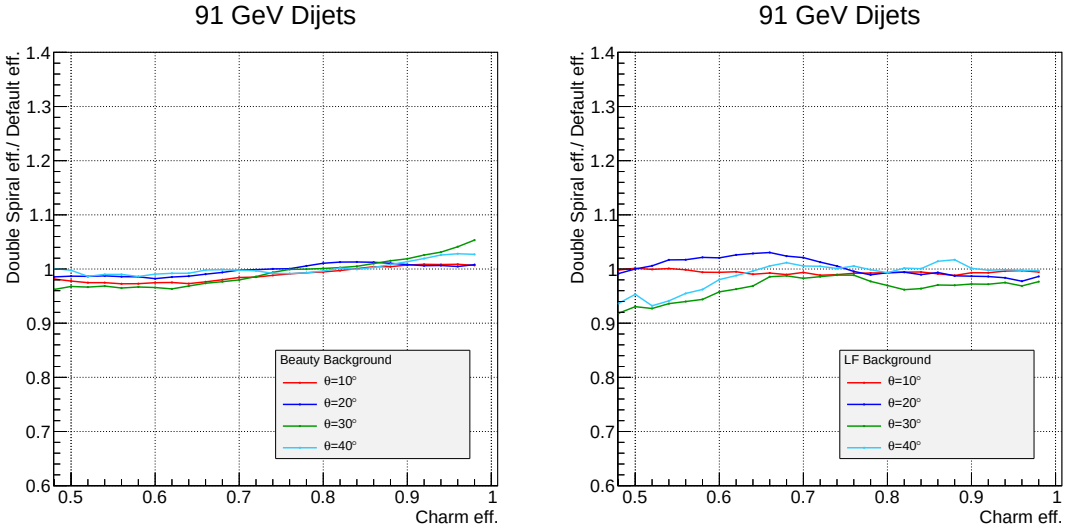
Figure A.28 – c-tag efficiency comparing the double layer and the default geometry in the forward region.

A.4.6 Dijets at 91 GeV: Double Layer vs. Default Geometry



(a) (b)

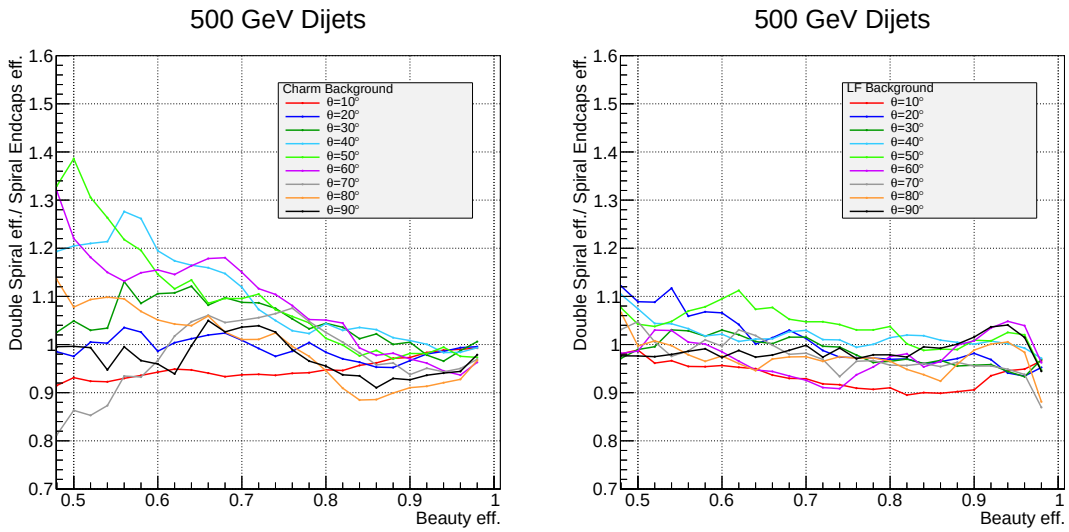
Figure A.29 – b-tag efficiency comparing the double layer and the default geometry in the forward region.



(a) (b)

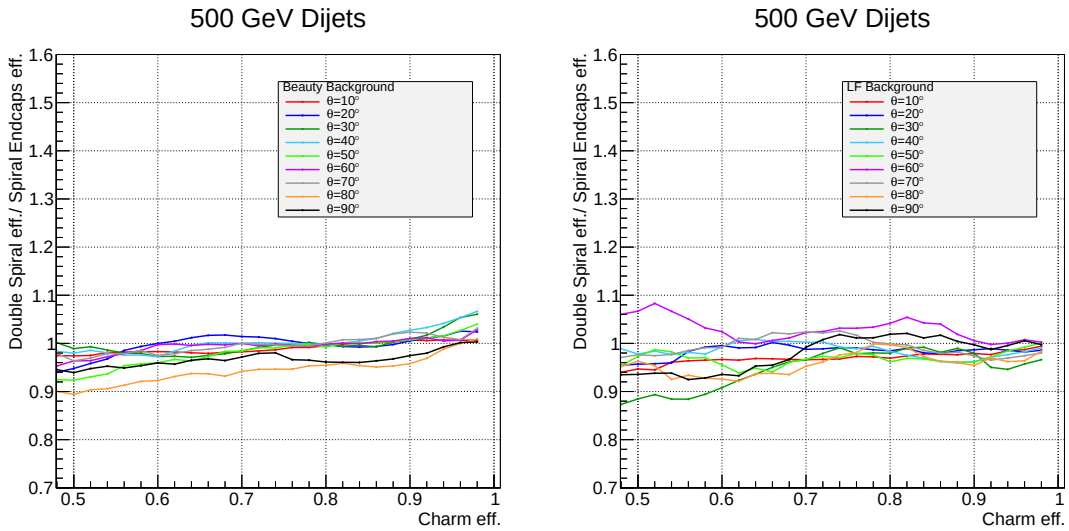
Figure A.30 – c-tag efficiency comparing the double layer and the default geometry in the forward region.

A.4.7 Dijets at 500 GeV: Double Layer vs. Spiral Endcaps Geometry



(a) (b)

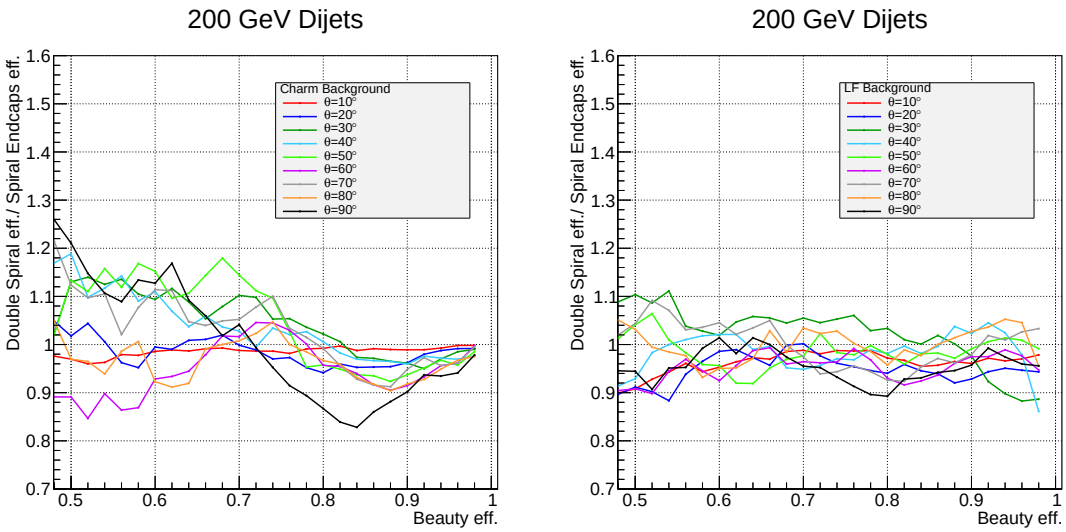
Figure A.31 – b-tag efficiency comparing the double layer and the spiral endcaps geometry.



(a) (b)

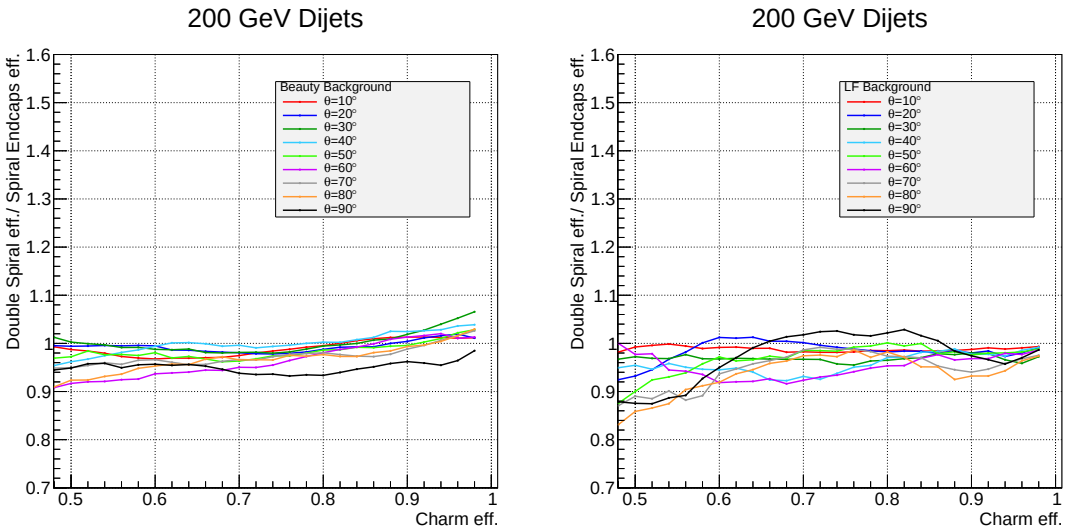
Figure A.32 – c-tag efficiency comparing the double layer and the spiral endcaps geometry.

A.4.8 Dijets at 200 GeV: Double Layer vs. Spiral Endcaps Geometry



(a) (b)

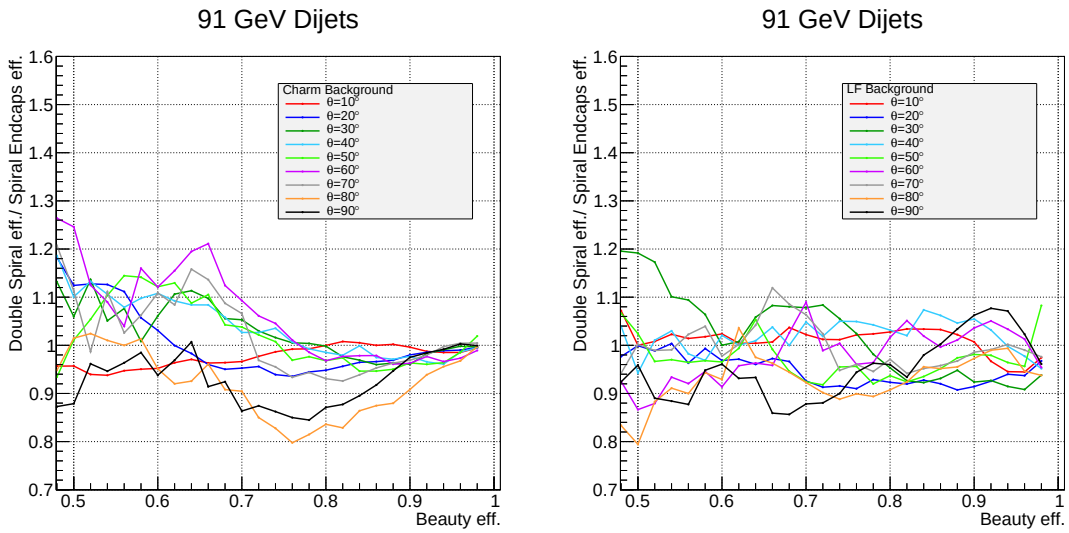
Figure A.33 – b-tag efficiency comparing the double layer and the spiral endcaps geometry.



(a) (b)

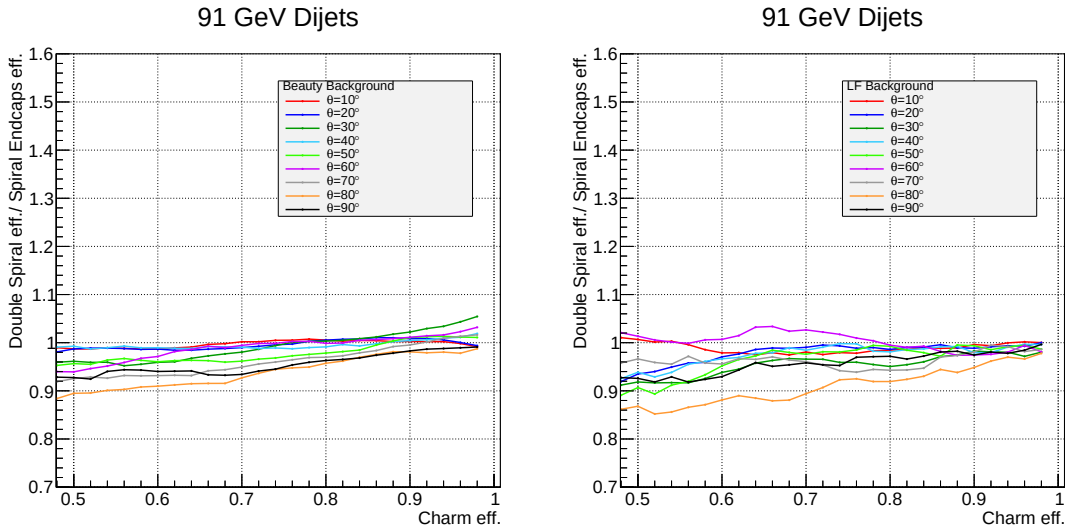
Figure A.34 – c-tag efficiency comparing the double layer and the spiral endcaps geometry.

A.4.9 Dijets at 91 GeV: Double Layer vs. Spiral Endcaps Geometry



(a) (b)

Figure A.35 – b-tag efficiency comparing the double layer and the spiral endcaps geometry.



(a) (b)

Figure A.36 – c-tag efficiency comparing the double layer and the spiral endcaps geometry.

Bibliography

[1] DIRAC. http://diracgrid.org/.

[2] GeomConverter. Website: http://lcsim.org/software/GeomConverter/.

[3] Linear Collider Detector Descriptions. Website: http://www.lcsim.org/detectors/.

[4] Linear Collider Simulations (LCSim). Website: http://lcsim.org/.

[5] Marlin, Modular Analysis & Reconstruction for the LINear collider. Website: http://ilcsoft.desy.de/portal/software_packages/marlin/.

[6] Simulator for the Linear Collider (SLIC). Website: http://lcsim.org/software/slic/.

[7] Standard Model. Website: http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[8] TMVA. http://tmva.sourceforge.net/.

[9] T. Abe et al. The International Large Detector: Letter of Intent. arXiv: 1006.3396 [hep-ex], 2010.

[10] S. Agostinelli et al. Geant4 – A Simulation Toolkit. *Nucl. Instrum. Methods Phys. Res., Sect. A*, 506(3):250–303, July 2003.

[11] H. Aihara, P. Burrows, M. Oreglia, E.L. Berger, V. Guarino, et al. SiD Letter of Intent. arXiv: 0911.0006 [physics.ins-det], 2009.

[12] J. Allison et al. Geant4 developments and applications. *IEEE T. Nucl. Sci.*, 53(1):270–278, Feb. 2006.

[13] J. Balbuena and G. Pellegrini. *Development of innovative silicon radiation detectors.* PhD thesis, Barcelona, Autonoma U., 2012. Presented 2012.

[14] K. Binder and D.W. Heermann. *Monte Carlo Simulation in Statistical Physics: An Introduction.* Springer Verlag, 2002.

[15] J. Brau, Y. Okada, N.J. Walker, A. Djouadi, J. Lykken, et al., editors. *International Linear Collider reference design report: ILC Global Design Effort and World Wide Study.* ILC-REPORT-2007-001. 2007.

[16] R. Brun and F. Rademakers. ROOT – an object oriented data analysis framework. *Nucl. Instrum. Methods Phys. Res., Sect. A*, 389(1-2):81–86, April 1997.

[17] D. Dannheim. Status for pixel detector R&D for future linear colliders. In *The 2013 European Physical Society Conference on High Energy Physics: EPSHEP 2013 Stockholm*, 2013.

[18] F. Duarte Ramos, H. Gerwig, and M. Villarejo Bermudez. CLIC inner detectors cooling simulations. 2013-08-13, Aug 2013.

[19] D. Bailey et al. The lcfivertex package: Vertexing, flavour tagging and vertex charge reconstruction with an ilc vertex detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 610(2):573 – 589, 2009.

[20] L. Linssen et al. CLIC e^+e^- Linear Collider Studies: Input to the update process of the European Strategy for Particle Physics. *arXiv:1208.1402 [physics.acc-ph]*, 2012.

[21] S. Agostinelli et al. Geant4-a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.

[22] F. Gaede. Lcio. http://lcio.desy.de/, August 2012.

[23] C. Grefe and A. Münnich. The CLIC_SiD_CDR Geometry for the CDR Monte Carlo Mass Production. LCD-Note-2011-009, 2011.

[24] D. Griffiths. *Introduction to Elementary Particles*. WILEY-VCH Verlag GmbH & Co. KGaA, second edition, 2008.

[25] G.F. Knoll. *Radiation detection and measurement*. Wiley, 2000.

[26] T. Lastovicka. Light Higgs Production and Decays to Pairs of Bottom and Charm Quarks at 3 TeV. LCD-Note-2011-036, Sep 2012.

[27] P. Lebrun, L. Linssen, A. Lucaci-Timoce, D. Schulte, F. Simon, S. Stapnes, N. Toge, H. Weerts, and J. D. Wells, editors. *The CLIC Programme: towards a staged e^+e^- Linear Collider exploring the Terascale*. CERN, 2012. ANL-HEP-TR-12-51, KEK Report 2012-2, MPP-2012-115, https://edms.cern.ch/document/1234246/.

[28] W.R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach*. Springer-Verlag, 2nd edition, 1994.

[29] L. Linssen, A. Miyamoto, M. Stanitzki, and H. Weerts, editors. *Physics and Detectors at CLIC: CLIC Conceptual Design Report*. CERN, 2012. ANL-HEP-TR-12-01, CERN-2012-003, DESY 12-008, KEK Report 2011-7, arXiv:1202.5940.

[30] F. Duarte Ramos. CLIC inner detector cooling and mechanics. *LCWS12 International Workshop on Future Linear Colliders*, 2012.

[31] P. Roloff. Flavour tagging results with increased material budget. http://indico.cern.ch/conferenceDisplay.py?confId=241270, March 2013.

[32] A. Sailer. Lcd seminar for summer interns. http://indico.cern.ch/conferenceDisplay.py?confId=264776, July 2013.

[33] G. Salam. Jets at hadron colliders. https://indico.cern.ch/conferenceDisplay.py?confId=115078, March 2011.

[34] T. Tanabe and T. Suehara. LCFIPlus. https://confluence.slac.stanford.edu/display/ilc/LCFIPlus, June 2013.

[35] P. Valerio, R. Ballabriga, and M. Campbell. Design of the 65 nm CLICpix demonstrator chip. LCD-Note-2012-018, 2012.

[36] E. van der Kraaij. *First Top Quark Physics with ATLAS – A Prospect.* PhD thesis, National instituut voor subatomaire fysica (Nikhef), Amsterdam, 2009.

[37] H. Yang, B.P. Roe, and J. Zhu. Studies of Boosted Decision Trees for MiniBooNE Particle Identification. *arXiv:physics/0508045v1 [physics.data-an]*, 2005.

Acknowledgements

I am obliged to pay my tribute to Prof. Pascal Frossard for his acceptance, guidance and support of the project.
I am pleased to present my gratitude to Lucie Linssen, first of all for accepting me in her very dynamic and creative group and second of all, for her support and guidance to realize and enrich this project.
I also acknowledge with thanks Philipp Roloff's excellent supervision of the project associated with scientific elaborations and nitty-gritties of the requisite basic knowledge and providing me a rich learning process through his praise worthy patience, precision and diligence.
Besides, I present my thanks and appreciation to Szymon Kulis and Pierpaolo Valerio for their sincere collaborations and sympathies throughout the implementation of the project.

Geneva, August 29, 2013 Niloufar Alipour Tehrani