Jet Energy Measurements at ILC: Calorimeter DAQ Requirements and Application in Higgs Boson Mass Measurements

Dissertation

zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und
Naturwissenschafften
Fachbereich Physik
der Universität Hamburg

vorgelegt von

Aliakbar Ebrahimi

aus Arsanjan, Iran

Hamburg
2017

Gutachterinen der Dissertation:	Prof. Dr. Erika Garutti Dr. Katja Krüger
Weitere Mitglieder der Prüfungskommission:	Prof. Dr. Daniela Pfannkuche Prof. Dr. Elisabetta Gallo Dr. Jenny List
Datum der Disputation:	01.11.2017

به پدر و مادر عزیزم ...

To my parents…

Abstract

The idea of spontaneous symmetry breaking as the mechanism through which elementary particles gain mass has been confirmed by the discovery of the Higgs boson at the CERN Large Hadron Collider. Studying the Higgs boson properties are of great importance to verify the Standard Model predictions. Any deviation from these predictions could uncover physics beyond the Standard Model. The mass of the Higgs boson is one of the important parameters of the Standard Model. The precise determination of the Higgs boson mass is of interest in its own right and also for other Higgs physics studies since it enters as parametric uncertainty into the extraction of the partial width from branching ratio measurements.

The International Linear Collider (ILC) is a future polarised e^+e^- collider designed for precision physics studies. The Higgs boson decay to a pair of bottom quarks $H \to b\bar{b}$ has the largest branching ratio of all Higgs decays, providing a large dataset for physics analyses. The possibility of measuring the Higgs boson mass in the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel is investigated in this thesis for centre-of-mass energies of 350 GeV and 500 GeV. Since the Higgs boson mass is reconstructed from two b jets, the jet energy resolution has a high impact on the measurement. A new method to estimate the jet energy resolution for each jet individually is developed in this thesis. The jet-specific energy resolution is then used in the analysis for the Higgs boson mass measurements. Various strategies for the Higgs boson mass measurement are investigated. For an integrated luminosity of 1000 fb^{-1} and a beam polarisation of $(-0.8,+0.3)$, statistical uncertainties of 42 MeV and 89 MeV are achieved for the centre-of-mass energies of 350 GeV and 500 GeV, respectively. Various sources of systematic uncertainties are also discussed.

These results are obtained using a full GEANT4-based simulation of the International Large Detector (ILD) concept. The jet energy resolution required for the Higgs boson mass measurement can only be achieved using the particle flow approach to reconstruction. The particle flow approach requires highly-granular calorimeters and a highly efficient tracking system. The CALICE collaboration is developing highly-granular calorimeters for such applications. One of the challenges in the development of such calorimeters with millions or read-out channels is their Data Acquisition System (DAQ) system. The second part of this thesis involves contributions to development of a new DAQ system for the CALICE scintillator calorimeters. The new DAQ system fulfils the requirements for the prototypes tests while being scalable to larger systems. The requirements and general architecture of the DAQ system is outlined in this thesis. The new DAQ system has been commissioned and tested with particle beams at the CERN Proton Synchrotron test beam facility in 2014, results of which are presented here.

Zusammenfassung

Die Idee der spontanen Symmetriebrechung, durch die Elementarteilchen Masse erwerben, wird durch die Entdeckung des Higgs-Bosons am CERN Large Hadron Collider bestätigt. Untersuchungen der Eigenschaften des Higgs-Bosons sind von großer Bedeutung, um die Vorhersagen des Standardmodells zu überprüfen. Jede Abweichung von diesen Vorhersagen könnte Physik jenseits des Standardmodells aufdecken. Die Masse des Higgs-Bosons ist einer der wichtigen Parameter des Standardmodells. Eine präzise Bestimmung der Higgs-Boson-Masse ist an sich und auch für andere Higgs-Physikstudien von Interesse da sie als parametrische Unsicherheit in die Extraktion der Partialbreite aus Messungen des Verzweigungsverhältnisses eingeht.

Der International Linear Collider (ILC) ist ein zukünftiger polarisierter e^+e^- Collider für Präzisionsphysik-Studien. Der Higgs-Boson-Zerfall in ein Bottom- und ein Anti-Bottom-Quark hat das größte Verzweigungsverhältnis aller Higgs-Zerfälle. Daeher bietet er einen großen Datensatz für Physikanalysen. In dieser Arbeit wird die Möglichkeit der Messung der Higgs-Boson-Masse am ILC im Kanal $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ für die Schwerpunktsenergien von 350 GeV und 500 GeV untersucht. Da die Higgs-Boson-Masse aus zwei b Jets rekonstruiert wird, hat die Jetenergieauflösung einen großen Einfluss auf diese Messung. Eine neue Methode zur Abschätzung der individuellen Jetenergieauflösung für jeden Teilchenjet wird in dieser Arbeit entwickelt. Die jetspezifische Energieauflösung wird dann in der Analyse zur Messung der Higgs-Boson-Masse verwendet. Es werden verschiedene Strategien für die Bestimmung der Masse mit kinematischen Fits untersucht. Für eine integrierte Luminosität von 1000 fb^{-1} und eine Strahlpolarisation von $(-0.8,+0.3)$ werden statistische Unsicherheiten von 42 MeV und 89 MeV für die Schwerpunktsenergien von 350 GeV bzw. 500 GeV erreicht. Verschiedene Quellen systematischer Unsicherheiten werden diskutiert.

Für diese Studie wird eine vollständige Geant4-Simulation das International Large Detector (ILD) Konzepts verwendet. Die für die Higgs-Boson-Massenmessung erforderliche Jetenergie-Auflösung kann nur mit dem *Particle Flow* Ansatz zur Ereignisrekonstruktion erreicht werden. Particle Flow Algorithmen erfordern Kalorimeter mit hoher Granularität und ein sehr effizientes Trackingsystem. Die CALICE-Kollaboration entwickelt Kalorimeter mit hoher Granularität für solche Anwendungen. Eine der Herausforderungen bei der Entwicklung solcher Kalorimeter mit mehreren Millionen Kanälen ist ihre Datenauslese. Der zweite Teil dieser Arbeit beinhaltet Beiträge zur Entwicklung eines neuen Datenauslesesystems für die CALICE-Szintillator-Kalorimeter. Das neue Auslesesystem erfüllt die Anforderungen für Tests von Kalorimeter-Prototypen, während es auf größere Systeme skalierbar ist. Die Anforderungen und die allgemeine Architektur des Datennahmesystem werden in dieser Arbeit skizziert. Das neue Datennahmesystem wurde im Jahr 2014 bei Teststrahl-Messungen am CERN Proton Synchrotron in Betrieb genommen und getestet. Die Ergebnisse sind hier dargestellt.

یا حق

اندازه‌گیری انرژی جت‌ها در برخورددهندهٔ خطیِ بین‌المللی:

کاربرد در اندازه‌گیری جرم بوزونِ هیگز و سامانهٔ داده‌گیری مورد نیاز

چکیده

کشف بوزون هیگز در برخورددهندهٔ هادرونی بزرگ (LHC) در سازمان پژوهش‌های هسته‌ای اروپا، سِرْنْ، ایدهٔ شکست خودبه‌خودیِ تقارن را تایید کرد. شکست خودبه‌خودی تقارن سازوکاری است که ذرات بنیادی به وسیلهٔ آن جرم‌دار می‌شوند. مطالعهٔ خصوصیات بوزون هیگز برای بازبینی پیش‌بینی‌های مدل استاندارد ذرات بنیادی اهمیت بالایی دارد. هرگونه انحراف از این پیش‌بینی‌ها می‌تواند نشانگر فیزیک فراسوی مدل استاندارد باشد. جرم بوزون هیگز یکی از کمّیّت‌های مهم مدل استاندارد است. تعیین دقیق جرم بوزون هیگز به خودی‌خود و همچنین برای دیگر مطالعات فیزیک هیگز اهمیت دارد زیرا در محاسبات تعیین پهنای واپاشی جزیی به وسیلهٔ آهنگ واپاشی وارد می‌شود.

برخورددهندهٔ خطی بین‌المللی یا آی-اِل-سی (ILC) یک برخورددهندهٔ قطبی‌شدهٔ الکترون-پوزیترون با درخشندگی بالاست که برای مطالعات دقیق فیزیک ذرات بنیادی در آینده طراحی شده است. در این پایان‌نامه، امکان اندازه‌گیری جرم بوزون هیگز در آی-اِل-سی با استفاده از کانال $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ برای انرژی‌های مرکز جرم ۳۵۰ و ۵۰۰ گیگاالکترون‌ولت بررسی شده است. از آن‌جا که جرم بوزون هیگز از دو جت ایجاد شده به وسیلهٔ کوراک‌های پایین بازسازی می‌شود، دقت اندازه‌گیری انرژی جت‌ها تاثیر به سزایی بر اندازه‌گیری جرم بوزون هیگز در این کانال دارد. در این پایان‌نامه روش تازه‌ای برای تخمین دقت اندازه‌گیری انرژی جت‌ها به صورت تک به تک ارائه شده است. سپس این تخمین‌ها در اندازه‌گیری جرم بوزون هیگز به کار گرفته شده‌اند. در این پژوهش، روش‌های گوناگونی برای اندازه‌گیری جرم بوزون هیگز بررسی شده‌اند. برای نمونه-داده‌ای با درخشندگی کل ۱۰۰۰ بر فمتوبارن و قطبش (۰/۳+ ,۰/۸-) برای انرژی‌های مرکز جرم ۳۵۰ و ۵۰۰ گیگاالکترون‌ولت، به ترتیب عدم قطعیت آماری ۴۲ مگاالکترون‌ولت و ۸۹ مگاالکترون‌ولت برای جرم بوزون هیگز به دست آمد. ریشه‌های عدم‌قطعیت‌های سیستمی نیز بررسی شده‌اند.

شبیه‌سازی کامل آشکارساز بزرگ بین‌المللی یا آی-اِل-دی (ILD) با استفاده از بستهٔ نرم‌افزاری GEANT4 برای این بررسی به کار گرفته شده است. دقت اندازه‌گیری انرژی جت مورد نیاز برای این اندازه‌گیری تنها با به‌کارگیری روش بازسازی رویداد موسوم به particle flow به دست می‌آید. از این رو، این روش اندازه‌گیری جرم بوزون هیگز تا حد زیادی به کارکرد روش بازسازی رویداد particle flow بستگی دارد. الگوریتم‌های particle flow به کالریمترهایی نیاز دارند که به بخش‌های بسیار کوچکی تقسیم شده باشند. یکی از چالش‌های فراروی ساخت چنین کالریمترهایی که از میلیون‌ها کانال برخوردار هستند، سامانهٔ داده‌گیری آنهاست. بخش دوم این پایان‌نامه به توسعهٔ یک سامانهٔ داده‌گیری جدید برای کالریمترهای CALICE اختصاص دارد. این سامانهٔ داده‌گیری نیازهای آزمایش کالریمترهای پیش‌نمونه را پاسخ می‌گوید و می‌تواند به سامانه‌های بزرگتر گسترش داده شود. نیازها و معماری کلی این سامانهٔ داده‌گیری در این پایان‌نامه ارائه شده است. سامانهٔ جدید در سال ۱۳۹۳ در مرکز سینکروترون پروتون سرن راه‌اندازی و آزمایش شد و نتایج آن در این جا ارائه شده است.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Our current knowledge of the basic building blocks of matter and their interactions is embodied in the Standard Model of particle physics. The discovery of the long-awaited Higgs boson by the ATLAS and CMS experiments at the CERN Large Hadron Collider (LHC) provides a first experimental indication that indeed the Brout-Englert-Higgs mechanism in its minimal or an extended version could explains the mass generation of the elementary particles. Albeit the Standard Model seems to be complete after the discovery of the Higgs boson, there are a number of open questions which are not answered by the Standard Model, suggesting that it is only part of a more complete theory. Any deviation of the Higgs boson properties from the Standard Model predictions can be a portal to physics beyond the Standard Model.

At the LHC only a subset of the Higgs decay modes can be observed and the precision of many of the measurements is not sufficient for investigating the deviations from the Standard Model at a level which is predicted by some of the theories beyond the Standard Model. A lepton collider such as the International Linear Collider (ILC) provides excellent experimental conditions for studying the Higgs properties to high precisions. ILC is a future polarised high-luminosity e^+e^- collider with a tunable centre-of-mass energy of 500 GeV and upgradable to 1 TeV.

Since the Higgs boson couples to all massive particles of the Standard Model, it is likely to provide answers to some of the open questions in fundamental physics. Provided that the masses of the other particles are measured, once the Higgs boson mass is known, the Standard Model predicts all couplings of the Higgs boson to the other particles without any free parameters. This in turn allows for prediction of the Higgs production cross-sections and decay branching ratios.

To extract the Higgs boson couplings from the measured branching ratios, the mass of the Higgs boson should be known to a high precision. The Higgs boson mass is one of the fundamental constants of Nature and is not predicted by the Standard Model, therefore, it needs to be experimentally measured. There are different ways to measure the Higgs

boson mass at ILC. According to the Standard Model, decay of the Higgs boson to a pair of bottom quarks has the largest branching ratio of all Higgs decays BR($H \to b\bar{b}$)$\approx 57.8\%$. Thus, this channel provides a large dataset for physics analyses. The clean experimental conditions at ILC combined with advanced detectors allow for exploiting this channel to measure the Higgs boson mass.

The measurement of the Higgs boson mass at ILC using the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel is studied in this thesis for centre-of-mass energies of 350 GeV and 500 GeV. Such measurements at ILC rely on highly-granular calorimeter. Contributions to development and commissioning of a new data acquisition system for such calorimeters is presented in the second part of this thesis.

This thesis is organises as following: the theoretical foundations for the Higgs boson mass measurement are elaborated in chapter 2. In this chapter, after a short introduction of the Standard Model the Higgs boson production and decay modes are introduced. The ILC project and the main components of the collider are presented in chapter 3. There are two general-purpose detector concepts being studied for ILC. A concise overview of the International Large Detector (ILD) concept which is used in the studies in this thesis is also presented. The various methods to measure the Higgs boson mass at ILC are discussed in chapter 4.

At the tree-level, there are four quark in the final state of the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel. The quarks generate jets of particles which are then detected in the detector. The most important subdetectors for jet measurements are calorimeters. The fundamental concepts of calorimetry and jet energy measurement is discussed in chapter 5. In addition, the particle flow approach to event reconstruction is explained. The particle flow approach enables reaching jet energy resolutions beyond the intrinsic calorimeter resolutions.

The Monte Carlo event generation and simulation procedures to create samples for this study are explained in chapter 6. Event preparation steps before the main analysis are also presented. The uncertainty on the jet energy measurement affects the outcome of the Higgs boson mass measurement using the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel. A new method to estimate the uncertainty on the jet energy for each jet individually is introduced in chapter 7. Various strategies for the Higgs boson mass measurement using the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel have been employed, including a χ^2-minimisation technique and kinematic fitting with various fit hypotheses. These strategies are discussed and their outcome is compared in chapter 8.

A new generation of calorimeters are necessary for the jet-specific energy resolution method and to reach a low relative jet energy uncertainty of $3-4\%$ which is required by the ILC physics programme. These calorimeter are optimised for particle flow approach to calorimetry which provides an unprecedented jet energy resolution. The Analog Hadronic Calorimeter (AHCAL) is one of the promising technologies for the hadronic calorimeters of the ILC experiments. In chapter 9, after a short presentation of various CALICE highly-granular calorimeter concepts, the AHCAL is introduced in detail.

One of the challenges in realisation of such highly-granular calorimeters with millions of read-out channels is their Data Acquisition System (DAQ) system. The second part of this thesis concerns the development of a DAQ system for the CALICE scintillator calorimeters. The architecture and various subsystems of a new DAQ system for the CALICE scintillator calorimeters is explained in chapter 10. Part of the DAQ subsystems are developed as the work carried on for this thesis. The new DAQ system has been commissioned and tested at the CERN PS test beam facility in 2014. The commissioning procedure and the test results are presented in chapter 11. A summary of the thesis and conclusions are given in chapter 12.

Chapter 2

Theoretical Foundations

The theoretical foundations for this thesis are described in this chapter. The two projects of this thesis concern the International Linear Collider, an international laboratory for high-energy physics experiments. The theoretical foundation of high-energy physics is known as the *Standard Model* of particle physics which is the fundamental theory of constituents of matter and their interactions.

A general overview of the Standard Model is given in section 2.1. Among the important measurements which will be performed at ILC is one of the subjects of this thesis, namely the precise measurement of the Higgs boson mass. In the Standard Model a phenomenon known as the Brout-Englert-Higgs (BEH) mechanism is responsible for the creation of the Higgs particle. The BEH mechanism and electroweak symmetry breaking are discussed in section 2.2.

The Higgs boson mass is measured in the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel. There are four quarks present in the final state of this process, each of which creates a spray of particles known as a jet. The basics of quarks and jets theory is summarised in section 2.3.

2.1 The Standard Model of Particle Physics

The fundamental structure of matter has been studied in great details throughout the twentieth century. The world as we know it is made of molecules which are formed by a number of atoms. Each atom is a system composed of a number of negatively charged *electrons* (e) which are bound to a nucleus. The nucleus is composed of positively charged *protons* (P) and electrically neutral *neutrons* (N). The electrons are bound to the nucleus by the *electromagnetic* force, while protons and neutrons in the nucleus are bound by the *strong* nuclear force.

A different force of Nature is observed in nuclear β-decays. Certain radioactive isotopes decay to an electron and a particle of very small mass called an *electron neutrino* (ν_e). The force responsible for this decay and nuclear fusion is called the *weak* force.

In addition to the other three forces, there exists another distinct force in Nature, the *gravity*. The gravity is the weakest force of the four and becomes important for macroscopic objects and distances. The gravitational interaction between microscopic objects is very small compared to the other interactions and can be neglected, nevertheless, gravity is very important in astrophysics and cosmology.

To the best of our current knowledge electrons and electron neutrinos are *elementary* particles, i.e. they are point-like particles without any internal structure or excited states. However, at higher energy scales, protons and neutrons exhibit their internal structure. They appear to be bound states of another family of elementary particles called *quark*s. Two *up-quark*s (u) and one *down-quark* (d) form a proton while two down-quarks and one up-quark form a neutron.

The findings at high-energy experiments have revealed that for each of the four elementary particles mentioned above (d, u, e and ν_e), there exist two more copies which differ primarily in mass. The first four elementary particles are collectively called the *first generation* while their heavier counterparts form the second and the third generations. The second generation includes the *strange quark* (s), the *charm quark* (c), the *muon* (μ) and the *muon neutrino* (ν_μ). The third generation elementary particles are the *bottom quark* (b), the *top quark* (t), the *tau* (τ) and the *tau neutrino* (ν_τ). All the elementary particles in the three generations are spin $1/2$, hence, they are called *fermions*.

All the findings gathered about the basic building blocks of matter are incorporated into a fundamental theory called the *Standard Model* of particle physics . The Standard Model strives to provide a mathematical description of all the phenomena observed in particle physics in terms of a small number of elementary particles and their interactions.

The mathematical framework of the Standard Model is Quantum Field Theory (QFT). The electromagnetic interactions are described by Quantum Electroynamics (QED), the strong force by Quantum Chromodynamics (QCD) and the weak interactions by the electroweak theory. In QFT each of the Standard Model elementary particles is described by a *field*. The kinematics and dynamics of the theory are given by a *Lagrangian density*, $\mathcal{L}$, or simply the Lagrangian. A Lagrangian is a function of the field $\phi(x)$ and its derivatives $\partial_\mu \phi(x)$.

Every fundamental conservation law is associated to a symmetry in the nature. This is one of the most fundamental principles in particle physics and the interactions among the elementary particles are described by the symmetries. In QFT the symmetries are expressed by requiring the Lagrangian to be invariant under certain symmetry transformations.

The symmetries are best described by the mathematical *groups*. The Standard Model is formulated based on the *gauge* groups:

$$G = SU(3) \times SU(2) \times U(1) \tag{2.1}$$

where the $SU(3)$ factor accounts for QCD and the $SU(2) \times U(1)$ factor for the electroweak theory. These factors are explained in more details in the following sections.

The six members of the quarks family, are electrically charged and also carry the *colour charge* which is the QCD equivalent of the electric charge. Therefore, the quarks interact via all three fundamental forces which are relevant for the elementary particles. The quarks are bound together by the strong force and form a group of particles know as *hadron*s.

The remaining six fermions are collectively called *lepton*s. The leptons do not carry the colour charge. Three members of the lepton family, electron, muon and tau-lepton, are electrically charged. Consequently, they undergo electromagnetic and weak interactions. The remaining three leptons are the neutrinos (electron neutrino, muon neutrino and tau neutrino) which are electrically neutral. Neutrinos interact only by the weak force.

The fermions interact by exchange of elementary spin 1 particles known as the *gauge boson*s. The gauge boson mediate an interaction by transferring a discrete amount of energy between the particles. The gauge boson of QCD which is the mediator for the strong interactions is a massless particle called *gluon* (g). Electromagnetic interactions are mediated by *photons* (γ), another massless particle which is the gauge boson of QED. The weak interactions have three massive mediators: the W^+ and W^- bosons which are electrically charged, and the Z boson which is electrically neutral.

The last member of the Standard Model elementary particles is a spin 0 particle called the *Higgs* boson (H). The fields associated with the fermions and gauge bosons have zero expectation value in the vacuum. In contrast, the vacuum expectation value of the Higgs field is not zero. This non-zero vacuum expectation value breaks the electroweak gauge symmetry and gives rise to the BEH mechanism through which all other elementary particles that interact with the non-zero Higgs field acquire mass.

The ATLAS and CMS experiments at the CERN Large Hadron Collider published evidence of the discovery of a new particle in 2012 [1, 2]. The new particle with a mass of about 125 GeV is compatible with the Standard Model Higgs boson. The mechanism by which the elementary particles acquire mass and the Higgs boson properties is discussed in the following section.

2.2 The Higgs Boson

First formulated by Glashaw, Weinberg and Salam, the electroweak theory unifies the electromagnetic and weak interactions in a single unified gauge theory under a $SU(2)_L \times U(1)_Y$ symmetry [3–5]. The electroweak theory extends the quantum electrodynamics to include the weak interactions. The Lagrangian for a free spin $1/2$ particle such as the electron is

$$\mathcal{L}_D = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \tag{2.2}$$

where γ_μ are the Dirac γ-matrices, $\psi(x)$ is the spinor field and m is the mass of the particle. $\mathcal{L}_D$ is invariant under a global gauge transformation $\psi \rightarrow e^{i\theta}\psi$ with θ being constant. However, $\mathcal{L}_D$ is not invariant under a *local* phase transformation of the field where θ is not constant, i.e. $\psi \rightarrow e^{iq\lambda(x)}\psi$.

The gauge invariance under $U(1)$ local phase transformation can be restored by introducing a new field A_μ and replacing the partial derivative ∂_μ with the so-called *covariant derivative* $D_\mu = \partial_\mu - iqA_\mu$. This replacement restores the $U(1)$ local gauge invariance provided that the new field transforms as $A_\mu \rightarrow A_\mu - \partial_\mu\lambda$.

The local gauge invariant Lagrangian for a spin $1/2$ particle becomes

$$\mathcal{L}_{QED} = \bar{\psi}(x)(i\gamma^\mu D_\mu - m)\psi(x) + \bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.3}$$

where a sum over repeated indices is implied and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The first term remains as the original Lagrangian ψ_D and describes the field of the free fermion. The second term describes the interaction of the fermion with the new field A_μ which can be identified as the massless photon. The last term describes the photon field.

The Lagrangian $\mathcal{L}_{QED}$ is invariant under the local gauge transformation assuming that the photon is massless. If the photon was massive, $\mathcal{L}_{QED}$ would include a term for the photon mass $\frac{1}{2}m_\gamma^2 A_\mu A^\mu$ which breaks the local gauge invariance. Extensive experimental evidences have shown that the photon is massless, therefore $\mathcal{L}_{QED}$ of equation 2.3 is valid and describes QED elegantly.

Using an approach similar to the case of QED, a Lagrangian for the weak interactions can be obtained by requiring the Lagrangian to be invariant under $SU(2)_L$ local phase transformation. This can be achieved by replacing the four-derivative term with a covariant derivative which is defined based on the generators of the $SU(2)_L$ group. The resulting Lagrangian satisfies $SU(2)_L$ local gauge symmetry if the gauge bosons of the interaction are massless.

However, the experimental observations have revealed that the gauge bosons of the weak interaction, $W^\pm$ and Z bosons, have large masses. Therefore, their masses should be included in the Lagrangian in which case the local gauge invariance will be broken.

Moreover, a similar problem arises with the masses of the fermions. Thus a new concept is required in order to restore the gauge invariance.

In the Standard Model the electroweak interactions are explained by a $SU(2)_L \times U(1)_Y$ gauge theory. The basic symmetry of the model does not allow for the fermions and bosons to be massive. The Electro-weak Symmetry-breaking (EWSB) provides a framework which generates the observed masses of the $W^{\pm}$ and Z bosons while conserving the structure of the gauge interactions [6–9]. The EWSB postulates the existence of a self-interacting complex scalar field with a new fundamental scalar particle, known as the *Higgs boson*.

In the Standard Model, the most general gauge-invariant renormalisable form of the scalar potential is given by:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{2.4}$$

where Φ is the Higgs field and λ is the Higgs self-coupling. The Higgs field is a self-interacting $SU(2)$-doublet defined as:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}\Phi^+ \\ \Phi^0 + ia^0 \end{pmatrix} \tag{2.5}$$

where Φ^+ is the complex charged component, and Φ^0 and a^0 are the CP-even and CP-odd neutral components of the Higgs doublet.

A negative μ^2 value in equation 2.4 causes the neutral part of the Higgs doublet to assume a non-zero vacuum expectation value which leads to the spontaneous symmetry breaking. By minimising the potential, the vacuum expectation value $\nu = -\frac{\mu}{\sqrt{\lambda}}$ and the mass of the physical Higgs boson $M_H = \sqrt{2\lambda}\nu$ are obtained.

The Higgs Lagrangian is given by

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \tag{2.6}$$

where $D_\mu = \partial_\mu + ig\sigma^a W^a_\mu/2 + ig' Y B_\mu/2$, W_μ and B_μ are the gauge fields associated with the $SU(2)_L + U(1)_Y$ local symmetry with g and g' couplings, σ^a with $a = 1, 2, 3$ are the Pauli matrices and $Y = 1/2$ is the hypercharge of the Higgs doublet. This leads to $W^{\pm}$ and Z bosons acquiring masses given by

$$M_W = \frac{g\nu}{2} \tag{2.7}$$

$$M_Z = \frac{\sqrt{g^2 + g'^2}\nu}{2}. \tag{2.8}$$

The fermions acquire mass by interacting with the Higgs field through the *Yukawa interactions*, Lagrangian of which is given by

$$\mathcal{L}_{Yukawa} = -\hat{h}_{d_{ij}}\bar{q}L_i\Phi d_{R_j} - \hat{h}_{u_{ij}}\bar{q}L_i\tilde{\Phi}u_{R_j} - \hat{h}_{l_{ij}}\bar{l}_{L_i}\Phi e_{R_j} + h.c. \tag{2.9}$$

where $\tilde{\Phi} = i\sigma_2\Phi^*$ is the conjugate Higgs doublet, qL, u_R and d_R are 3×3 are Yukawa coupling matrices for quarks $SU(2)$ doublet and singlets while l_L and e_R are the lepton $SU(2)$ doublet and singlet.

The coupling of the Higgs boson to the other fundamental particles of the Standard Model can be predicted in terms of the particles masses and the vacuum expectation value. The couplings are given by:

$$g_{Hf\bar{f}} = \frac{M_f}{\nu}, \tag{2.10}$$

$$g_{HVV} = \frac{2M_V^2}{\nu} \tag{2.11}$$

$$g_{HHVV} = \frac{2M_V^2}{\nu^2} \tag{2.12}$$

$$g_{HHH} = \frac{3M_H^2}{\nu} \tag{2.13}$$

$$g_{HHHH} = \frac{3M_H^2}{\nu^2} \tag{2.14}$$

where M_f is the fermion mass, V represents the vector bosons ($W^{\pm}, Z$) and M_V is the mass of the vector boson. The Higgs couplings to fermions are linearly proportional to the masses of the fermions while the couplings to the bosons are proportional to the square of the boson mass.

2.2.1 The Higgs Boson Mass

The mass of the Higgs boson is one of the important free parameters in the Standard Model. As mentioned in the previous section, the mass of the Higgs boson is $M_H = \sqrt{2\lambda}\nu$, where ν is the vacuum expectation value and λ is the Higgs self-coupling parameter in the Higgs potential (see equation 2.4). While the vacuum expectation value can be determined by only one parameter, namely the Fermi coupling as $\nu = (\sqrt{2}G_F)^{1/2}$, the parameter λ can not be known a priori, hence, the Higgs boson mass is a free parameter in the Standard Model. A precise measurement of the Higgs boson mass is necessary for validating the Standard Model consistency and also for precision Higgs physics.

After the discovery of the Higgs boson in 2012, the ATLAS and CMS experiments continued to collect collision data at the centre-of-mass energy of 8 TeV and published a combined measurement of the Higgs boson mass using datasets of 5 fb^{-1} at the centre-of-mass energies of 7 TeV and 20 fb^{-1} at 8 TeV for each of the experiments [10]. The combined measured mass of the Higgs boson is $125.09 \pm 0.21 \pm 0.11$ GeV. It is estimated

that with large data samples of approximately 300 fb^{-1} at the end of LHC operation, the ATLAS and CMS experiments will be able to determine the Higgs mass with an uncertainty of about 0.1 GeV [11–13]. However, a higher precision is required for precise Higgs physics and to look for any deviation from the Standard Model predictions.

The Higgs boson can decay to all other lighter fundamental particles of the Standard Model. Nevertheless, as shown in equation 2.10, the couplings are proportional to the mass of the particles involved in the decay, hence, the more massive particles have larger branching ratios. As a result, the largest branching ratio of the Higgs boson is to bottom quarks $BR(H \to b\bar{b}) = 57.8\%$. However, at hadron colliders such as the LHC due to large QCD backgrounds this channel has a very poor signal to background ratio and a poor mass resolution and can not be used for precise mass measurement.

Lepton colliders offer a clean experimental environment such that the $H \to b\bar{b}$ channel can be exploited for precise measurements of the Higgs boson properties, in particular the mass. Nevertheless, the measurement of hadronic final states comes with their specific complications which is explored in the following section.

2.3 Quarks and Jets

In the second half of the twentieth century, the high-energy experiments have revealed the structure of the atomic nuclei and the interactions of their constituents: quarks are bound together by the strong force. Moreover, a large number of composite particles which consist of quarks have been discovered and studied. These particles are collectively known as *hadrons*. The experiments have provided evidences that quarks are fractionally charged spin $1/2$ particles.

The strong interaction is a fundamental interaction in the Standard Model. The quantum field theory of the strong interactions is known as Quantum Chromodynamics (QCD). In this section the key concepts in QCD which are relevant to the work carried out in this thesis are presented.

2.3.1 Quantum Chromodynamics

The concept of Quantum Chromodynamics is very similar to that of QED except that QCD is formulated based on the $SU(3)$ gauge group. Thus, the QCD interactions are mediated by eight massless bosons corresponding to the eight generators of the $SU(3)$ local gauge symmetry. The bosons of the strong interaction are called *gluons*. Unlike photons, the gluons have self-interactions.

The Lagrangian of QCD is expressed as

$$\mathcal{L}_{QCD} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \sum_q \bar{q}_i(i\gamma^\mu D_\mu - m_q)_{ij}q_j \tag{2.15}$$

where a sum over repeated indices is implied, g_s is the gauge coupling, covariant derivative $D_\mu = \partial_\mu - ig_s T^a_{ij}A^a_\mu$ and field strength tensor $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc}A^b_\mu A^c_\nu$. The QCD Lagrangian is invariant under the $SU(3)$ local gauge transformation.

In $\mathcal{L}_{QCD}$, T^a_{ij} are 3×3 matrices generators of $SU(3)$, which means the wave function has three additional degrees of freedom, known as the *colour charge*s. The colour charges are named *red* (R), *green* (G) and *blue* (B). Only particles which carry colour charge couple to gluons, therefore, leptons do not participate in the strong interactions.

Unlike in QED where the photons are uncharged, the gluons carry the colour charge. Each gluon carries a colour charge and an anticolour charge simultaneously, creating nine different combinations corresponding to the generators of the $SU(3)$ gauge symmetry $R\bar{G}$, $G\bar{R}$, $R\bar{B}$, $B\bar{R}$, $G\bar{B}$, $B\bar{G}$, $G\bar{G}$, $R\bar{R}$ and $B\bar{B}$.

Although the existence of quarks is supported by a great number of experimental evidences, no free quark has ever been observed directly in an experiment. In QCD, this is explained by the so-called *colour confinement*. According to the colour confinement, colour-charged particles are confined to colourless states and therefore can not be isolated or directly observed. Currently, there is no analytical proof for the confinement, however, it is widely believed that the confinement is due to self-interactions of gluons.

QCD is an *asymptotically free* theory. This means that the strength of the QCD coupling is large at low-energy scales (large distances) and becomes considerably smaller at high energy scales (short distances). Due to this *asymptotic freedom* of QCD, the quarks can be considered as quasi-free particles inside hadrons rather than being strongly bound.

2.3.2 Hadronisation and Jets

Through a process known as *hadronisation*, the quarks turn into colourless hadrons which are observed in the detector. The QCD coupling is large for the energy scales of hadronisation. Therefore, these processes are not calculable using the perturbation theory. However, a number of phenomenological models have been developed. Experimental data can be reasonably described by these models, although each model has many free parameters and is far from a being a concrete theoretical description. These phenomenological models can be categorised into three main classes: independent hadronisation, string hadronisation [14–16] and cluster hadronisation [17, 18].

According to the colour confinement, only colourless states can exist as free particles. As a results, all bound states of quarks and antiquarks are colourless states. When a quark

and an antiquark are created in a process such as $e^+e^- \rightarrow q\bar{q}$, they will separate with equal and opposite momentum in the centre-of-mass reference frame. At short distances of the order of 1 fm they can move as quasi-free particles due to the asymptotic freedom of QCD. When the separation becomes larger, each of them will then materialise through the hadronisation process to a collimated spray of colourless particles called a *jet*. The gluons undergo hadronisation likewise.

The jet production mechanism can be explained in two steps. In the first step, the quarks which are accelerated in the production process in short distances ≤ 1 fm will emit gluons. Similar to the bremsstrahlung process in which charged particles radiate photons, coloured particles can radiate gluons. However, since the gluons carry colour charges, they can also emit further radiation. This process leads to a cascade of particles known as *parton showering*.

In the second step, when the created quarks separate further, a gluonic flux tube is build up between each pair which confines the two objects. At larger distances, around 1 fm, the energy stored in the colour field between them increases until it is sufficient to produce a new $q\bar{q}$ pair. This causes the colour field to break into smaller fields with lower energy, which is favourable. This process continues and creates a cascade of $q\bar{q}$ pairs until the energy of quarks and antiquarks is low enough to combine into colourless hadrons. The unstable hadrons which are produced in the hadronisation process decay to more stable particles that are then observed in the detector.

2.4 Physics Beyond the Standard Model

The Standard Model of particle physics has been extremely successful not only in describing many physical phenomena observed in experiments but also in predicting new phenomena. The current version of the Standard Model has a relatively large number of free parameters. These parameters and particle quantum numbers (electric charge, weak isospin and hypercharge) have no theoretical explanation and are chosen to match the experimental observations. Combining purely theoretical tools with experimental observation, the Standard Model has been constructed in such a way that it can explain and reproduce the experiments. However, there are a number of open questions which are not answered by the Standard Model, suggesting that it is only part of an ultimate theory of particle physics.

gravity although considered as one of the four fundamental forces of Nature, the Standard Model does not explain gravity. The attempts to include general relativity, the most successful theory of gravity to date, in the Standard Model has not been successful.

matter-antimatter asymetry The Standard Model predicts that matter and antimatter should have been created in roughly equal amounts after the Big Bang. However,

there is an imbalance in the observable Universe which is mostly made of baryonic matter and radiation. The Standard Model provides no explanation for this imbalance.

dark matter Astrophysical and cosmological observations have provided compelling evidences for the existence of some form of invisible matter known as *dark* matter. In order to be invisible, dark matter particles should not undergo electromagnetic and strong interactions. Although many candidates for dark matter particles have been proposed, none of them has been observed in the experiments. New physics beyond the Standard Model is necessary to explain the nature of dark matter.

neutrino masses the Standard Model does not consider mass for neutrinos. However, experiments have shown that the neutrinos are not massless. Neutrino masses can be added to the Standard Model as extra free parameters but these masses must be extremely small and the mechanism by which the neutrinos gain mass is unclear.

hierachy probelm The Higgs boson mass has contributions from quantum loop corrections. Assuming that the Standard Model is valid up to energy scales orders of magnitude higher than the electroweak symmetry breaking scale, very large quantum loop corrections become necessary at high energies and keeping the Higgs mass at the scale of 100 GeV becomes cumbersome. This situation is not considered as being natural.

There are a number of theoretical proposals which extend the Standard Model so that a few of the open question are answered. As mentioned earlier in this chapter, the Standard Model has three gauge symmetries which correspond to the three fundamental forces it covers. Each of these symmetries has a value of the coupling which changes with the energy at which it is measured. It has been noted that at an energy scale of around 10^{16} GeV the values for the three coupling converge. There are theories that attempt to unify the Standard Model symmetries into one simple symmetry with only one coupling constant at such high energies. These theories are collective known as *Grand Unified Theories* [19].

Another approach to extend the Standard Model is to add another class of symmetries to the Lagrangian by considering a super-partner for each of the Standard Model particles. The super-partners differs by half a unit of spin from their Standard Model partners, i.e. each fermion has a bosonic partner and each boson has a fermionic partner. There theories are known as *Supersymmetry* (SUSY) theories [19].

As explained in this chapter, the mass of the Higgs boson is one of the important parameters of the Standard Model. Knowing the Higgs boson mass, other properties of the Higgs boson can be predicted by the Standard Model. A precise measurement of the Higgs boson mass enables using the Higgs boson as a probe for the Standard Model and physics beyond the Standard Model.

Measurements at the hadron colliders such as LHC can not reach the level of uncertainty required by many of the beyond Standard Model theories. A lepton collider is the appropriate tool for conducting precision physics studies, including the Higgs boson mass measurement. The International Linear Collider (ILC) is a future lepton collider which is designed for such studies. The ILC machine and detectors are introduced in the following chapter.

Chapter 3

The International Linear Collider

The Standard Model of particle physics has been successful in describing the physics phenomena at both low and high energies. However, the Standard Model fails to answer a number of important open questions such as matter-antimatter asymmetry and the nature of the dark matter, as discussed in chapter 2, indicating that the theory is incomplete. More particles and interactions other than what is included in the Standard Model are required to expand the theory beyond its current limitations.

The quest for a more complete theory which answers the open questions needs to be pursued at facilities which offer higher energies and higher precision of the measurements. Currently, circular hadron colliders such as the Large Hadron Collider (LHC) at CERN are able to reach very high energies in the range of a few tera-electronvolts. The experiments at the LHC together with astro-particle experiments allow for searching for new physics at very high energies. However, regardless of the precision of the detectors of the hadron collider experiments, the precision of the measurements is limited due to the fact that the colliding particles are not elementary and the initial conditions of the collisions are unknown.

Lepton colliders, on the other hand, provide a clean environment for precise measurements. But the energy reach of lepton colliders is limited by synchrotron radiation in circular colliders and length of the machine in linear colliders. However, recent technological achievements allow for machines with centre-of-mass energies far beyond the previous machines such as the Large Electron-Positron Collider (LEP) at CERN.

Currently, there are a number of proposals for future lepton colliders such as the Compact Linear Collider (CLIC) [20], the Future Circular Collider (FCC) [21] and the Circular Electron-Positron Collider (CEPC) [22]. The most advanced proposal is a lepton collider for precision physics studies called the International Linear Collider (ILC). The ILC is a polarised e^+e^- collider with tunable centre-of-mass energies of $250 - 500$ GeV which can be upgraded to 1 TeV.

In this chapter, after a discussion on the physics case for a lepton collider such as ILC in section 3.1, a general overview of the ILC machine is given in section 3.2. There are two general-purpose detector concepts being designed and optimised for ILC. One of these concepts is the International Large Detector (ILD) concept. ILD and its various subdetectors are introduced in section 3.3.

3.1 Physics Case for a Lepton Collider

In contrast to the experiments at hadron colliders such as LHC where many overlapping events are present due to the QCD events, the experimental conditions provided by an e^+e^- collider are rather clean. Moreover, a linear lepton collider such as ILC has a low duty cycle of 0.5% with small event sizes. As a result, for detectors with today's advanced technology, a triggering system is not necessary and all collision data can be used in physics analyses. Moreover, the beam energy is tunable and the initial conditions of the colliding particles are known to a high level of precision. All these advantages promote high-energy e^+e^- linear collider machines as an essential tool for precision studies of fundamental particle physics. The study of the numerous observables available at an e^+e^- linear collider with an unprecedented precision can open a portal to new physics which can answer a number of the open questions.

Precise measurements of the properties of the Higgs boson are one of the main physics goals of such an e^+e^- linear collider. Mass, width, spin, CP and couplings of the Higgs boson will be measured to a great precision. A model-independent measurement of the Higgs couplings to the known standard model particles [23] allows for an investigation of the Standard Model prediction that the couplings are proportional to the mass of the particles. The measurement can also provide an answer to the important question whether the discovered Higgs boson is the only manifestation of the Higgs field or is part of an extended Higgs sector such as that predicted by for example SUSY theories. Hadron colliders such as LHC can only perform model-dependent measurements since the total Higgs production cross-section can not be measured without any assumption. Moreover, hadrons are not elementary particles and the initial conditions of their constituents, the actual colliding particles, are unknown and therefore the measurements rely on parton models.

Studying the top quark and gauge sector is an important tool to investigate the electroweak-symmetry breaking mechanism and physics beyond the Standard Model. These studies form an important part of the physics programme of a linear e^+e^- collider [24, 25].

Due to its very large mass, the top quark plays a special role in the Standard Model. Its strong coupling to the electroweak sector can be exploited to investigate the dynamics of the electroweak symmetry-breaking mechanism. Many of the Standard Model parameters such as the W and Z couplings and the Higgs boson mass are affected by the top quark mass via radiative corrections. While top quark mass measurements at hadron colliders have

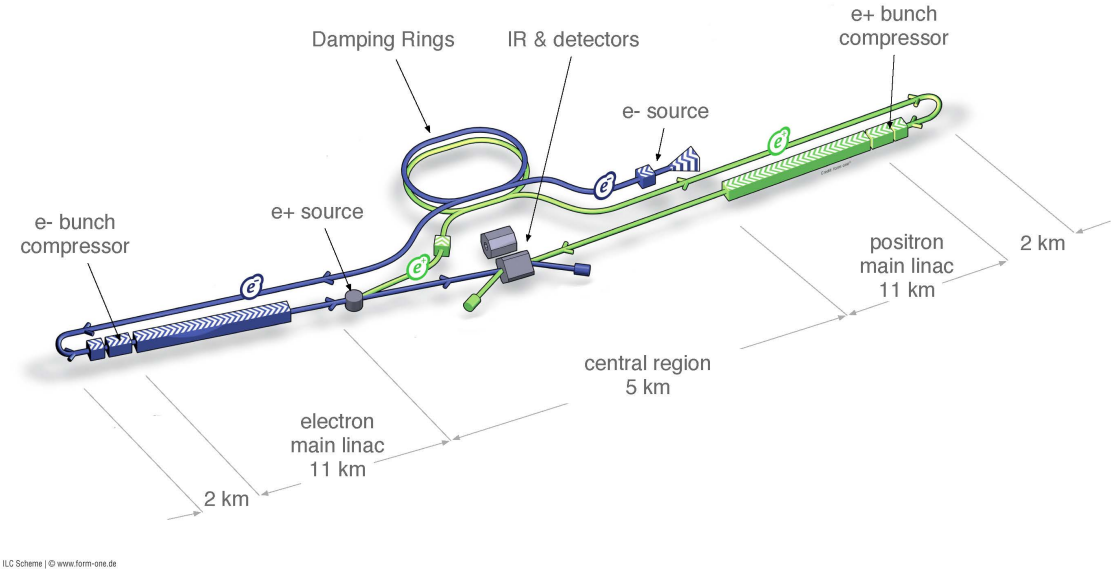


Figure 3.1 – A schematic overview of the International Linear Collider [27].

large theoretical and experimental uncertainties, the measurement using the production cross-section at the $t\bar{t}$ threshold at an e^+e^- linear collider provides a unique opportunity to determine the mass of the top quark to a high precision. Apart from the mass, top-antitop asymmetry and couplings to the gauge bosons measurements at a lepton collider provide a powerful probe for exploration of the Standard Model and beyond the Standard Model physics [26].

The physics programme for future lepton colliders has been studied over many years, in particular for ILC. The ILC physics programme is explained in detail in [24] and summarised in [25].

3.2 The International Linear Collider

One of the machines proposed for studying the physics cases mentioned in the previous section is the International Linear Collider (ILC). The design of ILC is based on more than two decades of active research and development for linear colliders. Under a mandate from the International Committee for Future Accelerators (ICFA), an international collaboration of more than 300 national laboratories, research institutes and universities produced and published a Reference Design Report (RDR) for ILC in 2007 [28–31]. After several years of active research and development to address the technological challenges which were identified in the RDR, the Technical Design Report (TDR) for ILC was published in 2013 [24, 27, 32–34].

In this section, after a general overview of the ILC collider, various components of the ILC machine are introduced. The ILC machine is presented in more details in [32] and [33].

3.2.1 General Overview

ILC is a linear collider with a total footprint of about 31 km. The centre-of-mass energy of the collisions is tunable from 200 to 500 GeV and an upgrade to 1 TeV is foreseen. The main parameters of the 500 GeV ILC machine are listed in table 3.1.

The main components of the ILC machine are illustrated in figure 3.1. A photocathode DC gun is the source of polarised electrons. The electrons are then bunched and pre-accelerated to 76 MeV using normal-conducting structures. In the next step a super-conducting linac is employed to accelerate the electrons to 5 GeV and a super-conducting solenoid rotates their spin vector to the vertical before they are injected to a damping ring.

The damping ring is used to reduce the emittance of the beams to fulfil the requirements of the luminosity production. The electrons are then transferred to the far-end of the main electron linac using a 5 GeV transport line. Prior to entering the main linac, a two-stage bunch compressor system is used to compress the bunches and accelerate them to 15 GeV. The beam is then accelerated to the maximum energy of 250 GeV in the main linac.

At the end of the main electron linac, a fraction of the beam is extracted to the positron source system to generate positrons and the remaining bunches are transported towards the interaction point. Using structures similar to the ones for the electron beam, the positrons are first accelerated and injected to the positron damping ring before they are transported to the main positron linac for acceleration to the maximum energy of 250 GeV.

The beam delivery system of ILC transports the electron and positron beams from the exit of the main linacs to the interaction point while collimating and focusing them to reach the luminosity goals and then transfers the spent beam to the main beam dumps. The beam delivery system also includes polarisation and energy monitoring systems.

3.2.2 Polarised Sources

One of the major advantages that is offered by ILC is polarisation of the beam particles. Using polarised beams with opposite signs for the electron and positron beams enhances the cross-section of interactions in which an electron and positron of opposite helicity annihilate into a vector boson. Moreover, using same-sign beam polarisation suppresses the Standard Model background in searches for new physics where a scalar particle acts as the propagator in a s-channel interaction.

Table 3.1 – The main parameters of the base-line 500 GeV ILC machine [33].

Parameter	Unit	Base-line machine		
Centre-of-Mass Energy	GeV	250	350	500
Electron linac rate	Hz	10	5	5
Number of bunches		1312	1312	1312
Bunch population	$\times 10^{10}$	2.0	2.0	2.0
Bunch separation	ns	554	554	554
Average total beam power	MW	5.9	7.3	10.5
Estimated AC power	MW	122	121	163
Electron polarisation	%	80	80	80
Positron polarisation	%	30	30	30
Fractional RMS energy loss due to beamstrahlung	%	0.97	1.9	4.5
Luminosity	$\times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$	0.75	1.0	1.8

At ILC, the polarised electron beam is produced by a DC photocathode gun. A circularly-polarised laser beam illuminates a Gallium arsenide photocathode and generates bunch trains of electrons. A maximum polarisation of 90% can be achieved using this technology. The electrons are then bunched into trains of 1312 bunches of 2.0×10^{10} electrons at a repetition rate of 5 Hz. Each bunch is 1 ns long at the source and the bunch repetition rate is 1.8 MHz. The electrons are pre-accelerated to 76 MeV in normal-conducting structures before being accelerated to 5 GeV in a super-conducting linac and injected into the damping ring.

After acceleration in the main electron linac, part of the electron beam is transported to a super-conducting helical undulator system where photons of $10 - 30$ MeV energy are generated. The photons are directed to impinge onto a rotating titanium-alloy target and produce electon-positron pairs. In the next step, the positrons are separated from electrons and remaining photons and are accelerated to 5 GeV before injection to the positron damping ring. The positron beam produced by such a system has a polarisation of greater than 30%. The bunch structure of the positrons is similar to that of the electrons. The positron production in the undulator system degrades for electron beam energy below 150 GeV. To address this issue, the electron source is required to run at 10 Hz repetition rate for centre-of-mass energies below 300 GeV in order to provide an additional electron pulse for positron production.

The electron and positron beams produced at the sources are longitudinally polarised. In order to preserve the polarisation in the damping rings, spin rotator systems are employed to change the spin vectors to vertical, parallel to the rotation axis of the damping ring, prior to injection to the damping ring. The spin vectors are rotated back to longitudinal after the damping ring.

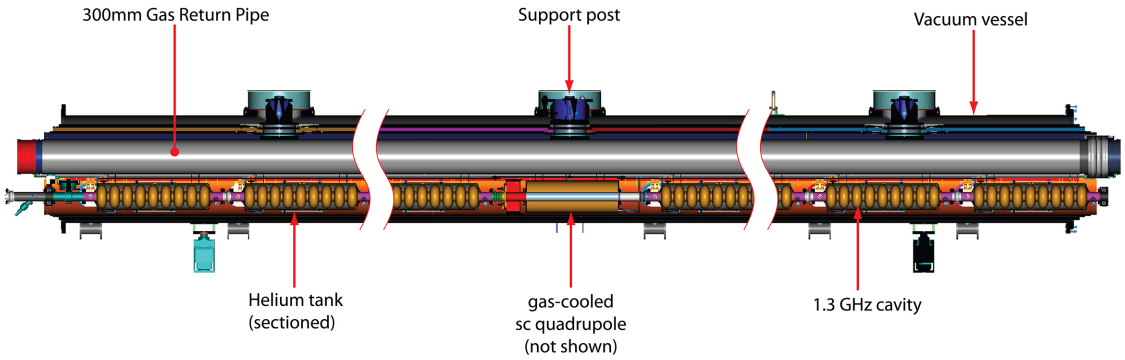


Figure 3.2 – A longitudinal cross-section of an ILC cryomodule [33].

3.2.3 Damping Rings and Ring to Main Linac

The electron and positron beams produced by the sources have very high emittance. ILC employs a damping system comprised of two separate rings for electrons and positrons to reduce the beam emittance to the required level. The damping rings with a circumference of 3.2 km operate at a beam energy of 5 GeV. Superferric wigglers along the rings perform the damping.

The Ring to Main Linac (RTML) systems are identical for electron and positron beams. Each RTML includes a $\sim$ 15km long transport line which transports the 5 GeV particle beam to the far-end of the main linac. As mentioned in section 3.2.2, the spin vector of the particles shall change back to the longitudinal after the damping ring. The desired orientation of the spin vectors is achieved by the spin rotator subsystem of the RTML. A two stage bunch-compressor subsystem reduces the beam bunch-length from several millimetres to a few hundred microns while accelerating the particles to an energy of 15 GeV. A 180 deg turn-around subsystem at the end of the RTML turns the direction of the beam towards the main linac.

3.2.4 Main Linacs

ILC employs 1.3 GHz Super-conducting Radio Frequency (SCRF) technology to accelerate polarised beams of electrons and positrons to the required energies. SCRF technology accelerates particles using resonant Radio Frequency (RF) cavities made of super-conducting material. An antenna is used to couple in a RF source which has the same frequency as that of the cavity mode. Under these circumstances high-amplitude resonant fields are formed. The electric fields accelerate charged particles when they pass through the cavities while the magnetic field deflects them.

Using super-conducting instead of normal-conducting material to build the cavities, a higher duty-cycle and a lower beam impedance can be achieved. Based on decades of research and development, ILC will use 1.3 GHz SCRF technology realised in nine-cell

niobium cavities. The nine-cell niobium cavities are cryogenically cooled to 2 K in order to maintain super-conductivity.

The cavities shall have an accelerating gradient of 31.5 GeV on average with a quality factor of $Q_0 \geq 10^{10}$ in order to reach the maximum centre-of-mass energy of 500 GeV. Approximately 7400 nine-cell cavities are required for each linac and to accommodate possible deviations from the maximum achievable gradients due to mass production imperfections, a random cavity-to-cavity spread of $\pm 20\%$ is tolerated.

Eight or nine cavities along with their corresponding support structures are packaged in a cryomodule, as shown in figure 3.2. Each linac of ILC is composed of approximately 850 cryomodules.

3.2.5 Beam Delivery System

After the particles are accelerated to the maximum energy in the main linacs, the Beam Delivery System (BDS) is responsible to bring the particles into collision under the required conditions. In addition to transporting the beams from the exit of the main linacs to the Interaction Point (IP), the BDS focuses the beams before the collision to fulfil the luminosity requirements. At the end of the BDS, the nominal bunch length is 300 μm.

After acceleration in the main linacs, the beams are surrounded by a beam-halo. The beam-halo particles can be a source of background in the detectors. The BDS includes a collimation system to remove the beam-halo particles. The collimation system is shielded using magnetised iron to deflect muons which can be produced in the collimation process.

ILC has a single interaction point where the two particle beams are brought into collision with a crossing-angle of 14 mrad. Such a geometry allows for collecting the spent beams in separate extraction lines and transporting them to the beam dumps. The BDS employs crab cavities to rotate the bunches in the horizontal plane such that heads-on collisions occur. A final focusing system uses strong quadropole magnets to focus the beams at the IP.

In addition to the aforementioned tasks, the BDS system is also responsible for beam monitoring purposes. The emittance of the beam is measured and corrected and key physics parameters such as the beam energy and polarisation are also measured in the BDS. These measurements are also crucial for the ILC physics programme. Each BDS is equipped with an upstream and a downstream polarimeter which use Compton scattering of high-power lasers to measure beam polarisation [35]. The design goal of the ILC polarimeters is to reach a relative accuracy of 0.25% [36]. A spectrometer based on beam position monitors which is located upstream of the interaction point and a synchrotron imaging detector which is located downstream are used for precise beam energy measurements to a relative accuracy of 10^{-4} [35]. An accurate beam energy measurement is crucial for the kinematic fitting as explained in chapter 8.

3.2.6 Machine Backgrounds

There are a number of machine backgrounds at ILC which need to be accounted for in the physics analyses. *beamstrahlung* [37] is one of the main sources of beam-induced backgrounds at ILC. Interaction of particles in a bunch with the electromagnetic field of an upcoming bunch generates synchrotron radiation which is known as beamstrahlung. The large amount of low-energy e^+e^- pairs which are produced as a result of beamstrahlung mainly effect the forward regions of the detectors. A so-called *anti-DID* system which is created by adding a dipole field to a conventional solenoid filed is foreseen to reduce the amount of the beamstrahlung backgrounds in the detector [38].

In addition to the e^+e^- pairs, interaction of bunches can produce a significant number of photons. Interaction of the photons is another important source of beam backgrounds at ILC. These $\gamma\gamma$ interaction can produce hadrons with low transverse momentum. The rate of these events varies with the machine parameters and has been studied for various machine designs [39]. The $\gamma\gamma$ background in the scope of the Higgs boson mass measurement analysis is addressed in sections 6.2.2 and 6.5.2.

3.2.7 Detectors

The requirements on the detectors for ILC are driven by its extensive physics programme, as explained in section 3.1. As general-purpose detectors, they should fulfil the requirements for all of these studies over the full energy range of ILC. The ILC physics programme requires significant improvements over the current technology of current collider detectors. Many technological challenges need to be overcome in order to provide adequate measurement precision for various physical observables.

Many of the physics processes of interest at ILC involve hadronic final states. The required precision on the jet energy measurements at ILC is 3 to 4 percent for 100 GeV jets [34]. This requirements is set by the need for separating W and Z di-jet final states. The requirement on the jet energy resolution is beyond what is achievable by the current state of the art detectors. The particle flow approach [40] has been developed as an alternative to the classical reconstruction approach, allowing for significantly better jet energy resolution. Particle flow reconstruction requires highly-segmented calorimeters and a precise tracking system.

The Higgsstrahlung process, in which a Higgs boson recoils off an accompanying Z boson that decays into two leptons, sets the requirements on the resolution of the tracking system. In order to be able to reconstruct the Higgs boson mass to the required precision using the Higgsstrahlung process, a track momentum resolution of $\Delta p/p = 5 \times 10^{-5}\ (\mathrm{GeV/c})^{-1}$ is required [34]. This momentum resolution fulfils the requirements of the particle flow algorithms (PFAs) as well.

An efficient flavour tagging and quark charge tagging is necessary to study the hadronic decays of the Higgs boson as well as the b-quark charge asymmetry. These studies set the requirements on the vertex detector. The vertex detector should be capable of measuring the impact parameter of a track to an accuracy better than $5\ \mu\mathrm{m} \oplus 10\ \mu\mathrm{m/p(GeV/c)}\ \sin^{3/2}(\theta)$.

The extensive research and development programs have shown that the aforementioned ambitious goals are achievable. In parallel to accelerator activities, the detector community pursued their research and development programs which has lead to detailed base-line designs for two general-purpose detectors for ILC. A push-pull approach will be followed to time-share the single interaction point of ILC between the two experiments. In the push-pull scheme, while one of the detectors is taking data at the IP, the other is parked in a nearby maintenance position. The data taking continues for short time intervals before one detector is pushed out of the IP and the other one is pulled in.

There are two detector concepts being developed for ILC [34]: the International Large Detector (ILD) [41] and the Silicon Detector (SiD) [42]. Both of these detectors are general-purpose detectors, designed to fulfil the requirements of the ILC physics programme.

SiD utilises a full silicon tracking system in a 5 T magnetic field, hence the name Silicon Detector. SiD exploits time-stamping of single bunch crossings to eliminate spurious background bursts. Highly-segmented calorimeters of the SiD allow for single particle reconstruction using the particle flow algorithm. SiD employs the Si-ECAL technology for its electromagnetic calorimeter and the AHCAL or DHCAL technology for its hadronic calorimeter (see chapter 9 for more details on the calorimeter technologies). The SiD concept is explained in [34] and [42]. The ILD detector concept is introduced in more details in the following section.

3.3 The International Large Detector

The International Large Detector (ILD) is a general purpose detector for ILC. ILD is optimised to perform a broad range of physics studies over a wide range of energies. The detector is designed around the particle flow concept to deliver high precision measurements.

At the core of ILD, around the beam-pipe, is a high-precision vertex detector. The vertex detector is surrounded by a Time-Projection Chamber tracking system which is augmented with layers of silicon detectors. A highly-segmented calorimeter system follows the tracking system. The entire tracking and calorimeter system is enclosed in a 3.5 T solenoid magnet. The detector is completed by an iron magnetic field return yoke which is instrumented to act as a tail-catcher for the calorimeter system and a muon detection system. Various subdetectors of ILD are presented in this section.

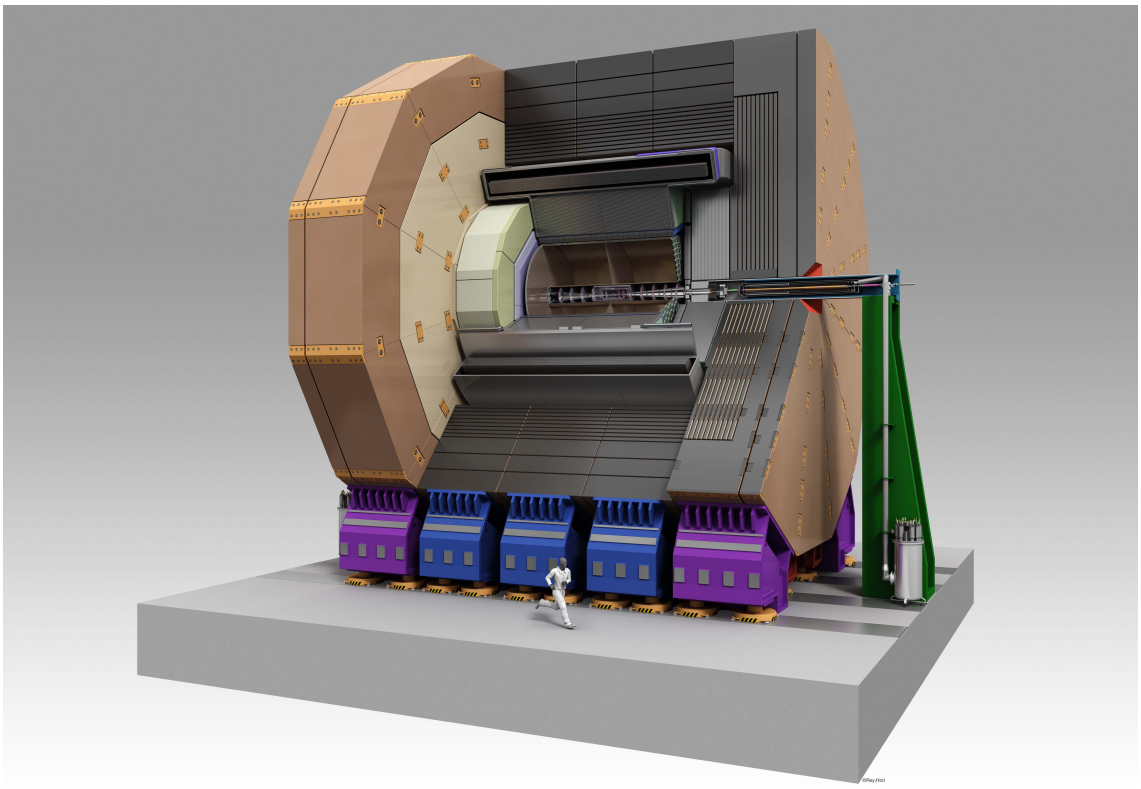


Figure 3.3 – A schematic overview of the International Large Detector detector [34].

3.3.1 ILD Coordinate System

In order to allow for extraction of the beams to a separated beam pipe after collisions, the interaction region at ILC has a horizontal crossing angle of 14 mrad between the beams. The mean beam direction is defined by the bisecting line of the smaller angle between the three-momenta of the two beams.

The coordinate system of ILD is a right-handed Cartesian coordinate system with its origin at the nominal interaction point [43]. The z-axis is along the mean beam direction pointing such that the z-component of the electron beam three-momenta is positive $p_z^- > 0$. The y-axis is vertically pointing upwards. In this coordinate system, the crossing angle is defined as the angle by which $\boldsymbol{p}^+$ has to be rotated around the y-axis such that it becomes anti-parallel to $\boldsymbol{p}^-$.

3.3.2 Tracking System

The tracking system of ILD is composed of multiple subdetectors. A vertex detector is located at the core of ILD surrounded by the main tracker which is a Time-Projection Chamber. The main tracker is augmented with a few layers of silicon tracking system. The various subsystems of the ILD tracking system are briefly introduced in the following.

Vertex Detector

The ILC physics programme relies on the identification of the heavy quark flavours. These flavours can be identified by reconstruction of the decay vertices of short lived particles such as D or B mesons which are formed by the hadronisation of heavy quarks. This procedure requires a very precise measurement of the tracks of charged particles in the close vicinity of the interaction point [34].

In addition to the performance requirements imposed by the physics programme, there are a number of requirements for realisation of the detector. The pixel detector is located at the core of ILD, therefore, not only the sensitive material should maintain a low material budget of less than $0.15\%X_0$ per layer, but also power consumption and dissipation should be as low as possible to reduce the amount of material needed for a cooling system inside the sensitive volume. Moreover, the detector should be tolerant to radiation from the beam backgrounds of a total annual dose of 1 kGy and fluence of approximately 10^{11} $\mathrm{n_{eq}}/\mathrm{cm}^2$. The spatial resolution near the interaction point should be better than 3μm.

ILD uses silicon pixel sensors for its vertex detector. Three layers of concentric nearly-cylindrical ladders which are equipped with silicon pixel sensors on both sides form the base-line design of the ILD vertex detector. The layers of the vertex detector surround the beam-pipe with the first layer placed at a radius of approximately 1.6 cm from the beam. The double-sided structure can be exploited to form mini-vectors using hits in the two adjacent sensors on a ladder. The mini-vectors can be used for easier alignment and to improve track reconstruction.

There are three different types of silicon pixel sensor technologies being investigated for ILD: CMOS Pixel Sensors (CPS) [44–46], Fine Pixel CCDs (FPCCD) [47–49] and Depleted Field Effect Transistors (DEPFET) [50, 51]. These technologies can potentially fulfil the requirements of the ILD vertex detector [34]. FPCCDs have a large number of pixels, leading to a slow read-out where low power consumption is considered. A power-pulsed read-out for CPS and DEPFET sensors is being studied.

Time-Projection Chamber

ILD uses a Time-Projection Chamber (TPC) as its central tracker. A TPC is a particle tracking system comprised of a sensitive volume of gas or liquid equipped with a position-sensitive electron collection system. A high-voltage electrode disk divides the cylindrical chamber into two halves and establishes an electric field between the disk and the end-plates where the position-sensitive system is located. Charged particles traversing the chamber ionise the gas or liquid along their track. The electrons and ions produced by the ionisation drift towards the electrodes in the electric field. The third dimension, which is along the axis of the cylinder, is determined using the drift time.

For the end-plates of the ILD TPC, Gas Electron Multiplier (GEM) [52] and Micro-MEsh GAseous Structure (MICROMEGAS) [53] technologies are considered. These technologies offer amplification of the gas signal while providing two-dimensional position information.

Using a TPC as the central tracker in a linear collider environment has a number of advantages over other competing technologies. The large number of three-dimensional space-points measured in a TPC facilitates track reconstruction with a very-high precision. The TPC has a rather low material budget, thus a low multiple scattering which results in a better momentum resolution. Moreover, the low material budget is required to improve the performance of the particle flow reconstruction. The TPC offers a good specific energy loss resolution of $dE/dx \simeq 5\%$ which can be exploited for particle identification.

The TPC of ILD is a 2350 mm long cylinder with an inner radius of 329 mm and outer radius of 1808 mm. The TPC is placed in a 3.5 T magnetic field. A gas mixture of Ar-CF4-isobutane called *T2K* is a promising candidate for this detector [54]. Studies have shown that using such a system, a point resolution of better than 100 μm is achievable [34].

Silicon Tracking System

The TPC of ILD is augmented by a silicon tracking system. Two layers of silicon tracker located between the vertex detector and the TPC form the Silicon Inner Tracker (SIT) while another layer placed between the TPC and the calorimeter system constitutes the Silicon External Tracker (SET). By providing precise space points with a radial resolution of 7 μm, the SIT and SET improve the overall momentum resolution of the TPC and facilitate linking of the tracks in the TPC to the tracks in the vertex detector on one end, and calorimeter hits on the other end of the track. Single-sided silicon sensors are used in these systems.

The silicon tracking system of ILD includes two more components: the forward tracker and end-cap components behind the endplates of the TPC. The end-cap components complete the spatial coverage of the tracking system. The forward tracker is comprised of seven disks of silicon pixel and strip detectors and covers the very small angles which are not covered by the TPC.

3.3.3 Calorimeters

ILD is optimised for particle flow reconstruction. In the particle flow approach individual particles are reconstructed and energy or momentum of individual particles is measured in the subdetector with the best resolution for that type of particle. The particle flow is one of the main topics of this thesis and is discussed in the following chapters. The particle flow algorithms rely heavily on pattern recognition techniques to distinguish

overlapping showers and require calorimeter systems with an unprecedented granularity. A nearly-cylindrical barrel calorimeter and two end-cap calorimeters constitute the calorimeter system of ILD. Each of these calorimeters consist of an electromagnetic calorimeter (ECal) and a hadronic calorimeter (HCal).

The electromagnetic calorimeter of ILD should have excellent granularity in three dimensions in order to fulfil the particle flow requirements. The main duty of the ECal is to measure the energy of photons. It should be able to distinguish showers created by different photons and other nearby particles. A large fraction of hadrons start showering in the ECal, therefore, the ECal plays an important role also in hadronic shower identification and separation. The baseline ECal of ILD is comprised of 30 active layers interleaved with tungsten absorbers with a total depth of 24 radiation length. While a silicon pad diode technology is considered as the baseline for the active layers, a scintillator-based technology is also being developed as an option. The two technologies can be combined in order to reduce the cost of the detector while fulfilling the requirements [34].

The main task of the HCal is to measure the energy of the hadrons. In the particle flow paradigm, charged hadrons are measured using the tracking system and only neutral hadrons are measured in the HCal. Hence, the HCal needs to be capable of providing the necessary topological information to separate showers of charged hadrons from showers of neutral hadrons. The HCal in the baseline design of ILD has 48 active layers interleaved with steel absorbers which correspond to 6 nuclear interaction length. Two options are considered for the active layers: one uses scintillator tiles as active material and the other is based on gaseous detector technology [34].

In addition to the aforementioned calorimeters, in the very forward regions ILD is equipped with a calorimeter which serves as a beamstrahlung monitor called BeamCal and another calorimeter which is used for luminosity monitoring, known as LumiCal. These forward detectors extend the detector coverage to low polar angles down to 5 mrad.

There are a number of technologies being developed for calorimetry at the linear collider experiments. These technologies are introduced in more details in chapter 9. The base-line design of ILD, which is also implemented in the detector simulation, employs silicon technology for the ECal and scintillator technology for the HCal.

3.3.4 Tail Catcher and Muon Tracker

The iron magnetic field return yoke of ILD in the barrel and end-caps is instrumented to act as a muon detection system and a tail catcher for the calorimeters. Such a system is essential to reach the measurement precision of ILD. The system is used to measure the part of the shower energy which leaks out of the calorimeter system, hence improving the energy resolution. It also identifies muons to supplement the tracker information [34].

In the barrel, one sensitive layer is placed between the coil and the flux return yoke, followed by 13 sensitive layers inside the yoke. The first 10 sensitive layers inside the yoke are separated 14 cm apart and act as a calorimeter. The remaining 3 layers are used only for muon tracking and are spaced by 60 cm of iron. In the end-caps, there are 10 sensitive layers placed every 14 cm followed by 2 layers separated by 60 cm.

There are two technologies considered for the sensitive layers of the tail catcher and muon tracker system. The first option is to use scintillator strips coupled to Silicon Photomultipliers (SiPM) as photo detectors. The second option is to use Resistive Plate Chambers (RPCs). Both options have been successfully tested by the research and development programs [55, 56].

As shown in this chapter, the clean experimental environment at ILC as a high-luminosity lepton collider combined with the advanced detectors such as ILD allows for precision physics studies. There are a number of different approaches to Higgs boson mass measurement at ILC which are introduced in the next chapter.

Chapter 4

Measuring the mass of the Higgs Boson at ILC

According to the Standard Model of particle physics, the Brout-Englert-Higgs mechanism gives mass to the charged leptons, quarks and vector bosons by adding a scalar Higgs field that breaks the symmetries which forbid these particles to be massive. This mechanism is explained in detail in chapter 2. The discovery of a boson compatible with the Standard Model Higgs boson at LHC gives experimental evidence for this idea and opens a new door to the Standard Model and beyond Standard Model studies.

The mass of the Higgs boson is one of the fundamental parameters in the Standard Model. While the Standard Model fully describes the interactions of the Higgs boson in terms of masses and couplings of the particles it decays to, there is no theoretical prediction for the Higgs boson mass. Therefore, it needs to be experimentally measured. Once the Higgs boson mass is determined, the Standard Model predicts all its couplings to the other elementary particles. This in turn allows for prediction of the Higgs boson production cross-sections and decay branching ratios. The decay width of the Higgs boson to WW^* and ZZ^* depends on the Higgs boson mass below the WW threshold. Every 100 MeV uncertainty on the Higgs boson mass adds $0.5 - 0.6\%$ uncertainty to the ratio of the $Hb\bar{b}$ and HWW^* couplings [24, p. 24] and each 200 MeV change in the Higgs boson mass shifts BR($H \to ZZ^*$) by 2.5% [57, p. 5]. Therefore, a precise measurement of the Higgs boson mass is crucial for extraction of the Higgs boson couplings from branching ratio measurements.

A precise measurement of the Higgs boson mass allows for tests of the electroweak theory and the Standard Model. In addition it is of high importance for probing the Higgs physics. Any deviation of the Higgs boson properties from the predictions of the Standard Model would be an evidence of new physics beyond the scope of the Standard Model. There are different models for physics beyond the Standard Model. These models result in different ways of deviations from the predicted Higgs boson couplings. In models which predict a composite Higgs boson, a uniform decrease in all Higgs couplings is expected. In

models which have more than one Higgs field, such as the supersymmetric models, it is expected that the couplings to down-type quarks and leptons and the effective couplings to gluons and photons exhibit the greatest deviations. As a result, searching for deviations of the Higgs boson couplings provides a way to know if the discovered Higgs boson is a fundamental scalar particle or a composite and whether there are more than one Higgs field.

One of the goals of the ILC physics programme is to measure the individual couplings of the Higgs boson with an uncertainty of less than 1% [25]. Measurement of the Higgs boson couplings to the desired precision requires a high precision measurement of the Higgs boson mass. In addition to the Higgs boson couplings, a precise measurement of the Higgs boson mass can also improve quantum corrections to many observables and enables more stringent tests of the electroweak theory.

The Higgs boson mass is currently measured with an uncertainty of approximately 0.2% at LHC [10]. It should be noted that only a subset of the Higgs boson decay modes can be observed at LHC. Hence, reaching better than 1% resolution in Higgs boson couplings is only possible at lepton colliders such as the ILC. Absence of backgrounds from the strong processes and calculable backgrounds from the electroweak interactions in addition to a well-defined initial state at ILC, enable measurements of Higgs boson properties with an unprecedented accuracy.

There are different ways to precisely measure the mass of the Higgs boson at ILC. These measurements are explained in this chapter.

4.1 Higgs Boson Mass Measurement Methods at ILC

At ILC the Higgs boson is produced by three major processes: $e^+e^- \rightarrow ZH$ known as *Higgsstrahlung*, $e^+e^- \rightarrow \nu_e\bar{\nu}_e H$ called WW*-fusion* and $e^+e^- \rightarrow e^+e^-H$ named ZZ*-fusion*. The cross-section of single Higgs boson production by these processes as a function of the centre-of-mass energy at ILC is shown in figure 4.1. The Higgsstrahlung process is the dominant Higgs boson production process at the centre-of-mass energy of 250 GeV and decreases with increasing centre-of-mass energy. The cross-section of the WW-fusion process increases steadily with the centre-of-mass energy until it becomes the dominant process at centre-of-mass energies above ~ 450 GeV. The ZZ-fusion cross-section remains very small for the entire range.

For all the aforementioned production processes, the detectors at ILC will be able to identify all of the major Higgs boson decay modes including $H \rightarrow b\bar{b}$, $H \rightarrow c\bar{c}$, $H \rightarrow gg$, $H \rightarrow WW^*$ and $H \rightarrow \tau\tau$. There are three possible methods for a precise measurement of the Higgs boson mass at ILC which are introduced in the following.

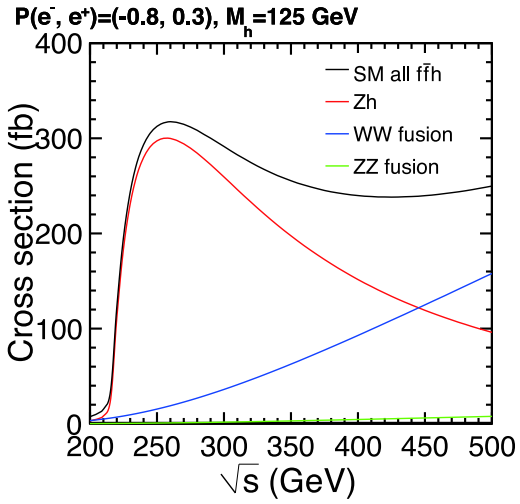


Figure 4.1 – Higgs boson production cross-section at ILC for the three major processes as a function of the centre-of-mass energy [57].

4.1.1 Model Independent Measurement

Precise knowledge of the four-momentum of the initial state and the low background environment at ILC allows for identifying Higgs events in the Higgsstrahlung process $e^+e^- \to ZH$ by reconstructing only the Z boson decay products and without looking into the Higgs boson decay products at all.

The measurement of the decay of Z boson to a pair of muons $Z \to \mu^+\mu^-$ can be used for a precise measurement of the Higgs boson mass. These decays can be reconstructed very precisely since the Z boson mass is known to a very good precision and the tracking detectors at ILC have a very high momentum resolution.

By identifying lepton pairs from Z boson decays, the mass of the particle recoiling against the lepton pair can be calculated as:

$$m_{rec} = \sqrt{(\sqrt{s} - E_Z)^2 - p_Z^2} \tag{4.1}$$

where $\sqrt{s}$ is the centre-of-mass energy and E_Z and p_Z are reconstructed energy and momentum of the Z, respectively.

The distribution of the recoil mass for $Z \to \mu^+\mu^-$ events is shown in figure 4.2. Using this method at the centre-of-mass energy of 250 GeV and with an integrated luminosity of 250 fb^{-1}, the Higgs boson mass can be measured with an uncertainty of about 40 MeV [58]. At the centre-of-mass energy of 500 GeV the uncertainty on the Higgs boson mass measured using the recoil technique is approximately 500 MeV for an integrate luminosity of 500 fb^{-1}. The reason for the deterioration of the uncertainty at higher energies is that due to ISR and beamstrahlung the centre-of-mass energy can be lower than the nominal value.

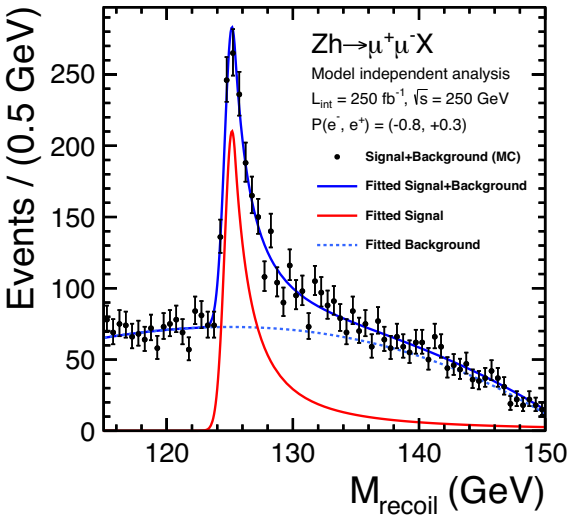


Figure 4.2 – Distribution of the recoil mass from $e^+e^- \to ZH$ followed by $Z \to \mu^+\mu^-$ for an integrated luminosity of 250 fb^{-1} at the centre-of-mass energy of 250 GeV. The input Higgs boson mass is 125 GeV [57].

4.1.2 $e^+e^- \to ZH$ with $Z \to ff$

The Higgsstrahlung process followed by $Z \to e^+e^-/\mu^+\mu^-/q\bar{q}$ and $H \to b\bar{b}/c\bar{c}/gg$ can also be employed for a precise measurement of the Higgs boson mass. In this method only the direction of the two jets produced by the decay of the Higgs boson and conservation of transverse momentum is used. In this way, the problems concerning the missing energy due to ISR and beamstrahlung are avoided.

Using this method on a $ZH \to \mu^+\mu^- b\bar{b}$ sample at the centre-of-mass energy of 500 GeV with an integrated luminosity of 500 fb^{-1} results in an uncertainty of roughly 230 MeV on the reconstructed Higgs boson mass [59].

4.1.3 $e^+e^- \to ZH$ with $ZH \to qqq'q'$

According to the Standard Model, the decay of the Z boson to a pair of quarks $Z \to q\bar{q}$ has the largest branching ratio of about 70% [19]. Thus, this channel provides much more statistics compared to the channels used for the previous methods, therefore, a better statistical uncertainty is expected.

The possibility of Higgs boson mass measurement at a linear e^+e^- collider with the centre-of-mass energy of 500 GeV has been studied for an integrated luminosity of 500 fb^{-1} [60]. Since the Higgs boson was not discovered at the time, the study was conducted for Higgs boson masses of 120 GeV, 150 GeV and 180 GeV. For the Higgs boson mass of 120 GeV, a statistical uncertainty of 45 MeV was achieved. This study was conducted using a fast simulation of the detector.

4.2 Summary and Discussion

At ILC, the mass of the Higgs boson can be measured to a high precision from different decay channels. Using the recoil technique enables a precise model-independent measurement of the Higgs boson mass. In order to collect sufficient data for this measurement, the machine should run at a low centre-of-mass energy where the Higgsstrahlung process has the highest cross-section, as shown in figure 4.1. However, low centre-of-mass energies are prohibitive for many other interesting physics cases such as top quark studies. Moreover, the instantaneous luminosity of ILC at the centre-of-mass energy of 500 GeV is more than a factor of two larger than at the centre-of-mass energy of 250 GeV. Therefore, the running time to reach a certain integrated luminosity at the centre-of-mass energy of 500 GeV is a factor of two shorter than at 250 GeV. Therefore, an alternative approach to measure the Higgs boson mass with a comparable uncertainty at higher energies is favourable. Moreover, such measurement at higher energies is also valuable for CLIC since it will not run at low energies required for the recoil mass technique [20].

The preliminary studies [59] have shown that at the centre-of-mass energy of 500 GeV, a measurement using $e^+e^- \to ZH$ with $Z \to ff$ can reach an uncertainty on the Higgs boson mass comparable to that of LHC measurements. However, due to larger branching fraction of $Z \to q\bar{q}$ and hence higher statistics, $e^+e^- \to ZH$ with $ZH \to qqq'q'$ has the potential for reaching a lower statistical uncertainty on the Higgs boson mass, as studied in [60].

Measurement of the Higgs boson mass using the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel is studied in this thesis. A fast simulation of the detector was employed for the previous study [60]. In the current study a realistic full simulation of the ILD detector concept is employed. The simulation has been benchmarked against performance of various subdetector prototypes at beam tests. In addition, a new reconstruction approach using advanced algorithms has been developed in the meantime which is exploited in this thesis.

The MC sample used in the previous study includes Standard Model backgrounds [60]. It is expected that performing the measurement using a full simulation has a small effect on the overall event selection and mainly affects the mass reconstruction. Therefore, the focus of this study is on the mass reconstruction and only signal events are used. Four quarks at the final state of this channel produce jets of particles in the detector. Since calorimetry is the most important tool in jet measurements, the concept of calorimetry and the calorimeter detectors are introduced in the next chapter, prior to discussing the main analysis.

Chapter 5

Calorimetry and Particle Flow Concept

Detectors for particle and nuclear physics experiments employ various techniques for particle detection and measurements. These techniques exploit various physical phenomena which occur when the particles interact with matter. The particle-matter interactions vary according to the fundamental properties of the particles involved. In case of charged particles, various detection techniques based on ionization have been developed. Such devices, commonly known as trackers, are able to measure particle momentum to a very high precision by using the curvature of the particle tracks in a magnetic field.

In case of photons and neutral hadrons which have no electric charge a different approach is required. The basic principle of neutral particle detection is to provide a bulk of material for the particle to interact with and measure the deposited energy of the entirely absorbed particle. This technology, commonly known as calorimetry, is widely used in high-energy physics experiments.

In this chapter, various interactions of particles with matter and basic principles of calorimetry are discussed in section 5.1. Different types of calorimeters have been developed over many years for different applications which are also briefly introduced in the same section. Basic characteristics of calorimeters such as energy resolution are presented in section 5.1.3.

In the recent years a new approach has been developed to improve overall event reconstruction and jet energy resolution of calorimeters. This technique, known as particle flow approach, is essential for the physics programme of the ILC. The basic concept of this technique is discussed in section 5.2.

5.1 Calorimetry at High-energy Physics Experiments

In high-energy and nuclear physics experiments, a calorimeter detector is an apparatus that detects and measures properties of particles by entirely absorbing them in a bulk of material. Although the main purpose of calorimeters is energy measurement, their application can be extended to position, direction and time measurements and particle identification.

The working principle of calorimeters follows a basic idea. While traversing a block of matter, charged and neutral particles interact with the matter and deposit their energy progressively through creation and absorption processes until being entirely absorbed. The deposited energy is then measured to determine the energy of the original particle.

Various techniques exist to measure the deposited energy. These techniques exploit different physical phenomena. In this section, first the main physical interactions of particles with matter which are relevant to calorimetry are discussed. The discussion is followed by a review of various calorimeter detectors and is concluded with main characteristics of these detectors.

5.1.1 Interaction of Particles with Matter

Particles traversing matter lose their energy by interacting with the matter. The type of interactions that take place depends on the properties of the particle such as charge and energy. These interactions are mostly of electromagnetic and strong natures.

Charged particles traversing a medium lose their energy by electromagnetic interactions which result in *excitation* or *ionisation* of the atoms. Excited atoms return to their ground states by radiating low-energy photons. If the energy of the particle is sufficient to liberate the atomic electrons from the Coulomb field of the atomic nuclei, it ionises the medium along its trajectory.

Electromagnetic Interactions

In the case of heavy charged particles, the Bethe equation describes the mean energy loss of the particles while traversing matter [19]:

$$\langle -\frac{dE}{dx} \rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} ln \frac{2 m_e c^2 \beta^2 \gamma^2 W_{Max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \tag{5.1}$$

where z is the charge of the incident particle, Z is the atomic number of the absorber material, A is the atomic weight of the absorber material, $\beta = v/c$ is the velocity of the incident particle, m_e is the electron mass, I is the mean excitation energy of the absorber material, W_{Max} is the maximum energy transfer in a single collision, γ is the Lorentz factor and

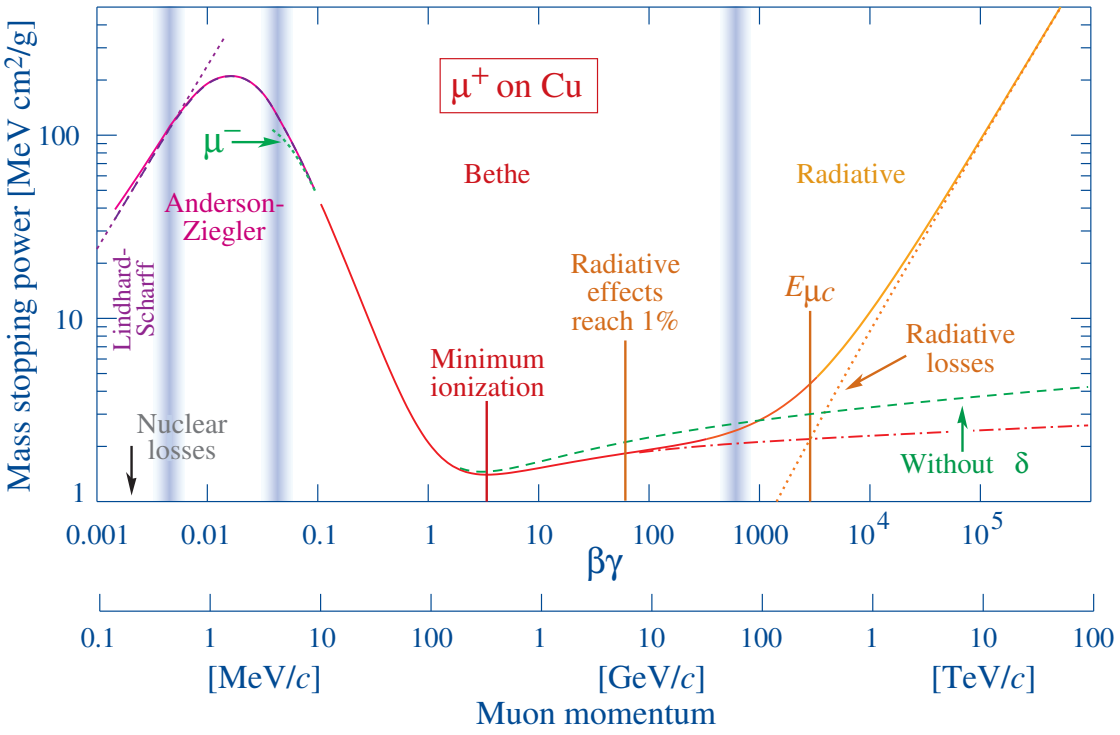


Figure 5.1 – Mass stopping power for positive muons in copper [19].

$\delta(\beta\gamma)$ is the density effect correction to the energy loss. The constant $K = 4\pi N_A r_e^2 m_e c^2$ in which N_A is Avogadro's number and r_e is the classical radius of an electron. The quantity $\langle -\frac{dE}{dx} \rangle$ is known as the *mass stopping power* and has units of $MeV\,g^{-1}cm^2$.

A plot of the mass stopping power for positive muons in copper as a function of their momentum is shown in figure 5.1. The Bethe formula describes the intermediate region between 10 MeV/c and 100 GeV/c, which is labelled as "Bethe" region. In order to properly describe slower particles, the Bethe formula of equation 5.1 needs to be corrected. These corrections mostly include terms with higher powers of particle charge z. By applying all corrections, the corrected Bethe formula can describe the energy loss of heavy charged particles, e.g. protons, down to a momentum of approximately 1 MeV.

In case of $e^{\pm}$ and photons, ionisation is the main source of energy loss at low energies. If the energy of the traversing charged particle is high enough, in the order of 100 MeV, the interaction with the Coulomb field of the atomic nuclei becomes the dominant source of energy loss. In such interactions the decelerating charged particle dissipates a fraction of its kinetic energy by radiating photons. This process is known as *bremsstrahlung*. While traversing the matter, the charged particle deviates from its original direction by undergoing multiple small-angle scatters, mostly due to Coulomb interactions with nuclei. This deviation is known as *Coulomb* or *multiple scattering*.

A useful quantity to characterise the energy loss via various interactions is the *radiation length*, X_0. One radiation length is the mean distance over which a high-energy electron loses 63.2% $(1 - 1/e)$ of its energy by bremsstrahlung, or alternatively, is $7/9$ of the mean

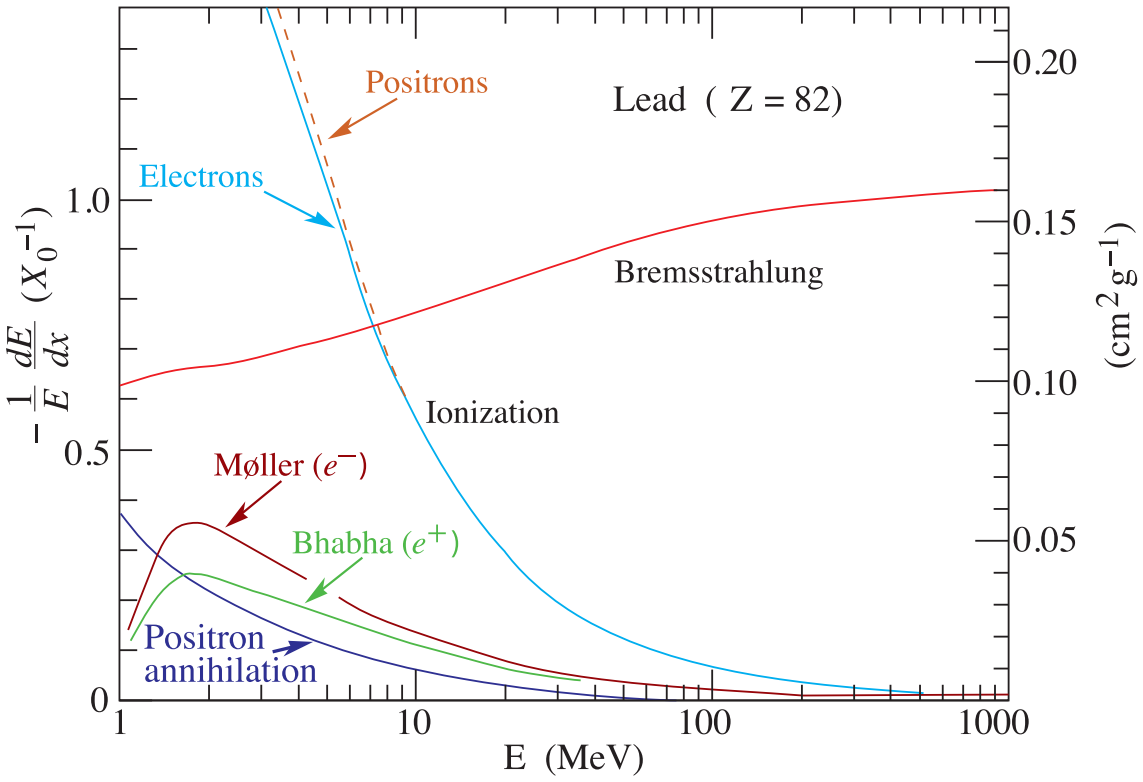


Figure 5.2 – Fractional energy loss per radiation length for $e^{\pm}$ in lead as a function of $e^{\pm}$ energy[19].

free path of a high-energy photon before undergoing pair production. The radiation length is usually measured in units of gcm^{-2}.

The fractional energy loss of an $e^{\pm}$ is shown in figure 5.2. Energy loss of an electron by ionisation is dominant at low energies and varies logarithmically with the energy. The bremsstrahlung process varies nearly linearly with the energy and is the dominant energy loss process at higher energies. The energy at which energy loss due to ionisation equals energy loss by bremsstrahlung is known as the *critical energy*, E_c. There are alternative definitions for the critical energy which result in slightly different values of E_c.

Low-energy photons lose their energy through the photoelectric effect, Compton scattering, Rayleigh scattering and photonuclear absorption, while the photoelectric effect is the dominant process. High-energy photons predominantly lose their energy by the so-called *pair production* process in which a photon annihilates in the Coulomb field of an atomic nucleus and creates an electron-positron pair.

Strong Interactions

In addition to the electromagnetic interactions, charged and neutral hadrons can undergo strong interactions while traversing matter. The strong interactions have a much larger

variety of possible processes than the electromagnetic interactions. These processes include *spallation*, *evaporation* and *fission*.

Spallation processes are one of the most probable interactions. In this process, an incoming hadron interacts with nucleons. The struck nucleon can in turn interact with other nucleons in the nucleus, leading to a cascade of energetic nucleons, pions, etc. within the nucleus. A number of these particles with energies larger than the nuclear binding energy can escape as secondary particles if they reach the boundaries of the nucleus. The particles which remain inside deposit their energy in the nucleus and bring it to an excited state.

The excited nucleus returns back to the ground state by subsequent decay processes such as multi-fragmentation, evaporation and fission [61]. If the excitation energy is larger than the separation energy, nucleons and light fragments such as α particles, deuterons and tritium can be emitted. Such an emission process is best describes as an evaporation form a hots system, hence the name evaporation process [62].

Nuclear fission is another process which contributes to de-excitation of the nuclei. In this process the excited nucleus splits into several smaller fragments. These fragments include lighter nuclei and neutrons.

5.1.2 Calorimeter Detectors

The basic principle of calorimeter detectors is to use a block of material to entirely absorb a particle and then measure the deposited energy. The number of particles that can be directly detected in a detector is limited since most of the particles have very short life times. The detectable particles can be categorised as below:

photons and $e^{\pm}$ which undergo electromagnetic interactions with the absorber material,

charged hadrons ($\pi^{\pm}$, $p^{\pm}$ and $K^{\pm}$) which interact with the absorber material via strong and electromagnetic interactions,

neutral hadrons (n, K_L) which interact with the atomic nuclei via the strong force,

muons which undergo electromagnetic interaction with the absorber material and lose energy predominantly by the ionisation process.

When a high-energy $e^{\pm}$ or photon enters a block of matter, the successive occurrence of bremsstrahlung and pair production processes creates a cascade of $e^{\pm}$ and photons with lower energies which is known as an *electromagnetic shower*. The shower development continues until the energy of the shower particles falls below the *critical energy*. At this stage, the ionisation process becomes the dominant source of energy loss (see section 5.1.1) and the number of new particles generated in the shower decreases. The depth at which no new particle is generated is known as the *shower maximum*. Some of the particle created

in the hadronic interactions such as π^0 and η develop electromagnetic sub-showers within the hadronic shower.

Heavier particles, such as muons and hadrons, have a very small cross-section for bremsstrahlung. Therefore, initialisation of an electromagnetic shower by these particles is very unlikely. However, hadrons can have a strong interaction with the atomic nuclei and subsequently initiate a so-called *hadronic shower*. The characteristic length describing the development of hadronic showers is known as the *interaction length*, λ_I, and is defined as the mean free path of a hadron before undergoing an inelastic interaction with an atomic nucleus.

Hadronic showers are much more complicated than the electromagnetic showers due to the complex nature of the strong interactions. As a result, in order to obtain best possible results, the calorimeter system of most high-energy physics experiments is comprised of two classes of calorimeters: hadronic calorimeters (HCal) and electromagnetic calorimeters (ECal). These two classes of calorimeters have different characteristics which are optimised for their purpose.

In general, there are two types of calorimeters: homogeneous calorimeters and sampling calorimeters. In sampling calorimeters, the material used to absorb particles (absorber) is distinct from the material used to generate the signal (active material). Typically, the two types of material are placed in an alternating arrangement. In homogeneous calorimeters, absorber and active materials are the same, i.e. the entire volume of the calorimeter is sensitive. The main advantage of the sampling calorimeters is that each of the two materials can be optimised for its task, without directly affecting the other. However, in the sampling calorimeters a fraction of the energy is deposited in the non-sensitive material and is lost. This introduces additional fluctuations which deteriorates the energy resolution. Calorimeters are usually segmented laterally and longitudinally into smaller cells in order to provide spatial information.

In order to determine the energy of an incident particle, a calorimeter basically measures the number of particles created in the shower generated by the incident particle. There are various techniques to perform this measurement. Scintillator material coupled to photo-detectors, silicon detectors and gaseous or liquid ionisation detectors can be used for this purpose. A number of calorimeters based on these detection techniques are presented in chapter 9.

5.1.3 Characteristic Properties of Calorimeters

A wide variety of calorimeters have been developed for different applications. Most of the calorimeters are optimised for the physics requirements of the experiment [63]. There are a number of factors which are common to all calorimeters and can be used to evaluate their performance.

Energy Resolution

The main task of calorimeters is energy measurement. The uncertainty with which the energy of an incident particle or a jet is measured is an important characterisation factor for calorimeters. There are three main sources of uncertainty which limit the precision of the calorimeters:

Statistical fluctuations this category includes statistical effects which are due to intrinsic fluctuations in shower development, sampling fluctuations, photoelectron statistics and dead material before and in the calorimeter. The contribution of the statistical fluctuations to the total relative energy resolution is typically in the order of a few percent for homogeneous and $10\%/\sqrt{E}$ for sampling electromagnetic calorimeters. The contribution is typically at the level of $60\%/\sqrt{E}$ for hadronic calorimeters [19].

Systematic uncertainty Calibration uncertainty and uniformity of the calorimeter are the main sources of systematic uncertainty in the energy measurement by calorimeters. Other factors such as damage due to radiation and difference in signals generated for electromagnetic and hadronic showers can also contribute to the systematics. Typically, the contribution of the systematic uncertainty to the relative energy resolution is around 1% [19].

Electronic Noise The intrinsic noise of the read-out electronics is another source of uncertainty on the energy measurement in calorimeters which depends on the read-out technique and circuitry. The higher the energy of a particle, the higher the generated signal. As a result, for energetic particles the electronic noise is negligible.

Based on the aforementioned sources of uncertainty, the total relative uncertainty on the measured energy can be parametrised as:

$$\frac{\sigma_E}{E(GeV)} = \frac{a}{\sqrt{E(GeV)}} \oplus b \oplus \frac{c}{E(GeV)} \tag{5.2}$$

where σ_E is uncertainty on the energy, E is the energy of the particle or jet, a is known as the stochastic term and represents the statistical fluctuations, b represents the systematics and c represents the electronic noise. The stochastic term is proportional to the inverse of the square-root of the energy while the noise term is inversely proportional to the energy. The systematic term has no energy dependency.

Spatial Resolution

As mentioned earlier in this chapter, in addition to energy measurement, calorimeters are often used for position and direction measurements, especially for neutral particles for which there is no tracker information. This information can be exploited for particle identification or particle-particle separation when several particles enter the calorimeter in close vicinity

of one other. Therefore, position resolution and angular resolution of calorimeters are important factors for the evaluation of the calorimeter performance, especially in the context of particle flow reconstruction.

A common way to determine the impact position of a particle is to find the average position of energy deposition in the calorimeter cells known as the *centre-of-gravity*:

$$x_c = \sum_i \frac{x_i E_i}{\sum_i E_i} \quad , \quad y_c = \sum_i \frac{y_i E_i}{\sum_i E_i} \tag{5.3}$$

where (x_c, y_c) is the coordinate of the centre-of-gravity, (x_i, y_i) is the coordinate of the cell i and E_i is the energy deposited in cell i.

The energy terms in equation 5.3 are subject to the energy resolution of equation 5.2. Consequently, the uncertainty on the position determination depends on the uncertainty on the energy measurement and better energy resolution results in better position resolution. Moreover, the (x_i, y_i) is the coordinate of a single point in a cell. Therefore, the energy of the particle is considered to be deposited at (x_i, y_i) regardless of the actual impact position on the cell. As a result, the positions determined using the centre-of-gravity method tend to be systematically shifted, especially for a small number of cells. Therefore, increasing lateral segmentation of the calorimeter improves the position resolution.

In order the determine the direction of a particle using a calorimeter, its position should be determined at more than one point. This requires longitudinal segmentation of the calorimeter in order to be able to determine the position at multiple depths. Higher longitudinal segmentation results in a better angular resolution.

Timing Resolution

Timing information from calorimeters can be used in addition to the energy and position information to enhance the signal to noise ratio. It can also be employed for pileup and background removal. The uncertainty on the calorimeter time measurements depends on multiple factors. The time of the first interaction of a particle after entering the calorimeter fluctuates on an event-by-event basis. The time period of shower development and signal generation also varies per event. In addition, the read-out electronics and data acquisition system have a characteristic time constant. All these factors contribute to the total timing resolution of a calorimeter.

The timing resolution can be improved by performing multiple independent measurements. Hence, the timing resolution of an entire shower created by a high energy particle is better than that of a shower created by a low energy particles. Timing resolution can also be improved by a finer segmentation of the calorimeter.

5.2 Particle Flow Approach to Event Reconstruction

High-energy physics experiments employ calorimeters to measure the energy of single particles as well as of jets. The classical approach to jet energy measurement is to sum up the total energy of all charged and neutral particles of a jet deposited in the electromagnetic and hadronic calorimeters. The uncertainty on the jet energy measurement using this approach follows the form of equation 5.2. In case of jets, the stochastic term a typically amounts to $50 - 100\%$ and the other two terms are a few percent.

The physics programme of the future experiments at a linear e^+e^- collider requires a jet energy resolution far beyond what can be achieved by the classical approach to jet energy measurement. Over the past decade, a new approach has been developed to address this issue. The overall concept of this new approach is presented in this section while its calorimeter requirements and application at the ILC are discussed in the following chapters.

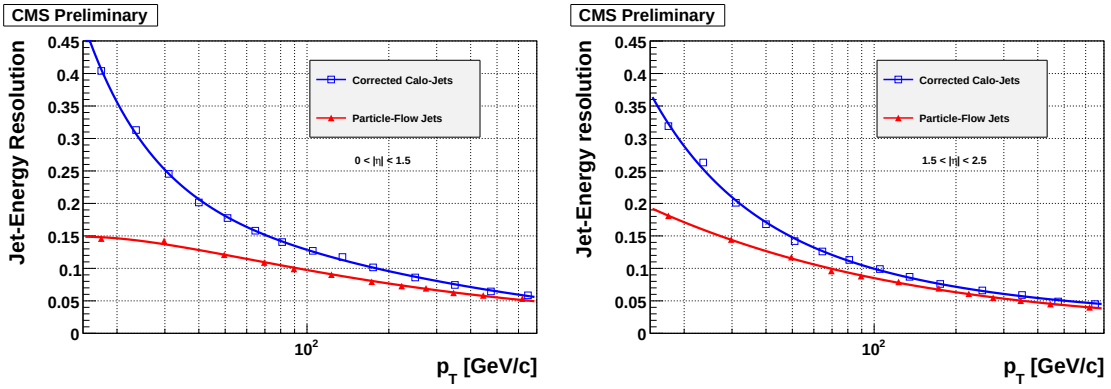
5.2.1 The Particle Flow Concept

A jet is a spray of well-collimated particles created as a result of the hadronisation of a coloured particle. The particle composition of jets has been studied in details at the LEP experiments. These studies show that on average approximately 62% of jet energy is carried by charged particles, 27% by photons and 10% by long-lived neutral hadrons. The remaining $\sim 1\%$ is carried by neutrinos (see chapter 7 for more details).

The charged particle content of jets is mainly composed of hadrons. Therefore, in the classical reconstruction approach where the energy of jets is meaured by summing up the energy depositions in the calorimeters, roughly 72% of the jet energy is measured by the HCal and the remaining by the ECal. Due to the reasons mentioned in section 5.1, the HCal has a poor energy resolution. Consequently, the poor resolution of the HCal is the dominant source of uncertainty on the jet energy measurement.

For many years, the experiments at LEP and other collider experiments used the tracker momentum information to replace the energy deposited in the calorimeters by charged particles. This approach, known as *energy flow*, improves the uncertainty on the jet energy measurements but due to limited granularity of calorimeters, a one-to-one association of calorimeter clusters and tracks was usually not possible. Nevertheless, to fulfil the requirements of the physics programme of the ILC and other future colliders, a greater improvement is necessary.

Over the last decade, the concept of energy flow reconstruction has been extended to the *particle flow* approach which optimises the jet energy resolution by reconstruction of each particle individually. In this approach, the jet energy is the sum of energies of all particles in the jet and the best available measurement for each particle is used: all charged particles are measured in the tracker, photons in the ECal and long-lived neutral hadrons in



(a) Barrel (b) End-cap

Figure 5.3 – CMS jet energy resolution as a function of transverse momentum [64].

the HCal. In this case, only long-lived neutral hadrons are measured with the poor energy resolution of the HCal.

The particle flow approach requires the reconstruction of all visible particles in a jet which means energy depositions in the calorimeter system should be correctly assigned to individual particles. Achieving this goal requires highly-segmented calorimeters and sophisticated reconstruction algorithms.

The Compact Muon Solenoid (CMS) experiment at the LHC employs the particle flow approach in jet energy measurements. Its large silicon tracker and segmented electromagnetic calorimeter enables CMS to exploit the particle flow techniques to improve the detector performance, as shown in figure 5.3. However, CMS can not benefit from the particle flow approach to the full extent since its electromagnetic and hadronic calorimeters are not sufficiently granular and the energy resolution of its hadronic calorimeter is very poor.

Confusion

The particle flow approach requires sophisticated pattern recognition algorithms to properly assign the energy deposited in calorimeters to individual particles. For the particle flow algorithms (PFA) to work properly, electromagnetic and hadronic calorimeters should be capable of providing sufficient spatial and topological information, which requires unprecedented longitudinal and lateral segmentation.

The PFAs are not perfect and can not always assign energy depositions to correct particles without ambiguity. This ambiguity, known as *confusion*, is an additional source of uncertainty on the jet energy measurement. The main sources of confusion can be categorised as:

- hits from charged hadrons that are reconstructed as a neutral hadron cluster,

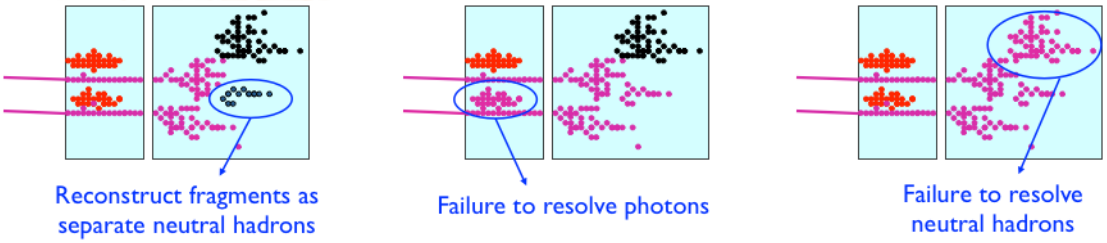


Figure 5.4 – Main sources of confusion in particle flow algorithms [40].

- hits from photons which are lost in charged hadrons,
- hits from neutral hadrons which are lost in charged hadron clusters.

These three different types of confusion are depicted in figure 5.4. If a charged hadron cluster has fragments which are separated from the main cluster, there is a probability for these fragments to be disassociated from the charged particle track and misidentified as a separate neutral hadron. The second type of confusion is due to failure to resolve photons when a photon cluster is close to a charged hadron track. The third type of confusion occurs if there is a neutral hadron cluster in a close vicinity of a charged hadron cluster. In this case, the particle flow algorithm might consider the two clusters as one, corresponding to the charged hadron. The confusion affects the reconstructed jet energy by causing energy loss due to failure to resolve the overlaps or by double counting the energy.

5.2.2 Implementation in PandoraPFA

PandoraPFA [40, 65] is a particle flow algorithm which was developed to study and employ the particle flow approach to calorimetry at the ILC. PandoraPFA reconstructs individual particles in an event by combining information from highly-granular calorimeters and the tracking system and processing them in a multi-stage process.

The tracking system is one of the sources of information about an event. In the first stage, PandoraPFA tracks are identified and classified based on their possible origin. The calorimeter hits are another source of information. PandoraPFA uses this information to form calorimeter hit objects which include information about the position and energy deposition of the hit as well as the geometry of the detector and hit density in the neighbouring region of the calorimeter. Isolated hits are removed at this stage. In the third stage selected calorimeter hits are clustered using a cone-based algorithm. First the photon clusters are identified and then the remaining hits are clustered. This clustering stage is designed in such a way that it splits up true clusters rather than merging clusters of more than one particle into one cluster. Topological signatures in the highly granular calorimeters are then used in the next stage to combine these clusters.

While an algorithm including the stages mentioned above performs well for jets with energies below 50 GeV, its performance degrades for jets of higher energies since the

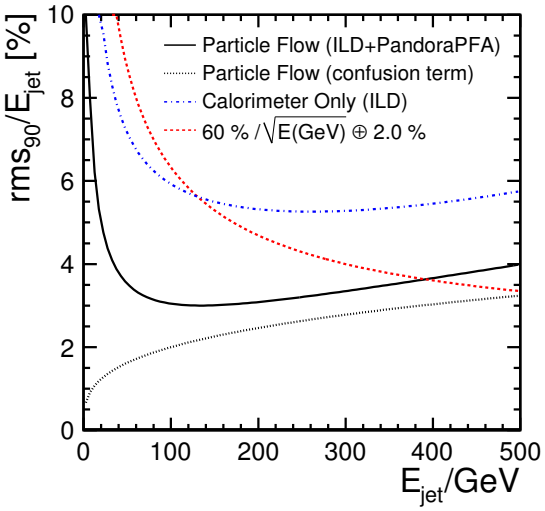


Figure 5.5 – Jet energy resolution obtained by PandoraPFA. The estimated contribution from the confusion and the typical jet energy resolution obtained by the traditional calorimetry approach are also shown [40].

overlap between hadronic showers from different particles increases with energy. Therefore, extra processing stages are required. In the next stage, PandoraPFA compares the cluster energy with the associated track momentum. If the energy and the momentum are inconsistent, re-clustering using different parameters is attempted until the energy and momentum match.

The next stage of PandoraPFA is to identify photon clusters using a sophisticated algorithm which exploits shower-profile information. At the end, an attempt to identify fragments of hadronic showers from charged particles which are misidentified as neutral clusters is made and these clusters are merged into the appropriate parent cluster. The details of these stages and other stages which are not mentioned there are explained in [65].

The estimated jet energy resolution obtained in ILD using PandoraPFA is displayed in figure 5.5. This figure shows that the jet energy resolution obtained by PandoraPFA is better than the results of the traditional approach, and the ILD relative jet energy resolution goal of $3-4\%$ is reached over a significant range of the jet energies.

The particle flow reconstruction enables jet energy measurements with an unprecedented resolution. Many of the physics analyses at ILC benefit from it for precise measurements. The effect of the particle flow reconstruction with PandoraPFA on the uncertainty of the jet energy reconstruction and its influence on the measurement of the Higgs boson mass will be discussed in the following chapters.

Chapter 6

Higgs Boson Mass Measurement at ILD - Simulation and Event Preparations

Event generation and simulation of the detector response plays a key role in any high energy physics experiment. In the early stages of a project, these simulations are used to explore possibilities that are offered by a particular detector design and its limitations. Simulations are also used to determine requirements needed for reaching the goals of a project. During the data taking phase and afterwards, simulations are used to model physics processes in order to design specific measurements and also to compare measurement results to theory expectations and to determine detector efficiencies.

In this chapter various steps and software tools which have been used to prepare the data sample for this analysis are introduced and briefly discussed. The iLCSoft software framework which is used in various stages of the analysis is briefly introduced in section 6.1. The simulation chain starts with the event generation using software packages that perform detailed cross-section calculations. Section 6.2 introduces the software packages and procedures which were used to generate the MC samples used in this analysis. The generated events then undergo a full GEANT4-based simulation of the ILD detector concept and digitisation as described in section 6.3. Section 6.4 explains how the events are reconstructed after detector simulation. Additional steps of data sample preparation before the main physics analysis are presented in section 6.5. All the software packages used for detector simulation, event reconstruction, event preparations and physics analysis are part of a large collection of software packages for linear colliders named iLCSoft.

6.1 iLCSoft Software Framework

Various software tools which are developed by the Linear Collider community are collected and managed within the iLCSoft project [66]. The iLCSoft package provides a complete framework which can be used for Monte Carlo studies and experiments. For instance,

concept studies for the ILD detector and analysis of data from beam tests of subdetector prototypes are performed within the iLCSoft framework.

Most of the tools in the iLCSoft project use a common Event Data Model (EDM) named Linear Collider I/O(LCIO), which provides a persistent solution for simulations and physics analysis [67]. This data model accounts for all possible data which can be produced by collider detectors. Using LCIO as a common EDM and file format, various detector development concepts and physics analysis groups are able to share software tools and algorithms.

The iLCSoft project provides Modular Analysis and Reconstruction for the Linear Collider (Marlin) as a modular C++ framework for event reconstruction and physics analysis [68]. Marlin uses LCIO as a transient data model and is configured using steering files in XML format. Using tools which are provided by the Marlin framework, users are able to develop custom modules which are tailored for their application and run it along other modules in a plug&play manner.

There are many reconstruction and analysis tools in the iLCSoft package which are accessible at the project website [66]. For the case of this thesis, iLCSoft version V01-17-08 was used for event reconstruction and version V01-17-11 for the analysis.

6.2 Event Generation

The first step in the simulation chain of high-energy physics experiments is to produce four-vectors of particles produced in a collision. This task is performed by event generator software packages. Event generators play the role of high-energy collider machines in the virtual world of computers. These tools take machine parameters such as centre-of-mass energy, polarisation, particle type etc. as input (see table 3.1 for the main parameters of the ILC base-line design) and by performing detailed cross-section calculations, generate a set of outgoing particles resulting from the collision of two beam particles. Due to the probabilistic nature of quantum mechanics, it is not possible to know the exact outcome of an event beforehand. Therefore, in event generator software packages Monte Carlo techniques are devised to reproduce quantum mechanical probabilities of an event.

For the case of this study, events produced by the ILC generator group for the TDR are used [34]. The events were generated by WHIZARD [69] version 1.95. WHIZARD is a general event generator for various kinds of scattering and decay processes at high-energy lepton and hadron colliders. Internally, this program uses the Optimized Matrix Element Generator (O'Mega) [70] for multi-leg tree-level matrix element generation.

The beam particles interact with the electromagnetic field of the oncoming beam and lose part of their energy by radiation which is known as beamstrahlung (see section 3.2.6). WHIZARD is able to consider the beam-spectrum of the ILC including beamstrahlung and

Initial State Radiation (ISR). Beam polarisation which, as described in subsection 3.2.2, is one of the powerful tools at the ILC, can be specified as well. While WHIZARD 1.95 includes a dedicated package with a few presets for a few electron-positron collider designs, for generation of the data samples for the current study, output of the Guinea-Pig [71] program was used as an alternative to generate e^+e^- beam spectra and beamstrahlung $\gamma\gamma$ spectra which match the ILC design. The generated spectra were manually passed on to WHIZARD [24].

WHIZARD 1.95 does not perform hadron decays and hadronisation internally and uses the Pythia Monte Carlo package [72] for these purposes, in particular Pythia version 6.4. The parameter set of Pythia was tuned to recommendations by the OPAL experiment at LEP. The difference between parameters recommended by OPAL and the default Pythia setting is explained in [34]. Particle multiplicities in the default settings and OPAL settings for a few selected species are shown in table 6.1. One of the main difference is in the number of neutral long-lived hadrons which is overestimated by the standard settings as compared to the observations at LEP. This parameter directly influences jet energy resolution due to the fact that neutral hadrons are measured with the least precision.

Table 6.1 – Predicted average number of various particle species in $e^+e^- \to q\bar{q}$ with $q = u, d, s$ collisions at the centre-of-mass energy of 500 GeV for default Pythia settings and OPAL settings [34].

	Standard tune	OPAL tune
All charged	37.4267	37.4975
π^0	17.2502	17.7834
$\pi^\pm$	31.1060	32.3830
$K^\pm$	3.7395	3.2706
P	2.5812	1.8439
N	2.5109	1.7778
K^0_S	1.8006	1.6120
K^0_L	1.8069	1.0164

6.2.1 Signal Processes

The aim of this study is to assess the precision of the Higgs boson mass measurement using the Higgsstrahlung process with hadronic decays of the Z and the H bosons. The leading order Feynman diagram for this process is illustrated in Fig. 6.1.

The signal events were generated for centre-of-mass energies of 350 GeV and 500 GeV. For each of the centre-of-mass energies, beam polarisations of $P_{e^-,e^+} = (-0.8, +0.3)$ was considered. The Higgs boson mass was set to 125.0 GeV. Since $H \to b\bar{b}$ has the largest branching fraction of Higgs decays [19], it has the potential for minimum statistical uncertainty and hence is the main interest of this study. The total numbers of generated

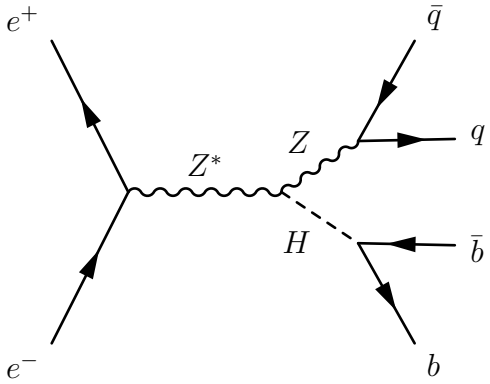


Figure 6.1 – Leading order Feynman diagram of the signal process $e^+e^- \to ZH \to q\bar{q}b\bar{b}$.

Table 6.2 – Total number of generated events for the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ process for each centre-of-mass energy and their corresponding integrated luminosities

Energy	Polarisation	Number of Events	Luminosity
350GeV	$(-0.8,+0.3)$	$132,300$	$1000\ \text{fb}^{-1}$
500GeV	$(-0.8,+0.3)$	$66,000$	$1000\ \text{fb}^{-1}$

events for each centre-of-mass energy and their corresponding integrated luminosities are shown in table 6.2.

6.2.2 Background Processes

There are two types of background events relevant for the Higgs boson mass measurement:

- other processes which have the same final state particles content as the signal and
- additional particles in the event which are not originating from the process of interest.

Other processes such as $e^+e^- \to ZZ$ and $e^+e^- \to W^+W^-$ can create final states containing four jets. Event selection in the presence of this type of backgrounds has already been investigated in the previous study with fast simulation [60]. Since these events are well modelled in fast simulations, minor changes are expected in full simulations. Thus, since the goal of this study is to estimate the best possible measurement of the Higgs boson mass, the focus in this analysis is set on how well the mass peak position can be understood. Once a full understanding of the signal events is obtained, the effect of background events can be studied and the effect of background reduction techniques can be evaluated.

As explained in section 3.2.6, interactions of bunches produce a large number of electron-positron pairs and photons. Interactions of the photons created in this process can create hadrons with low transverse momentum [39]. This effect can be an important source of background of second type above. The rate of hadron and jet production from this process depends on the beam parameters. For ILC running at the centre-of-mass energy

of 350 GeV on average 0.4 $\gamma\gamma \rightarrow hadrons$ events per bunch-crossing are expected while at the centre-of-mass energy of 500 GeV it raises to 1.2. In this thesis, the impact of the $\gamma\gamma \rightarrow hadrons$ events in the context of the Higgs boson mass measurement is studied.

The $\gamma\gamma \rightarrow$ hadron events were generated according to the cross section model given in [39]. Pythia was used to generate $\gamma\gamma \rightarrow hadrons$ background events for $\sqrt{s_{\gamma\gamma}} > 10$ GeV. For other $\gamma\gamma$ centre-of-mas energies, a phase space particle production model was employed [31]. These events are overlaid on the signal events according to a Poisson distribution with the mean equal to the average $\gamma\gamma \rightarrow hadrons$ events per bunch-crossing. During the mass-production of the sample for the centre-of-mas energy of 500 GeV, a wrong value of 1.7 for the average was used which overestimates the background.

Errors in the $\gamma\gamma$ removal procedure, which is discussed in detail in section 6.5.2, affect the signal jets and in turn the measured Higgs boson mass. In order to investigate the performance of the $\gamma\gamma$ removal procedure and its effect on the reconstructed Higgs boson mass, the analysis is performed on two samples of 500 GeV centre-of-mass energy with and without the overlay.

6.3 Event Simulation

The four-vectors of the collision products as provided by the event generator are passed on to the next step in the processing chain, the detector response simulation. In this step, interactions of the particles in the detector and the signals that they will produce are simulated. The detector simulation is crucial for detector design and optimisation, calibration, event reconstruction and evaluating physics analysis methods. Software tools which are able to simulate the interactions of particles with matter are needed for the detector simulation. The user should be able to describe geometrical volumes of the detector and materials which fill these volumes.

GEANT4 [73, 74] is a powerful particle-matter interaction simulation toolkit which provides these functionalities. This toolkit provides facilities for creating a virtual detector by describing geometries, materials and whether the geometrical volumes are sensitive parts of the detector or are dead materials. Once the detector is described, GEANT4 can propagate particles through the detector and track them until they are entirely absorbed or have left the detector. While the particles are being propagated, their physical interactions with the detector material are modelled and recorded. Models to describe the physical interactions of interest are defined by means of various physics lists. In addition, GEANT4 also provides visualisation and persistency tools.

A full GEANT4 simulation of the ILD detector is implemented in the MOKKA [75] software package that is included in iLCSoft. MOKKA has a realistic description of the ILD subdetectors including mechanical support structures, gaps and other non-instrumented material such as electronics and cabling. This level of details in the detector simulation al-

Table 6.3 – ILD detector options

Option	HCAL Technology	ECAL Technology
ILD_o1_v05	AHCAL	Si-ECAL
ILD_o2_v05	SDHCAL	Si-ECAL
ILD_o3_v05	AHCAL	Sc-ECAL

lows for detector validation and optimisation studies and physics analysis. The simulations used in this study have been performed for the TDR using MOKKA version 08-00-03 [34].

As presented in section 9.2, various calorimeter technologies are being studied for the ILC detectors. MOKKA offers multiple combinations of these technologies for ECal and HCal in terms of different detector options. These options and their corresponding combinations are listed in table 6.3. In this study option ILD_o1_v05 is used. This option employs the Si-ECAL which is a sampling calorimeter with tungsten absorbers and silicon active layers for the electromagnetic calorimeter, and the AHCAL sampling calorimeter technology with steel absorbers and scintillator active layers for the hadronic calorimeter (see chapter 9). To model the hadronic interaction of particles in the detector, QGSP_BERT [74] was used as the physics list in the simulation.

The outputs of the detector simulation are used as inputs to the digitiser program. A digitiser is a software to model how the interactions of particles in sensitive detectors are converted to the response of the detector readout electronics. Although digitisation is usually considered as part of the detector simulation, in iLCSoft it is performed as the first step of the event reconstruction due to technical reasons.

6.4 Event Reconstruction

The output of the digitisation step is similar to electronic signals which are recorded by the DAQ of a real experiment detector. At this stage, these signals should be interpreted to determine the original particles that passed the detector and their characteristics. This procedure is known as event reconstruction. Reconstruction of an event is mainly performed by track fitting and pattern recognition algorithms. Reconstruction usually involves multiple stages, in such a way that the output of earlier stages is used as input to later stages of the process. In addition to the detector hits, detector calibration databases and geometry databases are the other important sources of information in event reconstruction. The final output of the reconstruction is a collection of objects that correspond to properties of particles measured in the detector.

iLCSoft version v01-17-08 and the corresponding Marlin framework version are used for the reconstruction of the events for this study. In the first stage of reconstruction, track segments in individual tracking detectors (see section 3.3.2) are identified by means of pattern recognition algorithms. Then using a Kalman filter, a global track fit is performed

and trajectories of charged particles in the tracking detectors are identified [76]. Origin, direction of motion, charge and momentum of each charged particle can be inferred from its corresponding track.

The fitted tracks and calorimeter hits are then used as input information for the particle flow reconstruction of the event. The particle flow approach to event reconstruction is explained in section 5.2. An implementation of this approach is developed in the PandoraPFA project [40]. It has been a central component in demonstration of the ILC detector goals. Event reconstruction in PandoraPFA is performed in a complex multi-stage process, as explained in section 5.2. At the end of event reconstruction, collections of *particle flow objects* (PFOs) are formed which include the information about the reconstructed objects.

6.5 Event Preparation

Prior to the main analysis a few preparatory processing steps are performed on the reconstructed data samples. First for the sample with $\gamma\gamma \rightarrow hadron$ overlay, a procedure to remove the overlay is followed. Then, for all samples, the hits are clustered into jets. The details of these preparations are presented in this section.

6.5.1 Jet Finding

On tree-level, there are four primary quarks in the final state of the $e^+e^- \rightarrow ZH \rightarrow q\bar{q}b\bar{b}$ signal events which fragment and hadronise into hadronic jets. In addition, as discussed in section 6.2.2, additional particles are expected from $\gamma\gamma \rightarrow hadron$ background processes which might deteriorate the jet reconstruction. Therefore, jet finding plays an important role in the event preparations.

An implementation of the *FastJet* software package [77] is available in iLCSoft. FastJet includes many efficient implementations of jet finding algorithms. One of the widely used algorithms is the so-called longitudinally invariant k_t sequential jet clustering algorithm [78]. The k_t algorithm computes the k_t distance for each pair of particles i and j and the distance of each particle to the beam (z-axis) d_{iB}:

$$d_{ij} = min(p_{ti}^2, p_{tj}^2)\Delta R_{ij}^2/R^2 \tag{6.1}$$

$$d_{iB} = p_{ti}^2 \tag{6.2}$$

where p_{ti} is the transverse momentum and $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ in which y_i is the rapidity and ϕ_i is the azimuth of the particle i (see section 3.3.1 for the ILD coordinate system). R is a parameter of jet-radius and usually is taken to be of order of 1.

The two aforementioned distances can be used in two different ways in the k_t algorithm. In the *exclusive* mode, the minimum (d_{min}) of all d_{ij} and d_{iB} is found. If the d_{min} is a d_{ij},

then particles i and j are merged into a single particle and their four-momenta are summed up. This process is known as *E-scheme* recombination. If d_{min} happens to be a d_{iB} then the particle i is considered to be from beam background and is discarded. This process is repeated until all of the d_{ij} and d_{iB} are above some cut-off value.

The *inclusive* mode is similar to the *exclusive* mode except that if d_{min} is a d_{iB}, then particle i is declared to be a final jet and is removed from the list of particles. In this mode, this operation repeats until there is no particle left.

The following subsections explain how this algorithm and its variants are used in the current analysis.

6.5.2 $\gamma\gamma$ Overlay Removal

As explained in section 6.2.2 the interaction of photons can be a potential source of background at the ILC and this effect is included in the data sample by overlaying hadrons from the $\gamma\gamma$ processes on the events. A simple method to remove this overlay as part of the event preparations is explained here.

The hadrons that are created in the interaction of beam photons are very close to the beam direction with very small transverse momenta. Therefore, they represent a topology which propagates along the beam line in the detector. Considering this characteristic behaviour, a strategy for removing the overlaid hadrons is to use an *exclusive* jet finding algorithm to detect and remove these particles. This strategy, which was originally developed and used at hadron colliders to remove the remnants of the intersecting beam hadrons [77, 79], is applied in this study.

The *exclusive* mode of the longitudinally invariant k_t algorithm implemented in the FastJet package is used for $\gamma\gamma$ background removal. This algorithm takes the jet radius parameter R, and total number of expected jets as inputs. Four jets are expected from $ZH \rightarrow q\prime\bar{q}\prime q\bar{q}$ events and two additional jets are considered since previous analyses have shown that the removal procedure works best if more jets than the number of original partons are requested [80]. This prohibits false clustering of hard gluon radiation into the beam jets. Therefore, total number of 6 expected jets are used as input to the k_T algorithm.

Multiple values are tested for the jet radius parameter R. The effect of this parameter on the measurement of the Higgs boson mass is presented in section 8.5.3.

6.5.3 Forming Jet Objects

In the last step of data preparations, particles in each event are clustered into four jets. For the cases with $\gamma\gamma$ overlay, the 6 jets formed by the removal procedure are first unclustered

prior to this step. To form the jet objects for the analysis a modified version of the k_t algorithm is used which is specially designed for e^+e^- colliders and is often referred to as *Durham* algorithm [81]. The Durham algorithm has only a single distance:

$$d_{ij} = 2min(E_i^2, E_j^2)(1 - \cos(\theta_{ij})) \tag{6.3}$$

where θ_{ij} is angle between two particles i and j and E_i and E_j are their corresponding energies. Like the original k_t algorithm, the *E-scheme* is used to merge particles with minimum d_{ij} and this process is repeated until all particles are clustered into a number of jets which are requested.

Each of the jets which are found at this stage are then stored as a `ReconstructedParticle` object in the LCIO format. These objects include four-momenta of the jets. It is also possible to access the properties of individual particles which are clustered in a jet.

At the end of the event preparation procedure, the events are ready for the final analyses. Jet energy resolution is an important parameter in the measurement of the Higgs boson mass using the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel. A new method to estimated the jet energy resolution on a jet-by-jet basis is presented in the following chapter. The outcome of this method is then used in the Higgs boson mass measurement analysis.

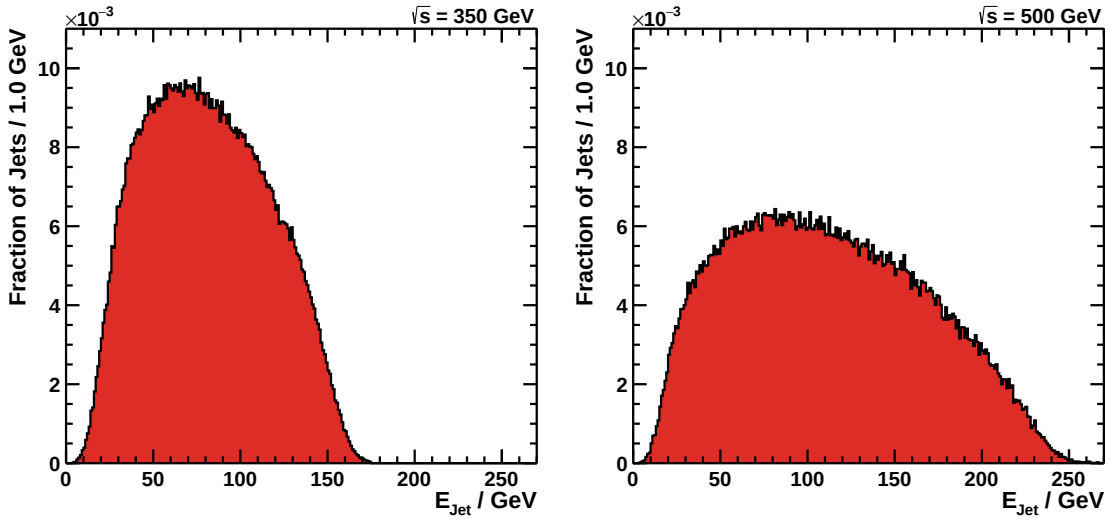
Chapter 7

Jet-Specific Energy Resolution

Particles which carry a colour charge can only exist in colourless states due to the QCD confinement, as explained in section 2.3. When a coloured particle is created, it undergoes hadronisation and creates a spray of well-collimated hadrons known as jets. The jets can be observed experimentally in particle detectors, and by analysing them properties of the original coloured particle can be inferred. Jet studies are also valuable to evaluate the current understanding of the underlying theory of strong interactions.

Jet measurements are essential to measurements which include hadronic final states. For many physics analyses, the key parameters of a jet are its energy and angles. Due to the intrinsic QCD effects such as parton showering and hadronisation, or limitation of the measurement instruments, the energy of a reconstructed jet differs from the true energy of the originating parton. These deviations of the measured energy from the true energy of the parton have direct impact on the results of physics analyses. Therefore it is of great importance to understand these deviations and keep them under control.

There are at least four jets in the final state of the signal process $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ of this study in which the Higgs boson mass is determined by computing the invariant mass of two jets. Invariant masses are also important to distinguish between the signal and backgrounds. Therefore the resolution with which the energies of these jets are measured is crucial for the purpose of this study. This chapter is concerned with the resolution with which the energy of a jet can be measured at ILD. The expected resolution varies from jet to jet due to the varying fractions of charged and neutral particles within the jets. In section 7.1, a new method for estimating the energy resolution of each individual jet is introduced. This method, called *error flow*, benefits from the ability of the particle flow approach to reconstruct single particles within a jet and estimates the energy resolution for each jet individually based on the particles content of that jet. The error flow method is presented in section 7.2. The subsections discuss various contributions to the total jet energy resolution. The estimated total uncertainty on the jet energy measurements is presented in section 7.3.



(a) (b)

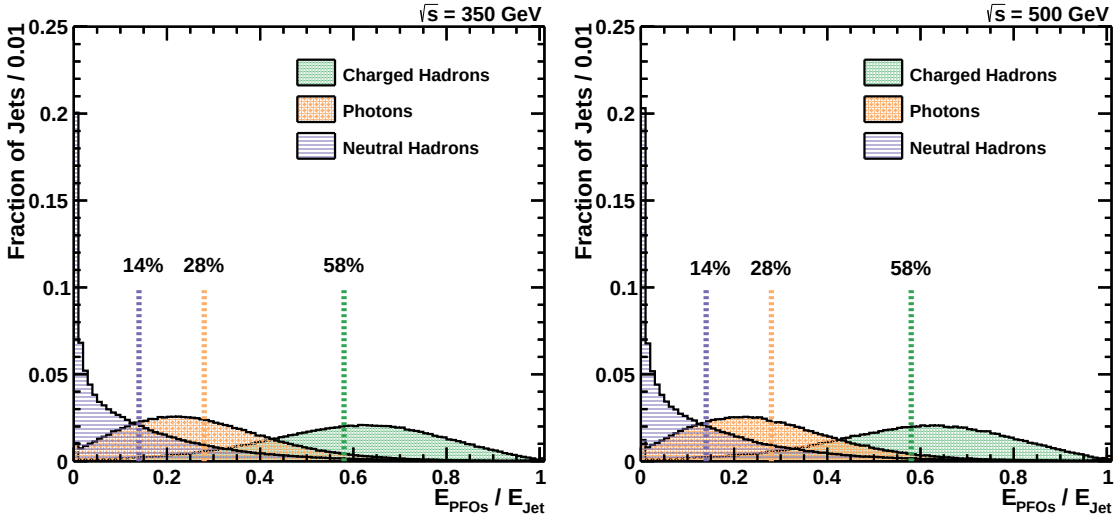
Figure 7.1 – Distribution of the reconstructed jet energies in the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ Monte Carlo samples for the centre-of-mass energies of a) 350 GeV and b) 500 GeV.

7.1 Jet Energy Distribution and Jet Constituent Multiplicity

In order to be able to develop a method for estimating the uncertainty on the measured energy of each jet individually, an insight into the particle content of jets is essential. The data sample prepared for the measurement of the Higgs boson mass (see section 6.2.1) is used to analyse the internal structure of jets. The results are presented in this section.

As explained in section 6.5.3, a jet clustering algorithm is used to cluster all particles in an event into 4 jets. The distributions of reconstructed energies of these jets are shown in figure 7.1 for centre-of-mass energies of 350 GeV and 500 GeV. In both cases, the energies of the jets range from about 10 GeV to half of the centre-of-mass energy. Since the available energy is higher than the total rest mass of the Z and H bosons combined, the bosons produced at these centre-of-mass energies and their jet products are boosted. Consequently, there are a significant fraction of jets which have energies above and below half the rest mass of the Z and H bosons. In each event two bosons are created, each of which produces two jets. Therefore in an ideal case, there should not be any jet which carries half of the centre-of-mass energy or more. But in some cases, the total reconstructed energy of the two jets produced by one of the bosons can be slightly higher than the true value due to measurement uncertainties and therefore one of them can have an energy equal to half of the centre-of-mass energy.

The particle constituents of a jet can be classified into three categories according to the subdetector in which their energy is measured in the particle flow approach. The fractions of jet energy carried by the particles of each category are shown in figure 7.2. For both



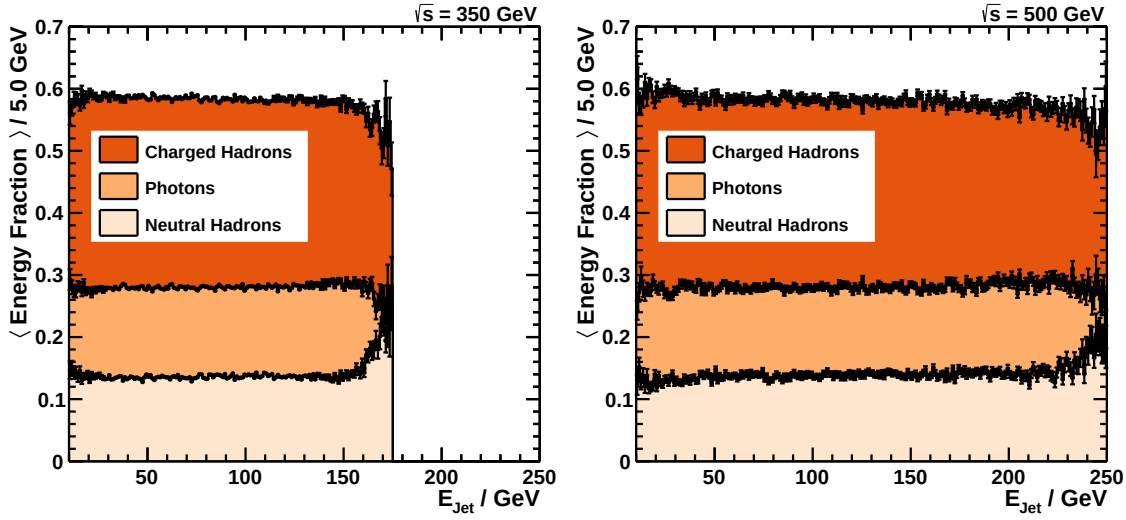
(a) (b)

Figure 7.2 – Fraction of the total jet energy in charged hadrons, photons and neutral hadrons for centre-of-mass energies of a) 350 GeV and b) 500 GeV. In both cases, on average around 58% of the jet energy is carried by charged hadrons, roughly 28% by photons and the remaining 14% by neutral hadrons. However, the spread around the average values is significant, indicating a large jet-to-jet difference.

centre-of-mass energies, the distributions of energy fractions carried by charged hadrons and photons are Gaussian-like. The mean of the distribution for the energy fraction of charged hadrons is roughly 58% while for the distribution of the fraction of energy carried by the photons the mean is around 28%. The remaining 14% of the jet energy is carried by neutral hadrons. A large number of jets have no neutral hadron content thus they can be measured with a better energy resolution. These jets can be candidates to determine the ass of the Higgs boson with a smaller uncertainty.

Figure 7.3 shows the mean energy fractions as a function of the jet energy for jet energies from 10 GeV to half of the centre-of-mass energy. The fraction of jet energy carried by different categories of particles remains relatively constant throughout the range of jet energies. The composition at the highest jet energies is distorted because of worse energy resolution of jets with large neutral hadron fraction, and the strongly decreasing number of jets with large true energy.

The importance of the jet particle multiplicity for the performance of PFAs is explained in section 5.2 and is studied in section 7.2.5. Distribution of the jet particle multiplicities in the Monte Carlo samples for the centre-of-mass energies of 350 GeV and 500 GeV is shown in figure 7.4. On average, each jet consists of around 20 particles. Average multiplicities of charged hadrons and photons in each jet are roughly 10 with a large spread around the mean value. The distribution of neutral hadron multiplicity is much narrower and peaks around 4.



(a) (b)

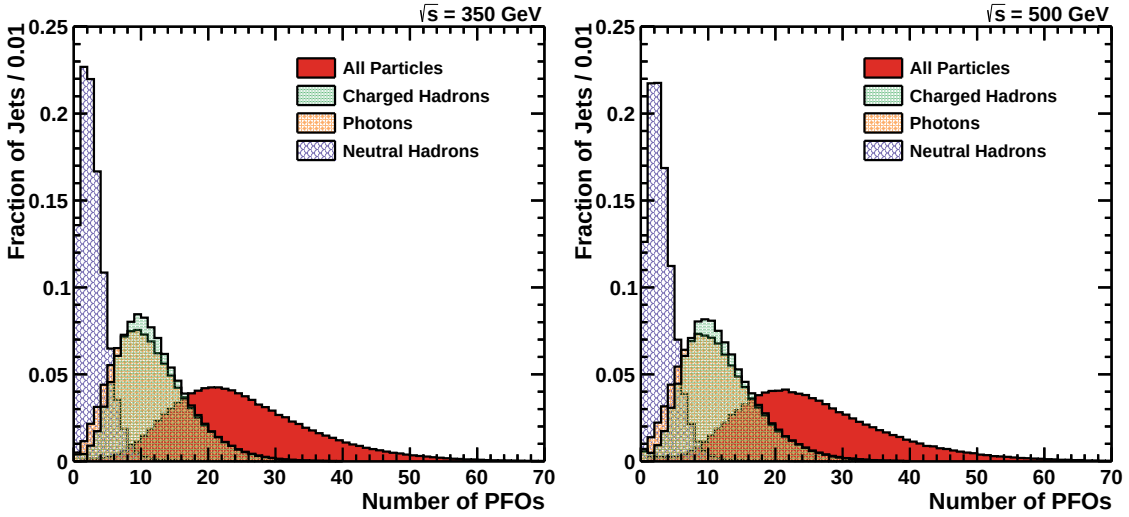
Figure 7.3 – Profile of the fractions of jet energy in the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ Monte Carlo samples for the centre-of-mass energies of a) 350 GeV and b) 500 GeV. The fractions remain relatively constant from the lowest to the highest jet energies.

The energy dependence of jet particle multiplicities in the range of 10 GeV to half of the centre-of-mass energy is shown in figures 7.5a and 7.5b. The total number of particles in a jet increases with the jet energy. In jets with nearly half of the centre-of-mass energy, the number of particles is larger than in jets with the lowest energies by a factor of 4. The number of neutral hadrons is increased by a factor of 5 while number of charged hadrons and photons are increased roughly by a factor of 3.

The aforementioned plots show that only a small fraction of the total jet energy is carried by the neutral hadrons. Since the energy of neutral hadrons is measured with poor resolution, the overall jet energy resolution is deteriorated. However, there are a large number of jets without any neutral hadron content or only a few. In principle, such jets are measured with a much lower uncertainty. By estimating the energy resolution for each jet individually, the effect of the neutral hadrons can be minimised. A method to estimate the jet-specific energy resolution is presented in the following.

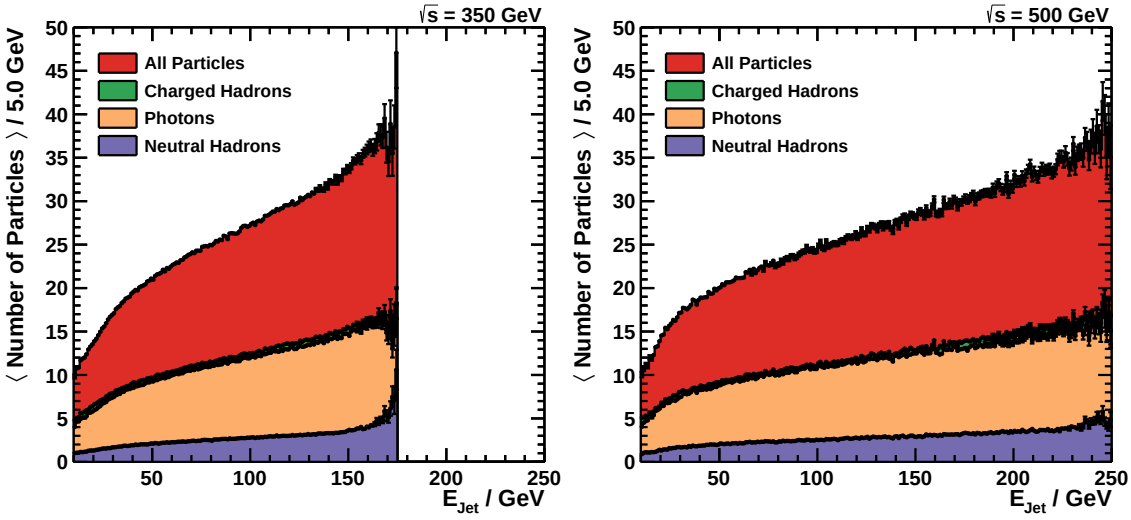
7.2 Jet-Specific Energy Resolution

Traditionally, an average Jet Energy Resolution (JER) is considered for a class of jets sharing similar properties such as energy, flavour or detection region. Due to the difference in instrumentation and performance of the detectors in the barrel and forward regions, different average values of the JER are considered for events taking place in these regions. The power of the PFA in the reconstruction of single particles in a jet can be exploited to develop a method to estimate the uncertainty on the jet energy measurement for each



(a) (b)

Figure 7.4 – Distribution of the jet particle multiplicities in the $e^+e^- \rightarrow ZH \rightarrow q\bar{q}b\bar{b}$ Monte Carlo samples for the centre-of-mass energies of a) 350 GeV and b) 500 GeV. On average, there are 20 particles in each jet. Average multiplicities of charged hadrons and photons in each jet are roughly 10 while for neutral hadron is around 4. There are many jets with 0 or 1 neutral hadrons.



(a) (b)

Figure 7.5 – Dependence of the average particle multiplicities per jet on the reconstructed jet energies in the $e^+e^- \rightarrow ZH \rightarrow q\bar{q}b\bar{b}$ Monte Carlo samples for centre-of-mass energies of a) 350 GeV and b) 500 GeV.

individual jet. This jet-specific energy resolution can be used in the physics analysis to improve the precision of measurements.

After a discussion about various sources which contribute to the total jet energy resolution, the details of how the main contributions are estimated is discussed in the following subsections. In chapter 8, the application of the estimated jet-specific energy resolution in the measurement of the Higgs boson mass in the $e^+e^- \rightarrow ZH \rightarrow q\bar{q}b\bar{b}$ channel is explained.

7.2.1 Sources of Uncertainty in Jet Energy Measurements

Jets are collimated showers of particles, created by the hadronisation of partons, as explained in section 2.3.2. There are multiple theoretical and experimental factors which impose limitations on measurements of the energy of these particle showers. Theoretical difficulties in the jet energy measurements mostly arise from the fact that the models used to describe the interactions do not perfectly describe the details involved or if they do, it can not be computed to the lowest levels. On the experimental aspects, intrinsic limitations of the measurement devices in one hand, and detection difficulties due to the nature of the particles and the complexity of the analysis techniques on the other hand, increase the uncertainty of the jet energy measurements.

The main contributions to the total uncertainty on the measured jet energy include, but are not limited to:

- the uncertainty due to the intrinsic QCD effects in the parton showering and hadronization process ($\sigma_{Had.}$) (see section 2.3.2);
- the uncertainty due to the limitations of the measurement instruments ($\sigma_{Det.}$);
- the uncertainty due to the particle recognition confusion in the PFA ($\sigma_{Conf.}$) (see section 5.2);
- the uncertainty due to semi-leptonic decays (σ_{ν});
- the uncertainty due to the misassignment of particles in the jet clustering ($\sigma_{Clust.}$) (see section 6.5.3).

There are other sources contributing to the total jet energy resolution, such as the uncertainties due to the calibration methods, which are of smaller magnitudes and are ignored in this analysis.

Here it is assumed that the contributions to the jet energy resolution are uncorrelated, and that the total uncertainty on the measured energy of a jet (σ_{Jet}) can be obtained by the

quadratic sum of the uncertainties listed above as:

$$\sigma_{Jet} = \sigma_{Had.} \oplus \sigma_{Det.} \oplus \sigma_{Conf.} \oplus \sigma_{\nu} \oplus \sigma_{Clust.}. \tag{7.1}$$

The uncertainties due to the measurement instruments $\sigma_{Det.}$, confusion $\sigma_{Conf.}$ and semi-leptonic decays σ_{ν} in the b-jets are expected to have the dominant contributions to the total uncertainty. The capability of the Particle Flow approach in reconstructing single particles in a jet can be employed to estimated these three uncertainties for each jet individually. The details of this method is described in the following sections.

7.2.2 Hadronization Uncertainty

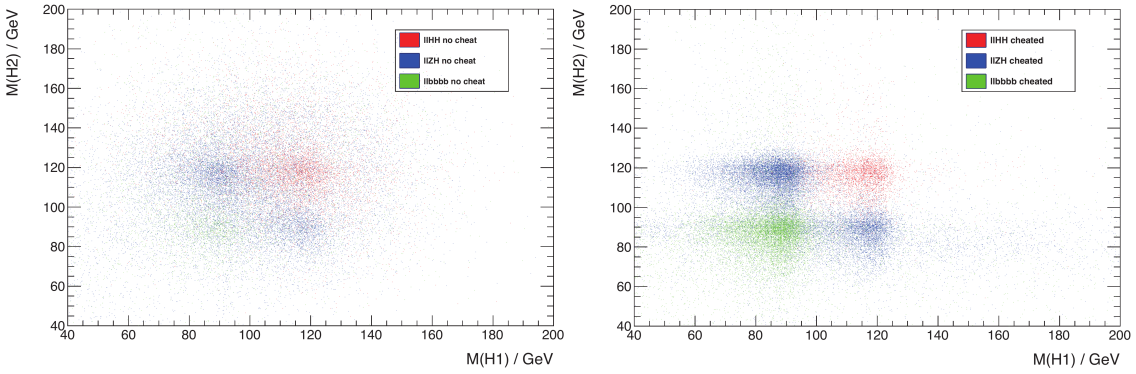
The uncertainties due to the parton showering and hadronization process arise because of limited theoretical knowledge of QCD phenomena. As discussed in section 2.3, these mechanisms can not be computed using the first principles and none of the models for the hadronisation process can perfectly describe the hadronic final states.

Mismodelling of the hadronisation process causes a bias or wrong estimation of the fluctuations. Therefore, the predictions of these models, which are implemented in the Monte Carlo event generator, contribute to the total uncertainty on the jet energy measurements. The hadronization uncertainties are usually small compared to the other sources.

7.2.3 Clustering Uncertainty

As explained in section 6.5.3, a jet finding algorithm is used to cluster particles in event into four jets. The jet finding algorithms are not perfect and can assign particles to wrong jets. This misassignment introduces another source of uncertainty on the jet energy measurement.

The performance of the Durham jet finding algorithm has been studied in the context of Higgs boson self-coupling measurement using $e^+e^- \to ZHH$ with $H \to b\bar{b}$ [24]. The invariant masses of two Higgs boson candidates using the Durham jet finding algorithm and Monte Carlo truth information are shown in figure 7.6. In the case where Monte Carlo truth information is used, the separation between signal and background is significantly improved. These results clearly demonstrate the importance of jet clustering. A new jet clustering algorithm is being developed which would improve the performance [24]. The uncertainty on the jet energy due to clustering should be studied in the future using the new algorithm.



(a) (b)

Figure 7.6 – Invariant mass of two Higgs boson candidates a) using Durham jet finding algorithm and b) using Monte Carlo truth information [24].

7.2.4 Detector Resolution

As described in section 6.5.3, jets are identified by means of jet finding algorithms which cluster particle candidates. In the recombination scheme used in the jet finding process, the four-momentum of a jet corresponds to the sum of the four-momenta of all particles included in the jet. In the particle flow approach, the energy or the momentum of individual particles are measured in the sub-detectors which offer the best possible resolution for that type of particle: charged particles are measured in the tracking system, photons in the electromagnetic calorimeter and neutral hadrons in the hadronic calorimeter.

Each of the sub-detectors has its own limitations and therefore its own uncertainty on the measurement of the properties of the particles. Hence, the total jet energy uncertainty due to the detector limitations can be described by sum of the uncertainties from the various sub-detectors:

$$\sigma_{Det.} = \sigma_{Tracker} \oplus \sigma_{ECAL} \oplus \sigma_{HCAL} \tag{7.2}$$

in which $\sigma_{Tracker}$ is the resolution of the tracking system, σ_{ECAL} is the single-particle energy resolution of the ECal and σ_{HCAL} is the single-particle energy resolution of the HCal.

The contribution of each sub-detector to the total jet uncertainty can be estimated by the sum of energy uncertainties of all the particles measured in that sub-detector. Thus, the equation 7.2 can be written as:

$$\sigma_{Det.} = \sum_{PFO_{Ch}} \sigma_{Tracker}.\boldsymbol{p}_{Ch} \oplus \sum_{PFO_{\gamma}} \sigma_{ECAL}.E_{\gamma} \oplus \sum_{PFO_{Ne}} \sigma_{HCAL}.E_{Ne} \tag{7.3}$$

where $\sum_{PFO_x}$ is sum over all particles in a jet that belong to a particular category of particles. $\boldsymbol{p}_{ch}$ is momentum of the charged particle and E_x is energy of the particle.

To compute the contribution from each of the sub-detectors, particle identification information as provided by PandoraPFA is employed. All of the particles clustered in a jet are checked to determine whether they are charged hadrons, photons or neutral hadrons. Then, the uncertainty of the particle is computed and added to the total detector uncertainty for the jet, according to the equation 7.3.

The relative uncertainty on the measured momentum in the tracking system is given by the Glückstern equation [82]:

$$\frac{\sigma_{p_T}}{p_T} = \frac{\sigma(x,y)p_T}{0.3L^2B}\sqrt{\frac{720}{N+4}} \tag{7.4}$$

where $\sigma(x,y)$ is spatial resolution of each hit, B is the magnetic field strength, L is the path length of the track and N is number of hits.

The energy resolution of calorimeters is explained in section 5.1.3. The relative energy resolution of a calorimeter is given by:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b \tag{7.5}$$

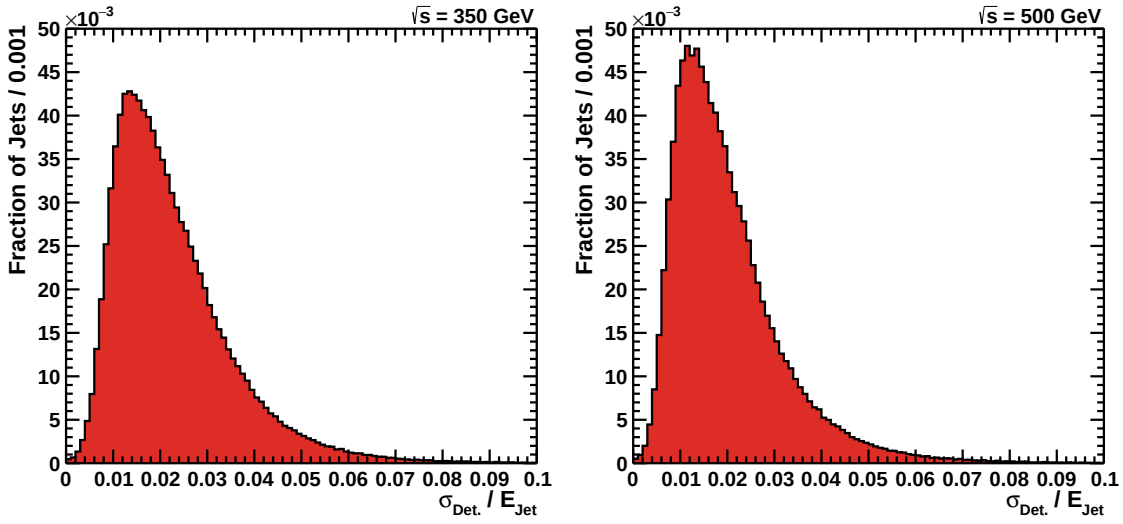
where σ_E is the absolute energy resolution of the calorimeter, E is the energy, a is the stochastic term and b is the constant term. The contribution from the noise term is small compared to the other two terms and decreases with particle energy, hence it is ignored here.

In this study the values which are used for the stochastic term and constant term in the resolution equation for the ECal and HCal are obtained from the CALICE test beams analysis for the AHCAL [83] and the Si-ECAL [84]. The published values are rounded up to the nearest integer (in percent) and are listed in table 7.1.

Table 7.1 – Calorimeter resolution terms used to estimate the detector resolution. These values are obtained from the CALICE test beam analysis for the AHCAL [83] and the Si-ECAL [84].

Calorimeter	Stochastic Term (%)	Constant Term (%)
HCal	50	1
ECal	16	2

The distribution of the relative jet energy resolution σ_E/E_{Jet} due to the detector single-particle resolutions is shown in figure 7.7. The mean of the distribution is around 2% while in few cases it can be as high as $6-7\%$. The distributions are asymmetric around the peak value with a tail towards higher values. The reason is that the calorimeter resolution deteriorates with the inverse of the square root of the energy, as can be seen in the dependence of the relative jet energy resolution on the jet energy in figure 7.8. The 2% relative detector resolution which is the peak of the distribution in figure 7.7 corresponds to the jets of 90 GeV. The jets with higher energies are measured with less uncertainty



(a) (b)

Figure 7.7 – Distribution of the relative jet energy resolution due to the detector limitations. On average, the detector contribution to the total uncertainty on the jet energy is about 2%.

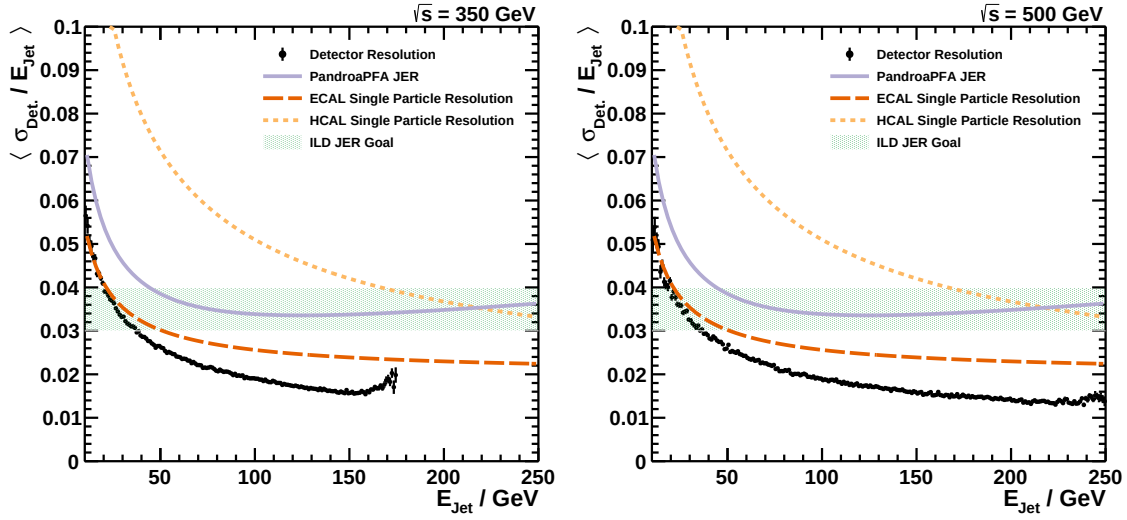
while the resolution for lower energy jets deteriorates rapidly. Figure 7.1 shows that a significant number of jets have energies below 90 GeV and therefore cause the asymmetry towards higher uncertainties in figure 7.7. The dashed lines in figure 7.8 represent the relative energy resolution functions of the ECal and the HCal for single particles which are obtained by inserting the values from table 7.1 into equation 7.5. The horizontal band between 3% and 4% represents the jet energy resolution goal at the ILD. The solid line indicates the jet energy resolution obtained by the PandoraPFA [40].

If a jet is entirely contained in the ECal or the HCal, then its energy resolution follows the intrinsic calorimeter resolution of the equation 7.5, or energy resolutions for single particles might add up to a worse energy resolution. The jet-specific energy resolution obtained using the error flow method is significantly better than the pure calorimetric resolution, as shown in figure 7.8.

7.2.5 Confusion Uncertainty

Misidentification of the particles and misassignment of tracks to calorimeter clusters cause an uncertainty in the jet energy measurement which is referred to as *confusion uncertainty*. The concept of confusion in the PFA is explained in details in section 5.2. A method is developed to estimate the confusion uncertainty based on the particle constituents of a jet. This method is presented in this section.

As described in section 5.2, there are three main sources contributing to the total confusion uncertainty:



(a) (b)

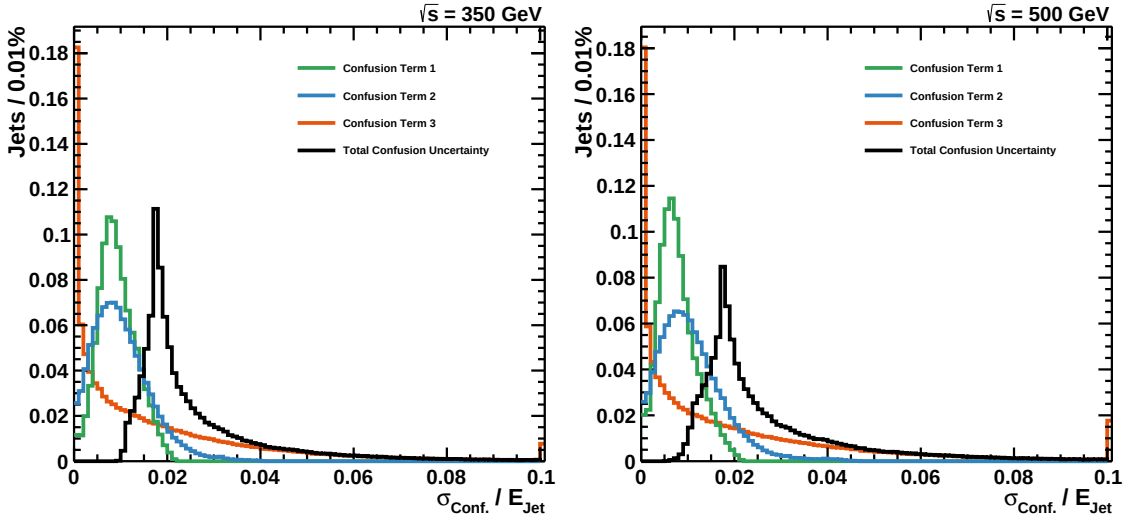
Figure 7.8 – Average relative uncertainty on the jet energy due to the detector limitations as a function of the jet energy. The intrinsic energy resolutions of the calorimeters for single particles are shown for comparison. The horizontal band represents the target region of the jet energy resolution at ILD. The solid line indicates the jet energy resolution of the PandoraPFA [40].

- calorimeter hits from charged hadrons that are reconstructed as a neutral hadron cluster ($\sigma_{Conf.,1}$) and thus double counted;
- calorimeter hits from photons which are lost in charged hadrons ($\sigma_{Conf,2}$);
- calorimeter hits from neutral hadrons which are lost in charged hadron clusters ($\sigma_{Conf.,3}$).

The total uncertainty on jet energy measurement due to the confusion is given by quadratic sum of the confusion terms above:

$$\sigma_{Conf.} = \sigma_{Conf,1} \oplus \sigma_{Conf,2} \oplus \sigma_{Conf,3}. \tag{7.6}$$

The contribution from each of the three confusion sources has been studied using Z decays to light quarks in [40] for jet energies of 45 GeV, 100 GeV, 180 GeV and 250 GeV and is listed in table 7.2. The values presented there are for average jets which have 62% of their energy carried by charged hadrons, 28% by photons and the remaining 10% by neutral hadrons. Using the PFA information about the particle content of an individual jet, the deviation of the jet energy fraction carried by each category of particle from the average fractions can be obtained. Then, these deviations are used to scale the tabulated confusion terms in order to estimate these terms for this individual jet. The following equation is



(a) (b)

Figure 7.9 – Distribution of the relative jet energy resolution due to the confusion.

used to scale the confusion terms:

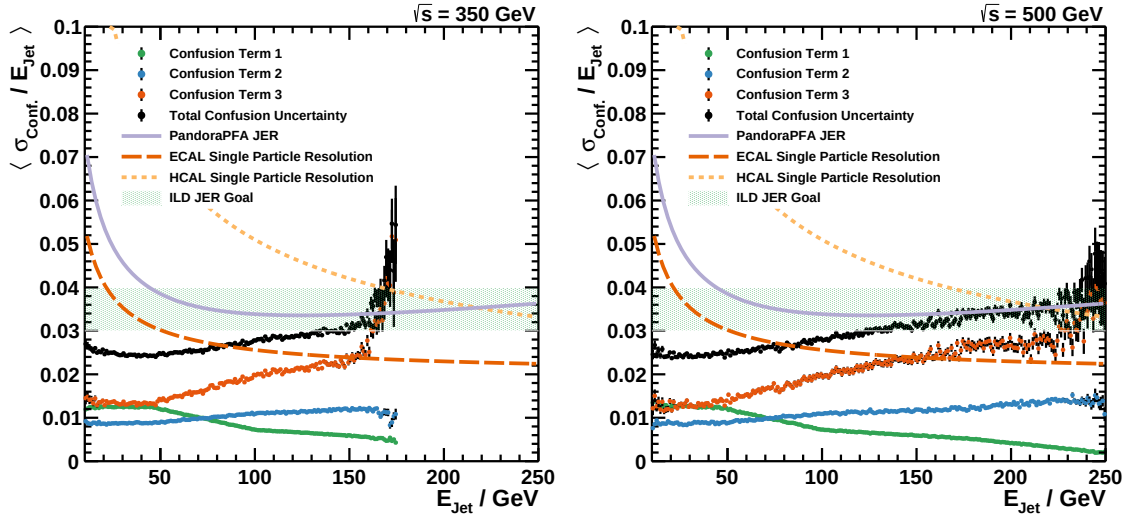
$$\sigma_{Conf,i} = \langle \sigma_{Conf,i} \rangle . (\frac{\Gamma_i}{\langle \Gamma_i \rangle}) \tag{7.7}$$

where $i \in [1, 3]$ defines the category of particles, $\langle \sigma_{Conf,i} \rangle$ is the average confusion term obtained from interpolation of the values given in [40], $\Gamma_i = E_i / E_{Jet}$ is fraction of the jet energy carried by a category of particles and $\langle \Gamma_i \rangle \in (0.62, 0.28, 0.1)$ is the average fraction of jet energy carried by a category of particles. The average confusion term $\langle \sigma_{Conf,i} \rangle$ is assumed to be constant for jets of energies below 45 GeV, since currently there is no other estimate available. This is a conservative assumption since the confusion is expected to decrease for less energetic jets.

Table 7.2 – Contributions to the total confusion from each of the three confusion sources, reproduced from [40].

Contribution	**Relative Jet Energy Resolution σ_{E_j}/E_j**			
	$E_j = 45$ GeV	$E_j = 100$ GeV	$E_j = 180$ GeV	$E_j = 250$ GeV
$\sigma_{Conf.,1}$	1.32 %	0.77 %	0.55 %	0.22 %
$\sigma_{Conf.,2}$	0.88 %	1.10 %	1.21 %	1.43 %
$\sigma_{Conf.,3}$	0.99 %	1.43 %	1.87 %	1.98 %
Total Confusion	1.87 %	1.98 %	2.31 %	2.53 %

The distribution of each of the aforementioned confusion uncertainties and the total uncertainty on the jet energy due to confusion is shown in figure 7.9. The distributions of $\sigma_{Conf,1}$ and $\sigma_{Conf,2}$ are relatively symmetric and peak around 1% while $\sigma_{Conf,3}$ distribution



(a) (b)

Figure 7.10 – Average relative uncertainty on the jet energy due to the confusion as a function of the jet energy. The intrinsic energy resolutions of the calorimeters for single particles are shown for comparison. The horizontal band represents the target region of the jet energy resolution at ILD. The solid line indicates the jet energy resolution of the PandoraPFA [40].

peaks at zero an decreases towards higher values with a mean of 2%. The distribution of the total confusion uncertainty $\sigma_{Conf.}$ has an average value of roughly 2%.

The changes in the mean confusion uncertainties as a function of the jet energy are shown in figure 7.10. $\sigma_{Conf,1}$ is decreasing with the jet energy. $\sigma_{Conf,1}$ is due to fragments of the charged hadron clusters which are reconstructed as neutral hadrons. The jets of higher energies are more collimated and on average they have fewer fragments separated from the main cluster. Hence, it is less likely that the pattern recognition algorithm of the PFA recognises part of the cluster of high energy charged hadrons as a neutral hadron and, consequently, $\sigma_{Conf,1}$ is decreasing with the jet energy.

The confusion terms $\sigma_{Conf,2}$ and $\sigma_{Conf,3}$ are increasing with the jet energy. The change in $\sigma_{Conf,2}$ over the range of jet energies is lower than the change in $\sigma_{Conf,3}$. The increase of these confusion terms as a function of the jet energy and the difference in the rate with which they increase can be explained by the particle contents of the jets. $\sigma_{Conf,2}$ is due to failure in resolving hits from photons which are close to the tracks of charged hadrons, while $\sigma_{Conf,3}$ are calorimeter hits from neutral hadrons which are lost in charged hadron clusters. It is shown in figure 7.5 that the total number of photons and charged hadrons in a jet increases roughly with the same rate as a function of the jet energy. Consequently, the density of these types of particles in a high energy jets is larger. Furthermore, the high energy jets are more densely collimated and hence the probability of assigning photon hits to charged hadron tracks is higher. Therefore, $\sigma_{Conf,2}$ increases from 0.8% for low energy jets to 1.5% for highest jet energies.

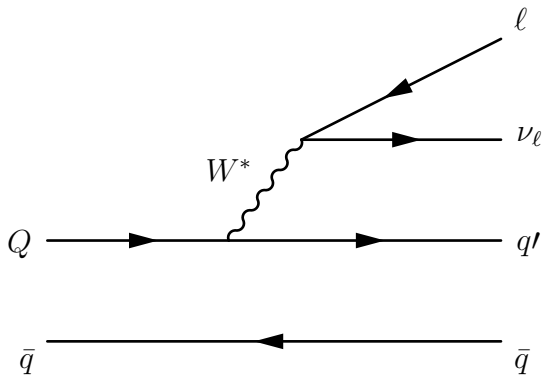


Figure 7.11 – Feynman diagram of the semi-leptonic decay

Figure 7.5 also shows that the multiplicity of the neutral hadrons in a jet increases with the energy of the jet. As a result, in addition to particles in high energy jets being more closely located in the jet cone, the significant increase in the number of charged and neutral hadrons in a jet, and subsequently the number of calorimeter clusters assigned to their tracks, increases the probability of loosing neutral hadron hits as part of charge hadron clusters and therefore a steeper increase in $\sigma_{Conf,3}$.

7.2.6 Semi-leptonic Decays

Heavy quarks can decay via charged weak interactions in a manner similar to the nuclear β-decay. Depending on the decay mode of the virtual W boson, the final state of the decay can consist of a charged lepton, its corresponding neutrino and one or more hadrons, $Q \to q\prime W^*, W^* \to l\nu_l$. This decay mode is referred to as semi-leptonic decay. A Feynman diagram of the semi-leptonic decay is shown in figure 7.11.

Due to the presence of a neutrino in the final state of a semi-leptonic decay and since the neutrinos are not measured in the detector, the energy of the jets containing a semi-leptonic decay differs from the energy of the initial quark. Nearly 10% of the b quarks decay in semi-leptonic mode [19]. The problem with the missing neutrino energy has no straightforward solution. A simple method to estimate the missing energy of the neutrino in semi-leptonic decays and the contribution of this method to the total jet energy uncertainty σ_ν is presented in this section.

The neutrinos created in semi-leptonic decays are always accompanied by a charged lepton. At ILD, lepton tracks within jets can be identified using dE/dx in the main tracker and shower shape and MIP signatures in the calorimeter and the muon system [85][86][87]. Considering this fact and keeping in mind that accompanying leptons are charged particles and their energy is measured to a high precision in the tracking tracking system, it is possible to construct an observable based on the energy of the charged lepton with which the missing energy of neutrino can be estimated. The total energy of the jet is then corrected

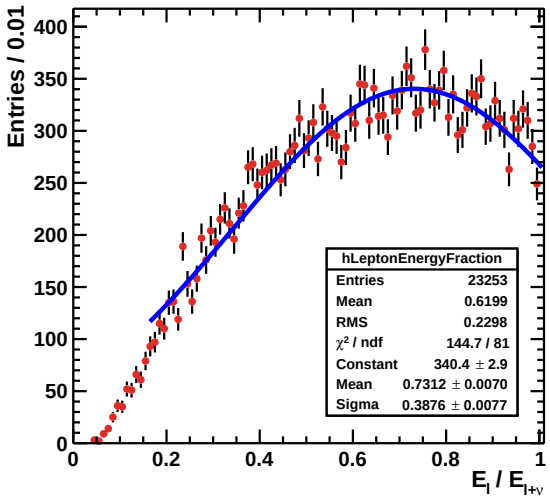


Figure 7.12 – Distribution of the energy of charged leptons relative to the total energy of the charged lepton and the corresponding neutrino in semi-leptonic decays at the MC truth level.

by adding the estimated energy of the neutrino to the jet energy:

$$E_{Jet}^{Est.} = E_{Jet}^{Meas.} + E_{\nu}^{Est.} \tag{7.8}$$

where $E_{Jet}^{Est.}$ is the estimated jet energy, $E_{Jet}^{Meas.}$ is jet energy measured in the detector and $E_{\nu}^{Est.}$ is estimated missing energy of the neutrinos.

Among other possibilities, the distribution of lepton energies relative to the total energy of the lepton-neutrino pairs can be used to estimate the missing energy of the neutrinos as:

$$x = \frac{E_\ell}{E_\ell + E_{\nu}^{Est.}} \rightarrow E_{\nu}^{Est.} = (\frac{1}{x} - 1)E_\ell \tag{7.9}$$

where the parameter x is defined to be a fraction of the energy of the lepton-neutrino pair carried by the charged lepton and E_ℓ is energy of the accompanying charged lepton.

Figure 7.12 shows the distribution of x for about sixty thousands $e^+e^- \rightarrow ZH \rightarrow q\bar{q}b\bar{b}$ events. Monte Carlo truth information is used to find semi-leptonic decays. A Gaussian function is fitted to the distribution. The expectation value $\langle x \rangle$ of this distribution can be used as an estimator for x. Then, equation for the true jet energy can be written as:

$$E_{Jet}^{Est.} = E_{Jet}^{Meas.} + \left(\frac{1}{\langle x \rangle} - 1\right) E_\ell. \tag{7.10}$$

Since this correction is based on statistical inference, there is an uncertainty on the estimated value which contributes to the total uncertainty on the jet energy measurements.

This uncertainty is estimated as:

$$\sigma_\nu^2 = \left(\frac{\sigma_{\langle x\rangle}}{\langle x\rangle^2}\right)^2 E_\ell^2 + \left(\frac{1}{\langle x\rangle} - 1\right)\Delta E_\ell^2 \tag{7.11}$$

in which $\sigma_{\langle x\rangle}$ is the width of the Gaussian fit and ΔE_ℓ is the uncertainty on the energy of the accompanying charged lepton. As mentioned earlier, E_ℓ is measured in the tracking system and its corresponding uncertainty is negligible here. Therefore, the second term can be omitted from the uncertainty equation.

By substituting the values for $\langle x\rangle$ and $\Delta\langle x\rangle$ from the fit in figure 7.12, the equations for the jet energy and its uncertainty can be written as

$$E_{Jet}^{Est.} = E_{Jet}^{Meas.} + 0.37 \cdot E_\ell \tag{7.12}$$

$$\sigma_\nu = 0.73 \cdot E_\ell \tag{7.13}$$

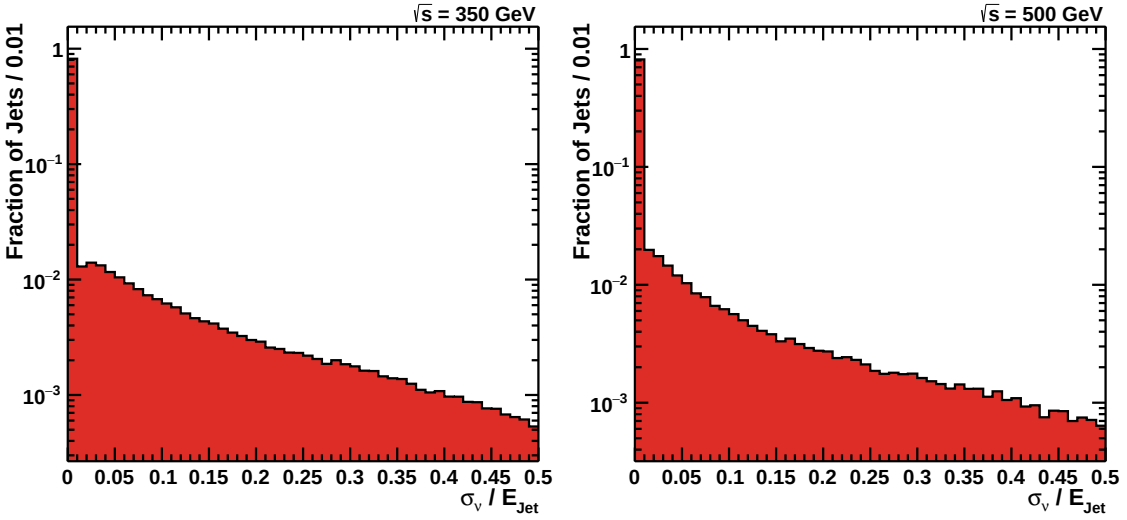
The distribution of the uncertainty contribution arising from the semi-leptonic correction is shown in figure 7.13. As expected, nearly 10% of the jets have semi-leptonic decays and hence non-zero σ_ν. The relative uncertainty due to the semi-leptonic decays as a function of the jet energy only for events with semi-leptonic decays is shown in figure 7.14. In a wide range of jet energies, the relative uncertainty due to the semi-leptonic decays is on average between 10% to 15%.

The application of the correction of the jet energy for missing neutrino energy using equation 7.12 is discussed in section 8.

7.3 Total Jet Energy Resolution

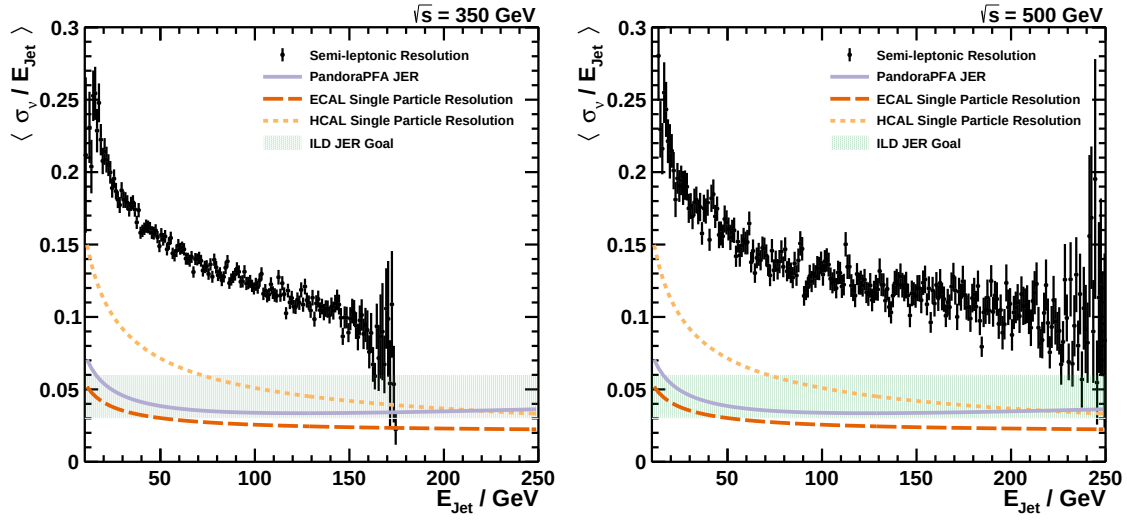
Various contributions to the total uncertainty on the measured jet energy have been discussed in the previous sections. The various contributions to the uncertainty on the measured jet energy can now be combined into a total uncertainty as the jet-specific energy resolution outcome of the error flow method. The distribution of the estimated total relative uncertainty on the measured jet energy is shown in figure 7.15. The distribution peaks approximately around $2-3\%$ with a tail towards higher uncertainties. It is worth noting that since contributions from hadronisation and jet clustering uncertainties are not included, it might be necessary to apply a fudge factor to account for these uncertainties.

The total relative uncertainty as a function of jet energy is plotted in figure 7.16. The sum of the uncertainties due to the detector limitations and due to the PFA confusion is also shown on the same plots. When only the uncertainties due to the detector limitations and confusion are considered, $\sigma_{Det.} \oplus \sigma_{Conf.}$, the total uncertainty on the jet energy is within the



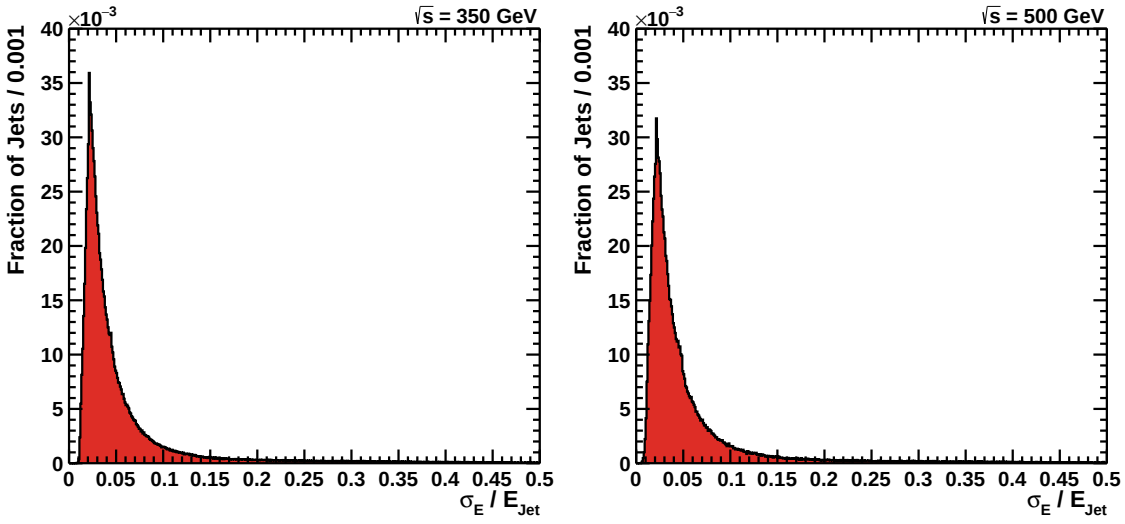
(a) (b)

Figure 7.13 – Distribution of the relative uncertainty on the jet energy due to the semi-leptonic correction. Nearly 10% of the decays are via semi-leptonic mode and hence have non-zero σ_ν.



(a) (b)

Figure 7.14 – Average relative uncertainty on the jet energy due to the semi-leptonic correction as a function of the jet energy for events with semi-leptonic decays. The intrinsic energy resolutions of the calorimeters for single particles are shown for comparison. The horizontal band represents the target region of the jet energy resolution at ILD. The solid line indicates the jet energy resolution of the PandoraPFA [40].



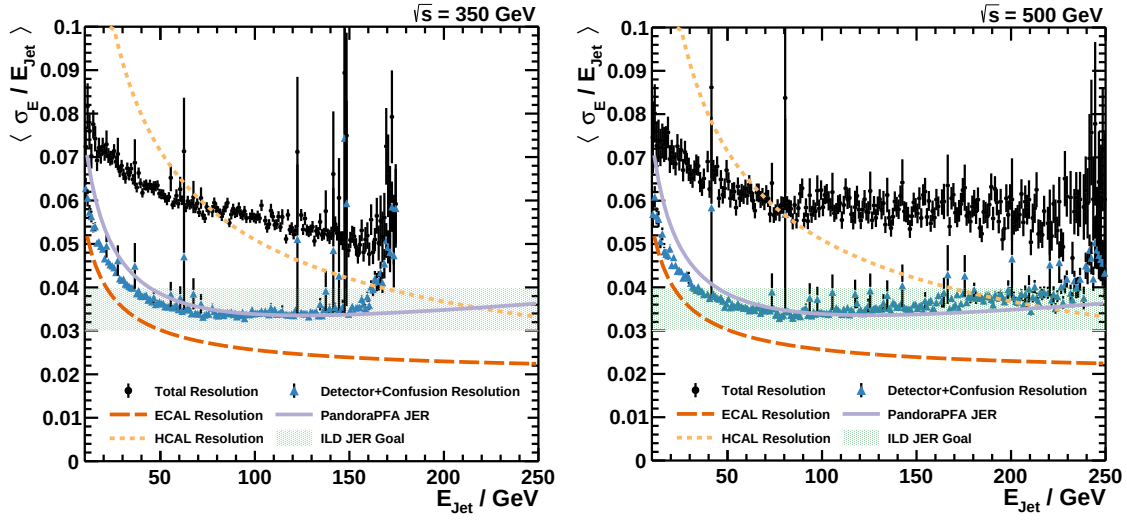
(a) (b)

Figure 7.15 – Distribution of the relative total jet energy resolution obtained using the error flow method for the centre-of-mass energies of 350 GeV and 500 GeV.

$3-4\%$ goal of ILD except for energies below 40GeV. The total uncertainty is compatible with the PandoraPFA jet energy resolution curve over a significant range of jet energies. This agreement is expected since the PandoraPFA curve is determined for uds jets. As explained in section 7.2.5, a in the parametrisation of the confusion uncertainty, a constant confusion is assumed for the jet energies below 40 GeV. Hence, the small discrepancy for the jets with energies below ~ 50 GeV is very likely to be due to the underestimation of confusion at these energies.

By adding the uncertainty due to the semi-leptonic decays, the total relative uncertainty deteriorates across the jet energy range by approximately 2%, bringing up the total relative uncertainty above the uncertainty goal for ILD. However, only a small fraction of quarks undergo semi-leptonic decays and in principle it is possible to identify such events and exclude them to meet the goal. Moreover, it is worth noting that the jet energy uncertainties plotted here are average values and there are a significant number of jets with lower uncertainties. The jet-specific energy uncertainty obtained by the error flow empowers selection of jets with uncertainties better than the $3-4\%$ goal.

PandoraPFA provides a global value for the jet energy resolution regardless of the jet composition. As introduced in this chapter, a new method for estimating the jet-specific energy resolution has been developed which uses the power of the particle flow algorithm in reconstructing individual particle in a jet. The jet-specific energy resolution can be exploited in analyses in order to improve the results.



(a) (b)

Figure 7.16 – Average total relative uncertainty on the jet energy as a function of the jet energy. Intrinsic single particle resolution of the calorimeters and the jet energy resolution from PandoraPFA [40] are shown for comparison. The horizontal band illustrates the target jet energy resolution at ILD. The triangles represent the sum of uncertainties due to the intrinsic detector resolution and particle flow confusion, obtained using the error flow method. The circles show the case where the correction for the semi-leptonic decays is included in the sum.

In the following section, the jet-specific energy resolution is used in the context of the Higgs boson mass measurement. The performance of this method is evaluated and is compared with the results obtained by using a fixed jet energy resolution.

Chapter 8

Higgs Boson Mass Reconstruction in the Fully Hadronic Channel

The motivation for a precise measurement of the Higgs boson mass is explained in chapter 4, where multiple possibilities for this measurement at ILC are also discussed. One of the promising approaches for measuring the Higgs boson mass to a high precision at ILC is to use the Higgsstrahlung process with hadronic decays of the Z and the Higgs bosons, $e^+e^- \to ZH \to q\bar{q}b\bar{b}$. In this chapter, the potential of the Higgs boson mass measurement using the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel as an alternative to the recoil technique at ILC is investigated using the ILD detector concept for the centre-of-mass energies of 350 GeV and 500 GeV.

The Monte Carlo samples used for the analysis are introduced in chapter 6. There are two samples for the centre-of-mass energy of 500 GeV, one with $\gamma\gamma$ background overlay and one without the overlay. In the event preparation a $\gamma\gamma$ background removal procedure was performed on the sample with the overlay. At the end of the event preparations of all samples, four jet objects where formed for each $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ event.

The first step of the analysis is to determine which two jets originate from the decay products of the Higgs boson and to compute their invariant mass. Then the Higgs boson mass can be extracted from the distribution of the invariant mass. A method to extract the Higgs boson mass and its corresponding uncertainty from the invariant mass distribution is presented in section 8.1.

One way to determine the jet pairs which are produced by the Higgs boson is to use a χ^2-minimisation technique. The extraction of the Higgs boson mass using this method is explained in section 8.2. Various issues arise from the heavy flavour quarks that are involved in this measurement. The flavour dependence of the jet energy scale and the Higgs boson mass reconstruction are explained in section 8.3.

Since the initial state of the collision is known to a high precision at ILC, the conservation laws can be exploited to adjust the four-momenta of the jets and thus improve the measurement results. Kinematic fitting is a mathematical tool which provides the framework for such improvements. The concept of kinematic fitting and its application on this analysis is discussed in section 8.4.

The outcome of the aforementioned studies are discussed in section 8.5. The reconstructed Higgs boson mass and its corresponding uncertainty obtained using various strategies is analysed in section 8.5.1. Throughout the analyses, the error flow method was used to provide a jet-specific energy uncertainty. The performance of the jet-specific energy uncertainty is discussed in section 8.5.2. As mentioned above, one of the Monte Carlo samples is overlaid with $\gamma\gamma$ background events. In section 8.5.3, the effectiveness of the $\gamma\gamma$ background removal procedure which is explained in section 6.5.2 is evaluated by using the outcome of the Higgs boson mass studies as a measure. The possible sources of systematic uncertainties are outlined in section 8.5.4. The chapter is concluded in section 8.6 with a discussion on the prospects of the Higgs boson mass measurement at ILC using only the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel and also combined with the recoil technique.

8.1 Extraction of the Mass from an Invariant Mass Distribution

In order to obtain the best estimate for the value of the Higgs boson mass, the distribution of the reconstructed invariant mass should be fitted with an appropriate function. Then the fit can be used to obtain the most-probable-value (position of the peak) and its corresponding uncertainty. An example of the invariant mass distributions is shown in figure 8.2. Usually, it is not possible to fit a well defined function on these distributions, except by narrowing down the fit range to a small region around the peak. Therefore, a more robust method for obtaining the Higgs boson mass is employed here. In this method, the following χ^2 for a number of test masses around the expected mass is computed:

$$\chi^2_i = \sum_j \frac{(M_j^{Fit} - M_i^{Test})^2}{(\sigma_j^M)^2} \tag{8.1}$$

where the sum j is over all events, M_j^{Fit} is the reconstructed invariant mass of event j, M_i^{Test} is the test mass and σ_j^M is the uncertainty on the reconstructed invariant mass. Then the minimum value is obtained by fitting a parabola on the computed χ^2 values as a function of test masses. The minimum χ^2_{min} is taken as the reconstructed mass of the particle and the width of the parabola in the range of $\chi^2_{min} \pm 1$ as the uncertainty on the reconstructed mass.

This method is also sensitive to the tails of the distribution, and applying cuts on the reconstructed invariant masses is necessary in order to be independent from backgrounds.

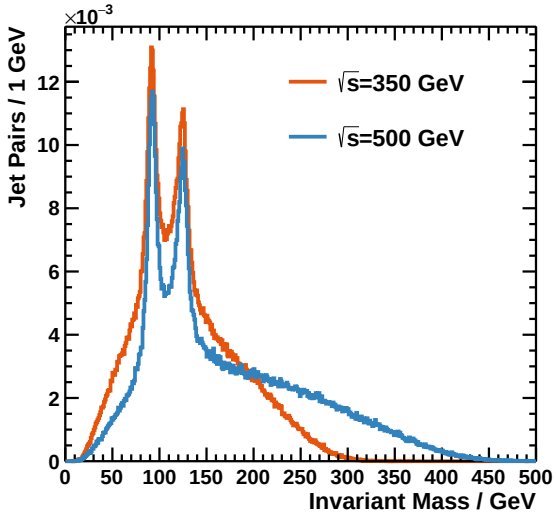


Figure 8.1 – Invariant mass distribution of all 6 possible permutations of the four jets for the centre-of-mass energies of 350 GeV and 500 GeV.

However, the range of invariant masses which is used to extract the mass is significantly larger than it would be for fitting an ordinary function such as a Gaussian. This method is used to extract Z and Higgs boson masses in this chapter.

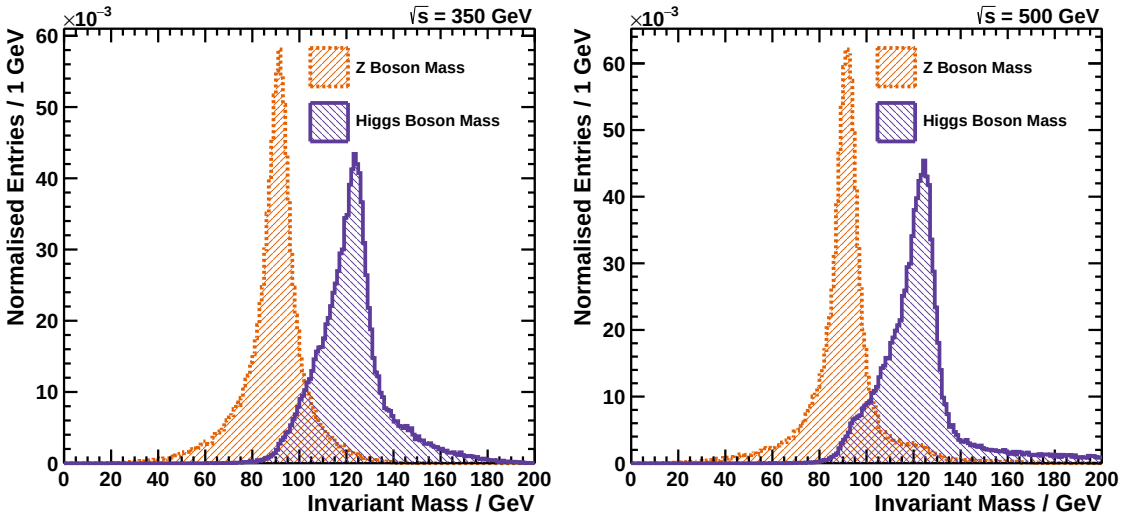
8.2 χ^2-Minimisation for Jet Pair Determination

At the tree level, in absence of hard gluon radiation, there are four jets expected in the final state of the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ process. A jet finding algorithm is used to cluster all particles within an event into four jets, as explained in section 6.5.3. Two of the jets belong to the decay products of the Z boson and the other two jets are produced by the two quarks created in the decay of the Higgs boson. In principle, it is possible to reconstruct the mass of the parent bosons by computing the invariant mass of the jets which originate from them, if the jets could be correctly assigned to the bosons. In practice, there is no evident way to determine from which of the bosons a jet is originating. Therefore, there is a three-fold ambiguity in the jet pairing in $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ events.

The invariant mass of all six possible permutations of two-jet subsets of the four jets in each event is computed and plotted in figure 8.1. The distribution has two peaks around 91 GeV and 125 GeV, corresponding to the rest masses of the Z and the Higgs boson, respectively. The broad shoulders are due to incorrect jet pairs.

One possible way to find the best pairing is to employ a χ^2-minimisation technique. The χ^2 is defined as:

$$\chi^2 = \frac{(M_{ij} - M_Z)^2}{\sigma_Z^2} + \frac{(M_{kl} - M_H)^2}{\sigma_H^2} \tag{8.2}$$



(a) (b)

Figure 8.2 – Distributions of the reconstructed masses of the Z and the H bosons computed using the χ^2-minimisation method for the centre-of-mass energies of a) 350 GeV and b) 500 GeV. In both cases, the distribution of the H mass is skewed due to the missing energy from the semi-leptonic decays of the heavy flavour quarks.

where M_{ij} and M_{kl} are the invariant masses of two-jet subsets of the total four jets in an event, M_Z and M_H are the expected masses of the Z and the Higgs bosons, and σ_Z and σ_H are the expected uncertainties on their masses, respectively. In this analysis, the expected masses of the bosons are set to the values used in the Monte Carlo generator for the data sample of this study$M_Z = 91.1882$ GeV and $M_H = 125.0$ GeV. The uncertainties on the masses are computed using the error flow result for the uncertainty on the jet energy and a fixed uncertainty on the jet azimuthal and polar angles of $\sigma_\theta = \sigma_\phi = 100$ mrad [1].

The χ^2 of equation 8.2 is computed for each of the 6 possible permutations of the four jets in an event. The jet pairing with the minimum χ^2 is chosen as the best permutation and M_{ij} is taken as the reconstructed mass of the Z boson and M_{kl} as the reconstructed mass of the Higgs boson. The distributions of the reconstructed masses of the Z and Higgs bosons are shown in figure 8.2. While the distribution of the Z boson mass is nearly symmetric around the simulation input mass, the distribution of the reconstructed Higgs boson mass is skewed. The skewness of the M_H distribution is more pronounced for the case of 500 GeV centre-of-mass energy. This skewness is possibly due to semi-leptonic decays of heavy flavour quarks. In the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ data sample which is used in for this study, the Higgs boson solely decays to via $H \to b\bar{b}$ channel, while the Z boson decays to b and other lighter quarks. Therefore, the effect of heavy flavour quarks is more visible on the reconstructed mass distribution of the Higgs boson. This issue is discussed in details in sections 7.2.6 and 8.3.

[1] This value was found in [88] to result in the best convergence rate of the kinematic fit. Nevertheless, it is a very large value and should be optimised in the future.

Table 8.1 – Reconstructed mass of the Z and the Higgs bosons obtained by the χ^2-minimisation method for an integrated luminosity of 1000 fb^{-1} and a beam polarisation of $(-0.8,+0.3)$.

		$\sqrt{s} = 500$ GeV w/ $\gamma\gamma$ Overlay	$\sqrt{s} = 500$ GeV w/o $\gamma\gamma$ Overlay	$\sqrt{s} = 350$ GeV w/o $\gamma\gamma$ Overlay
Without Corrections	M_H (GeV)	123.994 ± 0.091	123.375 ± 0.090	123.540 ± 0.044
	M_Z (GeV)	91.206 ± 0.102	91.058 ± 0.096	90.966 ± 0.048
Semi-lep. Correction	M_H (GeV)	124.692 ± 0.095	124.215 ± 0.093	124.005 ± 0.048
	M_Z (GeV)	91.291 ± 0.106	91.162 ± 0.101	91.074 ± 0.050

The mass extraction method which is introduced in section 8.1 is employed to determine the masses of the bosons from the distributions of the reconstructed invariant mass. In case of the Higgs boson the range of $M_{Rec.} = 125.0 \pm 10$ GeV and for the Z boson the range of $M_{Rec.} = 91.0 \pm 5$ GeV is considered. The results are noted in table 8.1. For a centre-of-mass energy of 350 GeV, the reconstructed Z boson mass is 90.966 GeV which is ~ 220 Mev smaller than the input mass of $M_Z = 91.1882$ GeV with an uncertainty of 48 MeV. For the case of 500 GeV centre-of-mass energy without $\gamma\gamma$ overlay, the reconstructed Z mass is 91.058 ± 0.096 GeV, which is nearly 130 MeV smaller than the simulation input mass. The reconstructed mass of the Higgs boson is smaller than the simulation input mass by approximately 1.5 GeV for the samples of 350 GeV and 500 GeV without $\gamma\gamma$ overlay. These results are an indication of flavour dependence of the mass reconstruction which is further investigated in section 8.3.

The reconstructed Higgs boson masses of the 500 GeV samples with and without $\gamma\gamma$ overlay differ by about 600 MeV. The reconstructed Z mass using the sample with $\gamma\gamma$ overlay is also larger than the reconstructed mass using the sample without the overlay. These differences suggests possible problems with the $\gamma\gamma$ removal procedure (see section 6.5.2) and is studied further in section 8.5.3.

8.3 Reconstruction of Heavy Flavour Jets

As shown in the previous chapter, the reconstructed mass of the Higgs boson is underestimated by about 1.5 GeV while the reconstructed mass of the Z boson is only ~ 200 MeV smaller than the simulation input mass. In the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ Monte Carlo samples which are used for this measurement the Higgs boson decays exclusively to $b\bar{b}$ hadrons while Z decays to $b\bar{b}$ hadrons as well as other hadrons composed of lighter quarks. This can be an indication of possible issues with heavy flavour quark measurements. Although the Z boson also decays to heavy flavour hadrons, the sum of the branching fractions of $Z \to b\bar{b}$ and $Z \to c\bar{c}$ is only $\sim 28\%$ of the total hadronic decays of the Z boson [19]. Therefore, possible issues with heavy flavours are obscured by the dominating light hadrons.

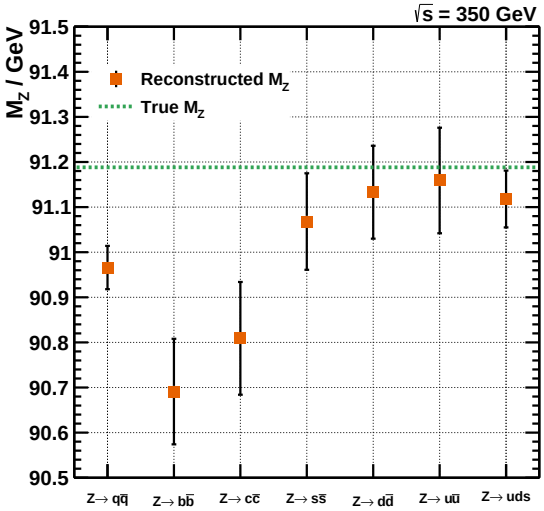


Figure 8.3 – Reconstructed Z boson mass using the χ^2-minimisation method for $Z \to q\bar{q}$, $Z \to b\bar{b}$, $Z \to c\bar{c}$, $Z \to s\bar{s}$, $Z \to d\bar{d}$, $Z \to u\bar{u}$ and $Z \to uds$ (Z boson decays to any of the light quark pairs) processes. While the reconstructed mass of the $Z \to d\bar{d}$ and $Z \to u\bar{u}$ samples are in agreement with the simulation input mass of Z boson, the reconstructed masses of $Z \to c\bar{c}$ and $Z \to b\bar{b}$ data samples are underestimated by $\sim$ 500MeV and $\sim$ 400MeV, respectively.

Since the mass of the Z boson is known to a very high precision, it can be employed to investigate possible issues with the heavy flavour measurements. To study these issues, the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ sample of 350 GeV centre-of-mass energy is divided to $Z \to b\bar{b}$, $Z \to c\bar{c}$, $Z \to s\bar{s}$, $Z \to d\bar{d}$, $Z \to u\bar{u}$ and $Z \to uds$ (Z boson decays to any of the light quark pairs) samples. The χ^2-minimisation method of section 8.2 is utilised to compute the reconstructed Z boson mass for each of these samples individually. The results are displayed in figure 8.3. It is clearly shown that while the reconstructed mass of the $Z \to q\bar{q}$ sample is $\sim$ 200 MeV smaller than the simulation input mass of the Z boson, for the $Z \to c\bar{c}$ and $Z \to b\bar{b}$ samples where heavy quarks are involved the reconstructed Z boson mass is underestimated by $\sim$ 500 MeV and $\sim$ 400 MeV, respectively. The reconstructed Z boson mass obtained from the $Z \to s\bar{s}$ is about 100 MeV smaller than the simulation input mass. The results obtained from the $Z \to d\bar{d}$ and $Z \to u\bar{u}$ samples are in agreement with the simulation input mass within the uncertainties. The Z boson mass from the $Z \to uds$ sample is very close to the simulation input mass. Since PandoraPFA is calibrated using $Z \to uds$ samples, it is expected that the reconstructed mass obtained using these channels is compatible with the input mass.

The results presented in this section show a clear correlation between the reconstructed mass and the flavour of the quarks to which the Z boson decays. These results explain the underestimated reconstructed mass of the Higgs boson which is computed in section 8.2. For the processes which involve heavy quark flavours, phenomena such as missing energy of neutrinos in the semi-leptonic decays affect the reconstructed mass of the parent particle. The issue with semi-leptonic decays is addressed in section 7.2.6 where the missing energy of neutrinos and its corresponding uncertainty on the jet energy is estimated. By applying

the semi-leptonic correction to the Monte Carlo samples and using the χ^2-minimisation method, the estimated Higgs boson mass improves by about 600 MeV, as the results noted in table 8.1 suggest. Nevertheless, the reconstructed Higgs boson mass is still smaller than the simulation input mass. Since the Z boson mass is known to a high precision, one possible way to correct for this underestimation is to exploit $Z \to b\bar{b}$ events to calibrate other decays to bottom quarks such as $H \to b\bar{b}$.

8.4 Kinematic Fitting

One of the major benefits of lepton colliders over hadron colliders is that the initial state of the colliding particles is known to a good precision. This allows for exploiting the conservation laws to improve the measurements beyond the detector jet energy resolution. A kinematic fitting technique is a well-established mathematical procedure that employs this information. In this section, after an introduction to the kinematic fitting concept and its formalism, the application of kinematic fitting in the precise measurement of the Higgs boson mass is presented.

8.4.1 Concept

In a given interaction, measured and unmeasured quantities are supposed to comply with certain kinematic requirements imposed by the conservation laws. The uncertainties on the measured quantities prevent these requirements from being fully satisfied. Kinematic fitting is a mathematical procedure which varies the quantities within their uncertainties in order to satisfy the kinematic requirements deduced from the event hypothesis. In this procedure, the least-square method is combined with the Lagrange multipliers technique to adjust the quantities [89].

Assuming that there are N measured quantities, a χ^2 function is constructed as:

$$\chi^2(\boldsymbol{\eta}) = (\boldsymbol{y} - \boldsymbol{\eta})^T \boldsymbol{V}^{-1} (\boldsymbol{y} - \boldsymbol{\eta}) \tag{8.3}$$

in which $\boldsymbol{y}$ is a N-dimensional vector of measurements, $\boldsymbol{\eta}$ is a N-dimensional vector of fit values of the measured quantities and $\boldsymbol{V}$ is the covariance matrix. There can be a number of unmeasured quantities as well. A new J-dimensional vector $\boldsymbol{\xi}$ is formed to hold the values of the J unmeasured quantities. The physical constraints imposed by the kinematic requirements are a function of the measured and unmeasured quantities. These constraints construct a K-dimensional vector $\boldsymbol{f}$, assuming there are K constraints. When a constraint is satisfied, the following relation holds:

$$\boldsymbol{f}_k(\eta_1, ..., \eta_N, \xi_1, ..., \xi_j) = 0, \quad k \in [1, K]. \tag{8.4}$$

The constraints are added to the χ^2 equation of 8.3 via the Lagrange multipliers technique. Hence, the final χ^2 equation to be minimised is:

$$\chi^2(\boldsymbol{\eta}, \boldsymbol{\xi}, \boldsymbol{f}) = (\boldsymbol{y} - \boldsymbol{\eta})^T \boldsymbol{V}^{-1} (\boldsymbol{y} - \boldsymbol{\eta}) - 2\boldsymbol{\lambda}^T \boldsymbol{f}(\boldsymbol{\eta}, \boldsymbol{\xi}) \tag{8.5}$$

in which $\boldsymbol{\lambda}$ is a K-dimensional vector of Lagrange multipliers. The partial second derivative of the χ^2 function is set equal to zero to find the critical points. The minimisation procedure is repeated iteratively until a minimum is found under some requirements such as a certain fit probability or variation of the χ^2.

A complete set of constraints form a fit hypothesis. If a hypothesis is correct and the corresponding measurement uncertainties are correctly modelled, a flat uniform distribution of the fit probabilities is expected. In addition, a pull distribution can be used as a measure of fit performance. In the case of kinematic fitting, the pull can be computed as:

$$pull(x) = \sum_i \frac{x_{Fit,i} - x_{Meas,i}}{\sqrt{\sigma^2_{Fit,i} - \sigma^2_{Meas,i}}} \tag{8.6}$$

where the sum is over all events, x_i^{Fit} is the kinematic variable after fit and $x_i^{Meas.}$ is the measured kinematic variable with an uncertainty of $\sigma_i^{Meas.}$. For a Gaussian distribution of a random variable x, in case of no bias, the pull distribution is expected to be Gaussian with mean zero and unit width.

8.4.2 Kinematic fitting in the context of $e^+e^- \to ZH \to q\bar{q}b\bar{b}$

iLCSoft provides an implementation of kinematic fitting in the MarlinKinfit package [90]. This package provides three different fit engines. In the case of this study the OPAL fit engine is used except for the cases with a soft mass constraint where the *new* fit engine is employed. The physical objects to perform the fit on are referred to as *fit objects* and are parametrised based on the measured quantities and their corresponding uncertainties.

After the event selection and preparations described in chapter 6, four jets are present in the final state of the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ events. Each of these jets forms a *jet fit object* which is parametrised as:

$$JFO_i(E_i, \theta_i, \phi_i, \sigma_{E,i}, \sigma_{\theta,i}, \sigma_{\phi,i}, m_i) \tag{8.7}$$

where JFO is a jet fit object, E is the measured energy of the jet with σ_E uncertainty, θ is the polar angle of the jet with σ_θ uncertainty, ϕ is the azimuthal angle of the jet with σ_ϕ uncertainty and m is the mass of the jet.

The measured quantities are provided by the jet clustering algorithm (see section 6.5.3). Previous studies have shown that for a standard parametrisation of jet uncertainties choosing $\sigma_\theta = \sigma_\phi = 100$ mrad and $\sigma_E = 120\%/\sqrt{E}$ give the best fit convergence [88]. In this study

the uncertainties on θ and ϕ are set to constant values of 100 mrad while the uncertainty of the jet energy σ_E is computed for each jet individually according to the method presented in chapter 7. A comparison of the results obtained by employing fixed and jet-specific energy uncertainties is presented in section 8.5.2.

As explained in section 8.4.1, additional information from the event hypothesis is provided to the kinematic fit in terms of constraints. In the case of this study, the following constraints can be deduced from the kinematics of the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ process:

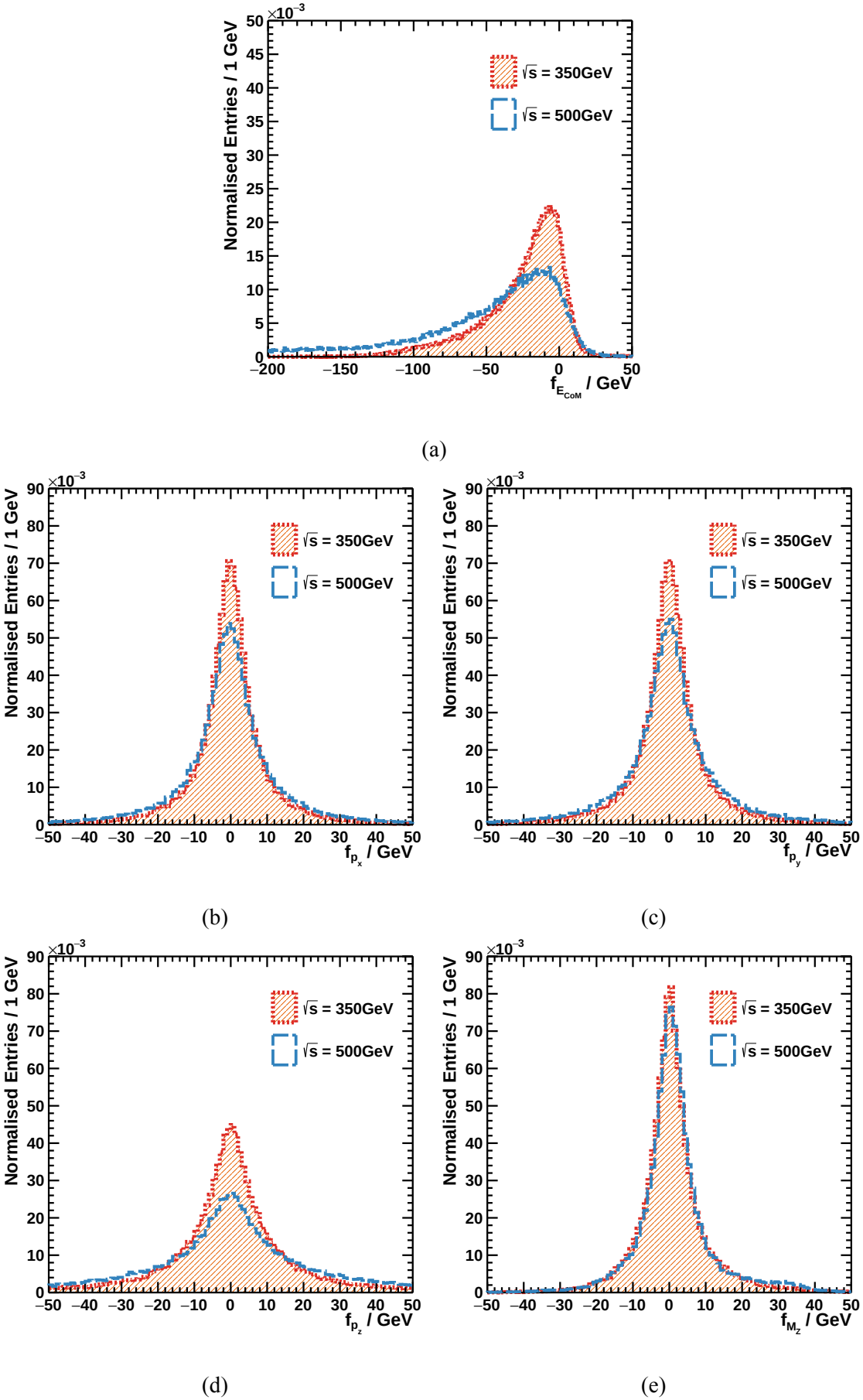
centre-of-mass energy conservation of energy imposes a constraint that the sum of the energies in a reconstructed event should be equal to the centre-of-mass energy of the collision. In the presence of ISR and beamstrahlung, the true centre-of-mass energy of the collision is smaller than the nominal. This has to be accounted for in the kinematic fit.

jet momenta in case of a head-on collision, conservation of momentum requires the vector sum of the momenta from all jets to be equal to zero. However, the beams at ILC collide with a crossing angle. Consequently, the constraint on the momenta needs to be modified, as explained in the following.

Z boson mass it is known from the event hypothesis that two of the jets should be originating from the Z boson. Since the mass of the Z boson has been measured to a high precision, it can be used as another constraint for the kinematic fit. Two types of mass constraints can be defined. The first type, known as a *hard* mass constraint, assumes the mass of the particle to be an exact value. However, the measured mass can deviate slightly from the nominal value within the natural width. To account for these deviations in the fit in case of the particles for which the natural width is sufficiently large to be resolved in the detector, the second type of mass constraint known as the *soft* mass constraint can be used.

The kinematic fitting procedure with various combinations of constraints is performed on the Monte Carlo samples of 350 GeV and 500 GeV without the $\gamma\gamma$ background overlay. For each event, kinematic fitting is performed on all six possible permutations of the four jets and the permutation with highest fit probability is taken as the best permutation from which the invariant mass is obtained.

Initial values of the various constraints used in the kinematic fitting are shown in figure 8.4 for the centre of mass energies of 350 GeV and 500 GeV. When a constraint is perfectly satisfied, it has a value of zero. It is clear from the distributions that for most of the events none of the constraints is satisfied. The constraints are discussed in detail in the following sections where the definition of each constraint and its value is explained. The results of kinematic fitting using these constraints are also presented.



(a)

(b)

(c)

(d)

(e)

Figure 8.4 – Distributions of the initial values of various constraints used in kinematic fitting for the centre-of-mass energies of 350 GeV and 500 GeV.

8.4.3 Constraint on the Momenta and Centre-of-mass Energy

The principles of conservation of energy and momentum provide a set of four constraints for the kinematic fitting, referred to as a $4C$ fit. From the conservation of energy, a constraint on the centre-of-mass energy ensures that the sum of energies of all fit objects in an event is equal to the centre-of-mass energy of the collision

$$f_{E_{CoM}} = \sum_i E_i - E_{CoM} \tag{8.8}$$

where the sum is over all fit objects, E_i is the total energy of each fit object and E_{CoM} is the nominal centre-of-mass energy of the collision. The parameter $f_{E_{CoM}}$ represents the difference between the sum of the measured energy of the all fit objects and the nominal centre-of-mass energy of the collision and should be zero in case of a perfectly fulfilled constraint.

The distribution of the initial $f_{E_{CoM}}$ for the permutations with the best fit probability is displayed in figure 8.4a. The distribution for both of the centre-of-mass energies is highly asymmetric with a long tail towards negative values. The sample of 500 GeV centre-of-mass energy has more events in the tail. According to equation 8.8, a negative $f_{E_{CoM}}$ value represent cases where the nominal centre-of-mass energy is larger than the sum of the measured energies of all fit objects. The distribution of $f_{E_{CoM}}$ is a clear indication of missing energy in the reconstructed events which is addressed in section 8.4.6.

In case of heads-on collisions, conservation of momentum requires the vector sum of all fit object momenta to be equal to zero:

$$\sum_i \boldsymbol{p}_i = 0 \tag{8.9}$$

where the sum is over all fit objects and $\boldsymbol{p}_i$ is the three-momentum of each fit object. Individual components of the momentum vector are expected to be zero:

$$p_x = p_y = p_z = 0. \tag{8.10}$$

However, the interaction region at ILC will have a horizontal crossing angle of 14 mrad between the beams which allows for extraction of the beams to separate beam pipes after the collision (see chapter 3). Therefore, the collisions are not head-on and this needs to be taken into account in the physics analyses, including kinematic fitting.

The coordinate system of ILD is described in section 3.3.1. In this coordinate system, the crossing angle is defined as the angle by which $\boldsymbol{p}^+$ has to be rotated around the y-axis such that it becomes anti-parallel to $\boldsymbol{p}^-$. Subsequently, the momentum constraint in the

kinematic fit is modified as below in order to include the effect of the crossing angle:

$$f_{p_x} = \sum_i p_{x,i} - \sin(\theta_c/2) \times E_{CoM}, \tag{8.11}$$

$$f_{p_y} = \sum_i p_{y,i}, \tag{8.12}$$

$$f_{p_z} = \sum_i p_{z,i}, \tag{8.13}$$

where the sum is over all fit objects, $p_{x,i}$, $p_{y,i}$ and $p_{z,i}$ are the Cartesian components of the fit object momentum, θ_c is the crossing angle and E_{CoM} is the nominal centre-of-mass energy of the collision. The parameters f_{p_x}, f_{p_y} and f_{p_z} represent the constraint values and should be zero when the constraint is perfectly fulfilled.

The distributions of f_{p_x}, f_{p_y} and f_{p_z} are shown in figures 8.4b, 8.4c and 8.4d, respectively. All three distributions are symmetric around zero while the distribution of f_{p_z} has longer tails on both sides as compared to the other two momentum constraints.

ISR and semi-leptonic decays are two possible sources of missing energy and hence the asymmetry of the $f_{E_{CoM}}$ and the tails of the f_{p_x}, f_{p_y} and f_{p_z} distributions. The methods of correcting for these missing energies are explained in section 8.4.5.

8.4.4 Mass Constraint

From the event hypothesis, it is known that two of the jets are produced by the decay products of the Z boson and therefore their invariant mass should be equal to the mass of the Z boson. This information can be exploited to set another constraint on the kinematic fit. This constraint is defined as:

$$f_{M_Z} = M_{i,j} - M_Z^{True} \tag{8.14}$$

where $M_{i,j}$ is the invariant mass of two of the jets which are assumed to originate from the Z boson, and $M_Z^{True} = 91.1882$ GeV is the simulation input mass of the Z boson. The value of the constraint on the Z boson mass f_{M_Z} represents the difference between the measured and simulation input mass of the Z boson and is equal to zero when the constraint is fulfilled.

The distribution of the initial values of f_{M_Z} (before performing kinematic fitting) for the permutation with best fit probability is shown in figure 8.4e. For both of the centre-of-mass energies, f_{M_Z} has a roughly symmetric distribution around zero, with a longer tail towards positive values.

After the kinematic fitting is performed, the four-momenta of the fit objects have been adjusted by the fit such that all the constraint values are equal to zero. Since the natural width of the Z boson can be resolved in the ILC detectors, a soft mass constraint can be

used to account for the width. Both hard and soft mass constraints are used in the kinematic fit hypotheses, as discussed in the following.

8.4.5 Corrections for ISR and Semi-leptonic Decays

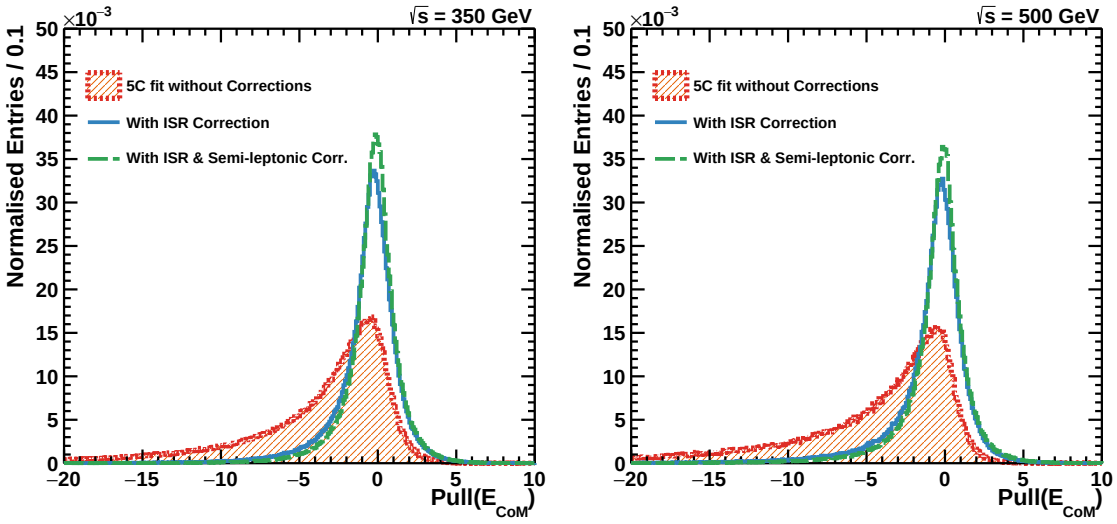
At ILC, a fraction of the momentum and energy of the beam particles can be carried away by photons along the beam due to the initial state radiation and beamstrahlung. Therefore, assuming a fixed constraint on the momenta and the centre-of-mass energy in kinematic fitting is not valid for the events with ISR.

A method has been developed to correct the effects of the ISR in kinematic fitting [91]. This correction is achieved by introducing a photon fit object which represents the ISR photons. Since the transverse momentum of the ISR photons is approximately zero, they affect only energy and longitudinal momentum conservation constraints. Therefore, the ISR photon fit object is parametrised by its longitudinal momentum $p_{z,\gamma}$ as the only free unmeasured parameter while $p_{x,\gamma} = p_{y,\gamma} = 0$ and $E_\gamma = |p_{z,\gamma}|$. The parametrisation uses an energy spectrum which follows a power law [92] for the probability of the ISR photon energy. This approach allows the constraints on the centre-of-mass energy and momenta to be recovered at the cost of introduction of a new parameter.

The performance of the ISR correction method has been investigated on a $e^+e^- \to u\bar{d}d\bar{u}$ sample at the centre-of-mass energy of 500 GeV using a full simulation of the ILD detector [91]. This study shows that by introducing the photon fit object the convergence rate of the kinematic fits with energy and momentum constraints improves without deteriorating the resulting resolution. The effectiveness of this method on $H \to b\bar{b}$ events has been demonstrated in [93].

Semi-leptonic decays are another source of problems with fixed constraints on momenta and centre-of-mass energy in kinematic fitting. A method to correct for the missing energy of undetected neutrinos in semi-leptonic decays is explained in section 7.2.6. This method has been used to correct jet energies prior to performing kinematic fitting.

The distribution of the pulls on the centre-of-mass energy with and without corrections is shown in figure 8.5. For the case of a kinematic fit with constraints on the centre-of-mass energy, the total momentum and the Z boson mass (a so-called $5C$ fit), using jet-specific energy resolution and without any corrections at the centre-of-mass energy of 350 GeV (500 GeV), the pull distribution is non-Gaussian with the most-probable-value of -3.47 (-4.13) and RMS width of 4.33 (4.71) and has a long tail towards negative values. By including the ISR in the fit, the pull distribution becomes more Gaussian-like, the most-probable-value moves to -0.69 (-0.91) and the distribution becomes narrower to a width of 2.21 (2.60). Applying the semi-leptonic correction further improves the pull distribution with a mean of -0.35 (-0.52) and width of 1.83 (2.17).



(a) (b)

Figure 8.5 – Distribution of the pull on the centre-of-mass energy for $5C$ kinematic fitting with and without ISR and semi-leptonic corrections for the centre-of-mass energies of a) 350 GeV and b) 500 GeV.

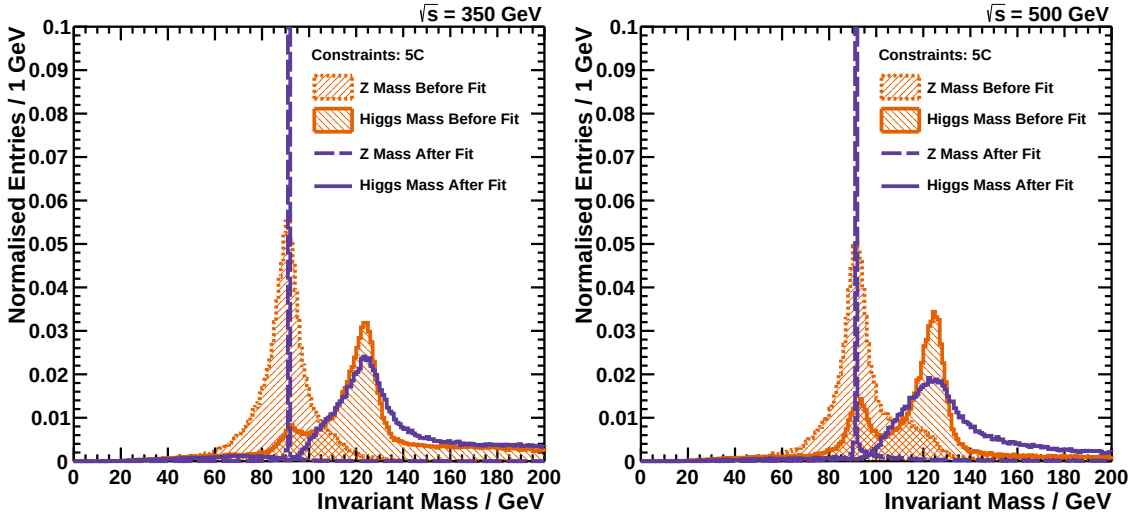
8.4.6 Hypothesis Testing

The various constraints discussed in the previous section are combined to form a set of fit hypotheses. These hypotheses have been employed in kinematic fits on the MC samples of 350 GeV centre-of-mass energy and 500 GeV centre-of-mass energy without $\gamma\gamma$ overlay and the outcome is presented in this section. The results obtained using the sample with $\gamma\gamma$ overlay is discussed in section 8.5.3.

$5C$ Hypothesis

Constraints on the centre-of-mass energy of the collision and the total momentum form a fit hypothesis, known as a $4C$ fit. These constraints apply to the sum of all four jets and the fit has no means to identify which pair belongs to which of the bosons, thus this fit hypothesis can not be used in the current analysis unless jet pairing is determined before performing the fit. The constraint on the Z boson mass enables the kinematic fit to identify the correct jet pairings.

By adding the constraint on the Z boson mass to the constraints on the centre-of-mass energy and total momentum, a $5C$ fit hypothesis is formed. The invariant mass distribution resulting from performing a $5C$ kinematic fit on the Monte Carlo samples of the two centre-of-mass energies is shown in figure 8.6. The filled histograms are the distributions of the invariant masses of the two bosons for the jet pairings with the highest fit probability but using initial parameters of the jets. The remaining two histograms show the result obtained by using the adjusted four-momenta for the same jet pairings. By performing the



(a) (b)

Figure 8.6 – Distributions of the reconstructed invariant mass of the Z and the H bosons using the $5C$ hypothesis for the centre-of-mass energies of a) 350 GeV and b) 500 GeV. The filled histograms are the invariant mass of the jet pairing with best fit probability but using initial jet parameters. The other two histograms shows the results of kinematic fitting on the same pairings.

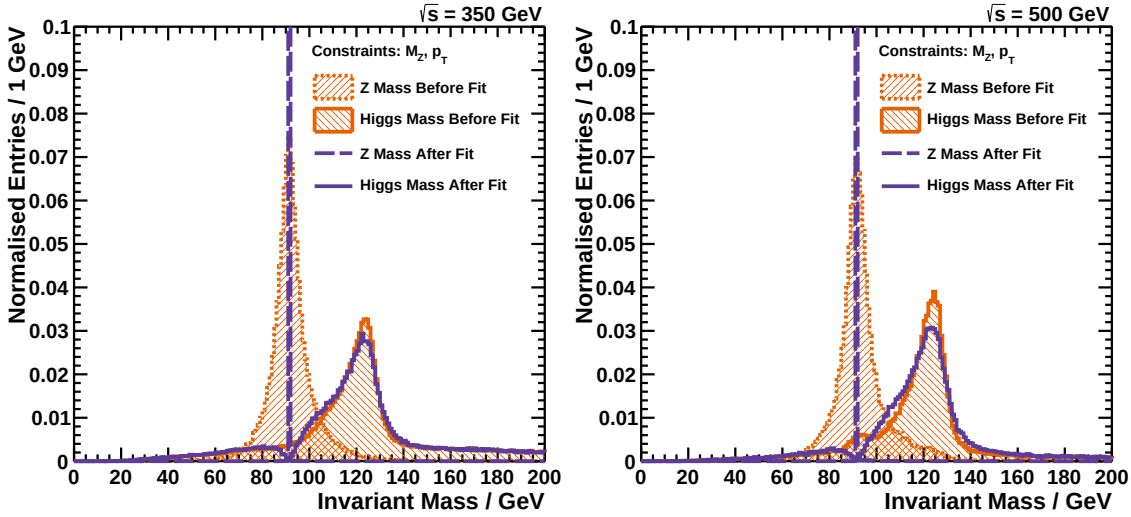
kinematic fit with a hard constraint on the Z boson mass, the fit forces the Z boson mass to be equal to the requested value. Therefore, all of the entries in the Z mass distribution are accumulated in one bin of the histogram.

The solid line histogram with a peak above ~ 120GeV is the distribution of the invariant mass associated with the Higgs boson after kinematic fitting. The tail towards higher invariant masses is due to wrong jet pairings. For both centre-of-mass energies, the tail ends at $M = E_{CoM} - M_Z$ as expected.

Another point worth noticing is that in the distribution of the Higgs boson mass before kinematic fitting in figure 8.6, there is a second peak around the Z boson mass which disappears in the distribution of the Higgs boson mass after performing kinematic fitting. Since the two distributions are for the same jet pairings and the pairs which form the second peak around the Z boson mass are mostly likely originating from the Z boson, this demonstrates that the adjustments performed by the kinematic fitting can lead to errors in assigning the jet pairs to the bosons or even errors in choosing the correct permutation.

Z Boson Mass and Transverse Momentum Hypothesis

The asymmetry in the distribution of the initial $f_{E_{CoM}}$ (see figure 8.4) indicates that the constraints on the centre-of-mass energy and the total momenta are not valid at least for a fraction of the events. This is expected due to the presence of ISR and beamstrahlung in the events. In order to further investigate this issue, the kinematic fitting is performed with



(a) (b)

Figure 8.7 – Distributions of the reconstructed invariant mass of the Z and the H bosons using the $3C$ hypothesis for the centre-of-mass energies of a) 350 GeV and b) 500 GeV. The filled histograms are the invariant mass of the jet pairing with best fit probability but using initial jet parameters. The other two histograms shows the results of kinematic fitting on the same pairings.

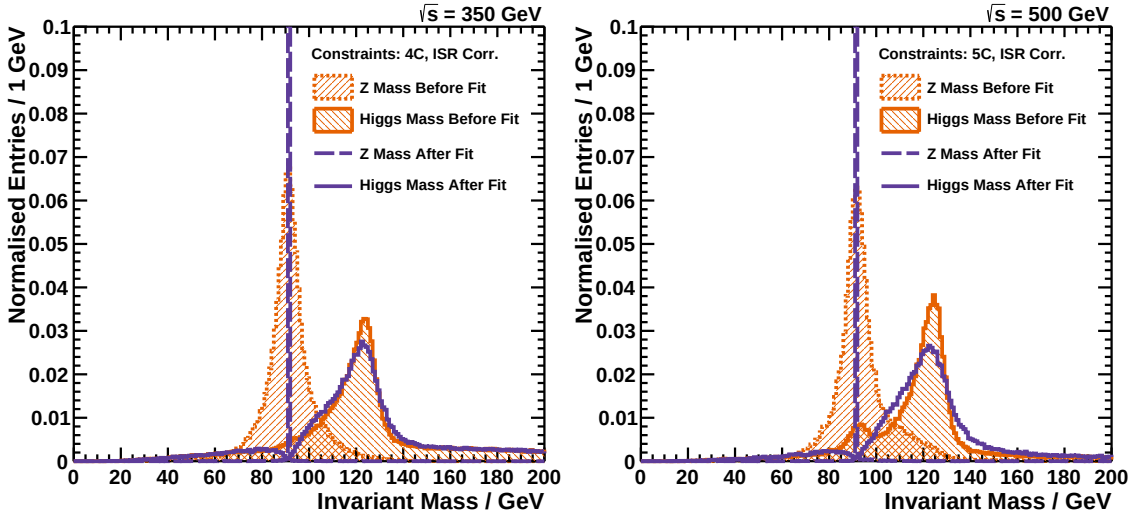
a hard constraint on the Z boson mass and a constraint only on the transverse momentum, thus a $3C$ hypothesis, results of which is shown in figure 8.7.

The distribution of the Higgs boson mass after kinematic fitting with constraints on the Z boson mass and the transverse momentum is improved and the Higgs boson mass obtained here is closer to the simulation input mass as compared to the case of the $5C$ kinematic fit. These results confirm that the constraints on the centre-of-mass energy and z-component of the momenta that is along the beams is not valid for all of the events, which in turn points to an amount of missing energy in the events.

$5C$ with ISR Correction

A fraction of the centre-of-mass energy is carried away by the Initial State Radiation photons. Therefore, setting a constraint on the centre-of-mass energy to be exactly 350 GeV or 500 GeV is not valid for all events. As explained in section 8.4.5 including an ISR photon fit object in kinematic fit can correct for the missing energy of the ISR.

By accounting for ISR in the fit, the peak of the distribution of the reconstructed Higgs boson mass improves significantly with respect to the fit with the same constraints but without the ISR correction, as shown in figure 8.8. The distributions of the invariant mass before fit are also improved, demonstrating that a better jet pairing is achieved using this fit hypothesis. The improvement in the case of the 500 GeV centre-of-mass energy is more noticeable.



(a) (b)

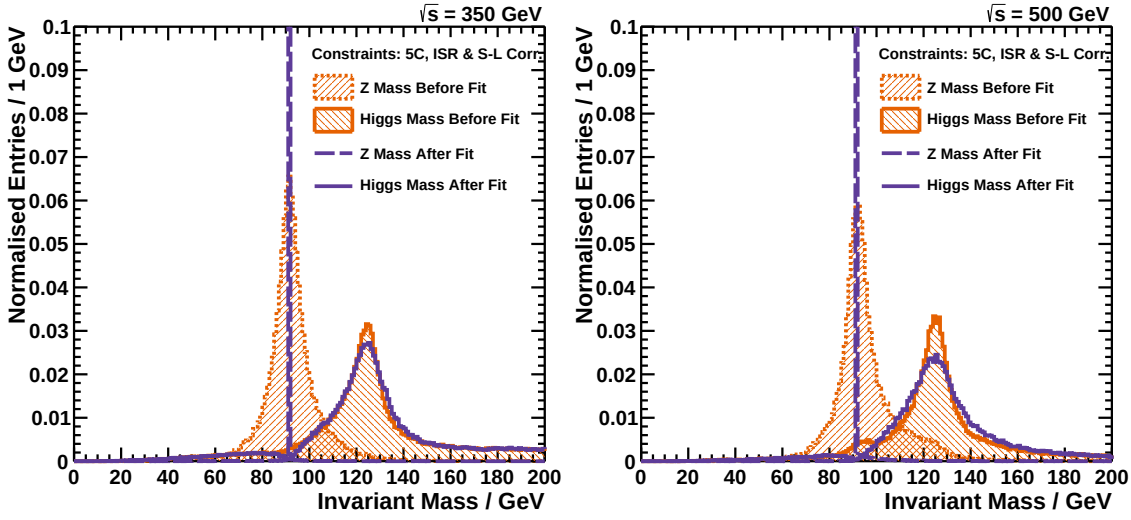
Figure 8.8 – Distributions of the reconstructed invariant mass of the Z and the H bosons using the $5C$ + ISR hypothesis for the centre-of-mass energies of a) 350 GeV and b) 500 GeV. The filled histograms are the invariant mass of the jet pairing with best fit probability but using initial jet parameters. The other two histograms shows the results of kinematic fitting on the same pairings.

The masses obtained after the ISR correction are smaller than all other results. The reason for this is that the missing energy of neutrinos in the semi-leptonic decays is also considered to be carried away by the ISR photons. By correcting jet energies for missing neutrino energies in the semi-leptonic decays the fit results improve, as shown in the following section.

$5C$ with ISR and Semi-leptonic Corrections

For the events with semi-leptonic decays, the constraint on the centre-of-mass energy is not valid since the missing energy of neutrinos produced in semi-leptonic decays are not considered. Therefore, the corrections mentioned in section 8.4.5 are necessary. By correcting measured jet energies according to equation 7.12 before performing kinematic fitting, an improvement in the results of the kinematic fit is expected. The outcome of applying the semi-leptonic correction and considering ISR in the $5C$ kinematic fit is displayed in figure 8.9. The distributions are improved compared to the case where only the ISR correction is performed. The correction works best for the case of the 350 GeV centre-of-mass energy. This results show a significant improvement over the fit where only ISR correction is applied.

The $5C$ kinematic fit with ISR and semi-leptonic corrections is also performed using a soft constraint on the Z boson mass. For both centre-of-mass energies, the reconstructed Higgs boson mass obtained by this fit agrees within the uncertainties with the correspond-



(a) (b)

Figure 8.9 – Distributions of the reconstructed masses of the Z and the H bosons using the $5C$ + ISR hypothesis and applying the semi-leptonic corrections for the centre-of-mass energies of a) 350 GeV and b) 500 GeV. The filled histograms are the invariant mass of the jet pairing with best fit probability but using initial jet parameters. The other two histograms shows the results of kinematic fitting on the same pairings.

ing fit with a hard constraint. The uncertainty on the reconstructed Higgs mass is also comparable to the case with the hard constraint on the Z boson mass.

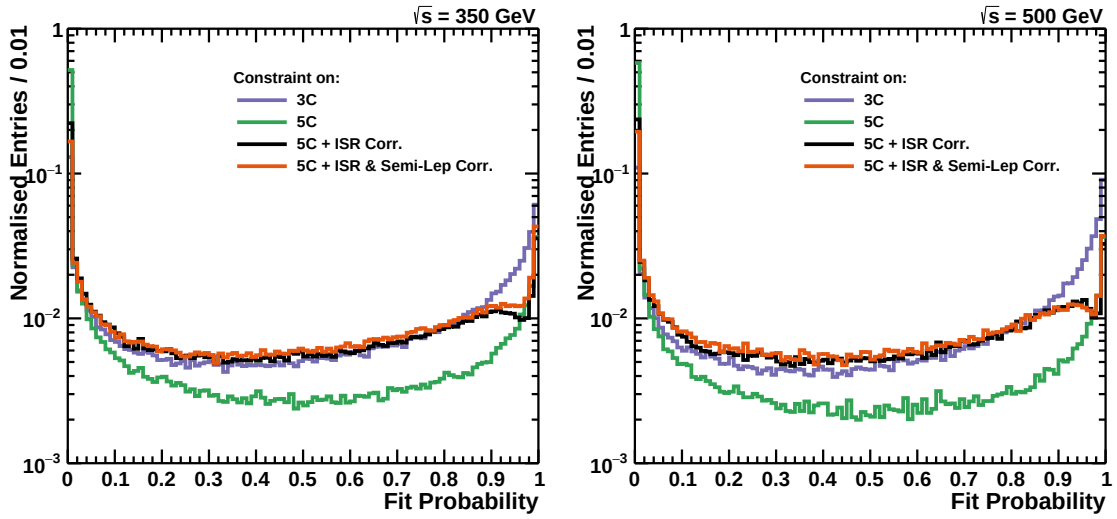
Fit Probability of Various Hypotheses

When a fit hypothesis is correct and measurement uncertainties are accurately modelled, a uniform distribution of fit probabilities is expected. The distributions of fit probabilities for each fit hypothesis are displayed in figure 8.10.

For the case of the $3C$ fit, the distribution is nearly flat between $\sim 15\%$ and $\sim 50\%$ and then slowly rises, creating peaks at the two ends of the distribution. Performing the $5C$ kinematic fitting significantly increases the events with zero fit probability and lowers the distribution in the range above $\sim 5\%$. However, applying the ISR correction in the $5C$ kinematic fit causes a significant improvement, as is clearly visible in figure 8.10.

The probability distribution for the $5C$ fit with ISR and semi-leptonic corrections is rather similar to the case with only ISR correction. The last two cases improve the fit probability distribution at higher probabilities as compared to the case of the $3C$ kinematic fit. This improvement is observed for both centre-of-mass energies. In case of 500 GeV, the distribution is improved at lower probabilities as well.

The kinematic fit is performed using fixed uncertainties on the polar and azimuthal angles. The large uncertainty values used here can lead to the peak at high probabilities.



(a) (b)

Figure 8.10 – Distributions of the fit probability for kinematic fitting with various constraints for the centre-of-mass energies of a) 350 GeV and b) 500 GeV.

Thus, further improvements can be expected once in addition to the jet energy uncertainty, angular uncertainties are also determined using the Error Flow approach.

8.5 Results and Discussion

The results of the aforementioned studies are summarised and discussed in this section. First, the reconstructed Higgs boson masses and their corresponding uncertainty obtained using various analyses are summarised and discussed. Next, the performance of the jet-specific energy resolution obtained by the error flow is compared with the case with fixed resolution. Finally, the effect of the $\gamma\gamma$ overlay and the performance of the removal procedure is discussed.

8.5.1 Higgs Boson Mass

There are three main steps in the Higgs boson mass estimation using the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel:

- to find the correct jet pairing,
- to find the best estimate of the jet four-vectors, and finally
- to extract the Higgs boson mass from the invariant mass distribution.

Various strategies to address the aforementioned challenges have been tested, as explained in the previous sections of this chapter. The first method which is presented in section 8.2 employs a χ^2-minimisation method to determine jet pairings according to the event hypothesis and then determines the mass of the parent boson by computing the invariant mass of the jet pair system. The second method benefits from the conservation laws in the context of kinematic fitting in order to further improve the precision of the measurement by adjusting the four-vectors of the jets. Various hypotheses are evaluated in the context of kinematic fitting.

These two methods are applied to two Monte Carlo samples with the centre-of-mass energies of 350 GeV and 500 GeV. The event preparation procedures are explained in details in chapter 6. The results obtained by performing the various methods on these samples are summarised in table 8.2 and in figure 8.11. In all cases, the reconstructed masses obtained from the kinematic fit on the two samples of 350 GeV and 500 GeV centre-of-mass energy without $\gamma\gamma$ overlay are in agreement within uncertainties.

The reconstructed mass of the Higgs boson obtained by using the χ^2-minimisation method is 123.542 ± 0.044 and 123.375 ± 0.090 for the centre-of-mass energies of 350 GeV and 500 GeV, respectively. These two values are approximately 1.5 GeV smaller than the simulation input mass of the Higgs boson. As discussed in section 8.3 for the case of the Z boson, one reason for this underestimation is due to the issues with heavy flavour quarks such as semi-leptonic decays. By applying the semi-leptonic corrections the underestimation becomes smaller by about 500 MeV.

For each kinematic fit hypothesis, the reconstructed Higgs boson masses obtained by performing the kinematic fit on the two samples are compatible within the uncertainties. The reconstructed masses using the $3C$ hypothesis are ~ 1.5 GeV smaller than the simulation input mass, very similar to the results obtained using the χ^2-minimisation method.

In the case of the $5C$ fit hypothesis where the centre-of-mass energy, the Z boson mass and the momenta are constrained, the reconstructed Higgs boson mass improves significantly and is only roughly 400 MeV smaller than the simulation input mass for both samples. Compared to the $3C$ fit hypothesis, the uncertainty on the reconstructed mass improves for the case of 350 GeV and deteriorates for the case of 500 GeV. This different behaviour can be explained by the different amount of ISR at the two centre-of-mass energies. Including the ISR correction in the $5C$ kinematic fitting improves the uncertainty by 17 MeV on the 500 GeV sample while the uncertainty of 350 GeV sample improves by only 1 MeV.

Although including ISR in the $5C$ kinematic fit improves the uncertainty, it causes the reconstructed Higgs boson mass to be smaller than the simulation input mass by about 800 MeV which is worse than the case of $5C$ without ISR correction. The reason is that the missing energy of neutrinos in semi-leptonic decays is also assumed to be carried away by the ISR photons. By correcting the jet energies for missing neutrino energies prior to performing the kinematic fit, the value of the reconstructed Higgs boson mass improves

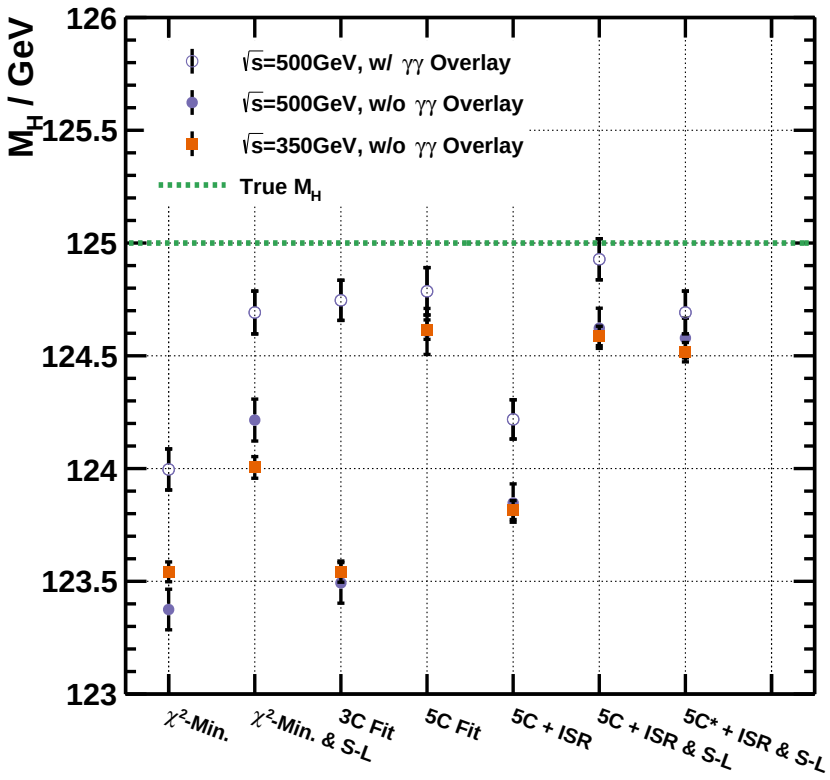
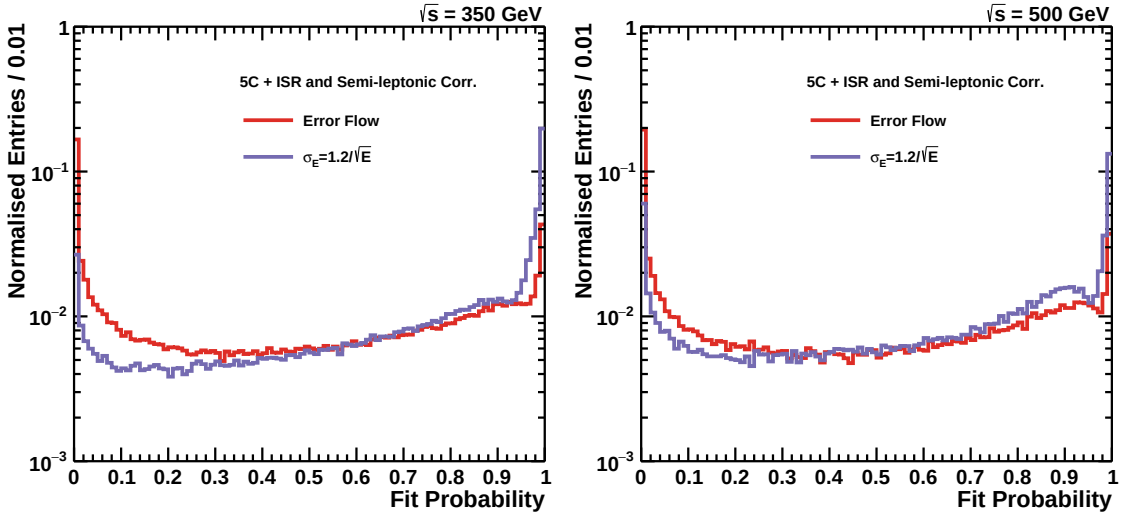


Figure 8.11 – Higgs boson mass obtained from the χ^2-minimisation method and kinematic fitting with various sets of constraints. The analysis is performed on the MC samples of 350 GeV and 500 GeV centre-of-mass energy. One of the samples of 500 GeV centre-of-mass energy is with $\gamma\gamma$ overlay and one is without $\gamma\gamma$ overlay.

Table 8.2 – Summary of the reconstructed Higgs boson mass obtained using the χ^2-minimisation method and kinematic fitting with various fit hypotheses for an integrated luminosity of 1000 fb^{-1} and a beam polarisation of $(-0.8,+0.3)$. The asterisk in the kinematic fit hypothesis titles indicates a soft constraints on the Z boson mass.

Sample	$\sqrt{s} = 500$ GeV w/ $\gamma\gamma$ Overlay	$\sqrt{s} = 500$ GeV w/o $\gamma\gamma$ Overlay	$\sqrt{s} = 350$ GeV w/o $\gamma\gamma$ Overlay
χ^2-min.	123.996 ± 0.091	123.375 ± 0.090	123.542 ± 0.044
χ^2-min. w/ Semi-lep. Corr.	124.692 ± 0.095	124.215 ± 0.093	124.005 ± 0.048
$3C$	124.746 ± 0.089	123.493 ± 0.090	123.543 ± 0.047
$5C$	124.786 ± 0.105	124.608 ± 0.102	124.616 ± 0.043
$5C$ + ISR	124.218 ± 0.087	123.847 ± 0.085	123.826 ± 0.042
$5C$ + ISR w/ Semi-lep. Corr.	124.928 ± 0.091	124.622 ± 0.089	124.595 ± 0.042
$5C^*$ + ISR w/ Semi-lep. Corr.	124.692 ± 0.095	124.579 ± 0.088	124.516 ± 0.043



(a) (b)

Figure 8.12 – Comparison of the kinematic fit probability using the jet-specific energy uncertainty and a fixed uncertainty of $120\%/\sqrt{E}$. The analysis is performed on the MC samples of 350 GeV and 500 GeV centre-of-mass energy with $5C$ fit hypothesis with ISR and semi-leptonic corrections.

significantly while the uncertainty slightly deteriorates. Nevertheless, the reconstructed Higgs boson mass obtained using this method is still smaller than simulation input mass by approximately 400MeV.

8.5.2 Error Flow Performance

As mentioned earlier, a previous study has shown that using a fixed relative jet energy uncertainty of $120\%/\sqrt{E}$ for all jets results in the best convergence rate of the kinematic fitting [88], but not the best mass resolution. In the studies which are discussed in section 8.4, the jet-specific energy uncertainty computed based on the error flow approach (see chapter 7) is employed in the kinematic fit. In order to evaluate the performance of the jet-specific energy uncertainty, the same procedures of section 8.4 are repeated with a fixed relative jet energy uncertainty of $120\%/\sqrt{E}$. In both cases the uncertainties on the polar and azimuthal angles are considered to be fixed and equal to 100 mrad for all jets.

The results show that the fit converges for 70% of the events for the case where the jet-specific energy uncertainty is used. The fit convergence rate is about 78% for the kinematic fit with a fixed uncertainty on the jet energy, which is higher than the fit with the jet-specific energy uncertainty. However, the distributions of fit probabilities displayed in figure 8.12 show that in the case of the fixed jet energy uncertainty, there is a significant increase in number of events with high fit probabilities, with a clear peak at the maximum probability. In the case where the jet-specific energy uncertainty is employed,

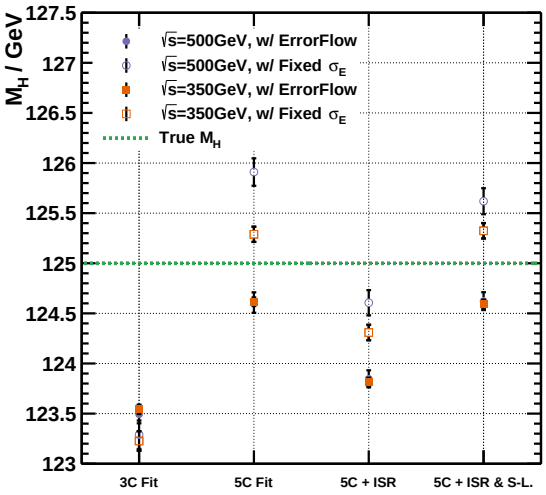


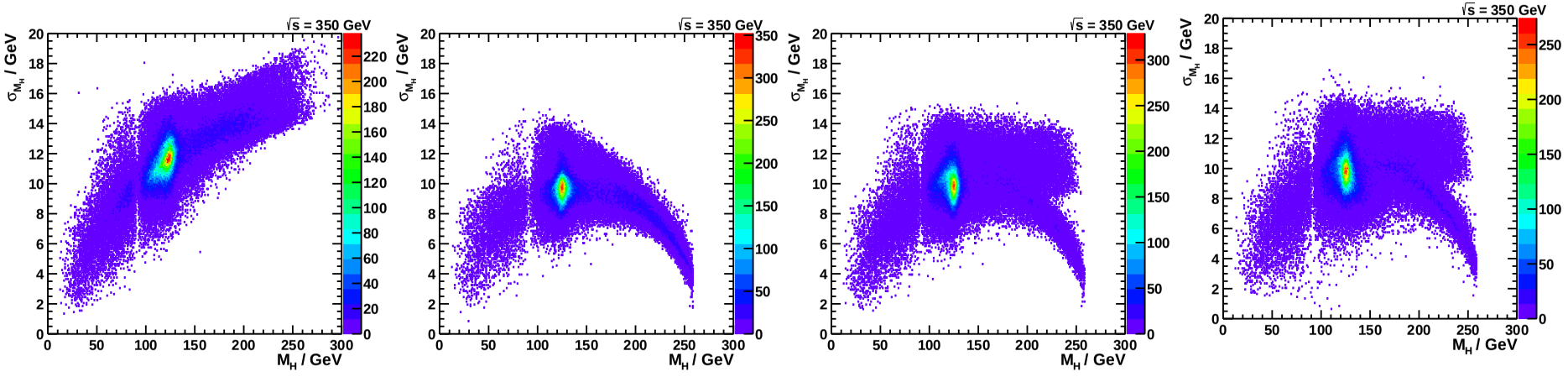
Figure 8.13 – Comparison of the Higgs boson mass obtained using kinematic fitting using the jet-specific energy uncertainty and a fixed uncertainty of $120\%/\sqrt{E}$. The analysis is performed on the MC samples of 350 GeV and 500 GeV centre-of-mass energy.

the distribution is improved, although it is still far from the expected flat distribution. The difference is more noticeable in the case of 350 GeV centre-of-mass energy.

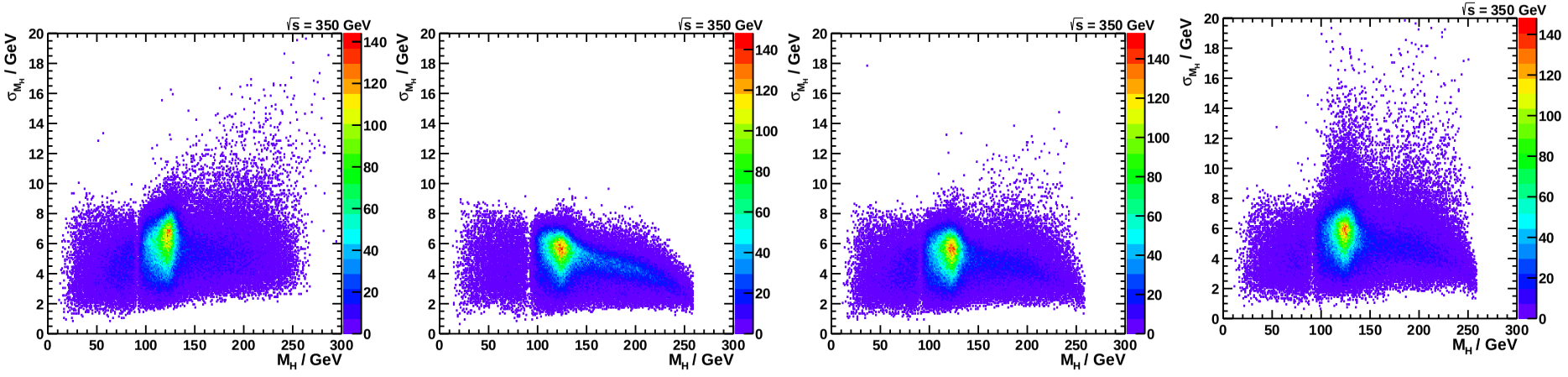
In both cases, the kinematic fit is performed with fixed uncertainties on the angular and azimuthal angles $\sigma_\theta = \sigma_\phi = 100$ mrad. This value was optimised to increase the converges rate [88] and is very likely to be unrealistically large, thus leading to the rise of the distributions towards the higher probabilities. Including a jet-specific angular uncertainty in the error flow should further improve the probability distribution. Moreover, as discussed in chapter 7, the energy uncertainties due to the hadronisation process and the jet clustering are not currently accounted for in the error flow procedure. It is expected that including these uncertainties should result in a more flat probability distribution.

For both of the centre-of-mass energies, the mass obtained using the fixed energy uncertainty using the kinematic fit with $5C$ hypothesis, with and without corrections, is larger than the mass obtained by using the jet-specific energy uncertainty by approximately $0.5 - 1.0$ GeV, as displayed in figure 8.13. Unlike the case with jet-specific energy resolution, the reconstructed masses for the two different centre-of-mass energies are not compatible within uncertainties. However, for the case of the kinematic fit with the $3C$ hypothesis all four reconstructed masses agree within their uncertainties.

The distributions of the uncertainty on the invariant mass of the Higgs boson as a function of the invariant mass for various fit hypotheses are displayed in figure 8.14. The top row shows the distributions for various fit hypotheses using the fixed jet energy uncertainty while their corresponding distributions using the jet-specific energy uncertainty are shown on the bottom. The distribution of the kinematic fit with $3C$ hypothesis using fixed jet energy uncertainty shows that the average uncertainty on the invariant mass increases with the invariant mass. By including the constraint on p_z in the $5C$ hypothesis,



(a) $3C$ with fixed σ_E (b) $5C$ with fixed σ_E (c) $5C$ + ISR with fixed σ_E (d) $5C$ + ISR and Semi-leptonic correction with fixed σ_E



(e) $3C$ with Error Flow (f) $5C$ with Error Flow (g) $5C$ + ISR with Error Flow (h) $5C$ + ISR and semi-leptonic correction with Error Flow

Figure 8.14 – Distributions of the uncertainty on the invariant mass of the Higgs boson as a function of the invariant mass for various fit hypothesis using fixed and jet-specific energy uncertainty.

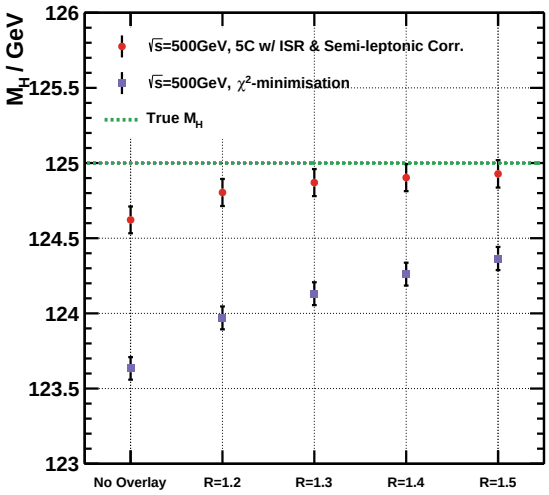


Figure 8.15 – Comparison of the effect of $\gamma\gamma$ background removal on the Higgs boson mass obtained using the χ^2 minimisation method and $5C$ kinematic fitting with ISR and semi-leptonic corrections.

the average uncertainty changes its behaviour for invariant masses above the simulation input Higgs boson mass and decreases with the invariant mass. By performing the $5C$ kinematic fit with ISR and semi-leptonic corrections, the distributions are slightly improved but the peculiar shape at higher values is still visible. In the cases which use the jet-specific energy uncertainty all distributions are significantly improved and are nearly flat. Moreover, the average uncertainty on the invariant mass per event is approximately 4 GeV smaller than the cases with fixed jet energy uncertainty.

8.5.3 Effect of the $\gamma\gamma$ Overlay

In order to evaluate the performance of the $\gamma\gamma$ background removal procedure explained in section 6.5.2, the analyses methods explained in the previous sections are applied to two Monte Carlo samples of 500 GeV centre-of-mass energy. One of the samples is with $\gamma\gamma$ background overlay and an attempt is made to remove the overlay in the event preparation phase as described in section 6.5.2. The second sample does not have $\gamma\gamma$-overlay. The samples are otherwise identical.

The results presented in table 8.2 and in figure 8.11 show that in all cases, the Higgs boson mass obtained from the sample with $\gamma\gamma$ background overlay is systematically larger than the mass measured using the sample without the overlay. In order to verify that this difference originates from the $\gamma\gamma$ background overlay, the χ^2-minimisation method and the $5C$ kinematic fit with the ISR and the semi-leptonic corrections are performed on the Monte Carlo sample of 500 GeV centre-of-mass energy with $\gamma\gamma$ overlay with various values for the R parameter of the exclusive jet finding algorithm (see section 6.5.2). The results

presented in figure 8.15 clearly show that the reconstructed Higgs boson mass depends on the value of R used in $\gamma\gamma$ background removal procedure.

The results presented in this section indicate that the approach which is used to remove the $\gamma\gamma$ background has limitations and needs to be improved. In order to formulate an efficient $\gamma\gamma$ background removal strategy, a better understanding of the $\gamma\gamma$ background at ILC is crucial. This issue is currently being studied in details and the outcome will be used to develop the necessary strategy and tools to remove $\gamma\gamma$ background at ILC.

8.5.4 Systematic Uncertainties

The uncertainties on the Higgs boson mass quoted throughout this chapter are statistical uncertainties. Although systematic uncertainties are not studied in detail here, a number of possible sources of systematic uncertainties are discussed in the following:

detector calibration one of the most important possible sources of systematic uncertainties is the calibration of the detector. A general-purpose detector such as ILD comprise multiple subdetectors. If not calibrated to a high accuracy, each of these subdetectors can contribute to the systematic uncertainties on the measurements.

flavour tagging the analysis performed here is on preselected $e^+e^- \rightarrow ZH \rightarrow q\bar{q}b\bar{b}$ signal samples in which the Higgs boson decays dedicatedly to b quarks. In principle, the Higgs boson can decay to all lighter quarks as well. Thus, errors in tagging of the quarks can lead to systematic uncertainties. A set of software tools for flavour tagging at linear e^+e^- colliders have been developed which demonstrate improved performance over previous flavour tagging software [94]. Further improvement of these software tools would minimise the uncertainty due to misidentification of the quark flavours.

particle identification the semi-leptonic correction relies on the identification of charged leptons in hadronic jets. Misidentification of the charged leptons can cause biased corrections and thus systematic uncertainty on the jet energy. New generation of particle identification tools for ILC are under active development [85–87]. Using current versions of these tools electrons and muons can be identified with better than 90% and 70% efficiency, respectively [85]. It is hoped that the efficiencies further improve in the future.

PFA callibration miscalibration of the PFA and unrealistic estimation of the confusion uncertainty can result in a biased energy measurement and incorrect estimation of the jet energy.

centre-of-mass energy the centre-of-mass energy of the collisions is used as a constraint on the kinematic fit. Hence, an incorrect measurement of the centre-of-mass energy can be a source of systematic uncertainty.

hadronisation model theoretical models used to describe the hadronisation process are developed based on phenomenology. The uncertainties of these model can affect jet measurements and in turn result in a biased Higgs boson mass measurement. The impact of the hadronisation models as a source of systematic uncertainty on the particle mass measurements has been investigated in the past, for instance for the W mass measurement at LEP [95]. The systematic uncertainty due to hadronisation was estimated to be $18 - 19$ MeV. The hadronisation uncertainties would be reduced as the theoretical models improve.

jet clustering the efficiency of the jet finding algorithm to correctly cluster particles into jets plays an important role in this analysis. Problems such as overlapping jets contribute to the uncertainty of the reconstructed Higgs boson mass, as explained in section 7.2.3.

invariant mass fitting the Higgs boson mass is extracted from the invariant mass distribution. The method used to extract the Higgs boson mass can give rise to systematic uncertainty.

$\gamma\gamma$ backgrounds remnants of background events clustered in jets can increase the measured jet energy and thus the Higgs boson mass. An example of this is the $\gamma\gamma$ background which is discussed earlier in this chapter.

8.6 Summary

The studies presented in this chapter demonstrate that with an integrated luminosity of 1000 fb^{-1} and a beam polarisation of $(-0.8,+0.3)$, the Higgs boson mass can be measured with a statistical uncertainty better than 100 MeV for the centre-of-mass energies of 350 GeV and 500 GeV. The uncertainty of 42 MeV obtained at the centre-of-mass energy of 350 GeV is more than a factor of two better than 89 MeV uncertainty at 500 GeV. This results are obtained by performing a $5C$ kinematic fit with ISR and semi-leptonic corrections. However, in both cases the reconstructed Higgs boson mass is smaller than the simulation input mass by approximately 400 MeV. The mass of the Z boson is known to a high precision. Therefore, $Z \to b\bar{b}$ events can be used to calibrate the jet energy scale for $H \to b\bar{b}$ events, hence further improving the Higgs boson mass measurement from $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ process. In addition, the semi-leptonic correction which is employed in this study is a simple approach and a more sophisticated approach could possibly improve the results. The semi-leptonic correction uses Monte Carlo truth information to identify leptons in the jets. An efficient particle identification capability is very important for exploiting this correction on experimental data.

A number of operating scenarios are considered for a 20 years physics programme at ILC [96]. For the H-20 operating scenario of ILC where an integrated luminosity of 200 fb^{-1} at 350 GeV and 4000 fb^{-1} at 500 GeV is foreseen, the combined uncertainty

on the Higgs boson mass measurement from the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel at these centre-of-mass energies is about 40 MeV. The combined uncertainty shrinks down to 26 MeV for the I-20 scenario.

Both of these scenarios include data taking at the centre-of-mass energy of 250 GeV. It is estimated that for the H-20 scenario where an integrated luminosity of 2000 fb^{-1} is foreseen, the Higgs boson mass can be measured with an uncertainty of approximately 14 MeV using the recoil technique at this centre-of-mass energy [58]. This uncertainty scales to about 28 MeV for the case of I-20 scenario which has only 500 fb^{-1} integrated luminosity at the centre-of-mass energy of 250 GeV. Combining these uncertainties with the results of the measurement using the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel, the uncertainty on the Higgs boson mass measurement becomes 13 MeV and 19 MeV for the H-20 and I-20 scenarios, respectively.

The combined uncertainty on the Higgs boson mass achieved in the H-20 scenario is about 6 MeV better then that of the I-20 scenario. However, the H-20 scenario envisions a factor of four longer data taking at the centre-of-mass energy of 250 GeV than the I-20 scenario. Therefore, since this centre-of-mass energy is prohibitive for many other interesting physics cases such as top quark studies, the I-20 scenario would be more favourable.

The outcome of this analysis prove that $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel can be exploited for a precise measurement of the Higgs boson mass. This channel provides a competitive alternative to the measurement using the recoil technique at the centre-of-mass energy of 250 GeV. The uncertainty on the Higgs boson mass obtained using this channel is also significantly better than the estimations for the measurements by the ATLAS and CMS experiments at the end of LHC operation which is estimated to be about 100 MeV for an integrated luminosity of approximately 300 fb^{-1} [11–13].

The particle flow reconstruction plays a central role in achieving the results presented in this chapter. Advanced detector technologies are essential for the particles flow reconstruction to perform efficiently. The particle flow algorithms require calorimeter systems with unprecedented granularity. Various technologies for such calorimeters are being actively developed for the ILC detectors. A number of these technologies are presented in the following chapter with an emphasis on the AHCAL.

Chapter 9

CALICE Analog Hadron Calorimeter Engineering Prototype

Lepton colliders such as the ILC provide a clean environment for precision physics studies. However, even though the final states of lepton collisions are more clean compared to hadron collisions, in many cases the precision of the measurement is limited by the jet energy resolution which in turn is limited by the energy resolution of the calorimeters.

The jet energy resolution achieved by traditional calorimetry is not sufficient for the physics goals of the ILC. A promising solution to improve jet energy resolution is the particle flow approach to calorimetry. The particle flow approach is introduced in chapter 5 and its application on a precise measurement of the Higgs mass at the ILC is studied in chapter 7 and chapter 8.

In order to achieve the unprecedented jet energy resolution required for the ILC physics programme, the sophisticated reconstruction techniques of the particle flow approach require highly segmented electromagnetic and hadronic calorimeters. The CALICE collaboration co-ordinates a number of research and development programmes for the realisation of such highly granular calorimeters.

In this chapter, after a discussion on the requirements for calorimeters at the ILC in section 9.1, various technologies under investigation by the CALICE collaboration are introduced in section 9.2. One of the very promising technologies for hadron calorimetry being developed by the CALICE collaboration is the Analog Hadronic Calorimeter (AHCAL). Currently, an engineering prototype of the CALICE AHCAL is under development. In section 9.3 the AHCAL engineering prototype is presented in detail.

9.1 Requirements for the ILD Calorimeters

Jet energy resolution has a direct impact on the sensitivity to many physics processes. For the detectors at the ILC the aim is to achieve a jet energy resolution which is better than $3-4\%$.

In the traditional calorimetric approach, all components of the jet are measured in the calorimeter systems. Therefore, approximately 70% of the jet energy is measured in the hadronic calorimeter which has a poor energy resolution and the remaining 30% is measured in the electromagnetic calorimeter. As discussed in chapter 5, the idea behind the particle flow approach is to employ the tracking system to measure the energy of the charged components of a jet, the ECal to measure the energy of photons and the HCal only to measure the energy of neutral hadrons. In this approach, only 10% of the jet energy is measured with the poor energy resolution of the HCal.

In order to be able to benefit from the particle flow approach the calorimeters need to be able to resolve energy deposits from different particles. Achieving this goal requires highly granular calorimeters. Sophisticated reconstruction software can then be applied on the raw calorimeter data to identify energy deposits from each individual particle.

Increasing the granularity of the calorimeters introduces new challenges for the calorimeter development. The higher granularity requires a higher number of read-out channels. Each read-out channel requires power supplies and physical links for control and acquisition of the signal from sensors. In the traditional calorimeters which have a small number of read-out channels, each read-out channel is supported by a number of dedicated power and signal lines which guide the signal out of the calorimeter before it is processed. In the highly granular calorimeters which have millions of read-out channels, the traditional approach is not feasible.

Apart from the particle flow requirements, a calorimeter for ILD should also fulfil the requirements imposed by ILD. These requirements include constraints on mechanics and power consumption and dissipation. In order to reduce the amount of non-instrumented areas in the detector, there is no active cooling system considered between the calorimeter absorbers at ILD. Therefore, the power dissipation by the front-end electronics should be kept at a minimum level. The beams at ILC are structured in bunch trains of roughly 1 ms length. Between bunch trains, there is a 199 ms long gap during which no collision takes place. In order to reduce power dissipation, it is foreseen to power-down the parts of the front-end ASICs which are not needed during the gap. This is known as *power-pulsing* operation scheme.

Based on the discussion above, the general requirements for the calorimeters at ILD are the followings:

- The calorimeter shall provide the required granularity for the particle flow approach.

- The calorimeter shall have minimum non-instrumented areas between the absorber layers.
- The calorimeter shall be capable of operation without active cooling inside absorber layers.
- The calorimeter shall be capable of operation in a magnetic filed with the strength of 4 T.
- The calorimeter shall fulfil the mechanical requirements of ILD.
- The calorimeter shall be able to operate according to the power-pulsing scheme.

9.2 The CALICE Collaboration

The CALICE collaboration is a co-ordinated group of physicists and engineers from around the world who work on the development of highly granular calorimeters for the future experiments at the electron-positron colliders [97].

The research projects of the CALICE collaboration consist of physics studies and technological development for highly granular calorimeters. The collaboration maintains a number of research and development programmes which study different technologies for electromagnetic and hadronic calorimeters. All of the calorimeters which are developed by the CALICE collaboration are sampling calorimeters (see chapter 5). Various absorber materials and active media are considered for these calorimeters.

The CALICE collaboration follows a multi-phase development strategy. In the first phase, the feasibility of constructing a highly granular calorimeter based on a certain technology and its performance is studied by constructing a *physics prototype*. After a concept is proven using the physics prototype, a *technological prototype* is constructed in order to study and overcome technological challenges of realisation of a full-scale calorimeter as a subdetector of a general purpose detector such as the ILD.

The focus of this chapter is on the CALICE AHCAL. Before introducing the AHCAL, other calorimeter technologies under development by the CALICE collaboration are briefly introduced here.

9.2.1 Silicon-tungsten Electromagnetic Calorimeter

The Silicon-tungsten Electromagnetic Calorimeter (Si-ECAL) is a highly granular sampling electromagnetic calorimeter being developed by the CALICE collaboration. The absorbers of the Si-ECAL are made of tungsten. Tungsten has a small Molière radius of ~ 9 mm and short radiation length of ~ 3.5 mm, therefore, electro-magnetic showers spread over a

smaller volume in tungsten which reduces the probability of overlapping showers from nearby particles.

The Si-ECAL employs silicon detectors as active material. The silicon technology is widely used in tracking and vertex detectors. Segmenting silicon detectors into small active regions can be achieved without great difficulty, qualifying it as a suitable technology for a highly granular calorimeter. Silicon detectors are thin and including them within the tungsten absorbers does not increase the Molière radius of the calorimeter greatly. However, it is an expensive technology and the cost of employing them in large-scale experiments can be prohibitive.

During the first development phase of the Si-ECAL, a physics prototype was built based on the aforementioned concept [98]. The Si-ECAL physics prototype consists of thirty calorimeter layers which total to 24 radiation length and features 6480 read-out channels. Each layer has an active area of 18×18 cm^2 which is segmented into pads of 1×1 cm^2.

The physics prototype of the CALICE Si-ECAL has been tested with electron and hadron beams at the DESY and CERN test beam facilities in 2006 and 2007. The test beam results show that the detector response is linear to within approximately 1% and a relative energy resolution of $(16.6 \pm 0.1)/\sqrt{E(GeV)} \oplus 1.1 \pm 0.1(\%)$ was achieved [84]. The test beams proved that the Si-ECAL technology can fulfil the requirements to achieve the physics goals of the ILC.

After the concept of the highly granular electromagnetic calorimeter using the silicon technology was proved by the Si-ECAL physics prototype, construction of the technological prototype of the Si-ECAL has started. The technological prototype has smaller pad size and the read-out electronics is integrated in the active layers. The space constraints are realistic for a subdetector of a general-purpose detector such as ILD.

9.2.2 Scintillator Strip Electromagnetic Calorimeter

The Scintillator Strip Electromagnetic Calorimeter (Sc-ECAL) is an electromagnetic calorimeter which exploits scintillator as active material. The concept of the Sc-ECAL is similar to the concept of the AHCAL which is discussed in details in section 9.3. The Sc-ECAL employs tungsten absorbers and scintillator material as active medium and offers a higher granularity than the AHCAL. The scintillators are read out by Silicon Photomultipliers (see section 9.3.5).

The first physics prototype of the Sc-ECAL was built, featuring 26 layers of tungsten-cobalt absorbers and plastic scintillator strips as active material. The prototype had a total radiation length of $18.5X_0$. The scintillator strips are 45×10 mm^3. The scintillator layers where placed alternating in horizontal and vertical orientation to reach an effective granularity of 1 cm^2. The prototype of the Sc-ECAL has been successfully tested at various test beam facilities. In the energy range between 1 to 6 GeV the energy measurement was

linear to within 1%. The energy resolution of the prototype was measured in various configurations and at different positions. The stochastic term of the relative energy resolution (see section 5.1.3) was between 13 and $14\%/\sqrt{E(GeV)}$ and the constant term between 3 and 4.5% [99].

The physics prototype proved that the Sc-ECAL technology fulfils the energy resolution requirements of the ILD. A technological prototype of the Sc-ECAL which is scalable to a full calorimeter and has integrated electronics is under development.

9.2.3 Digital Hadronic Calorimeter

The CALICE Digital Hadronic Calorimeter (DHCAL) is a highly segmented hadronic calorimeter based on the Resistive Plate Chamber (RPC) gaseous technology. The DHCAL has a simple 1-bit read-out system, hence the name digital. This read-out concept is based on the assumption that to first order the energy of the incident particle is proportional to number of pads with a signal over a given threshold.

As a proof of concept, a large-scale prototype of the DHCAL was constructed and tested at various test beam facilities [100]. The DHCAL prototype consists of 38 calorimeter layers with 1.74 cm thick steel absorbers as the main stack. The main stack is followed by a tail catcher and muon tracker comprised of 8 layers with 2.0 cm steel absorbers and 6 layers with 10 cm absorber plates. Each layer has an active area of 96×96 cm^2 segmented into 9216 read-out channels of 1×1 cm^2. The prototype features approximately $500,000$ read-out channels in total. This prototype has bees tested extensively at the CERN and Fermilab test beam facilities. A preliminary analysis of the tests with pions and positrons demonstrate that a relative hadronic energy resolution of $\frac{\sigma_E}{E} = \frac{55\%}{\sqrt{E}} \oplus 7.5\%$ and an electromagnetic energy resolution of $24\% - 14\%$ in the energy range of $2 - 25$ GeV is achieved [101].

9.2.4 Semi-digital Hadronic Calorimeter

Similar to the DHCAL, the CALICE Semi-digital Hadronic Calorimeter (SDHCAL) also employs gaseous sensors. The SDHCAL gaseous sensors are Resistive Plate Chambers (RPCs). Instead of the binary hit/no-hit read-out scheme of the DHCAL, the SDHCAL has a coarse read-out system based on 3 thresholds, hence the name Semi-digital.

A large prototype of the SDHCAL is built which contains 48 layers and measures 1.3 m^3 [102]. Each layer of the SDHCAL is made of 1 m^2 Glass Resistive Plate Chamber (GRPC) detector. The active layers are interleaved with 1.5 cm thick stainless steel absorbers. The required segmentation of the SDHCAL is achieved by using 1 cm^2 electronic read-out pads. The SDHCAL prototype features roughly $440,000$ read-out channels.

The SDHCAL prototype has been tested at the CERN SPS test beam facility in the energy range of $5 - 80$ GeV [103]. The first results show a linear detector response within $4 - 5\%$. An energy resolution of 7.7% was obtained at 80 GeV.

As an alternative to the RPCs, the performance of Micro-MEsh GAseous Structures (MICROMEGAS) as active sensors for the SDHCAL is also being studied.

9.3 The Analog Hadron Calorimeter

Particle detection using scintillation light is one of the well-established and most widely used techniques in nuclear and particle physics. This technique benefits from a phenomenon in which certain materials emit flashes of light when charged particles traverse them, known as scintillation process. By optically coupling scintillator material to a photodetector, the scintillation light can be converted to an electrical signal suitable for further analysis.

The CALICE Analog Hadronic Calorimeter (AHCAL) is a highly-granular calorimeter which employs the scintillation process in its read-out mechanism. Layers of absorber material are interleaved with layers of scintillator tiles which are used to sample the shower which develops in the absorber layers. Each scintillator tile is optically coupled to a dedicated photodetector which converts the scintillation photons to an electrical signal.

In order to evaluate the feasibility and performance of this concept for construction of a highly-granular calorimeter, a physics prototype was constructed and studied in multiple beam test campaigns. In this section, after a short introduction of the AHCAL physics prototype, the next generation of the CALICE AHCAL known as the technological prototype is described in detail.

9.3.1 The AHCAL Physics Prototype

The physics prototype of the CALICE AHCAL was constructed with two main goals. The first goal was to study the performance of a highly granular calorimeter using the scintillator technology. The large sample of hadronic showers collected using the AHCAL physics prototype was also used to study hadronic shower shapes and to examine the available simulation models. The second goal of the AHCAL physics prototype was to investigate the usage of novel Silicon Photomultipliers (SiPMs) as read-out sensors in a large-scale and to identify their critical operational issues. Development of quality control procedures and a reliable calibration system was also pursued.

The physics prototype of the CALICE AHCAL calorimeter employs plastic scintillators as active material. The scintillators are segmented into small tiles to reach the required granularity. Each scintillator tile is equipped with a Silicon Photomultiplier (SiPM) photo detectors to read out the scintillators. [104]

The AHCAL physics prototype is a $1 \times 1 \times 1$ m^3 detector consisting of 38 calorimeter layers. The 1.74 cm thick steel absorber plates are interleaved with 5 mm thick plastic scintillator tiles. The tiles measure 3×3 cm^2 in the core and 6×6 cm^2 or 12×12 cm^2 in the outer regions. In total there are 7608 scintillator tiles. The 5.3 neutron nuclear interaction length of the physics prototype is comparable to that of the ILD hadronic calorimeter.

The AHCAL physics prototype underwent the first test at the CERN test beam facilities in 2006 and 2007 where energy deposits of muons, electrons and pions where measured in a combined setup with the physics prototype of the CALICE Si-ECAL in front and a tail-catcher and muon tracker behind.

The test at CERN and further studies using the Fermilab test beam facilities demonstrated that the AHCAL physics prototype achieves the expected performance [83]. An intrinsic energy resolution of $\frac{\sigma_E}{E} = \frac{58\%}{\sqrt{E}} \oplus 1.6\%$ for hadrons was achieved. The linearity of the calorimeter response to hadrons predominantly contained in the calorimeter was found to be within 2% over the energy range of the study. The operational stability and the hadron shower physics results obtained using the AHCAL physics prototype proved the overall concept. Construction of the next generation prototype to overcome the technological challenges of a full-scale hadron calorimeter based on the AHCAL concept started in 2010. This prototype is introduced in details in the following section.

9.3.2 Motivations for the Engineering Prototype

After the general concept of a highly granular hadronic calorimeter using scintillator and SiPM technologies was successfully validated by the AHCAL physics prototype, construction of the next generation prototype started. The aim of the new prototype, which is referred to as the AHCAL engineering prototype, is to investigate and overcome technological challenges of building a full hadronic calorimeter for a general-purpose detector such as ILD using the AHCAL technology.

A full hadron calorimeter for the ILD will have approximately 4 million read-out channels in the barrel and a similar number of read-out channel in the end-caps. Mass production and assembly of hardware components of this large number of channels is a challenge which needs to be addressed. Production quality should be controlled and time constraints for production and assembly should be considered. Providing power and signal lines to the AHCAL channels in a manageable manner with introducing minimum non-instrumented material is another challenge. Data concentration should take place as early as possible in order to reduce the amount of service cables. A solution to this issues is to integrate the read-out electronics in the absorber gaps next to the active media.

Including all read-out electronics inside the absorber gaps introduces yet another challenge. In order to minimise the amount of dead material, active cooling is not available between the absorber gaps and therefore the heat produced by the front-end electronics can not be dissipated readily. Hence, the amount of power dissipated by the read-out electronics

should be kept as low as possible. Furthermore, the calorimeter should operate according to the power-pulsing scheme in order to further reduce heat dissipation.

Operation of such a calorimeter with a large number of read-out channels requires a dedicated data acquisition system which is designed specifically for this system. Although the requirements for data acquisition system for beam tests are rather different from the ILC requirements (see chapter 10), development of a data acquisition system which can later be scaled to a large calorimeter at a collider experiment needs to be pursued in the engineering prototype.

The engineering prototype of the CALICE AHCAL is aimed at addressing the aforementioned challenges. Moreover, the complete prototype will be used to study different aspects of hadron shower physics using test beam data.

9.3.3 Absorber Material

Materials with high atomic number and short nuclear interaction length λ_I , can contain a hadronic shower in a smaller volume as compared to a material with larger λ_I . Thus, when materials with shorter λ_I are used as absorbers, the detector volume can be kept small. However, a reasonable sampling of the electromagnetic component of the hadronic showers is also beneficial. Therefore, materials for which the ratio of the nuclear interaction length to the radiation length X_0 is large should be avoided. In addition, the mechanical properties of the absorber material should be considered. The absorber material should be suitable for building a rigid structure at a reasonable cost.

Considering these aspects, stainless steel which has $\lambda_I = 16.8$ cm, $X_0 = 1.8$ cm and $\lambda_I/X_0 = 9.3$ has been chosen as the material for the AHCAL absorbers at the ILD [34].

The possibility of using tungsten as the absorber material as an option for higher energy machines such as CLIC has been investigated using the AHCAL physics prototype. Tungsten has a much shorter λ_I and therefore, if used as absorber material, can fully contain the high-energy hadronic showers in a smaller volume, reducing the size and possibly the cost of the detector [105, 106]. However, new challenges arise due to very small electromagnetic subshowers in tungsten absorbers.

9.3.4 Active Material

Passage of particles through certain material can excite their atoms and molecules. These excited states are unstable and decay quickly to the ground state by means of photon emission. This process, which is known as scintillation, is widely used as a particle detection technique in nuclear and particle physics.

Various scintillating materials exist. However, not all scintillators are suitable for particle detection. A scintillator needs to satisfy the following general requirements in order to be qualified for particle detection:

- it shall have a high efficiency of converting the energy lost by the traversing particle to photons,
- it shall be transparent to its own scintillated light so that the light can be transmitted outside of the material,
- it shall have a short decay time to allow for fast response,
- the emitted scintillation light shall be in a spectral range which is consistent with the spectral response of the photon detectors.

Apart from the general requirements mentioned above, cost of the scintillator is another important factor for large instruments such as a hadron calorimeter for the ILD.

The AHCAL has opted for plastic scintillators as active material. Plastic scintillators are solutions of aromatic hydrocarbon compounds doped with low concentration of fluors (organic scintillators) in a solid plastic solvent. Plastic scintillators offer a very short decay time in the order of ~ 2 ns. Flexibility of the plastic scintillators is another major advantage of this material since they can be easily machined or moulded to the desired forms. Plastic scintillators are commercially available at low cost and therefore construction of a large detector by employing this material is affordable.

The scintillator tiles of the AHCAL engineering prototype measure $3 \times 3 \times 0.3$ cm^3. Previous studies have shown that a full hadronic calorimeter with such a granularity fulfils the requirements of the particle flow algorithms.

Considering the fact that a full hadronic calorimeter will consist of millions of scintillator tiles, it is necessary to have a homogeneous response from all channels in order to be able to simplify detector commissioning and calibration procedures. The scintillator tiles of the AHCAL physics prototype were equipped with wavelength-shifting fibres since the available photodetectors were not sensitive at the wavelength of the scintillation light. By employing new photodetectors which can directly detect the light from scintillator tiles, there is no need for the wavelength-shifting fibres in the AHCAL technological prototype [107]. A new tile design offers a 10% spread in the response of a large number of channels [108]. In this design, the tiles are machined from a large scintillator plastic plate to the required dimensions. A cut-out on one of the edges of the scintillator tile where maximum light collection is possible hosts a photodetector. To minimise the optical cross-talk between tiles, each tile is wrapped in a reflecting plastic foil.

In the previous designs, the photodetector was attached to the side of a scintillator tile. In order to simplify mass assembly, in a new design the photodectors are mounted on the surface of the electronics carrier board together with other electronics components and

the scintillator tiles with concave cavity are placed on top [109, 110]. In this design, the assembly procedure can be further simplified by producing the scintillator in a "mega-tile" format which covers multiple channels.

9.3.5 Photo Detection

Before the scintillation light can be used to gain information concerning the incident particles, it needs to be converted to a measurable electrical signal. This conversion is performed by means of a photo detector.

The traditional photo detectors used for scintillation light read-out are vacuum tubes which are know as Photomultiplier Tubes (PMTs). The PMTs are low-noise high-gain devices with very fast response time and because of their very high gain can be employed for single photon detection. However, the PMTs are relatively large and costly devices and require very high bias voltages in the orders of a few kilovolts. Moreover, traditional PMTs are sensitive to the magnetic field, therefore, they are not suitable for applications such as the ILD hadron calorimeter where the calorimeter is located inside a high-field magnet, unless they are properly shielded. These disadvantages, disqualify the PMTs as photo detectors for the AHCAL.

Over the last decades, various semi-conductor devices have been developed as an alternative to the PMTs. Avalance Photo-Diodes (APDs) are silicon-based devices which exploit the photo-electric effect to convert light to electrical current. The APDs are designed in such a way that electron-hole carriers are accelerated in a high electric field so that they are energetic enough to create more free electron-hole pairs and increase the current, which is known as avalanche phenomenon. The APDs benefit from the avalanche phenomenon to provide a built-in gain mechanism of around 100.

A gain of 100 is not sufficient for low-light applications. For the low-light applications where a gain of 10^5 to 10^6 is required, certain types of the APDs are operated with a reverse voltage above their minimum reverse voltage that makes the diode conduct, known as the *breakdown voltage*. This high-gain operation mode is referred to as Geiger mode and the APDs which are designed specifically for this mode are known as Geiger-mode Avalance Photo-Diode (GAPD). The GAPDs are capable of single-photon detection. When a photon impinges a GAPD a current pulse is generated which signals the arrival of the photon. The output signal is identical regardless of the number of incident photons. Therefore, a GAPD is a binary sensor and can not be used to detect multiple photons simultaneously. Moreover, the size of a GAPD is limited since the characteristic dark count rate of the device and its recovery time increase with the active area.

In order to overcome the limitations of the GAPDs mentioned above, an array of GAPDs can be connected in parallel to create a new large-area device which is capable of detecting multiple photons. Each element of the array is called a *pixel* and these devices

are named Silicon Photomultipliers (SiPMs) and have become known as an alternative to the traditional PMTs.

The performance of SiPMs is characterised by the parameters listed below:

Photon detection efficiency is a measure of the number of photons detected by a SiPM to the total number of incident photons.

Gain is defined as mean number of charge carriers generated in an avalanche initiated by a single photon.

Dynamic range is a measure of the ability of the device to detect multiple photons before getting saturated. This quantity is related to the total number of pixels of a device.

Recovery time is the time required to recharge a SiPM cell after a breakdown.

Breakdown voltage is the minimum reverse voltage at which the diodes of the SiPM conduct.

Dark count rate is the rate of pulses which are generated due to phenomena other than an incident photon.

Crosstalk is a signal which is generated due to secondary photons coming from avalanche development in a neighbouring cell.

Afterpulsing is a secondary avalanche event initiated by the carriers trapped from a previous event and released with a delay.

There are a few semi-conductor manufacturing companies which offer various types of SiPMs for different applications. The aforementioned parameters has been evaluated for a number of SiPMs to find the optimal candidate for the AHCAL application.

The SiPMs have a number of advantages compared to the PMTs. The bias voltages required for the operation of the SiPMs is a few tens of volts which is much lower than the bias voltage required for the PMTs. The dimensions of a SiPM are very small in the order of a few millimetres and they are available at low cost. Another important aspect for the AHCAL application is that the SiPMs are not sensitive to the magnetic field.

9.3.6 Read-out Chip

The very front-end electronics required for reading-out the SiPMs should be very compact (see section 9.3.7). This goal is achievable using a dedicated Application-specific Integrated Circuit (ASIC) for SiPM read-out and signal processing. The main requirements for an ASIC for the AHCAL are the followings:

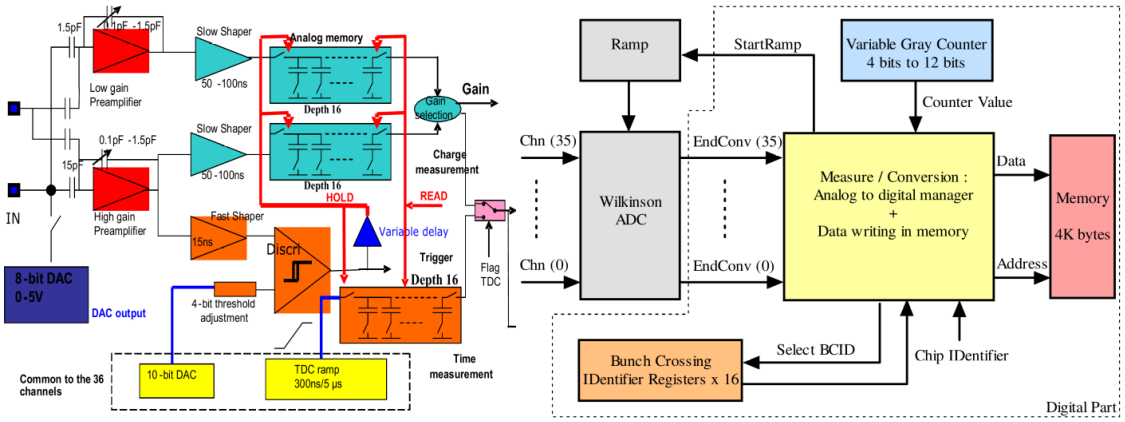


Figure 9.1 – Schematic diagram of one channel of the SPIROC ASIC [111].

- The ASIC shall read-out the analog output signal of the SiPMs.
- The ASIC shall process the analog signal of the SiPMs.
- The ASIC shall digitise the analog signal of the SiPMs.
- The ASIC shall provide timing information for the SiPM signals.
- The ASIC shall provide high-voltage adjustment for each read-out channel individually.
- The ASIC shall provide gain adjustment for each read-out channel individually.
- The ASIC shall be capable of internally storing all processed signals during one ILC bunch train.
- The ASIC shall send out the stored signals during the gap between bunch trains.
- The ASIC shall multiplex signal and control lines.
- The ASIC shall consume power no more than 25 mW per channel in order to avoid the need for active cooling.
- The ASIC shall provide a dynamic range of 1 to 2000 photoelectrons.
- The ASIC shall provide auto-triggering functionalities with an adjustable threshold.
- The ASIC shall be compact enough to fit in the gap between the absorber plates alongside the scintillator tiles.

The timing information can be used to improve energy measurements. Auto-trigger functionality is required in order to reduce the noise and thus lower the data volume stored locally before transmission.

The Silicon Photomultiplier Read-out Chip (SPIROC) has been developed by the OMEGA group [112] to fulfil the requirements of the AHCAL [113]. The SPIROC has

36 input channels and is capable of charge and time measurement on each channel. An internal array of analog memories with the depths of 16 samples is available to store charge and time measurements. The content of the analog memories are digitised using a 12-bit Wilkinson Analog-to-digital Converter (ADC) before they are stored in a SRAM.

The SPIROC consists of an analog and a digital part. The schematic diagram of the analog part of the SPIROC is shown in figure 9.1. A Digital-to-Analog Converter (DAC) with 8 bits resolution is used to adjust the high-voltage provided to the SiPMs within the range of 5 V. This allows for tuning the gain of the SiPMs channel by channel. Two different low-noise variable pre-amplifiers are available which enable the SPIROC to operate in low-gain or high-gain modes in order to provide the required dynamic range of 1 to 2000 photoelectrons.

The auto-trigger functionality of the SPIROC is provided by a fast shaper which feeds a discriminator. The threshold of the discriminator is adjusted by means of a DAC with 10 bits resolution. The threshold is common among all 36 input channels but it can be fine-tuned channel by channel using 4 additional bits. When a trigger occurs, the output of the two slow shapers which follow the pre-amplifiers are stored in two Switched Capacitor Arrays (SCAs) with the depth of 16 samples. An external trigger-validation signal can be used to reduce the noise from SiPMs. Time measurement is provided by means of a 300 ns long voltage ramp which is common to all 36 input channels. The voltage ramp can be changed to 4 μs in order to fulfil beam test requirements. The analog time information is stored in another SCA when a trigger occurs.

A Wilkinson ADC with a resolution of 12 bits is used to digitise the content of the SCAs before they are stored in a SRAM to be read-out of the ASIC. The SPIROC has a rather complex digital part which controls various features of the ASIC and communicates with the DAQ system.

The top-level finite state machine of the SPIROC has four main states:

Acquisition State in the acquisition state valid data are stored in the analog memories. An external validation signal can be provided to flag and remove noise events. The coarse time measurement(BXID) is stored in the digital memory and the fine time measurement(TDC) is stored in the analog memory.

Conversion State during this state the contents of the analog memories are converted to digital data before they are stored in the SRAM. In addition, the Bunch Crossing Identifier (BXID) and two additional bits to flag the gain mode of the ASIC (gain-bit) and whether the signal amplitude is above threshold (hit-bit) are stored in the SRAM.

Read-out State during this state the data stored in the SRAM are sent to the data acquisition system through a serial link. A single serial link can be used to read out up to 256 ASICs.

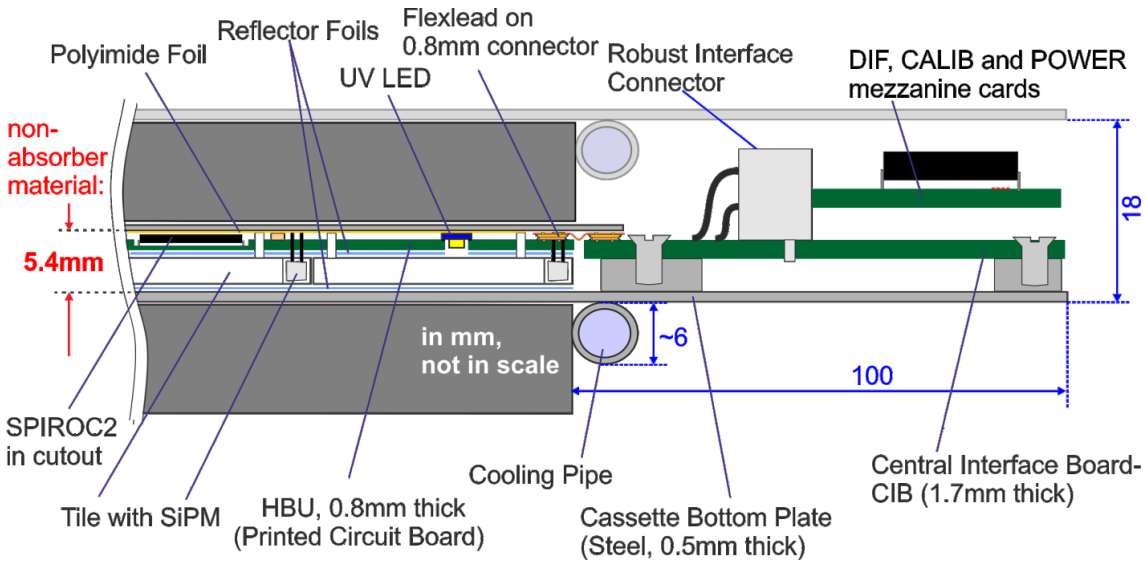


Figure 9.2 – Schematic diagram of the gap between the AHCAL absorbers [114].

Idle State when all aforementioned operations are completed the ASIC goes to the Idle state to save power.

There have been several iterations of the SPIROC ASIC to fix bugs and improve various functionalities. SPIROC2b has been used for a large prototype which has been intensively tested at DESY and CERN test beam facilities. Currently, a new iteration of the SPIROC code named SPIROC2e is being used for constructing a cubic-meter prototype of the AHCAL.

9.3.7 AHCAL Electro-mechanics

The physics programmes of the ILC defines the requirements for the detectors. In order to fulfil the requirements of the particle flow approach to achieve the necessary jet energy resolution for the physics programme, a calorimeter which is highly granular and dense with minimum dead material is necessary. The number of electronic read-out channels grows with the granularity of the calorimeter. Therefore, to avoid a large amount of dead material required for power and signal transfer, the read-out electronics should be located in a close vicinity of the sensors.

In order to fulfil the electro-mechanical requirements, the AHCAL read-out electronics should be mounted inside the gaps between the absorber plates alongside the scintillator tiles. A cross section of a gap between two absorber plates of the AHCAL is shown in figure 9.2. The gap between the absorber plates is 5.4 mm wide while 3 mm of it is filled with the scintillator tiles. Therefore, the thickness of the front-end electronics should not exceed 2.4 mm. To fit the required electronics for service and data transfer in such a narrow space is challenging.

As explained in section 9.3.6, most of the front-end electronics needed for biasing the SiPMs, acquisition control and signal processing is concentrated in a compact ASIC. A standard size Printed Circuit Board (PCB) is employed as the host of the ASICs, the SiPMs and other necessary electronic components as well as power and signal distribution lines. The scintillator tiles are assembled on the back side of the PCB. This standard size PCB which measures 36×36 cm is known as HCal Base Unit (HBU).

A picture of the HBU is shown in figure 9.3. Each HBU hosts four SPIROC ASICs and 144 read-out channels. Scintillator tiles and the SiPMs are assembled on the back side of the HBU.

Several HBUs can be connected together using flexible low-height connectors to form a full layer of the ILD hadron calorimeter. A full layer of the ILD consists of 18 HBUs arranged in 3 slabs of 6 HBUs. For the barrel AHCAL, necessary electronics and interfaces for a layer are placed at the two end-faces of the barrel where they are easily accessible for service lines and maintenance and active cooling is provided, as shown in figure 9.4.

Each layer of the AHCAL is equipped with a Central Interface Board (CIB) which is placed outside of the gap towards the endcaps. The CIB hosts three pluggable modules. The power module manages power regulation and distribution of the entire layer. The calibration module controls the LED calibration system which is introduced in section 9.3.8. The third module provides the interface between the layer and the data acquisition system as explained in section 10.3.1.

9.3.8 LED Calibration System

The gain of the SiPMs varies with changes in temperature and bias voltage, as explained in section 9.3.5. Therefore, a gain calibration and monitoring system is required particularly at the ILD since due to the lack of active cooling in the gaps between the absorber layer the temperature of the SiPMs will change over time.

The SiPMs feature a characteristic Single Pixel Spectrum (SPS) when exposed to low-intensity light conditions. The SPS consists of multiple peaks which correspond to the number of pixels fired on the SiPM. The distance between the peaks is a measure of the gain of the SiPM. This characteristic feature can be exploited to devise a calibration system. This calibration technique has the advantage that the calibration signal does not need to be constant over time and it is not necessary to know its size.

For the engineering prototype of the AHCAL, the HBU is equipped with a ultra-violet LED for each scintillator tile. The Detector Interface module (see section 10.3.1) provides the required driving pulse for the LED driver circuits on the HBU. The amplitude of the LED pulse can be adjusted via the data acquisition system.

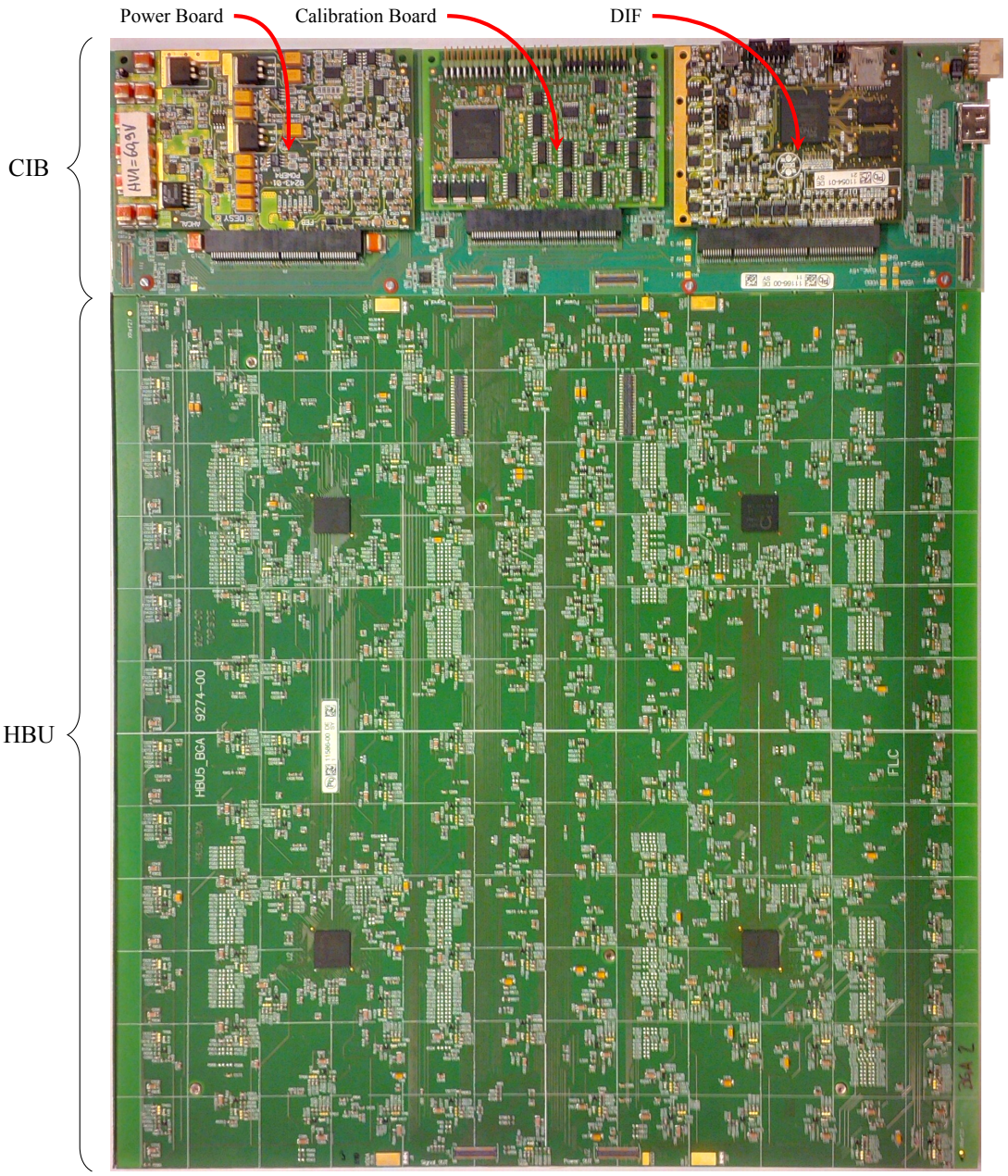


Figure 9.3 – A HCal Base Unit (HBU) with Central Interface Board (CIB) and power, calibration and Detector Interface (DIF) boards. Each of the white squares on the HBU depicts the area of one scintillator tile which is mounted on the rear side of the HBU. The four black squares are the read-out ASICs.

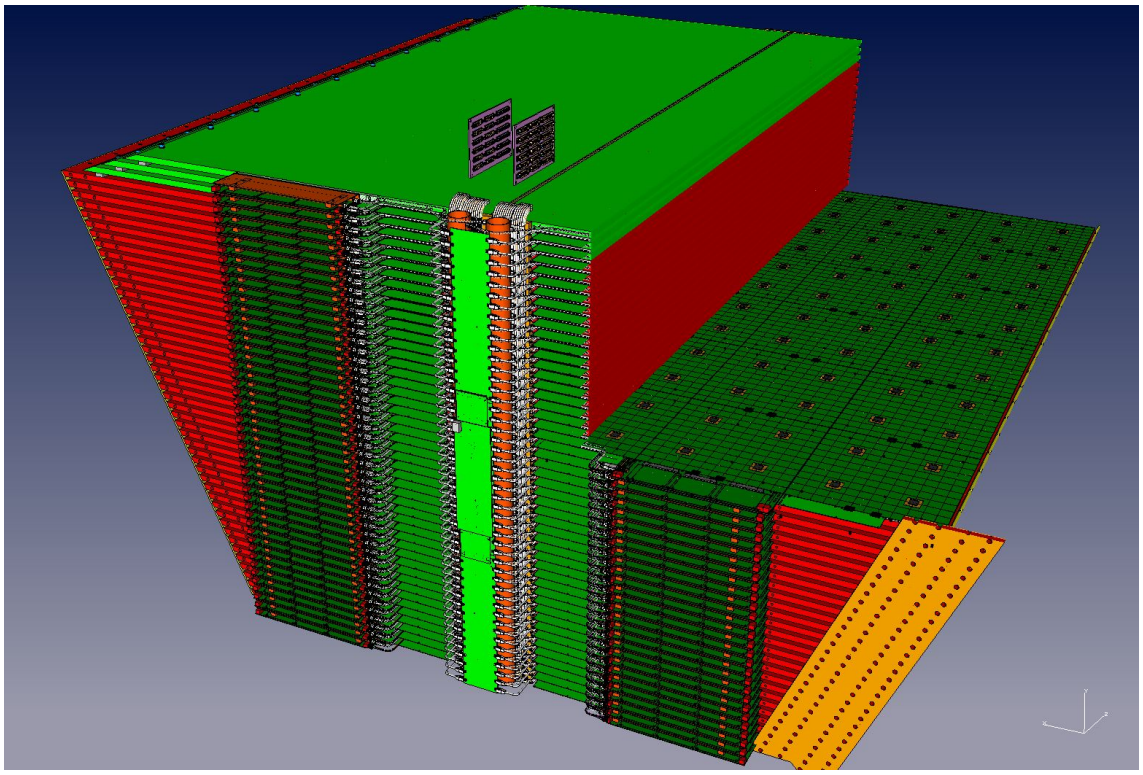


Figure 9.4 – Engineering CAD design of a section of the AHCAL for the ILD barrel [115].

Chapter 10

Data Acquisition System for the CALICE Scintillator Calorimeters

The CALICE Analog Hadronic Calorimeter (AHCAL) is a highly granular sampling calorimeter designed to fulfil the requirements for application of the particle flow reconstruction at ILC. The AHCAL uses plastic scintillator tiles as active material which are read out by novel Silicon Photomultipliers. A dedicated front-end read-out ASIC called SPIROC is used to read-out the Silicon Photomultipliers. The CALICE Scintillator Strip Electromagnetic Calorimeter (Sc-ECAL) is based on the same concept with very similar electronics. The details of the engineering prototype of the CALICE AHCAL and the Sc-ECAL are presented in chapter 9.

One of the challenges in realisation of such a highly granular calorimeter consisting of millions of read-out channels is to design and implement a Data Acquisition System (DAQ) which meets the needs of such a detector. The requirements for the calorimeter DAQ system at ILC are rather different from those of the LHC experiments. Therefore, a dedicated development is necessary. In contrast to the LHC experiments, the ILC calorimeters have a large number of read-out channels, but because of low occupancy and low event rate the data rate is relatively low. As a result, auto-trigger functionality and early zero suppression and data aggregation is necessary. Moreover, since the highly-granular calorimeters are in the development phase, the DAQ system should accommodate the needs for laboratory and beam tests.

In this chapter, after a discussion on the requirements of the DAQ system for the CALICE scintillator calorimeters, the architecture and implementation of the DAQ system for the engineering prototypes of the AHCAL and the Sc-ECAL are presented. The working principle of the AHCAL and Sc-ECAL is the same and both of the detectors utilise the same read-out ASIC as well as front-end electronics. Therefore, a DAQ system developed for the AHCAL can be employed to operate the Sc-ECAL without any modification. The same DAQ can also be used for operation of the two detectors in a combined setup.

The AHCAL has multiple running modes while operating as a stand-alone detector. Furthermore, it should be able to run with other types of detectors in a combined detector setup. Each of these modes has its own requirements which are presented in section 10.1. The overall architecture of the data acquisition system for the engineering prototypes of the CALICE scintillator calorimeters is introduced in section 10.2. The functionality, implementation and operation principles of the various scintillator DAQ hardware components are explained in section 10.3. The various software components of the DAQ system are presented in section 10.4. The AHCAL and the Sc-ECAL should be able to operate in a combined setup along with other types of detectors. The first steps in integrating the scintillator DAQ system into a higher level common data acquisition system are discussed in section 10.5.

10.1 Requirements for the Data Acquisition System

A calorimeter based on the CALICE scintillator technology for a general-purpose detector such as the ILD will have millions of read-out channels. The AHCAL and the Sc-ECAL should work based on the power-pulsing principle and should fulfil ILD requirements such as geometry and power consumption/dissipation. The calibration system of the AHCAL operates in an external-trigger mode which has different requirements than the data taking mode of the ILD. Moreover, the technological prototypes of the scintillator calorimeters which are in the development phase are regularly tested with particle beams. The operation of the detector at the beam tests imposes another set of requirements. Eventually, the AHCAL and Sc-ECAL will be integrated into the ILD detector and need to be operated along with other subdetectors of ILD. Therefore, their DAQ system should be able to operate in a slave mode as part of a higher level common data acquisition system for ILD. A common operation with other detectors during beam tests is also necessary.

In this section, the top level requirements for a DAQ system that addresses the aforementioned needs are presented. A more detailed analysis of these requirements and their impact on each of the DAQ subsystems are discussed in the following sections.

10.1.1 General Requirements

There are a number of requirements that are common among the various calorimeter operation conditions which are mentioned in the introduction of this sections. These requirements concern the basic functionality of the calorimeter and are as following:

1. The DAQ system shall physically interface with the power control system of the calorimeter modules.

2. The DAQ system shall provide power switching functionalities for the calorimeter.

3. The DAQ system shall physically interface with the very-front-end read-out ASICs.
4. The DAQ system shall be capable of configuring the very-front-end read-out ASICs.
5. The DAQ system shall be capable of configuring the various DAQ subsystems.
6. The DAQ system shall provide operating modes of the calorimeter:
   (a) external-trigger mode
   (b) auto-trigger mode
7. The DAQ system shall provide global clocks to the calorimeter subsystems.
8. The DAQ system shall ensure synchronous operation of the entire calorimeter.
9. The DAQ system shall acquire data from the calorimeter modules with 100% efficiency.
10. The DAQ system shall ensure the integrity of the acquired data.
11. The DAQ system shall store the acquired data in proper formats for instrument performance evaluations and data analysis.
12. The DAQ system shall provide/interface to online monitoring facilities for data quality assurance.
13. The DAQ system shall ensure stable operation of the detector over an extended period of time.
14. The DAQ system shall provide debugging functionalities.

In addition to the essential requirements mentioned above, the followings are also desired:

1. The DAQ system shall provide slow-control monitors for the detector layers and subsystems:
   (a) temperature monitors
   (b) voltage monitors
   (c) current monitors
2. The DAQ system shall store the slow-control data in proper formats for future reference.
3. The DAQ system shall be capable of automatic restart of calibration or data-taking runs.
4. The DAQ system shall be capable of performing parameter scans.

The slow-control monitoring is usually performed by a dedicated system. However, at the current stage of the scintillator calorimeter development it is considered to be part of the DAQ system. The developments regarding the slow-control monitoring can later be transferred to a dedicated system, should it be necessary.

10.1.2 Requirements for LED Calibration System

The CALICE scintillator calorimeters have a LED calibration system which is explained in section 9.3.8. This system is also used for laboratory tests without particle beams during the development phase. The LED calibration system needs special parameter settings and configuration. During the operation of the detector in the calibration mode, all memory cells of all detector channels are full. Therefore, a large data volume is expected. The following are the exclusive DAQ system requirements for the LED calibration system:

1. The DAQ system shall physically interface with the LED calibration board.
2. The DAQ system shall be capable of setting LED voltages per layer.
3. The DAQ system shall operate at maximum speed achievable in order to minimise the time needed for calibration.

10.1.3 Requirements for Beam Tests

During the research and development phase of the scintillator calorimeters, the detector prototypes are examined with particle beams both for evaluation of the detector performance and as a fixed target experiment for physics studies. These experiments take place at various test beam facilities with different conditions such as beam structure and beam rate. The DAQ system should be capable of operating the detector under these test beam specific conditions. The requirements for beam test operations are listed below:

1. The DAQ system shall physically interface with beam instrumentation.
2. The DAQ system shall perform measurements only within the time window when particle beam is available (within accelerator spill).
3. The DAQ system shall distribute the validation signal to the detector layers.
4. The DAQ system shall run fast enough to use the beam time as efficient as possible.

10.1.4 Requirements for ILC

Operating a full calorimeter as a sub-detector of a large general-purpose detector such as ILD sets additional requirements on the DAQ system. There is no active cooling between the calorimeter absorbers at ILD, therefore, power dissipation of the front-end electronics should be kept at a very low level which is defined by the detector requirements. The ILC detectors shall follow the power-pulsing operation scheme (see section 9.1) in order to reduce the power dissipation. The beam tests can also benefit from power-pulsing to avoid active cooling. The DAQ system should operate the detector according to the power-pulsing scheme.

At ILD data from each bunch-crossing is recorded and no trigger system is present. In addition, a full hadronic calorimeter based on the AHCAL concept will have approximately 8 million channels. However, due to low occupancy and low event rate, the data rate is relatively low. Therefore, the DAQ system should be able to cope with these conditions.

When integrated into a general-purpose detector, the scintillator calorimeters should operate along with other sub-detectors. Therefore, the DAQ system for the scintillator calorimeters should be able to operate in slave mode, while the top-level DAQ system of ILD is the master of operation. Although the requirements of the top-level DAQ system are not defined as of yet, a number of generic features required for integration into a higher level common DAQ system can already be considered. These requirements are also valid for combined operation of the prototypes in beam tests.

- The DAQ system shall operate at a power level required by the ILD.
- The DAQ system shall interface with a higher level DAQ system.
- The DAQ system shall be capable of operating the calorimeter according to the power-pulsing operation scheme.
- The DAQ system shall operate according to the ILC bunch and bunch-train structure.
- The final version of the DAQ system shall be scalable to a full detector with up to 8 million read-out channels.

10.2 The CALICE Scintillator DAQ System Architecture

During the development of the first generation prototypes, the CALICE collaboration initiated an effort to design a data acquisition system for calorimeters at the ILC [116, 117]. This effort resulted in a conceptual design which aimed at development of a system based on industrial standards.

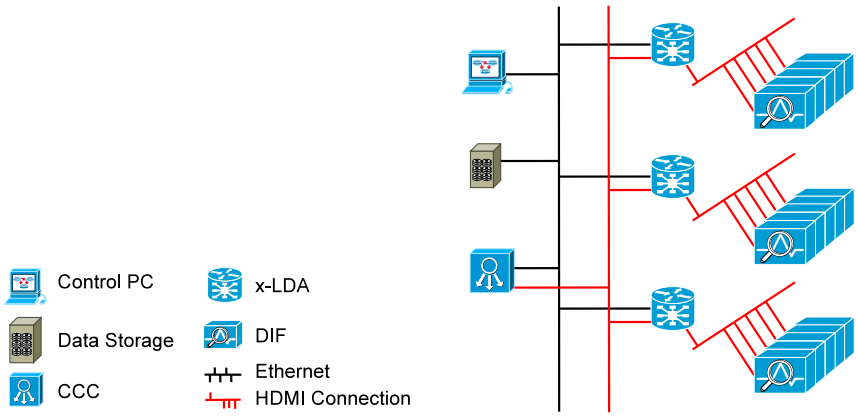


Figure 10.1 – A general overview of the data acquisition system for the CALICE scintillator calorimeters.

The architecture of the data acquisition system for the CALICE scintillator calorimeters is based on the original CALICE DAQ sytem conceptual design. A number of modifications has been made to the original design in order to accommodate further requirements or to benefit from new technological advancements. A general overview of the DAQ system for the CALICE scintillator calorimeters is presented in this section. The implementation of this architecture is explained in details in the following sections.

10.2.1 General Operation Overview

At the highest level, a detector data acquisition system is a mix of open-loop and closed-loop control systems which provides communication to and control over the detector. On one end of this system are the very-front-end electronics and a computer system is situated on the other side. The computer system provides human interaction with the detector as well as automated control for a number of tasks. The tasks of a data acquisition system are distributed over a number of subsystems which are specifically designed to fulfil one or more requirements.

An overview of the data acquisition system of the CALICE scintillator calorimeters is illustrated in figure 10.1. The system consists of four main subsystems. Each unit of the scintillator calorimeter (see chapter 9) is equipped with a Detector Interface (DIF). The DIF is the interface between the detector unit and the rest of the DAQ system. The DIF is explained in details in section 10.3.1. Synchronisation and control of all calorimeter modules is handled by the Clock and Control Card (CCC) which is introduced in section 10.3.2. The physical links from the DIFs are aggregated by another DAQ subsystem called x-Link and Data Aggregator (xLDA). There are two types of xLDA: Mini-LDA and Wing-LDA. The functionality and implementation of the two types of the xLDA is presented in section 10.3.3. The last subsystem which provides human interface and data storage is the DAQ computer system and software components running on them which are

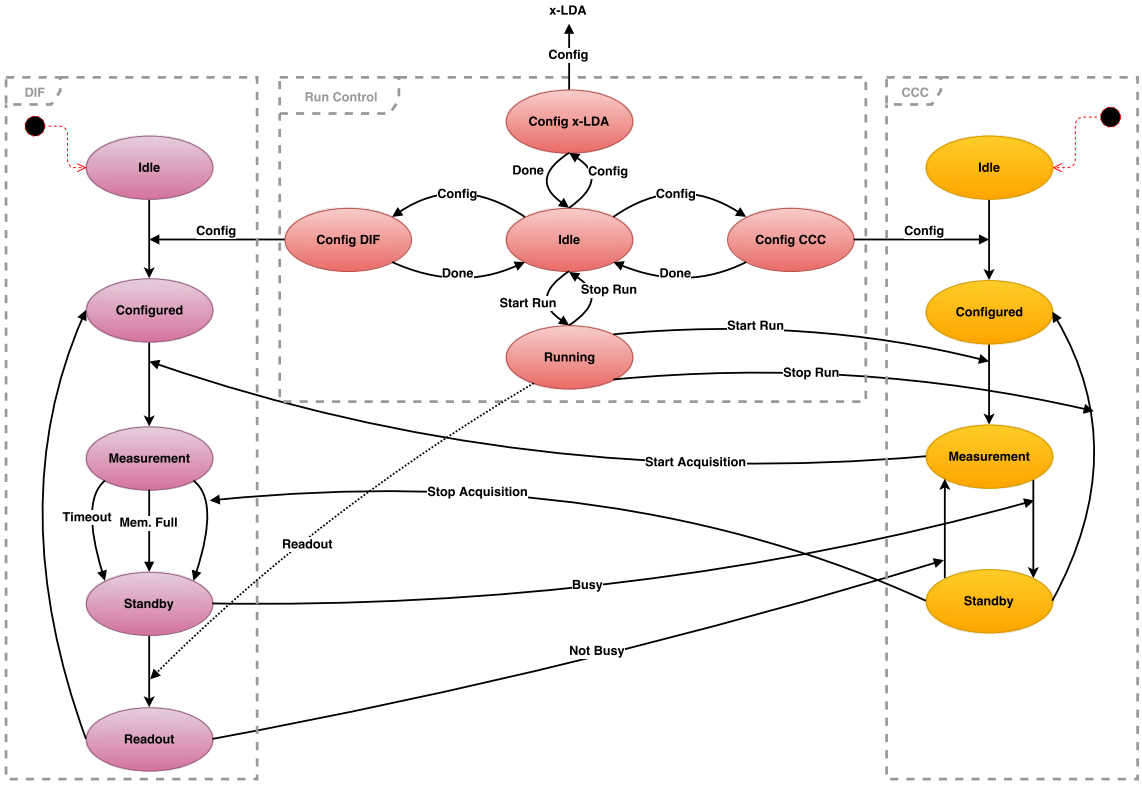


Figure 10.2 – A state diagram of the data acquisition system for the CALICE scintillator calorimeters.

described in section 10.4. The physical links between the DAQ subsystems are introduced in section 10.2.2.

The state diagram of the data acquisition system is shown in figure 10.2. The run control computer is the master of operations. It supervises correct initialisation and configuration of the detector and various DAQ subsystems. Once all subsystems are in the configured state, data taking can begin.

Data taking starts by sending a run start command from the run-control computer to the CCC. After receiving the start command, the CCC sends a start acquisition command to all calorimeter modules via the xLDA. The xLDA is transparent throughout the operation. The DIF receives the start acquisition command and dispatches it to all ASICs on the unit and data acquisition starts. When a all memory cells of an ASIC are full, it sends a busy signal to the CCC via the DIF. The CCC issues a stop acquisition command as soon as it receives a busy signal. The stop acquisition command is broadcast to all calorimeter modules and the acquisition stops regardless of the number of free memory cells on other ASICs. The data read-out starts automatically with the stop acquisition command. The DIFs read-out the data stored on all ASICs of their corresponding calorimeter module and transfer it to the data storage system via the xLDA. In addition to the busy signal, the data acquisition can be stopped due to a predefined time-out or with a manual stop command issued by the run-control computer. A data taking run is ended by a stop run command from the run-control computer to the CCC.

10.2.2 Physical Links and Communication Protocols

The calorimeter modules and DAQ subsystem are interconnected using two types of physical links. Ethernet is used as the physical link between the computer system, the CCC and the xLDAs, as illustrated in figure 10.1. The standard TCP/IP is used as the communication protocol over the Ethernet links, forming a local area network between the DAQ subsystems. The commands and data that are not time-critical are transferred via these links. The Ethernet equipment such as cables and switches are commercially available at low cost. The TCP/IP protocol is a well-established and widely used networking technology built-in the major operating systems. The TCP/IP protocol has a built-in error recovery system.

For the links which are used to transfer clocks and time-critical commands, the HDMI interface is used as the physical connection (see figure 10.1). The HDMI is an industrial standard for transferring audio/video signals and its cables are commercially available at low cost. An HDMI cable is composed of four shielded twisted pairs which are suitable for differential signalling plus a number of individual conductors. The individual conductors can be used as additional signal carriers or three of them can be combined to form an extra shielded twisted pair [118].

There are industrial standard protocols available for data transfer via HDMI connections but none of these protocols is suitable for the purpose of the data acquisition system. Therefore, the HDMI interface is solely used as a physical connection. A mixed-protocol scheme which is customised for the DAQ needs is used for communication over the HDMI links. Three of the twisted pairs are used to transfer clock, validation and busy signals as low-latency level signals. Two pairs of the remaining conductors are used for digital communication. The pin assignment is explained in detail in [119]. The digital communication links use the Universal Asynchronous Receiver/Transmitter (UART) protocol to transfer commands and data.

10.2.3 Commands and Signals

In the original CALICE conceptual design two types of frames are defined for the digital commands [120]. These types are briefly introduced here.

Fast Commands (FC) are used for time-critical commands which are used for synchronisation purposes and broadcast commands. A fast command frame is a 16 bits long word: The most significant byte of the word is called the Komma Character and the

15	...	8	7	...	0
Komma Character (K)			Command Word (D)		

least significant byte the Command Word which caries the command to be executed.

Block Transfer Commands (BTC) are used to transfer time-uncritical commands as well as configuration, status or measurement data. The length of a BTC is not fixed but only an even number of 16-bit words are allowed. Maximum length of a block is 1 kB. BTCs can be addressed to a single DIF or broadcast to all DIFs connected to a single xLDA or to all DIFs connected to all xLDAs. The structrure of a BTC frame is shown below:

Packet Type	Packet ID	Type Modifier	Specifier	Data Length	Data	CRC
16 bits	16 bits	16 bits	16 bits	16 bits	Data Length * 16 bits	16 bits

Packet type defines the type of the packet (block data, generic command, acknowledge, etc.) and also carries the DIF address to which the packet should be delivered. Packet ID is a counter of sent blocks used for identifying possible data losses. Type modifier is the definition of a command and takes an argument which is defined by the specifier. The data length carries the number of 16-bits long data words which are sent in the data block. At the end of the block comes a Cyclic Redundancy Check to detect transmission errors.

10.3 The DAQ Hardware Components

The main subsystems of the CALICE scintillator data acquisition system are briefly introduced in section 10.2. In this section a detailed specification of the hardware subsystem is presented. The DAQ software components are discussed in the following section.

10.3.1 Detector Interface (DIF)

Each unit of a scintillator calorimeter is equipped with a Detector Interface (DIF) card. The DIF is the interface between the ASICs on the calorimeter module and the rest of the DAQ system. The requirements for the DIF are listed below:

- The DIF shall match the geometry of the calorimeter module.
- The DIF shall dissipate power at a level no more than the ILD requirements.
- The DIF shall be capable of handling a full calorimeter layer with up to 72 ASICs.
- The DIF shall be capable of configuring the ASICs.
- The DIF shall physically interface with the xLDA.

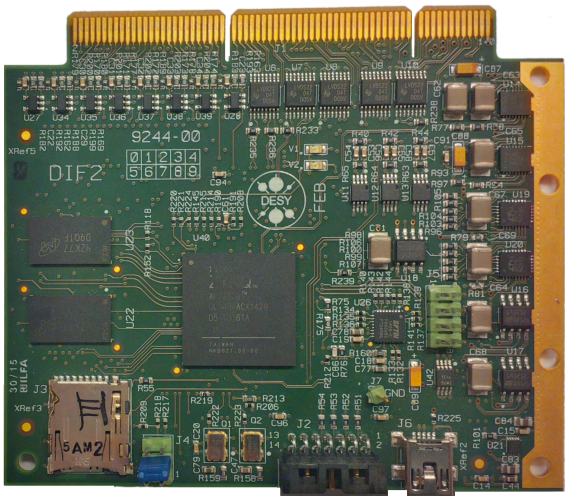


Figure 10.3 – The Detector Interface (DIF) board for CALICE scintillator calorimeters.

- The DIF shall physically interface with the Central Interface Board which provides connections to the ASICs.
- The DIF shall be capable of generating a clock for internal usage.
- The DIF shall receive a global clock on the interface to the xLDA.
- The DIF shall receive a validation level-signal on the interface to the xLDA.
- The DIF shall transfer the validation signal to all ASICs.
- The DIF shall provide the ASICs with required clocks generated using the global clock.
- The DIF shall provide the ASICs with required clocks generated using the internally-generated clock.
- The DIF shall be capable of bi-directional digital data transfer on the interface to the xLDA.
- The DIF shall be capable of decoding fast-commands and block transfer commands.
- The DIF shall send acknowledgement for the block transfer commands.
- The DIF shall receive start/stop acquisition commands.
- The DIF shall send a busy level-signal on the interface to the xLDA when one of the ASICs of the calorimeter module is full.
- The DIF shall read-out all the ASICs when a stop acquisition command is received.
- The DIF shall send read-out data to the xLDA.
- The DIF shall unset the busy signal when read-out is completed.

The DIF is realised in a plug-in card format which connects to the central interface board (CIB) on the calorimeter module, as shown in figure 10.3. The CIB provides the necessary physical interfaces to the ASICs on the calorimeter module and powers-up the DIF. A HDMI connector located on the CIB provides the interface to the xLDA. The logic to fulfil the requirements is handled by a FPGA.

10.3.2 Clock and Control Card (CCC)

The scintillator calorimeter consists of calorimeter modules with separated read-out electronics. The calorimeter modules are not interconnected, therefore, a mechanism to ensure synchronous operation of the entire calorimeter is necessary. In the CALICE scintillator DAQ system a device called Clock and Control Card (CCC) generates global clock and control signals to ensure synchronous operation of the calorimeter. The requirements for the CCC are listed here:

- The CCC shall physically interface with the beam instrumentation to receive clock, validation and spill signals.
- The CCC shall physically interface with the run-control computer.
- The CCC shall be configurable by the DAQ computer system.
- The CCC shall physically interface with up to 16 xLDAs.
- The CCC shall generate global clocks.
- The CCC shall distribute global clocks to the detector units via the xLDAs.
- The CCC shall receive start/stop run from the run-control computer.
- The CCC shall send start/stop acquisition commands.
- The CCC shall send stop acquisition signal when a busy signal is set.
- The CCC shall send start acquisition signal when the busy signal us unset.
- The CCC shall relay the validation signal from beam instrumentation to the calorimeter modules via the xLDAs.
- The CCC shall handle the spill signal from the machine.
- The CCC shall be capable of operating in slave mode by receiving the clock and the busy signal from other devices such as another CCC.

At the current stage of the development, a commercial FPGA development board called ZedBoard is used as a baseboard. The ZedBoard is a development board for the Xilinx® Zynq®-7000 All Programmable System-on-a-Chip which is introduced in the following section.

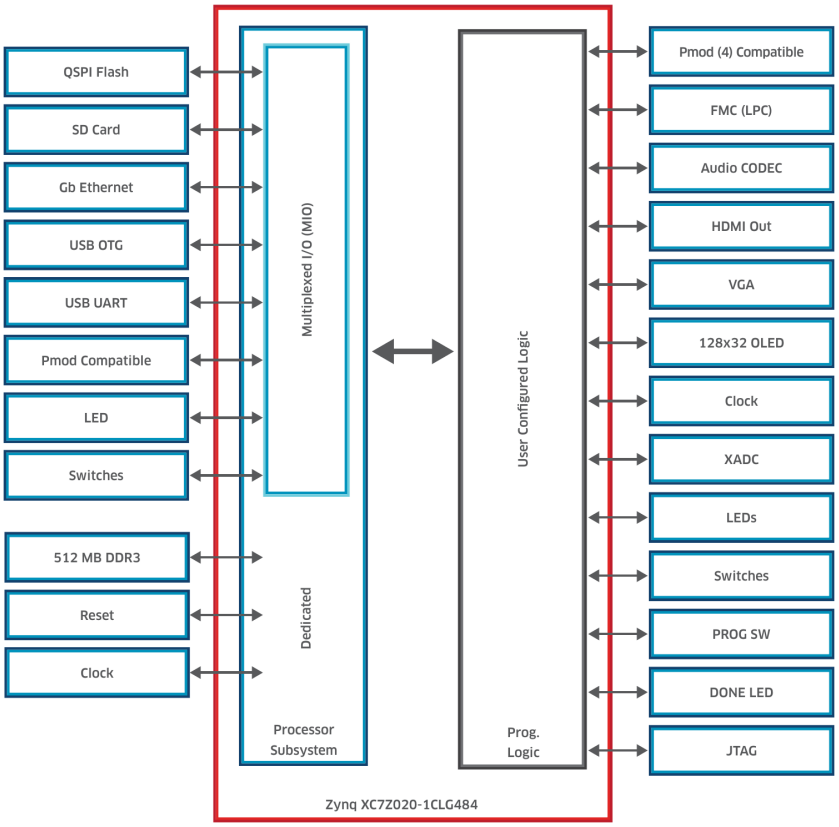


Figure 10.4 – A block diagram of the ZedBoard which shows how the peripherals are connected to the processing system and programmable logic of the Zynq-7000 SoC [121].

Zynq SoC and ZedBoard

The Zynq-7000 SoC family combines software programmability of an ARM®-based processor with hardware flexibility of a FPGA [122]. The Processing System (PS) and the FPGA which is called Programmable Logic (PL) are packaged in a single device, providing a very fast interconnect at a low power consumption.

The PS of the Zynq-7000 SoC is a dual-core ARM Cortex-A9 MPCore processor. The ARM Cortex-A9 processor family is a popular choice for applications which require low-power consumption and are thermally constrained. It is a mature technology which is widely used in the industry for devices such as smart phones and digital televisions. A complete technical specification of the Zynq-7000 SoC is available at [123].

A commercial development board for the Xilinx Zynq-7000 SoC, the so-called ZedBoard [124], is used as the development platform and as a base of the CCC and the Mini-LDA. On the ZedBoard, the Zynq-7000 SoC is supported with additional peripherals such as DDR SDRAM memory and Ethernet controller and interface. Therefore, the ZedBoard is capable of running an embedded operating system. The software programmability of the ARM processor and industry standard technologies such as networking functionalities which are built-in the operating system, make the Zynq-7000 a powerful platform for flexible and efficient designs.

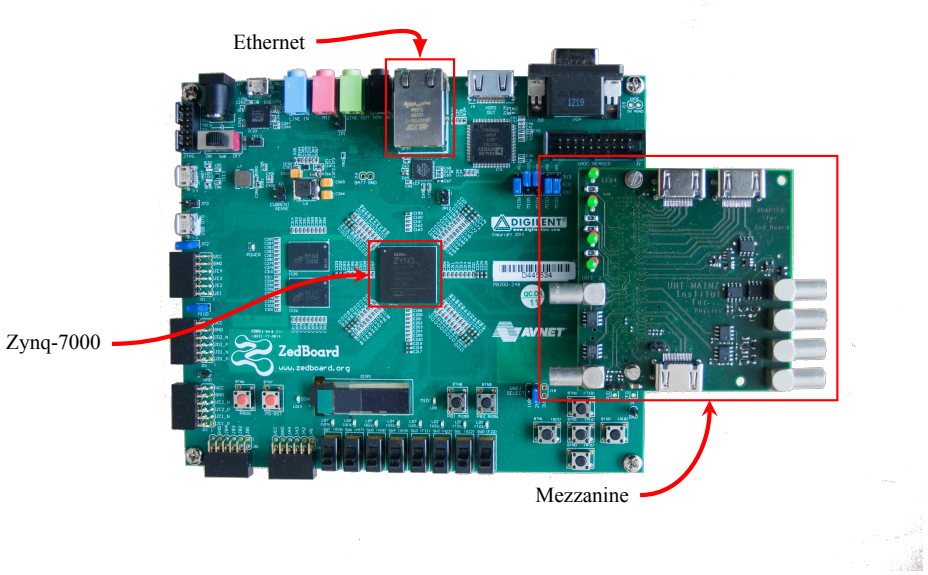


Figure 10.5 – The Clock and Control Card (CCC) for the CALICE scintillator calorimeters.

The Zedboard also offers a variety of input/output connectors which are interfaced with the PL and/or the PS (see figure 10.4). One of the most versatile connectors available on the ZedBoard is the FPGA Mezzanine Card (FMC). This connector enables the user to develop a custom expansion card according to their needs and readily connect it to the Zynq-7000 SoC. This possibility is used to realise the CCC and the Mini-LDA which is introduced in the following section.

The PS of the Zynq-7000 SoC can be used as a bare-metal system or to run an embedded operating system. Similar to a conventional FPGA, the PL of the Zynq-7000 SoC requires a firmware to be developed in a Hardware Description Language (HDL). In any case, the firmware for the PL and the software for the PS should be stored on a non-volatile memory and loaded at device power-up. On the ZedBoard a Quad-SPI flash memory or a SD memory card can be used as non-volatile memory.

The manufacturer of the Zynq-7000 SoC, Xilinx Inc, provides an integrated development environment called Vivado Design Suit. Vivado includes all necessary libraries and the tool-chain needed for embedded design for the Zynq-7000 SoC family. The firmware for the PL and software for the PS is developed in this environment.

CCC Implementation

The current version of the CCC hardware is realised by interfacing a mezzanine card to the FMC connector of a ZedBoard [125]. The mezzanine card is a passive board with a number of HDMI type A receptacle connectors, LEMO connectors and status LEDs (see figure 10.5). The HDMI connectors are used to connect to the xLDA or to a fan-out device that serves more than one xLDA. The LEMO connectors are used to connect to the beam instrumentation. The ZedBoard has an on-board Ethernet PHY controller with a RJ45

connector. This is used to connect the CCC to the run-control computer via a local area network.

The HDMI and LEMO connectors are connected to the PL of the Zynq-7000 SoC. The logic needed for fulfilling the requirements of the CCC is implemented in the firmware of the PL. An embedded Linux distribution is running on the PS which provides networking functionalities and runs the software component needed for fulfilling the CCC requirements. The embedded software components of the DAQ system are introduced in section 10.4.2.

The firmware of the PL and the embedded operating system and software of the PS are stored on a SD memory card together with a boot-loader. By switching on the ZedBoard, the boot-loader automatically programs the PL using the firmware and boots the operating system on the PS. The CCC software is launched automatically on operating system boot and runs in the background as a demon.

The CCC hardware based on the ZedBoard is intended for the development phase. From the point of view of mechanics and power supply, it is not well suited for experimental setups at beam tests or at the ILD. Therefore, another version of the CCC hardware has been designed and produced according to the VMEbus standard. This version also benefits from the Zynq-7000 SoC in form of a pluggable module called Mars module (see section 10.3.3). Hence, the firmware and software which is developed for the version based on the ZedBoard is portable to the VMEbus version with minimum modification.

10.3.3 x-Link and Data Aggregator (xLDA)

The scintillator calorimeters consists of a number of calorimeter modules. Each unit is equipped with a DIF module which is the interface to the rest of the DAQ system. The physical links from the calorimeter modules need to be aggregated by a central hub. Logically, this device should route commands from the run-control computer and the CCC to the appropriate destination(s) on the calorimeter modules and collect read-out data from them. The device which handles this task is called x-Link and Data Aggregator (xLDA) which is based on the LDA in the original CALICE conceptual design. The requirements for such a device are:

- The xLDA shall eventually match the ILD geometry.
- The xLDA shall physically interface with the run-control computer.
- The xLDA shall physically interface with the data storage system.
- The xLDA shall physically interface with the slow-control computer.
- The xLDA shall physically interface with the CCC.
- The xLDA shall physically interface with multiple calorimeter modules, the number of which is defined by the ILD geometry.

- The xLDA shall be capable of enabling/disabling interface ports to the calorimeter modules.
- The xLDA shall receive a global clock from the CCC.
- The xLDA shall receive a validation signal from the CCC.
- The xLDA shall distribute the global clock to the calorimeter modules.
- The xLDA shall distribute the validation signal to the calorimeter modules.
- The xLDA shall send a busy signal to the CCC when at least one calorimeter modules is busy.
- The xLDA shall be configurable by the DAQ computer system.
- The xLDA shall receive commands from the run-control computer.
- The xLDA shall receive commands from the slow-control system.
- The xLDA shall receive commands from the data-storage system.
- The xLDA shall transfer commands from the CCC to the calorimeter modules.
- The xLDA shall transfer commands from the run-control computer to the calorimeter modules.
- The xLDA shall transfer commands from the slow-control computer to the calorimeter modules.
- The xLDA shall aggregate data from the calorimeter modules.
- The xLDA shall decode the read-out data (the ASIC output is grey-coded).
- The xLDA shall transfer aggregated data to the data storage system.

Two types of the xLDAs are developed based on the Zynq-7000 SoC. The first type which is called Mini-LDA and can support up to 10 calorimeter units was used as the logic development platform and at the end of the development is used for smaller calorimeter setups in the laboratory and at beam tests. The second version that is called Wing-LDA supports up to 96 calorimeter units and is designed to mechanically match the geometry of the ILD. These two types of the xLDA are described in detail in the following sections.

Mini-LDA

As part of this thesis, the software and logic development for the xLDA was performed on a smaller type called Mini-LDA. Similar to the CCC, the Mini-LDA uses a ZedBoard as the base with a customised mezzanine card attached to the FMC connector on the ZedBoard, as shown in figure 10.6.

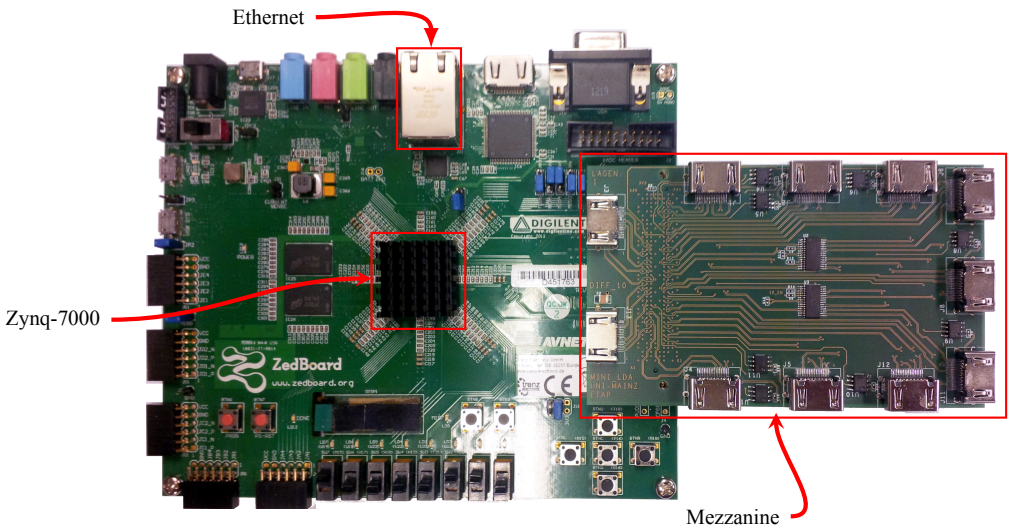


Figure 10.6 – The Mini-Link and Data Aggregator (Mini-LDA) for the CALICE scintillator calorimeters.

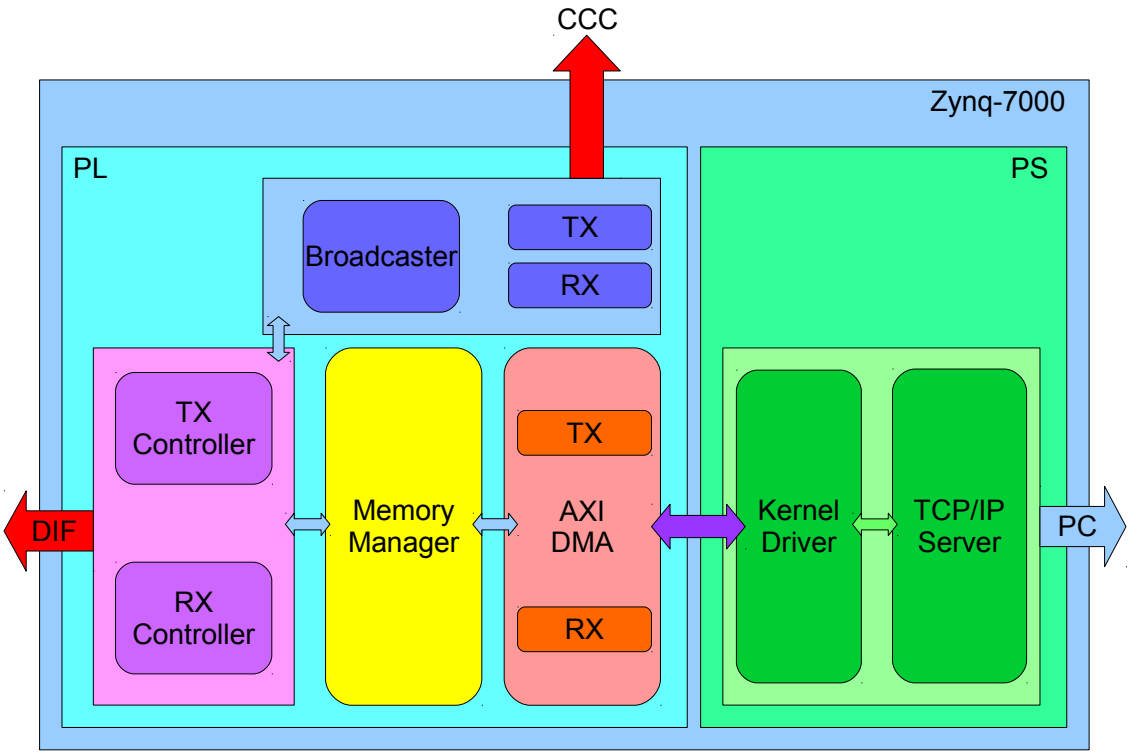


Figure 10.7 – A simplified block diagram of the Mini-LDA.

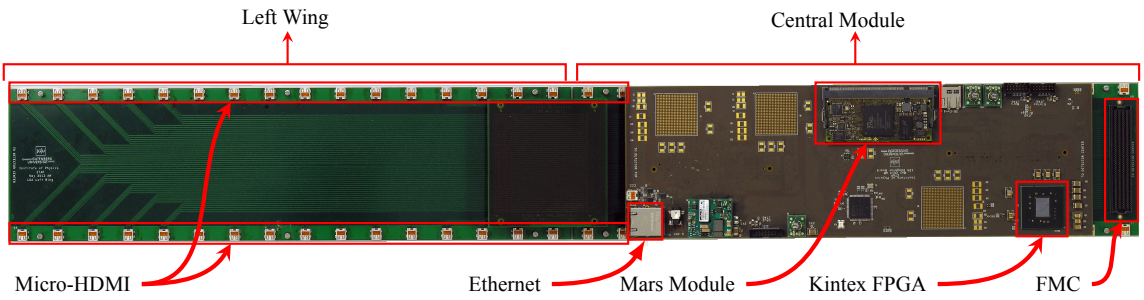


Figure 10.8 – The Wing-Link and Data Aggregator (Wing-LDA) for the CALICE scintillator calorimeters. Only the left wing is shown.

The mezzanine card has 11 HDMI type A receptacle connectors. One of these HDMI connectors is used for a physical link to the CCC and the remaining 10 ports are for physical links to the calorimeter modules. The RJ45 connector on the ZedBoard is used to establish a physical link to the DAQ local area network using the TCP/IP protocol.

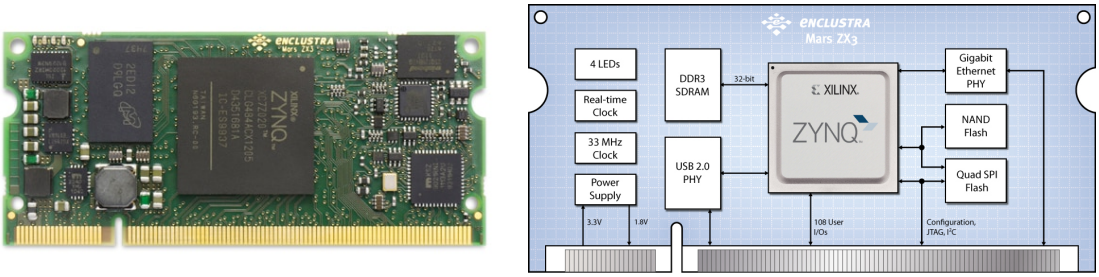
During the operation of a small AHCAL setup using a Mini-LDA, it was revealed that the setup is mechanically unstable due to the weight and the strain of the rigid halogen-free HDMI cables. The standard HDMI type A receptacle connectors do not provide a locking mechanism to mechanically secure the connected cable. Therefore, the cable plug might move slightly out of its proper position under cable stress and cause loss of some signals. A mechanical support for cables was employed to reduce the stress and mitigate the problem.

A simplified block diagram of the Mini-LDA firmware and embedded software is shown in figure 10.7. The HDMI ports on the mezzanine card are interfaced to the PL of the Zynq-7000 SoC on the ZedBoard. The logic implemented in the PL receives the global clock and validation signal on the link to the CCC and distributes it to the links to the calorimeter modules. Each of the links to the CCC and to the calorimeter modules is equipped with a UART receiver and a UART transmitter (on different pairs of wires) for digital communication. A complex memory manager module supervises data collection and routing. This module routes commands and configuration packets to the correct destination by decoding packet headers. It also buffers the incoming data from the DIF, decodes and encapsulates them in packets to be transferred to the PS. The link between the PL and the PS is established using an AXI DMA interface [126]. An embedded distribution of the Linux operating system runs on the PS which provides networking infrastructure and also hosts the dedicated xLDA software that is presented in section 10.4.2.

The Mini-LDA is intended to be used as an early-stage development platform and also for small calorimeter setups with up to 10 calorimeter modules.

Wing-LDA

A version of the xLDA which should meet the requirements of a calorimeter for the ILD is called Wing-LDA [125]. The Wing-LDA hardware, which is shown in figure 10.8, matches



(a) (b)

Figure 10.9 – a) A photo and b) the block diagram of the Mars ZX3 SoC module by Enclustra [127].

the geometry requirements of the ILD. The type A standard HDMI connectors are replaced by type D Micro HDMI on the Wing-LDA which is smaller and allows for the necessary layer spacing. In addition, the smaller connectors and cables reduces the problem with mechanical stress.

The current version of the Wing-LDA consists of four boards which are connected together using FMC connectors. Three of the boards are passive and solely host 96 type D Micro HDMI receptacle connectors used for establishing physical links to detector units. The remaining board is home to the logic devices as well as the connectors for the Ethernet connection and the physical link to the CCC. The power subsystem of the Wing-LDA is also located on this board.

Similar to the CCC and the Mini-LDA, a Zynq-7000 SoC is used to handle the main logic of the device. The Zynq-7000 SoC is added to the board in form of a commercial plug-in module called Mars ZX3 SoC module by Enclustra [127]. In addition to the Zynq-7000, the Mars module provides DDR SDRAM memory, non-volatile NAND and Quad-SPI flash memories, a Gigabit Ethernet PHY and a real-time clock. A picture of the Mars module and its block diagram are shown in figure 10.9. Debugging and commissioning of the Mars module on the Wing-LDA hardware was carried out as part of this thesis.

The Mars module has a SO-DIMM form factor with 200 pins and offers 108 user input/outputs interfaces with the Zynq-7000 SoC. The number of I/Os available on a Zynq-7000 SoC is not enough to serve the required number of HDMI ports. Therefore, additional logic units are required between the Zynq-7000 SoC and the HDMI ports. The Wing-LDA employs four Xilinx Kintex FPGAs for this purpose. Each of the Kintex FPGAs interface the Zynq-7000 SoC with up to 24 HDMI ports.

Figure 10.10 illustrates a block diagram of the firmware and software for the Wing-LDA. The overall structure is similar to that of the Mini-LDA. However, due to the aforementioned reasons, the firmware functionality is split between the Zynq-7000 SoC and the Kintex FPGAs, as depicted by the dashed line on the figure.

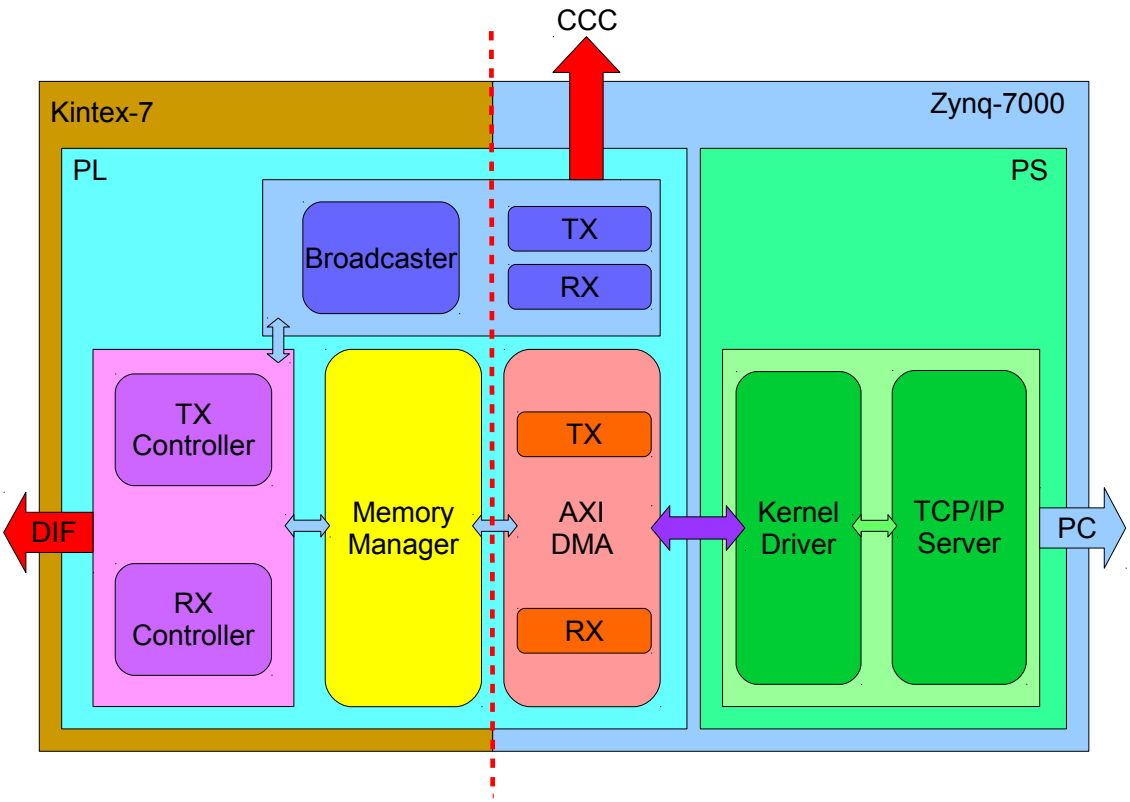


Figure 10.10 – A simplified Block diagram of the Wing-LDA.

10.4 The DAQ Software Components

In a typical data acquisition system for high-energy physics experiments computers are used for higher-level control of the DAQ system. The computer system requires dedicated software to address the needs of the experiment. The DAQ software consists of multiple components designed to perform specific tasks. In larger experiments, these components are usually distributed over multiple computers. The software components are used for configuration of the various detector subsystems, system control, data handling, monitoring, etc..

The DAQ system for the CALICE scintillator calorimeters includes a number of computers. These computers are used for run-control, slow-control, monitoring and data storage. The DAQ computers run software which is specifically developed for their purpose. Furthermore, other subsystems of the DAQ system also run an embedded operating system with dedicated software. The various software components of the DAQ system are presented in this section.

10.4.1 Main Configuration and Control Software

The DAQ software is used for initialisation and configuration of the calorimeter modules and DAQ subsystems. Calibration and data-taking runs are also controlled by the DAQ software at the highest level. Moreover, the DAQ software is the main interface for human interactions with the calorimeter system. The main requirements for the DAQ software are listed below:

- The DAQ software shall provide an interface for human interactions with the system.
- The DAQ software shall communicate with the DAQ subsystems.
- The DAQ software shall supervise correct calorimeter system initialisation.
- The DAQ software shall be capable of configuring the DAQ subsystems.
- The DAQ software shall be capable of configuring the calorimeter modules.
- The DAQ software shall send start/stop run command to the CCC.
- The DAQ software shall be capable of setting data-taking run parameters.
- The DAQ software shall provide monitoring functionalities.
- The DAQ software shall provide data storage functionality.
- The DAQ software shall be capable of interfacing with a higher-level common DAQ system.

The requirements mentioned above are usually fulfilled by various software components which run on different computers. At the current stage, the main software of the scintillator DAQ is able to fulfil the aforementioned requirements on a single computer.

The DAQ software for the CALICE scintillator calorimeters is developed on the Laboratory Virtual Instrument Engineering Workbench (LabVIEW) platform. LabVIEW is a system-design and development platform which is commonly used for automation, instrument control and data acquisition. LabVIEW libraries include a large number of functions to perform common tasks, facilitating the development process. As part of this thesis, an existing run-control software was extended to fulfil the requirements of the new DAQ system.

The scintillator DAQ software provides a Graphical User Interface (GUI). The GUI is used to

- initialise the system,
- power-up/down the calorimeter modules,

- configure the DAQ subsystems and the calorimeter modules,
- setting up calibration run parameters,
- setting up data-taking run parameters,
- start/stop data taking,
- data read-out,
- low-level (ADC and TDC) data monitoring,
- data storage.

The DAQ software communicates with the CCC and the xLDA via the TCP/IP protocol over the DAQ local area network. Currently, only one network interface is used for all communications. Should more bandwidth be required in the future, more network interfaces or optical links can be employed.

The LabVIEW DAQ software stores the read-out data as raw data or in a tab-separated ASCII file format. This format, although human readable, is not a suitable solution for large data storage since file size is inherently big and also slow to read and write. The desirable format for the CALICE calorimeters is the LCIO format since it is a binary format based on a well-defined event data model with data compression and can be directly used by the iLCSoft packages (see section 6.1). Data storage in LCIO format is achieved by using EUDAQ software, as explained in section 10.5.

The LabVIEW framework has many advantages for rapid system-design and modification during detector development. However, it has some disadvantages compared to hand-coded software in native languages such as C or C++. LabVIEW applications are usually slower than natively developed applications. The LabVIEW applications need the LabVIEW runtime environment which adds an overhead to memory and processor usage. For such reasons, ultimately the DAQ software for a large calorimeter needs to be implemented using a native programming language.

10.4.2 Embedded Software

As mentioned in sections 10.3.2 and 10.3.3, the CCC and the xLDA run an embedded Linux operating system. Although the operating system provides basic standard infrastructures such as networking and memory management, a dedicated software is required to fulfil the requirements of the devices. The following are the general requirements for such a software:

- The embedded software shall communicate with the DAQ computer subsystems using the networking infrastructures of the operating system.

- The embedded software shall transfer data from/to the PL of the Zynq-7000 SoC.
- The embedded software shall automatically run in the background after the boot process of the device is completed.
- The embedded software shall handle the commands from the computer subsystems.

To develop the embedded software for devices based on the Zynq-7000 SoC, Vivado Design Suit is used. The architecture of the processor of the personal computer on which Vivado development environment is hosted is different than the architecture of the ARM processor of the Zynq-7000 SoC. Therefore, the software which is developed for the Zynq-7000 SoC needs to be cross compiled [128].

The Embedded Software for the CCC

The CCC employs an embedded software called Piconf [129]. Piconf is an XML-based framework for automating common procedures such as device configuration by providing general purpose tools. Piconf communicates with the run-control computer using the TCP/IP protocol to receive commands and configuration data. Piconf then transfers them to the PL of the Zynq-7000 SoC. Piconf runs as a demon in the background and also provides a command line interface for user interactions.

The interface between the PS and the PL of the Zynq-7000 SoC is provided by the Advanced eXtensible Interface (AXI). There are a number of different AXI interfaces available with various working principles and data width. Piconf uses an interface called AXI General Purpose I/O (GPIO) which provides a software interface to the AXI interface. The commands and configuration are then transferred between the PL and the PS using this interface.

During the development phase it was discovered that data corruption can occur when data are transferred via the AXI GPIO interface. The probability of data corruption increases with the data rate. Since the data rate on the CCC is rather low this is not a significant issue at this stage but it needs to be solved to ensure stable operation of the system in long runs.

The Embedded Software for the xLDA

The task of the xLDA is more complex with respect to the CCC and the data transfer rate is much higher. The latest version of Piconf is not capable of large data transfer. Furthermore, the AXI GPIO interface which Piconf uses for communication between the PL and the PS of the Zynq-7000 SoC is not fast enough for this purpose. Also, the data transferred via this interface can be corrupted. Therefore, a dedicated Linux kernel driver and userland software is developed for the xLDA.

The kernel driver provides an interface for fast and secure transfer of a large volume of data using a shared memory controller available on the ZedBoard.

The userland software, called XLDAS, acts as a TCP/IP server. It listens on a specific TCP port and establishes a connection when a request is received from a client such as the run-control computer. XLDAS also employs the kernel driver to transfer command and data to/from the PL.

Since the main software of the DAQ system is used for data storage as well, currently, the XLDAS uses only one TCP/IP connection for all purposes including command and data transfer. In the future, when the storage is handled by a dedicated storage system, the xLDA should be able to establish other TCP/IP connections to the storage system. The various connections can use the same physical link or can be distributed over multiple links depending on the data rate requirements.

10.5 Integration into a Common DAQ System

Eventually, a full scintillator calorimeter will be integrated into a general-purpose detector such as the ILD and it should operate with other sub-detectors. Although a new DAQ system that can integrate all subdetectors directly can be developed, usually a modular approach is preferred. Therefore, the data acquisition system of the calorimeter should be capable of running in slave mode when the top-level DAQ system of the ILD is the master of operation. Furthermore, during the detector research and development phase, it is advantageous to operate various types of detector prototypes in a combined setup for beam tests. Hence, a DAQ system capable of integrating into a common DAQ system is required during the development phase as well.

Currently the design of the ILD is not finalised and its baseline design is undergoing an optimisation process. There are a few options available for each of the ILD sub-detectors. These options are based on different technologies with different read-out and operation schemes. Therefore, the requirements for a top-level ILD data acquisition system are not well-defined at this stage. Subsequently, in the design and implementation of the data acquisition system for the CALICE scintillator calorimeters only generic requirements for integration into a top-level common DAQ system are considered.

At the current stage, the CALICE collaboration and other Linear Collider detector research and development programs have opted for a generic data acquisition framework named EUDAQ as a top-level common DAQ system for beam tests. The general operational concept of the EUDAQ and integration of the DAQ system for the CALICE scintillator calorimeters is presented in the following sections.

10.5.1 EUDAQ

EUDAQ is a generic modular software for data acquisition written in C++ [130]. The primary goal of the software is to provide a generic light-weight easy-to-use platform for building DAQ systems. The genericness is achieved by separating hardware-specific parts of the software from the core functionalities in such a way that the core libraries can be used independently of the hardware.

EUDAQ consists of a number of modules which communicate internally using the TCP/IP protocol. The core module of EUDAQ is the run control which is the master of operation and manages the other modules. It also provides a user interface for controlling the entire DAQ system. Any hardware which produces data has a producer module. This module is specific to the hardware and is used to initialise, configure and control data taking of the hardware by receiving commands from the run-control module. The producer module also reads-out the data and sends it to another module called Data Collector which combines the data from all producers and sends it to the storage system. There is a Log Collector module for logging and a Monitor module for data monitoring purposes.

EUDAQ can store the read-out data in raw and LCIO format. The LCIO is the favourable format for the CALICE calorimeter data since it is the common file format for Linear Collider studies, as explained in section 6.1.

10.5.2 Integration with EUDAQ

The main software of the DAQ system for the CALICE scintillator calorimeters is capable of running in slave mode while an instance of EUDAQ version 1 acts as a common top-level DAQ system. The scintillator DAQ software establishes a standard TCP/IP connection on a specific port to communicate with the computer hosting an instance of the EUDAQ.

A dedicated EUDAQ producer module is developed for the CALICE scintillator calorimeters. Currently the EUDAQ can be used to perform:

- establish a TCP/IP connection to main scintillator DAQ software
- set LED voltages for calibration runs,
- start and stop a data-taking run,
- receive read-out data and monitoring information,
- store the data in raw and LCIO file formats,
- perform a scan of parameters such as LED voltages.

The calorimeter DAQ subsystems and calorimeter modules need to be initialised and configured using the main scintillator DAQ software before EUDAQ becomes the master of operations.

As part of the EUDAQ configuration process, a TCP/IP connection to the main scintillator DAQ software is established. The configuration process can also configure LED calibration runs by sending a path to the location where LED voltage setting files are stored to the main scintillator DAQ software. The main scintillator DAQ software then automatically sets the LED voltages for each calorimeter module according to the configuration files. This feature allows for parameter scans and automatic restart of runs with a defined number of read-out cycles.

Once the EUDAQ configuration process is completed, a calibration or data-taking run can be started and stopped by the EUDAQ. Currently, the main scintillator DAQ software receives the read-out data and then transfers it to the EUDAQ computer via the Ethernet connection. But in principle, the EUDAQ can be considered as the storage system and should receive the read-out data directly from the xLDA.

The general architecture and various subsystems of the data acquisition system for the CALICE scintillator calorimeters are introduced in this chapter. The full system has been commissioned and tested in the laboratory and at the CERN PS test beam facility which is presented in the next chapter.

Chapter 11

The AHCAL DAQ Characterisation and Performance

The architecture of the data acquisition system for the engineering prototype of the CALICE AHCAL is presented in chapter 10. During the development phase, individual components of the DAQ system were tested regularly to verify the functionality of the components. However, a full system test is feasible only when all components are already in an advanced stage of the development process.

The first tests of the data acquisition system for the CALICE scintillator calorimeters were performed in a laboratory using a single AHCAL module consisting of only one HBU. The goal of these tests was to operate the new DAQ system with all subsystems to verify its functionality and perform the necessary debugging and improvements. These tests are described in section 11.1.

Although the laboratory tests are advantageous for debugging purposes and general functionality tests, some aspects of the data acquisition system can only be tested with particle beams. The commissioning of the data acquisition system using particles was performed at the CERN Proton Synchrotron in 2014. The beam-line installation is presented in section 11.2.1. The calorimeter stack at the test beam consisted of several CALICE AHCAL and Sc-ECAL modules which is explained in details in section 11.2.2. The beam-line instrumentation and DAQ system setup for the 2014 test beam is discussed in section 11.2.3. A number of tests were performed to analyse the performance of the DAQ system. The results of these tests and a summary of the DAQ system performance are presented in section 11.2.4.

Apart from the commissioning of the DAQ system for the CALICE scintillator calorimeters, during the 2014 beam test at the CERN PS, combined running with the CALICE Silicon-tungsten Electromagnetic Calorimeter (Si-ECAL) was successfully tested. The two calorimeter technologies utilise different DAQ systems. Therefore, the combined operation

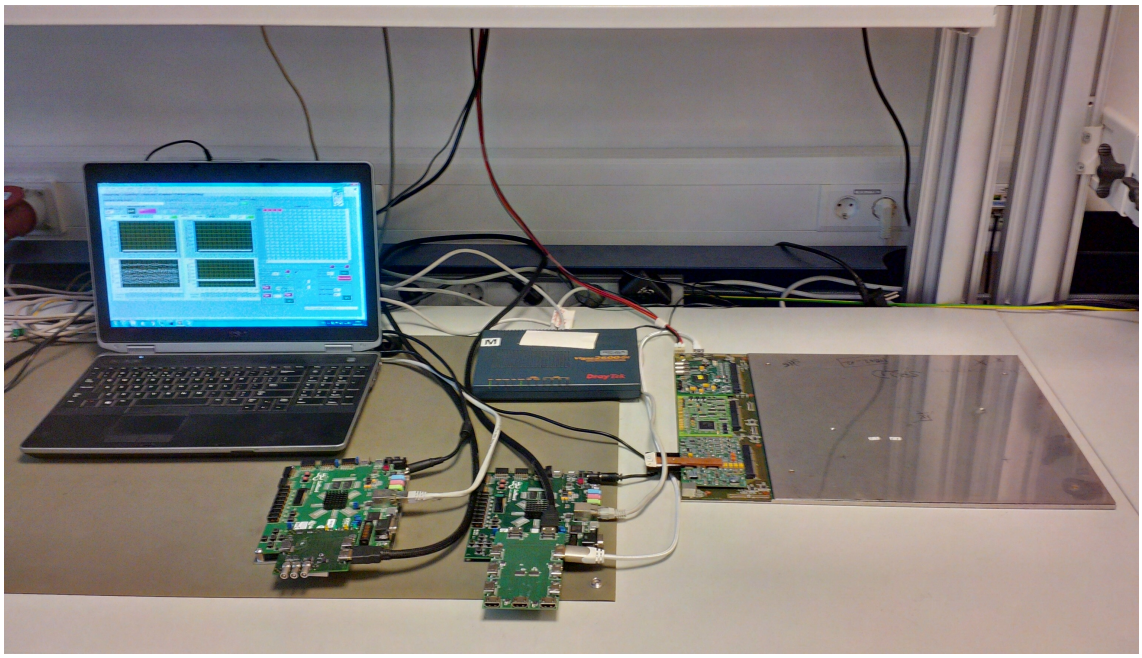


Figure 11.1 – Experimental setup in the laboratory.

of the two technologies required integration into a common data acquisition system. The combined operation using a common DAQ system is explained in section 11.2.5.

Since the 2014 beam tests at the CERN PS, the performance of the data acquisition system of the AHCAL engineering prototype has been improved. The current status of the DAQ system and the plans for the future are discussed in section 11.3.

11.1 Laboratory Tests with the LED Calibration System

Each subsystem of the data acquisition system for the CALICE scintillator calorimeters was developed individually according its functional requirements. During the early stages of the development, certain aspects of each component could be tested as a stand-alone device. These tests are necessary for initial performance evaluation and debugging purposes. When all subsystems pass the stand-alone tests successfully, full system tests involving all DAQ subsystems are required for further debugging and to verify the performance of the entire system.

The AHCAL engineering prototype is equipped with a LED calibration system (see section 9.3.8). The calibration is performed by means of LED light and no particle beam is required. Therefore, the LED calibration system can be used for various tests in the laboratory without the need for particle beams.

The first tests of the DAQ system for the CALICE scintillator calorimeter prototypes were performed in a laboratory setup. The experimental setup in the laboratory is shown in figure 11.1. An AHCAL module consisting of one HBU and a CIB was operated with the

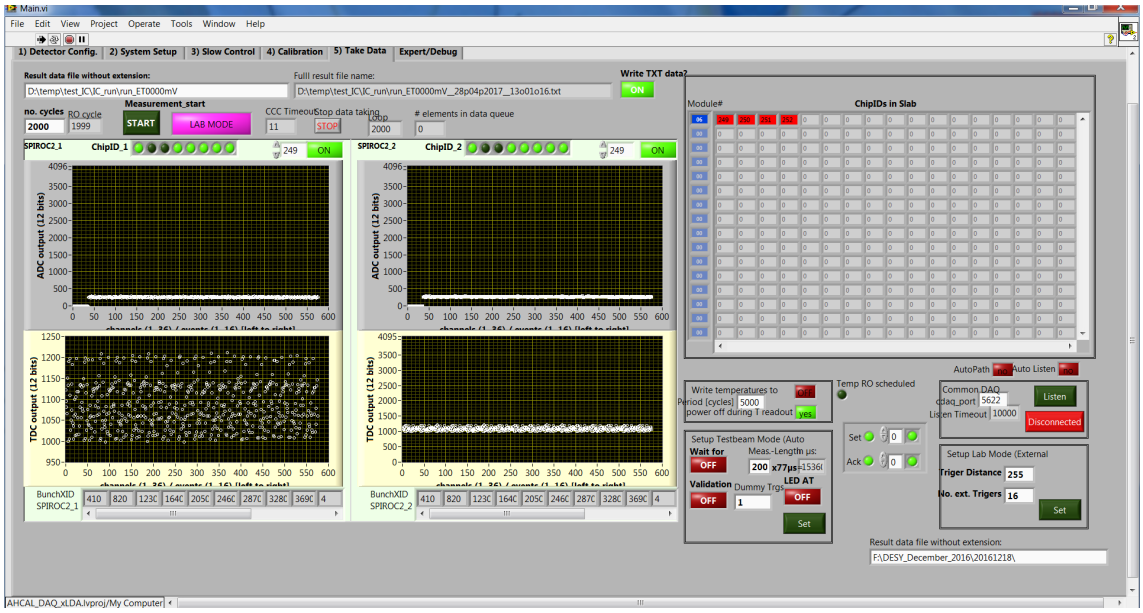


Figure 11.2 – The user interface of the main DAQ software.

new DAQ system. The HBU was equipped with four SPIROC2b ASICs and 144 scintillator tiles. A Mini-LDA was connected to a CCC and to the calorimeter module using two HDMI cables. A standard network switch was used to setup a Local Area Network between the CCC, Mini-LDA and the run-control computer.

The run-control computer was hosting the main DAQ software which is developed in the LabVIEW platform. The software has multiple tabs for initialisation, configuration, calibration, data taking and debugging. The data taking tab is shown in figure 11.2. In this tab, the user is able to choose external-trigger or auto-trigger modes and start a run for a specified number of read-out cycles. A list of all active ASICs is displayed on the top-right table. The ADC and the TDC values for two different ASICs can be displayed online.

Various bugs were found and fixed during the laboratory tests. At the end of the debugging phase, the calorimeter module was operated in calibration mode for many hours to verify the stable operation of the system. For externally triggered calibration runs, a read-out rate of ~ 4 read-out cycles per second was achieved. This read-out rate is about a factor of four faster than the rate achieved by the previously used single-layer DAQ system based on a direct USB link between the run-control computer and a DIF. More results on externally triggered calibration runs of multiple calorimeter modules are presented in section 11.2.4.

11.2 Beam Tests at the CERN Proton Synchrotron

In November and December 2014 a test beam campaign at the CERN Proton Synchrotron was organised with two main goals. The first goal was the commissioning of the new data

acquisition system using multiple AHCAL and Sc-ECAL modules at the test beam. The second main goal was to employ the new DAQ system to commission new calorimeter modules and take data with muon and pion beams. In addition to the main goal, a combined run with the CALICE Si-ECAL was successfully performed for the first time.

In this section, the details of the 12 days campaign in 2014 at the CERN PS test beam facility are presented.

11.2.1 Beam-line Installation

The test beam campaign of November-December 2014 took place at the T9 beam-line of the CERN Proton Synchrotron test beam facility. The T9 beam is a secondary beam produced by the primary proton beam of the PS accelerator impinging on a target. The collision of protons with the target material creates a beam composed of various particles such as electrons, positrons, pions, kaons and protons. The beam provided to the experimental area is composed of either positively or negatively charged particles. The particle composition of the beam depends on the beam energy. The momenta of the beam particles vary between 0.5 GeV and 10 GeV. The user can select the desired momentum

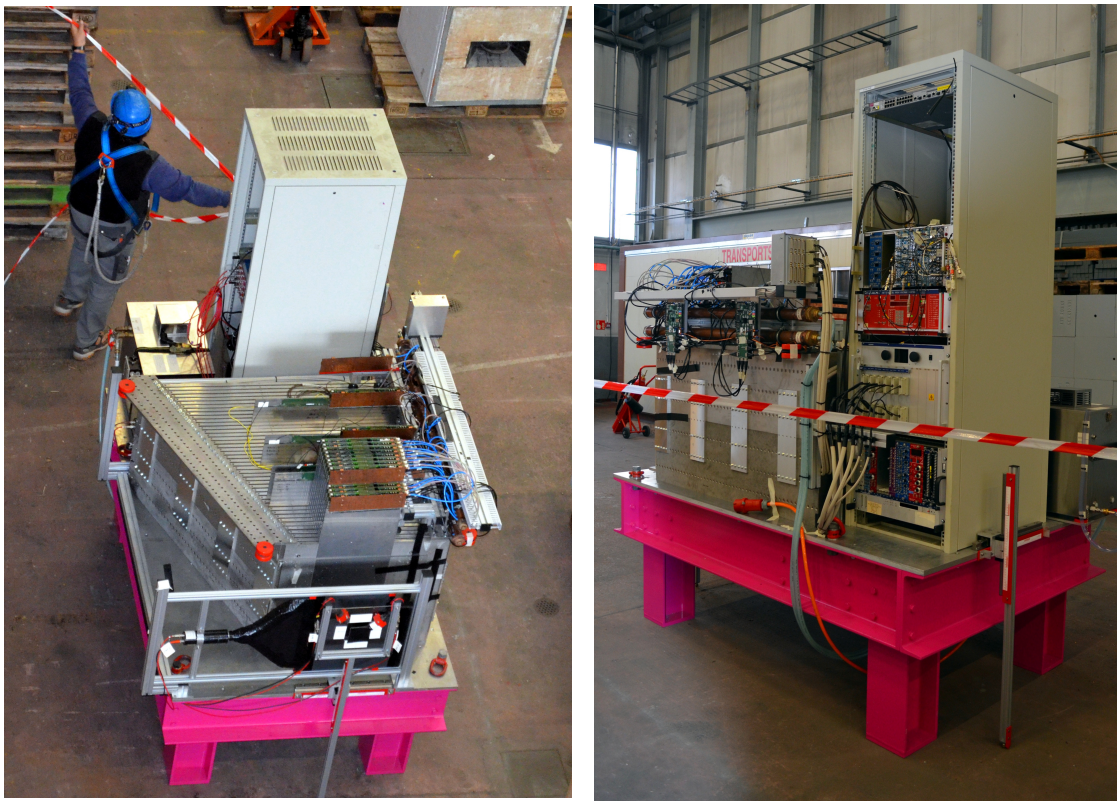
The T9 experimental area is about 4×12 m^2 and is equipped with a number of fixed and removable detectors which can be used to analyse the characteristics and composition of the beam. A gas Cherenkov detector was used to discriminate between electrons, pions and heavier particles. The information from the beam-line instrumentation could not be used directly in the AHCAL DAQ system.

The beam is delivered uniformly in bursts of 0.4 s length (one spill). Typically, there are $1-2$ spills every 15 seconds. The maximum particle rate of 10^6 particles per second is achieved for the 10 GeV beam of positively charged particles. The beam rate drops for the lower energies and is typically lower for the beams of negatively charged particles of the same momentum.

11.2.2 Detector Setup

The setup of the calorimeter technological prototype at the CERN PS consisted of 12 CALICE AHCAL and 3 CALICE Sc-ECAL modules. The modules were placed in a steel absorber structure similar to the half-octant structure planned for the ILD barrel, as shown in figure 11.3. The absorber plates are 17.2mm thick and separated by a 8.8 mm gap where the calorimeter modules are placed. The stack was mounted on a fixed platform.

The first three slots (slots $0-2$) where occupied by the Sc-ECAL modules, each consisting of one single ECal Base Unit (EBU). The slots $3-10$ were hosting AHCAL modules composed of one HCal Base Unit (HBU). Four large AHCAL modules comprising



(a) (b)

Figure 11.3 – The AHCAL setup at the CERN Proton Synchrotron in 2014.

2×2 HBUs were placed in slots 11, 13, 21 and 31. The stack was protected from the ambient light using a light-tight cover.

The calorimeter modules were equipped with different types of scintillator tiles/strips and SiPMs. In total, the calorimeter stack consisted of approximately 3800 read-out channels. A rack installed on the platform next to the stack was hosting a commercial power supply and various DAQ subsystems. The power supply was controlled by a dedicated software. For safety reason, the power supply control software was hosted on an separate computer.

Two sets of validation scintillator plates were installed on the stack. The first set consisted of two scintillator plates of 10×10 cm^2 installed in front of the stack and was used for validation of electrons and pions. The second set comprised two scintillator plates of about 30×30 cm^2, one of which was installed in front and the second on the backside of the stack. This set was used for validation of muons. The validation scintillator plates were read-out by PMTs and the coincidence of each set was used as a validation signal.

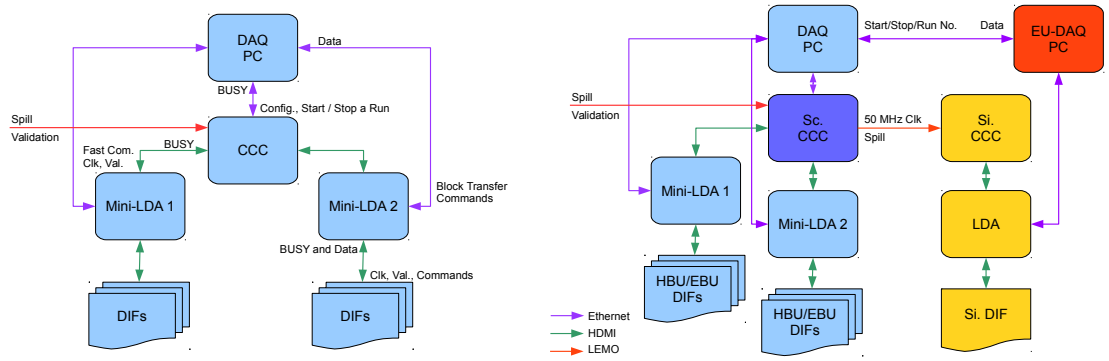


Figure 11.4 – Schematic diagram of the scintillator DAQ for the 2014 beam tests at the CERN PS a) in stand-alone configuration b) in combined configuration.

11.2.3 Data Acquisition System System

A diagram of the DAQ system setup for the 2014 beam tests at the CERN PS is illustrated in figure 11.4. The Wing-LDA was not ready to be commissioned during this beam tests. Therefore, two Mini-LDAs were employed to operate the 15 calorimeter modules. 8 modules (layers 0, 2, 4, ..., 14) were connected to the first Mini-LDA and the remaining 7 modules were connected to the second Mini-LDA. The reason for such a distribution of even and odd layers on the two Mini-LDAs was to mitigate the risk of loosing part of the calorimeter data in case of synchronisation problems. In such interleaved arrangement, it is possible to reconstruct a full shower using the data of calorimeter modules connected to only one Mini-LDA.

The two Mini-LDAs were connected to a CCC using HDMI cables. The spill signal provided by the beam facility instrumentation and the signals from the validation scintillators were processed using NIM electronics and were fed to the CCC using coaxial cables connected to the LEMO connectors on the CCC.

A commercial network switch was used to setup a Local Area Network between the computers, the Mini-LDAs and the CCC. The run control computer was placed in the beam area and was equipped with two Network Interface Cards (NIC). One of the NICs was used to connect to the DAQ network and the other was connected to the beam area Intranet to provide remote access to the computer. The power supply control computer was placed in the counting room and was directly connected to the power supply via a dedicated Ethernet cable.

11.2.4 Test Results

The DAQ system for the CALICE scintillator calorimeters was commissioned successfully during the 2014 beam test campaign at the CERN PS. In the first step, a number of externally triggered calibration runs were performed in order to verify performance of the DAQ system and data integrity. After the error-free operation of the entire system was assured, data taking with particle beams in auto-trigger mode was carried out.

In externally triggered calibration runs, parameters such as the number of triggers and the distance between successive triggers are set by the operator. Therefore, the theoretical expected outcome of a calibration run can be checked against the measurement results to verify the performance of the system.

The total number of events expected from the entire stack for an externally triggered calibration run of 1000 read-out cycles is obtained by

$$N_{evt/cycle} = N_{ASIC} \times N_{channel} \times N_{mem} = 108 \times 36 \times 16 = 62,208 \quad (11.1)$$

$$N_{evt/ASIC} = N_{cycle} \times N_{channel} \times N_{mem} = 1000 \times 36 \times 16 = 576,000 \quad (11.2)$$

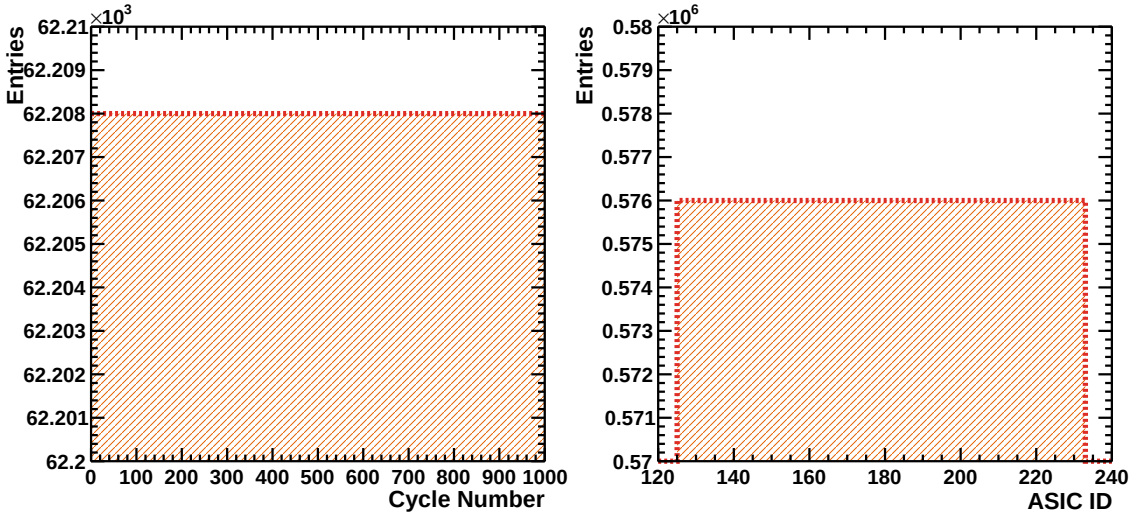
$$N_{evt/channel} = N_{cycle} \times N_{ASIC} \times N_{mem} = 1000 \times 108 \times 16 = 1,728,000 \quad (11.3)$$

$$N_{evt/mem} = N_{cycle} \times N_{ASIC} \times N_{channel} = 1000 \times 108 \times 36 = 3,888,000 \quad (11.4)$$

where $N_{ASIC} = 108$ is total number of ASICs in the setup, $N_{channel} = 36$ is number of read-out channels per ASIC, $N_{mem} = 16$ is number of analog memory cells for each read-out channel and $N_{cycles} = 1000$ is number of read-out cycles of the run.

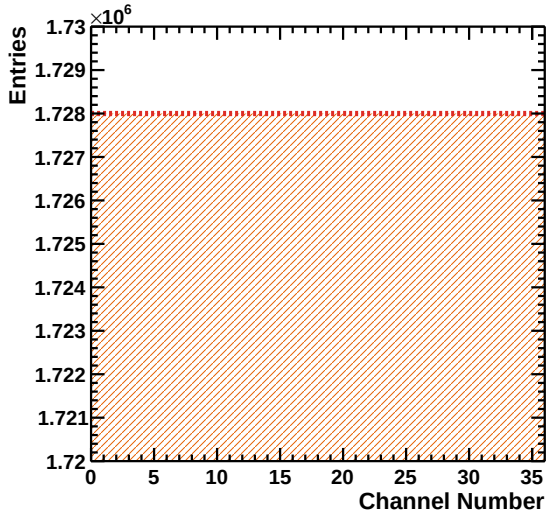
The number of events per read-out cycle, ASIC, read-out channel and ASIC memory cell for an externally triggered calibration run of 1000 read-out cycles are shown in figure 11.5. The numbers obtained from the calibration run are in exact agreement with the expectations listed above. The flat distributions over the entire range of the histograms prove the integrity of the acquired data.

The distribution of the BXIDs for the above mentioned run is shown in figure 11.5e. The dashed line histogram shows the entries from the calorimeter modules connected to the first Mini-LDA and the solid line histogram contains the entries from the calorimeter modules connected to the second Mini-LDA. The calibration runs are externally triggered, therefore, a fixed distance between the BXIDs is expected. The BXIDs of the two histograms are in agreement, demonstrating synchronous operation of the entire calorimeter system with modules distributed over two Mini-LDAs. In the case of this specific run, there were 40 ASICs on the modules which were connected to the first Mini-LDA and 68 ASICs on the layers which were connected to the second Mini-LDA. Therefore, for a calibration run of 1000 read-out cycles, $1,440,000$ and $2,4480,000$ events are expected on the first and the second Mini-LDAs, respectively. The results shown in figure 11.5e are in perfect agreement with the expectation, demonstrating that the data acquisition system preserved the integrity of the acquired data.

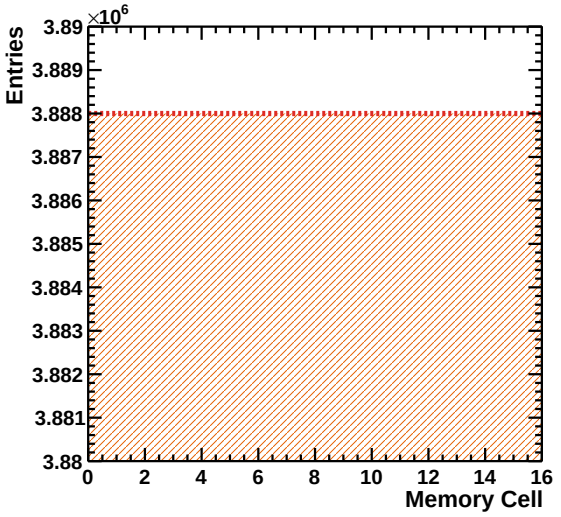


(a)

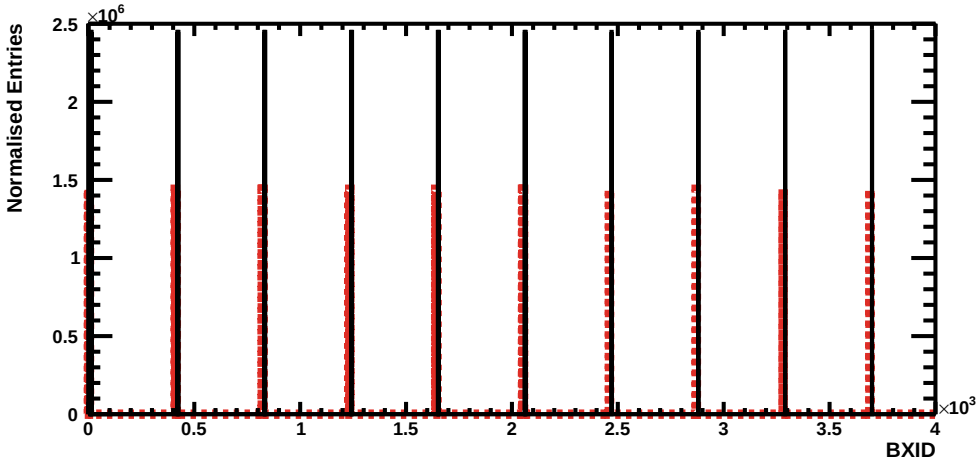
(b)



(c)

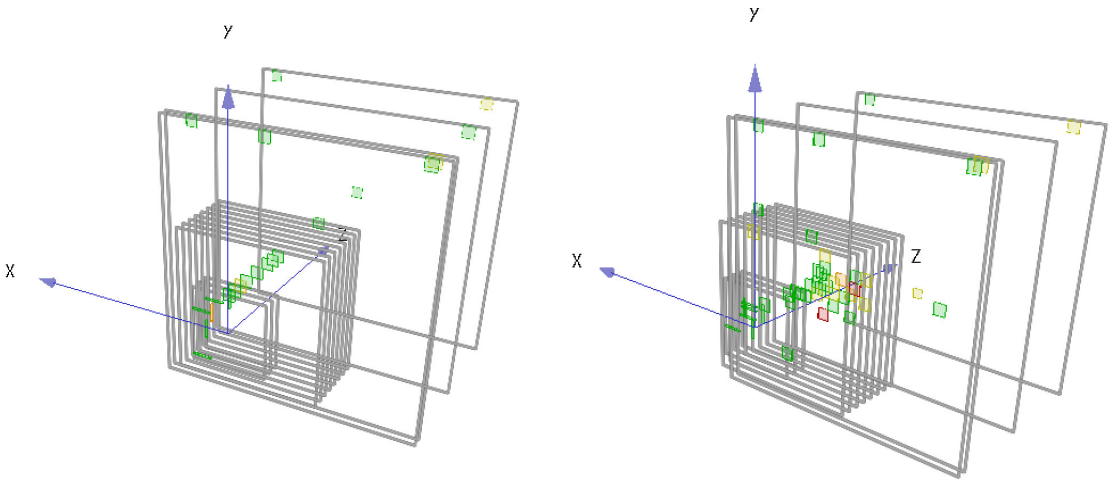


(d)



(e)

Figure 11.5 – Distributions of total number of events for a) each read-out cycle b) each read-out ASIC c) each read-out channel d) each ASIC memory cell d) BXID. The distributions are for an externally triggered calibration run of 1000 read-out cycles.



(a) (b)

Figure 11.6 – Event display of a) a muon and b) a pion in the calorimeter stack of the 2014 beam tests at the CERN PS.

After successful commissioning using calibration runs, data taking using particle beams in auto-trigger mode was carried out. More than 250 data taking runs with typically 1000 cycles each were completed using the new DAQ system. The DAQ operated stably and reliably during the entire beam test period.

The data taking was performed with electrons, muons and pions of different energies with various rates. In addition, daily calibration runs with multiple LED voltages were performed. A track of a muon striking all calorimeter layers and a pion showering in the calorimeter are shown in the event display images of figure 11.6. Such events indicate synchronous operation of the entire calorimeter in the auto-trigger mode.

The read-out rate of ~ 4 read-out cycles per second was achieved for the LED calibration runs where all 16 memory cells of all ASICs were full. The read-out rate for data taking with beam particles in the auto-trigger mode reached a rate of ~ 8 read-out cycles per second (~ 4 read-out cycles per spill of 0.4 s length).

During the entire 12 days beam test campaign the DAQ system operation was stable and reliable. No unexpected crash of the system was observed. Some instabilities in data-taking was observed but were later traced to be due to a wrong initialisation of the ASICs and not the DAQ system. After the successful operation of the scintillator DAQ system in stand-alone mode was verified, the first steps for the integration into a top-level common DAQ system were taken which is explained in the following section.

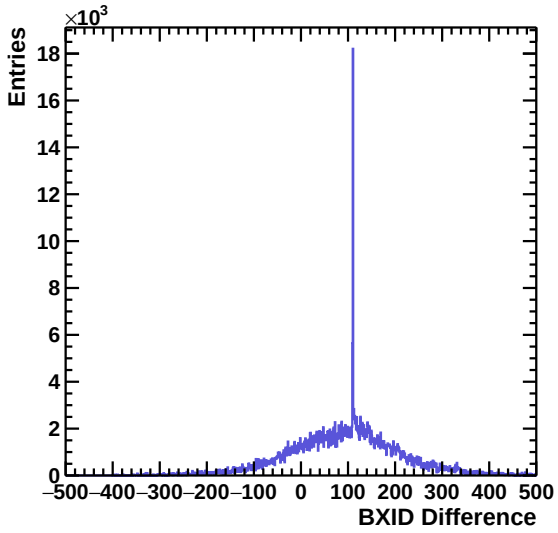


Figure 11.7 – Difference between the BXIDs of the AHCAL stack and the Si-ECAL layer.

11.2.5 Combined Run with the CALICE Si-ECAL

During the November-December 2014 beam tests at the CERN PS, common data taking of the CALICE scintillator calorimeters and a layer of the CALICE Si-ECAL was performed for the first time. The Si-ECAL layer was installed on a second platform placed between the beam pipe and the scintillator calorimeter stack. An instance of EUDAQ acted as the master of the operation while the dedicated DAQ systems of the scintillator calorimeters and the silicon layer were running in slave mode (see section 10.5 for more information about the common DAQ system).

The read-out concept of the Si-ECAL is different from the scintillator calorimeters. One of the main differences is that the Si-ECAL runs on a user-defined time-out scheme and does not provide a busy signal when the read-out ASIC is full. Also, the frequency of the main clock of the Si-ECAL is 50 MHz while the scintillator calorimeters run on a 40 MHz clock. The Si-ECAL uses a 2.5 MHz clock to generate the BXID while the scintillator calorimeters use a 250 kHz clock. Moreover, although the read-out ASIC of the Si-ECAL is capable of storing multiple events before being read-out, due to technical issues the Si-ECAL layer available for the beam test had a severe problem with re-triggering, such that one hit filled all memory cells. Consequently, only the first memory cell could be used. However, in order to collect data efficiently with the AHCAL, the decision was made to run the AHCAL as in the stand-alone setup and not to stop the data-taking after each event.

The scintillator DAQ CCC firmware was modified to provide two clocks with frequencies of 40 MHz and 50 MHz. The CCC started a data acquisition cycle and stopped the acquisition on a busy signal raised by the scintillator calorimeter.

The correlation of the BXIDs of the two system is shown in figure 11.7. Considering the factor of 10 difference between the two BXIDs, the sharp peak demonstrates synchronous operation of the three CALICE calorimeters in a combined setup operated by a common top-level data acquisition system. The different in start-up time for the Si-ECAL and the AHCAL after the start of data acquisition shifts the peak away from zero.

11.3 Current Status of the CALICE Scintillator DAQ System

The new data acquisition system for the CALICE scintillator calorimeters was successfully commissioned and tested during the beam tests at the CERN PS in November and December 2014. Although the DAQ system operation was stable and according to the expectations, further performance improvements were feasible by finalising the development of various subsystems. The progress of the DAQ system since the beam tests at the CERN PS in 2014 are summarised in this section.

The Wing-LDA was successfully commissioned and tested in March 2015 at the DESY test beam. The current version of the Wing-LDA employs one of the four Kintex FPGAs and is capable of servicing up to 24 calorimeter modules.

The DIF firmware and the DAQ software are further improved to optimise the data taking rate by eliminating or reducing the rather conservative user-defined delays. These delays were employed to reduce sources of possible failure during the first commissioning phase. A read-out rate of $25 - 30$ read-out cycles per second was achieved for an AHCAL module composed of one HBU at the DESY test beam facility [131].

The integration of the CALICE scintillator DAQ system into EUDAQ has been extended. A new LCIO data format has been implemented to reduce the data volume on the disk. Additional data such as temperature, LED voltages set for calibration runs and slow control parameters used to configure each ASIC are being stored in output LCIO files in addition to the acquired data.

The DAQ system for the CALICE scintillator calorimeters has been utilised in multiple test beam campaigns at the DESY and CERN test beam facilities since the first test beam in 2014. The calorimeter stacks have been operated reliably and efficiently using the new DAQ system. A large amount of calibration and shower data has been collected for various detector and shower physics analyses.

11.4 Summary

The new data acquisition system for the CALICE scintillator calorimeters was successfully commissioned and tested during the November-December beam test campaign at the CERN PS test-beam facility.

The DAQ system performed reliably and fulfils the general requirements listed in section 10.1.1. The DAQ system was employed for LED calibration runs and for data taking with particle beam, demonstrating that the main requirements mentioned in sections 10.1.2 and 10.1.3 are fulfilled. Although the read-out speed was satisfactory for the commissioning and first beam test, it could further increase. This has been addressed in later developments.

While the DAQ system can not be fully tested for operation at ILD, a number of the requirements listed in section 10.1.4 can be tested at the current stage. Integration into a higher-level DAQ system has been investigated in combined setups with other detectors and the development is in progress. Operation with the power-pulsing scheme has been successfully tested. The first tests of operation in a 1.5 T magnetic field has been performed. The power consumption of the DAQ system is being optimised to match the requirements of the ILC detectors.

Chapter 12

Summary and Conclusion

After the discovery of the Higgs boson at the CERN LHC experiments, investigating the properties of the newly discovered particle is of great importance. These measurements are crucial for verifying the Standard Model and also for searching for physics beyond the Standard Model. The mass of the Higgs boson is one of the important free parameters in the Standard Model. It is estimated that the LHC experiments can measure the Higgs boson mass with about 0.1% relative uncertainty by the end of the LHC operation. However, this level of uncertainty on the Higgs boson mass is not sufficient for determining the Higgs boson couplings to a precision required for verification of the models for physics beyond the Standard Model.

The International Linear Collider (ILC) is a future e^+e^- collider which provides excellent conditions for precision physics studies. The clean environment of a lepton collider combined with advanced detector technologies and software at ILC enables measurements with an unprecedented precision. At ILC, the mass of the Higgs boson can be measured using multiple methods. One of the very promising methods is to use the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel. Since there are many jets in the final state of this channel, this measurement depends strongly on jet energy measurements and its corresponding uncertainty.

The advanced calorimeters of the ILC experiments are designed for the particle flow approach to calorimetry which offers a better jet energy resolution compared to the traditional approach. In the particle flow approach, every particle within a jet is reconstructed individually. In this thesis, this powerful ability of the particle flow approach is exploited to develop a method to determine the jet energy resolution for each jet individually. This allows for obtaining an overall better jet energy resolution and also offers the possibility of selecting jets with lower uncertainty for further analysis.

In this thesis, the measurement of the Higgs boson mass from the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel is studied for two centre-of-mass energies of 350 GeV and 500 GeV. The study is performed on a GEANT4 simulation of the International Large Detector (ILD) detector.

ILD is a general-purpose detector optimised for the particle flow approach. At tree level, there are four quarks in the final state of the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel. Each of these quarks produces a jet which is then observed in the detector. The mass of the parent boson can be determined by computing the invariant mass of the jets originating from them, given that the jets are correctly associated to the bosons. But in practice there is no evident way to determine from which boson a jet originates. Two different methods are employed to determine the mass of the Higgs boson.

One method to determine the correct jet pairings is to use a χ^2-minimisation technique. In this method a χ^2 is computed for all possible jet pairings and the pairing with the minimum χ^2 is chosen as the correct pairing. The reconstructed Higgs boson masses obtained using this method for an integrated luminosity of 1000 fb^{-1} and a beam polarisation of (-0.8,$+0.3$) is 123.542 ± 0.044 and 123.375 ± 90 for the centre-of-mass energies of 350 GeV and 500 GeV, respectively. The reconstructed mass of the Higgs boson is around 1.5 GeV below the simulation input mass. By applying the same χ^2 technique on various samples in which the Z boson decays to different quark flavours, it is shown that the reason for the underestimation is related to the heavy flavour quarks. Performing a correction for the missing energy of neutrinos in semi-leptonic decays improves the results to 124.005 ± 0.048 GeV and 124.215 ± 0.093 GeV for the centre-of-mass energies of 350 GeV and 500 GeV, respectively.

At a lepton collider such as ILC, the initial state of the colliding particles is known. This allows for exploiting the conservation laws to improve the precision of the measurements. Kinematic fitting is a mathematical procedure which exploits this information and is employed in this thesis as another method for the Higgs boson mass measurement. For this purposes constraints on the Z boson mass, centre-of-mass energy and momentum are combined to form various fit hypotheses. In addition, corrections for the ISR and semi-leptonic decays are implemented. The best result for the reconstructed Higgs boson mass is obtained from a $5C$ kinematic fit with ISR and semi-leptonic corrections. In this case, for an integrated luminosity of 1000 fb^{-1} and a beam polarisation of (-0.8,$+0.3$), statistical uncertainties of 42 MeV and 89 MeV are achieved for the centre-of-mass energies of 350 GeV and 500 GeV, respectively. Compared to the results of the χ^2-minimisation method, the reconstructed Higgs boson mass obtained using kinematic fitting is approximately 1 GeV larger but is still about 400 MeV smaller than the true mass. The semi-leptonic correction which is employed in this study is a simple approach and a more sophisticated approach could possibly improve the results. In addition since the mass of the Z boson is known to a high precision, $Z \to b\bar{b}$ events can be used to calibrate the jet energy scale for $H \to b\bar{b}$ events, hence further improving the Higgs boson mass measurement from the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ process.

Multiple operating scenarios are envisioned for a 20 years physics programme at ILC. For the H-20 operating scenario of ILC where an integrated luminosity of 200 fb^{-1} at 350 GeV and 4000 fb^{-1} at 500 GeV is foreseen, the combined uncertainty on the Higgs boson mass measurement from the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel at these centre-of-mass energies is about 40 MeV. The combined uncertainty shrinks down to 26 MeV for the

I-20 scenario where an integrated luminosity of 1700 fb^{-1} at 350 GeV and 4000 fb^{-1} at 500 GeV is envisaged. These results prove that the $e^+e^- \to ZH \to q\bar{q}b\bar{b}$ channel can be exploited for a precise measurement of the Higgs boson mass which is significantly better than the estimations for the measurements by the ATLAS and CMS experiments at the end of LHC operation. This channel provides a competitive alternative to the measurement using the recoil technique at the centre-of-mass energy of 250 GeV.

The impact of the $\gamma\gamma \to hadrons$ backgrounds on the Higgs boson mass measurement is investigated using a Monte Carlo sample with $\gamma\gamma$ background overlay. A background removal procedure is applied on the sample before the Higgs boson mass reconstruction. For all reconstruction strategies, the reconstructed Higgs boson mass obtained from the sample with the $\gamma\gamma$ overlay is systematically larger than the mass measured using a similar sample but without the overlay. These results indicate that the $\gamma\gamma$ background removal procedure is limited and needs to be improved. Other sources of systematic uncertainties are identified and discussed, a number of which should be further investigated in the future.

The aforementioned measurements depend greatly on the particle flow approach and jet energy resolution. The particle flow approach requires highly granular calorimeter systems. The baseline design of ILD employs the CALICE Analog Hadronic Calorimeter (AHCAL) as its hadron calorimeter. One of the challenges in the realisation of such a calorimeter with millions of read-out channels is its Data Acquisition System (DAQ). The DAQ system of the AHCAL should fulfil the specific requirements imposed by the ILC machine parameters and by ILD. Moreover, since the AHCAL is in the research and development phase, the DAQ system should fulfil the requirements for beam tests of prototypes while being scalable to larger systems.

Based on the original CALICE DAQ design, a new DAQ system for the AHCAL and the other CALICE scintillator based calorimeter, the Scintillator Strip Electromagnetic Calorimeter (Sc-ECAL), is developed. The new DAQ system has been successfully commissioned and tested during a beam test campaign at the CERN CERN Proton Synchrotron in 2014. The detector prototype for the beam tests consisted of 12 layers of AHCAL and 3 layers of Sc-ECAL placed in a steel absorber structure. The new DAQ system was used to operate the detector prototype for externally triggered calibration and data taking with particle beams. The DAQ system performed reliably over an extended period of operation. Integration of the scintillator DAQ system into a higher-level DAQ system was also successfully tested in a combined setup with other detectors.

The development and commissioning of the new DAQ system is an important step towards a final DAQ system for the calorimeters at ILC. The system should be tested with larger detector setups in the future in order to verify the performance of the system and to identify and mitigate possible risks. Many components of the DAQ system will be placed in a ~ 4 T magnetic field in ILD. Therefore, it is necessary to fully test the DAQ system in a magnetic field.

Bibliography

[1] S. Chatrchyan et al. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Phys. Lett.* B716 (2012), pp. 30–61. doi: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[2] G. Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett.* B716 (2012), pp. 1–29. doi: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[3] S. L. Glashow. “Partial Symmetries of Weak Interactions”. In: *Nucl. Phys.* 22 (1961), pp. 579–588. doi: 10.1016/0029-5582(61)90469-2.

[4] S. Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. doi: 10.1103/PhysRevLett.19.1264.

[5] A. Salam. “Weak and Electromagnetic Interactions”. In: *Conf. Proc.* C680519 (1968), pp. 367–377.

[6] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (1964), pp. 321–323. doi: 10.1103/PhysRevLett.13.321.

[7] P. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Physical Review Letters* 13.16 (Oct. 1964), 508â509. issn: 0031-9007. doi: 10.1103/physrevlett.13.508. url: http://dx.doi.org/10.1103/PhysRevLett.13.508.

[8] P. W. Higgs. “Spontaneous Symmetry Breakdown without Massless Bosons”. In: *Phys. Rev.* 145 (1966), pp. 1156–1163. doi: 10.1103/PhysRev.145.1156.

[9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (1964), pp. 585–587. doi: 10.1103/PhysRevLett.13.585.

[10] G. Aad et al. “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments”. In: *Phys. Rev. Lett.* 114 (2015), p. 191803. doi: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex].

[11] A. Airapetian et al. *ATLAS: Detector and physics performance technical design report*. CERN-LHCC-99-14, ATLAS-TDR-14. 1999.

[12] G. L. Bayatian et al. “CMS technical design report, volume II: Physics performance”. In: *J. Phys.* G34.6 (2007), pp. 995–1579. doi: 10.1088/0954-3899/34/6/S01.

[13] F. Gianotti and M. Pepe-Altarelli. “Precision physics at the LHC”. In: *Nucl. Phys. Proc. Suppl.* 89 (2000), pp. 177–189. doi: 10.1016/S0920-5632(00)00841-0. arXiv: hep-ex/0006016 [hep-ex].

[14] B. Andersson et al. “Parton Fragmentation and String Dynamics”. In: *Phys. Rept.* 97 (1983), pp. 31–145. doi: 10.1016/0370-1573(83)90080-7.

[15] T. Sjostrand. “The Lund Monte Carlo for Jet Fragmentation and e+ e- Physics: Jetset Version 6.2”. In: *Comput. Phys. Commun.* 39 (1986), pp. 347–407. doi: 10.1016/0010-4655(86)90096-2.

[16] T. Sjostrand. “High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4”. In: *Comput. Phys. Commun.* 82 (1994), pp. 74–90. doi: 10.1016/0010-4655(94)90132-5.

[17] B. R. Webber. “A QCD Model for Jet Fragmentation Including Soft Gluon Interference”. In: *Nucl. Phys.* B238 (1984), pp. 492–528. doi: 10.1016/0550-3213(84)90333-X.

[18] G. Marchesini and B. R. Webber. “Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation”. In: *Nucl. Phys.* B310 (1988), pp. 461–526. doi: 10.1016/0550-3213(88)90089-2.

[19] K. A. Olive et al. “Review of Particle Physics”. In: *Chin. Phys.* C38 (2014), p. 090001. doi: 10.1088/1674-1137/38/9/090001.

[20] *Compact Linear Collider Website*. 2017. url: http://clic-study.web.cern.ch/.

[21] *Future Circular Collider Website*. 2017. url: https://fcc.web.cern.ch/.

[22] *Circular Electron Positron Collider Website*. url: http://cepc.ihep.ac.cn/.

[23] J. Tian and K. Fujii. “Measurement of Higgs boson couplings at the International Linear Collider”. In: *Nucl. Part. Phys. Proc.* 273-275 (2016), pp. 826–833. doi: 10.1016/j.nuclphysbps.2015.09.127.

[24] H. Baer et al. *The International Linear Collider Technical Design Report - Volume 2: Physics*. June 2013. arXiv: 1306.6352 [hep-ph].

[25] K. Fujii et al. *Physics Case for the International Linear Collider*. June 2015. arXiv: 1506.05992 [hep-ex].

[26] J. Reuter. “The BSM Physics Case of the ILC”. In: *Proceedings, International Workshop on Future Linear Colliders (LCWS15): Whistler, B.C., Canada, November 02-06, 2015*. 2016. arXiv: 1602.08352 [hep-ph]. url: http://inspirehep.net/record/1424538/files/arXiv:1602.08352.pdf.

[27] T. Behnke et al. *The International Linear Collider Technical Design Report - Volume 1: Executive Summary*. June 2013. arXiv: 1306.6327 [physics.acc-ph].

[28] J. Brau, Y. Okada, and N. Walker. *ILC Reference Design Report Volume 1 - Executive Summary*. Dec. 2007. arXiv: 0712.1950 [physics.acc-ph].

[29] A. Djouadi et al. *International Linear Collider Reference Design Report Volume 2: PHYSICS AT THE ILC*. Sept. 2007. arXiv: 0709.1893 [hep-ph].

[30] N. Phinney, N. Toge, and N. Walker. *ILC Reference Design Report Volume 3 - Accelerator*. Dec. 2007. arXiv: 0712.2361 [physics.acc-ph].

[31] T. Behnke et al. *ILC Reference Design Report Volume 4 - Detectors*. Dec. 2007. arXiv: 0712.2356 [physics.ins-det].

[32] C. Adolphsen et al. *The International Linear Collider Technical Design Report - Volume 3.I: Accelerator R&D in the Technical Design Phase*. June 2013. arXiv: 1306.6353 [physics.acc-ph].

[33] C. Adolphsen et al. *The International Linear Collider Technical Design Report - Volume 3.II: Accelerator Baseline Design*. June 2013. arXiv: 1306.6328 [physics.acc-ph].

[34] T. Behnke et al. *The International Linear Collider Technical Design Report - Volume 4: Detectors*. June 2013. arXiv: 1306.6329 [physics.ins-det].

[35] S. Boogert et al. "Polarimeters and Energy Spectrometers for the ILC Beam Delivery System". In: *JINST* 4 (2009), P10015. doi: 10.1088/1748-0221/4/10/P10015. arXiv: 0904.0122 [physics.ins-det].

[36] M. Beckmann et al. "Spin Transport and Polarimetry in the Beam Delivery System of the International Linear Collider". In: *JINST* 9 (2014), P07003. doi: 10.1088/1748-0221/9/07/P07003. arXiv: 1405.2156 [physics.acc-ph].

[37] J. E. Augustin et al. "Limitations on Performance of e+ e- Storage Rings and Linear Colliding Beam Systems at High Energy". In: *eConf* C781015 (1978), p. 009.

[38] B. Parker and A. Seryi. *Novel method of compensation of the effects of detector solenoid on the vertical beam orbit in a linear colliders*. Tech. rep. SLAC-PUB-11038. 2005.

[39] P. Chen, T. L. Barklow, and M. E. Peskin. "Hadron production in gamma-gamma collisions as a background for e+e- linear colliders". In: *Phys.Rev. D* 49 (1994), pp. 3209–3227. doi: 10.1103/PhysRevD.49.3209. arXiv: hep-ph/9305247.

[40] M. A. Thomson. "Particle Flow Calorimetry and the PandoraPFA Algorithm". In: *Nucl. Instrum. Meth.* A611 (2009), pp. 25–40. doi: 10.1016/j.nima.2009.09.009. arXiv: 0907.3577 [physics.ins-det].

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

[42] H. Aihara, P. Burrows, and M. Oreglia. *SiD Letter of Intent*. 2010. arXiv: 0911.0006 [physics.ins-det].

[43] A. Vogel. *The coordinate system for LDC detector studies*. Tech. rep. LC-DET-2005-009. 2005.

[44] C. Hu-Guo et al. "First reticule size MAPS with digital output and integrated zero suppression for the EUDET-JRA1 beam telescope". In: *Nucl. Instrum. Meth.* A623 (2010), pp. 480–482. doi: 10.1016/j.nima.2010.03.043.

[45] I. Valin et al. "A reticle size CMOS pixel sensor dedicated to the STAR HFT". In: *JINST* 7 (2012), p. C01102. doi: 10.1088/1748-0221/7/01/C01102.

[46] L. Greiner et al. “A MAPS based vertex detector for the STAR experiment at RHIC”. In: *Nucl. Instrum. Meth.* A650 (2011), pp. 68–72. doi: 10.1016/j.nima.2010.12.006.

[47] Y. Sugimoto. “Fine Pixel CCD Option for the ILC Vertex Detector”. In: *eConf* C050318 (2005), p. 0804.

[48] Y. Sugimoto et al. “R and D Status and Plan for FPCCD VTX”. In: *International Linear Collider Workshop (LCWS10 and ILC10) Beijing, China, March 26-30, 2010*. 2010. arXiv: 1007.2471 [physics.ins-det]. url: http://inspirehep.net/record/861406/files/arXiv:1007.2471.pdf.

[49] Y. Sugimoto et al. “Status of FPCCD Vertex Detector for ILD”. In: *International Workshop on Future Linear Colliders (LCWS11) Granada, Spain, September 26-30, 2011*. 2012. arXiv: 1202.5832 [physics.ins-det]. url: http://inspirehep.net/record/1090859/files/arXiv:1202.5832.pdf.

[50] O. Alonso et al. “DEPFET active pixel detectors for a future linear e^+e^- collider”. In: *IEEE Trans. Nucl. Sci.* 60 (2013), p. 1457. doi: 10.1109/TNS.2013.2245680. arXiv: 1212.2160 [physics.ins-det].

[51] J. J. Velthuis et al. “DEPFET, a monolithic active pixel sensor for the ILC”. In: *Nucl. Instrum. Meth.* A579 (2007), pp. 685–689. doi: 10.1016/j.nima.2007.05.278.

[52] F. Sauli. “GEM: A new concept for electron amplification in gas detectors”. In: *Nucl. Instrum. Meth.* A386 (1997), pp. 531–534. doi: 10.1016/S0168-9002(96)01172-2.

[53] Y. Giomataris et al. “MICROMEGAS: A High granularity position sensitive gaseous detector for high particle flux environments”. In: *Nucl. Instrum. Meth.* A376 (1996), pp. 29–35. doi: 10.1016/0168-9002(96)00175-1.

[54] M. Kobayashi et al. “Cosmic ray tests of a GEM-based TPC prototype operated in Ar-CF4-isobutane gas mixtures”. In: *Nucl. Instrum. Meth.* A641 (2011). [Erratum: Nucl. Instrum. Meth.A697,122(2013)], pp. 37–47. doi: 10.1016/j.nima.2011.02.042,10.1016/j.nima.2012.09.016. arXiv: 1008.5068 [physics.ins-det].

[55] V. Balagura et al. “Study of scintillator strip with wavelength shifting fiber and silicon photomultiplier”. In: *Nucl. Instrum. Meth.* A564 (2006), pp. 590–596. doi: 10.1016/j.nima.2006.04.030. arXiv: physics/0504194 [physics].

[56] J. Repond. *Overview of the DHCAL Project*. 2010. arXiv: 1005.0412 [physics.ins-det].

[57] G. Moortgat-Pick et al. *Physics at the e+ e- Linear Collider*. Apr. 2015. doi: 10.1140/epjc/s10052-015-3511-9. arXiv: 1504.01726 [hep-ph].

[58] J. Yan et al. “Measurement of the Higgs boson mass and $e^+e^- \to ZH$ cross section using $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ at the ILC”. In: *Phys. Rev.* D94.11 (2016), p. 113002. doi: 10.1103/PhysRevD.94.113002. arXiv: 1604.07524 [hep-ex].

[59] J. Tian. “A new method for Higgs mass measurement”. Linear Collider Workshop 2016. 2016.

[60] P. Garcia-Abia, W. Lohmann, and A. Raspereza. *Measurement of the Higgs Boson Mass with a Linear e+e- Collider*. 2007. arXiv: hep-ex/0505096.

[61] J. Pochodzalla et al. “Probing the nuclear liquid - gas phase transition”. In: *Phys. Rev. Lett.* 75 (1995), pp. 1040–1043. doi: 10.1103/PhysRevLett.75.1040.

[62] A. Ferrari et al. “Cascade particles, nuclear evaporation, and residual nuclei in high-energy hadron - nucleus interactions”. In: *Z. Phys.* C70 (1996), pp. 413–426. doi: 10.1007/s002880050119. arXiv: nucl-th/9509039 [nucl-th].

[63] C. W. Fabjan and F. Gianotti. “Calorimetry for particle physics”. In: *Rev. Mod. Phys.* 75 (2003), pp. 1243–1286. doi: 10.1103/RevModPhys.75.1243.

[64] C. Collaboration. *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*. Tech. rep. CMS-PAS-PFT-09-001. CERN, 2009.

[65] J. S. Marshall and M. A. Thomson. “The Pandora Software Development Kit for Pattern Recognition”. In: *Eur. Phys. J.* C75.9 (2015), p. 439. doi: 10.1140/epjc/s10052-015-3659-3. arXiv: 1506.05348 [physics.data-an].

[66] *iLCSoft project website*. 2016. url: http://ilcsoft.desy.de.

[67] F. Gaede et al. “LCIO: A Persistency framework for linear collider simulation studies”. In: *eConf* C0303241 (2003), TUKT001. arXiv: physics/0306114 [physics].

[68] F. Gaede. “Marlin and LCCD: Software tools for the ILC”. In: *Nucl. Instrum. Meth.* A559 (2006), pp. 177–180. doi: 10.1016/j.nima.2005.11.138.

[69] W. Kilian, T. Ohl, and J. Reuter. “WHIZARD: Simulating Multi-Particle Processes at LHC and ILC”. In: *Eur.Phys.J.C* 71:1742,2011 (Aug. 2007). doi: 10.1140/epjc/s10052-011-1742-y. arXiv: 0708.4233 [hep-ph].

[70] M. Moretti, T. Ohl, and J. Reuter. *O'Mega: An Optimizing Matrix Element Generator*. 2007. arXiv: hep-ph/0102195.

[71] D. Schulte. “Study of Electromagnetic and Hadronic Background in the Interaction Region of the TESLA Collider”. PhD thesis. Hamburg University, 1997. url: http://inspirehep.net/record/888433/files/shulte.pdf.

[72] T. Sjostrand, S. Mrenna, and P. Skands. “PYTHIA 6.4 Physics and Manual”. In: *JHEP* 0605:026,2006 (JHEP 0605:026,2006). doi: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175.

[73] S. Agostinelli et al. “GEANT4: A Simulation toolkit”. In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. doi: 10.1016/S0168-9002(03)01368-8.

[74] J. Allison et al. “Recent developments in G$_{EANT}$4”. In: *Nucl. Instrum. Meth.* A835 (2016), pp. 186–225. doi: 10.1016/j.nima.2016.06.125.

[75] P. Mora de Freitas and H. Videau. “Detector simulation with MOKKA / GEANT4: Present and future”. In: *Linear colliders. Proceedings, International Workshop on physics and experiments with future electron-positron linear colliders, LCWS 2002, Seogwipo, Jeju Island, Korea, August 26-30, 2002*. 2002, pp. 623–627. url: http://www-library.desy.de/cgi-bin/showprep.pl?lc-tool03-010.

[76] R. Fruhwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nucl. Instrum. Meth.* A262 (1987), pp. 444–450. doi: 10.1016/0168-9002(87)90887-4.

[77] M. Cacciari, G. P. Salam, and G. Soyez. *FastJet user manual*. Nov. 2011. doi: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[78] S. Catani et al. "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions". In: *Nucl. Phys.* B406 (1993), pp. 187–224. doi: 10.1016/0550-3213(93)90166-M.

[79] S. Catani, Y. L. Dokshitzer, and B. R. Webber. "The K^- perpendicular clustering algorithm for jets in deep inelastic scattering and hadron collisions". In: *Phys. Lett.* B285 (1992), pp. 291–299. doi: 10.1016/0370-2693(92)91467-N.

[80] F. J. Müller. "Development of a Triple GEM Readout Module for a Time Projection Chamber & Measurement Accuracies of Hadronic Higgs Branching Fractions in $\nu\nu$H at a 350 GeV ILC". PhD thesis. Hamburg: Hamburg U., 2016. doi: 10.3204/PUBDB-2016-02659. url: http://bib-pubdb1.desy.de/search?cc=Publication+Database&of=hd&p=reportnumber:DESY-THESIS-2016-018.

[81] S. Catani et al. "New clustering algorithm for multi - jet cross-sections in e+ e- annihilation". In: *Phys. Lett.* B269 (1991), pp. 432–438. doi: 10.1016/0370-2693(91)90196-W.

[82] R. L. Gluckstern. "Uncertainties in track momentum and direction, due to multiple scattering and measurement errors". In: *Nucl. Instrum. Meth.* 24 (1963), pp. 381–389. doi: 10.1016/0029-554X(63)90347-1.

[83] C. Adloff et al. "Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques". In: *JINST* 7 (2012), P09017. doi: 10.1088/1748-0221/7/09/P09017. arXiv: 1207.4210 [physics.ins-det].

[84] C. Adloff et al. "Response of the CALICE Si-W electromagnetic calorimeter physics prototype to electrons". In: *Nucl. Instrum. Meth.* A608 (2009), pp. 372–383. doi: 10.1016/j.nima.2009.07.026.

[85] M. Kurata. "High Level Reconstruction Tools". Asian Linear Collider Workshop 2015 (ALCWS2015). Tsukuba, Japan, 2015.

[86] M. Kurata. "Partilce ID Study and its Application". Asian Linear Collider Workshop 2015 (ALCWS2015). Tsukuba, Japan, 2015.

[87] S. Lukic. "PID Tools". ILD Software and Optimization Workshop. Hamburg, Germany, 2016.

[88] B. Hermberg. "Measuring Hadronic Jets at the ILC - From Particle Flow Calorimetry to the Higgs Self-Coupling". PhD thesis. Hamburg: U. Hamburg, Dept. Phys., 2015. url: http://bib-pubdb1.desy.de/search?cc=Publication+Database&of=hd&p=reportnumber:DESY-THESIS-2015-039.

[89] A. G. Frodesen, O. Skjeggestad, and H. Tofte. *Probability and Statistics in Particle Physics*. Bergen, Norway: Universitetsforlaget, 1979. url: http://www-spires.fnal.gov/spires/find/books/www?cl=QC174.8.F928.

[90] B. List and J. List. “MarlinKinfit: An Object-Oriented Kinematic Fitting Package”. 2009.

[91] M. Beckmann, B. List, and J. List. “Treatment of Photon Radiation in Kinematic Fits at Future e+ e- Colliders”. In: *Nucl. Instrum. Meth.* A624 (2010), pp. 184–191. doi: 10.1016/j.nima.2010.08.107. arXiv: 1006.0436 [hep-ex].

[92] F. A. Berends, R. Kleiss, and S. Jadach. “Radiative Corrections to Muon Pair and Quark Pair Production in electron-Positron Collisions in the Z(0) Region”. In: *Nucl. Phys.* B202 (1982), pp. 63–88. doi: 10.1016/0550-3213(82)90221-8.

[93] C. F. Dürig. “Measuring the Higgs Self-coupling at the International Linear Collider”. PhD thesis. Hamburg: Hamburg U., 2016. url: http://bib-pubdb1.desy.de/search?cc=Publication+Database&of=hd&p=reportnumber:DESY-THESIS-2016-027.

[94] T. Suehara and T. Tanabe. “LCFIPlus: A Framework for Jet Analysis in Linear Collider Studies”. In: *Nucl. Instrum. Meth.* A808 (2016), pp. 109–116. doi: 10.1016/j.nima.2015.11.054. arXiv: 1506.08371 [physics.ins-det].

[95] A. Straessner. “Measurement of the W boson mass at LEP”. In: *Proceedings, 39th Rencontres de Moriond, 04 Electroweak interactions and unified theories: La Thuile, Aosta, Italy, Mar 21-28, 2004*. 2004, pp. 3–10. arXiv: hep-ex/0405005 [hep-ex]. url: http://inspirehep.net/record/649688/files/Pages_from_C04-03-21_3.pdf.

[96] T. Barklow et al. *ILC Operating Scenarios*. June 2015. arXiv: 1506.07830 [hep-ex].

[97] *CALICE Collaboration Website*. 2016. url: https://twiki.cern.ch/twiki/bin/view/CALICE/WebHome.

[98] J. Repond et al. “Design and Electronics Commissioning of the Physics Prototype of a Si-W Electromagnetic Calorimeter for the International Linear Collider”. In: *JINST* 3 (2008), P08001. doi: 10.1088/1748-0221/3/08/P08001. arXiv: 0805.4833 [physics.ins-det].

[99] K. Francis et al. “Performance of the first prototype of the CALICE scintillator strip electromagnetic calorimeter”. In: *Nucl. Instrum. Meth.* A763 (2014), pp. 278–289. doi: 10.1016/j.nima.2014.06.039. arXiv: 1311.3761 [physics.ins-det].

[100] C. Adams et al. “Design, construction and commissioning of the Digital Hadron Calorimeter—DHCAL”. In: *JINST* 11.07 (2016), P07007. doi: 10.1088/1748-0221/11/07/P07007. arXiv: 1603.01653 [physics.ins-det].

[101] B. Bilki. *DHCAL Response to Positrons and Pions*. CALICE Analysis Note CAN-032. CALICE Collaboration, 2011.

[102] G. Baulieu et al. “Construction and commissioning of a technological prototype of a high-granularity semi-digital hadronic calorimeter”. In: *JINST* 10.10 (2015), P10039. doi: 10.1088/1748-0221/10/10/P10039. arXiv: 1506.05316 [physics.ins-det].

[103] V. Buridon et al. “First results of the CALICE SDHCAL technological prototype”. In: *JINST* 11.04 (2016), P04001. doi: 10.1088/1748-0221/11/04/P04001. arXiv: 1602.02276 [physics.ins-det].

[104] C. Adloff et al. “Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype”. In: *JINST* 5 (2010), P05004. doi: 10.1088/1748-0221/5/05/P05004. arXiv: 1003.2662 [physics.ins-det].

[105] C. Adloff et al. “Shower development of particles with momenta from 1 to 10 GeV in the CALICE Scintillator-Tungsten HCAL”. In: *JINST* 9.01 (2014), P01004. doi: 10.1088/1748-0221/9/01/P01004. arXiv: 1311.3505 [physics.ins-det].

[106] M. Chefdeville et al. “Shower development of particles with momenta from 15 GeV to 150 GeV in the CALICE scintillator-tungsten hadronic calorimeter”. In: *JINST* 10.12 (2015), P12006. doi: 10.1088/1748-0221/10/12/P12006. arXiv: 1509.00617 [physics.ins-det].

[107] F. Simon and C. Soldner. “Uniformity Studies of Scintillator Tiles directly coupled to SiPMs for Imaging Calorimetry”. In: *Nucl. Instrum. Meth.* A620 (2010), pp. 196–201. doi: 10.1016/j.nima.2010.03.142. arXiv: 1001.4665 [physics.ins-det].

[108] E. Garutti, S. Laurien, and M. Ramilli. “A novel design for a highly granular hadron calorimeter scintillator-SiPM tile”. In: *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. Nov. 2014, pp. 1–5. doi: 10.1109/NSSMIC.2014.7431159.

[109] G. Blazey et al. “Directly Coupled Tiles as Elements of a Scintillator Calorimeter with MPPC Readout”. In: *Nucl. Instrum. Meth.* A605 (2009), pp. 277–281. doi: 10.1016/j.nima.2009.03.253.

[110] Y. Liu et al. *A Design of Scintillator Tiles Read Out by Surface-Mounted SiPMs for a Future Hadron Calorimeter*. Dec. 2015. arXiv: 1512.05900 [physics.ins-det].

[111] OMEGA. *SPIROC 2 User Guide*. 2009.

[112] *OMEGA Website*. 2017. url: http://omega.in2p3.fr/.

[113] S. Conforti Di Lorenzo et al. “SPIROC: Design and performances of a dedicated very front-end electronics for an ILC Analog Hadronic CALorimeter (AHCAL) prototype with SiPM read-out”. In: *JINST* 8 (2013), p. C01027. doi: 10.1088/1748-0221/8/01/C01027.

[114] M. Terwort. “Realization and Test of the Engineering Prototype of the CALICE Tile Hadron Calorimeter”. In: *Proceedings, 2010 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC 2010): Knoxville, Tennessee, October 30-November 6, 2010*. 2010, pp. 838–842. doi: 10.1109/NSSMIC.2010.5873879. arXiv: 1011.4760 [physics.ins-det]. url: http://inspirehep.net/record/878337/files/arXiv:1011.4760.pdf.

[115] K. Gadow. private communication. 2017.

[116] M. Wing et al. *A proposed DAQ system for a calorimeter at the International Linear Collider*. Tech. rep. 2007. arXiv: physics/0611299.

[117] B. Hommels. “Data acquisition systems for future calorimetry at the International Linear collider”. In: *Proceedings, 2008 IEEE Nuclear Science Symposium, Medical Imaging Conference and 16th International Workshop on Room-Temperature Semiconductor X-Ray and Gamma-Ray Detectors (NSS/MIC 2008 / RTSD 2008): Dresden, Germany, October 19-25, 2008*. 2008, pp. 3542–3545. doi: 10.1109/NSSMIC.2008.4775100.

[118] *High-Definition Multimedia Interface Specification Version 1.4*. HDMI Licensing, LLC. 2009.

[119] T. A. D. Developers. *AHCAL DAQ - Operating Manual*. Version 1.4. 2014.

[120] M. Kelly. *Calice LDA Docs*. Tech. rep. Version 1.2. Manchester, 2008. url: http://www.hep.manchester.ac.uk/u/mpkelly/calice/lda/Calice_LDA_Overview.pdf.

[121] *Avnet ZedBoard Brochure English Version*. Avnet. 2012. url: http://zedboard.com/sites/default/files/Avnet%20ZedBoard%20Brochure%20English%20Version.pdf.

[122] *Zynq-7000 All Programmable SoC Overview*. DS190. V1.10. Xilinx. 2016.

[123] *Zynq-7000 All Programmable SoC Technical Reference Manual*. UG585. ver1.11. Xilinx. 2016.

[124] *ZedBoard Product Page*. 2016. url: http://zedboard.org/product/zedboard.

[125] A. Welker. “Auslese Kette für hochauflösende hadronische Kalorimeter mit Siliziumphotomultipliern”. MA thesis. Johannes Gutenberg-Universität Mainz, 2013.

[126] *AXI DMA Product Guide*. PG021. Version 7.1. Xilinx. 2016.

[127] *Mars ZX3 Product Brief*. Tech. rep. Enclustra, 2016.

[128] *Using Xilinx SDK*. UG782. ver2016.3. Xilinx. 2016.

[129] R. Spreckels. *Processing Instruction Configurator (piconf)*. url: https://sourceforge.net/projects/piconf/.

[130] S. Spannagel. “Test Beam Measurements for the Upgrade of the CMS Pixel Detector and Measurement of the Top Quark Mass from Differential Cross Sections”. PhD thesis. Hamburg: U. Hamburg, Dept. Phys., 2016. url: http://bib-pubdb1.desy.de/search?cc=Publication+Database&of=hd&p=reportnumber:DESY-THESIS-2016-010.

[131] J. Kvasnicka. “Data Acquisition System for the CALICE AHCAL Calorimeter”. In: *JINST* 12.03 (2017), p. C03043. doi: 10.1088/1748-0221/12/03/C03043. arXiv: 1701.02232 [physics.ins-det].

List of Acronyms

ADC Analog-to-digital Converter. 119, 145, 153

AHCAL Analog Hadronic Calorimeter. xiii, xvi, xvii, 2, 25, 54, 67, 106, 107, 109, 110, 112--118, 120, 121, 123, 125, 126, 129, 141, 151, 152, 154, 155, 160, 161, 165

APD Avalance Photo-Diode. 116

ARM Advanced RISC Machine. 136, 146

ASCII American Standard Code for Information Interchange. 145

ASIC Application-specific Integrated Circuit. xvi, 108, 117--122, 125, 127, 131, 133--135, 139, 153, 157--161

ATLAS A Toroidal LHC ApparatuS. 1, 7, 10, 11, 106, 165

AXI Advanced eXtensible Interface. 141, 146

BDS Beam Delivery System. 23

BEH Brout-Englert-Higgs. 1, 5, 7, 31

BTC Block Transfer Command. 133

BXID Bunch Crossing Identifier. xvii, 119, 157, 158, 160, 161

CALICE Calorimetry for an ILC detector. v, vii, xiv, xvi, xvii, 2, 3, 67, 107, 109--114, 125, 126, 128--135, 137, 138, 140, 141, 143--145, 147--149, 151, 152, 154, 157, 160--162, 165, 183

CCC Clock and Control Card. xiii, xvii, 130--132, 135--139, 141, 142, 144--146, 153, 156, 160

CCD Charge-coupled Device. 27, 178

CEPC Circular Electron-Positron Collider. 17

CERN Organisation Européenne pour la Recherche Nucléaire. v, vii, 1, 7, 17, 110--113, 120, 149, 161--163, 165

CERN PS CERN Proton Synchrotron. v, vii, xvii, 3, 151--157, 159--161, 165

CIB Central Interface Board. xvi, 121, 122, 134, 135, 152

CLIC Compact Linear Collider. 17, 35, 114

CMOS Complementary Metal–Oxide–Semiconductor. 27, 178

CMS Compact Muon Solenoid. xv, 1, 7, 10, 11, 46, 106, 165

CP Charge-Parity. 9, 18

CPS CMOS Pixel Sensors. 27

CRC Cyclic Redundancy Check. 133

DAC Digital-to-Analog Converter. 119

DAQ Data Acquisition System. v, xiv, xvii, 3, 54, 119, 125--133, 135, 138, 139, 141, 143--145, 147--149, 151--157, 159--162, 165

DDR SDRAM Double Data Rate Synchronous Dynamic Random-access Memory. 136, 142

DEPFET Depleted Field Effect Transistor. 27

DESY Deutsches Elektronen-Synchrotron. 110, 120, 161, 183

DHCAL Digital Hadronic Calorimeter. xiii, 25, 111

DIF Detector Interface. xiii, xvi, xvii, 121, 122, 130, 131, 133--135, 138, 141, 153, 161

DMA Direct Memory Access. 141

EBU ECal Base Unit. 154

ECal electromagnetic calorimeter. 29, 42, 45, 54, 66--68, 108

EDM Event Data Model. 50

EWSB Electro-weak Symmetry-breaking. 9

FC Fast Command. 132

FCC Future Circular Collider. 17

FMC FPGA Mezzanine Card. 137, 139, 141, 142

FPCCD Fine Pixel CCD. 27

FPGA Field-programmable Gate Array. 135--137, 142, 161

GAPD Geiger-mode Avalance Photo-Diode. 116

GEM Gas Electron Multiplier. 28

GPIO General Purpose I/O. 146

GRPC Glass Resistive Plate Chamber. 111

GUI Graphical User Interface. 144

HBU HCal Base Unit. xvi, 121, 122, 151--155, 161

HCal hadronic calorimeter. 29, 42, 45, 46, 54, 66--68, 108

HDL Hardware Description Language. 137

HDMI High-Definition Multimedia Interface. 132, 135, 137, 138, 141, 142, 153, 156

I/O Input/Output. 142, 146, 179

ICFA International Committee for Future Accelerators. 19

ILC International Linear Collider. i, v, vii, ix, xiii, xv, xix, 1, 2, 5, 15, 17--27, 30, 32, 33, 35, 37, 45, 47, 48, 50--52, 54--56, 79, 80, 87, 89--91, 104--108, 110, 114, 118, 120, 125, 129, 162--165

iLCSoft ILC Sofware Package. 49, 50, 53--55, 86, 145

ILD International Large Detector. v, vii, ix, xi, xv, xvi, xix, 2, 18, 25--30, 35, 48--50, 53--55, 59, 68, 69, 71, 72, 75--77, 79, 89, 91, 104, 108--111, 113--116, 121, 123, 126, 129, 133, 138, 139, 141, 142, 147, 154, 162--165

IP Interaction Point. 23, 25

ISR Initial State Radiation. xiii, xvi, 33, 34, 51, 87, 90--96, 98--100, 102, 103, 105, 164

JER Jet Energy Resolution. 62

LabVIEW Laboratory Virtual Instrument Engineering Workbench. 144, 145, 153

LAN Local Area Network. 153, 156

LCIO Linear Collider I/O. 50, 57, 145, 148, 161

LED Light-emitting diode. xiii, 121, 128, 137, 148, 149, 152, 159, 161, 162

LEP Large Electron-Positron Collider. 17, 45, 51, 105, 180

LHC Large Hadron Collider. v, vii, ix, 1, 7, 11, 15, 17, 18, 31, 32, 35, 46, 106, 125, 163, 165

Marlin Modular Analysis and Reconstruction for the Linear Collider. 50, 54

MC Monte Carlo. xv, 2, 35, 49, 50, 60--63, 65, 66, 73, 79, 80, 82, 83, 85, 87, 92, 98--101, 103, 105, 165

MICROMEGAS Micro-MEsh GAseous Structure. 28, 112

Mini-LDA Mini-Link and Data Aggregator. xvii, 130, 136, 137, 139--142, 153, 156, 157

MIP Minimum Ionising Partilce. 72

NIC Network Interface Card. 156

NIM Nuclear Instrumentation Module. 156

O'Mega Optimized Matrix Element Generator. 50

OPAL Omni-Purpose Apparatus at LEP. 51, 86

PandoraPFA Pandora Particle Flow Algorithm. xv, 47, 48, 55, 67--69, 71, 75--77, 84

PCB Printed Circuit Board. 121

PFA particle flow algorithm. 24, 25, 46, 61, 62, 64, 68, 69, 71, 74, 104

PFO Particle Flow Object. 55

PL Programmable Logic. 136--138, 141, 146, 147

PMT Photomultiplier Tube. 116, 117, 155

PS Processing System. 136--138, 141, 146, 149, 154, 162

QCD Quantum Chromodynamics. xi, 6, 7, 11--13, 18, 59, 64, 65

QED Quantum Electroynamics. 6--8, 11, 12

QFT Quantum Field Theory. 6

RDR Reference Design Report. 19

RF Radio Frequency. 22

RMS Root Mean Square. 91

RPC Resistive Plate Chamber. 30, 111, 112

RTML Ring to Main Linac. 22

Sc-ECAL Scintillator Strip Electromagnetic Calorimeter. xiii, 54, 110, 111, 125, 126, 151, 154, 165

SCA Switched Capacitor Array. 119

SCRF Super-conducting Radio Frequency. 22

SD Secure Digital. 137, 138

SDHCAL Semi-digital Hadronic Calorimeter. xiii, 54, 111, 112

SET Silicon External Tracker. 28

Si-ECAL Silicon-tungsten Electromagnetic Calorimeter. xiii, xiv, xvii, 25, 54, 67, 109, 110, 113, 151, 154, 160, 161

SiD Silicon Detector. 25

SiPM Silicon Photomultiplier. 30, 110, 112, 113, 117--119, 121, 125, 155

SIT Silicon Inner Tracker. 28

SoC System-on-a-Chip. xvii, 135--139, 141, 142, 146

SPI Serial Peripheral Interface. 137, 142

SPIROC Silicon Photomultiplier Read-out Chip. xvi, 118--121, 125, 153

SPS Single Pixel Spectrum. 112, 121

SRAM Static Random-Access Memory. 119

SUSY Supersymmetry. 14, 18

TDC Time-to-digital Converter. 119, 145, 153

TDR Technical Design Report. 19, 50, 54

TPC Time-Projection Chamber. 25--28

UART Universal Asynchronous Receiver/Transmitter. 132, 141

USB Universal Serial Bus. 153

VMEbus Versa Module Europa bus. 138

Wing-LDA Wing-Link and Data Aggregator. xvii, 130, 139, 141--143, 156, 161

xLDA x-Link and Data Aggregator. xiii, 130--135, 137--139, 141, 145--147, 149

Acknowledgements

I would like to thank a few people without their help and support completing this thesis would not have been possible.

First of all, I would like to thank Erika Garutti for giving me the great opportunity to write my Ph.D. thesis under her supervision. Furthermore, I am grateful to Ties Behnke for giving me the possibility to do my Ph.D. as a member of the DESY FLC group.

I would like to thank Katja Krüger for her supervision and kind support and advice throughout my work and Jenny List for the supervision of the analysis part of my thesis. I am very grateful for all I have learned from them.

Many people have contributed to the development of the data acquisition system for the CALICE scintillator calorimeters. I am thankful for what I have learned form all of them, in particular Mathias Reinecke and Jiři Kvasnička.

I would like to thank the members of the FLC group for all their help and great time we had together, in particular Oskar Hartbrich, Benjamin Hermberg, Eldwan Brianne, Ambra Provenza, Dimitra Tsionou and Adrián Irles.

I am very grateful to my entire family, especially my loving parents and brothers for their endless support throughout my life. Without all your support and confidence I would have not been able to achieve my goals.

The research leading to these results has received funding from the European Union Seventh Framework Programme [FP7/2007-2013] under Grant Agreement no. 256984, and is supported by a Marie Curie Early Initial Training Network Fellowship of the European Community's Seventh Framework Programme under contract number (PITN-GA-2011-289355-PicoSEC-MCNet).

Hamburg, July 2017

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hamburg, den ...../...../..... Unterschrift .....