Detector and Trigger Studies Towards Discovering the Higgs Boson Produced via Vector Boson Fusion Using the CMS Detector

Anastasios Papageorgiou

High Energy Physics Department
The Blackett Laboratory
Imperial College
London

A thesis submitted for the degree of
Doctor of Philosophy
of the University of London
and the Diploma of Imperial College.

October 2010
Declaration

The work presented in this thesis was conducted by the author and any contributions from other members of the collaboration are appropriately referenced.

Specifically for the work included in this thesis, the author was responsible for and performed the following tasks:

- Developed the analysis and the selection criteria for the extraction of single pion responses from the H2 2006 combined test beam data. Also developed the simulation to provide reconstructed calorimeter objects including all the effects of the complete official CMS simulation and reconstruction software and produced the data used for the comparison analysis. The author showed that the data collected from the test beam and the simulated data were not in very good agreement and the simulation needed to be improved.

- Implemented and optimised a new L1 and HLT trigger, VBF jets plus lepton, for the $2 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ instantaneous luminosity regime. The author demonstrated a large improvement in the selection of VBF $H \rightarrow \tau\tau \rightarrow l + \tau$-jet preselected events for no significant increase in the trigger rate.

- Evaluated the performance of the newly developed trigger by using it for signal and background event selections. In the process, the author had to reprocess all the officially produced data samples to enable the newly implemented trigger to execute. Moreover, the author carried out the development and optimisation of selection criteria aiming to reduce the large backgrounds to a level where the Higgs boson could be detected. The study showed that the newly implemented trigger can play an important role in the discovery of a low mass Higgs boson in high luminosity regimes.
Abstract

Detection of the Higgs boson is one of the primary goals of the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC). To achieve this goal a good understanding of both the detector and its simulation is required. In this thesis the response of the CMS calorimeters to single pions extracted from test beam data will be presented, as well as a comparison with the response from simulated data. Also, a new trigger is proposed, aiming to increase the selection efficiency of the Standard Model Higgs boson, produced via the Vector Boson Fusion (VBF) mechanism and decaying into two taus with a lepton and a $\tau$-jet in the final state. Finally, the effect of this trigger on the trigger rate, as well as on the signal and background processes is investigated.

Data collected from the H2 beam line at CERN were used to calculate the average energy response and resolution of pions with momenta ranging from 2 GeV/c to 300 GeV/c. The results were then compared with data produced using the full CMS detector simulation. It was found that the average responses of the hadronic calorimeter were in good agreement, but the spread was underestimated by the simulation. Also the electromagnetic calorimeter response to pions was consistently underestimated in the detector simulation.

The Higgs boson decaying into two taus is a very promising channel for the detection of a low mass Higgs boson ($m_H < 150$ GeV/c$^2$). A trigger of two, separated in $\eta$, jets (VBF jets) and a lepton was developed and tuned for the $2 \times 10^{33}$ cm$^{-2}$s$^{-1}$ instantaneous luminosity regime. The trigger efficiency for a preselected VBF produced $H \to \tau\tau \to l + \tau$-jet sample was calculated to be $(76.1\pm0.7)$% with a $(0.6\pm0.3)$ Hz QCD rate. The combination of the new trigger with the existing ones provided an improvement of $\sim 33\%$ relative to the latter.

The performance of the trigger on the signal and background events passing selection criteria designed for Higgs boson detection was evaluated. The combined trigger was found to be $\sim 47\%$ more efficient for final signal selection than the existing ones, with only a $\sim 9\%$ increase in background events.
Acknowledgements

The road towards the completion of my PhD has been long and challenging, but most of all it has been an extremely valuable experience. First of all, I would like to thank my supervisor, Konstantinos Fountas, for the academic guidance and continuous encouragement and motivation I received during my many years of study.

I would also like to thank Alexandre Nikitenko, David Colling and Chris Seez for their support and guidance especially during the early stages of my PhD. Moreover I would like to thank my friend and colleague Konstantinos A. Petridis for his invaluable help and sharing of knowledge regarding my VBF Higgs analysis, as well as my friend and colleague Mark Pesaresi for his worryingly similar passion to solve insignificant computer problems and for looking out for me in general.

I would like to thank my sister Anthi and my partner Kalyani for their unreserved support and help in correcting this thesis.

Finally, I would like to give my deepest thanks to my mother, Kyveli, and father, Vangelis, for their never-ending belief in me, without whom none of this would have been possible. This thesis is truly for you.
Contents

Declaration 1

Abstract 2

Acknowledgements 3

Contents 4

List of Figures 10

List of Tables 18

Chapter 1. Theoretical Background 22

1.1 Quantum Field Theory and Gauge Invariance 22

1.2 The Standard Model 24

1.2.1 Quantum Electro-Dynamics 24

1.2.2 Quantum Chromo-Dynamics 26

1.2.3 Electro-Weak Theory 28

1.2.4 Electro-Weak Symmetry Breaking and Particle Masses 30
## Contents

2.2.6 The Muon System 56

2.2.7 The Trigger and Data Acquisition Systems 59

2.2.7.1 The Level-1 Trigger 59

2.2.7.2 The High Level Trigger 62

2.3 The CMS Computing Model 63

Chapter 3. Trigger and Offline Analysis Objects 65

3.1 Electrons 65

3.1.1 L1 $e/\gamma$ Trigger Object 65

3.1.2 HLT Electron 67

3.1.3 Offline Electron 70

3.2 Muons 73

3.2.1 L1 Muon Trigger Object 73

3.2.2 HLT Muon 76

3.2.3 Offline Muon Reconstruction 77

3.3 Jets 78

3.3.1 L1 Jet object 78

3.3.2 HLT and Offline Jet object 79

3.4 $\tau$-jets 80

3.4.1 Offline $\tau$-jet 80
Chapter 4. ECAL and HCAL Responses for Charged Pions and Comparisons with CMSSW Simulated data

4.1 H2 Beam Line

4.2 Test Beam 2006 Experimental Setup

4.3 Particle ID and Beam Cleaning

4.4 Data Sets

4.5 Energy Response

4.6 Comparison of Test Beam and Simulated Data

4.6.1 Detector Simulation and Reconstruction

4.6.2 Simulated Data Production

4.6.3 Comparison between Simulated and Test Beam Results

4.7 Summary and Conclusions

Chapter 5. Motivation and Development of Vector Boson Fusion Jets plus Lepton Trigger

5.1 Vector Boson Fusion Higgs

5.1.1 Production Cross Section

5.1.2 Event Kinematics at 10 TeV and 14 TeV CoM Energy

5.1.2.1 Vector Boson Fusion Jets

5.1.2.2 τ-jet and Lepton

5.2 Higgs Boson Trigger Study

5.2.1 Current Trigger Selections

5.2.2 The VBF Jets plus Lepton Trigger
## List of Figures

1.1 Second order electroweak correction diagrams to the $W^\pm$ and $Z^0$ masses from $H^0$ loops (a), (b) and to the coupling constants through fermion loops (c). 34

1.2 95% Confidence Level from Tevatron direct searches for the Higgs boson [22]. 35

1.3 Data fitted with the Electro-Weak theory predictions for the search of the Higgs boson. 36

1.4 Dominant Higgs boson production mechanisms at LHC. 37

1.5 Higgs boson production cross section at LHC as a function of the Higgs mass [5]. 38

1.6 Higgs boson branching ratios and natural width as a function of the Higgs mass [5]. 38

2.1 Schematic representation of the CERN accelerator complex. 42

2.2 The total and elastic $pp$ collision cross sections as a function of the laboratory beam momentum and the centre-of-mass energy. The high energy data points are calculated from cosmic ray data [38]. 43

2.3 Schematic of the CMS detector [39]. 45
2.4 A quadrant of the CMS silicon tracker in r-z view.

2.5 Illustration of the pixel detectors in the CMS tracker.

2.6 Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits.

2.7 Material budget of the CMS tracker in units of radiation length as a function of \( \eta \) for the different sub-detectors.

2.8 Transverse quadrant section of ECAL. The barrel extends from 1.29 to 1.75 meters from the interaction point and the endcap with the preshower module from 3.0 to 3.9 meters. The dashed lines show the coverage in \( \eta \).

2.9 The ECAL energy resolution, \( \sigma(E)/E \), as a function of electron energy measured in a 3×3 crystal array. Results were obtained from test beam data [39].

2.10 Transverse quadrant section of HCAL.

2.11 The energy resolution of HCAL (3×3 HCAL towers) plus ECAL (7×7 ECAL crystals) centred about the beam direction, as a function of charged pion momentum. Data were obtained from the 2004 test beam [45].

2.12 Transverse quadrant section of Muon system.

2.13 The muon transverse momentum resolution as a function of the muon transverse momentum (\( p_T \)) using the muon system only, the inner tracking only and both systems.

2.14 Inclusive proton-proton cross sections for basic physics processes. Interaction rates for the nominal luminosity are given on the right hand scale. The L1 and HLT input/output rates are also denoted in the graph. For low particle mass the processes are described primarily by the SM, while the processes and particles on the right are predicted by more exotic models (such as SUSY) [48].
2.15 Schematic overview of the L1 Trigger. 61

2.16 Schematic flow of event data in the CMS Computing Model. 63

3.1 Schematic representation of the L1 e/γ isolation and lateral shower shape algorithms. 66

3.2 Schematic representation of the Hybrid clustering algorithm. [53]. 67

3.3 Schematic representation of the Island clustering algorithm on the left and the super-clustering performance for electrons of $p_T = 30$ GeV/c [53] on the right. 68

3.4 Schematic representation of the Multi5×5 clustering algorithm [55]. 71

3.5 Four hit patterns created by muon tracks bent in the magnetic field. 75

3.6 Schematic representation of the L1 jet clustering algorithm. A 4×4 trigger tower window forms the trigger calorimeter region and in turn 3×3 trigger calorimeter regions form the L1 jet. 79

3.7 Comparison of $E_T$ (left) and $\phi$ (right) resolutions using Particle Flow reconstruction (red line) and calorimeter based reconstruction (black line) of single taus with $p_T = 50$ GeV/c. No jet energy corrections are applied in either case [62]. 81

3.8 Schematic representation of the signal cone and the isolation annulus. The signal cone is formed around the track of the leading PFlow charged hadron. For a tau to be considered isolated, no particle above a predefined $p_T$ is allowed in the isolation annulus. 82

4.1 Schematic representation of the Combined 2006 Test Beam experimental setup. 87

4.2 ADC output of trigger scintillator counters. The output of S1, S2 and S4 was used to reject events with no particles. The output of S3 was useful for studies which required a very focused beam. 89
4.3 ADC output of the BH scintillator counters. The BH scintillators were used to reject events with additional particles coming from the beam halo. The fits was performed at the gaussian part of each distribution. 90

4.4 ADC output of MV and CK scintillator counters. MV2 and MV3 are part of the MVW and their output was used in the LVE mode, while MVB was used in the high momentum mode. The output of CK3 was used to reject events containing electrons in the pion runs. The fits was performed at the gaussian part of each distribution. 92

4.8 The response of the pions measured by the test beam data (hollow) and MC (filled) is shown as a function of the $\pi$ beam momentum. Pions with $p_\pi < 8\text{ GeV/c}$ are measured only by the ECAL+HCAL system, while higher momentum pions are separated depending on if they started their hadronic shower in ECAL (blue) or not (red). 103

4.9 Comparison of ECAL and HCAL responses from simulated and test beam data for $p_\pi = 50\text{ GeV/c}$. 104

5.1 Pseudorapidity distribution of the two VBF jets, the $\tau$-jet and the lepton originating from the $\tau$ decay. The high $\eta$ nature of the jets is evident, while the Higgs boson decay products are located more centrally. These distributions correspond to the 14 TeV CoM scenario, but they are very similar to the 10 TeV generated data. 108

6.1 E/P distributions of reconstructed electrons matching MC electrons (solid blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines. 132

6.2 H/E distributions of reconstructed electrons matching MC electrons (solid blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines. 133
6.3 $\Delta \phi_{\text{in}}$ distributions of reconstructed electrons matching MC electrons (solid blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines. 133

6.4 $\Delta \eta_{\text{in}}$ distributions of reconstructed electrons matching MC electrons (solid blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines. 134

6.5 $\sigma_{\eta\eta}$ distributions of reconstructed electrons matching MC electrons (solid blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines. 134

6.6 Tracker isolation distributions of reconstructed electrons matching MC electrons (solid dark blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) as well as MC matched electrons coming from the $b/c \rightarrow e$ process (dash-dot light blue line) in the barrel (left) and the endcap (right). Electrons included in the plots have already satisfied the electron identification criteria presented in Section 6.3.1.1. The threshold value is shown by the vertical back line. 137

6.7 ECAL isolation distributions of reconstructed electrons matching MC electrons (solid dark blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) as well as MC matched electrons coming from the $b/c \rightarrow e$ process (dash-dot light blue line) in the barrel (left) and the endcap (right). Electrons included in the plots have already satisfied the electron identification criteria presented in Section 6.3.1.1. The threshold values are shown by the vertical dashed black lines. 137
6.8 HCAL isolation distributions of reconstructed electrons matching MC electrons (solid dark blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) as well as MC matched electrons coming from the $b/c \rightarrow e$ process (dash-dot light blue line) in the barrel (left) and the endcap (right). Electrons included in the plots have already satisfied the electron identification criteria presented in Section 6.3.1.1. The threshold values is shown by the vertical dashed black lines. 138

6.9 On the left $E_T$ distributions of offline reconstructed electrons for the signal and background processes. On the right $\eta$ distributions of offline reconstructed electrons for the signal and background processes. Electrons included in these plots have already satisfied the electron identification and isolation criteria presented earlier. The $E_T$ threshold value is shown by the vertical dashed black line. The drop of efficiency in $\eta$ is due to the ECAL calorimeter gaps. 138

6.10 Muon reconstruction efficiency as a function of the MC matched $p_T$ (left) and $\eta$ (right). Efficiency remains around 80% $p_T > 10$ GeV/c, while lower momentum muons result in the apparent drop in efficiency vs $\eta$. Muons included in these plots have satisfied the identification and isolation criteria described in the text. 140

6.11 Reconstruction purity for muons that have satisfied the identification and isolation criteria described in the text as a function of the reconstructed $p_T$ (left) and $\eta$ (right). The muon purity remains over 99.9% for the full range of $p_T$ and $\eta$ of the reconstructed muons. 141

6.12 On the left $p_T$ distributions of offline reconstructed muons for the signal and background processes. On the right $\eta$ distributions of offline reconstructed muons for the signal and background processes. Muons included in these plots have already satisfied the electron identification and isolation criteria presented earlier. The $p_T$ threshold value is shown by the vertical back line. 141
6.13 Variables used for electron rejection of $\tau$-jets. Offline reconstructed $\tau$-jets satisfy the transverse energy requirement of $E_T^{\tau\text{-jet}}>15$ GeV.

6.14 The $E_T$ distributions for the most energetic VBF jet (left) and the second most energetic VBF jet (right) for the signal (solid blue line), Z+jets (dashed red line), W+jets (dotted green line) and $t\bar{t}$+jets (dash-dot purple line).

6.15 The pseudorapidity gap between the two VBF jets (left) and the invariant mass of the two VBF jets (right) for the signal (solid blue line), Z+jets (dashed red line), W+jets (dotted green line) and $t\bar{t}$+jets (dash-dot purple line).

6.17 Transverse invariant mass of reconstructed lepton and MET for the signal (solid blue line), Z+jets (dashed red line), W+jets (dotted green line) and $t\bar{t}$+jets (dash-dot purple line).

6.18 Transverse opening angle between the reconstructed lepton and the $\tau$-jet for the signal (solid blue line), Z+jets (dashed red line), W+jets (dotted green line) and $t\bar{t}$+jets (dash-dot purple line). The lepton and $\tau$-jet coming from the Z decay tend to be back-to-back.

6.19 Distributions of Higgs and Z bosons masses calculated using the lepton and $\tau$-jet pair (left) and the collinear approximation for the neutrina (right).

6.20 Z (a) and H (b) offline mass reconstruction. The large positive tail signifies overestimated MET.

6.21 Graphical comparison of the cross section reductions due to the different selection criteria for the signal and background processes using the VBF jets plus lepton trigger (solid blue line) and the single lepton triggers (dashed red line). The cross sections of the combined trigger follow the trend on the single lepton triggers at the beginning, while at the end they follow that of the VBF jets plus lepton trigger.
6.22 Higgs boson reconstructed mass using the collinear approximation described in Section 6.3.8. The distributions were extracted before applying the VBF jet selection criteria and then scaled to the cross section expected by the combination of the VBF jets plus lepton and the single isolated lepton triggers.
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The LHC nominal beam parameters for proton and lead ion collisions.</td>
<td>42</td>
</tr>
<tr>
<td>2.2</td>
<td>Summary of the different charged tracker module parameters.</td>
<td>48</td>
</tr>
<tr>
<td>2.3</td>
<td>Main parameters of the CMS superconducting magnet.</td>
<td>56</td>
</tr>
<tr>
<td>4.1</td>
<td>Main parameters of the H2 beam.</td>
<td>85</td>
</tr>
<tr>
<td>4.2</td>
<td>Summary of the particles, their momenta and the number of events used in this analysis.</td>
<td>94</td>
</tr>
<tr>
<td>4.3</td>
<td>Average response and standard deviation of pions interacting in the HCAL and pions starting their hadronic shower in the ECAL using test beam data. The error on the average response obtained by factoring the error in the fit and by varying the fitting range is about 3% for low momentum pions and decreases gradually to about 0.2% at $p_\pi = 300\text{GeV}/c$. The error in the standard deviation obtained in the same way is about 3% for the all pion momenta.</td>
<td>98</td>
</tr>
<tr>
<td>4.4</td>
<td>Summary of the particles, their momenta and the number of events produced for this analysis.</td>
<td>101</td>
</tr>
</tbody>
</table>
4.5 Average response and standard deviation of pions interacting in the HCAL and pions starting their hadronic shower in the ECAL using MC data. The error on the average response obtained by factoring the error in the fit and by varying the fitting range is about 2% for low momentum pions and decreases gradually less than 0.1% at $p_\pi = 300$ GeV/c. The error in the standard deviation obtained in the same way is about 2% for the all pion momenta.

5.1 The Next to Leading Order (NLO) cross sections for the Higgs boson production via the VBF and the branching ratios for the $H \rightarrow \tau\tau$ decay channel for different masses of the Higgs boson. The cross sections were generated according to calculations from [67], using the code provided by M.Spira [68] with the CTEQ6M parton distribution function. The scale dependency from renormalisation and factorisation is of the order of $\pm 2\%$ for all values.

5.2 Summary of the kinematic selections and their corresponding rates of the trigger paths used for the VBF Higgs channel [71].

5.3 Electron HLT selection thresholds.

5.4 Muon HLT selection thresholds.

5.5 VBF jet specific condition thresholds.

5.6 Table summarising the performance of the VBF jet plus lepton trigger on QCD events with $\hat{p}_T > 15$ GeV/GeV/c for the $2 \times 10^{33}$ cm$^{-2}$s$^{-1}$ luminosity regime. The error on the rate represents the statistical uncertainty. A total rate of $(0.6 \pm 0.2)$ Hz was obtained.

5.7 Table summarising the performance of the single lepton triggers on QCD events with $\hat{p}_T > 15$ GeV/GeV/c for the $2 \times 10^{33}$ cm$^{-2}$s$^{-1}$ luminosity regime. The error on the rate represents the statistical uncertainty. A total rate of $(62.9 \pm 4.8)$ Hz was obtained.
5.8 Table summarising the performance of the VBF jet plus lepton trigger on pre-selected VBF $H \to \tau\tau \to l+\tau$-jet events. The pre-selection criteria are presented in Section 5.2.4. The error on the rate represents the statistical uncertainty. A efficiency of $(76.1 \pm 0.7)$Hz was obtained. 120

5.9 Additional efficiency by using VBF jets plus lepton trigger and a combination of the VBF jets plus lepton trigger with the single isolated lepton triggers, compared to using only the single isolated lepton triggers. 121

6.1 Summary of the MC samples used in this analysis. $\epsilon_{\text{fl}}$ denotes the generator level pre-selection efficiency, $\sigma$ the process cross section and $L_{\text{int}}$ the integrated luminosity of the sample. 125

6.2 Cumulative cross sections of the signal and background events for each step of the VBF jets plus lepton trigger path. The numbers in the brackets denote the efficiency of the selection criterion against the previous one. 127

6.3 Cumulative cross sections of the signal and background events for single isolated lepton trigger paths. The numbers in the brackets denote the efficiency of the selection criterion against the previous one. The $Z$ and $W^\pm$ boson HLT rates are in agreement with those presented in the Technical Design Report (see text for more information). The high efficiency both in signal and background is due to the absence of the VBF jets selection criteria. 128

6.4 Cumulative cross sections of the signal and background events for the combined VBF jets plus lepton and single isolated lepton trigger paths. The numbers in the brackets denote the efficiency of the selection criterion against the previous one. 129

6.5 Table summarising the selection efficiencies of different electron identification criteria for reconstructed electrons matching MC electron, MC $\tau$-jets and MC QCD jets. 132
6.6 Table summarising the photon conversion selection efficiencies. Electrons used had already satisfied the identification criteria described in this section.

6.7 Table summarising the selection efficiencies of different electron isolation criteria for signal reconstructed electrons matching MC electron, MC $\tau$-jets, MC QCD jets as well as MC matched electrons coming from $b/c \to e$ process. Electrons included in the table have already satisfied the electron identification criteria presented in the previous section.

6.8 Table summarising the performance of $\tau$-jet identification and isolation criteria as well as lepton rejections for $\tau$-jets with $E_T > 15$ GeV matched to MC $\tau$-jets, QCD jets, electrons and muons. Efficiencies are relative to the previous selection. The rejection factors of fake $\tau$-jets are of the order of $10^{-3}$ for a $\tau$-jet efficiency of $\sim 35%$.

6.9 Cumulative cross sections of the signal and background events, satisfying the VBF jets plus lepton trigger, for each step of the offline selection criteria.

6.10 Cumulative cross sections of the signal and background events, satisfying the single isolated lepton triggers, for each step of the offline selection criteria.

6.11 Cumulative cross sections of the signal and background events, satisfying the combination of the VBF jets plus lepton and single isolated lepton triggers, for each step of the offline selection criteria.
Chapter 1

Theoretical Background

The Higgs boson is the only particle predicted by the Standard Model (SM) of particle physics which is yet to be detected. It was originally suggested by Peter Higgs in 1964 \[1\] and further described through its interactions by Gerald Guralnik, Carl Hagen and Tom Kibble \[2\] later the same year. The proposal for the existence of the Higgs boson was motivated by experimental evidence that the weak force bosons (W$^\pm$ and Z$^0$ bosons) have mass, while the mathematical theory describing the interactions of the fundamental particles, Gauge Quantum Field Theory, predicts them to be massless. At the same time all massive SM particles acquire their masses through their interaction with the Higgs boson field, making it a fundamental constituent of the Standard Model. Although its existence has been predicted by theorists over 40 years ago, its detection remains for experimentalists a challenge yet to overcome\[1\].

1.1 Quantum Field Theory and Gauge Invariance

Quantum Field Theory (QFT) was first developed as an extension to relativistic wave mechanics incorporating the quantum nature of the electromagnetic field. It was later used to describe other fields and particles. QFT, unlike Quantum Mechanics (QM), is a multi-particle/field theory providing the mathematical formalism for

\[1\]This chapter draws from \[3, 4, 5\].
the creation and annihilation of particles through the excitation and de-excitation of their field. This is a fundamental requirement for any mathematical model used to describe particle physics, since interactions where the incoming and outgoing particles differ are experimentally observed. The probability of an initial state $i$ evolving to a final state $f$ is calculated by the quantum mechanical amplitude of the scattering process. This amplitude is called the $S$-matrix. It is easier to present the $S$-matrix using the Heisenberg representation of QM where the operators vary with time, while the states are time independent. In this picture the Hamiltonian ($\hat{H}$) splits into the free field ($\hat{H}_0$) and the interaction part ($\hat{H}_I$):

$$\hat{H} = \hat{H}_0 + \hat{H}_I.$$ (1.1)

The initial and the final states are taken at times $-\infty$ and $+\infty$ where the particles are considered free (non-interacting) fields. The $S$-matrix then is defined as:

$$S_{fi} = \langle f, t \rightarrow \infty | i, t \rightarrow -\infty \rangle.$$ (1.2)

The $S$-operator is then given by:

$$S_{fi} = 1 - i \int_{-\infty}^{\infty} dt_1 \hat{H}_I(t_1) + \frac{(-i)^2}{2!} \int_{-\infty}^{\infty} dt_1 dt_2 \hat{T}[\hat{H}_I(t_1)\hat{H}_I(t_2)] + \cdots$$

$$= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} dt_1 \cdots dt_n \hat{T}[\hat{H}_I(t_1) \cdots \hat{H}_I(t_2)]$$ (1.3)

where $\hat{T}$ is the time order operator and $\hat{H}_I$ depends on the specific theory. It is then easy to read the different processes from the different orders of the perturbation series. The higher orders correspond to radiative corrections and in a perturbative theory contribute to a small degree.

The mathematical formulation commonly used to describe quantum field theories is the Lagrangian. The Lagrangian formalism is preferred as it can demonstrate the symmetries of the theory. The Lorentz covariant form of the Euler-Lagrange equation is given by:

$$\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0$$ (1.4)
where $\mathcal{L}$ is the Lagrangian density ($L = \int \mathcal{L}(x,t)d^4x$), $\phi$ is the field and $\partial_\mu \equiv \partial/\partial x_\mu$.

Some examples are the Klein-Gordon and the Dirac Lagrangian densities given by:

$$\mathcal{L}_{\text{KG}} = \frac{1}{2} \partial^\mu \phi \partial_\mu \phi - \frac{1}{2}m^2 \phi^2$$  \hspace{1cm} (1.5)

$$\mathcal{L}_{\text{Dirac}} = \overline{\phi}(i\gamma^\mu \partial_\mu - m)\phi$$ \hspace{1cm} (1.6)

where $m$ is the mass of the field’s particle, $\gamma_\mu$ are the Dirac matrices satisfying the anti-commutator relationships $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$ and $\bar{\phi} \equiv \phi^\dagger \gamma^0$. By applying Equations 1.5, 1.6 to 1.4, we get:

$$\partial^\mu \partial_\mu \phi + m^2 \phi = (\Box + m^2)\phi = 0 \hspace{1cm} (1.7)$$

$$(-i\gamma^\mu \partial_\mu + m)\phi = 0. \hspace{1cm} (1.8)$$

These are the Klein-Gordon and the Dirac equations of the free field motion.

### 1.2 The Standard Model

The Standard Model is a combination of gauge invariant QFTs that describe the interactions of the elementary particles via the electroweak and the strong forces. The SM symmetry group is $\text{SU}(3)_c \times \text{SU}(2)_L \times \text{U}(1)_Y$ where the generators of $\text{SU}(3)_c$ group give the different QCD colour charges, the generators of $\text{SU}(2)_L$ give the isospin charges and the $\text{U}(1)_Y$ generator gives the weak hypercharge. As mentioned earlier, the weakly interacting bosons are massive, while no mass term for these bosons can be included into the Lagrangian density without breaking the gauge invariance. This implies that the symmetry must be broken. The electroweak symmetry breaking takes place in the SM through the Higgs mechanism. This is a minimal way to provide masses for the weakly interacting bosons, while all other particles acquire their masses through the interaction with the Higgs field.

#### 1.2.1 Quantum Electro-Dynamics

Quantum Electro-Dynamics (QED) was historically the first complete QFT in particle physics and is based on the $\text{U}(1)$ symmetry group. As the name suggests, it
provides the mathematical framework to describe the interactions of particles carrying electric charge. To describe the fermion fields, the Dirac Lagrangian density (Equation 1.6) provides a good basis, since it describes the kinematics of spin ±1/2 particles. At the heart of every SM QFT is a symmetry which allows the physics interaction to be invariant to the particular phase that the field has. This is called gauge symmetry and in the simplest Abelian case, U(1), the global phase transformation of a fermion field (ψ) is as follows:

$$\psi(x) \rightarrow \psi'(x) \equiv e^{-iq_{em}\theta} \psi(x)$$  \hspace{1cm} (1.9)

where $q_{em}$ is the charge magnitude and $\theta$ is an arbitrary real number. By substituting Equation 1.9 into Equation 1.6 it can be shown that $\mathcal{L}$ is invariant under a global phase transformation. According to Noether’s theorem [6, 7], for every symmetry in a system there is a corresponding conserved quantity. Using Equation 1.4 it can be shown that:

$$-i\theta \partial \mu \left( \frac{\partial \mathcal{L}}{\partial (\partial \mu \psi)} q_{em} \psi \right) = 0$$  \hspace{1cm} (1.10)

If we then define a current $J^\mu$ equal to the quantity in the brackets we have:

$$\partial \mu J^\mu = 0.$$  \hspace{1cm} (1.11)

This shows that the current $J^\mu$ is a conserved quantity (Noether current).

If we now consider a local gauge transformation of a fermion field (ψ) we get:

$$\psi(x) \rightarrow \psi'(x) \equiv e^{-iq_{em}\theta(x)} \psi(x)$$  \hspace{1cm} (1.12)

where $\theta(x)$ is an arbitrary function of $x$. It can be shown that $\mathcal{L}$ is not invariant under a local Gauge transformation. The important thing to note though is that only derivatives of $\mathcal{L}$ appear in the Euler-Lagrange equation (Equation 1.4). Therefore, any additional term to $\mathcal{L}$ that doesn’t appear in the final derivative will leave the equations of motion unchanged. To achieve this, we introduce the covariant derivative which is defined as

$$D_{\mu} = \partial_{\mu} + ig_{em}q_{em}A_{\mu}$$  \hspace{1cm} (1.13)
where \( g_{\text{em}} \) is the electric charge coupling constant and \( A_\mu \) is a vector potential - in this case of the photon - which transforms as:

\[
A_\mu (x) \rightarrow A'_\mu (x) = A_\mu + \frac{1}{g_{\text{em}}} \partial_\mu \theta(x) .
\] (1.14)

Using Equations 1.9, 1.13 and 1.14 we can deduce that the covariant derivative of a fermion field (\( \psi \)) transforms as

\[
D_\mu \psi (x) \rightarrow [D_\mu \psi (x)]' = \left[ \partial_\mu + ig_{\text{em}}q_{\text{em}}A'_\mu \right] e^{-ig_{\text{em}}q_{\text{em}}(x)} \psi (x) = e^{-ig_{\text{em}}q_{\text{em}}(x)} D_\mu \psi (x) .
\] (1.15)

It can then be shown that \( \mathcal{L} \) stays invariant under local gauge transformations. The introduction of the gauge boson field \( A_\mu \) implies that kinetic terms of this field have to be included in the Lagrangian density. These terms have to be gauge invariant themselves to give rise to the field tensor \( F_{\mu\nu} \) and the photon kinetic Lagrangian term of:

\[
F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad \text{(1.16)}
\]

\[
\mathcal{L}_\gamma = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad \text{(1.17)}
\]

where the \(-\frac{1}{4}\) ensures that the photon equations of motion coincide with Maxwell’s equations. A mass term of the photon potential \( A_\mu \) of the form of \(-\frac{1}{2} m_\gamma^2 A_\mu A^\mu \) would violate gauge invariance unless \( m_\gamma = 0 \).

Combining the Equations 1.6, 1.13 and 1.17 we get the gauge invariant QED Lagrangian density to be:

\[
\mathcal{L}_\gamma = \bar{\psi} (i \gamma^\mu D_\mu - m) \psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} .
\] (1.18)

### 1.2.2 Quantum Chromo-Dynamics

Quantum Chromo-Dynamics (QCD) is a QFT describing the interactions of the particles carrying colour (chromo) charge, governed by the non-Abelian SU(3) symmetry group. In an analogous way to the QED Lagrangian density (Equation 1.18) we can define the QCD Lagrangian density as:

\[
\mathcal{L}_{\text{QCD}} = \bar{\psi} (i \gamma^\mu D_\mu - m) \psi - \frac{1}{4} G_{\mu\nu}^i G_{ij}^{\mu\nu} .
\] (1.19)
where $D_\mu$ is the covariant derivative of the SU(3) symmetry group and $G_{\mu\nu}$ is the QCD strength tensor. Unlike the U(1) case where the symmetry group is one-dimensional, the phase transformation now occurs in three-dimensions. Therefore the local gauge transformation of the fermion field $\psi$ is:

$$\psi(x) \rightarrow \psi'(x) = e^{-iT_j \theta_j(x)} \psi(x) \quad (1.20)$$

where $T_j \ (j = 1, 2, \ldots, 8)$ are $3 \times 3$ matrices, forming a faithful representation of the eight SU(3) generators. Historically, the Gell-Mann matrices are used to represent the SU(3) generators. Since the symmetry group is non-Abelian the generators $T_j$ must satisfy the commutator relations determined by the Lie algebra:

$$[T_j, T_k] = i f_{jkl} T_l \quad (1.21)$$

where $f_{jkl}$ are the structure constants of the group and are anti-symmetric under an interchange of any pair of indices.

The covariant derivative then can be defined as:

$$D_\mu \equiv \partial_\mu + i g_s T_j G_{j\mu} \quad (1.22)$$

where $g_s$ is the coupling constant of the strong force and $G^\mu_j$ are the eight vector potentials of the gluon bosons, which transform under local a gauge transformation as:

$$G_{j\mu}(x) \rightarrow G'_{j\mu}(x) \equiv G_{j\mu} + \frac{1}{g_s} \partial_\mu \theta_j(x). \quad (1.23)$$

The eight field strength tensors $G_{j\mu\nu}$ are then defined as:

$$G_{j\mu\nu} \equiv \partial_\mu G_{j\nu} - \partial_\nu G_{j\mu} - g_s f_{jkl} G_{k\mu} G_{l\nu}. \quad (1.24)$$

The last term is a result of the non commuting SU(3) generators and indicates field self interaction. Combining Equations 1.19 and 1.22 we obtain a more analytical form of the QCD Lagrangian density given by:

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}^a (i \gamma^\mu \partial_\mu - m) \psi^a - g_s T_j G_{j\mu} \bar{\psi}^a \gamma^\mu \psi^a - \frac{1}{4} G_{j\mu\nu} G_{j}^{\mu\nu} \quad (1.25)$$

where $\psi^j$ are the quark fermion fields with different colour charge.
1.2.3 Electro-Weak Theory

The Electro-Weak theory describes the interactions governed by the unified electroweak force. It is based on the $SU(2)_L \times U(1)_Y$ symmetry group, where $SU(2)_L$ generators provide the isospin charge and $U(1)_Y$ provides the hypercharge. The hypercharge is related to the electromagnetic charge according to:

$$ q_{em} = I_3 + \frac{Y}{2} $$

(1.26)

where $I_3$ is the charge of the third isospin generator. The subscript $L$ on the $SU(2)$ symmetry group signifies that only left-handed fermions interact under the weak force. This is not a mathematical requirement, but rather an experimental input to the theory. It reflects the parity-violating nature of the weak interaction since no right-handed weakly governed interactions have been observed. The left-handed ($\psi_L$) and right-handed ($\psi_R$) fermion fields are defined as:

$$ \psi_L = \frac{(1 - \gamma^5)}{2} \psi $$

(1.27)

$$ \psi_R = \frac{(1 + \gamma^5)}{2} \psi $$

(1.28)

where $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$.

The leptonic fermion fields are given by:

$$ \psi_L = \left( \begin{array}{c} \nu_l \\ l \end{array} \right)_L \quad \psi_R = l_R $$

(1.29)

where $l$ is the lepton and $\nu_l$ is the corresponding lepton neutrino. In the case of the quarks, the free quantum states and the weakly interacting quantum states are not the same. The mixing of the states is described by the Cabibbo-Kobayashi-Maskawa matrix [8, 9] and the weak quark doublet of the first generation can be written as:

$$ \psi_L = \left( \begin{array}{c} u \\ d' \end{array} \right)_L \quad \psi_R = u_R \text{ or } d'_R $$

(1.30)

where $u$ is the up quark and $d'$ mixed state of the down quark.

The local gauge transformation of the electroweakly interacting fermion field $\psi$ then is:

$$ \psi(x) \rightarrow \psi'(x) \equiv e^{-i\frac{\theta_j}{2}\theta_j(x)}e^{-i\frac{Y}{2}\omega(x)\psi(x)} $$

(1.31)
1.2 The Standard Model

where $\tau_j$ ($j = 1, 2, 3$) are the Pauli matrices, which form a faithful representation of the SU(2) symmetry group. As with the SU(3) case, the group is non-Abelian and the generators satisfy the commutator relations:

$$\left[ \frac{\tau_j}{2}, \frac{\tau_k}{2} \right] = i\epsilon_{jkl} \frac{\tau_l}{2}$$  \hspace{1cm} (1.32)

where $\epsilon_{jkl}$ are the anti-symmetric structure constants of the SU(2) group.

The covariant derivative for the left-handed fermions can then be defined as:

$$D_\mu \equiv \partial_\mu + ig\frac{\tau_j}{2}W_{j\mu} + ig'\frac{1}{2}YB_\mu$$  \hspace{1cm} (1.33)

where $W_{j\mu}$ and $B_\mu$ are the isospin and the hypercharge vector field potentials respectively, while $g$ and $g'$ are the isospin and hypercharge coupling constants that are related to each other and the electric charge through:

$$\frac{g'}{g} = \tan \theta_W$$  \hspace{1cm} (1.34)

$$e = g \sin \theta_W = g' \cos \theta_W$$  \hspace{1cm} (1.35)

where $e$ is the electric charge and $\theta_W$ is the Weinberg angle. The covariant derivative for the right-handed fermions follows a U(1) symmetry and is the QED derivative with a different charge (see Equation 1.13).

The isospin ($W_{j\mu}$) and the hypercharge ($B_\mu$) vector potentials transform under local gauge transformation as:

$$W_{j\mu}(x) \rightarrow W'_{j\mu}(x) \equiv W_{j\mu} + \frac{1}{g}\partial_\mu \theta_j(x)$$  \hspace{1cm} (1.36)

$$B_\mu(x) \rightarrow B'_\mu(x) \equiv B_\mu + \frac{1}{g'}\partial_\mu \omega(x).$$  \hspace{1cm} (1.37)

The field strength tensors are then defined as:

$$W_{j\mu\nu} \equiv \partial_\mu W_{j\nu} - \partial_\nu W_{j\mu} - g\epsilon^{jkl}W_{k\mu}W_{l\nu}$$  \hspace{1cm} (1.38)

$$B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu.$$  \hspace{1cm} (1.39)

The complete Electro-Weak Lagrangian density is then given by:

$$\mathcal{L}_{EW} = i\bar{\psi}_L^a \gamma^\mu (\partial_\mu - g\frac{\tau_j}{2}W_{j\mu} - g'\frac{1}{2}YB_\mu)\psi^a_L$$

$$+ i\bar{\psi}_R^a \gamma^\mu (\partial_\mu - g'\frac{1}{2}YB_\mu)\psi^a_R$$

$$- \frac{1}{4} W_{j\mu\nu}W_{j}^{\mu\nu} - \frac{1}{4} B_{\mu\nu}B^{\mu\nu}.$$  \hspace{1cm} (1.40)
It is worth noting that no mass term can be entered in the Electro-Weak Lagrangian since the fermion mass term \( m \bar{\psi} \psi = m (\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L) \) is not invariant under gauge transformations. Also, the bosons are unable to acquire mass as explained in Section 1.2.1. This implies that the electroweak symmetry must be broken.

The weak bosons, \( W^\pm \) and \( Z^0 \), as well as photon can be written as a linear combination of the three \( W^j_{\mu} \) and the \( B_{\mu} \) fields as:

\[
W^\pm_{\mu} = \frac{1}{\sqrt{2}} (W^1_{\mu} \mp iW^2_{\mu}) \tag{1.41}
\]

\[
Z_{\mu} = B_{\mu} \sin \theta_W - W^3_{\mu} \cos \theta_W \tag{1.42}
\]

\[
\gamma_{\mu} = B_{\mu} \cos \theta_W + W^3_{\mu} \sin \theta_W \tag{1.43}
\]

where the \( \theta_W \) is the Weinberg angle (see Equation 1.34).

### 1.2.4 Electro-Weak Symmetry Breaking and Particle Masses

As discussed in the previous section, there are clear indications that the Electro-Weak symmetry must be broken. We start with a general Lagrangian density of a complex scalar field \( \phi \), which can be expressed as:

\[
\mathcal{L} = (D_{\mu} \phi)(D^{\mu} \phi^*) - \mu^2 \phi \phi^* - \lambda (\phi \phi^*)^2 \tag{1.44}
\]

where \( \mu^2 \) would be interpreted as the mass of the field quanta and \( \lambda \) would be interpreted as a form of self-interaction. The potential of this field can then be defined as

\[
V(\phi) \equiv \mu^2 \phi \phi^* + \lambda (\phi \phi^*)^2 = \mu^2 |\phi|^2 + \lambda (|\phi|^2)^2. \tag{1.45}
\]

The potential is unstable (no-minimum) for \( \lambda < 0 \), so \( \lambda > 0 \) is preferred. The choice of \( \mu^2 < 0 \) provides us with the desired potential such that the minimum is not at zero and therefore can break the symmetry. The minimum of such a potential is then at

\[
|\phi| = \frac{v}{\sqrt{2}} \tag{1.46}
\]

with

\[
v \equiv \sqrt{-\frac{\mu^2}{\lambda}}. \tag{1.47}
\]
In order to give mass to the three weak bosons the \( \phi \) field must have at least three scalar fields plus one for its own quantum. Thus, the simplest choice is a doublet of one charged and one neutral complex scalar field.

\[
\phi = \left( \begin{array}{c} \phi^+ \\ \phi^0 \end{array} \right). \tag{1.48}
\]

The field at its minimum state (vacuum) using Equation 1.46 can then be written as

\[
\phi_{\text{vac}} = \left( \begin{array}{c} 0 \\ \frac{v}{\sqrt{2}} \end{array} \right). \tag{1.49}
\]

An SU(2) \( \times \) U(1) symmetry breaking will provide three massless scalar bosons called “Gladstone” bosons \([10]\). These extra degrees of freedom are then used to provide mass for the weak bosons.

To be locally gauge invariant under electroweak symmetry the field must satisfy Equation 1.44 with the electroweak covariant derivative given by Equation 1.33.

Thus for the field hypercharge \( Y = 1 \) we have

\[
(D^\mu \phi_{\text{vac}})^\dagger D_\mu \phi_{\text{vac}} = -\frac{i}{2\sqrt{2}} \left( g(W^1_\mu + iW^2_\mu)v \right) \left( g' B_\mu - gW^3_\mu \right) v \times \frac{i}{2\sqrt{2}} \left( g(W^1_\mu - iW^2_\mu)v \right) \left( g' B_\mu - gW^3_\mu \right) v \]

\[
= \frac{1}{8} v^2 g^2 (W^1_\mu + iW^2_\mu)(W^1_\mu - iW^2_\mu) + \frac{1}{8} v^2 (g' B_\mu - gW^3_\mu)(g' B_\mu - gW^3_\mu). \tag{1.50}
\]

If we then substitute for \( W^\pm_\mu, Z^0_\mu \) and \( \gamma_\mu \) from Equations 1.41, 1.42 and 1.43 respectively into Equation 1.51 we get

\[
(D^\mu \phi_{\text{vac}})^\dagger D_\mu \phi_{\text{vac}} = \frac{v^2 g^2}{4} W^-_\mu W^+\mu + \frac{v^2 g^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu. \tag{1.51}
\]

As we can see no \( m^2 \gamma_\mu \gamma^\mu \) term appears, signifying that the U(1) generator remains unbroken. The mass of the \( W^\pm \) and \( Z^0 \) bosons can then be read

\[
m_{W^\pm} = \frac{g v}{2}, \tag{1.52}
\]

\[
m_{Z^0} = \frac{g v}{2 \cos \theta_W}. \tag{1.53}
\]
1.2 The Standard Model

As discussed in the previous section a mass term of the form \((m\bar{\psi}\psi)\) for the fermions violates \(SU(2)_L\) gauge invariance. However, the Yukawa coupling Lagrangian term given by Equation 1.54 remains gauge invariant:

\[
\mathcal{L}_{\text{Yuk}} = -g_f \bar{\psi} \phi \psi = g_f (\bar{\psi}_L \phi \psi_R + \bar{\psi}_R \phi^\dagger \psi_L)
\] (1.54)

where \(g_f\) is the coupling constant of the fermion with the Higgs field. For the first generation of leptons we get:

\[
-g_e (\bar{\psi}_L \phi \psi_R + \bar{\psi}_R \phi^\dagger \psi_L) = -g_e \left[ (\bar{\nu}_e_L \ e_L) \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} e_R + \bar{e}_R (0 \ v/\sqrt{2}) (\nu_{eL}) \right]
\] (1.55)

where the mass of the electron is

\[
m_e = \frac{g_e v}{\sqrt{2}}. \tag{1.56}
\]

The rest of the fermions acquire mass in an analogous way.

1.2.5 Shortcomings of the Standard Model

The Standard Model has proven to be one of the most successful models in particle physics. It has been experimentally tested extensively over the past decades without any evidence disproving it. Nevertheless, there are still issues that it provides very little or no information. For example, the SM does not include one of the four observed forces i.e. the gravitational force. Also, it does not provide any understanding into why the masses of the fermions are what they are. Furthermore, the SM does not provide a true unification of the forces, since the strengths of their coupling constants are not expected to meet. Moreover, the magnitude of the matter-antimatter asymmetry predicted by the SM can not explain the one that is observed in the universe.

The Higgs boson’s self couplings in the SM predict radiative corrections to the Higgs boson mass of the order of \(\Lambda^2\), where \(\Lambda\) is the energy scale of the interaction. If we assume that the SM is the theory that is valid up to the Grand Unified Theory
(GUT) scale of $10^{16}$ GeV, then the Higgs boson mass should be close to this scale, unless an unnatural fine-tuning of the perturbative expansion occurs. One possible solution is the SUperSYmmetric (SUSY) [11] model where these divergences cancel.

Lastly the SM predicts phenomena up to an energy scale $\Lambda = 1$ TeV, whereas the GUT scale is $\Lambda_{\text{GUT}} \sim 10^{16}$ GeV. This creates a problem of hierarchy, since it is unnatural to have a gap on new physics processes for an energy scale $10^3$ GeV $< \Lambda < 10^{16}$ GeV.

### 1.3 Detecting the Higgs Boson

#### 1.3.1 Direct constraints

The mass of the Higgs boson ($m_H = -2\mu^2$) is not directly constrained by an experimentally measured quantity. The $W_LW_L$ scattering amplitude includes $W^\pm$ boson and $H$ boson exchanges. It can be shown that for a high Higgs boson mass ($m_H$), the process violates unitarity [12]. This requirement places an upper bound of

$$m_H^2 \leq \frac{8\sqrt{2}\pi}{3G_F} \sim (760 \text{ GeV}/c^2)^2 \quad (1.57)$$

where $G_F$ is the Fermi coupling constant.

Further bounds (lower and upper) can be placed on the Higgs boson mass based on the Higgs boson’s self-interaction [13] [14]. These bounds depend on the SM energy scale $\Lambda$ and for $\Lambda \sim 1$ TeV we get $60 \text{ GeV}/c^2 \lesssim m_H \lesssim 700 \text{ GeV}/c^2$, while for $\Lambda \sim 10^{19}$ GeV the bounds become $130 \text{ GeV}/c^2 \lesssim m_H \lesssim 190 \text{ GeV}/c^2$.

#### 1.3.2 $W^\pm$ and $Z^0$ Radiative corrections

Radiative corrections of the weak coupling constant and the $W^\pm$ boson mass, involving the Higgs boson and massive fermions, are predicted by the Electro-Weak theory. The second order diagrams of such loops are shown in Figure [1.1]. These are of the order of $G_F/\sqrt{2} = g^2/8m_W^2$. After substituting leading-order expressions
for the electromagnetic coupling constant $\alpha$, the electroweak mixing angle and the $Z$ boson mass for the weak coupling constant and the $W^{\pm}$ boson mass, the relation becomes [5]:

$$G_F \frac{\sqrt{2}}{\sin^2 2\theta_W m_Z^2} = 2 \frac{2 \pi \alpha}{\sin^2 2\theta_W m_Z^2} [1 + \Delta r_\alpha + \Delta r_t + \Delta r_H]$$

(1.58)

where $\Delta r_\alpha$ describes the shift in the electromagnetic coupling constant $\alpha$ at the scale $m_Z^2$, $\Delta r_t$ is the top/bottom quark contribution to the $W^{\pm}$ and $Z$ boson masses. These depend on the fourth power of the top mass while $\Delta r_H$ is the contribution of the Higgs boson to the masses which depends logarithmically on the Higgs boson mass at leading order [15]

$$\Delta r_H = \frac{G_F m_Z^2 (1 + 9 \sin^2 \theta_W)}{24 \sqrt{2} \pi^2} \log \frac{m_W^2}{m_H^2} + \text{h.o.t} \quad (m_H^2 \gg m_W^2).$$

(1.59)

Figure 1.1: Second order electroweak correction diagrams to the $W^{\pm}$ and $Z^0$ masses from $H^0$ loops (a), (b) and to the coupling constants through fermion loops (c).

1.3.3 Past and Current Searches for the Higgs Boson

Many searches for the Higgs boson have been carried out since it was first suggested in the 1960’s. However, the mass of the particle has proved to be the biggest obstacle in this task up to now.

Direct searches for the Higgs boson at the Large Electron Positron (LEP) [16] experiments have set a lower bound on the Higgs boson mass, excluding masses below $m_H = 114.4$ GeV/$c^2$ (95% confidence level) [17]. This mass bound is based on the
main Higgs boson production mechanism at LEP, which is via the vector boson bremsstrahlung of a $H^0$ boson - Higgsstrahlung process $e^+e^- \rightarrow HZ$. Recent results from the Tevatron [18] experiments, the Collider Detector at Fermilab (CDF) [19] and D$\phi$ [20], have further excluded the Higgs mass in the range between 160 GeV/$c^2$ and 170 GeV/$c^2$ [21](see Fig. 1.2).

Indirect searches for the Higgs boson are based on the Electro-Weak theory predictions. The Electro-Weak theory predicts corrections to the $W^\pm$ and the $Z^0$ boson masses and coupling constants through higher order diagrams as described in Section 1.3.2. Using the electroweak precision data collected by the LEP, the Stanford Large Detector (SLD) [23] experiments for the $Z^0$ boson [23] and the data collected by the LEP, Tevatron experiments for the $W^\pm$ boson [25], the mass of the Higgs boson can be correlated, among other quantities, to the mass of the top fermion. Since the contribution of the top quark mass ($m_t$) on these corrections is quadratic, the Higgs mass prediction strongly depends on it. The $\Delta\chi^2$ fit to the electroweak data for $m_t = 173.1 \pm 0.6$ (stat.)$\pm 1.1$ (syst.) GeV/$c^2$ [26] as a function of the Higgs boson mass is shown in Figure 1.3.
1.3 Detecting the Higgs Boson

(a) $\Delta \chi^2$ for a global fit to electroweak data as a function of $m_H$, with the blue band indicating the theoretical uncertainty [27].

(b) 68% Confidence Level, given the latest measurements of W and top masses [27].

Figure 1.3: Data fitted with the Electro-Weak theory predictions for the search of the Higgs boson.

1.3.4 Higgs Boson at the Large Hadron Collider

The Large Hadron Collider (LHC) has the detection of the Higgs boson as a primary objective. The centre of mass energy for the $pp$ enables the probing of a Higgs boson with a mass spanning up to 1 TeV/$c^2$. As described in Section 1.3.3, current precision electroweak data favour a low mass SM Higgs. The detection effectiveness depends on the production mechanisms and the decay channels of the Higgs. The dominant production mechanisms of the Higgs at the LHC are: the gluon fusion ($gg \rightarrow H^0$ via top or bottom quark loop), Vector Boson Fusion (VBF $qq \rightarrow H^0 qq$, via $W^\pm$ or $Z^0$ fusion), vector boson bremsstrahlung (higgsstrahlung $q\bar{q}' \rightarrow H^0W, Z$) and the associated production with top or bottom quarks ($gg, qq \rightarrow H^0b\bar{b}$ and $gg, qq \rightarrow H^0t\bar{t}$) are shown in Fig. 1.4.

Figure 1.5 shows the cross section of these mechanisms as a function of the Higgs mass. The dominant mechanism is the gluon fusion with the cross section starting from $\sim 30$ pb for a Higgs of 115 GeV/$c^2$ mass. This decreases slowly with increasing Higgs mass, except for a small increase when the Higgs mass exceeds the $2m_t$. 

---

Figure 1.4: Production mechanisms of the Higgs at the LHC.

Figure 1.5: Cross section of Higgs production mechanisms as a function of the Higgs mass.
1.3 Detecting the Higgs Boson

Figure 1.4: Dominant Higgs boson production mechanisms at LHC.

threshold. For small Higgs masses the VBF mechanism is an order of magnitude lower in cross section, but the difference diminishes as the Higgs mass reaches the 1 TeV/c² mass, where they become approximately the same. This mechanism has a clear signature due to the existence of two high energy forward jets which can be used to select signal over background events. The Higgsstrahlung and the associated fermion productions are more important for a low mass Higgs, as they have two to three orders of magnitude smaller cross section with increasing Higgs mass.

The various decay channels of the Higgs boson are shown in Figure 1.6(a). Just above the limit that the LEP has set for the Higgs mass, 114.4 GeV/c², the branching ratios in decreasing order are the decays to: two bottom quarks (b¯b), two W bosons, two taus (τ⁺τ⁻), two charm quarks (c¯c), two guons (gg), two Z bosons and finally two photons (γγ). The branching ratio of the WW channel rises quickly until it reaches the 2m_W resonance and causes the suppression of the competing ZZ channel. Thereon above the 2m_Z resonance, the branching ratio of the ZZ channel increases. Finally when the mass of the Higgs exceeds the threshold of 2m_t, ~350 GeV/c², the branching ratio to t¯t increases rapidly.

Higgs boson decay channels to purely hadronic final states are not preferred for dis-
1.3 Detecting the Higgs Boson

\[ \sigma (pp \to H^+X) [\text{pb}] \]
\[ \sqrt{s} = 14 \text{ TeV} \]
\[ M_t = 174 \text{ GeV} \]

CTEQ6M

\[ \text{gg} \to H \]
\[ \text{qq} \to H \text{qq} \]
\[ \rightarrow \text{HW} \]
\[ \text{qq} \to \text{HZ} \]
\[ \text{gg}, \text{qq} \to \text{Htt} \]
\[ \text{gg}, \text{qq} \to \text{Hbb} \]

\[ M_H [\text{GeV}] \]

0 200 400 600 800 1000

\[ 10^{-4} \]
\[ 10^{-3} \]
\[ 10^{-2} \]
\[ 10^{-1} \]

\[ 1 \]

\[ 10 \]
\[ 2 \]

Figure 1.5: Higgs boson production cross section at LHC as a function of the Higgs mass \cite{5}.

covery, especially at low Higgs masses. This is because of the high cross sections of the Quantum Chromo-dynamic (QCD) processes at the LHC. An additional constraint is the very small natural width of the Higgs at low masses (Fig 1.6(b)), which poses a high resolution requirement. These facts make the Higgs boson’s decays to weak bosons which further decay to leptons \( H \to ZZ \to 4l, H \to WW \to 2l2\nu \) and to \( \gamma\gamma \), significant, despite their smaller branching ratios.

\[ \text{BR}(H) \]
\[ \text{bb} \]
\[ \tau^+ \tau^- \]
\[ \text{cc} \]
\[ \text{gg} \]
\[ \text{WW} \]
\[ \text{ZZ} \]
\[ \text{tt} \]
\[ \gamma\gamma \]

Figure 1.6: Higgs boson branching ratios and natural width as a function of the Higgs mass \cite{5}.
Chapter 2

The CMS Experiment at the LHC

2.1 The Large Hadron Collider

2.1.1 Overview

The Large Hadron Collider (LHC) \[28\] is the highest energy particle accelerator constructed up to date. The project is carried out at the European Organisation for Nuclear Research (CERN) \[29\]. The LHC accelerator ring spans the Swiss-France border a few kilometres outside Geneva and uses a 27 km circumference tunnel built for the LEP experiments. The tunnel for the ring is located 50 to 175 meters below ground level. It is designed to operate as a proton-proton (p⁺-p⁺) collider and a lead-ion (Pb^{82⁺}-Pb^{82⁺}) collider, at 14 TeV and 1,150 TeV centre of mass energy respectively. The accelerated bunches of particles will collide at four Interaction Points (IP), where the main detectors are located: A Large Ion Collider Experiment (ALICE) \[30\], A large Toroidal LHC ApparatuS (ATLAS) \[31\], the Compact Muon Solenoid (CMS) \[32\] and the Large Hadron Collider beauty (LHCb) \[33\] (see Figure 2.1). Two of those, ATLAS and CMS, are general purpose detectors, whereas LHCb and ALICE are dedicated physics detectors. The former measures the matter-antimatter asymmetry using B quarks, whereas the latter investigates the existence of quark-gluon plasmas.
2.1.2 Physics at the LHC

Apart from the Higgs boson searches described in detail in Section 1.3.4, the LHC will be able to probe more areas of particle physics. The production of a large number of B quarks will enable precise measurement of Charge conjugation-Parity violation ($\mathcal{CP}$) in the hadronic sector. This will provide a better understanding of the matter-antimatter asymmetry. Since the energy regime that LHC will operate at will be one order of magnitude higher than that of the current experiments, the collider will test theories extending the SM. One example is SUperSYmmetry (SUSY), which predicts that every SM particle has a corresponding SUSY particle. When the LHC will operate as a lead ion collider, it will assist the search for a new state of matter that is believed to have existed shortly after the Big Bang. In this state, called a “quark-gluon plasma”, the quarks are not confined into hadrons, but move freely.

2.1.3 Operation and Design Parameters

2.1.3.1 LHC and Injector Chain

The particles, $p^+$ or $^{208}\text{Pb}^{82+}$, are accelerated and collide in bunches. The stream of 3564 bunches separated by 25 ns is referred to as the “beam”. Before they reach the LHC acceleration ring, the particle beams have passed through a number of acceleration units (Figure 2.1). The particle beams reach their maximum energy through five stages of acceleration:

- The LINear ACcelerator 2 (LINAC2) \cite{34} is the primary source of protons, whereas LINAC3 is the primary source of the lead ions. The protons are accelerated to 50 MeV and are then injected into the Proton Synchrotron Booster (PSB) \cite{35}, while the lead ions are accelerated to 4.2 MeV/nucleon for injection into the Low Energy Ion Ring (LEIR) \cite{36}.

- Protons and lead ions are accelerated further from the PSB and LEIR to 1.4 GeV and 72.2 MeV/nucleon respectively, before they are injected into the Proton Synchrotron (PS) \cite{37}.
- In the PS, protons are boosted to 25 GeV and lead ions to 5.9 GeV/nucleon and are then injected into the Super Proton Synchrotron (SPS) \[37\].

- In the SPS, protons are accelerated to 450 GeV and the lead ions to 177.4 GeV/nucleon and are then injected into the LHC ring.

- In the LHC, the protons are accelerated to their maximum designed energy of 7 TeV and lead ions are accelerated to 2,759 GeV/nucleon (574 TeV per \(^{208}\text{Pb}^{82+}\)).

The LHC ring comprises \(\sim 9,300\) superconductive magnets. Of these, 1,232 are the main dipole RF magnets (responsible mainly for the acceleration of the beams), 858 are the quadrupole magnets (responsible for focusing) and the remaining magnets are either in charge of correcting the curvature or occupy the short straight sections of the ring. These magnets operate at a range of magnetic fields with the maximum, 8.33 T, used by the dipole magnets. The SPS injects particle beams into the LHC ring in two opposite directions. Each superconducting magnet has two beam pipes so that the two opposing particle beams can be accelerated before they collide. A special design of a triplet of quadrupole magnets has been put into place on both sides of each IP to allow the beams to collide. The superconductivity of the magnets relies on the temperature at which they operate. For this reason, the LHC will operate at 1.9 K. To make this possible, the magnets will be fitted inside a vessel filled with liquid helium.

### 2.1.3.2 Beam Parameters and Performance.

The LHC has been designed to accelerate proton beams to 7 TeV and lead ion beams to 2.76 TeV/nucleon. The parameters of the two beams are summarised in Table 2.1. The number of events, \(N\), generated per second at each IP is given by:

\[
N = \mathcal{L}\sigma
\]  \hspace{1cm} (2.1)
2.1 The Large Hadron Collider

Figure 2.1: Schematic representation of the CERN accelerator complex.

<table>
<thead>
<tr>
<th>Beam Parameters</th>
<th>( p^+p^+ )</th>
<th>( \text{Pb}^{82+}\text{Pb}^{82+} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy per nucleon</td>
<td>[TeV]</td>
<td>7</td>
</tr>
<tr>
<td>Number of bunches</td>
<td>2808</td>
<td>592</td>
</tr>
<tr>
<td>Number of particles per bunch</td>
<td>( 1.15 \times 10^{11} )</td>
<td>( 7.0 \times 10^7 )</td>
</tr>
<tr>
<td>RMS bunch length</td>
<td>[cm]</td>
<td>7.55</td>
</tr>
<tr>
<td>Bunch separation</td>
<td>[ns]</td>
<td>25</td>
</tr>
<tr>
<td>Luminosity</td>
<td>[cm(^{-2}\text{s}^{-1})]</td>
<td>( 1.0 \times 10^{34} )</td>
</tr>
</tbody>
</table>

Table 2.1: The LHC nominal beam parameters for proton and lead ion collisions.

where \( \mathcal{L} \) is the luminosity of the accelerator in cm\(^{-2}\text{s}^{-1}\) and \( \sigma \) is the cross section of the two colliding particles in cm\(^{-2}\). The total cross section at a centre of mass energy of 14 TeV for the \( pp \) collision (Fig. 2.2) is calculated to be \( \sim 110 \text{mb} \). However, many of these collisions are elastic and aren’t registered by the detectors. The inelastic cross section of \( pp \) is \( \sim 60 \text{mb} \). For the lead ion collision at a centre of mass energy of 1,150 TeV, the cross section is \( \sim 514 \text{b} \), not including Coulomb scattering. Given that only 2808 and 592 bunches are filled at any time with protons and lead ions respectively, under nominal operation, it results in \( \sim 19 \text{ pp inelastic collisions or } \sim 0.1 \text{ lead ion interactions per bunch crossing.} \n
The LHC started circulating beams in August 2008, with a public announcement on the 10th of September. It has resumed operation since late 2009, when it started providing the first collision data. On the 30th of March 2010, the LHC started
operating at a centre of mass of 7 TeV. By October 2010 the LHC had delivered approximately $10 \, \text{pb}^{-1}$ of integrated luminosity to the CMS detector with peak instantaneous luminosity of about $50 \times 10^{30} \, \text{cm}^{-2} \text{s}^{-1}$.

2.2 The CMS detector

2.2.1 Overview

The CMS detector \[39\] was designed to meet the challenging conditions created by the LHC. At the design luminosity, the CMS detector will observe $\sim 6.0 \times 10^8$ inelastic events/s with bunch crossings every 25 ns. These give rise to a series of experimental challenges:

1. The large event rate must be reduced to a manageable level of approximately 100 Hz by the trigger (on-line event selection). Clean signatures must be used to separate the signals of interesting physics processes from the background. The best candidates are high transverse momentum ($p_T$) leptons, photons and jets together with large amounts of missing transverse energy ($E_T$).
2. The number of inelastic pp collisions per bunch crossing (∼19) causes mixing of the interaction products under study with other interactions in the same bunch crossing. This problem is called pile-up and becomes more severe if the detector response is longer than 25 ns. The effect of pile-up can be reduced by using highly granular detectors with good time resolution.

3. The short separation of bunch crossings imposes extra limitations on the read-out and trigger systems. It is not feasible to have a trigger decision in the timing between bunch crossings, while a decision is needed for every bunch crossing. Therefore a pipelined trigger and readout system is required which can hold the data while the trigger decision is taken.

4. The large flux of particles coming from the pp collision products imposes a radiation hardness requirement on the detector elements and the associated electronics. This is especially important for the parts of the detector close to the beam pipe and the forward regions.

5. The high energy particles produced also affect the detector design. A large magnetic field is required in order to have precise momentum measurements of the charged particles. In order for high energy particles to deposit all their energy the calorimeters have to provide the appropriate material depth. Such considerations lead to very large detectors.

The overall layout of the CMS detector is shown in Figure 2.3. It consists of (from the beam pipe outward) the pixel detector, the silicon microstrip charge particle tracker, the lead tungstate (PbWO₄) crystal electromagnetic calorimeter, the plastic scintillator/brass hadronic calorimeter, the superconducting solenoid coil and the iron yoke where the muon detectors are placed. The most important feature driving the detector design and layout is the 4 Tesla superconducting solenoid magnet. The large magnetic field allows precise measurement of charged particles’ transverse momenta on the inside, while the return field is used to provide an extra measurement for muons that pass through the calorimeters. The overall dimensions of the CMS detector are 21.6 m in length, 15 m in diameter and 12,500 tons in weight.
The coordinate system used by the CMS detector has its origin centred at the
nominal interaction point. The $y$-axis points vertically upward, the $x$-axis points
towards the centre of LHC and the $z$-axis points along the beam direction towards
the west. Using cylindrical coordinate system conventions, the azimuthal angle $\varphi$
is measured from the $x$-axis in the $x$-$y$ plane and the polar angle $\theta$ is measured from
the $z$-axis. Pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$. Therefore, the transverse
quantities such as $p_T$, $E_T$ and $\not E_T$ are computed using the $x$-$y$ components.

2.2.2 The CMS Silicon Tracker

The CMS detector has at its heart an all-silicon tracking detector. The layout
of the tracker was modified in 1999 [40], replacing the originally proposed Micro
Strip Gas Chambers [41] with extra layers of silicon strip detectors, in order to cope
with the high track multiplicity expected at the high beam luminosities. The current
2.2 The CMS detector

Figure 2.4: A quadrant of the CMS silicon tracker in r-z view.

design, measuring 1.1 m in radius and 5.4 m in length, comprises three layers of pixel detectors and ten layers of microstrip detectors in the barrel, whereas each endcap comprises two pixel disks, three inner and nine outer microstrip disks (see Fig. 2.4). Due to the proximity of the tracker to the beam, it is required to be radiation hard. The expected radiation dose for the lifetime of the detector (500 fb$^{-1}$) at a radius of 4 cm is up to $3.2 \times 10^{15}$ fast hadrons (equivalent to 1 MeV neutron) per cm$^2$.[42]

2.2.2.1 The Pixel Detector

The purpose of the pixel detector (see Fig. 2.5) is to provide a precise measurement of 3D tracking points in $r - \varphi$ and $z$. Therefore it is responsible for good impact parameter measurement and secondary vertex reconstruction. It consists of three concentric cylindrical layers in the barrel and two endcap disks on either side. The barrel layers are placed at mean radii of 4.4 cm, 7.3 cm and 10.2 cm, respectively and are 53 cm long. The endcap disks extend from 6 to 15 cm in radius and are placed on either sides at positions $z=34.5$ cm and 46.5 cm. The pseudorapidity coverage of the pixel detector is up to $|\eta| < 2.5$.

To achieve optimal vertex resolution, a rectangular pixel shape of $100 \times 150 \mu m^2$ was adopted. The Lorentz drift in the 4 Tesla magnetic field leads to the signal charge spreading to more than one pixel. This fact is used to enhance the spatial resolution
by the analogue signal interpolation. Since the Lorentz drift is perpendicular to the
field, the detectors are deliberately not tilted in the barrel layers but are tilted at
20° in the endcap disks. The resulting position resolution is \( \sim 15\, \mu m \) in both \( r - \varphi \)
and \( z \) coordinates and the typical longitudinal vertex position resolution, made up
from pixel triplet hits, is 30-40\( \mu m \).

### 2.2.2.2 The Silicon Strip Detector

The silicon strip tracker is located outside the pixel detector in the radial region
between 20 cm and 116 cm (see Fig. 2.6). It measures 5.6 m in length and comprises
four different subsystems: the Tracker Inner Barrel (TIB), the Tracker Inner Disks
(TID), the Tracker Outer Barrel (TOB) and the Tracker EndCaps (TEC). The TIB
extends from 20 cm to 55 cm in radius with \( |z| < 65 \, \text{cm} \) and is composed of 4 barrel
layers, with silicon strip sensors 320\( \mu m \) thick and 12 cm long, and with a strip pitch
varying from 80\( \mu m \) to 120\( \mu m \). The TID extends, like the TIB, from 20 cm to 55 cm
in radius with \( 70 \, \text{cm} < |z| < 110 \, \text{cm} \) and is composed of 3 disks on either side. It
uses silicon strip sensors which are 320\( \mu m \) in thickness, 12 cm in length and have a
strip pitch varying from 100\( \mu m \) to 140\( \mu m \). The TOB extends from 55 cm to 116 cm
2.2 The CMS detector

<table>
<thead>
<tr>
<th>part</th>
<th>detectors</th>
<th>thickness (µm)</th>
<th>average pitch (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIB (2 inner/2 outer layers)</td>
<td>2724</td>
<td>320</td>
<td>81/118</td>
</tr>
<tr>
<td>TID (1/2/3 rings)</td>
<td>816</td>
<td>300</td>
<td>97/128/143</td>
</tr>
<tr>
<td>TOB (4 inner/2 outer layers)</td>
<td>5208</td>
<td>500</td>
<td>183/123</td>
</tr>
<tr>
<td>TEC (4 inner rings)</td>
<td>2512</td>
<td>300</td>
<td>96/126/128/143</td>
</tr>
<tr>
<td>TEC (3 outer rings)</td>
<td>3888</td>
<td>500</td>
<td>143/158/183</td>
</tr>
</tbody>
</table>

Table 2.2: Summary of the different charged tracker module parameters.

in radius with $|z| < 118$ cm and is composed of 6 barrel layers. It uses silicon strip sensors which are 500 µm in thickness, 9 cm in length and have a strip pitch varying from 120 µm to 180 µm. Finally, the TEC extends from 25 cm to 110 cm in radius with 120 cm < $|z| < 280$ cm and is composed of 9 disks, carrying 7 rings of silicon detectors on each side with thicknesses 320 µm (4 inner rings) and 500 µm (rings 5-7) and a strip pitch varying from 100 µm to 180 µm. See Table 2.2 for a summary.

Additionally the first two layers or disks of TIB, TID and TOB as well as rings 1, 2 and 5 of the TEC carry a second strip detector module which is mounted back-to-back with a stereo angle of 100 mrad in order to provide a measurement of the third coordinate ($z$ in the barrel and $r$ on the disks). The achieved position resolution is between 23-34 µm in $r - \varphi$ and 230 µm in $z$ in the TIB, 35-52 µm in $r - \varphi$ and 530 µm in $z$ in the TOB and it varies with pitch in the TID and TEC. The tracker layout ensures more that 8 hits in the silicon strip tracker in the full range of $|\eta| < 2.4$ with at least ~4 of them coming from the “stereo” modules.

The performance of the tracker (combining pixel and strip detectors) for muons with $p_T$ 10 to 100 GeV/c gives a reconstruction efficiency >95% with a $p_T$ resolution of 1-2% and transverse and longitudinal impact parameters of 10-30 µm and 10-100 µm, respectively. The material budget of the CMS tracker in units of radiation length is shown in Figure 2.7. It increases from 0.4 $X_0$ at $\eta = 0$ to about 1.8 $X_0$ at $|\eta| \sim 1.4$, beyond which it falls to about 1 $X_0$ at $|\eta| \sim 2.5$.

2.2.3 The Electromagnetic calorimeter

The CMS Electromagnetic calorimeter (ECAL) is a homogeneous scintillating calorimeter made out of lead-tungstate (PbWO₄) crystals. Its purpose is to measure
2.2 The CMS detector

Figure 2.6: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits.

Figure 2.7: Material budget of the CMS tracker in units of radiation length as a function of $\eta$ for the different sub-detectors.

the energy of electromagnetic objects, photons and electrons, as well as the electromagnetic component of the jets. The ECAL was designed to be able to detect the Higgs boson through the $H \rightarrow \gamma \gamma$ decay channel. This requires a very granular detector with excellent resolution.

The physical properties of PbWO$_4$ [38] make it an appropriate choice for the CMS ECAL. Its high density (8.3 g/cm$^3$) results in a short radiation length (0.89 cm) and
a small Molière radius (2.0 cm), which in turn allows for a compact and granular calorimeter. Furthermore, the PbWO$_4$ crystals are radiation hard (up to 10 Mrad) and have a fast response - about 80% of the light emitted from the scintillation decay is within the LHC bunch crossing time (25 ns). The relatively low light output (30 $\gamma$/MeV) requires the use of photo-detectors with internal gain that can operate in a high magnetic field and high radiation environment. The configuration of the magnetic field and the expected level of radiation lead to different choices - Avalanche Photo-Diodes (APD) in the barrel and Vacuum Photo-Triodes (VPT) in the endcaps. The larger surface coverage of the VPT on the back of the endcap crystals compensates for the lower internal gain compared to the APD. Since the light yield of the crystals and the amplification of the APD are both temperature dependent, the system’s temperature has to be kept constant to high precision. The nominal operating temperature of the CMS ECAL is designed to be 18.0±0.1°C.

The ECAL barrel (EB) extends from 1.29 m to 1.75 m in radius and is 6.1 m long, covering the pseudorapidity range of $|\eta| < 1.479$ (see Figure 2.8). It is formed by 360 crystals in $\varphi$ and (2×85) crystals in $\eta$, resulting in a total of 61200 PbWO$_4$ crystals. To avoid particle trajectories aligning with the ECAL cracks, the crystals are mounted at a 3° angle with respect to the vector from the nominal interaction vertex to the crystal surface. Each crystal approximately corresponds to 0.0174×0.0174 rads, in $\Delta\eta \times \Delta\varphi$, resulting in 22×22 mm$^2$ at the front surface and 26×26 mm$^2$ at the rear. The crystal is 230 mm long corresponding to a radiation length of 25.8 $X_0$.

Each ECAL endcap extends from 3.15 m to 3.9 m in $z$, covering the pseudorapidity range 1.479 < $|\eta|$ < 3.0 (see Figure 2.8) and consists of 7324 crystals. The crystals are arranged in an $x$-$y$ configuration, all pointing at $|z|=1.3$ m from the nominal interaction vertex, resulting in off pointing angles ranging from 2 to 8 degrees. The front surface of the crystals is 28.6×28.6 mm$^2$ and the rear 30×30 mm$^2$, while their length is 220 mm (24.7 $X_0$). In front of each endcap crystal calorimeter there is a 20 cm thick preshower detector covering the pseudorapidity range 1.653 < $|\eta|$ < 2.6 (see Figure 2.8). The preshower is a sampling calorimeter which consists of two disks of lead absorbers, 2 $X_0$ and 1 $X_0$ respectively, followed by two layers of silicon...
2.2 The CMS detector

Figure 2.8: Transverse quadrant section of ECAL. The barrel extends from 1.29 to 1.75 meters from the interaction point and the endcap with the preshower module from 3.0 to 3.9 meters. The dashed lines show the coverage in $\eta$.

The principal aim of the preshower is to identify neutral pions by providing better spatial resolution for the pairs of photons they decay into.

The resolution of the ECAL can be determined by the width of the gaussian distribution parametrised as a function of the reconstructed energy according to Equation 2.2:

$$\left( \frac{\sigma}{E} \right)^2 = \left( \frac{S}{\sqrt{E}} \right)^2 + \left( \frac{N}{E} \right)^2 + C^2$$  \hspace{1cm} (2.2)

The $S$ term is called the stochastic term, arising mainly from fluctuations in the lateral shower containment, photostatistics and the fluctuations in the energy deposited in the preshower absorber (where present) with respect to that measured in the silicon detector. The $N$ term is called the noise term and it accounts for the noise in the electronics, the digitisation, as well as the noise due to pile-up. The $C$ term is called the constant term and it accounts for the non-uniformity of the longitudinal light collection, intercalibration errors and energy leakage from the back of the crystal. The ECAL energy resolution has been measured from test-beam results. Figure 2.9 shows the ECAL energy resolution as a function of electron energy.
measured by summing a $3 \times 3$ crystal array around the crystal, with the electron impacting within $4 \times 4$ mm of its centre. Equation 2.2 was used to fit the data giving values for $S=2.8\% \text{GeV}^2$, $N=0.12 \text{GeV}$ and $C=0.3\%$.

### 2.2.4 The Hadronic calorimeter

The CMS Hadronic calorimeter (HCAL) \cite{44} consists of a sampling calorimeter which uses brass as the absorber material and plastic scintillator as an active material in the barrel, endcaps ($|\eta|<3$) and a Cerenkov-based steel/quartz fibre calorimeter in the forward region ($3<|\eta|<5$). The purpose of the HCAL is to measure the energy of hadronic jets, while an asymmetry in the energy measured in the transverse plane signifies the existence of neutrinos or other exotic particles. The design of the HCAL is strongly influenced by physical constraints. The fact that the HCAL barrel is located between the ECAL (1.77 m outer radius) and the magnet coil (2.95 m inner radius) limits the total amount of material that can be used as an absorber. Furthermore, the HCAL has to operate inside the magnetic field. Combining the two, the absorber is required to be dense and non-magnetic - hence the choice of brass.
The HCAL barrel (HB) covers the pseudorapidity range $|\eta| < 1.4$ (see Figure 2.10). It consists of 36 identical azimuthal wedges, divided equally in the positive and negative $z$ direction. Each wedge has 4 segmentations in $\varphi$ and 16 segmentations in $\eta$, resulting in $0.087 \times 0.087$ rads ($\Delta \eta \times \Delta \varphi$) and a total of 2,304 HB towers. The tower absorber is formed by two steel plates enclosing 14 brass plates of about 5 cm thickness. The tower active material consists of 17 layers of plastic scintillator, with the first located in front of the steel plate sampling showers that develop in the inert material between EB and HB. The HB towers have single longitudinal read-out, with the exception of the towers close to the endcap transition which have two. The effective thickness of HB, measured in interaction lengths ($\lambda_I$), increases from $5.82 \lambda_I$ at $\eta = 0$ to $10.6 \lambda_I$ at $\eta = 1.3$, while the ECAL adds about $1.1 \lambda_I$.

The HCAL endcaps (HE) cover the pseudorapidity range $1.3 < |\eta| < 3.0$ (see Figure 2.10). Each HE has 14 segmentations in $\eta$. The first 5 (lower $\eta$) towers have 72 segmentations in $\varphi$ and the remaining 9 have 36 segmentations in $\varphi$, resulting in $\sim 0.087 \times 0.087$ rads ($\Delta \eta \times \Delta \varphi$) for the lower $\eta$ region and $\sim 0.17 \times 0.17$ rads for the higher $\eta$ region. The total number of HE towers is 1368. The HE towers are formed by 79 mm thick brass absorber plates with 9 mm thick gaps for the plastic scintillators. The effective thickness of HE including the ECAL is about $10 \lambda_I$.

The effective thickness of EB and HB is not sufficient to contain the central hadronic showers. In order to solve this problem, an extra sampling calorimeter - the Outer hadronic calorimeter (HO) - is placed outside the magnet solenoid, covering the pseudorapidity range $|\eta| < 1.26$ (see Figure 2.10). It comprises five rings of active material along $\eta$, matching the HB segmentation, $0.087 \times 0.087$ rads in $\Delta \eta \times \Delta \varphi$. The HO makes use of the solenoid coil as an absorber equal to $1.4 / \sin(\theta) \lambda_I$ for all five rings, while the central ring has a second layer of plastic scintillators located on the other side of a 19.5 cm thick iron piece (part of the magnet return yoke). The total effective thickness of the calorimeter is thus extended to a minimum of $11.8 \lambda_I$.

The HCAL forward calorimeter (HF) is located at $z = \pm 11.2$ m covering the pseudorapidity range $3.0 < |\eta| < 5.0$. The design of the HF is motivated by the large particle fluxes which deposit an average of 760 GeV per $pp$ interaction into the two
2.2 The CMS detector

HF calorimeters as opposed to 100 GeV deposited into the rest of the detector. The extreme conditions that the calorimeter will be exposed to were the reason that quartz fibres were chosen as the active material. The signal is generated when charged particles passing through the steel absorber generate Cerenkov light, which makes the HF more sensitive to the electromagnetic component of the showers. The fibres are positioned parallel to the beam line inside the 5 mm thick holes that are grooved in the steel absorber, and are grouped to form $0.175 \times 0.175$ rads ($\Delta \eta \times \Delta \phi$) towers.

The CMS HCAL is a non-compensating calorimeter resulting in a non-linear energy response. The resolution of the HCAL+ECAL combined calorimeter as a function of energy is described by Equation 2.3:

$$\left( \frac{\sigma}{E} \right)^2 = \left( \frac{\alpha}{\sqrt{E}} \right)^2 + \beta^2 \quad (2.3)$$

where the $\alpha$ and $\beta$ terms correspond to the $S$ and $C$ of Equation 2.2, respectively.

The HCAL+ECAL energy response as a function of charged pion momentum is shown in Figure 2.11. Fitting the data with Equation 2.3 results in $\alpha = 1.21 \text{GeV}^{1/2}$ and $\beta = 0.095$ [45].
2.2 The CMS detector

2.2.5 The Superconducting Magnet

The CMS superconducting magnet [46] has been designed to provide sufficient bending power both inside and outside of the magnet coil. The choice of a high magnetic field, 4 T, is dictated by the physics requirement to determine accurately the sign of muons with momenta up to 1 TeV/c, which results in a momentum resolution of $\Delta p/p \sim 10\%$ at $p = 1$ TeV/c.

The superconducting magnet extends from 2.95 m to 3.15 m in radius and is 12.5 m in length. The superconducting coil is composed of a four layer winding made of reinforced conductor. The conductor is made from a Niobium-Titanium (NbTi) Rutherford-type cable co-extruded with high purity aluminium and mechanically reinforced with an aluminium alloy. The parameters of the coil are summarised in Table 2.3. The field is returned by the iron yoke which comprises 5 barrel wheels and 6 endcap disks.
2.2 The CMS detector

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnetic field</td>
<td>4 T</td>
</tr>
<tr>
<td>Inner bore</td>
<td>5.9 m</td>
</tr>
<tr>
<td>Magnetic length</td>
<td>12.5 m</td>
</tr>
<tr>
<td>Nominal Current</td>
<td>19.14 kA</td>
</tr>
<tr>
<td>Stored Energy</td>
<td>2.6 GJ</td>
</tr>
</tbody>
</table>

Table 2.3: Main parameters of the CMS superconducting magnet.

2.2.6 The Muon System

The CMS muon detection system [47] consists of three different types of gaseous detectors and is placed in the magnet return yoke. The primary functions of the muon detection system are to identify muons, provide a momentum measurement and to trigger. The choice of the muon detectors is driven by the large area that they cover and the radiation environment.

The Drift Tube system (DT) is located on the barrel section of the CMS detector covering the pseudorapidity range $|\eta| < 1.2$ (see Figure 2.12). The low rate expected, both from prompt muons and neutron-induced background, and the uniformity of the magnetic field make DTs a good choice for the barrel. The DT system consists of 4 concentric cylindrical layers, called “stations”, around the beam line divided over the five iron barrel wheels. Every barrel wheel is divided into 12 sectors, each covering $30^\circ$ azimuthal angle ($\phi$). One DT chamber is fitted in every sector of the three inner stations, while the top and the bottom sectors of the outer station are fitted with an additional DT chamber. This results in a total of 250 DT chambers arranged such that high $p_T$ muons cross at least 3 out of the 4 stations. Each DT chamber of the three inner stations is composed of 3 superlayers of drift tubes - 2 to measure the $\phi$ coordinate (outside) and one to measure $z$ (inside) - while the DT chambers of the outer station have only 2 superlayers for the $\phi$ coordinate measurement. Each superlayer consists of four layers of drift tubes. The resolution for a single point is $\sim 200\mu m$ and $\sim 1$ mrad in the $\phi$ direction.

The Cathode Strip Chamber system (CSC) is located in the iron disks perpendicular to the beamline, covering the pseudorapidity range $0.9 < |\eta| < 2.4$ (see Figure 2.12). The higher background rates and the non-uniform magnetic field favour the use of
CSCs in the endcap region. The CSC system consists of four stations at each end. The inner CSC station is composed of three rings of chambers, while the three outer stations are composed of two rings each. All rings of the first disk and the outer rings of the other three consist of 72 chambers, while the inner rings consist of 36 chambers, resulting in a total of 540 chambers at each cap. The outer disk will have only the inner ring installed during the early years of CMS operation. The chambers are trapezoidal in shape and each consists of six active gas layers. The single point resolution for each chamber is $\sim 200 \mu m$ and $\sim 10 \text{ mrad}$ in the $\phi$ direction.

The Resistive Plate Chambers system (RPC) is located both on the iron barrel wheels and the endcap disks, covering the pseudorapidity range $|\eta| < 2.1$. However in the beginning of CMS operation, only the RPC up to $|\eta| = 1.6$ will be installed (see Figure 2.12). The fast response of the RPCs and their large coverage makes them ideal for the triggering process. The RPCs also help to resolve ambiguities in DT and CSC track reconstruction from multiple hits in a chamber. The RPC system in the barrel consists of six radial stations placed within the same five wheels as the DT system - two for the inner two stations and one for each the outer ones. At the endcap it consists of a plane in front of each of the outer three CSC stations.
2.2 The CMS detector

The CMS detector

\[ \frac{\Delta p_T}{p_T} \]

Muon system only
Full system
Inner tracker only

(a) Muon resolution for \(|\eta| < 0.8\).

(b) Muon resolution for \(1.2 < |\eta| < 2.4\).

Figure 2.13: The muon transverse momentum resolution as a function of the muon transverse momentum \((p_T)\) using the muon system only, the inner tracking only and both systems.

The muons produced in the central region of the detector can be measured independently in the silicon tracker and the muon systems. The muon momentum measurement using only the muon system is determined by the bending angle at the exit of the magnet coil and the interaction point, which is known with high accuracy (see 2.2.2.1). The resolution of this momentum measurement, from the muon system only, is dominated by the multiple scattering process in the detector material before the inner muon station for \(p_T^{\mu} < 200\) GeV/c. For higher muon momentum the chamber spatial resolution becomes dominant. The silicon tracker provides a better resolution measurement for low momentum muons. As Figure 2.13 demonstrates, the combination of silicon tracker and muon system can improve the momentum resolution even further, especially for high \(p_T\) muons.
2.2 The CMS detector

2.2.7 The Trigger and Data Acquisition Systems

As discussed in Section 2.1.3.2, the LHC will produce a high rate of interactions, \( \sim 10^9 \) per second at nominal luminosity, but only about 100 crossings per second can be stored in archival media. This limitation is imposed by the current technology for transferring and storing data, subsequently imposing a need for event reduction by at least a factor of \( 10^6 \). Since the events produced are going to be predominantly QCD processes, a selection mechanism is essential to enhance the acquisition efficiency of other physics processes (see Figure 2.14 for relative cross sections as a function of jet \( E_T \) or particle mass). This task is performed by the trigger system, which is the first step of the physics event selection process. The rate reduction is achieved in two stages - the Level-1 (L1) Trigger and the High-Level Trigger (HLT). The L1 trigger is formed by on-detector electronics and is responsible for reducing the rate to about 100 kHz. The information that it uses to form physics objects like electrons, muons and jets, is limited to coarse measurements from the calorimeters and the muon system. If any of the conditions are satisfied, the information is then passed to the HLT, which has access to the complete data read-out and can execute complex algorithms to reduce the rate by a further factor of \( 10^3 \).

2.2.7.1 The Level-1 Trigger

The Level-1 trigger [48] has to reduce the rate to 100 kHz and needs to reach a decision at the LHC bunch crossings rate, i.e. 40 MHz. The size of the tracker buffer was designed to be able to keep information from the last 192 bunch crossings. This constrains the maximum available time for the data transit and L1 decision to \( \sim 4 \mu s \). For this reason the L1 trigger has to be on the detector and it consists of custom made electronic boards implemented mainly in Field Programmable Gate Array (FPGA) technology.

The L1 trigger system is divided into three major subsystems: the L1 calorimeter trigger, the L1 muon trigger and the L1 global trigger. Figure 2.15 shows a schematic representation of the data flow in the L1 trigger. The decision is formed
Figure 2.14: Inclusive proton-proton cross sections for basic physics processes. Interaction rates for the nominal luminosity are given on the right hand scale. The L1 and HLT input/output rates are also denoted in the graph. For low particle mass the processes are described primarily by the SM, while the processes and particles on the right are predicted by more exotic models (such as SUSY) [48].
using coarsely reconstructed physics objects, such as photons/electrons, muons and jets, which are called trigger primitives. During the first step the Regional Calorimeter Trigger (RCT) sums the energy of $3 \times 3$ trigger towers to form photon/electron candidates, while a $4 \times 4$ trigger region array is used as the building block for the jets. A trigger tower comprises an ECAL $5 \times 5$ crystal array and an HCAL tower, while the trigger region consists of a $3 \times 3$ trigger tower matrix. The objects are created along with information about isolation and the lateral shape of the electromagnetic shower. These objects are then passed to the Global Calorimeter Trigger (GCT) which sorts them according to $E_T$ and creates a list of a maximum of 4 isolated and 4 non-isolated photon/electron candidates, 4 central and 4 forward jets, 4 $\tau$ jets, together with missing $E_T$ and the sum of all HCAL towers ($H_T$). Similarly, the information from the three muon systems is used to form muon candidates on a local level, before the information is passed to the Global Muon Trigger (GMT). The purpose of the GMT is to improve trigger efficiency, reduce trigger rates and suppress background by making use of the complementarity and redundancy of the three muon systems. The Global Trigger (GT) is responsible for making the final decision based on combinations of objects that cannot be done at the subsystem level. Once the decision is reached, the GT contacts the Data Acquisition system
(DAQ) to read-out the high resolution data that are held in the pipelined memories, or dump them accordingly. The GT decision is transmitted to the Timing Trigger and Control system (TTC), which is responsible for synchronising the decision between the different subdetectors.

The output of the L1 trigger is compressed into a series of boolean decisions corresponding to the different criteria and thresholds, together with the objects satisfying them. This information is then passed to the High Level Trigger for further processing, using the more detailed detector information.

### 2.2.7.2 The High Level Trigger

The High Level Trigger\footnote{49} is responsible for reducing the rate to approximately 100 Hz using all detector information and complex algorithms, in order to select the events to be stored for further physics analysis. In contrast to the L1, the HLT is executed in commercial computing systems consisting of about 1000 CPU’s, allowing the HLT to benefit from the evolution in computer technology. The maximum available time for the HLT decision is $\sim 40$ ms at start-up and will increase to $\sim 1$ s for the nominal luminosity runs.

The HLT is organised in a series of algorithmic sequences optimised to perform the more time consuming operations at the end. For this reason the reconstruction of the various objects happens in stages with intermediate trigger decisions. Once a decision is negative, the remainder of the sequence is skipped. The large time needed to unpack and process the silicon tracker and pixel information causes a division, mostly logical, of the HLT sequence. At the first stage, called L2, information from only the calorimeter and muon systems is used to form photon/electron candidates (ECAL clusters), jets and muons. At the next stage, called L3, tracker information matching the L2 objects topologically is unpacked and tracks are reconstructed. When the algorithms require it, pixel information is unpacked to form primary vertices and pixel track reconstruction is performed between L2 and L3, called L2.5.
The HLT algorithmic sequences are implemented using the same reconstruction software as used for offline analysis (CMSSW). This ensures compatibility of the algorithms and reduces the complexity of the system.

### 2.3 The CMS Computing Model

The unprecedented amount of data that will be produced by the LHC experiments cannot be transferred using conventional network technology. This fact was understood from the early stages of the LHC development and the LHC Computing Grid (LCG) [50] project was initiated. CMS will be using Grid computing resources, services and toolkits as its base for data transfer and analysis.

The CMS Computing Model [51] makes use of the hierarchy of computing Tiers as has been proposed by the M0dels of Networked Analysis at Regional Centres working group (MONARC) [52]. This model is composed of a Tier-0 computing centre at CERN, which will be directly connected to the experiment for initial processing and data archiving. Considerable processing will be done at Tier-1 centres and analysis will be done at Tier-2 centres, all interconnected using Grid technology.

Explicitly, the CMS computing model (see Figure 2.16) consists of:
2.3 The CMS Computing Model

- The Tier-0 centre at CERN is responsible for the safe-keeping of the first copy of the RAW experimental data. The first reconstruction pass will be made and stored there. The RAW and the reconstructed datasets will be distributed from the Tier-0 to the Tier-1 centres which will together hold a second copy. Finally the Tier-0 will reprocess the data during LHC down-times.

- There are 7 expected Tier-1 centres. Each of these centres is responsible for the safe-keeping of a share of the second copy of the RAW and the reconstructed data. Products of large scale data reprocessing will be held there. Lastly, they are responsible for the distribution of these data products to Tier-2 centres together with the safe-keeping of the simulated data produced at these centres.

- Tier-2 centres are expected to number about 50. Each centre is responsible for holding data to serve analysis.
Chapter 3

Trigger and Offline Analysis Objects

The information collected from the different CMS sub-detectors needs to be combined in order to understand the outcome of the $pp$ collision taking place at the centre of the CMS detector. The trigger objects as well as the physics analysis objects formed by the several reconstruction algorithms correspond to different particles and physics objects. In this chapter the objects that were used in the rest of this thesis will be described together with some of the baseline selection criteria.

3.1 Electrons

Electrons in CMS deposit most of their energy in the ECAL calorimeter. In order to discriminate them from photons, a charged track associated with the ECAL energy deposit is required. The different reconstruction algorithms in the triggers and the offline process will be presented in this section.

3.1.1 L1 $e/\gamma$ Trigger Object

Due to the lack of tracker information in the L1 trigger [48], photons and electrons cannot be discriminated. Therefore, a single algorithm is used to form electromagnetic trigger objects. The basis of the $e/\gamma$ trigger object is a sliding $3\times3$ trigger
3.1 Electrons

Figure 3.1: Schematic representation of the L1 e/γ isolation and lateral shower shape algorithms.

tower ($\Delta \eta \times \Delta \phi = 0.087 \times 0.087$ rads for each trigger tower) window spanning across the complete $\eta, \phi$ coverage of the CMS ECAL. Initially, the $3 \times 3$ trigger tower window is centred around the ECAL trigger cells with the largest energy deposits. The HCAL tower right behind the ECAL cells and the remaining 8 neighbouring trigger towers are used to check the lateral shower shape and isolation, as well as to reject hadronic particles.

A graphical representation of the algorithm is shown in Figure 3.1. The Fine Grain (FG) algorithm provides information that reflects the lateral shape of the electromagnetic shower. Electrons and converted photons are expected to lose energy due to bremsstrahlung radiation in the $\phi$ direction. Therefore, a strip of $2 \times 5$ ($\eta \times \phi$) ECAL crystals is required to contain $\gtrsim 90\%$ of the sum of the $5 \times 5$ trigger tower crystals. If this is not the case, the FG veto bit is activated and the $e/\gamma$ object is flagged as non electromagnetic. In addition, the longitudinal shower profile is checked by calculating the ratio of the $E_T$ deposits in the HCAL to the ECAL (H/E) portions of the calorimeter in the central trigger tower of the $3 \times 3$ window. The HCAL to ECAL ratio veto is activated for H/E greater than 5%. The isolated $e/\gamma$ trigger candidates need to pass additional vetos. The first is based on all 8 neighbouring trigger towers passing the FG and the H/E vetos and the second is based of the $E_T$ sum of the “L” shape strips shown in Figure 3.1 formed by each 5 tower corner. This is called the Isolation Veto (IV) which is activated when none of these “L” shape strips are below the programmable 1.5 GeV threshold. The $e/\gamma$
trigger candidate $E_T$ is calculated as the sum of the central tower and the highest tower in $E_T$ of the four adjacent neighbouring trigger towers.

### 3.1.2 HLT Electron

The HLT electron reconstruction starts only when the L1 e/γ trigger object requirements are satisfied. To minimise the reconstruction time, the L1 e/γ trigger objects are used to “seed” the ECAL regional readout. The selection criteria discussed in this section were developed and tuned for an ideally aligned CMS detector in the $10^{32} \text{cm}^{-2}\text{s}^{-1}$ LHC luminosity regime. Most of the selection criteria should remain unchanged for higher luminosities except for the $E_T$ selection which will be discussed later (see Section 5.2.1). To collect the energy deposited in the ECAL, two different clustering algorithms are used - the Hybrid and the Island algorithms for the barrel and the endcap regions, respectively.

#### The Hybrid Algorithm

The Hybrid clustering algorithm \cite{hybrid_algorithm} makes use of the lateral shower shape knowledge in the $\eta$ and $\phi$ directions in the barrel. In the $\eta$ direction, a domino of $1 \times 3$ or $1 \times 5$ ECAL crystals is used, while the energy lost due to bremsstrahlung radiation is collected by dynamically searching for crystals with a high energy deposition in the $\phi$ direction (see Figure 3.2).

![Figure 3.2: Schematic representation of the Hybrid clustering algorithm. \cite{hybrid_algorithm}.](image-url)
The Hybrid clustering algorithm starts by finding the highest energy deposit crystals and forming seeds which should satisfy $E_{\text{seed}}^{T} > 1.5 \text{ GeV}$. By default $1\times3$ crystal dominoes in the $\eta$ direction are assigned to the cluster with the centre of the domino being aligned in $\eta$ with the seed crystal. A quantity called $E_{\text{wing}}$ is used to judge the energy containment of the dominoes. When the domino’s central crystal exceeds the value $E_{\text{wing}} > 1 \text{ GeV}$, then a $1\times5$ crystal domino is formed. The default value of steps in $\phi$ is 10 in both directions, but dominoes which contain energy less than $E_{\text{thres}} > 0.1 \text{ GeV}$ are neglected. Additional dominoes are clustered in $\phi$, provided that the deposited energy in the seed of the domino is larger than $E_{\text{seed}} > 0.35 \text{ GeV}$.

**The Island Algorithm**

The Island clustering algorithm \[53\] is used in the endcap region and it starts by finding seed crystals which satisfy $E_{\text{seed}}^{T} > 0.18 \text{ GeV}$. A list of seeds is formed by removing the ones that are adjacent to higher energy seeds.

Beginning with the most energetic seed, the algorithm assigns crystals to the cluster in both directions of $\phi$ and stops if a higher energy or “zero-suppressed” crystal is found. The algorithm then makes a step in $\eta$ and repeats the search in $\phi$. The assignment of crystals in the $\eta$ direction stops if a higher energy or “zero-suppressed” crystal is found. The $\eta$ search is performed in both directions. The Island algorithm
3.1 Electrons

is illustrated graphically in Figure 3.3. All crystals assigned to a cluster are removed from the list of available crystals and they cannot be assigned to a different cluster. This implies that if a seed crystal is used in a cluster it won’t be used to seed another. This way the problem of crystal energy double counting is avoided.

Energy clustering using the Island algorithm has certain advantages. Firstly, the energy from electromagnetic showers formed by electrons and their soft bremsstrahlung radiated photons will not split if the showers are close enough. On the other hand, showers formed by photons coming from a $\pi^0$ decay will be separated provided that the opening angle is large enough. Finally, crystals with an energy deposit lower than the seed threshold, caused by noise or low energy pile-up events, will not be clustered.

In order to collect the bremsstrahlung energy radiated outside of the main shower, a cluster of clusters (“super-cluster”) is formed within a windows of $0.14 \times 0.4$ rads in $\eta \times \varphi$. The effect of the formation of super-clusters is a significant reduction of the negative tail and is illustrated in Figure 3.3.

HCAL Isolation

In order to discriminate electrons from hadronic particles, a minimum HCAL activity behind the electron candidate is required. The ratio of HCAL energy over ECAL energy is set to be less than 0.05 or the sum of the HCAL towers’ $E_T$ within an $\eta$-$\varphi$ cone of 0.15 around the super-cluster position has to be less than 3 GeV.

Track Reconstruction

In order to distinguish an electron from a photon, a charged track pointing to the super-cluster is required. The position of the super-cluster is calculated by an energy weighted average of the individual crystal positions and is propagated back to the pixel layers taking into account the magnetic field and the beam spot in both the positive and the negative charge scenarios. The energy weighted average position of the super-cluster has the advantage that it accounts for the bremsstrahlung radiated
energy and therefore the propagation to the pixel layers remains unaffected by it. The first pixel hit is searched in a $\Delta \varphi_1$ window of 40 mrad, $\pm 25$ mrad$\mp 15$ mrad for positively and negatively charged candidates, respectively. Once this is found, a trajectory based on the momentum calculated by the ECAL super-cluster and the position of the first pixel hit is estimated for both charge scenarios. This trajectory is then used to search for a second compatible hit in a tighter window ($\Delta \varphi_2 = \pm 1$ mrad, $\Delta z_2 = \pm 0.05$ cm). Two compatible hits in the three pixel layers are required in the barrel. In the endcap, to account for the reduced silicon pixel detector coverage, the second compatible hit can be searched for in the first silicon strip detector layer. Once two compatible hits are found, they are used as the seeds for the reconstruction of the track based on the Kalman Filter (KF) \cite{54} algorithm. The main reason that the KF algorithm is used for the HLT instead of the Gaussian Sum Filter (GSF) \cite{53}, which is used for the offline electron, is the faster execution time.

### Charged Tracker Isolation and Identification

In order to further reduce the QCD particles “faking” electrons, a few tracker based isolation and identification selection criteria are applied to the basic electron candidate. The sum of the $p_T$ of all the tracks above 1.5 GeV within a $\eta$-$\varphi$ annulus between 0.02 and 0.2 around the matched track is checked. This sum is required to be equal to or less than 0.06 times the $p_T$ of the matched track. As an identification criterion, a quantity defined as the difference between the inverse super-cluster energy and the inverse matched charged track momentum ($1/E - 1/p$) is checked. This was found to perform better than the usual $E/p$ quantity. In this trigger configuration, a loose threshold value for the selection is set to $1/E - 1/p < 999$.

### 3.1.3 Offline Electron

The offline electron reconstruction shares many similarities with the HLT. However, the execution time of the algorithms is not a limiting factor any more, hence some algorithms are modified while some extra selection criteria are applied. The areas where these changes occur are: the clustering algorithms, the charged track reconstruction algorithm, the ECAL isolation and the electron identification criteria.
In contrast to the HLT electron reconstruction, the algorithm used for the offline endcap clustering is the Multi5×5 [55]. This is the equivalent of the Hybrid algorithm for the barrel, which cannot be applied due to the different geometry. The seeds are listed as described earlier in this chapter. For a crystal to be considered as a seed, it must satisfy $E_{\text{seed}}^{T} > 0.18 \text{ GeV}$. The clustering starts with the highest $E_{T}$ seed, where a 5×5 crystal array is assigned to it. The 8 neighbouring crystals to the seed cannot be used to form another cluster, although the 16 outer crystals can, if they satisfy certain conditions. Each of the outer crystals is checked against the $E_{\text{seed}}^{T}$ threshold of 0.18 GeV, while it is compared with each four adjacent neighbours to establish if it is a local maximum. If both conditions are satisfied, a new 5×5 crystal array is formed assigning the crystals that are not already allocated to a cluster. The Multi5×5 algorithm is illustrated graphically in Figure 3.4. To recover the energy lost due to bremsstrahlung radiation, a rectangular window of $\Delta\eta \times \Delta\varphi = 0.14 \times 0.6 \text{ rads}$ is formed around every seed in decreasing energy order and the clusters that fall inside are combined into a super-cluster. Each cluster can be assigned to only one super-cluster.

The Offline Hybrid Algorithm Selections

While the algorithm remains unchanged (see Section 3.1.2), the values of the selection criteria are adjusted. The seed threshold is $E_{T} = 1 \text{ GeV}$, the $E_{\text{wing}}$ threshold is,
by default, 0 GeV ensuring that all dominoes are 1×5 size and the default value of \( \varphi \) steps in both directions becomes 17.

### The Offline Track Reconstruction

The search for pixel seeds is done in the same way as described in Section 3.1.2. The search ranges are slightly modified. The \( \Delta \varphi \) window for the first pixel becomes 200 mrad (±125 mrad–±75 mrad for positively and negatively charged candidates, respectively), the \( \Delta \varphi \) window for the second pixel becomes ±2 mrad and the \( \Delta z \) windows becomes ±0.09 cm.

These matched pixels are used again to seed the reconstruction of a charged track this time based on the Gaussian Sum Filter (GSF) algorithm. The advantage of the GSF algorithm is that it can account for the non-Gaussian track propagation caused by the bremsstrahlung radiation emission [56]. This fact ensures that the track can be reconstructed up to the outer layers of the silicon tracker.

### Isolation and Identification

The charged tracker and HCAL isolations are performed in a similar manner to the HLT. An ECAL isolation is introduced at the offline level. To ensure that the energy of the electron is excluded from the calculation, crystals within the two different shape regions are formed around the most energetic cluster of the super-cluster and excluded. First there is a rectangular strip in \( \varphi \) of a \( \Delta \eta \) width and then a cone in \( \eta - \varphi \).

The identification criteria are mainly applied at the offline reconstruction level. A \( \Delta \eta, \Delta \varphi \) test is performed between the calculated super-cluster \( \eta, \varphi \) position and the \( \eta, \varphi \) values given by the inner part of the matched track (\( \Delta \eta_{\text{in}}, \Delta \varphi_{\text{in}} \)). A few lateral shower shape checks are performed. The lateral shape of the electromagnetic shower calculated as the energy weighted average in the \( \eta \) direction is checked (\( \sigma_{\eta \eta} \)) and the ratio of the maximum containment 2×5 and 1×5 crystal arrays in the \( \eta \) direction against the 5×5 crystal matrix is calculated. All the threshold and selection values will be presented later.
3.2 Muons

Muons in CMS are identified using the dedicated Muon Detector System (see Section 2.2.6) and their properties are measured using a combination of the Inner tracker and the Muon System. The reconstruction algorithms in the trigger and offline process will be presented in this section.

3.2.1 L1 Muon Trigger Object

The L1 muon trigger object is constructed by combining the objects formed from the different muon detector subsystems (see Section 2.2.6). The drift tubes, the cathode strip chamber and the resistive plate chambers all form muon candidate objects which are combined together with information from the calorimeters in the Global Muon Trigger. As mentioned before, no charge tracker information is available for the L1 trigger.

Drift Tube Trigger

The Drift Tube (DT) chambers provide data for track reconstruction and triggering. The three levels of logical DT trigger.

In the first level, reconstruction is performed in single Super Layers (SL). This stage is called Bunch and Track Identifier (BTI) and rough muon track fits are formed using at least three hits. The position and direction of the muon trigger candidate are calculated by fitting a straight line within a programmable angular acceptance in both $\theta$ and $\varphi$. In the $\theta$ direction, only tracks coming from the vertex are formed. In the same level, the muon track trigger candidates are assigned to a bunch crossing (b.x.) for triggering purposes.

Since only three hits are required in the first level of the algorithm, many false triggers are generated. To deal with this issue, two more reconstruction levels are deployed in the bending plane, the TRACO, stands for “TRAcker COrrrector”, and the chamber Trigger Server (TS). The TRACO is responsible for improving the
3.2 Muons

angular resolution of the muon trigger candidate, while also converting the $\varphi$ separation between the two SL into a $p_T$ quantity. The TS is responsible for track selection decision and therefore, the di-muon candidate formation. In the last step of the reconstruction, the track segments are passed to the regional trigger system called the Drift Tube Track Finder (DTTF) which forms the final muon track trigger candidates and calculates the position, direction and transverse momentum values.

Cathode Strip Chamber Trigger

The purpose of the Cathode Strip Chamber (CSC) trigger is to reconstruct tracks in the CSC endcap muon system, while providing a measurement of the azimuthal angle ($\varphi$), the pseudorapidity ($\eta$) and the transverse momentum ($p_T$) for each muon candidate. Despite the reduced magnetic field in the endcap, the CSC, by using information of up to three stations, can achieve measurements of similar precision to that of DTTF. The hits on the 6 layer CSC system are called Local Charged Tracks (LCT). The hits on the anode and the cathode are reconstructed separately.

The purpose of the CSC anode trigger algorithm is to assign the LCT to a specific bunch crossing. Typically, the first two layers are used to determine the timing, while the rest are used to establish the existence of a muon track. On the other hand, the purpose of the CSC cathode trigger algorithm is to measure the $\varphi$ angle with high precision. This allows for a good muon transverse momentum measurement up to high momenta. The cathode and anode segments are combined in the CSC Track Finder algorithm, which forms the final muon track trigger candidates.

Resistive Plate Chambers Trigger

The Resistive Plate Chambers (RPC) trigger is based on the PAttern Comparator (PAC) trigger \[57\]. The PAC algorithm establishes patterns based on the spatial and time coincidence of the hits on the 4 RPC layers and compares them with pre-defined patterns. Figure 3.5 demonstrates examples of such patterns. Each pre-defined spatial pattern of hits corresponds to a muon of a specific transverse
Figure 3.5: Four hit patterns created by muon tracks bent in the magnetic field.

momentum. These spatial patterns depend additionally on the direction of the muon ($\eta, \varphi$). Due to the dense material between the muon stations (iron yoke), the muons tend to follow a trajectory affected by multiple scattering and energy losses. Muons of lower transverse momenta are affected more, therefore the number of possible patterns increases as the muon momentum decreases. However, for low momentum muons, the matching can be accurate using patterns with a reduced number of hits, since they bend more.

Global Muon Trigger

The purpose of the Global Muon Trigger (GMT) is to combine the muon candidates produced by the DT, CSC and RPC systems. This process is crucial to achieve good reconstruction efficiency, while controlling the trigger rate and reducing redundancy. The GMT input consists of the best four candidates produced by each trigger system. A spatial matching for the candidates from the overlapping trigger system is performed. If a candidate cannot be successfully matched, some quality criteria checks are performed to decide where it will be processed further. The muon candidates are sorted according to their transverse momentum and quality. As mentioned before, the GMT also receives information from the calorimeters. The muon track is propagated back to the calorimeters and the energy deposit for a windows
3.2 Muons

of $\Delta \eta \times \Delta \varphi = 0.35 \times 0.35$ is used to determine the isolation and whether it is a Minimum Ionising Particle (MIP).

3.2.2 HLT Muon

The HLT Muon reconstruction initiates only when L1 muon candidates have satisfied the conditions. The reconstruction is done in logical levels. In the first level, also called “L2”, the muon detector systems and the calorimetric information are used, while in the next level, “L3”, the inner silicon tracker information is used. Various charged track reconstructions and isolations will be presented here.

Muon System Track

The different muon detector systems are readout in the regions that the L1 candidates “fired” the trigger and the information is used to create tracks. The track trajectory is built using a Kalman Filter (KF) algorithm [54]. The seeding is based on hits in the inner chamber layers and at each step the trajectory is extrapolated to the next chamber layer, where the hit that minimises the $\chi^2$ test is chosen. When the measurement satisfies $\chi^2 < 1000$, the new hit is included in the fit and the track parameters are updated while the track is extrapolated further. The iterative fitting process is performed twice - the first time from the inner to the outer chamber layers and the second time backwards (outer to inner layers). The second time, the $\chi^2$ criterion is tighter ($\chi^2 < 25$), and thus serves the purpose of smoothing the track. Finally the track is propagated to the inner layers of the pixel silicon tracker and a vertex constraint is applied in order to improve momentum resolution.

Calorimeter Isolation

As in the L1 calorimetric isolation, the muon candidate is propagated back to the calorimeters. Towers consisting of ECAL and HCAL energy deposits are formed in a $\eta$-$\varphi$ cone of 0.24 around the track. Towers include ECAL and HCAL energy contributions only if their $E_T$ is greater than 0.2 GeV and 0.5 GeV, respectively. The
3.2 Muons

Combined tower energy is calculated using a weighting factor of 1.5 for the ECAL energy deposits and 1 for the HCAL ones. A veto area is formed in a $\eta$-$\phi$ cone of 0.07 around the track in the ECAL and 0.1 in the HCAL to account for the muon energy deposited on the calorimeters. The energy of the towers outside the veto cone and within the 0.24 isolation cone is summed and compared with a threshold value. The threshold value varies from $E_T < 2$ GeV to 4 GeV for different $\eta$ ranges.

**Inner Track**

The muon candidates are used to define Regions Of Interest (ROI) in the inner silicon tracker. Track reconstruction is performed in these ROI using the KF algorithm and the tracks are then matched to muon system tracks. Once an inner track and a muon system track are matched, the complete set of hits is refitted and the track trajectory is extrapolated to the vertex.

**Tracker Isolation**

Tracker isolation is done in a similar manner as the calorimeter one. To reduce the time of data unpacking and algorithm execution, the tracker isolation is performed on pixel tracks only. Tracks with $\Delta z < 0.2$ cm that are outside the 0.01 $\eta$-$\phi$ veto cone and within the 0.24 isolation cone are collected. The $p_T$ sum of these pixel tracks has to be less than $\sim 1$ GeV/c for the muon candidate to be considered isolated.

**3.2.3 Offline Muon Reconstruction**

The offline muon reconstruction is almost identical to the HLT. As with the case of the electron, since reconstruction time is not an issue, some processes are slightly modified. Firstly, the muon system track seeds are searched for in the whole detector instead of the regions within which the L1 muon candidates lie. The standard Combinatorial Track Finder (CTF) \cite{58} reconstructed tracks are used in the inner tracker and those inside the ROI are checked for compatibility with the muon system tracks. Also, tracker isolation is performed using the full inner tracks. Finally, an
extra algorithm is used to reconstruct muons called Muon Identification (MI) \cite{59}. Muons with low transverse momenta, especially those with $p_T < 7\text{ GeV}/c$, would usually not reach the muon stations or will not have muon system tracks. To account for these muons, the MI algorithm reconstructs muon tracks starting from the silicon inner tracking system. For each track the energy deposits in the calorimeters are checked against those expected for a muon. In addition, the MI algorithm is able to use compatible hits in the muon chambers that are not associated with a muon system track. This improves the momentum resolution measurement.

## 3.3 Jets

The outgoing partons of an interaction are seen by the detector as jets of particles which are products of the hadronisation process they undergo. Since these jets of particles deposit their energy in the CMS calorimeters, the jet finding algorithms search for calorimeter, mainly HCAL, regions with high energy activity. The jet finding algorithms in the different levels of trigger and offline reconstruction will be presented here.

### 3.3.1 L1 Jet object

The L1 jet finding algorithm starts by computing the sums of transverse energies in a 4×4 trigger tower window, called a trigger calorimeter region, except in the HF, where HCAL towers are measured individually. Each trigger tower comprises an ECAL 5×5 crystal array and an HCAL tower. The basis of the L1 jet trigger object is a 3×3 trigger calorimeter region window spanning across the complete $\eta, \varphi$ of the CMS barrel and endcap calorimeter (see Figure \ref{fig:3.6}), while a 3×3 HF tower window is used in the forward region. The window is centred around the calorimeter region/tower with the highest transverse energy deposit. The jet $\varphi$ window size remains the same, but the $\eta$ binning increases in size from the central to the forward region due to the calorimeter segmentation. The L1 jet is assigned the sum of the transverse energies of the 3×3 trigger calorimeter region, as well as the $\eta$ and $\varphi$ position of the central region.
The HLT and offline jet objects are almost identical. The jets are reconstructed using the “iterative cone” algorithm [60] with a $\eta$-$\phi$ cone of $\Delta R < 0.5$, where $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$, on the zero-suppressed calorimeter readout. The reconstructed jets are then corrected by using the “L2+L3” calibration factors [61]. The L2+L3 calibration is obtained by comparing the reconstructed jets with the jets produced by using the same iterative cone algorithm on stable MC particles. The jets are then matched spatially ($\Delta R < 0.25$) and the calibration factors are calculated in order to bring the ratio of reconstructed transverse energy to generated transverse energy to one. The small differences observed between the HLT and the offline reconstructed jets in the data presented later (see Section 5.2.6) are due to the fact that a more recent L2+L3 calibration was used during the offline reconstruction.
3.4 \( \tau \)-jets

The products of the tau lepton hadronic decay give a similar signature in the CMS detector as the standard QCD jets - hadronic and electromagnetic energy deposits originating from the charged and neutral pions, respectively. The main difference lies in the fact that the \( \tau \)-jets are more collimated compared to the QCD jets. This in turn means that the tracker and ECAL signal cones are narrower, while the sub-detectors are expected to have no activity above the noise level around the reconstructed tau. Since the \( \tau \)-jet trigger \[48\] was not used in any of the studies presented it is not included in this chapter.

3.4.1 Offline \( \tau \)-jet

The CMS reconstruction software incorporates two different algorithms for the \( \tau \)-jet. In both cases the iterative cone algorithm is used for clustering. However, in one case calorimeter towers are used as building blocks, while in the other case Particle Flow (PFlow) objects are utilised. In this analysis the latter was used due to its superior energy and position resolution as demonstrated in Figure 3.7.

The PFlow \( \tau \)-jet reconstruction and identification process is divided into two stages. In the first stage, relatively simple and robust methods are used in order to suppress the QCD background while preserving a large portion of the genuine taus, thereby serving as the reconstruction basis. In the second stage, more advanced identification criteria are applied in order to reject electrons and muons falsely reconstructed as \( \tau \)-jets. Before any clustering is performed, the Particle Flow objects are formed. The idea behind the PFlow algorithm is to combine information from different sub-detectors, such as the silicon charged tracker, the ECAL and the HCAL, in order to improve the measurement resolution.

PFlow objects

The PFlow algorithm starts by linking all the iterative KF tracks \[63\] with calorimetric energy clusters. The linkage is done by extrapolating the tracks from the
3.4 $\tau$-jets

![Graph showing $E_T$ and $\phi$ resolutions comparison](image)

Figure 3.7: Comparison of $E_T$ (left) and $\phi$ (right) resolutions using Particle Flow reconstruction (red line) and calorimeter based reconstruction (black line) of single taus with $p_T = 50$ GeV/c. No jet energy corrections are applied in either case [62].

last measured tracker hit to the ECAL and HCAL, at the depth where they are expected to have started depositing energy. When the extrapolated point falls within the boundaries of the calorimetric cluster, the momentum/energy compatibility is checked. If the calibrated ECAL and HCAL energies match the track momentum within the sub-detector measurement resolutions, a charged PFlow candidate is formed. If on the other hand, there is an excess of energy, PFlow photons and neutral hadrons are formed with the remaining ECAL and HCAL energy respectively. Once all tracks are used, the remaining ECAL and HCAL clusters that have not been linked give rise to additional PFlow photons and PFlow neutral hadrons. The PFlow charged candidates that match the electron identification criteria are assigned as such and removed from the list.

The PFlow $\tau$-jet Reconstruction and Identification

Once all PFlow candidates are built, the iterative cone algorithm with a $\eta-\phi$ cone of 0.5 is used to form the PFlow jets. Subsequently a PFlow charged hadron with $p_T > 5$ GeV/c is required to be found within an $\eta-\phi$ cone around the jet direction of $\Delta R < 0.1$. The highest PFlow charged hadron within the cone is assigned as the “leading” PFlow charged hadron. An inner “signal” cone, of $\Delta R_s$, is formed around
the leading PFlow charged hadron, where all the tau decay products are expected to lie. A larger cone, outside the signal cone, is then defined as the isolation cone of $\Delta R_{\text{iso}}$ size. Typical values for $\Delta R_s$ and $\Delta R_{\text{iso}}$ are 0.07 and 0.5, respectively.

A graphical representation of the tau reconstruction with the signal and isolation cones is shown in Figure 3.8. The isolation annulus in between the two cones, $\Delta R_s < \Delta R < \Delta R_{\text{iso}}$, is expected to have minimal activity. For the $\tau$-jet to be considered isolated, there should be no PFlow charged hadron with $p_T^{\text{trk}} > 1 \text{ GeV/c}$ and no PFlow photon with $E_T > 1.5 \text{ GeV}$.

As mentioned earlier, the first stage of reconstruction is followed by some advanced identification criteria. These mainly consist of criteria to reject electrons and muons, while also checking the track multiplicity. Electron rejection is based on the fact that they deposit most of their energy in the ECAL, while $\tau$-jets deposit a large fraction of their energy in the HCAL. Criteria using the opposite logic to electron identification, as discussed earlier, are applied on the $\tau$-jet candidate. These include the ratios of either ECAL energy or maximum $3 \times 3$ HCAL tower energy of the PFlow tau over the track transverse momentum of the charged hadron. In a second approach, the reconstruction of an electron is attempted. Particular importance is given to the criteria that forced the potential electron reconstruction to fail originally.
The multivariate analysis takes into account the track quality reconstruction, the difference in the momentum measurement from the inner to the outer part of the track and the compatibility of the latter with the ECAL cluster linked to it. A 90-95% efficiency for electrons is achieved across the tracker acceptance [62]. The very high muon identification efficiency provides a nearly optimal rejection of muon for PFlow tau candidates. The tracks matched with hits on the muon chambers as well as the tracks that have no signal in the muon system due to a possible gap between chambers, but with their calorimeter deposits fitting the minimum ionising particle scenario, are considered. The PFlow tau candidate is rejected if it either overlaps with an identified muon or matches a track with at least one compatible hit in the muon chambers. The muon rejection achieves an efficiency of over 99% for taus with a muon efficiency of less than 1% [62].
Chapter 4

ECAL and HCAL Responses for Charged Pions and Comparisons with CMSSW Simulated data

As discussed in Section 2.2.3, the CMS electromagnetic calorimeter is designed for optimal photon and electron energy measurements. This results in the ECAL being non-compensating, with the response difference between electrons and charged pions being substantial and energy dependent. Furthermore, the CMS hadronic calorimeter is designed to measure jets with $E_T$ up to several TeV in magnitude and is unsuitable for low energy hadron measurements. Since the particle content of the jets formed during the fragmentation process of the quarks and gluons contains a significant number of low energy particles at energies of a few GeV, it is essential to understand the response of the calorimeters to them. This information can be used to improve the linearity and resolution of jets, which is an important factor in separating signal from background events. This is achieved by using the silicon strip tracker which provides excellent momentum resolution measurements (see Section 2.2.2.2), and the ECAL and HCAL fine granularity to allow for energy response correction within a jet. Finally, understanding of the calorimeter responses is essential for improving the performance of the Particle Flow techniques described in Section 3.4.1.
4.1 H2 Beam Line

| Maximum momentum | 360 GeV/c for hadrons, electrons and muons  
| Δp/p | ~0.1%  
| Beam height in EHN1 | 2460 mm  
| Beam length | 615 m |

Table 4.1: Main parameters of the H2 beam.

In order to understand the response of the calorimeters to low energy particles, dedicated experimental setups, also known as test beams, have been commissioned. Results presented in this chapter have been collected during the 2006 combined test beam runs using segments of the ECAL and HCAL calorimeters in an arrangement resembling the CMS detector configuration which was located at the CERN H2 beam line. The response of charged pions will be presented in the momentum range \( 2 \leq p \leq 300 \text{ GeV/c} \). At the end, the response distributions are compared with Monte Carlo data, simulated and reconstructed using the official CMS tools (CMSSW 1.3.0) modified to match the detector geometry of the test beam.

4.1 H2 Beam Line

The CERN H2 beam line is part of the SPS North Area and provides protons with momentum \( p = 400 \text{ GeV/c} \) (primary beam) as well as hadrons, electrons and muons with momentum range \( 10 \leq p \leq 360 \text{ GeV/c} \) (secondary beam). The secondary beam is produced when the 400 GeV/c protons hit the T2 target located 590.9 m upstream of the calorimeters. The main parameters of the beam are summarised in Table 4.1. For the purposes of the test beam, an additional beam mode was configured. A Very Low Energy (VLE) mode (see Figure 4.1) was needed to provide particles with momenta 1 to 9 GeV/c. The VLE beam was produced with the addition of a T22 target located 97.0 m upstream of the calorimeters. In the high energy mode, the T22 target and the VLE beam dump were removed from the beam line.
4.2 Test Beam 2006 Experimental Setup

The combined 2006 test beam experimental setup is depicted schematically in Figure 4.1. An ECAL barrel supermodule (85×20 crystals in $\Delta \eta \times \Delta \varphi$) was used as the ECAL calorimeter. For the HCAL, two barrel wedges (4×16 towers in $\Delta \eta \times \Delta \varphi$ each) with the corresponding outer calorimeter placed behind each of them and a segment of endcap were used. For more details about the ECAL and HCAL calorimeter properties see Sections 2.2.3 and 2.2.4, respectively. A block of aluminium was placed between the HB and HO detection systems to simulate the effect of the magnet.

A number of additional detectors were placed along the beam line to assist the process of particle identification and beam cleaning:

- Three Cerenkov counters (CK 1 to CK 3): CK1 was located on the high energy mode part of the beam line, CK2 on the VLE beam line, while CK3 was crossed by both (see Figure 4.1). The CK2 detector is a 1.85 m long Cerenkov counter filled with CO$_2$ and was used to identify electrons in the VLE mode. CK2 was operated at 0.7 bar pressure so that electrons would give a signal with efficiency close to 99%. The CK3 detector is also a 1.85 m long Cerenkov counter filled with Freon 134a. CK3 was operated under different pressures depending on the beam properties; for beams with $p \leq 3$ GeV/c, to tag electrons, it was set at 0.88 bar and for beams with $p > 4$ GeV/c, to separate pions from kaons and protons, it was set at 1.2 bar.

- Eight Wire Chambers planes: WC 1 to WC 3 were located upstream and WC A to WC E were located downstream. Their purpose was to determine the beam position. The spatial resolution of each WC was $\sim 350 \mu$m in both the $x$ and $y$ coordinates.

- Four Scintillation counters (S 1 to S 4) were used for trigger purposes and beam cleaning. For an event to trigger the calorimeter read-out, the coincidence of the S 1, S 2 and S 4 counters was required. The S 4 counter was also used to
reject events with more than one particle, since it gave a clean signal separation. The S 3 counter had a smaller surface area and was used to select focused events.

- Four Beam Halo scintillation counters (BH 1 to BH 4) were arranged such that a $7 \times 7 \text{cm}^2$ gap was left for the beam to pass through. The BH counters were used to reject events with additional particles coming from the beam halo at angles larger than those that the S 4 could detect.

- Muon Veto scintillation counters were placed at different parts of the beam line. Upstream (UpMV) was placed behind the beam dump of VLE to ensure that muons didn’t pass through. Between the HCAL barrel/endcap and the HO detectors, a set of eight MV movable counters (muon veto wall covering a plane of 226 cm vertically and 100 cm horizontally) was placed in order to veto events, mainly consisting of charged pions converted into muons. Because showers developed from high energy hadrons are likely to not be fully contained within the HCAL barrel/endcap, the muon veto wall was not very effective. To deal with this, two additional MV counters ($80 \times 80 \text{cm}^2$) were placed, one in front and one behind a block of a 80 cm thick iron absorber.

- Two Time Of Flight counters (TOF1 and TOF2) were placed approximately 55 m apart and were used to identify hadrons in the VLE mode. The scintillator plates were $10 \times 10 \text{cm}^2$ in area and 2 cm thick, while the discrimination of the analogue pulses provided a time resolution of about 300 ps.
4.3 Particle ID and Beam Cleaning

As described in the previous section, a number of detectors were used to select the events used for analysis through particle identification and rejection of multiple particle events. For the rejection of multiple particle events, two methods were used. The first method targeted events with additional particles collinear to the beam using the S4 scintillator counter. The second method used the four beam halo scintillation counters BH1 to BH4 for events where additional particles came from the beam halo.

The output of the four scintillator counters is shown in Figure 4.2. In particular from Figure 4.2(d) events with signal less than 200 counts can be seen, signifying that there was no charged particle present (noise). The output of the S4 counter was used in conjunction with that of S1 and S2 to identify events with no particles (see Figures 4.2(a), 4.2(b)). The output of S4 was also used to distinguish multiple particle events from single particle ones. A Gaussian fit was made on the rising curve of the Analog To Digital (ADC) counter to give the mean and the standard deviation (σ) expected for a single particle (see Figure 4.2). The enhanced positive tail on the distribution is the result of the non-linear response of the ADC counter, whereas at about 500 counts the hint of two particle events signal can be seen. For an event to be selected, it had to satisfy \(200 < S4_{ADC} < 350\) counts, where 350 corresponds approximately to two standard deviations from the mean. The output of S3 (Figure 4.2(c)) was important for studies that needed a very focused beam and it was not used in this study.

The outputs of the BH1 to BH4 scintillator counters are shown in Figure 4.3. Since no particle is expected, the events selected are the ones with an ADC count within three standard deviations from the mean in the positive direction. Analytically, for the BH1 scintillator counter \(BH1_{ADC} < 340\) counts, for BH2 scintillator counter \(BH2_{ADC} < 120\) counts, for BH3 scintillator counter \(BH3_{ADC} < 170\) counts and for BH4 scintillator counter \(BH4_{ADC} < 260\) counts.
Figure 4.2: ADC output of trigger scintillator counters. The output of S1, S2 and S4 was used to reject events with no particles. The output of S3 was useful for studies which required a very focused beam.
Figure 4.3: ADC output of the BH scintillator counters. The BH scintillators were used to reject events with additional particles coming from the beam halo. The fits was performed at the gaussian part of each distribution.
The pion runs are contaminated mostly with electrons and muons. To separate events that contained pions from those that contained electrons or muons, the Cerenkov counters and the Muon Veto counters were used.

For data taken in the LVE mode configuration, the Muon Veto Wall (MVW) counters were used to identify muons. This was primarily done because the momenta of the produced muons were not sufficient to penetrate the ECAL/HCAL calorimeters and the iron absorber, while the pions were expected to be fully contained. For higher momentum runs, the MVW counter measurements were affected by hadronic showers from pions which were not fully contained within the calorimeter. Therefore the MV counter behind the iron absorber was used. Figures 4.4(a), 4.4(b) and 4.4(c) illustrate how the muon events were detected depending on the output of the different MV counters.

Electron contamination was more severe in the VLE mode pion runs. To identify events with electrons, the Cerenkov counter (CK 3) was used. Its output is shown in Figure 4.4(d). Since only electrons are expected to have signal above the noise level, the selected events had CK 3 \text{ADC} < 500 counts.

It is worth noting that the calibration of the counters could change slightly over time depending on parameters such as temperature. Therefore, the selections presented in this section are a good baseline, but they were altered for different runs to better match the specific calibration.

As an example, the energy deposited in the ECAL against that deposited in the HCAL for a pion run at $p_\pi = 3$ GeV/c before and after the beam cleaning is shown in Figure 4.5. From Figure 4.5(a), electron contamination is apparent by the number of events depositing around 3 GeV in the ECAL, while the muons are concentrated around events that have minimum ionising deposits both in the ECAL and the HCAL. On the right of the Figure 4.5(a), events with ECAL energy around 6 GeV can be seen, signifying multiple particle events - in this case two electron events give the cleanest signature. Finally, around (0,0) there are events with no particles. In Figure 4.5(b) it can be seen that most of the effects described above have been removed, apart from the events at (0,0) which will be further discussed later in this chapter (see Section 4.6.3).
Figure 4.4: ADC output of MV and CK scintillator counters. MV 2 and MV 3 are part of the MVW and their output was used in the LVE mode, while MV B was used in the high momentum mode. The output of CK 3 was used to reject events containing electrons in the pion runs. The fits were performed at the gaussian part of each distribution.
4.4 Data Sets

Officially reconstructed data sets were used in this analysis. The data sets were produced using the CMS reconstruction software CMSSW_0.8.1 and used the version 6 and 6.2 of the test beam reconstruction and calibration software. The ECAL and HCAL were calibrated using electrons at $p_e = 50$ GeV/c. The calibration of the HCAL barrel was performed without the ECAL supermodule in front and it was done by aiming the electron beam at the centre of each HCAL tower. The ECAL calibration was performed in a similar manner.

For low momentum beams both $\pi^-$ and $\pi^+$ pion runs were examined, whereas for high momentum only $\pi^-$ were used. The pion momenta and number of events included in the analysis are summarised in Table 4.2. Only runs where the particle beam was aimed at the calorimeter barrel were used.

4.5 Energy Response

The CMS reconstruction software CMSSW_0.8.4 was used for the analysis of the data sets. The ECAL energy deposits were calculated as the sum of a $7 \times 7$ crystal
### 4.5 Energy Response

<table>
<thead>
<tr>
<th>Particle</th>
<th>Generated Momentum [GeV/c]</th>
<th>Number of events</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi^+$</td>
<td>3</td>
<td>45,319</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>4</td>
<td>55,387</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>5</td>
<td>46,244</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>6</td>
<td>45,903</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>7</td>
<td>54,124</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>8</td>
<td>45,856</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>9</td>
<td>23,425</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>2</td>
<td>207,163</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>2.5</td>
<td>82,590</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>3</td>
<td>125,320</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>4</td>
<td>175,118</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>5</td>
<td>140,879</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>6</td>
<td>131,907</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>7</td>
<td>91,197</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>9</td>
<td>90,515</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>15</td>
<td>76,696</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>20</td>
<td>37,923</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>30</td>
<td>76,241</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>50</td>
<td>76,407</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>100</td>
<td>37,759</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>150</td>
<td>37,609</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>200</td>
<td>37,859</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>300</td>
<td>31,154</td>
</tr>
</tbody>
</table>

**Table 4.2:** Summary of the particles, their momenta and the number of events used in this analysis.
array formed around the most energetic crystal, while the HCAL energy deposits were calculated as the sum of a $3 \times 3$ tower array centred behind the central ECAL crystal. The choice of the array sizes is based on maximum containment of the pion showers while minimising the noise entering the sum and has been used in similar test beam studies [45] [65].

Figures 4.6 and 4.7 show the ECAL and HCAL responses for different pion momenta. Pions interact in two distinct modes when passing through the ECAL: in one mode they pass as Minimum Ionising Particles (MIPs) with an average energy deposit of $\sim 400$ MeV, while in the other mode the pions start their hadronic shower inside the $1.1 \lambda_I$ ECAL. The two modes are very clear for high momentum pions as can be seen in the left Figure 4.7(b) and it remains observable, although the two modes overlap down to $p_\pi = 4$ GeV (see left plot of Figure 4.6(b)). For $p_\pi \leq 3$ GeV, the two modes become indistinguishable since they fully overlap (see left plot of Figure 4.6(a)).

The effect of the pion response in the ECAL can be seen in the HCAL response. For $p_\pi \geq 4$ GeV, a second peak arises corresponding to the MIP pions in the ECAL, though the HCAL resolution makes it less obvious than in the ECAL (see middle plot of Figure 4.6(b)). The effect can be seen clearly in left plot of Figure 4.7(a).

It is important to understand these different modes in order to provide the response of a pion of a given momentum. In this analysis, the response of the pion is defined as the mean of the distribution, which assumes a Gaussian shape distribution. For high momentum pions separate responses can be given for pions that passed as MIPs from the ECAL and for those that deposited substantial energy in it. Pions that deposited less than 1.6 GeV ($\mu_{\text{MIP}} + 3\sigma_{\text{MIP}}$, with both $\mu_{\text{MIP}}$ and $\sigma_{\text{MIP}} \sim 400$ MeV) in the $7 \times 7$ ECAL crystal array were considered as MIPs. In this case, only the response of the HCAL was considered to be the best estimate. For pions that started their hadronic shower in the ECAL, the sum of the ECAL and HCAL response was considered. This distinction was essential since the mean and the spread of the two distributions differ substantially (see middle and right plot of Figure 4.7(b)). It is feasible due to the high granularity of the ECAL which allows the matching of tracks with single ECAL clusters even in a particle jet environment. As shown in the middle plot of Figure 4.7(a), the HCAL response distribution for low momentum pions that passed...
4.5 Energy Response

(a) The ECAL, HCAL and the combined ECAL+HCAL responses can be seen from left to right for π⁻ with \( p_{\pi^-} = 3 \text{ GeV/c} \).

(b) The ECAL, HCAL and the combined ECAL+HCAL responses can be seen from left to right for π⁻ with \( p_{\pi^-} = 4 \text{ GeV/c} \). In the right and middle a second distribution caused by the two different pion interaction modes can be seen emerging.

Figure 4.6: ECAL and HCAL responses for π⁻ momenta of 3 GeV/c and 4 GeV/c.
4.5 Energy Response

(a) The HCAL, HCAL without ECAL interaction and the ECAL+HCAL responses can be seen from left to right for $\pi^-$ with $p_{\pi^-} = 7$ GeV/c.

(b) The ECAL, HCAL and the combined ECAL+HCAL without ECAL interaction responses can be seen from left to right for $\pi^-$ with $p_{\pi^-} = 50$ GeV/c. The pions passing as MIPs form the ECAL can be seen in plot on the left.

Figure 4.7: ECAL and HCAL responses for $\pi^-$ momenta of 7 GeV/c and 50 GeV/c.
### Table 4.3

<table>
<thead>
<tr>
<th>Particle</th>
<th>Momentum [GeV/c]</th>
<th>HCAL response with MIP in ECAL (σ) [GeV]</th>
<th>ECAL+HCAL response with interaction in ECAL (σ) [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>π⁻</td>
<td>2</td>
<td>−</td>
<td>1.02 (0.93)</td>
</tr>
<tr>
<td>π⁻</td>
<td>2.5</td>
<td>−</td>
<td>1.31 (0.99)</td>
</tr>
<tr>
<td>π⁻</td>
<td>3</td>
<td>−</td>
<td>1.54 (1.05)</td>
</tr>
<tr>
<td>π⁺</td>
<td>3</td>
<td>−</td>
<td>1.48 (1.06)</td>
</tr>
<tr>
<td>π⁻</td>
<td>4</td>
<td>−</td>
<td>2.13 (1.23)</td>
</tr>
<tr>
<td>π⁺</td>
<td>4</td>
<td>−</td>
<td>2.07 (1.21)</td>
</tr>
<tr>
<td>π⁻</td>
<td>5</td>
<td>−</td>
<td>2.68 (1.32)</td>
</tr>
<tr>
<td>π⁺</td>
<td>5</td>
<td>−</td>
<td>2.66 (1.25)</td>
</tr>
<tr>
<td>π⁻</td>
<td>6</td>
<td>−</td>
<td>3.39 (1.42)</td>
</tr>
<tr>
<td>π⁺</td>
<td>6</td>
<td>−</td>
<td>3.37 (1.40)</td>
</tr>
<tr>
<td>π⁻</td>
<td>7</td>
<td>−</td>
<td>3.91 (1.49)</td>
</tr>
<tr>
<td>π⁺</td>
<td>7</td>
<td>−</td>
<td>3.92 (1.47)</td>
</tr>
<tr>
<td>π⁺</td>
<td>8</td>
<td>6.04 (2.43)</td>
<td>4.34 (1.66)</td>
</tr>
<tr>
<td>π⁻</td>
<td>9</td>
<td>7.39 (2.58)</td>
<td>4.88 (1.65)</td>
</tr>
<tr>
<td>π⁺</td>
<td>9</td>
<td>7.36 (2.49)</td>
<td>4.82 (1.82)</td>
</tr>
<tr>
<td>π⁻</td>
<td>20</td>
<td>18.58 (4.80)</td>
<td>13.14 (3.61)</td>
</tr>
<tr>
<td>π⁻</td>
<td>30</td>
<td>28.40 (5.86)</td>
<td>21.86 (4.55)</td>
</tr>
<tr>
<td>π⁻</td>
<td>50</td>
<td>49.58 (8.99)</td>
<td>39.35 (7.16)</td>
</tr>
<tr>
<td>π⁻</td>
<td>100</td>
<td>101.45 (15.99)</td>
<td>84.10 (11.39)</td>
</tr>
<tr>
<td>π⁻</td>
<td>150</td>
<td>151.62 (21.83)</td>
<td>129.58 (15.44)</td>
</tr>
<tr>
<td>π⁻</td>
<td>200</td>
<td>202.54 (29.24)</td>
<td>175.45 (18.63)</td>
</tr>
<tr>
<td>π⁻</td>
<td>300</td>
<td>310.70 (42.06)</td>
<td>274.40 (27.15)</td>
</tr>
</tbody>
</table>

Table 4.3: Average response and standard deviation of pions interacting in the HCAL and pions starting their hadronic shower in the ECAL using test beam data. The error on the average response obtained by factoring the error in the fit and by varying the fitting range is about 3% for low momentum pions and decreases gradually to about 0.2% at $p_\pi = 300$GeV/c. The error in the standard deviation obtained in the same way is about 3% for the all pion momenta.

as MIPs from the ECAL have a substantial non-Gaussian negative tail. Therefore the response for pions with $p_\pi \leq 7$GeV/c was considered to be the sum of the ECAL and HCAL components (see right plots of Figures 4.6(a), 4.6(b) and 4.7(a)). The energy responses of π⁻’s and π⁺’s for all the different momenta are summarised in Table 4.3.

Some additional things to note:

- From Table 4.3, it can be seen with the average HCAL response for a $p_\pi = 50$GeV/c pion being ~49.6GeV and considering that pions deposit on average 400 MeV...
4.5 Energy Response

in the ECAL, the calorimeter response adds up to 50 GeV. This provides a good test for the energy calibration.

- Also from Table 4.3 it can be seen that for pions with $p_\pi > 50$ GeV/c the average HCAL response is overestimated, which is caused by the non-linear response of the HCAL detector.

- Another demonstration of HCAL’s non-linearity is the positive tails that appears in all distributions that include the HCAL response in Figures 4.6 and 4.7.

- From Table 4.3 it can be seen that the average $\pi^-$ and $\pi^+$ responses are in good agreement.

- A small contamination of electrons with ECAL energy deposits of $\sim$4 GeV can be seen in the left plot of Figure 4.6(b). These events were excluded from the calculation of the pion average response.

- During operation the CMS detector will apply a suppression on the calorimeter data to reduce the event data size. This is called zero suppression and is done by keeping a record only of detector elements with energy deposits greater than the noise levels. This would have an effect, especially on low energy pion detection, where the energy deposited in the HCAL is comparable to the noise level. However, low energy pions usually appear in particle jets, where the density of particles is higher than the granularity of the HCAL. This means that the same HCAL tower will sample energy deposited by more than one hadron, thus minimising the effect of the zero suppression.

- Finally, a small number of events seem to include particles that deposit energies close to the noise level both in the ECAL and the HCAL (see peak close to zero in the middle plot of Figure 4.7(b)). This has been observed in analysis of previous test beam data [45], but due to the lack of a comprehensive muon veto system in the setup, it was considered to be muon contamination in the beam. This could not be the case here and it will be discussed later in the comparison with the simulated data (Section 4.6.3).
4.6 Comparison of Test Beam and Simulated Data

The aim of this study was to assess the agreement of the official Monte Carlo (MC) simulation tools used by the CMS experiment and the test beam data. At the time that the study was conducted, there were no MC simulated data sets. The tools available to produce MC data sets provided only the test beam geometry and the Geant4 [66] simulation of the detector elements without the expected detector noise. Since this made it impossible to compare the two, the MC data sets produced for this study used the full CMS simulation, digitisation and reconstruction of the calorimeters with the ECAL and HCAL responses given in physical energy.

4.6.1 Detector Simulation and Reconstruction

To provide a realistic representation of the test beam setup, the file providing the geometry of the calorimeters, all the additional detectors and absorbers was used. This also meant that both the tracker simulation and the simulation of the magnetic field had to be disabled. The only compatible calorimeter modules to provide digitisation and reconstruction were those developed for the dedicated ECAL test beam, while there was nothing equivalent for the HCAL detector. Since the ECAL digitisation and reconstruction modules were aimed at mimicking the test beam data by including hard-coded constants and calibration factors, they were not suitable for use. Instead, the modules included in the official CMS tools (CMSSW) were configured to provide the digitisation and reconstruction. This included the modules simulating the electronic noise for the ECAL and the HCAL detectors. Finally, zero suppression on the calorimeter data was disabled since the data provided by the test beam did not include any.

4.6.2 Simulated Data Production

Production of MC data was done using CMSSW_1.3.0. A “particle gun” was used to produce $\pi^-$ data sets for the full range of momenta 2–300 GeV/c, while a few data
4.6 Comparison of Test Beam and Simulated Data

<table>
<thead>
<tr>
<th>Particle</th>
<th>Generated Momentum [GeV/c]</th>
<th>Number of events</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi^+$</td>
<td>5</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>9</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>2</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>3</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>4</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>5</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>6</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>7</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>9</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>15</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>20</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>30</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>50</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>100</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>150</td>
<td>40,000</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>300</td>
<td>25,000</td>
</tr>
</tbody>
</table>

Table 4.4: Summary of the particles, their momenta and the number of events produced for this analysis.

sets of $\pi^+$ were produced for comparison. The pion momenta and number of events produced for this analysis are summarised in Table 4.4. The pions were aimed at a section of the calorimeter similar to the one used for the test beam runs, while a spread in $\eta$ and $\phi$ was introduced to match that of the test beam pions.

4.6.3 Comparison between Simulated and Test Beam Results

Analysis of the MC data was done using the techniques described in Section 4.5. The energy responses of $\pi^-$'s and $\pi^+$'s for the different momenta are summarised in Table 4.5. The combination of the TB and MC responses can be seen graphically in Figure 4.8.

From Figure 4.8 it can be seen that the response of pions is systematically underestimated when the hadronic shower has started in the ECAL. This signifies that the response of pions in the ECAL is not simulated correctly. Figure 4.9(a) shows the ECAL response distribution of $\pi^-$ with $p_{\pi^-} = 50\,\text{GeV/c}$ from test beam and
Table 4.5: Average response and standard deviation of pions interacting in the HCAL and pions starting their hadronic shower in the ECAL using MC data. The error on the average response obtained by factoring the error in the fit and by varying the fitting range is about 2% for low momentum pions and decreases gradually less than 0.1% at $p_\pi = 300$ GeV/c. The error in the standard deviation obtained in the same way is about 2% for the all pion momenta.

<table>
<thead>
<tr>
<th>Particle</th>
<th>Momentum [GeV/c]</th>
<th>HCAL response with MIP in ECAL ($\sigma$) [GeV]</th>
<th>ECAL+HCAL response with interaction in ECAL ($\sigma$) [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi^-$</td>
<td>2</td>
<td>–</td>
<td>0.82 (0.95)</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>3</td>
<td>–</td>
<td>1.30 (1.09)</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>4</td>
<td>–</td>
<td>1.84 (1.21)</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>5</td>
<td>–</td>
<td>2.43 (1.26)</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>5</td>
<td>–</td>
<td>2.39 (1.31)</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>6</td>
<td>–</td>
<td>2.99 (1.31)</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>7</td>
<td>–</td>
<td>3.61 (1.41)</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>9</td>
<td>7.12 (2.64)</td>
<td>4.64 (1.61)</td>
</tr>
<tr>
<td>$\pi^+$</td>
<td>9</td>
<td>7.06 (2.78)</td>
<td>4.57 (1.56)</td>
</tr>
<tr>
<td>$\pi^-$</td>
<td>15</td>
<td>13.12 (3.57)</td>
<td>8.58 (2.64)</td>
</tr>
<tr>
<td>$\pi^-\pi$</td>
<td>20</td>
<td>18.68 (4.43)</td>
<td>12.46 (2.92)</td>
</tr>
<tr>
<td>$\pi^-\pi$</td>
<td>30</td>
<td>29.14 (5.43)</td>
<td>20.29 (3.78)</td>
</tr>
<tr>
<td>$\pi^-\pi$</td>
<td>50</td>
<td>50.40 (7.26)</td>
<td>36.71 (6.13)</td>
</tr>
<tr>
<td>$\pi^-\pi$</td>
<td>100</td>
<td>102.77 (11.55)</td>
<td>78.89 (15.59)</td>
</tr>
<tr>
<td>$\pi^-\pi$</td>
<td>150</td>
<td>154.72 (14.96)</td>
<td>122.15 (16.04)</td>
</tr>
<tr>
<td>$\pi^-\pi$</td>
<td>300</td>
<td>309.13 (25.21)</td>
<td>249.87 (50.25)</td>
</tr>
</tbody>
</table>

simulated data. It can be seen that in the simulated data, pions deposit no more than 39 GeV in the ECAL, while in the test beam pions can deposit up to 50 GeV.

Also from Figure 4.8, it can be seen that the HCAL average response of the simulated and the test beam data are in close agreement, approximately within 2%. However, as seen from the second column of Tables 4.3 and 4.5 the spreads of the corresponding distributions differ significantly. The difference is of the order of 20% for $p_\pi = 50$ GeV/c and increases to 40% for $p_\pi = 100$ GeV/c. Figure 4.9(b) demonstrates the difference of the HCAL response for the simulated and the test beam pions. This signifies that the HCAL resolution in the simulation is better than that given by the test beam data.

As discussed before, a small number of pions deposit a small amount of energy both in the ECAL and the HCAL. Figure 4.9(b) shows that these pions are also observed in the simulated data, while their normalised number is of the same order
Figure 4.8: The response of the pions measured by the test beam data (hollow) and MC (filled) is shown as a function of the $\pi$ beam momentum. Pions with $p_\pi < 8 \text{ GeV/c}$ are measured only by the ECAL+HCAL system, while higher momentum pions are separated depending on if they started their hadronic shower in ECAL (blue) or not (red).
4.7 Summary and Conclusions

The energy response of pions with momenta ranging from $p_\pi = 2 \text{ GeV/c}$ to $p_\pi = 300 \text{ GeV/c}$ was calculated using data from the 2006 Combined Test Beam. The results demonstrated the non-linear response of the CMS hadronic calorimeter. The simulated data were produced using the official CMS simulation and reconstruction software.
4.7 Summary and Conclusions

CMSSW_1.3.0. Comparison between the two showed a good agreement in the HCAL average response, while a deviation was observed in the HCAL resolution. Also it was shown that the simulation underestimates the response of pions in the ECAL. At the time that the results were reported steps were taken to correct the simulated HCAL resolution and it was done in CMSSW_1.5.2. The deviation in the response of the pions at the ECAL was already known from previous test beam analysis, but it was not thought to be a priority. Particle Flow techniques (see Section 3.4.1) rely heavily on knowing the responses of the calorimeters to different particles. Therefore the understanding of these responses is very important from the beginning of the CMS operation. Although this analysis is a good baseline to understand the response of the calorimeters to pions, there are a few limitations compared to the full CMS detector. The presence of the tracker, approximately 0.3 $\lambda_I$, will contribute to more pions developing hadronic showers earlier. Also the operation of the calorimeters inside the 4 T magnetic field will increase the electronic noise. Finally, a similar analysis using data from the LHC collisions will require dedicated runs collecting unsuppressed calorimeter information.
The detection of the Higgs boson is one of the primary goals of the CMS experiment. Precision electroweak data from LEP and SLD combined with measurements of the top quark and the $W^\pm$ boson masses from the LEP and Tevatron experiments favour a light Standard Model Higgs boson of $m_H < 160 \text{ GeV}/c^2$ (see Section 1.3.3). In this mass range, the decay channel of the Higgs boson to a pair of $\tau$ leptons is particularly important for its detection. The $\tau$ lepton pair decay channel is preferred to the dominant $b\bar{b}$, see Figure 1.6(a), because the purely hadronic final state of the latter would cause great difficulty in distinguishing it from the large QCD background. Neglecting neutrinos which cannot be detected by CMS, a tau lepton’s decay branching ratio to electrons is 17.8%, to muons is 17.4% and to hadrons ($\tau$-jet) is 64.8%. Therefore, a lepton and a $\tau$-jet are the most likely final state of a $\tau$ lepton pair with branching ratio of 45.6%, followed by two $\tau$-jets and two leptons final states, with branching ratios of 42% and 12.4%, respectively. To assist further in increasing the Higgs boson signal to background event ratio, the Vector Boson Fusion (VBF) production mechanism is favoured. The VBF Higgs boson production mechanism has the second highest cross section - approximately an order of
magnitude below the gluon fusion (see Figure 1.5) - but it provides a characteristic identification signature due to the two forward jets (see Figure 1.4).

To maximise the number of Higgs boson events saved on disk for offline analysis, an efficient trigger is needed. In previous analyses, the single lepton and lepton plus τ-jet triggers were used. However to keep the trigger rate under control, the minimum $E_T$ of these objects has to increase with increasing LHC luminosity, thus rejecting many Higgs boson signal events. In this chapter, the development and implications of a dedicated Level-1 (L1) and High Level Trigger (HLT) aiming at increasing the trigger efficiency will be discussed. In order to keep the lepton $E_T$ thresholds at a low value, two well separated jets (VBF jets) were required. Also, the basic characteristics of the Higgs boson produced via the VBF mechanism with final state of a lepton, a τ-jet and the corresponding neutrinos will be presented.

## 5.1 Vector Boson Fusion Higgs

### 5.1.1 Production Cross Section

The Higgs boson production cross sections via the VBF mechanism for $m_H = 115$ GeV/$c^2$ and 135 GeV/$c^2$ and the corresponding branching ratios to a τ lepton pair for 10 TeV and 14 TeV Centre of Mass (CoM) energy are summarised in Table 5.1. The samples used for this study were generated with CoM energy of 10 TeV. This decision was made mainly due to the sample availability of the dominant background events for the CMS simulation and reconstruction software version used (CMSSW_2_2_13). The differences between the Higgs boson events produced via the VBF channel at 10 TeV and 14 TeV CoM energy will be discussed in Section 5.1.2.

### 5.1.2 Event Kinematics at 10 TeV and 14 TeV CoM Energy

#### 5.1.2.1 Vector Boson Fusion Jets

A characteristic identification signature of the VBF events is the existence of two well separated final state jets formed by the two leading quarks. The main difference
Table 5.1: The Next to Leading Order (NLO) cross sections for the Higgs boson production via the VBF and the branching ratios for the $H \rightarrow \tau \tau$ decay channel for different masses of the Higgs boson. The cross sections were generated according to calculations from [67], using the code provided by M.Spira [68] with the CTEQ6M parton distribution function. The scale dependency from renormalisation and factorisation is of the order of ±2% for all values.

\[
\begin{array}{|c|c|c|c|c|}
\hline
m_H \ [\text{GeV}/c^2] & 115 & 135 \\
\hline
\text{Centre of mass energy, } \sqrt{s} \ [\text{TeV}] & 14 & 10 & 14 & 10 \\
\hline
\text{VBF production, } \sigma \ [\text{pb}] & 4.67 & 2.64 & 3.97 & 2.22 \\
\hline
\text{Branching ratio, BR}(H \rightarrow \tau \tau) & 7.409 \times 10^{-2} & 4.532 \times 10^{-2} \\
\hline
\sigma \times \text{BR}(H \rightarrow \tau \tau \rightarrow l + \tau\text{-jet}) \ [\text{fb}] & 157.8 & 89.21 & 82.06 & 45.89 \\
\hline
\end{array}
\]

Figure 5.1: Pseudorapidity distribution of the two VBF jets, the $\tau$-jet and the lepton originating from the $\tau$ decay. The high $\eta$ nature of the jets is evident, while the Higgs boson decay products are located more centrally. These distributions correspond to the 14 TeV CoM scenario, but they are very similar to the 10 TeV generated data.

The jet $p_T$ distributions at parton level of the two VBF jets can be seen in Fig-
for a CoM energy of 10 TeV and 14 TeV. The differences between their means and shapes are minimal. Since the spectrum of the QCD jets is dropping exponentially, the kinematic threshold for the most energetic jet is set high. To maintain a reasonable efficiency the second jet has a lower threshold. The thresholds are decided in such a way as to preserve approximately 70% efficiency of each jet. The rapidity gap distributions of the VBF jet for CoM energy of 10 and 14 TeV are shown in Figure 5.2(b). The difference here is more noticeable, while the efficiency for a selection condition of $\Delta \eta_{j_1,j_2} > 3.6$ is $\sim 73\%$ and $\sim 69\%$ for 14 TeV and 10 TeV CoM energy, respectively. This results in a $\sim 5\%$ reduction in the signal efficiency. Finally, the invariant mass of the two VBF jets can be seen in Figure 5.2(c). A slightly softer spectrum is observed for the Higgs generated at the 10 TeV CoM energy. This results in a difference in selection efficiency changing from $\sim 8\%$ at $m(j_1^{\mu}, j_2^{\nu}) = 400$ GeV/$c^2$ to $\sim 22\%$ at $m(j_1^{\mu}, j_2^{\nu}) = 900$ GeV/$c^2$.

5.1.2.2 $\tau$-jet and Lepton

The $\tau$ lepton has a lifetime of $3 \times 10^{-13}$ s, which corresponds to a distance $\sim 90 \mu m$ in the lab frame [38]. Hence the $\tau$ lepton is always detected through its decay products. Most frequently it decays to hadrons, with a 49.2% probability of a single charged hadron in the final state and 14.6% probability of three charged hadrons (one and three prong decays). This fact can be used as an identification criterion to suppress QCD jets by requiring one or three charged tracks associated with the tau candidate.

A comparison of the $\tau$-jet $p_T$, $\eta$ and $\varphi$ distributions at 10 TeV and 14 TeV CoM energy can be seen in Figure 5.3(a). The 10 TeV CoM energy $\tau$-jet $\eta$ distribution is narrower, signifying that the Higgs boson decays more centrally. The same can be deduced from the electron and muon $\eta$ distributions in Figures 5.3(b) and 5.3(c). This is caused by the lower average energy in the VBF interaction, resulting in a less boosted Higgs boson. However, the $p_T$ distribution of the particles depends mostly on the Higgs boson mass. Hence the $p_T$ spectra of the $\tau$-jet, electron and muon are very similar for the 10 TeV and 14 TeV CoM cases (see Figure 5.3). As expected, the decay products of the Higgs boson have a uniform distribution in $\varphi$ as demonstrated in Figures 5.3(a), 5.3(b) and 5.3(c).
5.1 Vector Boson Fusion Higgs

(a) The $p_T$ distributions for the VBF jet with the higher transverse energy on the left and the VBF jet with the lower transverse energy at parton level on the right for 10 TeV (dashed blue line) and 14 TeV (solid red line) CoM energy. The substantial difference in these distributions motivates the choice of different kinematic thresholds for the most and the least energetic VBF jets.

(b) The pseudorapidity gap between the two VBF jets for the 10 TeV (dashed blue line) and 14 TeV (solid red line) CoM energy cases. This is the characteristic signature of the VBF production mechanism. An offline cut of about 600-900 GeV/c$^2$ is used depending on the analysis scenario.

(c) The invariant mass of the two VBF jets for the 10 TeV (blue line) and 14 TeV (solid red line) CoM energy cases. This condition is important in order to reject di-jet events.

Figure 5.2: VBF jet kinematics for the 10 TeV and 14 TeV CoM energy cases.
Figure 5.3: Kinematic distributions of the visible final state decay products of the Higgs boson.

(a) The $p_T$ (left), $\eta$ (middle) and $\phi$ (right) distributions for the $\tau$-jet in the 10 TeV (dashed blue line) and 14 TeV (solid red line) CoM cases.

(b) The $p_T$ (left), $\eta$ (middle) and $\phi$ (right) distributions for the electron in the 10 TeV (dashed blue line) and 14 TeV (solid red line) CoM cases.

(c) The $p_T$ (left), $\eta$ (middle) and $\phi$ (right) distributions for the muon in the 10 TeV (dashed blue line) and 14 TeV (solid red line) CoM cases.
5.2 Higgs Boson Trigger Study

One of the challenges of the online selection system is to trigger on the events that will be used in physics analyses while keeping the QCD background at a manageable rate. Given the small cross section of the VBF Higgs signal process ($H \rightarrow \tau\tau \rightarrow l + \tau\text{-jet}$), saving the maximum number of these events will be essential for its discovery.

5.2.1 Current Trigger Selections

The proposed trigger paths for the Higgs signal events are motivated by the lepton and $\tau$-jet final state. In previous studies the trigger paths used were those of the single isolated electron, the single isolated muon, as well as the combined isolated electron plus $\tau$-jet and isolated muon plus $\tau$-jet [69, 70]. The single $\tau$ trigger is not used since the L1 $E_T$ trigger threshold is set to 100 GeV. The trigger objects and their respective kinematic thresholds ($E_T/p_T$) of the L1 trigger and HLT for instantaneous luminosity $\mathcal{L} = 2 \times 10^{33}$ cm$^{-2}$s$^{-1}$ are summarised in Table 5.2. Note that the values of the kinematic thresholds and rates can be used for reference purposes only since these triggers were developed and analysed with the previous official CMS reconstruction software, ORCA (Object-oriented Reconstruction for CMS Analysis) which became obsolete in 2006. The triggers were re-developed for the current official CMS reconstruction software (CMSSW) and their performance was analysed, but only for the early collision data scenario and hence a lower luminosity regime ($\mathcal{L} = 10^{32}$ cm$^{-2}$s$^{-1}$). Also, since there were no plans to use the combined e-$\tau$ and $\mu$-$\tau$ triggers in this luminosity regime, the available algorithms were not fully implemented and therefore were not used for this study.

5.2.2 The VBF Jets plus Lepton Trigger

As shown in Section 5.1.2.2, the $E_T$ spectrum of the leptonically decaying $\tau$ drops exponentially. Since the offline thresholds used for the offline reconstructed electrons and muons are $E_T^{\text{thres}} = 15$ GeV and $p_T^{\text{thres}} = 10$ GeV respectively (see Section 5.2.4),


5.2 Higgs Boson Trigger Study

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Level-1 object used</th>
<th>( E_{T}/p_{T} ) Threshold [GeV]</th>
<th>Rate (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 isolated ( e/\gamma )</td>
<td>-</td>
<td>22</td>
<td>((4.2\pm0.1)\times10^{3})</td>
</tr>
<tr>
<td>L1 isolated ( \mu )</td>
<td>-</td>
<td>14</td>
<td>((2.7\pm0.1)\times10^{3})</td>
</tr>
<tr>
<td>L1 isolated ( e/\gamma-\tau )</td>
<td>-</td>
<td>14, 52</td>
<td>((5.4\pm0.2)\times10^{3})</td>
</tr>
<tr>
<td>L1 isolated ( \mu-\tau )</td>
<td>-</td>
<td>7, 40</td>
<td>((1.2\pm0.1)\times10^{3})</td>
</tr>
<tr>
<td>HLT isolated ( e )</td>
<td>isolated ( e/\gamma )</td>
<td>26</td>
<td>23.5\pm6.7</td>
</tr>
<tr>
<td>HLT isolated ( \mu )</td>
<td>isolated ( \mu )</td>
<td>19</td>
<td>25.8\pm0.8</td>
</tr>
<tr>
<td>HLT isolated ( e-\tau )</td>
<td>isolated ( e/\gamma-\tau )</td>
<td>16, 52</td>
<td>&lt;1.0</td>
</tr>
<tr>
<td>HLT isolated ( \mu-\tau )</td>
<td>isolated ( \mu-\tau )</td>
<td>15,40</td>
<td>&lt;1.0</td>
</tr>
</tbody>
</table>

Table 5.2: Summary of the kinematic selections and their corresponding rates of the trigger paths used for the VBF Higgs channel [71].

The single lepton triggers reject a large number of signal events. To recover these events by lowering the kinematic thresholds of the trigger electron and muon, a new trigger path is proposed which makes use of the two forward jets characteristic identification signature (VBF jets).

5.2.2.1 The Algorithm

The VBF jet plus lepton trigger algorithm requires the coincidence of a lepton trigger candidate and two trigger jets that satisfy conditions mimicking the offline selection criteria (see Section 5.2.4). The algorithm requires the following:

- An isolated lepton trigger candidate satisfying the kinematic threshold, \( E_{T,l} \geq E_{T,l}^{\text{thres}} \).

- A pair of jets neither of which is collinear to the lepton and satisfying their respective kinematic thresholds, \( \Delta R_{l,j} > 0.5, E_{T,j_1} \geq E_{T,j_1}^{\text{thres}} \) and \( E_{T,j_2} \geq E_{T,j_2}^{\text{thres}} \).

- A rapidity gap between the two jets greater than a pre-defined value, \( \Delta \eta_{j_1,j_2} > \Delta \eta_{j_1,j_2}^{\text{thres}} \).

- The lepton trigger candidate lying between the two jets in \( \eta \).

- The invariant mass of the two jets to be greater than a pre-defined value, \( m(j_1^{\mu},j_2^{\mu}) > m(j_1^{\mu},j_2^{\mu})^{\text{thres}} \).
The algorithm iterates through all the possible combinations of lepton candidates and jet candidates. For simplicity, the algorithm remains the same for the L1 trigger and the HLT, with the only differences being the candidates used and the selection criteria values.

### 5.2.3 L1 and HLT Configuration

The lepton trigger candidates are reconstructed in the same way for the single isolated lepton trigger paths and the VBF jets plus lepton. They only differ in the value of the kinematic selections. For the L1 trigger objects, this is a trivial decision since the reconstruction algorithms are fixed. For the HLT trigger objects, all other selection criteria were kept the same to enable a fair comparison without the need of analysing individual selection values. The selection values are summarised in Table 5.3 and Table 5.4 for the electron and the muon candidates, respectively. The remaining selection values of the VBF jets plus lepton trigger are summarised in Table 5.5. The VBF jet selections are motivated by the offline jets (see Section 5.2.4).
Table 5.5: VBF jet specific condition thresholds.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e/\gamma$ $E_T$</td>
<td>12 GeV</td>
</tr>
<tr>
<td>$\mu$ $p_T$</td>
<td>8 GeV/c</td>
</tr>
<tr>
<td>Jet$_1$, Jet$_2$ $E_T$</td>
<td>45, 20 GeV</td>
</tr>
<tr>
<td>$\Delta\eta_{j_1,j_2}$</td>
<td>3.4</td>
</tr>
<tr>
<td>Jet invariant mass $m(j_1^\nu,j_2^\nu)$</td>
<td>400 GeV/c$^2$</td>
</tr>
</tbody>
</table>

### 5.2.4 Pre-selected Signal Sample

In order to quantify the improvement in the triggering efficiency of the signal $H \rightarrow \tau\tau \rightarrow l + \tau$-jet using the developed VBF jets plus lepton trigger, an offline pre-selected sample was used. The selection criteria for the various analysis objects were motivated by previous studies [70, 72] and they will be justified in Chapter 6. The threshold values chosen for this study aimed at ensuring high purity for the different analysis objects, with a cumulative efficiency of approximately 5%. The selection criteria and their corresponding values are summarised below.

**Offline reconstructed electron selections:**

- Kinematic: $E_T^e > 15$ GeV.
- Barrel-Endcap gap rejection: $|\eta_e| < 1.4442$ or $|\eta_e| > 1.56$.
- Tracker isolation: no track with $p_T > 1$ GeV within an annulus of $0.015 < \Delta R < 0.5$.
- “Robust” identification criteria for barrel electrons: $H/E < 0.015$, $\Delta\varphi_{in} < 0.02$ rad, $\Delta\eta_{in} < 0.025$, $\sigma_{\eta} < 0.0092$ and for endcap: $H/E < 0.018$, $\Delta\varphi_{in} < 0.02$ rad, $\Delta\eta_{in} < 0.04$, $\sigma_{\eta} < 0.025$.

**Offline reconstructed muon selections:**
5.2 Higgs Boson Trigger Study

- Kinematic: \( p_T^\mu > 10 \text{ GeV/c} \).

- “Global” muon: reconstructed track both in the inner tracker and the muon stations.

- Tracker isolation: no track with \( p_T^{\text{trk}} > 1 \text{ GeV} \) within an annulus of \( 0.01 < \Delta R < 0.5 \).

- HCAL isolation: \( \Sigma E_{\text{HCAL}} < 3 \text{ GeV} \) within an annulus of \( 0.1 < \Delta R < 0.3 \).

- ECAL isolation: \( \Sigma E_{\text{ECAL}} < 1 \text{ GeV} \) within an annulus of \( 0.07 < \Delta R < 0.3 \).

Offline particle-flow \( \tau \)-jet selections:

- Kinematic: \( E_T^\tau > 15 \text{ GeV} \).

- Collinearity: \( \Delta R_{\tau-l} > 0.5 \) for both the electron and the muon.

- Leading Track: a reconstructed track with \( p_T^{\text{ltr}} > 5 \text{ GeV/c} \).

- Tracker isolation: no track with \( p_T^{\text{trk}} > 1 \text{ GeV/c} \) within an annulus of \( 0.15 < \Delta R < 0.5 \).

- Electron and muon rejection: passing electron and muon rejection criteria as described in Section 3.4.1.

Offline VBF jet selections:

- Kinematic: \( E_T^{\text{j1}} > 55 \text{ GeV/c} \) and \( E_T^{\text{j2}} > 25 \text{ GeV/c} \).

- Collinearity: \( \Delta R_{j-l} > 0.5 \), where \( l \) could be a fully reconstructed electron, muon or \( \tau \)-jet.

- Rapidity gap: \( \Delta \eta_{j_1,j_2} > 3.6 \).

- Invariant mass: \( m(j_1^{\mu}, j_2^\nu) > 600 \text{ GeV/c}^2 \).

These selections serve as a good baseline and will be discussed further in Chapter 6.
5.2 Higgs Boson Trigger Study

5.2.5 Simulated Data Sets and Software Environment

The trigger rates for the VBF jets plus lepton trigger were measured using an officially generated PYTHIA 6.4 [73] sample of QCD events with $\hat{p}_T > 15$ GeV/c as a base, where $\hat{p}_T$ is the transverse momentum scale of the process. The data set was originally produced using the CMS simulation and reconstruction software CMSSW 2.1.7 and was later reprocessed in order to incorporate the newer data formats with CMSSW 2.2.1. Due to a problem with the geometry configuration used in the reprocessing, the L1 software emulator failed in a large number of events that accessed the “RAW” muon system information. To solve this issue, the data set was privately reproduced from the generator level information and stored at the local Tier-2 analysis centre at Imperial College. The focus on the early LHC data meant that all data sets available used $\sqrt{s} = 10$ TeV, where $\sqrt{s}$ denotes the centre of mass energy. This study was conducted with the $\mathcal{L} = 2 \times 10^{33}$ cm$^{-2}$s$^{-1}$ instantaneous luminosity regime in mind, at which point the LHC is expected to operate in the designed centre of mass energy, $\sqrt{s} = 14$ TeV. Given that on an average only one third of the available proton energy is carried by the interacting partons, the percentage difference in the total available energy is expected to be less than 10%.

The trigger efficiencies were measured on an offline pre-selected PYTHIA VBF H$\rightarrow$ττ$\rightarrow$l + τ-jet data set with $m_H = 135$ GeV/c$^2$ (as presented in Section 5.2.4). Since there was a problem with the officially produced data set, the sample was privately produced with CMSSW 2.2.13 using the corrected data card and stored at the Imperial College Tier-2 centre. The centre of mass energy was kept at $\sqrt{s} = 10$ TeV for consistency. To simulate τ decay polarisation correctly, the TAUOLA [74] package was used.

Analysis was performed using the CMS reconstruction software CMSSW 2.2.13. The sizes of the two samples were 8,031,780 events for the QCD data sample and 91,747 for the VBF H$\rightarrow$ττ$\rightarrow$l + τ-jet before the pre-selections.
5.2 Higgs Boson Trigger Study

5.2.6 Trigger Rate

Traditionally the rate of the different trigger paths is measured using Minimum Bias events. However, due to the complexity of the VBF jets plus lepton trigger the QCD events with low \( \hat{p}_T \) will not pass. The trigger rates of the VBF jets plus lepton trigger using QCD events with \( \hat{p}_T > 15 \text{ GeV/c} \) for the \( \mathcal{L} = 2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1} \) instantaneous luminosity regime are presented in Table 5.6. The configurations used are described in Sections 5.2.1 and 5.2.3. The overall QCD rate was found to be \((0.6 \pm 0.3) \text{ Hz}\) for the \(2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}\) instantaneous luminosity regime. This rate is considered acceptable given that the HLT rate at the start up will be 100 Hz, while there are plans to upgrade it later on. The available events were scaled to simulate the expected rate of interactions in the \(\mathcal{L} = 2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}\) instantaneous luminosity regime. A drawback of this method is that it cannot correctly calculate the events failing the isolation criteria due to increased detector activity. Therefore, this trigger rate represents the upper limit of the possible rate, while at the same time it includes a fraction of the single lepton’s trigger rate. For comparison, the single isolated lepton trigger rates are summarised in Table 5.7. It is worth noting that the default selection criteria for these triggers were not tuned for the \(2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}\) instantaneous luminosity regime and thus the total rate is higher than the \(\sim 50 \text{ Hz}\) that was shown in Section 5.2.1.

5.2.7 Trigger Efficiency

The efficiency performance of the VBF jets plus lepton trigger on a VBF \( H \rightarrow \tau \tau \rightarrow l + \tau\)-jet pre-selected sample using the selection criteria presented in Section 5.2.4 is shown in Table 5.8. An efficiency of \((76.1 \pm 0.7)\%\) was obtained. As seen from Table 5.8, the efficiency of each individual selection is above 90%. The least efficient selection criterion is the rapidity gap requirement at L1. This is caused by the crude position resolution of the L1 jet finder. The HLT lepton efficiency is also not 100% since more strict quality criteria are required for the offline reconstructed lepton. This is due to some minor differences in the reconstruction of the two objects - mainly in the case on the electron. As discussed in Sections 3.1.2 and 3.1.3, the search windows of compatible pixels in the HLT are smaller than those in the offline.
Table 5.6: Table summarising the performance of the VBF jet plus lepton trigger on QCD events with $\hat{p}_T > 15\text{GeV}\text{GeV/c}$ for the $2\times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ luminosity regime. The error on the rate represents the statistical uncertainty. A total rate of $(0.6 \pm 0.2)\text{Hz}$ was obtained.

<table>
<thead>
<tr>
<th>VBF Jets plus Lepton</th>
<th>Rate (Hz)</th>
<th>Relative Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 Isolated Lepton:</td>
<td>40717±121</td>
<td>0.014</td>
</tr>
<tr>
<td>$E_T^{e/\gamma} &gt; 12\text{GeV}$ or $p_T^{\mu} &gt; 8\text{GeV}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 VBF Jets: $E_T^{h_1} &gt; 45\text{GeV}$, $E_T^{h_2} &gt; 20\text{GeV}$</td>
<td>2130±27</td>
<td>0.052</td>
</tr>
<tr>
<td>L1 Jet Rapidity Gap:</td>
<td>542±14</td>
<td>0.254</td>
</tr>
<tr>
<td>$\Delta \eta_{j_1,j_2} &gt; 3.4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 Lepton between VBF Jets</td>
<td>397±12</td>
<td>0.733</td>
</tr>
<tr>
<td>L1 Jet Invariant Mass:</td>
<td>246±9</td>
<td>0.621</td>
</tr>
<tr>
<td>$m(j_1^\mu,j_2^\nu) &gt; 400\text{GeV}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLT Isolated Lepton:</td>
<td>3.3±1.1</td>
<td>0.013</td>
</tr>
<tr>
<td>$E_T^{e} &gt; 15\text{GeV}$ or $p_T^{\mu} &gt; 10\text{GeV}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLT VBF Jets: $E_T^{h_1} &gt; 43\text{GeV}$, $E_T^{h_2} &gt; 20\text{GeV}$</td>
<td>1.5±0.7</td>
<td>0.444</td>
</tr>
<tr>
<td>HLT Jet Rapidity Gap:</td>
<td>1.1±0.4</td>
<td>0.725</td>
</tr>
<tr>
<td>$\Delta \eta_{j_1,j_2} &gt; 3.4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLT Lepton between VBF Jets</td>
<td>0.9±0.3</td>
<td>0.859</td>
</tr>
<tr>
<td>HLT Jet Invariant Mass:</td>
<td>0.6±0.3</td>
<td>0.682</td>
</tr>
<tr>
<td>$m(j_1^\mu,j_2^\nu) &gt; 400\text{GeV}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.7: Table summarising the performance of the single lepton triggers on QCD events with $\hat{p}_T > 15\text{GeV}\text{GeV/c}$ for the $2\times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ luminosity regime. The error on the rate represents the statistical uncertainty. A total rate of $(62.9 \pm 4.8)\text{Hz}$ was obtained.

<table>
<thead>
<tr>
<th>Isolated Electron OR Muon</th>
<th>Rate (Hz)</th>
<th>Relative Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 Isolated Lepton:</td>
<td>6247±47</td>
<td>0.002</td>
</tr>
<tr>
<td>$E_T^{e/\gamma} &gt; 22\text{GeV}$ or $p_T^{\mu} &gt; 14\text{GeV}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLT Isolated Lepton:</td>
<td>62.9±4.8</td>
<td>0.010</td>
</tr>
<tr>
<td>$E_T^{e} &gt; 26\text{GeV}$ or $p_T^{\mu} &gt; 19\text{GeV}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 5.2 Higgs Boson Trigger Study

<table>
<thead>
<tr>
<th>VBF Jets plus Lepton</th>
<th>Cumulative Efficiency</th>
<th>Relative Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 Isolated Lepton: $E_T^{e/\gamma} &gt; 12\text{ GeV}$ or $p_T^{\mu} &gt; 8\text{ GeV}$</td>
<td>0.959±0.003</td>
<td>0.959</td>
</tr>
<tr>
<td>L1 VBF Jets: $E_{T1}^{j} &gt; 45\text{ GeV}$, $E_{T2}^{j} &gt; 20\text{ GeV}$</td>
<td>0.950±0.004</td>
<td>0.990</td>
</tr>
<tr>
<td>L1 Jet Rapidity Gap: $\Delta \eta_{j1,j2} &gt; 3.4$</td>
<td>0.864±0.006</td>
<td>0.910</td>
</tr>
<tr>
<td>L1 Lepton between VBF Jets</td>
<td>0.830±0.006</td>
<td>0.961</td>
</tr>
<tr>
<td>L1 Jet Invariant Mass: $m(j_{T1}^{\mu},j_{T2}^{\nu}) &gt; 400\text{ GeV}$</td>
<td>0.825±0.007</td>
<td>0.995</td>
</tr>
<tr>
<td>HLT Isolated Lepton: $E_T^{e} &gt; 15\text{ GeV}$ or $p_T^{\mu} &gt; 10\text{ GeV}$</td>
<td>0.779±0.007</td>
<td>0.944</td>
</tr>
<tr>
<td>HLT VBF Jets: $E_{T1}^{j} &gt; 43\text{ GeV}$, $E_{T2}^{j} &gt; 20\text{ GeV}$</td>
<td>0.779±0.007</td>
<td>1.000</td>
</tr>
<tr>
<td>HLT Jet Rapidity Gap: $\Delta \eta_{j1,j2} &gt; 3.4$</td>
<td>0.767±0.007</td>
<td>0.985</td>
</tr>
<tr>
<td>HLT Lepton between VBF Jets</td>
<td>0.762±0.007</td>
<td>0.993</td>
</tr>
<tr>
<td>HLT Jet Invariant Mass: $m(j_{T1}^{\mu},j_{T2}^{\nu}) &gt; 400\text{ GeV}$</td>
<td>0.761±0.007</td>
<td>0.998</td>
</tr>
</tbody>
</table>

**Table 5.8:** Table summarising the performance of the VBF jet plus lepton trigger on pre-selected VBF $H \rightarrow \tau \tau \rightarrow l + \tau$-jet events. The pre-selection criteria are presented in Section 5.2.4. The error on the rate represents the statistical uncertainty. A efficiency of $(76.1\pm0.7)\text{Hz}$ was obtained.

Reconstruction. Also, the tracking algorithm used for the HLT electron (KF) can fail in some cases where the GSF tracking algorithm successfully reconstructs the track. The offline reconstructed jets seem to be in good agreement with the trigger jets, especially the HLT ones. This is also assisted by the fact that they use the same type of energy calibration (L2+L3).

A combination of the VBF jets plus lepton trigger with the single lepton triggers can improve the triggering efficiency further. Table 5.9 summarises the additional efficiency by using a logical OR combination of the single lepton triggers and the VBF jets plus lepton trigger as well as the efficiency of the single lepton triggers. A gain of $\sim 33\%$ over the current triggers was observed when the combined trigger was used. This is a substantial improvement which will be even more important for higher luminosity regimes where the single lepton triggers will be forced to have an even higher kinematic threshold.
5.3 Summary and Conclusions

The basic characteristics of the VBF $H \rightarrow \tau \tau \rightarrow l + \tau$-jet channel for $m_H = 135 \text{ GeV}/c^2$ as well as a comparison of the VBF jets and the properties of the decay products between 10 TeV and 14 TeV centre of mass energies were presented. The study was performed assuming the 10 TeV centre of mass energy operation regime, mainly due to the availability of the background samples needed for the analysis presented in the next chapter. It should be noted that the VBF jets plus lepton trigger described in this chapter cannot be implemented using the current structure of L1. However, given that the trigger provides sufficient gain all selections apart from the the VBF jets invariant mass could be easily incorporated in the L1 programmable integrated circuits. In this case the VBF jets invariant mass selection can be omitted and a higher $\Delta \eta$ selection can be used since the two are correlated. The algorithm developed for the HLT could run on the currently implemented trigger as a separate path.

A detailed description of the reconstruction algorithms for all the analysis objects used at L1, HLT and offline level was presented. The current trigger paths as well as their default kinematic thresholds were discussed in Section 5.2.1. The motivation and performance of the VBF jets plus lepton trigger for the $2 \times 10^{33} \text{cm}^{-2}\text{s}^{-1}$ luminosity regime on VBF $H \rightarrow \tau \tau \rightarrow l + \tau$-jet and QCD data samples were also presented. Given the trigger configuration described in Section 5.2.3 and the offline pre-selection conditions (Section 5.2.4), a signal efficiency of $(76.1 \pm 0.7)\%$ with a QCD rate of $(0.6 \pm 0.3) \text{ Hz}$ for the $2 \times 10^{33} \text{cm}^{-2}\text{s}^{-1}$ instantaneous luminosity regime.

<table>
<thead>
<tr>
<th>Trigger Type</th>
<th>Single Electron OR Muon Trigger</th>
<th>VBF Jets plus Lepton</th>
<th>Single Electron OR Muon OR VBF Jets plus Lepton Trigger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency</td>
<td>0.647(\pm)0.008</td>
<td>0.761(\pm)0.008</td>
<td>0.861(\pm)0.007</td>
</tr>
<tr>
<td>Efficiency Gain ($\Delta \epsilon/\epsilon$)</td>
<td>(~18%)</td>
<td>(~33%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.9: Additional efficiency by using VBF jets plus lepton trigger and a combination of the VBF jets plus lepton trigger with the single isolated lepton triggers, compared to using only the single isolated lepton triggers.
was obtained. Finally, the efficiency of the logical OR between the single isolated lepton triggers and the VBF jets plus lepton trigger was compared to that of the single isolated lepton triggers alone. It was shown to increase from \((64.7 \pm 0.8)\%\) for the current trigger to \((86.1 \pm 0.7)\%\) for the combined trigger - an improvement of \(\sim 33\%\). Since the rates scale linearly with the instantaneous luminosity for a certain set of selections, the kinematic thresholds for the design LHC luminosity regime, \(10^{34}\text{cm}^{-2}\text{s}^{-1}\), will be forced to rise. Given the very low rate of the developed VBF jets plus lepton trigger, it will be able to operate in that luminosity regime with only minor adjustments.
Chapter 6

Selection of VBF $H \rightarrow \tau \tau \rightarrow l + \tau$-jet using VBF Jets plus Lepton Trigger

As demonstrated in the previous chapter, the use of the VBF jets plus lepton trigger can increase the triggering efficiency of the VBF Higgs boson offline pre-selected events significantly compared to the single isolated triggers, while maintaining a very low QCD rate. This fact can prove very important for the detection of the Higgs boson in the higher LHC luminosity regimes ($2 \times 10^{33}$ and $10^{34} \text{cm}^{-2}\text{s}^{-1}$). However, the detection of the Higgs boson poses a far greater challenge than just the online selection efficiency. Its small production cross section (see Section 5.1.1) dictates the need for a set of stringent selection criteria that can suppress the large background processes sufficiently. The main backgrounds of the VBF $H \rightarrow \tau \tau \rightarrow l + \tau$-jet signal include processes with the same final state $Z \rightarrow \tau \tau + 2$ jets as well as processes with large cross sections that can mimic the signal final state by mis-reconstruction of the lepton or the $\tau$-jet, such as QCD events with a lepton or photon, W+jets and $t\bar{t}$+jets events.

This chapter focuses on evaluating the overall gain of the use of the VBF jets plus lepton trigger for the VBF $H \rightarrow \tau \tau \rightarrow l + \tau$-jet process as well as the implications of its use on the background processes. The set of selection criteria chosen for this analysis along with the motivation behind them will be presented.
6.1 Simulated Data Sets and Software Environment

The signal and background processes were simulated using the PYTHIA 6.4 \cite{73} and the MadGraph v4 \cite{75} event generators. Unlike Pythia, MadGraph automatically creates the amplitudes for the relevant subprocesses by using their leading order matrix elements. Its output is then interfaced with PYTHIA, which is responsible for the simulation of the parton showers and the hadronisation process. The TAUOLA package was used to simulate the $\tau$ polarisation faithfully.

Due to the problem with the L1 muon trigger mentioned before (see Section\ref{sec:5.2.5}), all the officially produced data samples were privately reprocessed from the generator level information using the CMS detector simulation and reconstruction software CMSSW 2.2.13. The data were stored at the local Tier-2 analysis centre at Imperial College where the analysis was performed.

The $Z$+jets background with $Z \rightarrow ll$, where $l = e, \mu, \tau$, was generated using the MadGraph event generator and the parton shower and hadronisation processes were performed in PYTHIA. The $\tau$ decays were simulated using the TAUOLA package. To reduce the file size without affecting the final offline selection, the data sample was required to have an invariant mass of the lepton pair of $m(l,l) > 50$ GeV/c$^2$ at the generator level. The $W$+jets background was generated using the same techniques as the $Z$+jets. The $W^\pm$ bosons were also forced to decay leptonically at the generator level while no other pre-selections were applied. The $t\bar{t}$+jets events were also generated using the MadGraph and PYTHIA event generators. The $t$ quarks in this case were each forced to decay to the dominant decay mode of a $W^\pm$ boson and a $b$ quark at the generator level, while the $W^\pm$ boson decays were not constrained. No other pre-selection was applied.

Two kinds of QCD samples were studied in this analysis in order to investigate the effect of the processes on the electron identification and isolation criteria. The QCD background processes were generated using the PYTHIA event generator. The $\gamma$+jets background was generated using a $p_T > 15$ GeV/c requirement. The
other data samples involved \(b\) and \(c\) quarks with an electron in the final state. The BCtoE \((b/c \rightarrow e)\) samples were generated in \(\hat{p}_T\) bins of \(\hat{p}_T \in (20,30)\) GeV/c, \(\hat{p}_T \in (30,80)\) GeV/c and \(\hat{p}_T \in (80,170)\) GeV/c. The \(b/c \rightarrow e\) sample events were required to contain a generator level electron that originated from a \(b\) or a \(c\) quark with \(E_T > 10\) GeV and within the tracker acceptance \(|\eta| < 2.5\). No QCD sample with muons in the final state was used due to lack of availability.

Finally, the signal samples of VBF \(H \rightarrow \tau\tau\) with \(m_H = 115\) and \(135\) GeV/c\(^2\) were generated using the PYTHIA event generator. As mentioned in Section 5.2.5, the signal sample with \(m_H = 135\) GeV/c\(^2\) had the taus decaying to a lepton/\(\tau\)-jet pair and was privately produced. For the sample with \(m_H = 115\) GeV/c\(^2\), the tau decays were not constrained and it was reprocessed as described earlier. A summary of all the data samples used in this analysis together with their cross sections and integrated luminosities is shown in Table 6.1.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Event Generator</th>
<th>(\epsilon_{\text{flt}})</th>
<th>(\sigma \times \epsilon)</th>
<th>(\mathcal{L}_{\text{int}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBF (H \rightarrow \tau\tau) ((l\tau\text{-jet}, m_H = 135) GeV/c(^2))</td>
<td>PYTHIA</td>
<td>-</td>
<td>45.9 fb</td>
<td>(2 \times 10^3) fb(^{-1})</td>
</tr>
<tr>
<td>VBF (H \rightarrow \tau\tau) ((m_H = 115) GeV/c(^2))</td>
<td>PYTHIA</td>
<td>-</td>
<td>195.6 fb</td>
<td>1070 fb(^{-1})</td>
</tr>
<tr>
<td>(Z + \text{jets} (Z \rightarrow ll, \text{no } \nu\bar{\nu}))</td>
<td>MadGraph</td>
<td>-</td>
<td>3.7 nb</td>
<td>0.334 fb(^{-1})</td>
</tr>
<tr>
<td>(W + \text{jets} (W \rightarrow l\bar{\nu}_l))</td>
<td>MadGraph</td>
<td>-</td>
<td>40 nb</td>
<td>0.242 fb(^{-1})</td>
</tr>
<tr>
<td>(t\bar{t} + \text{jets} (t \rightarrow bW))</td>
<td>MadGraph</td>
<td>-</td>
<td>317 pb</td>
<td>3.2 fb(^{-1})</td>
</tr>
<tr>
<td>QCD (b/c \rightarrow e) ((\hat{p}_T \in (20, 30)) GeV/c)</td>
<td>PYTHIA</td>
<td>(4.8 \times 10^{-4})</td>
<td>192 nb</td>
<td>10.4 pb(^{-1})</td>
</tr>
<tr>
<td>QCD (b/c \rightarrow e) ((\hat{p}_T \in (30, 80)) GeV/c)</td>
<td>PYTHIA</td>
<td>(2.4 \times 10^{-3})</td>
<td>240 nb</td>
<td>8.3 pb(^{-1})</td>
</tr>
<tr>
<td>QCD (b/c \rightarrow e) ((\hat{p}_T \in (80, 170)) GeV/c)</td>
<td>PYTHIA</td>
<td>(1.2 \times 10^{-2})</td>
<td>23 nb</td>
<td>46 pb(^{-1})</td>
</tr>
<tr>
<td>(\gamma + \text{jets} (\hat{p}_T &gt; 15) GeV/c)</td>
<td>PYTHIA</td>
<td>-</td>
<td>288 nb</td>
<td>3.5 pb(^{-1})</td>
</tr>
</tbody>
</table>

Table 6.1: Summary of the MC samples used in this analysis. \(\epsilon_{\text{flt}}\) denotes the generator level pre-selection efficiency, \(\sigma\) the process cross section and \(\mathcal{L}_{\text{int}}\) the integrated luminosity of the sample.
6.2 Trigger

The single isolated lepton (electron, muon) triggers and the VBF jets plus lepton trigger were used in this analysis and were discussed in detail in the previous chapter (see Sections 5.2.1, 5.2.2, 5.2.3). The efficiencies of these triggers for the signal and all the background samples are summarised in Tables 6.2, 6.3 and 6.4. By comparing Tables 6.2 and 6.3 it can be seen that the efficiency of the VBF jets plus lepton trigger is, as expected, lower than the single isolated lepton triggers for all the samples. This is due to the requirement of the two jets with a rapidity gap and invariant mass constraints on samples that were not pre-selected using any offline criteria. It can also be seen that the efficiency of the individual selections for the VBF Higgs with \( m_H = 115 \text{ GeV}/c^2 \) is consistently lower than that of the VBF Higgs with \( m_H = 135 \text{ GeV}/c^2 \). This can be understood since the two taus in the VBF Higgs sample with \( m_H = 115 \text{ GeV}/c^2 \) are not constrained on their decay modes and therefore the final state of two \( \tau \)-jets is also included, unlike in the VBF Higgs sample with \( m_H = 135 \text{ GeV}/c^2 \). However, since the VBF Higgs cross section for \( m_H = 115 \text{ GeV}/c^2 \) is higher than that for \( m_H = 135 \text{ GeV}/c^2 \) there are more events passing the L1 and HLT selection criteria. The Z+jets sample has an increased efficiency compared to the W+jets sample mainly due to the existence of a larger number of electrons and muons in its final state, which subsequently enhances the reconstruction efficiency of at least one lepton in the event while the second lepton can be tagged as the second VBF jet. Once the existence of the two additional jets is ensured in the L1 trigger, the selection efficiencies follow a similar pattern for both samples. The relatively high efficiency of the \( tt\bar{t} \) sample on the VBF jets plus lepton trigger can be explained by the existence of the two \( b \) jets in the event. These jets do not follow the same kinematic distributions as the VBF jets and their high selection efficiency is due to the relaxed selection criteria applied.

The two QCD samples (\( b/c \rightarrow e \) and \( \gamma+\text{jets} \)) were mainly included to provide an estimate on the additional trigger rate. The \( b/c \rightarrow e \) sample provides an HLT trigger rate of \( \sim 0.16 \text{ Hz} \) and accounts approximately for 1/4 of the total QCD rate (the other 3/4 coming from \( b/c \rightarrow \mu \) and QCD plus lepton events). This is consistent with
Table 6.2: Cumulative cross sections of the signal and background events for each step of the VBF jets plus lepton trigger path. The numbers in the brackets denote the efficiency of the selection criterion against the previous one.
Table 6.3: Cumulative cross sections of the signal and background events for single isolated lepton trigger paths. The numbers in the brackets denote the efficiency of the selection criterion against the previous one. The Z and W± boson HLT rates are in agreement with those presented in the Technical Design Report (see text for more information). The high efficiency both in signal and background is due to the absence of the VBF jets selection criteria.
Table 6.4: Cumulative cross sections of the signal and background events for the combined VBF jets plus lepton and single isolated lepton trigger paths. The numbers in the brackets denote the efficiency of the selection criterion against the previous one.
the 0.6 ± 0.3 Hz calculated in the previous chapter (0.6/4=1.5). The γ+jets rate is three orders of magnitude smaller and it is not expected to affect the HLT trigger rate. Due to the relatively small number of QCD events passing the L1 and HLT selection criteria, these samples were only used to check the electron identification and isolation criteria and were not considered in the rest of the offline selection analysis.

In Table 6.3 the HLT rates of the single isolated lepton triggers for the W+jets and Z+jets processes of ∼21 Hz and ∼3 Hz are consistent with those provided by the CMS physics performance Technical Design Report (TDR) [71] of ∼23.7 Hz and ∼2.9 Hz, respectively. Finally in Table 6.4 the efficiencies of the combined single lepton and VBF jets plus lepton trigger can be seen. This illustrates the fact that the additional efficiency for the background processes of the combined trigger compared to the single lepton triggers is very small.

### 6.3 Offline Selection

Offline selection criteria and threshold values were motivated by those used in previous studies [70, 72]. However the threshold values were adjusted as well as some extra criteria were used to better fit the characteristics of this study. The chosen threshold values were selected, mainly visually, by plotting the quantity for the signal and the background processes. The main criterion was to ensure a relatively high signal efficiency (>75-80%) while reducing the background events as much as possible.

#### 6.3.1 Electron Selection

The primary objective of the electron selection criteria is to maintain a high efficiency of signal events and at the same time reject objects faking electrons.
6.3 Offline Selection

6.3.1.1 Identification

Tight electron identification criteria are needed in order to reduce the number of jets faking electrons and select well reconstructed electrons. The offline electron identification criteria have been described in detail in Section 3.1.3. The distributions of $E/P$, $H/E$, $\Delta \varphi_{\text{in}}$, $\Delta \eta_{\text{in}}$ and $\sigma_{\eta\eta}$ are shown in Figures 6.1, 6.2, 6.3, 6.4 and 6.5 respectively for offline reconstructed electrons in the barrel and the endcap regions matched to MC electrons, MC $\tau$-jets and MC QCD jets. The vertical dashed lines denote the position of the thresholds for the different quantities. As seen from Figures 6.1 and 6.2 the $E/P$ and $H/E$ quantities are very effective discriminants against jets, while the $\sigma_{\eta\eta}$ (see Figure 6.5) plays a more important role in the barrel region mainly due to the geometry of the ECAL cluster ($\eta - \varphi$). The $\Delta \varphi_{\text{in}}$ and $\Delta \eta_{\text{in}}$ variables as seen in Figures 6.3 and 6.4 can further increase the purity of offline reconstructed electrons while maintaining a very high efficiency for the true electrons. Table 6.5 summarises the values of the selections and the efficiencies for reconstructed electrons with $E_T > 15$ GeV matched with MC electrons, $\tau$-jets and QCD jets. An efficiency of 69.8% is obtained for the true electrons while rejections of more than 97% and 99% are achieved for the $\tau$-jets and QCD jets, respectively.

The relatively low efficiency of the $E/P$ criterion is mainly due to electrons that lose a substantial amount of energy through bremsstrahlung radiation which the GSF tracking algorithm (see Section 3.1.3) can not account for.

The $\gamma + \text{jets}$ sample provides a different source of “fake” electrons. These come from isolated photons that convert into a pair of electrons in the first layers of the charged tracker. The identification algorithm for converted photons is described in [76]. Since photons are massless, the opening angle between the two electrons is small and is used for calculating the invariant mass of the electron pair. The angular difference is parametrised as $\Delta \cot \theta$, where $\theta$ is the one of the detector coordinate system. All standard KF charged tracks with $p_T > 0.3$ GeV/c in a radius $\Delta R < 0.3$ around the GSF electron track are considered. An excess of tracks with $\Delta \cot \theta \sim 0$ is expected for the converted photons. The threshold is set to $|\Delta \cot \theta| < 0.045$ and the number of tracks satisfying this condition are counted ($N_{\text{conv}}$). When $N_{\text{conv}} > 1$, to
Table 6.5: Table summarising the selection efficiencies of different electron identification criteria for reconstructed electrons matching MC electron, MC $\tau$-jets and MC QCD jets.

<table>
<thead>
<tr>
<th>Selection</th>
<th>Relative Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>electron $\tau$-jet QCD jet</td>
</tr>
<tr>
<td>$0.8&lt;E/P&lt;1.2$ (barrel), $0.8&lt;E/P&lt;1.5$ (endcap)</td>
<td>0.757 0.195 0.149</td>
</tr>
<tr>
<td>$H/E&lt;0.015$ (barrel), $H/E&lt;0.018$ (endcap)</td>
<td>0.965 0.376 0.197</td>
</tr>
<tr>
<td>$</td>
<td>\Delta\varphi_{\text{in}}</td>
</tr>
<tr>
<td>$</td>
<td>\Delta\eta_{\text{in}}</td>
</tr>
<tr>
<td>$\sigma_{\eta\eta}&lt;0.012$ (barrel), $\sigma_{\eta\eta}&lt;0.025$ (endcap)</td>
<td>0.98 0.644 0.362</td>
</tr>
<tr>
<td>Total Efficiency</td>
<td>69.8% 2.8% 0.6%</td>
</tr>
</tbody>
</table>

Figure 6.1: E/P distributions of reconstructed electrons matching MC electrons (solid blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines.
6.3 Offline Selection

Figure 6.2: \(H/E\) distributions of reconstructed electrons matching MC electrons (solid blue line), MC \(\tau\)-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines.

Figure 6.3: \(\Delta \varphi_{\text{in}}\) distributions of reconstructed electrons matching MC electrons (solid blue line), MC \(\tau\)-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines.
6.3 Offline Selection

Figure 6.4: $\Delta\eta_n$ distributions of reconstructed electrons matching MC electrons (solid blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines.

Figure 6.5: $\sigma_{\eta\eta}$ distributions of reconstructed electrons matching MC electrons (solid blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) in the barrel (left) and the endcap (right). The threshold values are shown by the vertical dashed black lines.
Table 6.6: Table summarising the photon conversion selection efficiencies. Electrons used had already satisfied the identification criteria described in this section.

<table>
<thead>
<tr>
<th>Selection</th>
<th>electron</th>
<th>photon</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{conv}} = 1$</td>
<td>0.96</td>
<td>0.43</td>
</tr>
<tr>
<td>Pixel hit</td>
<td>0.92</td>
<td>0.21</td>
</tr>
</tbody>
</table>

account for the KF track matching the GSF track, the electron is rejected. To further reduce the number of electrons originating from converted photons, a requirement for a hit in the first layer of the pixel tracker can be used. The performance of these selections are summarised in Table 6.6 for reconstructed electrons that pass the identification criteria described above, with $E_T > 15 \text{ GeV}$ matched to MC electrons and photons. An efficiency of 92% is obtained for the true electrons while converted electrons have an efficiency of 21%.

6.3.1.2 Isolation

To further reduce the number of jets faking electrons and electrons coming from QCD processes, isolation criteria are applied in three sub-detectors - the charged tracker, the ECAL and the HCAL. The quantities used were the $\Sigma p_T^{\text{trk}}$, the $\Sigma E_T^{\text{ECAL}}$ and the $\Sigma E_T^{\text{HCAL}}$ which denote the summation of the track $p_T$’s, the ECAL crystal $E_T$’s and HCAL tower energies respectively for the corresponding isolation annuli. All plots presented in this section correspond to objects that have already passed the identification criteria described in Section 6.3.1.1. The isolation in the charged tracker used tracks with $p_T > 1 \text{ GeV/c}$ in an annulus between $0.015 < \Delta R < 0.6$. This selection criterion was particularly effective for electrons coming from QCD processes as seen in Figure 6.6. For example, in $b/c \rightarrow e$ the electrons pass the identification criteria but are not isolated. The ECAL isolation is performed in an annulus between $0.045 < \Delta R < 0.6$ in the barrel and $0.1 < \Delta R < 0.6$ in the endcap excluding also a $\phi$ window of 0.5 rad to account for the bremsstrahlung radiation. The ECAL isolation is mainly used to reject electrons with photons inside the isolation annulus. The distributions of $\Sigma E_T^{\text{ECAL}}$ for the various objects are shown in Figure 6.7. The HCAL isolation is performed in a cone of radius $\Delta R < 0.5$ and the distributions of $\Sigma H^{\text{HCAL}}$. 
6.3 Offline Selection

Table 6.7: Table summarising the selection efficiencies of different electron isolation criteria for signal reconstructed electrons matching MC electron, MC \( \tau \)-jets, MC QCD jets as well as MC matched electrons coming from \( b/c \rightarrow e \) process. Electrons included in the table have already satisfied the electron identification criteria presented in the previous section.

<table>
<thead>
<tr>
<th>Selection</th>
<th>Relative Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Sigma p_T^{\text{trk}} &lt; 1 \text{ GeV/c} )</td>
<td>0.757</td>
</tr>
<tr>
<td>( \Sigma E_{\text{ECAL}}^{T} &lt; 2 \text{ GeV (barrel)} )</td>
<td>0.892</td>
</tr>
<tr>
<td>( \Sigma E_{\text{ECAL}}^{T} &lt; 2.5 \text{ GeV (endcap)} )</td>
<td>0.984</td>
</tr>
<tr>
<td>( \Sigma H_{\text{HCAL}}^{T} &lt; 2 \text{ GeV} )</td>
<td>9.84</td>
</tr>
<tr>
<td>Total Efficiency</td>
<td>66.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>electron (signal)</th>
<th>electron ((b/c \rightarrow e))</th>
<th>(\tau)-jet</th>
<th>QCD jet</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.066</td>
<td>0.48</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>0.255</td>
<td>0.627</td>
<td>0.464</td>
<td></td>
</tr>
<tr>
<td>0.798</td>
<td>0.624</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>1.3%</td>
<td>18.8%</td>
<td>1.6%</td>
<td></td>
</tr>
</tbody>
</table>

are shown in Figure 6.8. It should be noted that the HCAL isolation is correlated with the H/E criterion used in the electron identification and therefore it does not have the expected rejection power against QCD jet events.

The performance of the isolation criteria and their corresponding thresholds for the different objects are summarised in Table 6.7. The efficiency for \( b/c \rightarrow e \) electrons passing the isolation criteria is of the order of 1%. The total efficiency for the signal electrons passing the identification and isolation criteria is 46.4%. The corresponding efficiency for the \( \tau \)-jets is 0.5% and that for the QCD jets is of the order of \( 10^{-4} \).

6.3.1.3 Kinematic selections

The transverse energy and the pseudorapidity of the offline reconstructed electrons for the signal and the background processes are shown in Figure 6.9. The offline electrons have been selected using the identification and isolation criteria discussed in Sections 6.3.1.1 and 6.3.1.2, while the requirement of a single electron satisfying these criteria in the event is imposed. In the left plot of Figure 6.9, it can be seen that the electron transverse energy distribution in the W+jets background peaks around 40 GeV before dropping rapidly. This is due to the W\(^\pm\) boson mass (\( m_W \sim 80 \text{ GeV/c}^2 \)) limiting the available energy for the electron in the \( W \rightarrow e\bar{\nu}_e \) two-body decay. The negative tail of the distribution is populated by the forward electrons as well as the electrons originating from the \( \tau \) decay of the W\(^\pm\) boson.
6.3 Offline Selection

**Figure 6.6:** Tracker isolation distributions of reconstructed electrons matching MC electrons (solid dark blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) as well as MC matched electrons coming from the $b/c \rightarrow e$ process (dash-dot light blue line) in the barrel (left) and the endcap (right). Electrons included in the plots have already satisfied the electron identification criteria presented in Section 6.3.1.1. The threshold value is shown by the vertical back line.

**Figure 6.7:** ECAL isolation distributions of reconstructed electrons matching MC electrons (solid dark blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) as well as MC matched electrons coming from the $b/c \rightarrow e$ process (dash-dot light blue line) in the barrel (left) and the endcap (right). Electrons included in the plots have already satisfied the electron identification criteria presented in Section 6.3.1.1. The threshold values are shown by the vertical dashed black lines.
6.3 Offline Selection

Figure 6.8: HCAL isolation distributions of reconstructed electrons matching MC electrons (solid dark blue line), MC $\tau$-jets (dashed red line) and MC QCD jets (dotted purple line) as well as MC matched electrons coming from the $b/c \rightarrow e$ process (dash-dot light blue line) in the barrel (left) and the endcap (right). Electrons included in the plots have already satisfied the electron identification criteria presented in Section 6.3.1.1. The threshold values is shown by the vertical dashed black lines.

Figure 6.9: On the left $E_T$ distributions of offline reconstructed electrons for the signal and background processes. On the right $\eta$ distributions of offline reconstructed electrons for the signal and background processes. Electrons included in these plots have already satisfied the electron identification and isolation criteria presented earlier. The $E_T$ threshold value is shown by the vertical dashed black line. The drop of efficiency in $\eta$ is due to the ECAL calorimeter gaps.
The electron transverse energy distribution in the Z+jets background is very similar to that of the W+jets process. The only differences are that it drops after $E_T = m_Z/2 \sim 45$ GeV and the peak is suppressed due to the single electron requirement which affects the $Z \rightarrow ee$ decay channel.

The offline reconstructed electron pseudorapidity distributions seen in the right plot of Figure 6.9 have two regions with reduced number of entries around $\eta = 0$ and $|\eta| = 1.5$. This is an effect of the gaps in the CMS electromagnetic calorimeter at the centre of the detector and at the transition from the barrel to the endcap region. In these regions, the electrons are most likely to fail the identification criteria due to energy leakage, either because the cluster and track matching criteria or the H/E criterion fails. In order to maintain a high efficiency in signal electron reconstruction, the $E_T$ threshold is set at 15 GeV.

### 6.3.2 Muon Selection

The muon selection criteria do not need to be as stringent as those for the electrons. The muon offline reconstruction is described in detail in Section 3.1.3. To ensure high purity, only muon candidates with a reconstructed track with $|\eta| < 2.4$, both in the inner silicon tracker and the muon system were selected. This criterion will be referred to as the identification criterion. Also a charged tracker isolation was applied in a cone of radius $0.01 < \Delta R < 0.5$ using tracks with $p_T > 1$ GeV/c. Muon candidates with any track within the isolation cone were rejected.

Figure 6.10 shows the reconstruction efficiency as a function of the true $p_T$ (left plot) and the true $\eta$ (right plot) of muon candidates passing the identification and isolation criteria with $p_T > 10$ GeV/c and $|\eta| < 2.4$ matched to MC muons with $p_T > 5$ GeV/c and $|\eta| < 2.4$. It can be seen that the efficiency remains about 80% for the full range of muon transverse momentum ($p_T > 10$ GeV/c) while the lower efficiency in the $\eta$ plot is due to muons with $5$ GeV/c < $p_T < 7$ GeV/c which do not reach the muon detector system.
6.3 Offline Selection

Figure 6.10: Muon reconstruction efficiency as a function of the MC matched \( p_T \) (left) and \( \eta \) (right). Efficiency remains around 80% \( p_T > 10 \text{ GeV/c} \), while lower momentum muons result in the apparent drop in efficiency vs \( \eta \). Muons included in these plots have satisfied the identification and isolation criteria described in the text.

The purity of the muon candidates that have passed the identification and isolation criteria is shown in Figure 6.11 as a function of the reconstructed \( p_T \) and the reconstructed \( \eta \). It can be seen that the achieved purity is more than 99.9% for the full spectrum of the muon candidate transverse momentum and \( \eta \).

The offline reconstructed muons’ transverse momentum and the pseudorapidity distributions are very similar to these of the electrons and can be seen in Figure 6.12. In the left plot of Figure 6.12, it can be seen that the high muon reconstruction efficiency implies that very often both muons in the \( Z \to \mu\mu \) channel are reconstructed. Thus the requirement of a single muon in the event provides an extra suppression factor (see Section 6.3.1.3). Due to the almost uniform muon efficiency across \( \eta \) (see Figure 6.10) the signal muon \( \eta \) distribution seen in the right plot of Figure 6.12 resembles that of the MC muon \( \eta \) distribution shown in previous the chapter (Figure 5.3(c)). The \( p_T \) threshold for the offline muon was set at 10 GeV/c.

6.3.3 \( \tau \)-jet Selection

The \( \tau \)-jets used in this analysis were produced by the Particle-Flow (PFlow) reconstruction algorithm described in Section 3.4.1. The values of the different selection
### 6.3 Offline Selection

Figure 6.11: Reconstruction purity for muons that have satisfied the identification and isolation criteria described in the text as a function of the reconstructed $p_T$ (left) and $\eta$ (right). The muon purity remains over 99.9% for the full range of $p_T$ and $\eta$ of the reconstructed muons.

Figure 6.12: On the left $p_T$ distributions of offline reconstructed muons for the signal and background processes. On the right $\eta$ distributions of offline reconstructed muons for the signal and background processes. Muons included in these plots have already satisfied the electron identification and isolation criteria presented earlier. The $p_T$ threshold value is shown by the vertical back line.
criteria were chosen to minimise the number of fake $\tau$-jets, while maintaining a reasonable efficiency for the true $\tau$-jets. The $\tau$-jet $E_T$ selection value was set at $E_T > 15\text{ GeV}$. The identification criteria applied are very loose, aiming mainly to reject QCD jets by requiring a PFlow charged hadron with $p_T^{\text{ldg}} > 5\text{ GeV/c}$. Also, the number of PFlow charged hadrons within the signal cone ($\Delta R_s = 0.07$) is counted and only $\tau$-jets with one or three PFlow candidates are selected (one or three prong $\tau$-jets). Furthermore, in order to reject events where the electron or the muon in the event is identified as a $\tau$-jet, a collinearity test is performed between the reconstructed lepton and the $\tau$-jet and only $\tau$-jets with $\Delta R_{lt} > 0.3$ are selected. Finally, $\tau$-jets were selected only in the pseudorapidity region of $|\eta| < 2.4$. This was done to avoid muon reconstruction inefficiency for $|\eta| > 2.4$, which hampers the muon rejection algorithm (discussed later).

The $\tau$-jet isolation was performed using the PFlow charged hadrons for tracker isolation and the PFlow $\gamma$ candidates for ECAL isolation. The isolation annuli for both criteria were chosen to be $\Delta R_s < \Delta R < \Delta R_{\text{iso}}$, with $\Delta R_s = 0.07$ as mentioned in the previous paragraph and $\Delta R_{\text{iso}} = 0.5$. Only PFlow $\gamma$ candidates with $E_T > 1.5\text{ GeV}$ were considered for the PFlow $\gamma$ isolation, while for the PFlow hadron isolation only PFlow candidates with $p_T^{\text{trk}} > 1\text{ GeV/c}$ were considered. Table 6.8 summarises the identification and isolation efficiencies for the $\tau$-jets with $E_T > 15\text{ GeV}$ matched to MC $\tau$-jets and MC QCD jets.

Rejecting electrons and muons which are reconstructed as $\tau$-jets is very important especially for background processes such as $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$. The leptons in these cases are well isolated in the charged tracker and the ECAL, thus making it very easy to pass the $\tau$-jet identification and isolation criteria.

As discussed in Section 3.4.1, a number of variables can be used to perform electron rejection. The ones used in this study are the ratios of the ECAL and HCAL energy of the leading charged hadron over the corresponding track transverse momentum ($E/P$ and $H/P$ respectively). Figure 6.13 shows the reconstructed $\tau$-jet distributions of $E/P$ and $H/P$ matched to true MC $\tau$-jets and electrons. In Figure 6.13(a), it can be seen that the $E/P$ distribution for electrons peaks around one as expected.
A number of electrons with E/P close to zero can be noticed. These are mainly electrons that have fallen into the ECAL detector cracks and therefore deposited a minimum amount of energy in the ECAL. To reject these electrons, an additional selection is performed rejecting all $\tau$-jets that coincide with an ECAL crack. In Figure 6.13(b) it can be seen that most electrons deposit a very small fraction of their energy in the HCAL. Since the $\tau$-jet leading charged hadron is predominantly a pion, the shape of the H/P distribution resembles the H/P distribution extracted from single pions as seen in Chapter 3. The pions that start their hadronic shower in the ECAL form the low H/P peak, while the ones that passed as MIPs form the peak close to one. The values chosen for the E/P and H/P selections were 0.8 and 0.2, respectively. The efficiencies of the different $\tau$-jet selection criteria applied on $\tau$-jets with $E_T^{\tau-jet} > 15$ GeV matched to a MC electrons are summarised in Table 6.8.

The muon rejection algorithm searches for compatible hits in the muon chambers. Due to the very high muon reconstruction efficiency up to $\eta = 2.4$, a rejection efficiency of less than 1% is achieved, while maintaining $\sim 99\%$ of the real $\tau$-jets. The efficiency of the different $\tau$-jet selection criteria applied on $\tau$-jets with $E_T > 15$ GeV matched to a MC muon are summarised in Table 6.8.
6.3 Offline Selection

<table>
<thead>
<tr>
<th>Selection</th>
<th>Relative Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau$-jet</td>
</tr>
<tr>
<td>Fiducial</td>
<td>0.863</td>
</tr>
<tr>
<td>Lepton separation</td>
<td>0.995</td>
</tr>
<tr>
<td>Leading Track</td>
<td>0.861</td>
</tr>
<tr>
<td>1,3 prong</td>
<td>0.89</td>
</tr>
<tr>
<td>PF Had iso</td>
<td>0.807</td>
</tr>
<tr>
<td>PF $\gamma$ iso</td>
<td>0.798</td>
</tr>
<tr>
<td>$e$-rejection</td>
<td>0.85</td>
</tr>
<tr>
<td>$\mu$-rejection</td>
<td>0.99</td>
</tr>
<tr>
<td>Total Efficiency</td>
<td>35.7%</td>
</tr>
</tbody>
</table>

Table 6.8: Table summarising the performance of $\tau$-jet identification and isolation criteria as well as lepton rejections for $\tau$-jets with $E_T > 15$ GeV matched to MC $\tau$-jets, QCD jets, electrons and muons. Efficiencies are relative to the previous selection. The rejection factors of fake $\tau$-jets are of the order of $10^{-3}$ for a $\tau$-jet efficiency of $\sim 35\%$.

6.3.4 Jet Selections

6.3.4.1 VBF Jet Selections

The QCD jets used in this analysis were produced by the “standard” iterative cone algorithm with 0.5 $\eta$-$\phi$ cone size. The transverse energy distributions of the two most energetic reconstructed jets that are not collinear with a MC lepton or $\tau$-jet are shown in Figure 6.14. The jets shown in Figure 6.14 incorporated the energy calibration that was discussed in Section 3.3.2. The jet transverse energy spectrum of both jets for the Z+jets and W+jets processes is very similar. Since there are no pre-selection criteria applied in either of the samples, the peaks of the distributions are dominated by low transverse energy jets. The $t\bar{t}$+jets background process includes two high energy jets coming from the $b$ quarks ($t \rightarrow bW$) which are shown clearly in the right and left plots of Figure 6.14. Therefore, the transverse energy threshold does not provide an effective rejection criterion as seen before at the trigger level (see Table 6.2). The transverse energy thresholds used for the most and the least energetic jet in the event were set at 60 GeV and 27 GeV respectively (dotted vertical lines in the plots). These values ensured a selection efficiency of about 75% for each jet when both jets are reconstructed while providing a large suppression of the background Z+jets and W+jets processes.
6.3 Offline Selection

![Figure 6.14:](image)

The $E_T$ distributions for the most energetic VBF jet (left) and the second most energetic VBF jet (right) for the signal (solid blue line), Z+jets (dashed red line), W+jets (dotted green line) and $t\bar{t}$+jets (dash-dot purple line).

The pseudorapidity gap and the invariant mass distributions of the reconstructed jet pair are shown in Figure 6.15. The distributions for the Z+jets and W+jets are again very similar. The pair of reconstructed jets in the $t\bar{t}$+jets background process tend to have a small pseudorapidity gap as seen in the left plot of Figure 6.15. The invariant mass of the jet pair is a very important criterion to reduce the large background, which is why a high threshold value was chosen (right plot of Figure 6.15). The values for the pseudorapidity gap and the invariant mass thresholds were 4.2 and 900 GeV/c$^2$ respectively.

It is important, especially for the background processes, that the reconstructed lepton and $\tau$-jet are not tagged as one of the VBF jets. Therefore any jet within an $\eta$-/$\varphi$ cone of 0.5 from a reconstructed lepton or $\tau$-jet was not considered.

6.3.4.2 Central Jet Veto

To further reduce the background processes, a requirement on no additional jets in the central region of the detector was placed. The Central Jet Veto (CJV) rejection criterion based on calorimeter jets suffers from electronic noise (fake jets) and pile-up effects. To reduce these effects, information from the charged tracker and the vertex detectors was used. The CJV jet was required to have at least one reconstructed
track associated with it, while the highest transverse momentum associated track should be coming from the interaction vertex ($\Delta z < 0.2$ cm). Finally, the additional jet was required to have transverse energy of $E_{T}^{CJV} > 20$ GeV and to lie within the VBF selected jets satisfying $\eta_{j}^{\min} + 0.5 < \eta_{CJV} < \eta_{j}^{\max} - 0.5$.

### 6.3.5 Missing Transverse Energy

The missing transverse energy ($E_{T}^{\text{miss}}$ or MET) is a detector effect caused by neutral particles (mainly neutrinos) depositing a very small fraction of their energy before exiting the detector. Since the colliding protons have no momentum component in the transverse ($x$-$y$) plane, the vector summation of all the particles’ transverse momentum components should balance out. In practice, the calorimeters provide an energy measurement - equivalent to a momentum measurement in the high momentum limit ($p^2 c^2 \gg m^2 c^4$) with $c$ set to unity - for all known particles except neutrinos and muons. Thus, the calorimetric MET vector including the correction for the muons is defined as:

$$
\vec{E}_{T}^{\text{miss}} = -\sum_{i} \vec{E}_{T i}^{\text{calo}} - \sum_{i} \vec{p}_{T i}^{\mu}
$$

(6.1)
where $E_{\text{calo}}^{\text{i}} = E_{\text{calo}}^{\text{i}} \sin \theta_i \cos \phi_i \hat{x} + E_{\text{calo}}^{\text{i}} \sin \theta_i \sin \phi_i \hat{y}$ of the $i^{th}$ calorimetric object and $\vec{p}_{T_i}$ is the transverse momentum of the $i^{th}$ reconstructed muon. The calorimetric objects could be either individual calorimeter towers or clustered into jets. The true MET calculated from the vector sum of the MC neutrinos for the signal and the $Z+\text{jets}$ with $Z \rightarrow \tau \tau$ background process is shown in Figure 6.16(a). More than 60% of the signal events have $E_T^{\text{miss}} < 50$ GeV, while the $Z+\text{jets}$ with $Z \rightarrow \tau \tau$ background process has significantly less MET, mainly due to the back-to-back topology of the $Z$ decay products (see Section 6.3.7). The difference of the true and the reconstructed MET divided by the true MET is shown in Figure 6.16(b). It can be seen that the MET calculated using the corrected calorimeter tower energy is on an average close to the true MC MET. However it gets overestimated for small values of true MET and will be discussed later in Section 6.3.8. The MET using uncorrected calorimeter towers corresponds to about 50% of the true MET, but it was preferred for this study as both the spread and positive tail of the distribution seen in Figure 6.16(b) were smaller. The performance of the reconstructed MET measurement provided by the calorimeters and the muon detectors makes it hard to apply a threshold criterion on the calculated transverse energy. However an improvement could be seen in the future by applying particle flow techniques. For this study, a combination of the reconstructed MET with the reconstructed lepton was used as a selection criterion.

### 6.3.6 MET-Lepton Selection

The leptons in the $W+\text{jets}$ and the $t\bar{t}+\text{jets}$ samples usually originate from the leptonic decay of the $W^\pm$ bosons. This implies that the invariant mass of the lepton with the neutrino should be equal to the $W^\pm$ boson mass, while for the signal it is expected to drop exponentially after $m_\tau \sim 1.8 \text{ GeV}/c^2$. Since the MET is used, the calculation of the invariant mass in the transverse plane is preferred. Figure 6.17 shows the invariant mass in the transverse plane of the reconstructed lepton and MET for signal and background events that have a reconstructed lepton and $\tau$-jet. As expected, the samples with a $W^\pm$ boson form a peak lower than the $W^\pm$ boson’s mass, accounting for the missing longitudinal component and the uncorrected reconstructed MET. For the $W+\text{jets}$ background sample, the signature is more prominent.
6.3 Offline Selection

![Figure 6.16: True MET seen in (a) and reconstructed MET performance in (b)](image)

(a) True MET calculated by the sum of the neutrino momenta for the VBF Higgs → ττ process divided by the true MET for reconstructed MET using uncorrected (solid blue line) and corrected (dashed red line) processes.

(b) The difference between the true and reconstructed MET divided by the true MET for reconstructed MET using uncorrected (solid green line) and corrected (dashed purple line) calorimeter tower energies.

since only one neutrino is expected in the event, while that for the t¯t+jets distribution is more complex due to the presence of two W± bosons in the event. The threshold value used for this study was $m_T(l,E_{T}^{\text{miss}}) < 40$ GeV/c$^2$.

### 6.3.7 Additional Selection Criteria

Some additional selections were performed on the reconstructed objects. To reduce the Z→ ee and Z→ µµ backgrounds, the reconstructed leptons in the event were counted. Only events with one reconstructed lepton passing the identification criteria in the event were selected. This selection criterion was more effective for the muon decays due to the high muon reconstruction efficiency. For the electron decays an additional selection criterion was applied. To reject events where the second electron was not identified either due to a bad track momentum measurement or rejected for entering the calorimeter in an ECAL gap, the collection of the super-clusters was used. The invariant mass of the well reconstructed electron and each super-cluster in the event was calculated and events that had at least one super-cluster satisfying $70$ GeV/c$^2 < m(e,SC) < 110$ GeV/c$^2$ were rejected.
6.3 Offline Selection

![Figure 6.17](image)

**Figure 6.17:** Transverse invariant mass of reconstructed lepton and MET for the signal (solid blue line), Z+jets (dashed red line), W+jets (dotted green line) and $t\bar{t}$+jets (dash-dot purple line).

In the signal events, the lepton and the $\tau$-jet are required to be oppositely charged. This requirement holds for the Z+jets background process but does not hold for the W+jets and $t\bar{t}$+jets background processes. Therefore, only events where the reconstructed lepton and the reconstructed $\tau$-jet pair were oppositely charged were selected (OS$_{\tau,l}$). The charge of the lepton was determined by its track, while for the $\tau$-jet the sum of the track charges inside the isolation cone was used.

Finally to further reduce events from the Z+jets background process, a selection in the transverse opening angle between the reconstructed lepton and the $\tau$-jet was introduced. Due to the Z boson’s spin its decay products are expected to be back-to-back in the transverse plane. The transverse opening angle between the reconstructed lepton and $\tau$-jet ($\Delta\phi_{\tau,l}$) for the signal and the background processes is shown in Figure 6.18. As expected, the Z+jets events are concentrated close to $\Delta\phi_{\tau,l} \sim \pi$. The selection threshold was placed at $\Delta\phi_{\tau,l} < 2.7$ rad which preserves more than 75% of the signal events while rejecting $\sim 80\%$ of the Z+jets background.
Figure 6.18: Transverse opening angle between the reconstructed lepton and the \( \tau \)-jet for the signal (solid blue line), Z+jets (dashed red line), W+jets (dotted green line) and \( t\bar{t} \)+jets (dash-dot purple line). The lepton and \( \tau \)-jet coming from the Z decay tend to be back-to-back.
6.3.8 Mass Reconstruction

The mass reconstruction of the Higgs boson for the H→ l + τ-jet from its decay products can be done in different ways. For the Z→ ττ→ e + τ-jet channel the Z boson mass is reconstructed from the invariant mass of its visible decay products \cite{72}. Figure 6.19(a) shows the invariant mass of the MC lepton and the MC τ-jet for the Z→ l + τ-jet and H→ l + τ-jet processes. This method has the advantage of being independent of the MET measurement, but due to the large amount of missing transverse energy in the signal case (see Figure 6.16(a)) the resolution is degraded, making the two mass resonances very hard to separate. An alternative method is to include the MET in the calculation. The outgoing neutrinos from the τ lepton decay tend to be collinear to the visible decay products due to the Lorentz boost. To use this information, the MET is projected into the directions of the lepton and the τ-jet and the missing energy in the lepton and τ-jet direction can be expressed as:

\[ E_{\text{miss}}^{\tau_j} = \frac{E_{\text{miss}}^y x_{\tau_j} - E_{\text{miss}}^x y_{\tau_j}}{x_{\tau_j} y_{\tau_l} - x_{\tau_l} y_{\tau_j}} \] (6.2)

\[ E_{\text{miss}}^{\tau_l} = \frac{E_{\text{miss}}^x - E_{\text{miss}}^l x_{\tau_l}}{y_{\tau_j}} \] (6.3)

where \( x_{\tau_j,\tau_l} \) and \( y_{\tau_j,\tau_l} \) are the unit vectors in the lepton and τ-jet direction and are defined as:

\[ x_{\tau_j,\tau_l} = \sin(\theta_{\tau_j,\tau_l}) \cos(\phi_{\tau_j,\tau_l}) \] (6.4)

\[ y_{\tau_j,\tau_l} = \sin(\theta_{\tau_j,\tau_l}) \sin(\phi_{\tau_j,\tau_l}) \] (6.5)

This method is called the collinear approximation and it only provides a valid estimation when the lepton and the τ-jet are not back-to-back. The invariant mass of the reconstructed lepton, τ-jet and MC MET for \( \Delta \phi_{\tau,l} < 3\) rad is shown in Figure 6.19(b).
6.3 Offline Selection

(a) The calculated mass of the Higgs and Z bosons decaying to $\tau\tau$ channel, using MC par- collinear approximation using the recon- structed lepton, $\tau$-jet and MC MET (solid blue line) and MC lepton, MC $\tau$-jet and MC MET (black dots). Spread and peak are in agree- ment.

Figure 6.19: Distributions of Higgs and Z bosons masses calculated using the lepton and $\tau$-jet pair (left) and the collinear approximation for the neutrino (right).

It can be seen that the collinear approximation is working well, restoring the calculated mass to $m_H \sim 135 \text{GeV/c}^2$ at which it was generated. Also in Figure 6.19(b), the collinear approximation is used with the MC true lepton and $\tau$-jet (black dots). Both the peak and the spread of the distributions are very similar, demonstrating that the reconstructed lepton and $\tau$-jet of the signal events are well reconstructed and the degradation in the mass reconstruction presented later originates from the reconstructed MET.

The invariant mass of the Z and the Higgs bosons using only offline reconstructed objects is shown in Figures 6.20(a) and 6.20(b) respectively. It can be seen that despite using the corrected MET, which brings the average calculated boson mass close to the generated mass, the spread also increases substantially. Also the positive tail of the distributions is more prominent when using the corrected MET, which is due to overestimation of the MET especially for low values of missing energy. The difference in performance of the collinear approximation is also clear between the Higgs and the Z boson. This is again due to the higher amounts of missing energy.
Figure 6.20: Z (a) and H (b) offline mass reconstruction. The large positive tail signifies overestimated MET.

in the signal process, resulting in better measurement resolution.

It is worth noting that bad MET reconstruction can result in negative neutrino energies given by Equations 6.2 and 6.3, causing the collinear approximation to fail. Only events that have positive neutrino energies for both the lepton and the \( \tau \)-jet components were selected. A selection could also be placed in a mass window around the expected Higgs mass to reduce the background processes further. However this would require a better MET resolution possibly obtained by using particle-flow techniques.

### 6.4 Results

#### 6.4.1 Selection Efficiency

The selection criteria described in Sections 6.3.1–6.3.8 were applied to the signal and background samples. The large difference in the cross section of signal and background events (\( \sigma_{W^++jets}/\sigma_{signal} \sim 10^6 \)) implies that it would have required \( 10^{11} \)
W+jets events to provide similar statistical uncertainty for each selection criterion. To solve the problem of limited statistics, the selection efficiency of some criteria was calculated using an uncorrelated sample of relaxed selection criteria. The cumulative efficiency of a criterion was then calculated as:

\[ \epsilon_{\text{criterion}_n} = \frac{N_{\text{criterion}_n}}{N_{\text{relaxed criterion}_{n-1}}} \epsilon_{\text{criterion}_{n-1}} \]  

(6.6)

where \( N_{\text{criterion}_n} \) is the number of events satisfying this criterion and \( N_{\text{relaxed criterion}_{n-1}} \) is the number of events satisfying the previous uncorrelated relaxed criterion.

In practice the \( \tau \)-jet selection criteria were applied in a sample where the isolation criteria of the leptons were relaxed. Similarly the VBF jet and CJV selections were applied in a sample where both the lepton and the \( \tau \)-jet isolation criteria were relaxed, while all the efficiencies of the selection criteria described in Section 6.3.5–6.3.8 were obtained by a sample that had a reconstructed lepton and \( \tau \)-jet. In all cases the events should have first satisfied the corresponding trigger path.

The only sample that had sufficient events passing all the selection criteria using both methods was the VBF Higgs sample and therefore it was used to test the validity of this method. The efficiencies of the individual selections were found to be less than 5% different, while the total efficiencies were less than 10% different. This was considered to be acceptable for the purposes of this study.

The cumulative cross sections expected at the LHC for the signal and the background processes satisfying the VBF jets plus lepton trigger path at \( \sqrt{s} = 10 \) TeV for each selection step are summarised in Table 6.9. As expected, the lepton reconstruction efficiency is fairly high since most of the criteria were satisfied at the HLT level. The \( \tau \)-jet kinematic criteria can be easily met since the reconstructed object at this level could be either the reconstructed lepton or one of the VBF jets. The \( \tau \)-jet isolation criteria reject a large fraction of the W+jets sample since no event with a lepton and a real \( \tau \)-jet is expected. The resulting signal to signal plus background ratio is \( S/(S+B) = 0.75 \) and 0.63 for \( m_H = 115 \) GeV/c\(^2 \) and \( m_H = 135 \) GeV/c\(^2 \) respectively.
Table 6.9: Cumulative cross sections of the signal and background events, satisfying the VBF jets plus lepton trigger, for each step of the offline selection criteria.
### 6.4 Results

<table>
<thead>
<tr>
<th>Selection</th>
<th>Cumulative Cross Section [fb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H \rightarrow \tau\tau$</td>
<td></td>
</tr>
<tr>
<td>$(m_H = 115 \text{ GeV}/c^2)$</td>
<td></td>
</tr>
<tr>
<td>$H \rightarrow l + \tau$-jet $(m_H = 135 \text{ GeV}/c^2)$</td>
<td></td>
</tr>
<tr>
<td>$Z$+jets $(Z \rightarrow ll)$</td>
<td></td>
</tr>
<tr>
<td>$W$+jets $(W \rightarrow l\bar{\nu}_l)$</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$+jets $(t \rightarrow bW)$</td>
<td></td>
</tr>
<tr>
<td><strong>Production</strong> $\sigma$</td>
<td>196</td>
</tr>
<tr>
<td><strong>Single Lepton L1+HLT</strong></td>
<td>39.2</td>
</tr>
<tr>
<td><strong>Lepton Kin</strong></td>
<td>38.8</td>
</tr>
<tr>
<td><strong>Lepton ID</strong></td>
<td>33.7</td>
</tr>
<tr>
<td><strong>Lepton Iso</strong></td>
<td>27</td>
</tr>
<tr>
<td><strong>$\gamma$-rej</strong></td>
<td>26.1</td>
</tr>
<tr>
<td><strong>Lepton Counting</strong></td>
<td>22.3</td>
</tr>
<tr>
<td><strong>$\tau$ Kin</strong></td>
<td>22.2</td>
</tr>
<tr>
<td><strong>$\tau$ ID</strong></td>
<td>15.1</td>
</tr>
<tr>
<td><strong>$\tau$ Iso</strong></td>
<td>8.16</td>
</tr>
<tr>
<td><strong>$\tau$ e-rej</strong></td>
<td>6.29</td>
</tr>
<tr>
<td><strong>$\tau$ $\mu$-rej</strong></td>
<td>5.77</td>
</tr>
<tr>
<td><strong>OS$_{\tau,l}$</strong></td>
<td>5.26</td>
</tr>
<tr>
<td><strong>$\Delta\phi_{\tau,l}$</strong></td>
<td>4.27</td>
</tr>
<tr>
<td><strong>VBF Jet $E_T$</strong></td>
<td>3.55</td>
</tr>
<tr>
<td><strong>$\Delta\eta_{j_1,j_2}$</strong></td>
<td>1.15</td>
</tr>
<tr>
<td><strong>$m(j_1,j_2)$</strong></td>
<td>0.62</td>
</tr>
<tr>
<td><strong>Central Jet Veto</strong></td>
<td>0.59</td>
</tr>
<tr>
<td><strong>Valid Mass $(E_\nu &gt; 0)$</strong></td>
<td>0.328</td>
</tr>
<tr>
<td><strong>$m(e,SC)$</strong></td>
<td>0.328</td>
</tr>
<tr>
<td><strong>$m_T(t,E_T^{\text{miss}})$</strong></td>
<td>0.269</td>
</tr>
<tr>
<td><strong>Total Efficiency</strong></td>
<td>0.00137</td>
</tr>
</tbody>
</table>

Table 6.10: Cumulative cross sections of the signal and background events, satisfying the single isolated lepton triggers, for each step of the offline selection criteria.
## 6.4 Results

<table>
<thead>
<tr>
<th>Selection</th>
<th>Cumulative Cross Section [fb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H \rightarrow \tau\tau ) (( m_H = 115 \text{ GeV}/c^2 ))</td>
<td>196</td>
</tr>
<tr>
<td>( H \rightarrow l + \tau^- ) jet (( m_H = 135 \text{ GeV}/c^2 ))</td>
<td>45.9</td>
</tr>
<tr>
<td>( Z + \text{jets} ) (( Z \rightarrow ll ))</td>
<td>3.7 \times 10^6</td>
</tr>
<tr>
<td>( W + \text{jets} ) (( W \rightarrow l\bar{\nu}_l ))</td>
<td>4 \times 10^7</td>
</tr>
<tr>
<td>( t\bar{t} + \text{jets} ) (( t \rightarrow bW ))</td>
<td>3.17 \times 10^5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>production ( \sigma )</th>
<th>Combined L1+HLT ( \sigma )</th>
<th>Lepton Kin ( \sigma )</th>
<th>Lepton ID ( \sigma )</th>
<th>Lepton Iso ( \gamma)-rej ( \sigma )</th>
<th>Lepton Counting ( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>44.2</td>
<td>16.8</td>
<td>15.5 \times 10^6</td>
<td>10.4 \times 10^6</td>
<td>1.1 \times 10^6</td>
<td>3.8 \times 10^4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \tau ) Kin ( \sigma )</th>
<th>( \tau ) ID ( \sigma )</th>
<th>( \tau ) Iso ( \sigma )</th>
<th>( \tau ) e-rej ( \sigma )</th>
<th>( \tau ) ( \mu )-rej ( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>10.9</td>
<td>4.81 \times 10^5</td>
<td>6.77 \times 10^6</td>
<td>2.86 \times 10^4</td>
</tr>
<tr>
<td>17.1</td>
<td>8.01</td>
<td>3.46 \times 10^5</td>
<td>1.23 \times 10^6</td>
<td>2.37 \times 10^4</td>
</tr>
<tr>
<td>9.27</td>
<td>4.46</td>
<td>1.87 \times 10^5</td>
<td>3.68 \times 10^4</td>
<td>2.01 \times 10^3</td>
</tr>
<tr>
<td>7.2</td>
<td>3.78</td>
<td>1.1 \times 10^5</td>
<td>3.36 \times 10^4</td>
<td>1.18 \times 10^3</td>
</tr>
<tr>
<td>6.64</td>
<td>3.72</td>
<td>3.65 \times 10^3</td>
<td>9.02 \times 10^3</td>
<td>426</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OS(( \tau , l )) ( \sigma )</th>
<th>( \Delta \phi_{\tau , l} ) ( \sigma )</th>
<th>VBF Jet ( E_T ) ( \sigma )</th>
<th>( \Delta \eta_{j_1, j_2} ) ( \sigma )</th>
<th>( m(j_1, j_2) ) ( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.08</td>
<td>3.49</td>
<td>3.43 \times 10^3</td>
<td>6.99 \times 10^4</td>
<td>374</td>
</tr>
<tr>
<td>4.98</td>
<td>2.74</td>
<td>652</td>
<td>5.66 \times 10^3</td>
<td>310</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Central Jet Veto ( \sigma )</th>
<th>Valid Mass ( (E_\nu &gt; 0) ) ( \sigma )</th>
<th>m(( e, \text{SC} )) ( \sigma )</th>
<th>m_T(( l, E_T^{\text{miss}} )) ( \sigma )</th>
<th>Total Efficiency ( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.836</td>
<td>0.462</td>
<td>0.462</td>
<td>0.391</td>
<td>0.002</td>
</tr>
<tr>
<td>0.502</td>
<td>0.265</td>
<td>0.265</td>
<td>0.216</td>
<td>0.0471</td>
</tr>
<tr>
<td>0.513</td>
<td>0.0616</td>
<td>0.026</td>
<td>0.0138</td>
<td>3.72 \times 10^{-9}</td>
</tr>
<tr>
<td>3.46</td>
<td>0.346</td>
<td>0.346</td>
<td>0.0744</td>
<td>1.86 \times 10^{-9}</td>
</tr>
<tr>
<td>1.15</td>
<td>0.184</td>
<td>0.133</td>
<td>0.036</td>
<td>1.14 \times 10^{-7}</td>
</tr>
</tbody>
</table>

Table 6.11: Cumulative cross sections of the signal and background events, satisfying the combination of the VBF jets plus lepton and single isolated lepton triggers, for each step of the offline selection criteria.
6.4 Results

Figure 6.21: Graphical comparison of the cross section reductions due to the different selection criteria for the signal and background processes using the VBF jets plus lepton trigger (solid blue line) and the single lepton triggers (dashed red line). The cross sections of the combined trigger follow the trend on the single lepton triggers at the beginning, while at the end they follow that of the VBF jets plus lepton trigger.

The cumulative cross sections expected at the LHC for the signal and the background processes satisfying the single isolated lepton trigger paths at $\sqrt{s} = 10$ TeV for each selection step are summarised in Table 6.10. As expected, the main background reduction is achieved by the VBF jets selection criteria since no selection was performed at the trigger level. The rest of the selection criteria function in a similar manner as in the VBF jets plus lepton trigger case. The final efficiency is lower for all the samples resulting in $S/(S+B) = 0.74$ and 0.56 for $m_H = 115$ GeV/c$^2$ and $m_H = 135$ GeV/c$^2$ respectively. For the $m_H = 135$ GeV/c$^2$ case the signal increases by $\sim 40\%$ while the background increases only by $\sim 8\%$. The cross section reductions due to the different selection criteria for the VBF jets plus lepton trigger and the single lepton triggers can be seen in Figure 6.21. As expected the cross sections...
remain higher for the single lepton trigger paths until the VBF jet selection criteria are applied. After that point the VBF jets plus lepton trigger path becomes more efficient.

The corresponding cumulative cross sections for events satisfying the combined VBF jets plus lepton OR single isolated lepton trigger are summarised in Table 6.11. The selection efficiencies for the signal and background processes are higher than the previous two cases, resulting in \( S/(S+B) = 0.76 \) and \( 0.64 \) for \( m_H = 115 \text{ GeV}/c^2 \) and \( m_H = 135 \text{ GeV}/c^2 \) respectively. The gain in signal (\( m_H = 135 \text{ GeV}/c^2 \)) from the use of the combined trigger over the single isolated lepton trigger is \( \sim 47\% \) while the background increases by \( \sim 9\% \).

### 6.4.2 Reconstructed Mass

The mass of the Higgs boson was calculated using the method described in Section 6.3.8. The VBF jet selections are not expected to influence the kinematic distributions of the lepton or the \( \tau \)-jet, while the dependence on the MET was considered acceptable to the purpose of this section. Therefore, the calculated mass distributions for the background processes were extracted before applying the VBF jet selections. Due to their similar shape the W+jets and the \( t\bar{t} \)+jets distributions were combined. Figure 6.22 shows the expected mass shape of the signal (\( m_H = 135 \text{ GeV}/c^2 \)) and the background processes at the LHC using the combined trigger (VBF jets plus lepton OR single isolated lepton) and the offline selection criteria presented in Section 6.4.1. As mentioned before, the uncorrected MET was used in the calculation resulting in the Higgs invariant mass being lower than generated. Since the purpose of this study was to evaluate the possible gains in selection efficiency by using the VBF jet plus lepton trigger, no further analysis regarding uncertainties was performed.

### 6.5 Summary and Conclusions

In this chapter the reconstruction and selection of the VBF \( H \rightarrow \tau\tau \rightarrow l + \tau\)-jet process was presented from the trigger level to offline. A set of offline selection criteria
was implemented in order to suppress the background processes, while maintaining a reasonable signal efficiency. Although this study was not focused on presenting a Higgs discovery technique, it was considered to be essential that the set of selection criteria used should provide similar background rejection power. Therefore, in addition to the selection criteria aimed at ensuring purity that were used in the previous chapter (Chapter 4), a set of selections aimed at increasing the signal to background ratio was developed using topological and kinematic signatures that favour the signal over the background events. The final suppression factors achieved for the Z+jet and W+jet background processes were of the order of $10^9$, while that for $t\bar{t}$+jets process was of the order of $10^7$. The corresponding signal efficiency was of the order of $10^{-3}$.

The signal over signal plus background ratio of VBF Higgs boson generated with $m_H = 135 \text{ GeV}/c^2$ was improved by the use of the VBF jets plus lepton trigger path from $S/(S+B)=0.56$ for the single isolated lepton case to $S/(S+B)=0.64$ for

![Figure 6.22: Higgs boson reconstructed mass using the collinear approximation described in Section 6.3.8. The distributions were extracted before applying the VBF jet selection criteria and then scaled to the cross section expected by the combination of the VBF jets plus lepton and the single isolated lepton triggers.](image-url)
the combined trigger. More importantly, the signal efficiency increased by \( \sim 47\% \) by using the combined trigger over the single lepton trigger with an increase in background of only \( \sim 9\% \). This is a very substantial improvement and can prove crucial in the discovery of the Higgs boson.
Chapter 7

Conclusions

Detecting the Higgs boson is one of the primary goals of the CMS experiment. Before any discovery can be made, a good understanding of the detector and its simulation will be required. In a hadronic environment the triggering system is essential. Understanding and optimising the trigger helps collecting quickly interesting data that can lead to the confirmation or rejection of particle physics models.

The work presented in this thesis contributes to the understanding of the CMS detector and its simulation as well as proposing a new trigger which can assist in the discovery of a low mass Standard Model Higgs boson.

Single Pion responses

The response of the CMS calorimeter to single pions is particularly important for understanding the jet response, especially for \( \tau \)-jets which have a high probability of decaying into a single pion (and a neutrino). The results presented in Chapter 3 demonstrate the performance of the ECAL and the HCAL calorimeters for incident pions with a momentum range \( 2 \, \text{GeV/c} \leq p_{\pi} \leq 300 \, \text{GeV/c} \). Since the CMS calorimeter is non-compensating, the calibrated electromagnetic calorimeter registered a smaller amount of energy than that deposited by the pion. The calibrated HCAL was found to correctly measure pions with \( p_T = 50 \, \text{GeV/c} \), however it overestimated the response of the more energetic pions.
A detector simulation using the official CMS simulation and reconstruction software adapted for the test beam conditions was developed. Data samples corresponding to the test beam data were produced for comparative analysis. The comparison between the test beam and the simulated data showed that the response of the ECAL to the pions in the simulation is even smaller than that observed in the test beam data. Also it was found that the HCAL resolution in the simulated data was significantly better than that in the test beam data - \(~20\%\) for \(p_\pi = 50\text{ GeV/c}\) and increasing to \(~40\%\) for \(p_\pi = 100\text{ GeV/c}\). It was concluded that although the average responses from the test beam and simulated data were in good agreement, the simulation needed to be improved to provide more realistic pion responses.

**VBF jets plus lepton Trigger**

In Chapter 5, the motivation and implementation of the VBF jets plus lepton trigger are discussed. A Higgs boson decaying into two taus is a very promising channel for the detection of a low mass SM Higgs boson \((m_H < 150\text{ GeV/c}^2)\). The VBF jets plus lepton trigger is motivated by Higgs boson discovery with the VBF production mechanism. The idea was to use the characteristic signature of the VBF channel, two jets separated in \(\eta\), to reduce the lepton \(E_T\) threshold at L1 and the in the HLT. The single isolated electron and the single isolated muon HLT trigger \(E_T\) thresholds, at instantaneous luminosity of \(2 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}\), are 26 GeV and 19 GeV, respectively [71]. The suggested trigger was implemented and optimised for the \(2 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}\) instantaneous luminosity regime. It was found that the VBF jets plus lepton trigger could improve the selection efficiency of offline preselected Higgs events over current triggers by 15\% for a QCD rate of \((0.6 \pm 0.3)\text{ Hz}\). If the VBF jets plus lepton trigger was used in conjunction with the current triggers, the selection efficiency was improved by \(~33\%\) over the current triggers.

**Signal and Background reconstruction and selection performance using VBF jets plus lepton Trigger**

The production of the weak vector bosons \(Z^0\) and \(W^\pm\) as well as the \(t\bar{t}\) production are the dominant background processes in this search channel. In Chapter 5, the VBF
jets plus lepton trigger was used as a filter on the background events. It was found to provide a small additional efficiency compared to the current single isolated lepton triggers. Offline selection criteria used in previous studies [69, 70, 72] and additional ones were implemented and optimised for this analysis. The final efficiency for the Z+jets and W+jets background processes was of the order of $10^{-9}$ while that for the $t\bar{t} + jets$ process was of the order of $10^{-7}$. The signal efficiency for a Higgs boson of mass $m_H = 135 \text{ GeV}/c^2$ was of the order of $10^{-3}$.

The signal over signal plus background ratio for a VBF Higgs boson generated with $m_H = 135 \text{ GeV}/c^2$ was improved by the use of the VBF jets plus lepton trigger path from $S/(S+B)=0.56$ for the single isolated lepton case to $S/(S+B)=0.64$ for the combined trigger. More significantly, the use of the VBF jets plus lepton trigger improved the signal selection efficiency by $\sim 47\%$. This can prove to be an important factor for an early Higgs boson discovery.
References


[22] CDF and D0 Collaboration, “Combined CDF and D0 Upper Limits on Standard Model Higgs- Boson Production with up to 6.7 fb$^{-1}$ of Data,” 1007.4587.

[23] “SLD Experement.”


[27] “The LEP Electroweak Working Group.”
http://lepewwg.web.cern.ch/LEPEWWG/


[29] “European Organisation of Nuclear Research.”
http://public.web.cern.ch/public/


[34] “CERN Hadron LINACS Project.”
http://linac2.home.cern.ch/linac2/default.htm

   \url{http://ab-dep-op-leir.web.cern.ch/ab-dep-op-leir/}

[37] “The Proton Synchrotron Project.”
   \url{http://ab-dep-op-ps.web.cern.ch/ab-dep-op-ps/}


[58] CMS Collaboration, “Track Reconstruction in the CMS tracker,”.


[69] C. Foudas, A. Nikitenko, and M. Takahashi, “Observation of the standard model higgs boson via $h \rightarrow \tau \tau \rightarrow$lepton+jet channel.”. CMS NOTE 2006/088.

[70] M. Vazquez Acosta et al., “Search for the standard model higgs boson produced in vector boson fusion and decaying into $\tau \tau$ pair in cms with 1 fb$^{-1}$,”. CMS NOTE 2007/035.


[75] J. Alwall et al., “MadGraph/MadEvent v4: The New Web Generation,” 