UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA
Facoltà di scienze MM. FF. NN.

Study of electron fake rate from QCD jets in CMS ECAL calorimeter

Cristian Consonni
Università degli studi di Milano-Bicocca
mat. 075356
a.a. 2007/2008

Relatore: Prof. Paganoni Marco
Correlatore: Dott. Salerno Roberto

Contents

Chapter 1

Introduction

The Large Hadron Collider (LHC) is a particle accelerator designed to collide opposing beams of protons or heavy ions. It has been built to verify various predictions of the Standard Model (SM) of high-energy physics, to test other theories like supersimmetry (SUSY) and to discover new physical phenomena.

One of the open questions of the modern physics is the origin of the masses of the particles. In the SM the solution to this problem is a Higgs mechanism leading to an additional scalar field whose quantum, the Higgs boson, is a new particle to be introduced. Up to date, no successful experimental observation of this particle has been performed.

Thanks to the 7 TeV energy achieved by each of the proton beams and to a nominal luminosity of $10^{34}\ cm^{-2}\ s^{-1}$ at full performance, new experimental studies about the Higgs mechanism and the observation of the Higgs boson become possible, providing the conditions for a confirmation of the theory. The search for the Higgs boson is therefore considered to be one of the most important task of the LHC and a goal for its experiments.

In this thesis, after a brief introduction about the LHC given in chapter 2 and in particular about the CMS experiment detector (sec. 2.1), a background process which affects one of the numerous decay channel of the Higgs boson is studied.

CMS is a general-purpose particle physics detector of the LHC. The large variety of decays of the Higgs depending on its mass requires well-signed and low-background decay channels for its identification. Dominant events present a final hadronic state with a strong QCD background, so leptonic and isolated photons final states are the clearest way to identify a possible Higgs event. The search of Higgs decays in leptonic states then requires for CMS an excellent tracking system, an electromagnetic calorimeter with optimal energy resolution and a good muon tracking system.

The decay channel $H \rightarrow WW^{(*)} \rightarrow e\nu e\nu$ is a well signed candidate for Higgs boson identification requiring two reconstructed electrons to be identified.
Signals satisfying definite conditions (e.g. a reconstructed track in the tracker matching an energy deposit in the ECAL) are identified as electrons, nevertheless other events may give an electron signal leading to a contamination coming, for example, from particle jets.
In this work the W+jets background is considered with the supplementary condition that at least one jet is misidentified as an electron. In this case, two

electrons are present in the event: one real, prompt electron from the W boson decay, generally isolated and a non-prompt electron following the misreconstruction of other particles. The latest, called a *fake* electron, constitutes a background for the Higgs search.

The CMS subdetector which measure the energy of electrons is the electromagnetic calorimeter ECAL and hence is the detector that interest our analysis, ECAL is described in detail in section 2.1.1.

As explained in section 3.1, a direct measure of the fake rate for the W+jets would be hard to compute thus an estimate of the probability is given using QCD jets and extrapolating the fake rate for the W+jets background.
The fake rate is determined as a function of jet transverse momentum p_T and in section 3.1.2 the effect of the application of different electron selection procedures on the fake electrons is studied.
In section 3.2 the definition of the integrated fake rate is given and the selection-based analysis is performed.

Finally, the estimated value of the fake rate for W+jets is extrapolated from the computed value.

Chapter 2

The Large Hadron Collider

The LHC is a circular proton-proton collider with separate magnetic fields and vacuum chambers in the main arcs and with common sections only at the insertion regions where the experimental detectors are located.

The LHC is designed to reach a center of mass energy of 14 TeV and a luminosity of 10^{34} cm^{-2} s^{-1} [1]. The collider, located at CERN (European Council for Nuclear Research), is contained in a circular tunnel with a circumference of 27 kilometers, at a depth ranging from 50 to 175 meters underground.

As depicted in figure 2.1 there are four main experiments[1] at LHC:

- CMS (Compact Muon Solenoid) a general purpose detector, section 2.1.
- ATLAS (A Toroidal LHC ApparatuS) is designed as a general-purpose detector. It is 44 meters long and 25 meters in diameter, weighing about 7,000 tonnes.
- LHCb (Large Hadron Collider beauty) a specialist b-physics experiment, aimed at measuring the parameters of CP violation in the interactions of b-hadrons, heavy particles containing a bottom (or beauty) quark.
- ALICE (A Large Ion Collider Experiment) is optimized to study heavy ${}^{82}_{208}\mathrm{Pb}$ ion collisions. The resulting temperature and energy density are expected to be large enough to generate a quark-gluon plasma, i.e. a state of the matter with deconfined quarks and gluons.

The LHC operating temperature is below 2 K to exploit superconducting magnets with NbTi windings and produce the magnetic field required to bend the proton beam.

The particles to be injected in the LHC are prepared in bunches formed in the Proton Synchrotron (PS) with an energy of 26 GeV and a nominal bunch

[1] There are also two minor experiments:

- TOTEM (Total Cross Section, Elastic Scattering and Diffraction Dissociation) aims at the measurement of the total cross section, the elastic scattering and the diffractive processes. It shares intersection point IP5 with the CMS experiment.
- LHCf (Large Hadron Collider forward) is a special-purpose LHC experiment for astroparticle physics (cosmic rays). It is intended to measure the energy and the number of neutral pions (π^0) produced by the collider. It shares IP1 with the ATLAS detector.

spacing of 25 ns. Every bunch is composed of 1.15×10^{11} protons with a nominal length of 1.6 ns. The particles are subsequently accelerated at 450 GeV/c in the Super Proton Synchrotron (SPS) and then injected in the LHC in a large number: 2,808 bunches for each proton beam[2]. Each of the 1,232 dipoles in radio frequency cavities gives the beam an impulse of 0.5 $MeV/cycle$. The LHC collision rate is 40 MHz with an estimate of ~ 20 collisions per crossing.

The LHC can also collide heavy ion beams with an energy of 2.76 TeV per nucleon, yielding a total center-of-mass energy of 1.15 PeV and a nominal design luminosity of 1.0×10^{27} cm^{-2} s^{-1}.

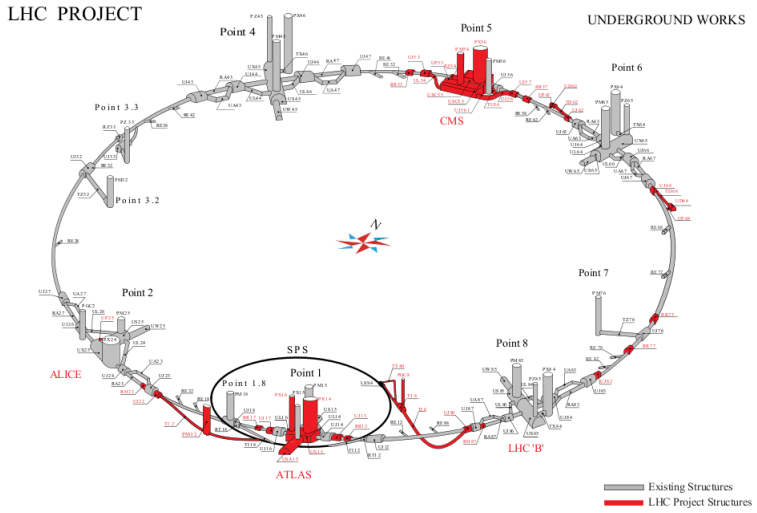


Figure 2.1: The LHC ring with experiments.

2.1 The CMS detector

CMS, which is the acronym of Compact Muon Solenoid, is a large general-purpose particle physics detector built at the interaction point number 5 of the LHC[3].

The physical standpoint for CMS design has been the search for the Higgs boson [2]. The large variety of decays of the Higgs depending on its mass requires well-signed and low-background decay channels for identification in the range of Higgs masses established from previous experiments[4] ($114.4\ GeV/c^2 \leq m_H$

[2]Filling the LHC requires 12 cycles of the SPS synchrotron, and each SPS fill requires 3 to 4 cycles of the PS synchrotron. The SPS and PS cycling times are 21.6 s and 3.6 s, respectively, yielding a total LHC filling time of approximately 4 minutes per beam. The minimum time required for ramping the beam energy in the LHC from 450 GeV/c to 7 TeV is approximately 20 minutes [1].

[3]The location is between Segny and Cessy (France)

[4]The limits follow the results of LEP, the Large Electron Positron collider, and other experiments. The lower limit has been directly established by LEP at a 95% confidence level. An indirect search has been performed in which the consistency between precise measurements of the electroweak sector of the SM and given Higgs masses were tested. Radiative loop corrections must be applied in the current energy scale of the experiments and affect the

$\leq$ 182 GeV/c^2). The branching ratios for the Higgs decay channels depend on m_H, but dominant events present final hadronic states with a strong QCD background, thus leptonic and isolated photons final states are the clearest way to identify possible Higgs events. These events present a non-negligible cross section for m_H values of about 2 m_W ($\sim$ 200 GeV/c).

The main feature of the CMS detector is that it is embedded in a 4 T axial magnetic field generated by a superconducting solenoid[5] 13 m long and with an internal diameter of 5.9 m.

The search of Higgs decays in leptonic states requires for CMS an excellent tracking system, an electromagnetic calorimeter with optimal energy resolution and a good muon tracking system. The structure of CMS is designed to guarantee a good level of hermeticity covering almost 4π. CMS is composed of a cylindrical central section (the barrel) closed with two caps (the endcaps).

The detector is cylindrical, it is 21 metros long, it has a diameter of 16 meters and it weighs approximately 12, 500 tonnes.

The coordinate system in CMS is chosen with the z axis along the beam direction, with the x axis directed towards the LHC ring center and the y axis orthogonal to them. Due to the cylindrical structure of the detector a more convenient coordinate system is (r, Φ, η) where r is the distance from the z axis, Φ is the azimuthal angle in the xy plane measured from the x axis and η is the pseudorapidity from the polar angle θ formed with the z axis defined as:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$$

The CMS is constituted by several subdetectors with different tasks as presented in figure 2.2. Particles passing through the detector from the interaction vertex encounter:

- The **tracker**, a detector which tracks charged particles. It is a cylindrical detector of 5.5 m in length and 1.1 m in radius. The tracker tasks are to determine the interaction point, to measure with good accuracy the momentum of charged particles and to identify the presence of secondary vertexes. It is made of pixel detectors and silicon strip trackers [3]. The tracking region is $|z| < 2.7\ m$ and $|\eta| < 2.5$.

  The pixel detectors provide in general 2 or 3 hits per track, each with a three dimensional precision of about 10 μm in the transverse plane (r, Φ), and 15 μm in z.

- The **electromagnetic calorimeter (ECAL)**: its main aim is to measure the energy of electrons and photons. A more detailed description is presented in section 2.1.1.

- The **Hadron Calorimeter (HCAL)**: a detector to measure the energy of both charged and neutral hadrons. It is a sampling calorimeter made of a cartridge brass plane as absorber and of a plastic scintillator as active element. It is placed in the region 1.8 $m < r < 2.9\ m$ and a forward region extends the coverage up to $|\eta| < 5$.

observed electroweak parameters. These corrections logarithmically depends on the Higgs boson mass. From these tests the upper limit is obtained at a 95% confidence level.

[5]The presence of the solenoid leaded to a compact design of the detector, explaining the origin of its name.

- The **solenoid**: it is a superconducting solenoidal coil. The return yoke for the magnetic field is an iron structure which hosts the muon detector. It is placed in the region 2.9 $m < r < 3.8$ m and $|\eta| < 1.5$.
- The **muon system**: it is the outermost detector. The bending of muons in the magnetic field allows the reconstruction of the momentum and trajectory of muons. It is the most voluminous detector extending in the region 4.0 $m < r < 7.4$ m and $|\eta| < 2.4$.

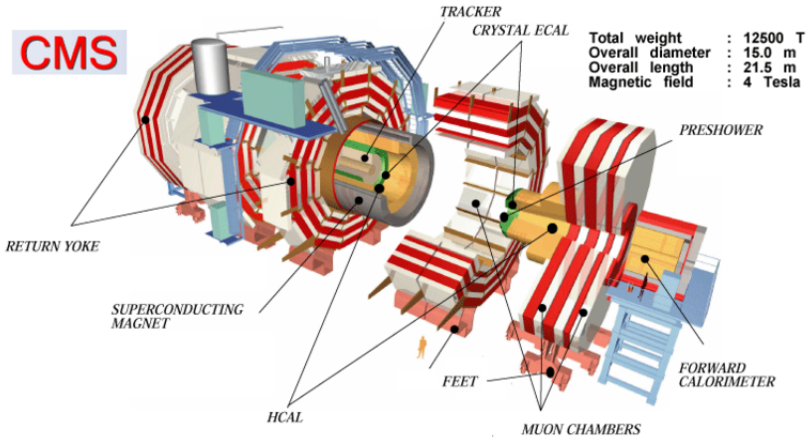


Figure 2.2: The CMS detector and subdetectors.

2.1.1 The Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) measures the energy of electrons and photons. The design of ECAL was driven by the requirements imposed by the search of the Higgs boson in the channel $H \rightarrow \gamma\gamma$, where a peak in the diphoton invariant mass placed at the Higgs mass has to be distinguished from a continuous background [4]. Furthermore a fine granularity is essential to obtain a good π^0/γ separation.

In order not to deteriorate the energy resolution the ECAL is placed inside the solenoid. The electromagnetic calorimeter is an homogeneous calorimeter made of $PbWO_4$ mono-crystals about 22 cm long organized in a barrel and two endcap sections. The scintillating medium $PbWO_4$ has been chosen for its high density ($\rho = 8.3$ g/cm^3), its radiation short length ($X_0 = 0.89$ cm), its small Molière radius ($R_m = 2.19$ cm) and its fast response ($\sim 80\%$ of light is collected in 25 ns). Another characteristic of the ECAL crystals is the radiation hardness. In fact, when exposed to high radiation doses crystals lower their response to a given deposited energy. This effect develops on a short time scale ($\sim$ hours) and need to be monitored and corrected. The radiation dose rate in LHC is up to 0.3 Gy/h in the barrel and up to 15 Gy/h in the endcaps. This adds important contribution to the relative error of the ECAL calorimeter in determining the deposited energy. A drawback of $PbWO_4$ crystals is the low

light yield, approximately 80 γ/MeV, which requires an internal amplification in photodetectors.

The barrel is divided into 36 supermodules. Each supermodule contains $1,700$ crystals and covers a pseudorapidity region from 0 to 1.479 and an azimuthal arc of 20°. Given the granularity of the ECAL, this corresponds to 85 crystals along η and 20 crystals along ϕ. The geometrical shape of the crystals is slightly different in the various pseudorapidity regions: the whole η range is covered by crystal of 17 different shapes, groups of five crystals have the same shape.

The supermodules are symmetrically disposed in the positive in the positive and negative η regions. Each supermodule is divided along η into four modules which are made of submodules, the basic assembling units containing 5×2 crystals each. The crystals axes are oriented with a 3° tilt with respect to the pointing geometry to avoid that particles can directly escape into the dead region between the crystals.

The granularity of the barrel is $\Delta\Phi \times \Delta\eta = 0.0175 \times 0.0175$ and the crystals are grouped, from the readout point of view, into 5×5 arrays corresponding to the trigger towers. The scintillation light is read by APD, Avalanche Photo Diodes.

The endcaps cover the region in η between 1.48 and 3.0. The transition region from the barrel to the endcaps is crucial for hermeticity, which is guaranteed by a conical cut that intercepts the last crystal ring in the barrel to give an overlap of half crystal between the endcap and the barrel.

The crystals in the endcaps have all the same shape and are mounted in a non-pointing geometry, with a granularity varying from $\Delta\Phi \times \Delta\eta = 0.0175 \times 0.0175$ to $\Delta\Phi \times \Delta\eta = 0.05 \times 0.05$. The light signal is read by Vacuum Photo Triodes.

To improve the vertex identification, a preshower is designed to cover the region $1.6 < |\eta| < 2.6$. The preshower is made of two lead converters of thickness $2X_0$ and X_0 followed by silicon strips with a pitch of less than 2 mm. The strips following the first and second absorber are disposed orthogonally. In order to reduce the leakage current in the detector and to reduce the effect of the radiation damage, the silicon strip is operated at a temperature of -5 ℃.

The design of the endcaps ensures a precise calorimetry up to $|\eta| = 2.6$ and the forward region up to $|\eta| = 3.0$, even with a loosen resolution, improves the coverage and therefore the missing transverse energy measurements.

Chapter 3

Fake rate analysis

Higgs decay channel $H \rightarrow WW^{(*)} \rightarrow e\nu e\nu$ provides a well defined signature for Higgs boson search at LHC. This process is dominant for masses higher than 140 GeV/c^2 and one of the W bosons involved may be virtual.

A signal satisfying definite conditions (e.g. a reconstructed track in the tracker matching an energy deposit in the ECAL) is identified as an electron, nevertheless other events may give an electron signal leading to a misidentification. Particle jets are a source of false electron signals defined *fake* electrons.

In particular the W+jets background becomes dangerous when at least one jet is misidentified as an electron. In this case, two electrons are present in the event: one real, prompt electron from the W boson decay, generally isolated and a fake, non-prompt electron from the hadronic decay of short lived mesons (D, B) or the decay in the detector of long lived particles (pions, kaons), or from misidentified hadrons due to e.g. photon conversions or $\pi^{\pm}/\pi^0$ overlap.

The aim of this study is to give an estimate of the probability to have fake electrons from QCD jets using a data-driven technique and extrapolating the fake rate for the W+jets background.
From Monte Carlo truth can be shown [5] that fake electrons in QCD jets follow the misreconstruction of charged pions $\pi^{\pm}$ (66.0%), charged kaons $K^{\pm}$ (14.2%), protons (11.4%), and photons γ (6.13%).

The definition given in section 3.1 can be used to analyze real data and to measure the fake rate.

We also give the integrated fake rate computed as the probability that jets with a momentum higher than a given value can give a fake electron.

3.1 Fake Rate Definition

The jet faking probability is measured from the rate by which jets are identified as electrons, namely:

$$f(\text{jet} \rightarrow \text{fake electron}) = \frac{\#\text{jets matching electron}(l)}{\#\text{reconstructed jets}(d)} = \frac{\mathrm{n_j(e)}}{\mathrm{n_j}} \qquad (3.1)$$

where the numerator is defined as the number of reconstructed jets that match a reconstructed electron, called l-objects. The denominator contains reconstructed jets, called d-objects. The jet reconstruction is performed with the

Iterative Cone Algorithm[1]. All the objects are required to lie inside ECAL acceptance ($|\eta| < 2.5$) and to have a transverse momentum greater than 10 GeV/c. The same definition is used in [6].

A direct measure with a data-driven technique of the fake rate for the W+jets would be hard to compute. It is more difficult to define a pure sample of these events that require a W boson (and then an electron) and a jet signal free from any background. For this reason the large QCD multi-jet data statistics has been used, allowing a precise determination of the jet faking probability as a function of jet transverse momentum p_T.

A sample of about 1 million QCD events, simulated with the PYTHIA [7] generator, has been used and the entire interaction with the detector has been reconstructed. Data has been processed using the CMSSW framework [8] and the analysis has been performed using ROOT [9].

3.1.1 Fakeable Jet Selection Criteria

To perform the fake rate analysis the following procedure is adopted: to every electron with $p_T > 10\ GeV/c$ is associated a jet minimizing the distance

$$\Delta R_{ele-jet} = \sqrt{\Delta\eta^2 + \Delta\Phi^2}$$

where $\Delta\eta = \eta_{ele} - \eta_{jet}$ and analogously $\Delta\Phi = \Phi_{ele} - \Phi_{jet}$.

The W+jets events need to pass one of the Single Lepton Triggers, while QCD events are kept if they fire the prescaled Single Jet Trigger (in the next referred as High Level Trigger, HLT). The prescaling refers to the fact that not all events that satisfy the trigger condition are kept: the limited bandwidth requires a reduction of the event rate because the high cross section of QCD events would fill completely the available band excluding other processes.

The leading jet fires the HLT and it is assumed to play the role of the electron coming from the W-boson decay in the W+jets events that satisfy the Single Lepton Trigger. Thus, all the events have to satisfy the additional requirement that the leading jet lies within $|\eta| < 2.5$.

In order not to introduce a trigger bias, the leading jet is excluded from the fake rate calculation in the QCD events.

Table 3.1 shows the applied HLT thresholds with their names in the CMS Trigger menu.

Using the Monte Carlo truth the jet origin can be identified, then a flavour is assigned to each jet matching the initial partons from the primary physics processes. Possible flavours can be gluon, charm, bottom and light [appendix A]. If the flavour assignment algorithms can not determine unambigously the jet flavour, this is marked as "failed". This component is not shown in the flavour decomposition, but it will be always present in the total faking probability.

[1] A list of objects (whatever particles or HCAL calorimetric towers) ordered in E_T (transverse energy) is prepared. The jet direction (θ, η, Φ) is computed as $\sin\theta = \sum E_T/E$, $\eta = \sum(E_{T_i} \cdot \eta_i)/\sum E_T$ and $\Phi = \sum(E_{T_i} \cdot \Phi_i)/\sum E_T$ where η_i, Φ_i and the sums refer to the jet components. Then the following steps are performed:

- A cone in (η, Φ) is constructed with $R_{cone} = 0.5$ around the highest E_T object;
- A protojet is defined and a new cone is constructed around the protojet;
- The procedure is iterated until the protojet direction variation is less than $\Delta R = 0.01$.

Name	Threshold (GeV/c)
CandHLT1jetPE7	30
HLT1jetPE5	60
HLT1jetPE3	110
HLT1jetPE1	150

Table 3.1: Thresholds for the Single Jet Trigger (High Level Trigger, HLT).

Figure 3.1 shows the correlation between the faking jet transverse momentum and the corresponding misreconstructed electron momentum, computed as the ratio between the electron and the jet p_T, yielding to a distribution of events in function of the jet p_T. Figure 3.2 shows the same results averaged on each momentum bin.

Soft jets are misreconstructed as electrons with a close transverse momentum: for the first three bins ($p_T : 10-15, 15-20, 20-25\ GeV/c$) the ratios are $r_{10-15} = 0.7377 \pm 0.0010$, $r_{20-25} : 0.5783 \pm 0.0010$. and $r_{20-25} = 0.4856 \pm 0.0010$.

Hard jets are misreconstructed as electrons with a lower p_T: for the last three bins ($p_T : 85-90, 90-95, 95-100\ GeV/c$) the ratios are $r_{85-90} = 0.3115 \pm 0.0013$, $r_{90-95} = 0.2966 \pm 0.0014$, and $r_{95-100} = 0.2866 \pm 0.0016$.

We can explain this behavior mainly noting that for soft jets all the particles composing the jet contribute to the electron misreconstruction and so the reconstructed momenta are similar. On the contrary fake electrons in high p_T jets are likely to be real, soft electrons from semileptonic b quark decays. We consider these electrons as fake electrons because they do not come from the signal we are studying, thus contributing to the background.

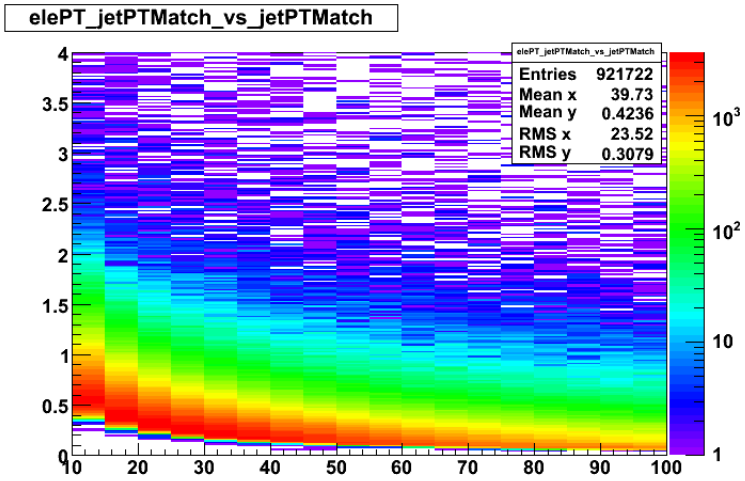


Figure 3.1: Correlation between the faking jet transverse momentum p_T^{jet} and the ratio between electron transverse momentum and the matching jet transverse momentum $\frac{p_T^{ele}}{p_{T\ match}^{jet}}$. The applied HLT sets a threshold at 30 GeV/c.

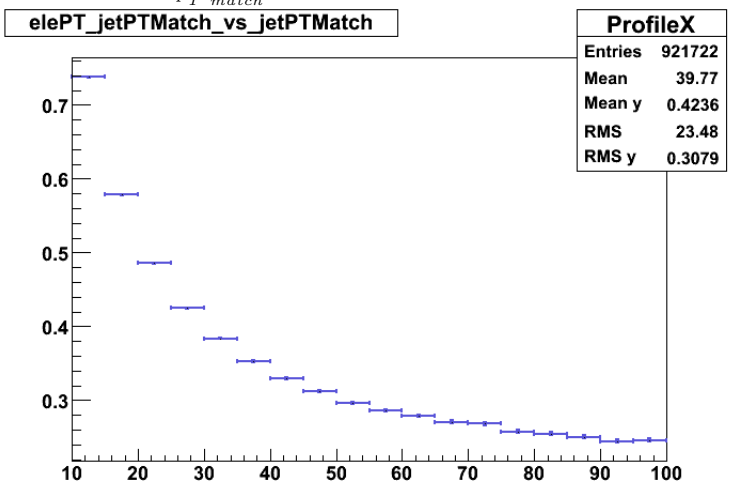


Figure 3.2: Correlation between the faking jet transverse momentum p_T^{jet} and the ratio $\frac{p_T^{ele}}{p_{T\ match}^{jet}}$. The values presented in the previous plot are averaged in each bin.

3.1.2 Electron Selections

An electron signal is defined as a reconstructed track matching an energy deposit in the ECAL calorimeter, but additional electron requirements have to be introduced to limit the background. This data reduction in a hadron collider is useful to reject fake electron contamination from jets.

The various steps through which electron selection is performed are:

1. Electron pre-selection

   1b. Ambiguity solving

2. Electron identification

3. Track based isolation

The effect of the various electron selection steps on the fake rate are shown presenting the single selection steps and the sequence of selections steps. Pre-selection criteria being always applied, the fake rate is computed for electron pre-selection (*single_resol*), electron pre-selection and identification (*single_eleIdTight*), electron pre-selection and track isolation (*single_tkIso*) and the sequence of all selections (*sequence_tkIso*).

Electron Pre-Selection

In the pre-selection step, a loose geometrical and energy momentum matching is imposed between the reconstructed electron track and the corresponding supercluster.

Electron candidates are defined as:

- a reconstructed electron track initiated by the reconstruction of a supercluster in the ECAL matched with hits in the pixel detector;
- an η geometrical matching[2] $|\Delta\eta| = |\eta_{SC} - \eta_{in}^{extrap.}| < 0.02$;
- an energy-momentum matching between the supercluster and the track, $E/p < 3$;
- a Φ geometrical matching $|\Delta\Phi| = |\Phi_{SC} - \Phi_{in}^{extrap.}| < 0.1$;
- a ratio of the energy deposited in the HCAL tower just behind the electromagnetic seed cluster over the energy of the seed cluster $H/E < 0.2$.

If the leading jet is close to the misreconstructed electron, within a cone with $\Delta R < 0.3$, then it is rejected. This avoids the introduction of systematic effects in the analysis.

Ambiguities may arise when a given track matches with different supercluster seeds (i.e. with different energy deposits in the ECAL) or, conversely, when a given supercluster seed results in more than one trajectory match. These ambiguities must be resolved to avoid double counting of electrons, thus only the pre-selected electron with the best E/p ratio is kept. Electrons that pass the ambiguity solving selection have an "Ambiguity bit" set.

Parameters applied in the first stage of selection are shown in table 3.2.

[2]The geometrical matching is performed taking the track parameters at the interaction vertex (η_{in}, Φ_{in}) and extrapolating to ECAL assuming a perfect helix, and matching the resulting ($\eta_{in}^{extrap.}$, $\Phi_{in}^{extrap.}$) to the energy-weighted position of the supercluster (η_{SC}, Φ_{SC})

Parameters	
Reconstructed track	
Electron p_T	$p_T^{ele} > 10\ GeV/c$
Acceptance	$\eta_{ele} < 2.5$
Distance jet-leading jet	$\Delta R_{ele-jet} > 0.3$
η Geometrical matching	$\lvert\Delta\eta\rvert < 0.02$
Φ Geometrical matching	$\lvert\Delta\Phi\rvert < 0.1$
HCAL/ECAL E	$H/E < 0.2$
Ambiguity solving bit	

Table 3.2: The pre-selection requirements for reconstructed objects (l).

Figures 3.3 and 3.4 show the fake rate estimate as a function of p_T for different HLT thresholds. The straight line shows the total fake rate computed as:

$$f = \sum_i s_i \cdot w_i$$

where f is the fake rate, s_i is the fake rate for the i-th flavour and w_i is the relative weight[3]. The flavour decomposition is also shown.

The fake rate is not particularly affected by the different HLT thresholds remaining on the order of $\mathcal{O}(10^{-1})$ on all the p_T range studied.

The jet origin influences the faking probability, table 3.3 shows some values of electron fake rate for different p_T values and jet flavours, in fact b-jets present an higher faking probability due the presence of real electrons in hard jets.

Table 3.4 shows the percentage composition of the faking objects, computed ad the ratio between the number of faking object with an assigned flavour and the total number of faking objects. We can note that, unless an higher faking probability, heavy quarks (b, c) contribute less to the total fake rate because the event rate for these type of interactions is lower.

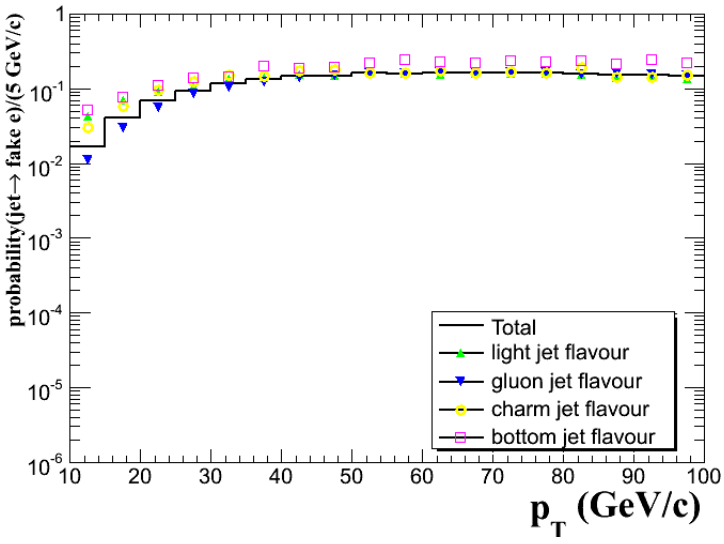
[3]The weight is calculated as: $w_i = \frac{\#\ fake\ events\ with\ flavour\ \text{“}i\text{”}}{\#\ total\ events}$.

CandHLT1jetPE7 (30 GeV/c)				HLT1jetPE5 (60 GeV/c)		
	a	b	c	a	b	c
T	$6.81 \cdot 10^{-02}$	$1.60 \cdot 10^{-01}$	$1.56 \cdot 10^{-01}$	$7.82 \cdot 10^{-02}$	$1.60 \cdot 10^{-01}$	$1.56 \cdot 10^{-01}$
l	$9.07 \cdot 10^{-02}$	$1.63 \cdot 10^{-01}$	$1.50 \cdot 10^{-01}$	$1.05 \cdot 10^{-01}$	$1.63 \cdot 10^{-01}$	$1.50 \cdot 10^{-01}$
g	$5.51 \cdot 10^{-02}$	$1.55 \cdot 10^{-01}$	$1.51 \cdot 10^{-01}$	$6.46 \cdot 10^{-02}$	$1.55 \cdot 10^{-01}$	$1.51 \cdot 10^{-01}$
c	$9.59 \cdot 10^{-02}$	$1.63 \cdot 10^{-01}$	$1.89 \cdot 10^{-01}$	$9.54 \cdot 10^{-02}$	$1.63 \cdot 10^{-01}$	$1.89 \cdot 10^{-01}$
b	$1.09 \cdot 10^{-01}$	$2.20 \cdot 10^{-01}$	$2.31 \cdot 10^{-01}$	$1.23 \cdot 10^{-01}$	$2.20 \cdot 10^{-01}$	$2.31 \cdot 10^{-01}$
HLT1jetPE3 (110 GeV/c)				HLT1jetPE1 (150 GeV/c)		
T	$7.93 \cdot 10^{-02}$	$1.62 \cdot 10^{-01}$	$1.57 \cdot 10^{-01}$	$7.94 \cdot 10^{-02}$	$1.60 \cdot 10^{-01}$	$1.60 \cdot 10^{-01}$
l	$9.73 \cdot 10^{-02}$	$1.62 \cdot 10^{-01}$	$1.50 \cdot 10^{-01}$	$1.07 \cdot 10^{-01}$	$1.65 \cdot 10^{-01}$	$1.49 \cdot 10^{-01}$
g	$6.93 \cdot 10^{-02}$	$1.58 \cdot 10^{-01}$	$1.52 \cdot 10^{-01}$	$6.68 \cdot 10^{-02}$	$1.51 \cdot 10^{-01}$	$1.55 \cdot 10^{-01}$
c	$1.04 \cdot 10^{-01}$	$1.53 \cdot 10^{-01}$	$1.89 \cdot 10^{-01}$	$1.05 \cdot 10^{-01}$	$1.69 \cdot 10^{-01}$	$1.95 \cdot 10^{-01}$
b	$1.12 \cdot 10^{-01}$	$2.24 \cdot 10^{-01}$	$2.32 \cdot 10^{-01}$	$1.10 \cdot 10^{-01}$	$2.41 \cdot 10^{-01}$	$2.44 \cdot 10^{-01}$

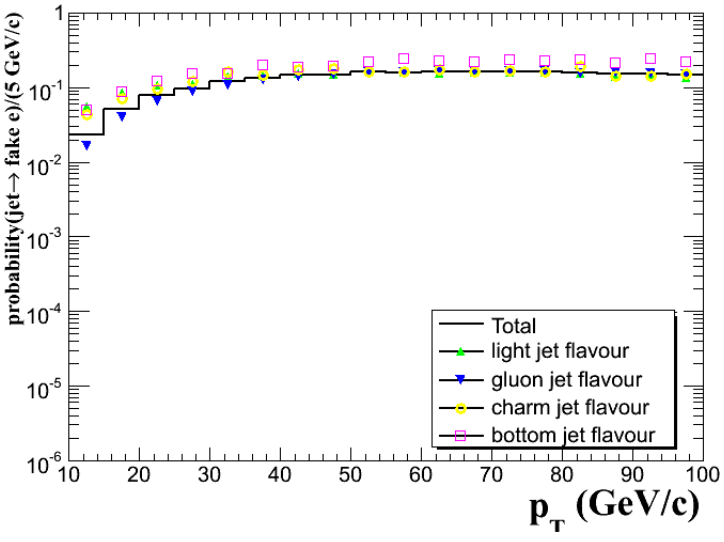
Table 3.3: The total fake rate after the electron pre-selection step with flavour decomposition (total (T), light (l), gluon (g), charm (c) and bottom (b)). Results are shown for different values of p_T: (a) 22.5 GeV/c, (b) 52.5 GeV/c, (c) 82.5 GeV/c

CandHLT1jetPE7 (30 GeV/c)				HLT1jetPE5 (60 GeV/c)			
	a	b	c		a	b	c
light	35.3%	29.8%	31.2%	light	33.4%	29.8%	31.2%
gluon	52.5%	58.3%	54.2%	gluon	54.3%	58.3%	54.2%
charm	7.7%	6.8%	8.7%	charm	6.8%	6.8%	8.7%
bottom	4.4%	5.0%	5.9%	bottom	5.4%	5.0%	5.9%
HLT1jetPE3 (110 GeV/c)				HLT1jetPE1 (150 GeV/c)			
light	28.5%	27.7%	31.2%	light	30.6%	28.5%	28.2%
gluon	58.0%	60.4%	54.1%	gluon	56.5%	56.8%	55.5%
charm	7.7%	6.1%	8.8%	charm	7.5%	7.3%	9.2%
bottom	5.7%	5.8%	5.9%	bottom	5.4%	7.4%	7.1%

Table 3.4: Percentage flavour composition of faking objects after the electron pre-selection step. Results are shown for different values of p_T: (a) 22.5 GeV/c, (b) 52.5 GeV/c, (c) 82.5 GeV/c

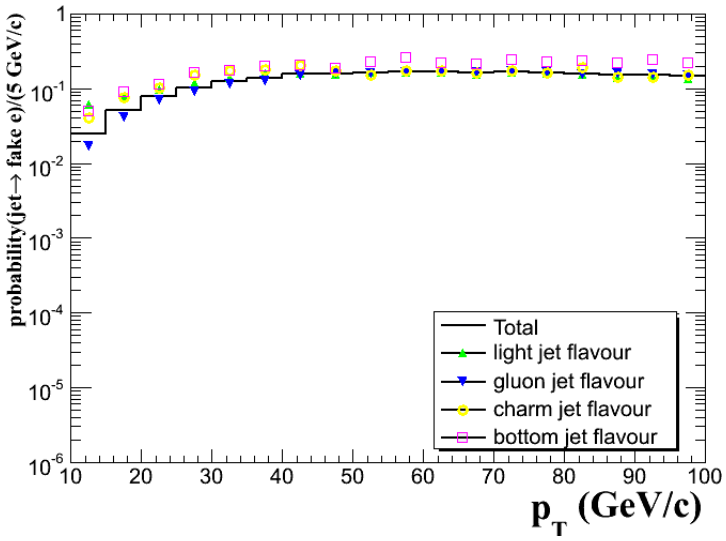


(a)

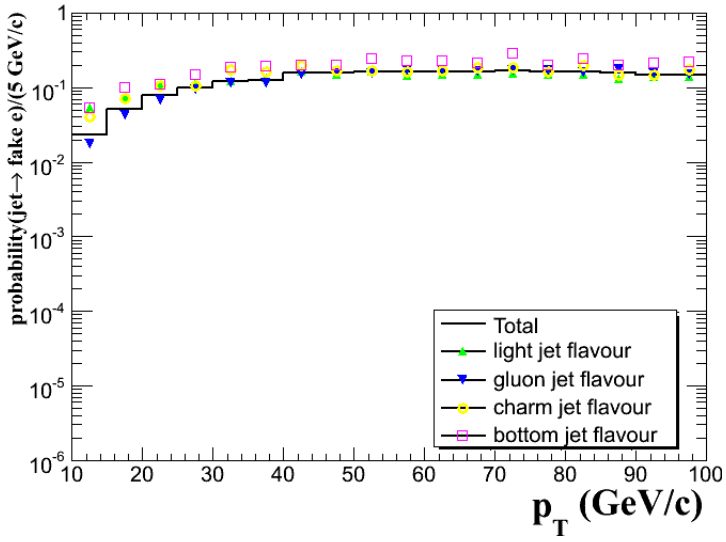


(b)

Figure 3.3: The electron fake rate vs jet p_T after the first electron selection (*single_resol*) with flavour decomposition. The HLT is set at 30 (a) and 60 GeV/c (b).



(a)



(b)

Figure 3.4: The electron fake rate vs jet p_T after the first electron selection (*single_resol*) with flavour decomposition. The HLT is set at 110 (a) and 150 GeV/c (b).

Electron Identification

Further electron identification requirements must be imposed to suppress the possible background [10]. Elaborate reconstruction procedures, involving recognition of distinct track-supercluster patterns, are used to distinguish between well measured and badly measured electrons. For this reason, four electron classes are defined: golden electron (well measured electrons), big brem electrons (electrons which show a large bremsstrahlung radiation), narrow electrons and showering electrons (badly measured electrons) [appendix B].
The procedure of electron identification uses different variables depending on the electrons class. Used variables are:

- H/E, the ratio between the HCAL energy deposit and the ECAL supercluster energy deposit;
- $|\Delta\Phi| = |\Phi_{SC} - \Phi_{in}^{extrap.}|$ and $|\Delta\eta| = |\eta_{SC} - \eta_{in}^{extrap.}|$, defined using vertex parameters;
- Σ_9/Σ_{25}, the ratio between the energy deposited in a 3×3 crystal matrix and a 5×5 crystal matrix in ECAL;
- $\sigma_{\eta\eta}$, a weighted average of the dispersion in η of crystals around the most energetic one: $\sigma_{\eta\eta}^2 = \sum_{crystals} \left(\frac{E_i}{E_{seed}} (\eta_i - \eta_{seed})^2 \right)$;

Electrons satisfying the electron identification requirements have a "EleIdTight bit" set.

Table 3.5 summarizes the selections applied.

Parameters	
Pre-selection requirements	
EleIdTightBit	
Electron p_T	$p_T^{ele} > 10\ GeV/c$
Acceptance	$\eta_{ele} < 2.5$
Distance jet-leading jet	$\Delta R_{ele-jet} > 0.3$

Table 3.5: The eleID requirements for reconstructed objects (l).

Figure 3.5 shows the total fake rate estimate as a function of p_T for different thresholds of the HLT after the electron isolation selection. The measured fake rate is of the order of $\mathcal{O}(10^{-3})$.

Electron Isolation

For the electrons the isolation criterion is obtained from tracks originating from a common (primary) vertex. Using track measurements at a primary vertex avoids the complications due to severe external bremsstrahlung, photon conversion and early showering in the tracker material. The isolation procedure works in the following way:

- The sum of all tracks with transverse momentum $p_T > 1.5\ GeV/c$ within an isolation cone in the (η, Φ) plane is taken. The radius of the cone is given by $R_{cone} = \sqrt{\Delta\eta^2 + \Delta\Phi^2} < 0.3$, centered on each electron.

- Reconstructed tracks must also satisfy a condition on the impact parameter ΔIP_L, the difference between the z position of the reconstructed track primary vertex and the z position of the electron, namely $|\Delta IP_L| = |z^e_{IP} - z^{tr}_{IP}| < 0.1$;
- To avoid counting the electron itself in the isolation procedure[4], a veto cone is then defined around the direction of each reconstructed electron candidate with $R_{veto} = 0.015$.

Isolated electrons are defined as the ones for which the sum of the transverse momentum of the tracks contained in the isolation cone divided by the electron p_T is below a certain threshold, $\left(\sum_{tck} \left(p_T^{tck}/p_T^{ele}\right) < 0.05\right)$.
The isolation algorithm is performed separately for each reconstructed electron in the event.

Table 3.6 shows the selections applied in the *single_tkIso* stage of selection.

Parameters	
Pre-selection requirements	
tkIsoBit $R_{cone} < 0.3$	$\sum_{tck} \left(p_T^{tck}/p_T^{ele}\right) < 0.05$
Track momentum	$p_T > 1.5\ GeV/c$
Interaction vertex distance	$\lvert\Delta IP_L\rvert < 0.1$
Veto cone	$R_{cone} < 0.015$
Electron p_T	$p_T^{ele} > 10\ GeV/c$
Acceptance	$\eta_{ele} < 2.5$
No closest leading jet	$\Delta R_{ele-jet} > 0.3$

Table 3.6: The isolation requirements for reconstructed objects (l).

Figure 3.6 shows the total fake rate estimate as a function of p_T for different thresholds of the HLT. We measure a fake rate of the order of $\mathcal{O}(10^{-3})$. Furthermore we can note that with this algorithm the electron fake rate decrease when the jet p_T increases. This can be explained noting that the isolation algorithm has a higher rejection power for high values of p_T [10].
A reduction of the fake rate from b quarks is also noted as we expected according to the origin proposed.

[4]The tracks summed in the previous steps are reconstructed in a completely separate and parallel way from the track associated to the electron candidate, therefore a track associated to the electron itself can be contained within the list of the isolation tracks. This is avoided with the veto cone.

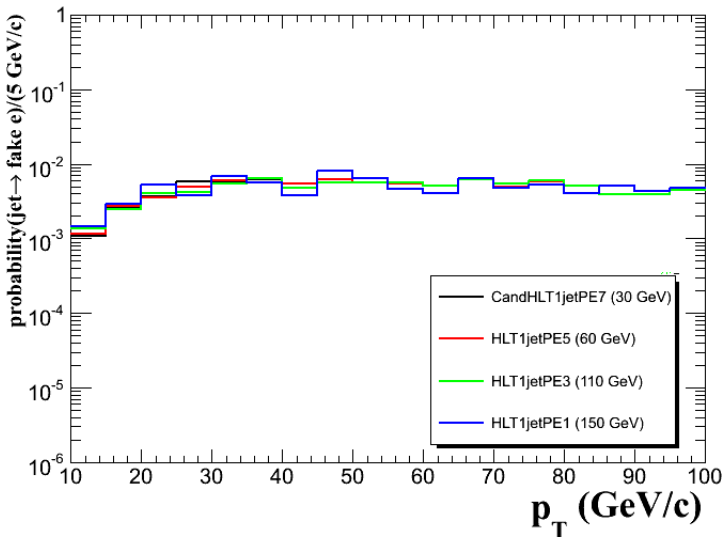


Figure 3.5: The electron fake rate vs jet p_T for jets matching electrons that satisfy the pre-selection criteria and the identification algorithm (*single_eleId*).

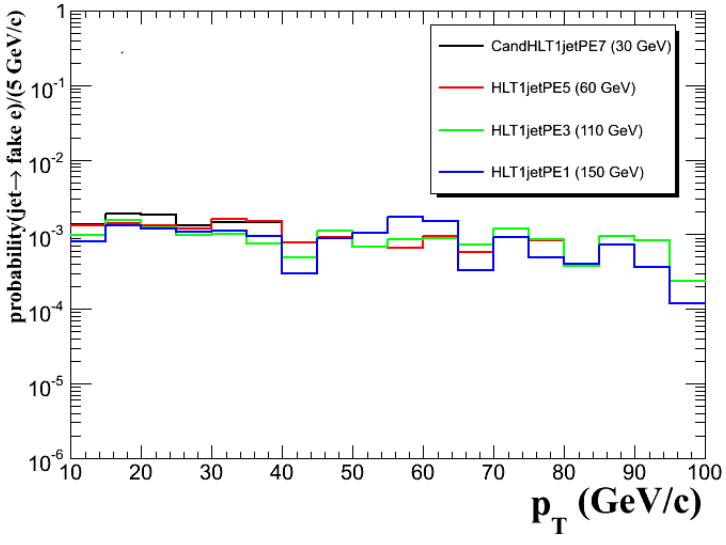


Figure 3.6: The electron fake rate vs jet p_T for jets matching electrons that satisfy the pre-selection criteria and the isolation algorithm (*single_tkIso*).

Selections Sequence

Now we present the fake rate after the application in sequence of all the selection criteria, thus we are considering jets matching electrons which passed the *pre-selection*, the *identification* and the *isolation* selection steps.

Table 3.7 gives an outline of all selections applied.

Parameters	
Reconstructed track	
Ambiguity solving bit	
EleIdTight bit	
η Geometrical matching	$\lvert\Delta\eta_{in}\rvert < 0.02$
Φ Geometrical matching	$\lvert\Delta\Phi\rvert < 0.1$
HCAL/ECAL E	$H/E < 0.2$
tkIsoBit $R_{cone} < 0.3$	$\sum_{tck}\left(p_T^{tck}/p_T^{ele}\right) < 0.05$
Track momentum	$p_T > 1.5\ GeV/c$
Interaction vertex distance	$\lvert\Delta IP_L\rvert < 0.1$
Veto cone	$R_{cone} < 0.015$
Electron p_T	$p_T^{ele} > 10GeV/c$
Acceptance	$\eta_{ele} < 2.5$
Distance jet-leading jet	$\Delta R_{ele-jet} > 0.3$

Table 3.7: The requirements for reconstructed objects (l) after that all selections are applied.

Figures 3.7 and 3.8 show the electron faking probability from QCD jets for electron satisfying all selections. The measured fake rate is of the order of $\mathcal{O}(10^{-4})$.

This analysis is affected by a lack of statistics, in fact, we have discontinuous data for the flavour composition of the faking jets. The faking jets appearing in the total probability can be assigned to the "failed" flavour. The effect is more important for higher HLT thresholds and for higher jet p_T values.

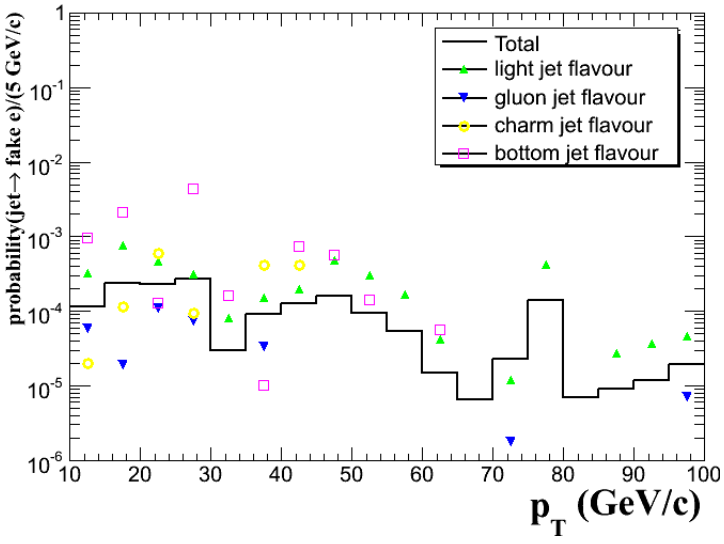
Table 3.8 presents the fake rate for several p_T values with flavour decomposition and table 3.9 shows the percentage flavour composition of the faking objects.

CandHLT1jetPE7 (30 GeV/c)				HLT1jetPE5 (60 GeV/c)		
	a	b	c	a	b	c
T	$2.29 \cdot 10^{-04}$	$9.36 \cdot 10^{-05}$	$6.88 \cdot 10^{-06}$	$3.54 \cdot 10^{-05}$	$9.36 \cdot 10^{-05}$	$6.88 \cdot 10^{-06}$
l	$4.53 \cdot 10^{-04}$	$2.97 \cdot 10^{-04}$	–	$7.69 \cdot 10^{-05}$	$2.97 \cdot 10^{-04}$	–
g	$1.07 \cdot 10^{-04}$	$4.71 \cdot 10^{-07}$	–	$7.15 \cdot 10^{-08}$	$4.71 \cdot 10^{-07}$	–
c	$5.98 \cdot 10^{-04}$	–	–	$1.85 \cdot 10^{-05}$	–	–
b	$1.25 \cdot 10^{-04}$	$1.41 \cdot 10^{-04}$	–	$1.38 \cdot 10^{-04}$	$1.41 \cdot 10^{-04}$	–
HLT1jetPE3 (110 GeV/c)				HLT1jetPE1 (150 GeV/c)		
T	$6.35 \cdot 10^{-05}$	$1.22 \cdot 10^{-04}$	$6.95 \cdot 10^{-06}$	$1.75 \cdot 10^{-04}$	$5.99 \cdot 10^{-05}$	$2.15 \cdot 10^{-05}$
l	$2.47 \cdot 10^{-04}$	$3.62 \cdot 10^{-04}$	–	$6.60 \cdot 10^{-04}$	$1.33 \cdot 10^{-04}$	–
g	$8.53 \cdot 10^{-07}$	$1.51 \cdot 10^{-06}$	–	$3.49 \cdot 10^{-06}$	$7.98 \cdot 10^{-06}$	–
c	$3.67 \cdot 10^{-05}$	–	–	$1.57 \cdot 10^{-04}$	–	–
b	$1.39 \cdot 10^{-05}$	$4.03 \cdot 10^{-04}$	–	$6.06 \cdot 10^{-05}$	–	–

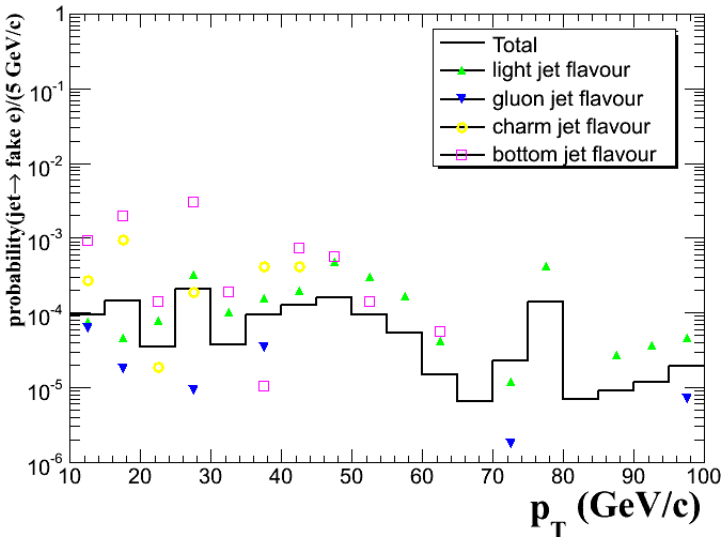
Table 3.8: The total fake rate after all the selection steps and the flavour decomposition (total (T), light (l), gluon (g), charm (c), and bottom (b)). Results are shown for different values of p_T: (a) 22.5 GeV/c, (b) 52.5 GeV/c, (c) 82.5 GeV/c

CandHLT1jetPE7 (30 GeV/c)				HLT1jetPE5 (60 GeV/c)			
	a	b	c		a	b	c
light	52.6%	93.0%	–	light	54.0%	93.0%	–
gluon	30.4%	0.3%	–	gluon	0.1%	0.3%	–
charm	14.3%	–	–	charm	2.9%	–	–
bottom	1.5%	5.5%	–	bottom	13.4%	5.5%	–
HLT1jetPE3 (110 GeV/c)				HLT1jetPE1 (150 GeV/c)			
light	90.5%	82.4%	–	light	85.8%	61.3%	–
gluon	0.9%	0.8%	–	gluon	1.3%	8.0%	–
charm	3.4%	–	–	charm	5.1%	–	–
bottom	0.9%	13.9%	–	bottom	1.3%	–	–

Table 3.9: The percentage flavour composition of fake objects after all selections steps. Results are shown for different values of p_T: (a) 22.5 GeV/c, (b) 52.5 GeV/c, (c) 82.5 GeV/c

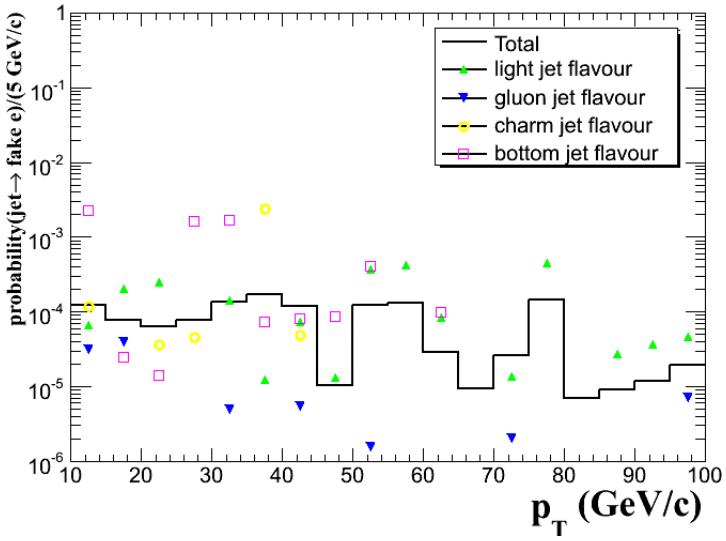


(a)

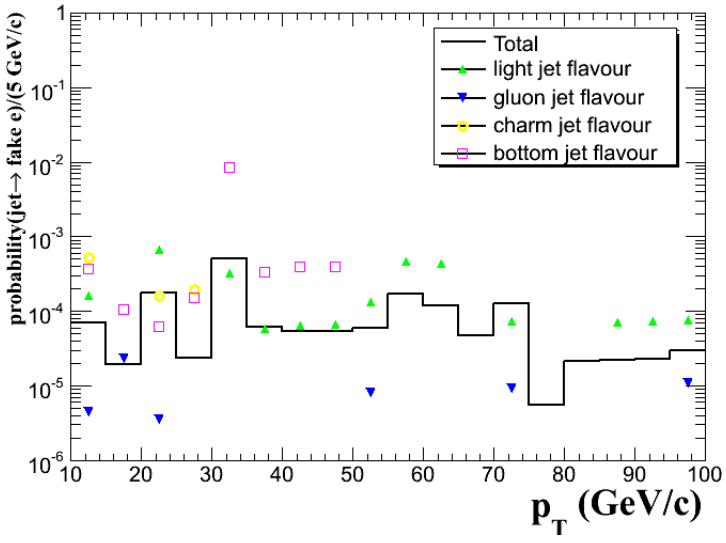


(b)

Figure 3.7: The electron fake rate vs jet p_T after all the stages of selection with flavour decomposition. The HLT is set at 30 (a) and 60 GeV/c (b).



(a)



(b)

Figure 3.8: The electron fake rate vs jet p_T after all the stages of selection with flavour decomposition. The HLT is set at 110 (a) and 150 GeV/c (b).

3.2 Integrated Probability

The integrated fake rate is computed using the following formula:

$$\mathcal{F}(\mathrm{p_T}, \eta) = \frac{\int \mathrm{n_j(e)}}{\int \mathrm{n_j}} \tag{3.2}$$

where the integrations are performed starting from an assigned value up to the maximum possible value. The fake rate is calculated for starting values in the range $p_T = 10 - 100\ GeV/c$ with a step of 5 GeV/c. The ratios are shown as a function of the lower integration limit.

We present the integral fake rate computed for the electron pre-selection step (*Int_single_resol*) and for the sequence of all the selections described before (*Int_sequence_tkIso*).

Electron Pre-Selection

Figure 3.9 shows the integral fake rate after the first selection step, the electron pre-selection. The flavour decomposition of the fake rate is shown.

Figure 3.10 shows the integral fake rate after the first selection step with the application of different HLT thresholds. Only the total faking probability is presented. We can see that the HLT does not modify the fake rate behavior and the integral faking probability is of the order of $\mathcal{O}(10^{-1})$.

Some values token at different p_T bins of the integral faking probability rate are reported in table 3.10. Once again we see that the highest faking probability comes from b quarks.

	CandHLT1jetPE7 (30 GeV/c)			HLT1jetPE5 (60 GeV/c)		
	a	b	c	a	b	c
T	$9.44 \cdot 10^{-02}$	$1.60 \cdot 10^{-01}$	$1.53 \cdot 10^{-01}$	$1.16 \cdot 10^{-01}$	$1.60 \cdot 10^{-01}$	$1.53 \cdot 10^{-01}$
l	$1.11 \cdot 10^{-01}$	$1.56 \cdot 10^{-01}$	$1.42 \cdot 10^{-01}$	$1.30 \cdot 10^{-01}$	$1.56 \cdot 10^{-01}$	$1.42 \cdot 10^{-01}$
g	$8.25 \cdot 10^{-02}$	$1.58 \cdot 10^{-01}$	$1.53 \cdot 10^{-01}$	$1.05 \cdot 10^{-01}$	$1.58 \cdot 10^{-01}$	$1.53 \cdot 10^{-01}$
c	$1.24 \cdot 10^{-01}$	$1.65 \cdot 10^{-01}$	$1.61 \cdot 10^{-01}$	$1.41 \cdot 10^{-01}$	$1.65 \cdot 10^{-01}$	$1.61 \cdot 10^{-01}$
b	$1.42 \cdot 10^{-01}$	$2.27 \cdot 10^{-01}$	$2.28 \cdot 10^{-01}$	$1.67 \cdot 10^{-01}$	$2.27 \cdot 10^{-01}$	$2.28 \cdot 10^{-01}$
	HLT1jetPE3 (110 GeV/c)			HLT1jetPE1 (150 GeV/c)		
T	$1.41 \cdot 10^{-01}$	$1.62 \cdot 10^{-01}$	$1.53 \cdot 10^{-01}$	$1.39 \cdot 10^{-01}$	$1.60 \cdot 10^{-01}$	$1.54 \cdot 10^{-01}$
l	$1.45 \cdot 10^{-01}$	$1.55 \cdot 10^{-01}$	$1.42 \cdot 10^{-01}$	$1.38 \cdot 10^{-01}$	$1.46 \cdot 10^{-01}$	$1.39 \cdot 10^{-01}$
g	$1.33 \cdot 10^{-01}$	$1.61 \cdot 10^{-01}$	$1.53 \cdot 10^{-01}$	$1.34 \cdot 10^{-01}$	$1.62 \cdot 10^{-01}$	$1.56 \cdot 10^{-01}$
c	$1.63 \cdot 10^{-01}$	$1.65 \cdot 10^{-01}$	$1.61 \cdot 10^{-01}$	$1.59 \cdot 10^{-01}$	$1.70 \cdot 10^{-01}$	$1.66 \cdot 10^{-01}$
b	$1.97 \cdot 10^{-01}$	$2.30 \cdot 10^{-01}$	$2.29 \cdot 10^{-01}$	$1.99 \cdot 10^{-01}$	$2.28 \cdot 10^{-01}$	$2.19 \cdot 10^{-01}$

Table 3.10: The integral fake rate after the electron pre-selection step with flavour decomposition (total (T), light (l), gluon (g), charm (c), and bottom (b)). Results are shown for different values of p_T: (a) 22.5 GeV/c, (b) 52.5 GeV/c, (c) 82.5 GeV/c

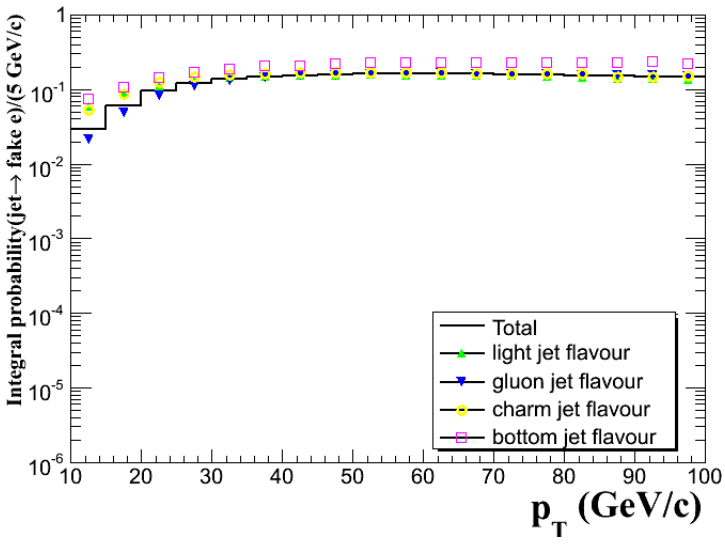


Figure 3.9: The integrated electron fake rate with flavout decomposition after the first electron selection (*Int_single_resol*) as a function of the lower integration limit. The HLT threshold is set at 30 GeV/c.

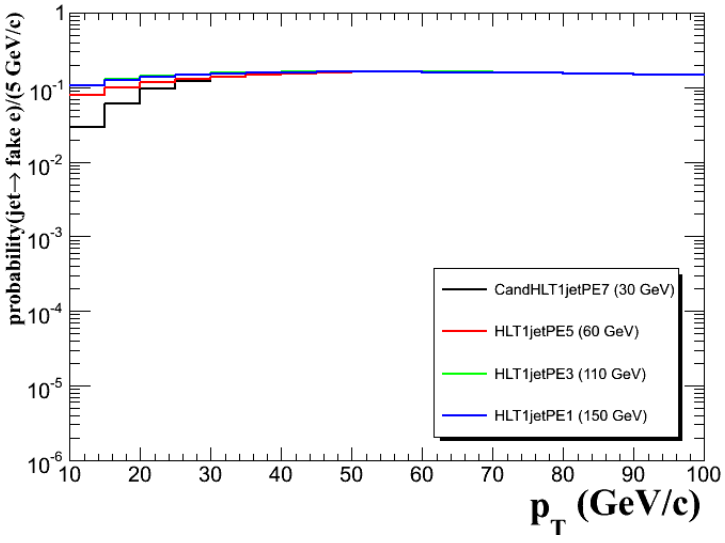


Figure 3.10: The integrated electron fake rate after the first electron selection (*Int_single_resol*) with different HLT thresholds as a function of the lower integration limit.

Selections Sequence

Figure 3.11 shows the integral fake rate and the flavour decomposition after all selection step: the electron pre-selection, the electron identification and the electron isolation.

Figure 3.12 shows the integral fake rate with the application of different HLT thresholds. Only the total faking probability is shown. We can see that the HLT does not modify the fake rate behavior and the integral faking probability is of the order of $\mathcal{O}(10^{-4})$.

Table 3.11 presents in detail the integral faking rate probability at different p_T bins. Once again we see that the highest faking probability comes from b quarks.

Also in this case we have to consider "failed" flavour jets to account for the total faking probability.

	CandHLT1jetPE7 (30 GeV/c)			HLT1jetPE5 (60 GeV/c)		
	a	b	c	a	b	c
T	$1.88 \cdot 10^{-04}$	$5.11 \cdot 10^{-05}$	$1.05 \cdot 10^{-05}$	$9.23 \cdot 10^{-05}$	$5.11 \cdot 10^{-05}$	$1.05 \cdot 10^{-05}$
l	$3.27 \cdot 10^{-04}$	$1.48 \cdot 10^{-04}$	$2.24 \cdot 10^{-05}$	$1.81 \cdot 10^{-04}$	$1.48 \cdot 10^{-04}$	$2.24 \cdot 10^{-05}$
g	$7.06 \cdot 10^{-05}$	$4.07 \cdot 10^{-07}$	$1.14 \cdot 10^{-06}$	$5.96 \cdot 10^{-06}$	$4.07 \cdot 10^{-07}$	$1.14 \cdot 10^{-06}$
c	$3.37 \cdot 10^{-04}$	–	–	$1.26 \cdot 10^{-04}$	–	–
b	$1.02 \cdot 10^{-03}$	$4.80 \cdot 10^{-05}$	–	$7.36 \cdot 10^{-04}$	$4.80 \cdot 10^{-05}$	–
	HLT1jetPE3 (110 GeV/c)			HLT1jetPE1 (150 GeV/c)		
T	$7.91 \cdot 10^{-05}$	$6.18 \cdot 10^{-05}$	$1.05 \cdot 10^{-05}$	$1.07 \cdot 10^{-04}$	$6.50 \cdot 10^{-05}$	$2.38 \cdot 10^{-05}$
l	$1.31 \cdot 10^{-04}$	$1.70 \cdot 10^{-04}$	$2.25 \cdot 10^{-05}$	$1.74 \cdot 10^{-04}$	$1.32 \cdot 10^{-04}$	$5.39 \cdot 10^{-05}$
g	$1.26 \cdot 10^{-06}$	$7.18 \cdot 10^{-07}$	$1.15 \cdot 10^{-06}$	$1.82 \cdot 10^{-06}$	$2.77 \cdot 10^{-06}$	$2.55 \cdot 10^{-06}$
c	$2.01 \cdot 10^{-04}$	–	–	$3.21 \cdot 10^{-05}$	–	–
b	$3.81 \cdot 10^{-04}$	$7.83 \cdot 10^{-05}$	–	$7.33 \cdot 10^{-04}$	–	–

Table 3.11: The integral fake rate after all the selection steps with flavour decomposition (total (T), light (l), gluon (g), charm (c), and bottom (b)). Results are shown for different values of p_T: (a) 22.5 GeV/c, (b) 52.5 GeV/c, (c) 82.5 GeV/c

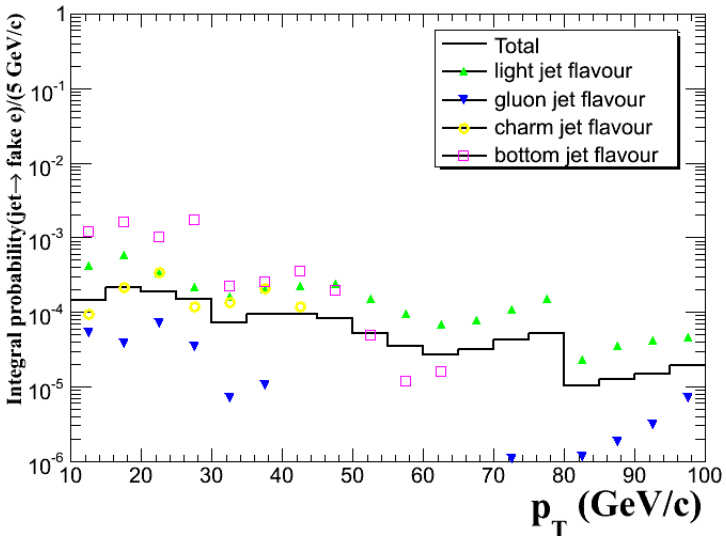


Figure 3.11: The integrated electron fake rate flavour decomposition for jets matching electrons which passed all the selection criteria as a function of the lower integration limit. The HLT threshold is set at 30 GeV/c.

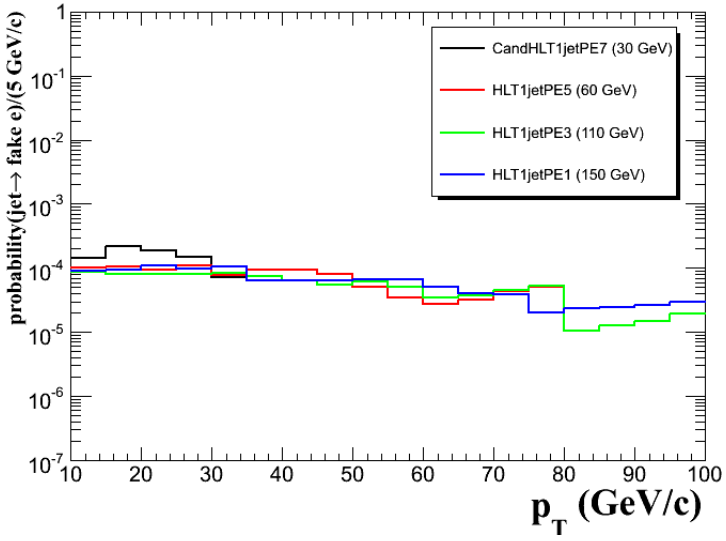


Figure 3.12: The integrated electron fake rate for jets matching electrons which passed all the selection criteria as a function of the lower integration limit. Only the total integral fake rate is shown with different HLT thresholds.

Chapter 4

Summary and Conclusions

The aim of this study was to give an estimate of the fake rate probability for the W+jets background signal affecting the $H \rightarrow e\nu e\nu$ channel of Higgs boson decay. A data driven method has been used based on QCD generated jets to provide a wider event statistics. In QCD events a High Level Trigger (HLT) is set on the leading jet momentum and four different trigger thresholds have been used respectively at 30, 60, 110 and 150 GeV/c.

In chapter 3.1 we have defined "fakeable" objects to be all the reconstructed jets respecting definite geometrical and energetic requirements and "faking" objects to be the jets close enough to a reconstructed electron. The faking probability rate has been computed as the ratio between the number of faking and fakeable objects, obtaining a function of the jets' transverse momentum.

In section 3.1.1 we have shown the correlation in momentum between faking jets and fake electrons and we have introduced the jet flavour, a property depending on the jet origin useful to spot out the physical processes that have a higher faking probability.

In section 3.1.2 we have applied different selection requirements to the electrons thus establishing different constraints on the numerator objects. In the first stage only pre-selection requirements were assumed, consisting substantially in an energy-momentum and geometrical matching between a reconstructed track and an ECAL supercluster energy deposit. The fake rate measured in this stage was of the order of $\mathcal{O}(10^{-1})$. Further selections, the electron identification and the electron isolation, have been applied and the effect on fake rate has been studied.
After all selection requirements the electron fake rate from QCD jets is decreased to the order of $\mathcal{O}(10^{-4})$.
This work has also shown that the application of different prescaled Single Jet Trigger thresholds does not vary the fake rate using these selections.

We can now give an estimate of the number of fake events φ for the W+jets background extrapolating from the QCD value with the following formula:

$$\varphi_{W+jets}(\text{jet} \rightarrow \#\text{ fake electron}) = \mathcal{L} \times \sigma(W + jets) \times BR(W \rightarrow e\nu) \times \varepsilon_t \times f_{QCD}$$

Where $\mathcal{L}$ is the integrated luminosity, σ_{W+jets} is the cross section for the process $H \rightarrow W + jets$, $BR_{W \rightarrow e\nu}$ is the branching ratio for the decay of the W boson in a couple (e,ν), ε_t the trigger efficiency, namely the rate with which

$W \to e\nu$ pass the Single Lepton Trigger and f_{QCD} is the estimated fake rate from QCD jets.

In chapter 3.2 the integrated fake rate has been calculated as the ratio between the number of all jets with a momentum higher than a given value that match a reconstructed electron over the number of jets with the momentum higher than the same value. After all selection cuts the integral fake rate has been found to be of the order of $\mathcal{O}(10^{-5})$.

In [6] has been shown that the event selection for the Higgs channel considered could be extended from a double electron signal $(l + l)$ to an electron plus jet signal $(l+d)$, where l and d were respectively a numerator and a denominator object in eq. 3.1. The number of fake events estimated from $(l + d)$ has then been compared with the $(l+l)$-selected events from W+jets and the distibution have been found to be in agreement, validating the estimation method.

Appendix A: Jet Flavours

Jets are classified in three different classes:

- `CaloJets`, jets produced from calorimetric towers. They carry energy-momentum data. Informations on the energy fraction deposited in each calorimeter module is available.
- `GenJets`: jets produced from particles at Monte Carlo level with detailed data about the constituents of the jets themselves.
- `PFJets`: Particle Flow jets (jets that contains a "list" of particles divided in photons, charged hadrons, neutral hadrons, electrons, muons).

There is no unambiguous answer to establish the correct underlying flavour of a reconstructed jet. For the flavour assignment three definitions are used, reflecting three different points of view:

`Physics definition`: matches reconstructed jets to "initial" partons from the primary physics process (within $\Delta R < 0.3$ of reconstructed jet cone). No matching (and thus the flavour "failed" is assigned) if hard radiation occurred and the parton direction changed significantly.

`Algorithmic definition`: Tries to find the parton that most likely determines the properties of the jet and assign that flavour as true flavour. Here, the final state partons (after showering or bremsstrahlung radiation processes) are analyzed (also within $\Delta R < 0.3$ of reconstructed jet cone). Jets from radiation are matched with full efficiency. If there is a b (c) quark within the jet cone, it is labeled as b (c). Otherwise, the flavour of the hardest parton is assigned.

`Energetic definition`: is applied to GenJets; a variable is built for each jet computing the fraction of the energy of the jet which comes from b and c hadrons. The two variables can be used to attribute a flavour to the GenJet, finally a matched CaloJet (for example by ΔR) can get the same flavour of the matched GenJet.

All the three definitions can be applied to GenJets while only the first two (physics and algorithmic definitions) can be applied to CaloJets and PFJets.

Appendix B: Electron Classes

Electron classes are defined to distinguish between well measured and "badly measured" electrons. The observables used to perform electron classification are:

- the "measured bremsstrahlung fraction" f_{brem}, defined as $f_{brem} = \frac{(p_{in}-p_{out})}{p_{in}}$ (p_{in} is the momentum of the electron reconstructed at the interaction vertex and p_{out} the momentum reconstructed at the last hit, this quantity is seen to be strongly correlated to the quantity of material traversed.);
- the Φ match between the reconstructed track and the supercluster total energy collected by the supercluster with the momentum measured at the track origin which is sensitive to the energy lost in the tracker material;
- the pattern of clusters in the superclusters;

Four electron classes are thus defined[1], the parameters for each class are given below:

- `golden electrons`, the class of the best measured electrons, it contains the electrons least affected by radiation emission, with a reconstructed track well matching the supercluster and a well behaved supercluster pattern. Golden electrons are the 20% of the total electron population. Definition parameters are:
  1. supercluster formed by a single cluster seed (no bremsstrahlung cluster);
  2. measured bremsstrahlung fraction $f_{brem} < 0.2$;
  3. supercluster position and track extrapolation from last point matching in Φ within ± 0.15 rad;
  4. a E_{SC}/p_{in} value above 0.9.

[1] The case where an electron is impacting in the immediate vicinity (less than about half the η width of a crystal) of ECAL inter-module borders are considered separately. In such cases, a significant fraction of the electron shower can leak behind the ECAL. The fiducial regions are defined to group such electrons with those impacting in the barrel-endcaps transition regions $1.4442 < |\eta_{SC}| < 1.560$. Both this groups of electrons are grouped in a *"boundary electrons"* class

- `big brem electrons`, this class contains electrons with a good matching between E_{SC} and p_{in}, a well behaved supercluster pattern, and no evidence of energy loss effects from secondary photon conversion despite a very large measured bremsstrahlung fraction[2]. The class is defined as:

  1. supercluster formed by a single cluster seed;
  2. measured bremsstrahlung fraction $f_{brem} > 0.5$;
  3. $0.9 < E_{SC}/p_{in} < 1.1$.

- `narrow electrons` have a significantly large bremsstrahlung fraction (but not as for "Big Brem" electrons), a well behaved supercluster, i.e. bremsstrahlung photons are merged inside the single cluster, and a relaxed track supercluster geometrical matching. It is defined as:

  1. supercluster formed by a single cluster seed;
  2. $0.9 < E_{SC}/p_{in} < 1.1$;
  3. a measured bremsstrahlung fraction and/or a Φ matching outside of the previous classes.

- `showering electrons` class contains electrons that failed to enter any of the previous classes. They are the 50% of the electron population.

It is possible to show that showering electron class distribution is directly correlated to the η distribution of the material thickness. On the contrary, golden electrons distribution is anti-correlated with the material thickness.

[2]Electrons emitting bremsstrahlung radiation in a single step, either at the beginning or at the end of the tracker silicon layer fall in this category.

Bibliography

[1] Lyndon Evans, Philip Bryant, "LHC Machine", JINST, Institute of Physics Publishing and SISSA (August 2008) 164 pp.

[2] CMS Collaboration, "The Compact Muon Solenoid - Technical Proposal", CERN/LHCC 94-38 (December 1994) 290pp.

[3] CMS Collaboration, "The Tracker Project - Technical Design Report", CERN/LHCC 98-6 (April 1998) 625pp.;
idem, "Addendum to the Tracker TDR", CERN/LHCC 2000-016 (Feb 2000) 82pp.

[4] CMS Collaboration, "The Electromagnetic Calorimeter Project", Technical Design Report, CERN/LHCC 97-23 (15 December 1998).

[5] Luca Mucibello, "Studio dell'identificazione di elettroni a CMS", Diploma thesis (August 2008) 95 pp.

[6] CMS Collaboration, "Search Strategy for a Standard Model Higgs Boson Decaying to Two W Bosons in the Fully Leptonic Final State", CMS AN-2008/39 (October 2008) 69 pp.

[7] The PYTHIA home page, available on the web, http://projects.hepforge.org/pythia6/

[8] The CMS Offline Work Book, available on the web, https://twiki.cern.ch/twiki/bin/view/CMS/WorkBook.

[9] Root manual, available on the web, http://root.cern.ch/.

[10] CMS Collaboration, "Electron reconstruction in CMS", CMS NOTE-2006/040, CMS NOTE-2006/040, (February 2006) 25 pp.