Facultad de Ciencias

MEJORA DEL B-TAGGING EN EL HLT PARA LA FASE-1 DEL DETECTOR DE PÍXELES

(Improvement of tracking for b-tagging at HLT by exploiting phase-1 pixel detector)

Trabajo de Fin de Grado
para acceder al

GRADO EN FÍSICA

Autor: Sara Ruiz Daza

Director: Rocío Vilar Cotabitarte

Co-Director: Silvio Donato

Octubre -2018

Agradecimientos

En primer lugar, agradecer a la Universidad de Cantabria por los años de formación. A las personas que apostaron por el Doble Grado en Física y Matemáticas y han hecho posible que haya llegado hasta aquí.

A la Universidad de Bergen, por descubrirme el mundo de la física de partículas. A la Universidad de Zurich, por darme la oportunidad de hacer unas prácticas e iniciar este trabajo. Al CERN, por hacerme pasar el mejor verano de mi vida. A toda la gente que conocí allí y que me devolvió la ilusión en la física.

En especial, al grupo de trabajo de partículas de la Universidad de Zurich por acogerme. A Florencia Canelli por hacerme sentir como en casa en Suiza. Y por supuesto, a Silvio Donato, el mejor supervisor que pude tener. Sin él, este trabajo no existiría. Thank you so much Silvio for all your effort, kindness, patience and teaching.

Al IFCA, a todo el grupo de partículas por orientarme tanto este último año. A su gente en el CERN por ser tan buenos guías en mi visita en febrero. No olvidaré nunca vuestra ilusión. Son muchas las personas que han aportado su granito de arena a este trabajo. En especial, gracias a Rocío Vilar por darme la oportunidad de desarrollarlo, por todo tu trabajo, motivación, apoyo y recibirme siempre con una sonrisa.

Al Intituto de Física Nuclear Polaco (IFJ PAN) por permitirme participar en su programa de verano. A mi supervisor allí, Andrzej Bozek, por mostrarme que la física abarca campos inimaginables. A todos los estudiantes con los que compartí la experiencia. He aprendido mucho con vosotros y os recordaré con cariño.

Finalmente, a mi familia, compañeros y amigos, simplemente por dejarme formar parte de vuestra vida, muchas gracias.

Resumen

El objetivo de esta tesis es entender y mejorar el rápido filtrado y guardado de datos en el experimento del *Solenoide de Muones Compacto* (CMS). En particular, el principal propósito de mi trabajo ha sido mejorar un algoritmo capaz de reconstruir el vértice primario tras una colisión para aumentar la eficiciencia del proceso de filtrado de datos.

La tesis comienza con una explicación teórica al Modelo Estándar. Después, se explica el dispositivo experimental utilizado, el experimento de CMS. En él, colisiones ocurren cada 25 ns, lo que supone 600 millones de colisiones por segundo. Dada la enorme cantidad de datos que este proceso implica, solo podemos guardar unos pocos. Se trata de almacenar aquellos que presentan las mejores características para ser estudiados. Para ello, el experimento tiene su propio sistema de filtrado, el *High Level Trigger* (HLT). Este sistema se encarga de correr diferentes algoritmos para filtrar los datos interensantes.

El HLT se encarga de una primera reconstrucción de eventos. De esta forma, en el segundo capítulo del trabajo se presentan los objetos físicos involucrados. El tercer capítulo introduce el CMS Trigger System.

Además, determinar con precisión el vértice primario (PV) ayuda a entender la física detrás de la colisión. El cuarto capítulo explica el algoritmo usado por el HLT para encontrar el PV y agilizar la reconstrucción de trazas. Tiene en cuenta los factores que contaminan la resolución del detector, como alta luminosidad y energía en las colisiones, pequeño espaciado entre colisiones y alto apilamiento. Este algoritmo es llamado *Fast Primary Vertex* (FPV)

El FPV usa los pixel clusters detectados por CMS. Mi trabajo comienza con una selección de esos pixel clusters de acuerdo a su coordenada ϕ, anchura, longitud, coordenada ρ y carga de los clusters y momento del jet. A continuación, los pixel clusters son proyectados sobre línea de haz de protones en la dirección η del jet. En la región con más proyecciones de clusters se encuentra el FPV.

Finalmente, el trabajo termina mostrando algunos de los resultados y conclusiones obtenidas al aplicar el FPV a datos generados por Monte Carlo. La eficiencia final obtenida ha sido del 82%. El lenguaje de programación usado ha sido *Python* y ROOT el software manejado en el análisis de datos.

Palabras clave: Solenoide Compacto de Muones (CMS), High Level Trigger (HLT), Fast Primary Vertex (FPV), Apilamiento.

Abstract

The aim of this thesis is to understand and improve the way that events are quickly filtered and saved in the *Compact Muon Solenoid (CMS)* experiment at the *Large Hadron Collider (LHC)*. Particularly, the main objective of my work has been to improve an algorithm to reconstruct the primary vertex of a collision to make better the trigger performance.

This thesis begins explaining the theoretical framework for the research performed, the Standard Model, then it explains the experimental setup used, the CMS experiment. The collisions happen every 25 ns, providing in total around 600 million of collisions per seconds at the beginning of each run. Given the large quantity of data, we can not store all of it. We try to keep only those events that are more similar to the particular phenomena we are interested to study. To do this, the experiment has a trigger system, the *High Level Trigger (HLT)*. This system runs different algorithms to select the interesting events.

A quick event reconstruction is executed at the HLT. The second chapter will present the physics objects involved in the event reconstruction. The third chapter will introduce the CMS Trigger System.

Furthermore, having an event with a precise primary vertex (PV) determination helps to understand the physics behind the collision. At the HLT the primary vertex is found using tracks originating from the primary vertex. The chapter four is dedicated to describe the algorithm used to find the PV at the HLT in order to speed up the tracking. It takes into account the factors that would pollute the resolution of the detector, like high luminosity and energy collisions, small bunch spacing and high pile-up. This algorithm is called *Fast Primary Vertex* (FPV).

The FPV algorithm uses the pixel clusters detected by CMS. My work begins with a selection of these pixel clusters according to their ϕ coordinate, cluster width and length, cluster ρ coordinate, cluster charge and jet momentum. Afterwards, the pixel clusters are projected back to the beam line, along the jet η direction of the jet. The FPV is enclosed in the region with more pixel clusters projected.

Finally the thesis ends with some results and conclusions obtained when the FPV is applied to $t\bar{t}$ events generated with Monte Carlo. The final efficiency obtained was of 82% with pile-up 35. The programming language was developed in *Python* and ROOT was the scientific software framework used in the analysis.

Keywords: Compact Muon Solenoid (CMS), High Level Trigger (HLT), Fast Primary Vertex (FPV), Pile-up.

Contents

Chapter 1

Introduction

From the beginning of time, several questions have worried our minds: What is beyond the universe? Where does mass come from? What is dark matter? Particle Physics tries to answer all these questions.

1.1 Standard Model

Particle Physics studies the basic constituents of matter and their interactions. From the electron discovered by Thomson in 1897 to the Higgs boson observed at the Large Hadron Collider (LHC) in 2012, the Standard Model has been developed starting from $\sim$ 1960.

The Standard Model [1] is the more fundamental theory in Particle Physics. It describes 24 fundamental particles (6 quarks, 6 leptons, 4 bosons and their antiparticles) described in 3 different generations according to their masses. They interact via 4 kind of bosons as shown in Figure 1.1.

There are 6 quarks: up (u), down (d), charm (c), strange (s), top (t) and bottom (b); and 6 leptons: electron (e), electron neutrino (ν_e), muon μ, muon neutrino (ν_μ), tau (τ) and tau neutrino (ν_τ). Each of these particles have their associated antiparticle. Quarks and leptons are fermions with half-integer spin and they obey Fermi–Dirac statistics and Pauli exclusion principle. While bosons have integer spin and follow Bose–Einstein statistics.

The four known fundamental forces in the universe are: the electromagnetic, weak, and strong interactions and the gravitational force. The last force can be neglected at the collider energy and it is very difficult to be included in the SM, although many theoretical efforts are aiming to do this. The strong interaction is mediated by the exchange of gluons (g). The quanta of the weak interaction are the W^+, W^- and Z bosons. Finally, photons act as the exchange

particle for the electromagnetic force.

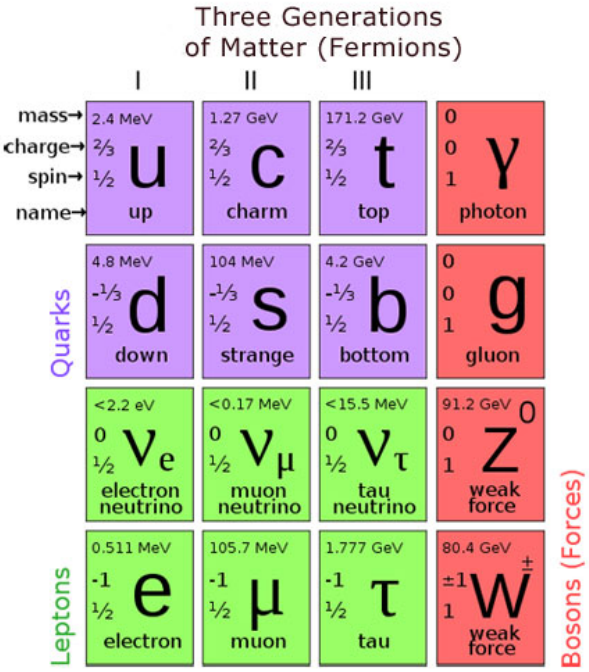


Figure 1.1: Standard Model particles.

1.2 The LHC accelerator and the CMS experiment

We are interested in studying collisions with high energy in order to know more about the interactions and particles that happened at the beginning of the universe.

For the collision we can use whatever is better for the experiments, and all of them have advantages and disadvantages. CERN decided to build an hadron collider in order to reach high energy and high luminosity, allowing to discover the Higgs boson and to search for new particles.

The Large Hadron Collider (LHC) collides proton beams at high energy. Experiments like CMS or ATLAS investigate these collisions. The following sections summarize the main features of the LHC and its experiments.

1.2.1 The LHC

The LHC (Large Hadron Collider) is a proton-proton (or heavy ion) collider, located at the European Organization for Nuclear Research (CERN, Switzerland). Originally, it was designed to produce collisions with a center of mass energy of $\sqrt{s} = 14\,\mathrm{TeV}$ [3] and luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Currently, is operating at the energy of $6.5\,\mathrm{TeV}$ per circulating beam, that means

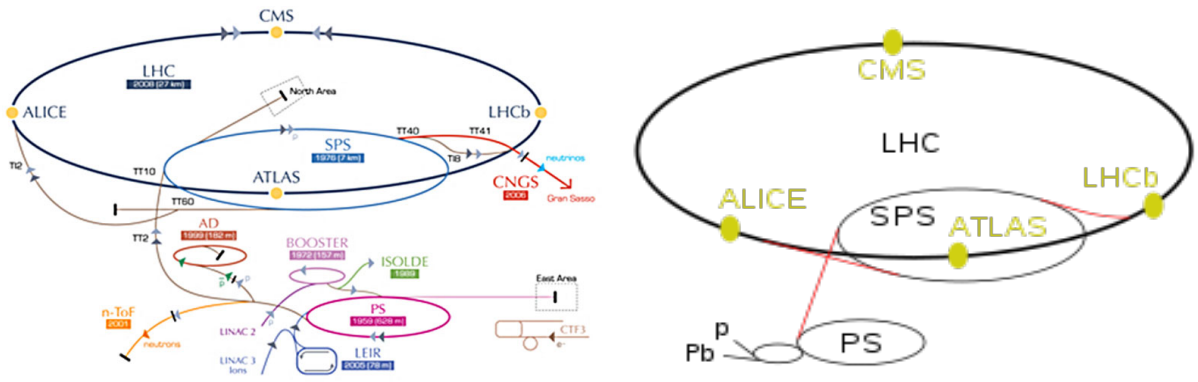
Parameter	Nominal value
Energy per proton	7 TeV
Design Luminosity	$10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$
Number of collisions per second	600 millions
Bunch spacing	25 ns
Dipole Field at 7 eV	8.33 T

Table 1.1: Nominal Beam Parameters at the LHC [2].

$\sqrt{s} = 13\,\mathrm{TeV}$. The instantaneous luminosity record was smashed, reaching $2.06 \cdot 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, i.e, about twice the nominal value.

The LHC accelerator is buried 100 metres underground on a 27 Km long ring tunnel close to Geneva. Protons follow an acceleration chain before to be injected onto the LHC. Foremost, they are obtained from hydrogen gas. Then, protons are injected into the first accelerator, Linac 2, it accelerates the protons to the energy of 50 MeV. Later, into the Proton Synchrotron Booster (PSB), where protons reach an energy of 1.4 GeV. The chain is followed by the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) as shown in Fig 1.2a. The SPS [4] provides accelerated protons with energy up to 450 GeV and velocity of 0.9998 times the speed of light. Once the protons are in the LHC, they reach a velocity of 99.9999964% the speed of light.

In order to accelerate protons and collimated in beams, 16 radiofrequency (RF) cavities are placed at the LHC. The RF are metallic chambers that contain an electromagnetic field. They are are grouped in fours cylindrical refrigerators called cryomodules (2 per beam). Superconducting magnets bend the particles' trajectory [5]. LHC parameters are given in Table 1.1.



(a) (b)

Figure 1.2: Acceleration chain of protons at CERN and the four main experiments at the LHC.

Accelerated protons run in opposite directions in separate beam pipes at ultra-high vacuum, and they collide in four points at the LHC. Four main detectors (ATLAS, CMS, LHCb and

ALICE) and "other minor experiments" study the physics behind these collisions as shown in Fig. 1.2b.

A Large Ion Collider Experiment (ALICE) is specialized in studying heavy-ions collisions [6] in order to recreate the plasma of particles generated in the first instances of the Big Bang. The Large Hadron Collider beauty (LHCb) [7] measures the parameters of CP violation in the interactions of b-hadrons [8] and their studies can help to understand the Matter-Antimatter asymmetry of the universe.

The Compact Muon Solenoid (CMS) [9] and A Toroidal ApparatuS (ATLAS) [10] are two general-purpose detectors. Differently than LHCb and ALICE, they can explore many sectors of particle physics (e.g., the Higgs boson). Thereby, CMS and ATLAS will be the ones set to discover new particles with high mass. And though they have similar physics goals, they were technically designed in a different way.

1.2.2 The CMS Experiment

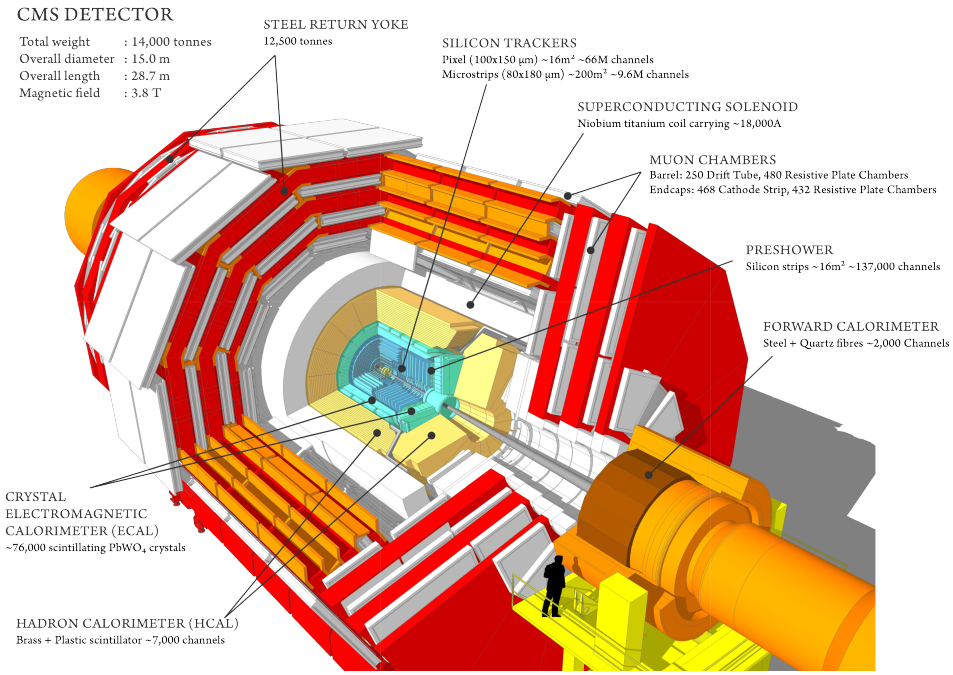


Figure 1.3: Schematic view of the CMS detector.

Once the LHC has been introduced and some of the main experiments carry on there mentioned, in this section more details about the CMS Experiment will be given. All the data used in this thesis was simulated with a Monte Carlo (MC) event generator, modeling real events measured by the CMS detector. Thereby, I have only used simulated data so far, although it can be used

with real data. Real data for the CMS experiment can be used freely by its members. IFCA (Institute of Physics of Cantabria) is a centre associated with the CMS collaboration.

The Compact Muon Solenoid (CMS) experiment was designed to study proton-proton collisions at the LHC. Over 4000 people compose the worldwide CMS Collaboration, including physicists, engineers, computer scientists, technicians and students. It is installed close to the French village of Cessy, about 100 meters underground.

The main feature of CMS is the large superconductive solenoid, 13 m length and 6 m in diameter, which generates a 3.8 T magnetic field. With 15 m tall, 22 m long and a weigh of 14000 tons is twice dense and compact than ATLAS, the other main general purpose detector at LHC.

The detector has a cylindrical structure. It is composed of successive layers as shown in Figure 1.3. Each of these layers is specialized to detect different particles and measure their properties. The CMS experiment is divided into a superconducting solenoid, the silicon trackers, the electromagnetic and the hadronic calorimeters and the muon system. Figure 1.4 represents the particle interactions with the CMS sub-detectors.

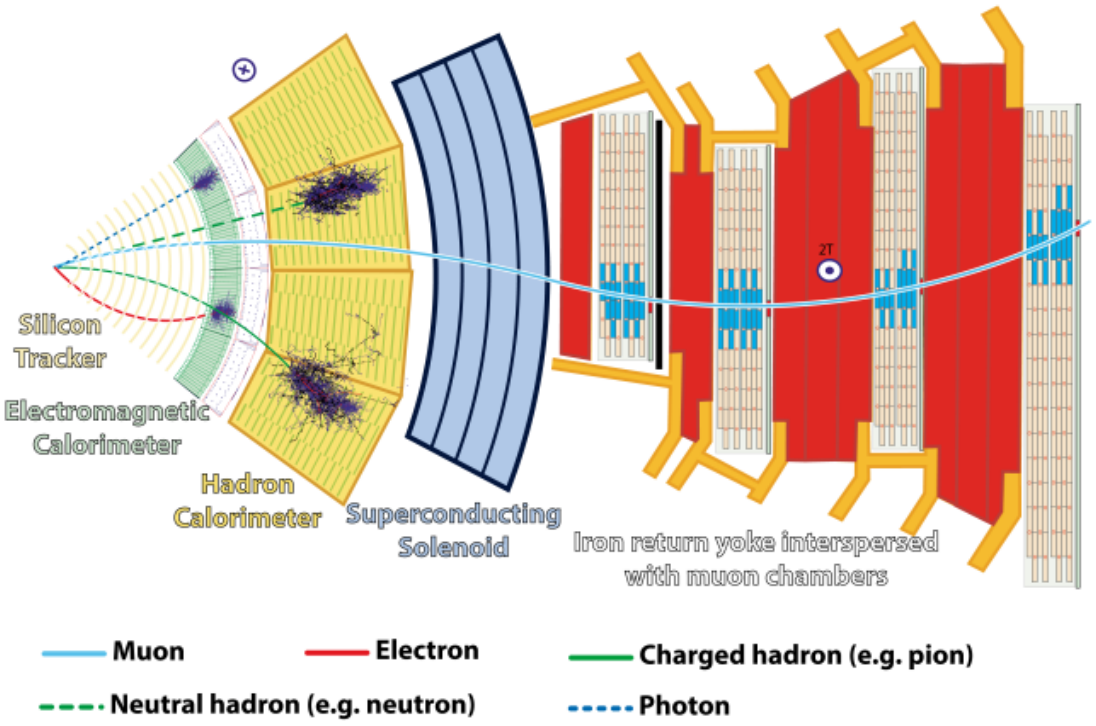


Figure 1.4: Slice showing CMS sub-detectors and how particles interact with them.

The coordinate system adopted by CMS has the origin centered at the nominal collision point. The x-axis points to the center of the LHC ring, the y-axis points vertically upward and the z-axis points along the direction of the beam toward the Jura mountains. The azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle θ is measured from the z-axis. The pseudo-rapidity describes the angle of a particle relative to the beam line and it is defined as:

$$\eta = -\ln(\tan(\theta/2)) \tag{1.1}$$

The reason why the pseudo-rapidity is preferred over the polar angle θ is because pseudo-rapidity differences are invariant with respect to Lorentz boosts along the z–axis for massless particles. The momentum component transverse to the beam direction is denoted by p_T and the transverse energy by E_T. Both of them are computed from the x- and y-components.

- **Superconducting solenoid**

  The superconducting solenoid provides a magnetic field of 3.8T (100.000 times stronger than the Earth's magnetic field). It bends the paths of charged particles to measure their momentum. The CMS magnet is cooled to 4K [11].

- **The silicon trackers**

  The charged tracker system records the positions of any charged particles like muons, electrons and charged hadrons. When these particles pass through the silicon trackers, a signal is induced. Afterwards, these signals will be reconstructed with a high efficiency. Silicons detectors are chosen due to their fast response and spatial resolution [12].

  The flux of particles changes as function of the distance to the beam line. Hence, different sub-detectors have been positioned as shown in Figure 1.5.

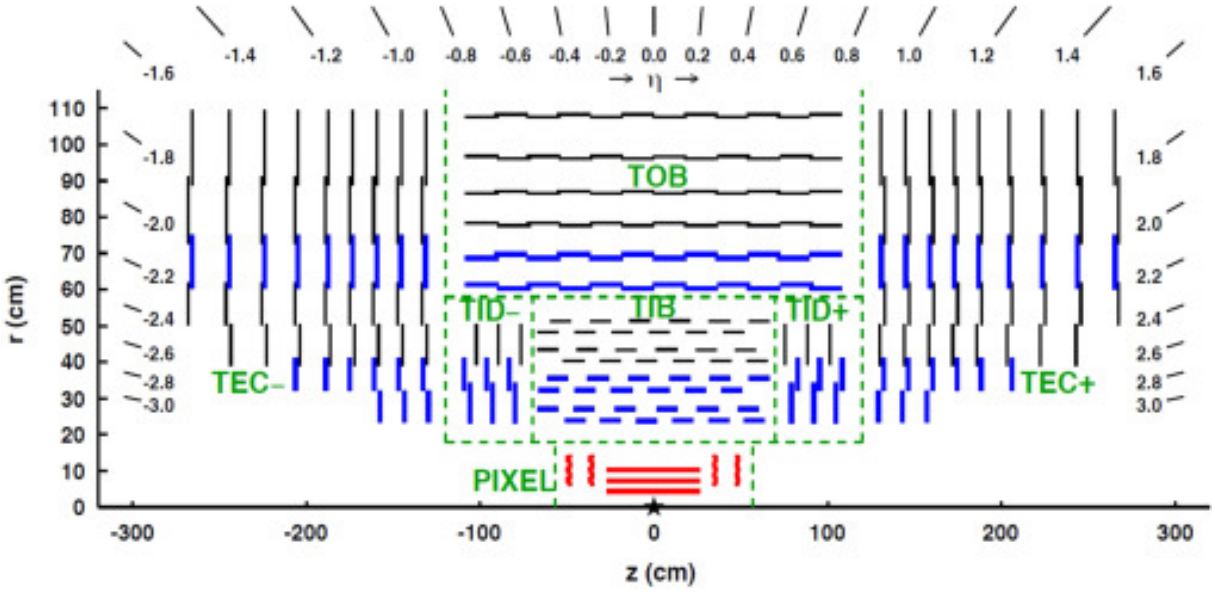


Figure 1.5: Tracker slice in the r-z plane.

- The pixels. The pixels are detectors placed at the innermost layer (30 < r < 160 mm). They are resistant to high radiation rates ($\sim 10^7$ particles $\mathrm{cm}^{-2}\mathrm{s}^{-1}$) and they have an almost square shape of 100×150 μm. The pixel detector was designed with three barrel layers and two end-cap disks. In 2016, the pixel detector have been upgraded with a new detector with 4 barrel and 3 endcap layers [13], as shown in Figure 1.6. That experiment upgrade improved the efficiency required by high luminosity and pile-up.

- The strip detectors. The strip detectors are located at the outermost region (20cm < r < 110cm) and made up of around 15200 modules. It is divided into substructures: TIB (Tracker Inner Barrel), TID (Tracker Inner Discs), TOB (Tracker Outer Barrel) and TEC (Tracker End Caps). The TIB (4 layers) and TOB (2 layers) composed the barrel tracker region and provide a stereo measurement in r-ϕ and r-z coordinates with stereo angle of 100 mrad. While the TID (3 disks) and TEC (9 disks) are the external endcaps.

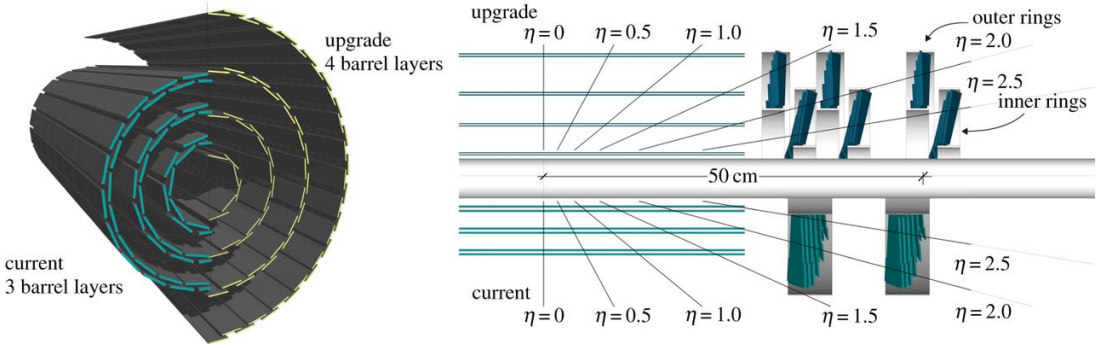


Figure 1.6: CMS pixel detector upgrade [14]

.

- **The electromagnetic calorimeter (ECAL)**

  The electromagnetic calorimeter (ECAL) measures the energy of electrons and photons. It is composed of exactly 75848 lead tungstate ($PbWO_4$) crystals. 61200 of them are placed in ECAL Barrel (EB), from 0 < |η| < 1.48 and 14648 in the ECAL Endcaps (EE), from 1.48 < |η| < 3 as shown in Figure 1.7 [15].

  The main feature of a lead tungstate crystal is its ability to emit light when electrons or photons pass through it. When these particles interact with the crystals, an electromagnetic shower is produced.

  $PbWO_4$ crystals have a short radiation length of 0.89 cm, a small Molière radiius of 2.2 cm and a fast response: 80% of the light is emitted within 25 ns [16].These features make the ECAL a compact calorimeter with good energy and angular resolution.

  In order to measure the light signal emitted by the crystals, photodetectors need to be placed. The EB employs silicon avalanche photodiodes (APDs). However, in the EE it is not possible the use of silicon photodiode because of the high radiation and the light is

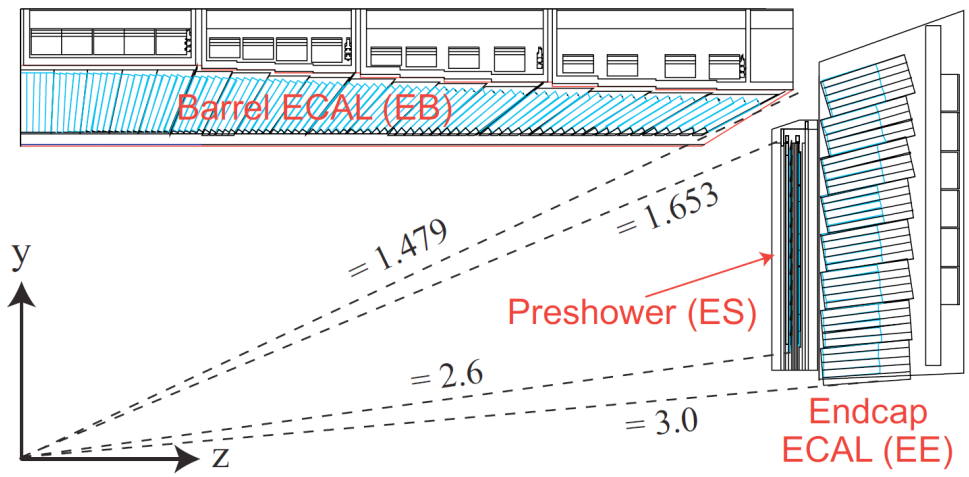


Figure 1.7: CMS Electromagnetic Calorimeter.

measured by vacuum phototriodes (VPTs). Finally, the silicon preshower detectors are in the internal face of the endcaps [17].

- **The hadron calorimeter (HCAL)**

  The hadron calorimeter (HCAL) measures the energy of hadrons. It covers a pseudorapidity region $|\eta| < 5$ and it is divided into four sub-detectors as shown in Figure 1.8. The innermost layers of the HCAL are the hadron barrel (HB) and the hadron endcap (HE). The hadronic outer calorimeter (HO) sourronds the HB and hadronic forward calorimeter (HF) is placed about 11.2 m from the collision point [18].

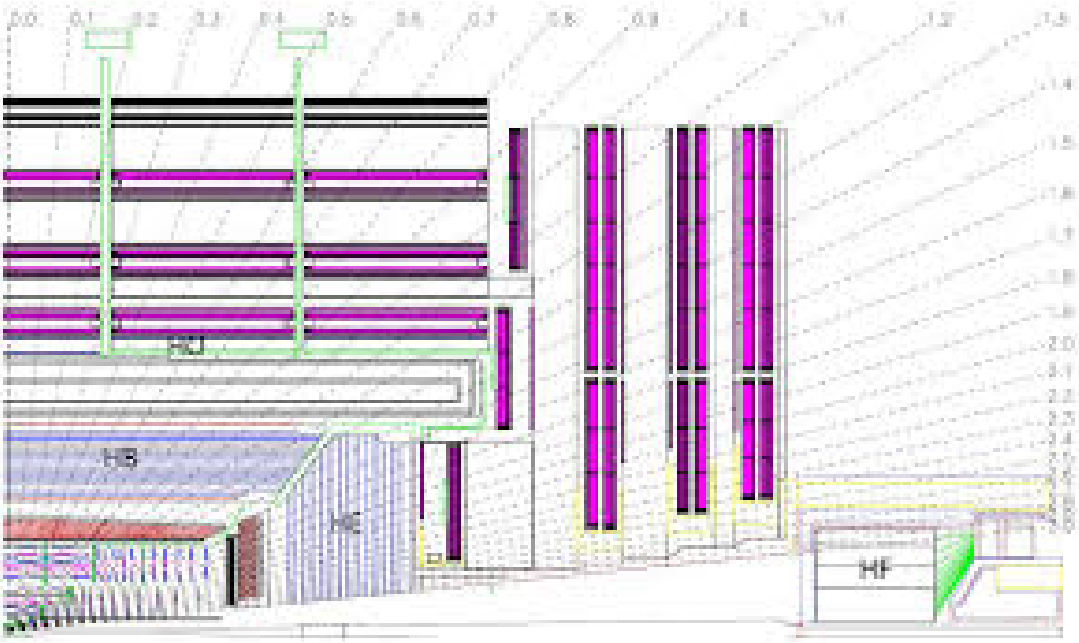


Figure 1.8: Layout of the CMS Hadron Calorimeter.

In HCAL, plastic scintillators have been chosen to produce the light pulses. Then, this visible-light signal is collected up by optic fibres and amplify by photodetectors [19].

The probability that an hadron goes through the whole CMS detector without being detected is approximately 10^{-9}.

- **The muon chambers**

  The moun chambers surround all the detectors introduced above. Muons are charged particles able to cross the whole detector because of their small interaction with the material. Neutrinos can also escape from the CMS detectors, though their presence can be inferred from the missing transverse energy.

  There are 1400 muon chambers, divided into three kinds of detectors: 250 drift tubes (DTs), 540 cathode strip chambers (CSCs) and 610 resistive plate chambers (RPCs) [20]. The structure is shown in 1.9.

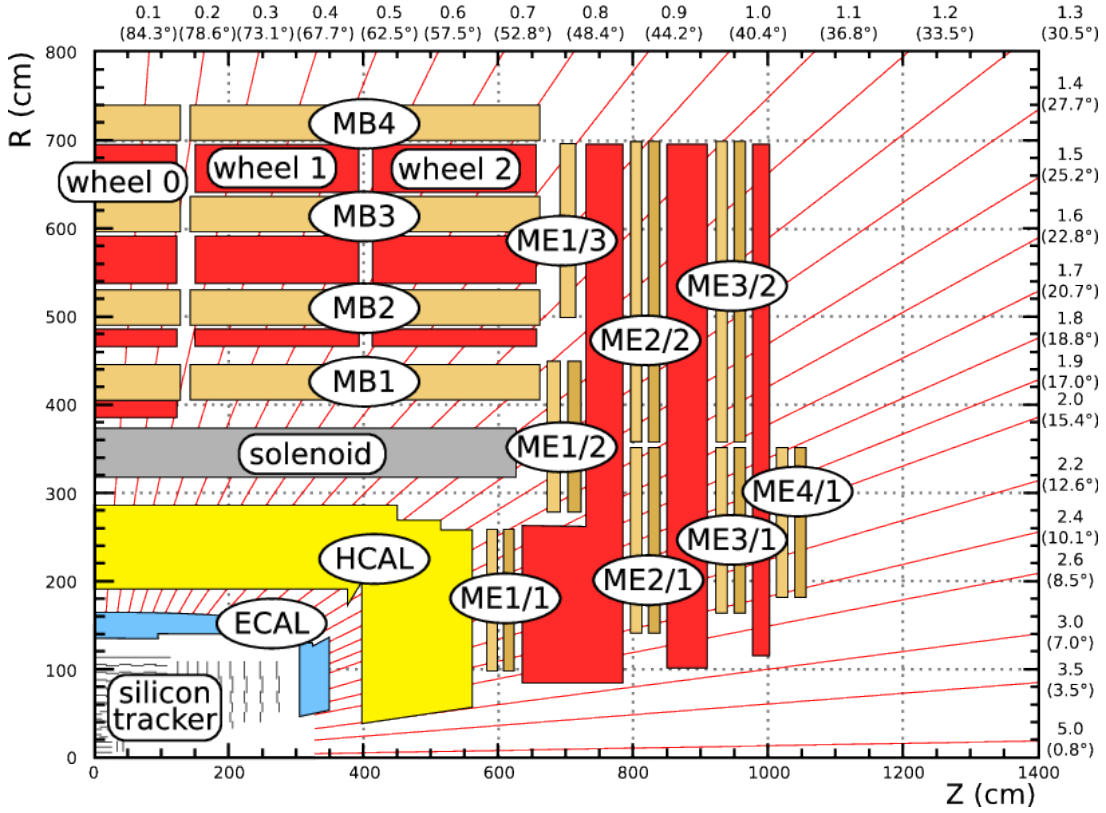


Figure 1.9: The CMS muon chambers.

Chapter 2

Physics objects and event reconstruction

After a brief introduction to CMS, this new chapter describes some of the physics objects involved in the event reconstruction. An event in nuclear and particle physics is defined as the data produced in a collision between proton bunches.

This chapter begins with an introduction to the pile-up effect due to its high importance in the event reconstruction. Afterwards, clusters, jets and tracks will be introduced. A detailed description of the primary vertex and its reconstruction will be given as well.

2.1 Pile-up

The pile-up is defined as the number of proton-proton collisions per bunch crossing. Pile-up increased in the last years in order to have more collisions to observe rare process. For instance, in Run 1 (2009-2013) it was around 20, while in Run 2 (2015-2018) the CMS reached a pile-up of 60.

In this thesis, all the data used was simulated with pile-up 35. In other words, an average of 35 proton-proton collisions per event happen.

2.2 Clusters and tracks reconstruction

Clusters are the footprints (individual points) left by the charged particles on the silicon trackers (pixels and strip detectors). Tracks are the trajectories of charged particles. Tracks are used to

find the primary vertex.

The reconstruction process of tracks is done in several interactions and is called tracking. The main idea of tracking is to associate sets of clusters to a determinate charged particle to fit their trajectories and measure their properties (e.g., p_T, ϕ, η) [21]. The tracking time is highly influenced by the pile-up conditions which increase the number of tracks significantly.

The track finder algorithm begins with the easiest tracks to be reconstructed. For instance,the first interaction reconstructs the tracks with large p_T or close to the beam spot. Afterwards, their associated clusters are removed of the following iterations. Therefore, the combinatorial complexity is reduced for each step. Once the easiest tracks are reconstructed, the track finder algorithm continue looking for tracks with low p_T or larger impact parameter.

2.3 Jets

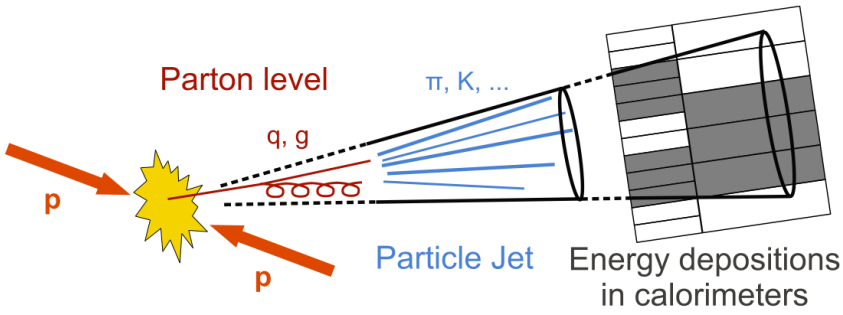


Figure 2.1: pp-collision and resulting collimated set of particles, a jet [22].

A jet is a collimated set of particles produced by the hadronization of a quark or gluon. Jets are generated at the LHC as illustrated in Figure 2.1.

Based on the information of the sub-detectors, jets are reconstructed at CMS with two different methods: the calorimeter-based and the Particle Flow (PF). More details about both methods can be found elsewhere [23].

2.4 Primary vertex

The objective of this thesis is to show an algorithm for the HLT able to reconstruct the primary vertex in short time and with a high efficiency. A definition of primary vertex will be given with a description of the main algorithm used at CMS to find its position. In Chapter 4 a new algorithm

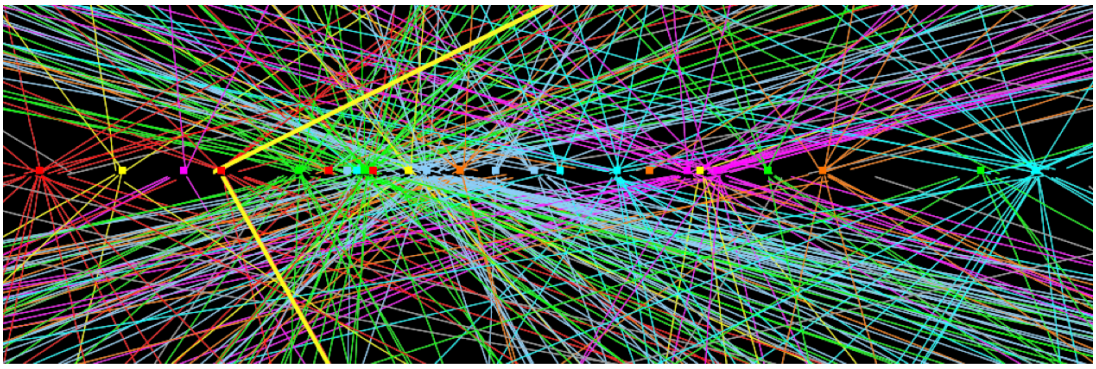


Figure 2.2: Event with several vertices due to a high pile-up [24].

.

performed will be shown.

The primary vertex is defined as the hardest proton-proton collision point in the beam-line. After a proton-proton collision, other pp collisions occur (pile-up). Thereby, for each proton-proton bunch crossing, several vertices are detected as shown in Figure 2.2.

The main algorithm used at CMS to find the primary vertex is called the *Adaptive Vertex Fitter* (AVF). It is based on a combinatorial Kalman fitter approach [25]. Mathematically is equivalent to the global least-squares minimization.

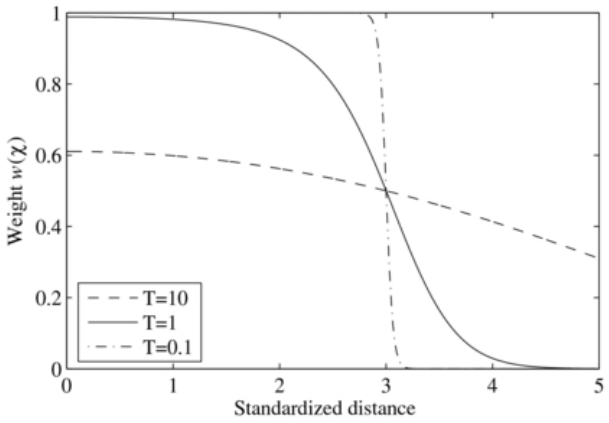


Figure 2.3: Weight-function used in the Adaptive Vertex Fitter with $\chi^2_{\text{cutoff}} = 9$ for T = 0.1, 1 and 10 [26].

The AVF is an iterative re-weighted fit which weights the contribution of each tracks according to their standardized (χ^2_i) to the vertex. The weight-function is:

$$\omega(\chi^2_i) = \frac{1}{1 + \frac{\chi^2_{cutoff} - \chi^2_i}{2T}} \tag{2.1}$$

where χ^2_{cutoff} is the distance where (χ^2) drops to 0.5. T is the global temperature and it is

positive. Figure 2.3 shows how the sharpness of the function changes as function of T.

The AVF starts from a collection of tracks (with $\omega(\chi_i^2)$) and a pre-defined value of T. At each iteration, the temperature is lowered and the track weights are updated using the equation 2.1. The candidate vertex position is also updated. When the iterative fit converges, one vertex have been reconstructed.

2.5 The b tagging

b tagging is the technique used to identify jets originated from the hadronization of bottom quarks. b-jets have a long life-time, high mass and large transverse momentum ($p \sim 50\,\mathrm{GeV}$). B hadrons contained in b quarks decay in secondary vertices, displaced from the primary vertex of $\sim 3000\ \mu$m. Figure 2.4 shows the presence of a secondary vertex originated from a b-hadron decay.

Secondary vertex can be directly reconstructed using the adaptive vertex fitter. It allows to identify b jets. In case of no secondary vertex are reconstructed, CMS uses several algorithms relying on the impact parameter, defined in Figure 2.4. The simplest are: the track counting algorithm and the jet probability algorithm. More details about both algorithms can be found elsewhere [27].

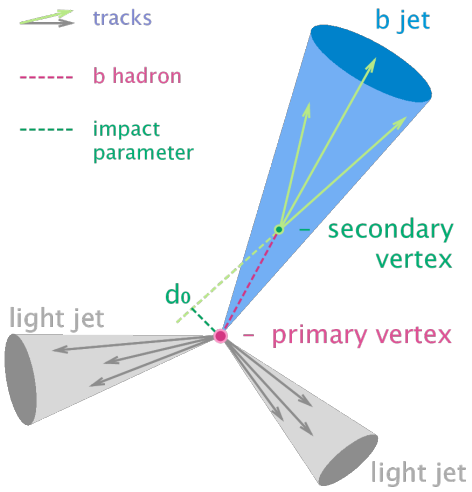


Figure 2.4: Sketch of the cascade decay of the b-hadron with the impact parameter definition. [28].

Chapter 3

CMS Trigger System and Data Acquisition

After the CMS detector has been introduced and all its layers described, it is necessary to explain what happens with all the data collected (or produced). In this way, the first part of the chapter introduces the CMS Trigger System and its two levels: Level-1 Trigger (L1) and High Level Trigger (HLT). These two levels provide a high selectivity of events for later off-line analysis. The second part of the chapter will describe the event reconstruction at HLT.

This chapter is really important due to the fact that all the research and algorithm improvement presented have been designed to be used and increase the efficiency of the HLT.

3.1 General Description

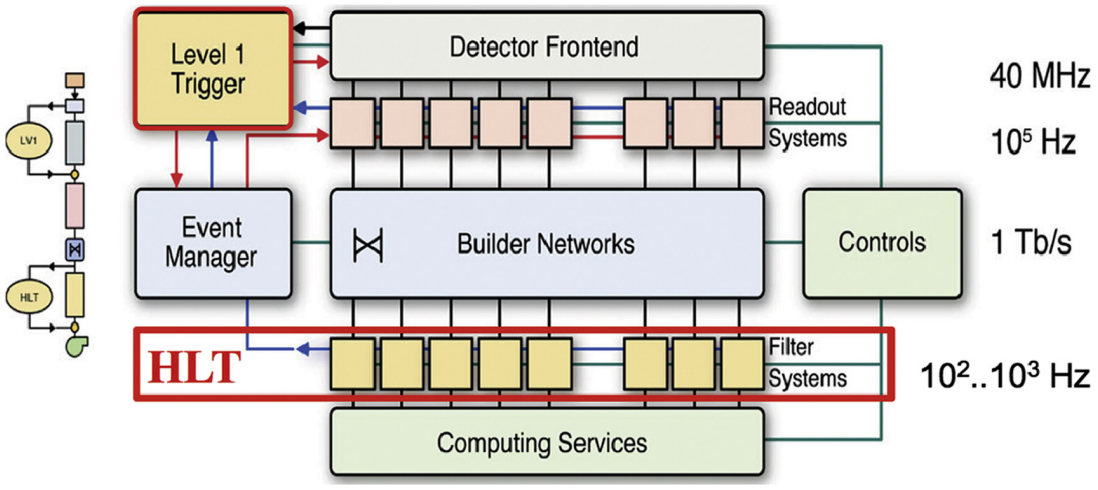


Figure 3.1: The CMS Data Acquisition system.

Due to the high number of collisions ($\sim 10^9\,\mathrm{interactions/s}$) and the fact that not all of them can be stored, the CMS experiment needs a trigger system to select, in a short amount of time, the potentially interesting events for further studying. The proton beams are organized in bunches spaced by 25 ns, hence a fast response is needed in order to reduce the rate of events to be written to disk.

The CMS trigger system is performed in two steps: the Level-1 Trigger (L1) and the High Level Trigger (HLT) [29]. This complex challenge of filtering data is shown in Figure 3.1.

3.1.1 Level-1 trigger (L1)

The Level-1 system (L1) is the hardware trigger in charge of the first event selection. The L1 trigger reduces the rate of events accepted from 40 MHz down to $\sim$ 100 kHz [30]. The decision time trigger each event is less than $1\,\mu$s . The L1 trigger process envolves information from the hadron and electromagnetic calorimeters and muon chambers as shown in Figure 3.2.

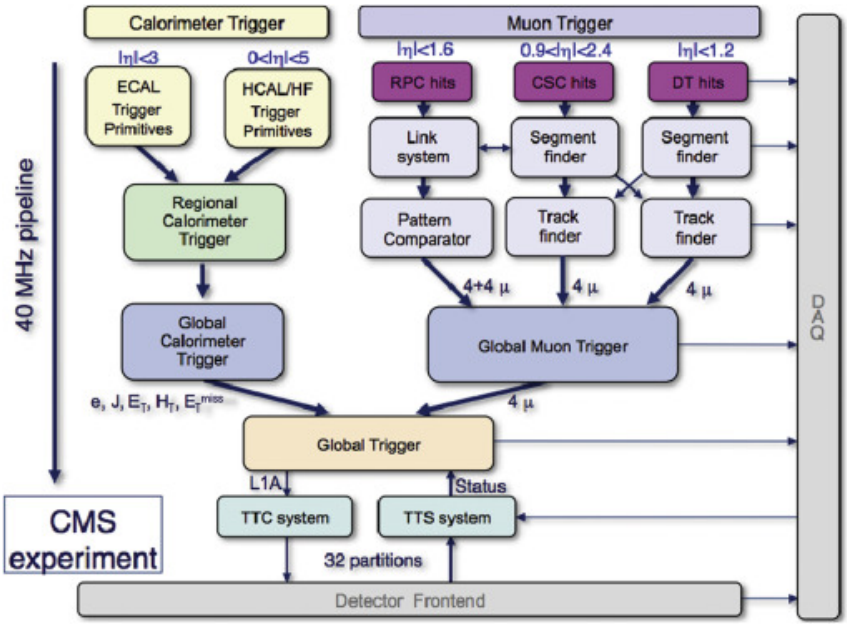


Figure 3.2: The CMS Level-1 Trigger.

On the one hand, data from the Electromagnetic Calorimeter (ECAL), Hadron Calorimeter (HCAL) and the Hadronic Forward Calorimeter (HF) are combined and processed in the Regional Calorimeter Trigger (RCT) and Global Calorimeter Trigger (GCT). At the end of these steps, events with energetic electrons, photons, jets or large missing transverse energy $(E_T{}^{miss})$ or transverse momenta (H_T) are accepted, according to criteria defined in the L1 trigger menu.

On the other hand, data from the muon detectors (resistive plate chambers (RPC), cathode strip chambers (CSC) and drift tubes (DT)) are combined and processed in the Global Muon

Trigger (GMT). The GMT choses the best muon candidates.

Then, the GCT and the GMT information is combined in the Global Trigger (GT). Finally, the Trigger Control System (TCS) takes the final decision of L1 trigger to keep or throw away the event. The Trigger, Timing and Control (TTC) sends out the information. Nevertheless, that amount of data is still high and it needs to be reduced further by the High Level Trigger.

3.1.2 High Level Trigger (HLT)

The High Level Trigger (HLT) is the software trigger in charge of reducing the event rate from 100 kHz to $\sim$1000 Hz, i.e. decrease the output rate by a factor of 100. Thereby, at the end of the trigger process only 1000 events/s are stored, the remaining events are rejected and will never be studied. Eventually, the Event Filter Farm (EVF) assembles the pieces of information from events fragments and builds up complete events.

It was redesigned and updated to lead with higher pile-up, higher center-of-mass energy, higher luminosity and smaller brunch spacing for Run 2, in 2015 [31]. All these parameters were almost doubled and the HLT and L1 had to be adapted [32].

3.2 Event reconstruction at High Level Trigger (HLT)

In the previous chapter, some information about the objects reconstruction was shown, in particular the Adaptative Vertex Fitter algorithm to reconstruct the primary vertex.

In order to reduce the tracking time at the HLT, the Fast Primary Vertex algorithm is implemented. We use the FPV to reconstruct only the tracks originating from the primary vertex. As a result, we reconstruct less tracks and the tracking is sped up.

However, the final primary vertex used in b-tagging is found using all tracks, with a resolution of $\sim$0.010 mm. The FPV resolution would be worse ($\sim$ 3 mm).

The main idea of the Fast Primary Vertex algorithm is that all tracks contained in a jet have a direction close to it. That means that the primary vertex can be found projecting the pixel clusters, associated to the jet, to the beam line with η-direction of the jet. The way to select and project jets, and how to extract the position of the FPV performed in this thesis, will be described in detail in Chapter 4.

Chapter 4

Fast Primary Vertex

In this chapter, I will explain the algorithm developed to find the primary vertex quickly at the HLT in order to reduce the tracking time. The $t\bar{t}$ events studied have been simulated using Monte Carlo (MC) event generators. The pile-up simulated was 35, i.e. 35 proton-proton collisions per event happen. In total, 9000 events have been reproduced.

As it was mentioned in section 3.2, the main idea of the Fast Primary Vertex is to assume that tracks contained in a jet have approximately the same direction of the jet. The process followed until the localization of the Fast Primary Vertex will be described in details in this Chapter.

4.1 Jets and clusters selection

First of all, we only consider energetic jets. The reason is that these jets contain a larger amount of tracks originated from the primary vertex. If we include soft tracks, we would consider a lot of hits from pile-up including very few hits from the primary vertex (the number of tracks and hits is roughly proportional to the jet p_T.Thereby, the transverse jet momentum acceptance was limited to $p_T > 40$ GeV.

Furthermore, given the pixel detector structure: pixel barrel and pixel end-cap (with different orientation of the pixel sensors), we have to distinguish between them. We will accept jets with $|\eta| < 1.8$ for the barrel and $|\eta| < 2.6$ for the barrel+end-cap. These given parameters are illustrated in Figure 1.6.

For each jet, we consider hits with $\Delta\phi$(jet, cluster) < 0.2 as compatible to it. This threshold allows to reduce the contamination of hits from pile-up, retaining a large fraction of hits associated to the jet tracks.

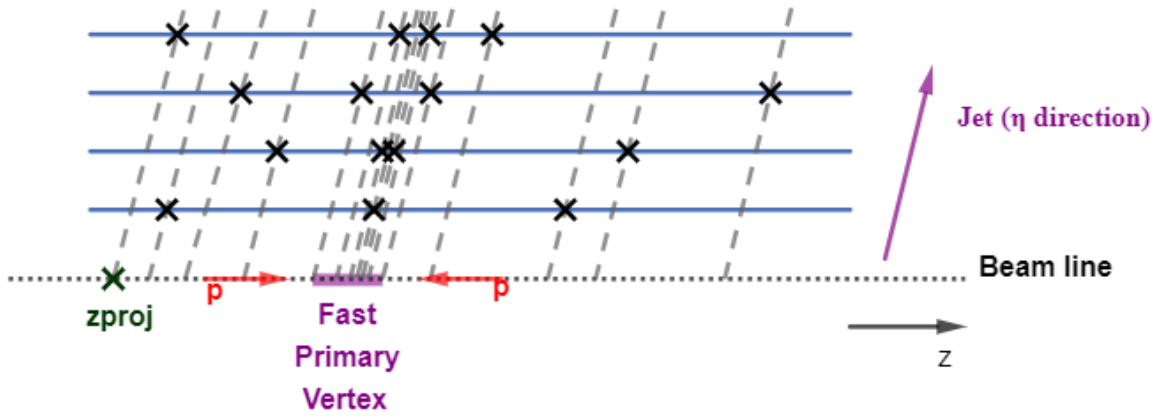
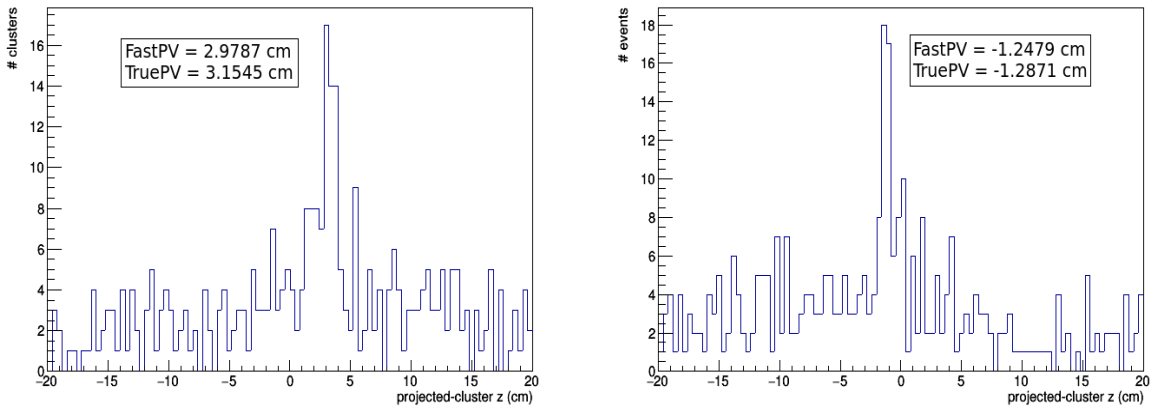


Figure 4.1: Sketch of the Fast Primary Vertex Algorithm. Clusters in the pixel barrels are projected to the beam line along the jet η direction. The region with the highest concentration of projected pixel clusters encloses the primary vertex.

If we project directly the clusters along z using the jet η-direction and study the peak in the z projection distribution, we expect a large number of projected hits close to the primary vertex. The sketch in Figure 4.1 illustrates this idea. Figure 4.2 shows the results obtained for two random events: the Fast Primary Vertex is defined as the peak in that distribution.



(a) (b)

Figure 4.2: z distribution of the projected clusters for two simulated $t\bar{t}$ events with center of mass energy of by $\sqrt{s}$ = 13 TeV and pile-up 35.

Now, we want to know how much efficient is this algorithm. Figure 4.3 shows the histogram of the vertex error distribution, defined as the difference of the Fast Primary Vertex (FastPV) and the MC-truth position (TruePV). We start defining efficiency as the fraction of events having |FastPV - TruePV|<1.5 cm (the reason why 1.5 cm is chosen as cut off will be explained in section 4.6. With this definition, an efficiency of 55 % has been obtained.

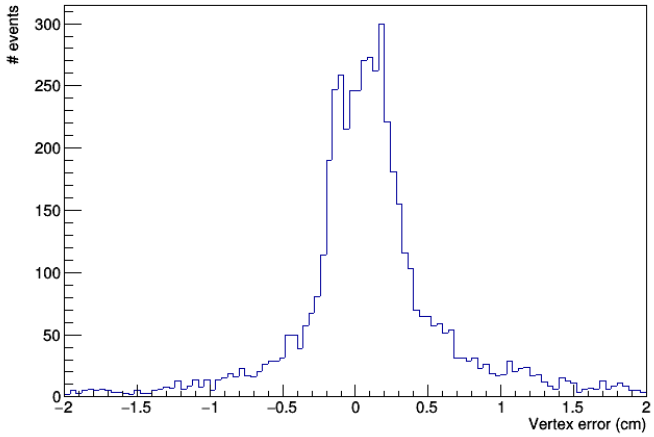


Figure 4.3: Difference between the Fast Primary Vertex and the True Primary Vertex. The plot has been produced using 9000 events simulation with pile-up 35 and $\sqrt{s}$ = 13 TeV.

4.2 Cluster weight

In order to improve the algorithm efficiency, we propose a new version by weighting the clusters depending on [33]:

- Jet momentum;
- $\Delta\phi$(jet, cluster) coordinate;
- Cluster size X and size Y (cluster width and length);
- Cluster ρ coordinate;
- Cluster charge

This section will explain the way these weights have been computed.

First of all, we need to distinguish between pixel clusters with a z projection close ("signal") or far ("background") from the true primary vertex. We will consider a cluster as signal if its z-projection is close (< 0.2 cm) to the real primary vertex. And a cluster belongs to the background if its distance to the real primary vertex is larger than 2 cm. From the distributions (signal and background) of the variables proposed to weight, we can obtain the ratio of signal and background. This parameter will give an idea of the likelihood of the clusters of belonging to the jet tracks. Thus, we can discuss the kind of weight to be used.

4.2.1 Jet momentum

A larger number of signal clusters will be recorded in high momentum jets and they will have a better position resolution with low pile-up hit contamination. This idea can be observed in

Figure 4.4 where the signal dominates at higher momentum.

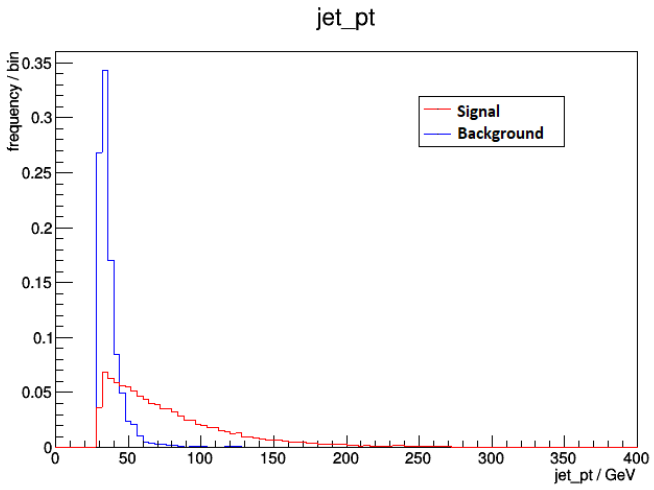


Figure 4.4: $\mathrm{Jet_{pt}}$ distribution for the signal and background.

From the ratio between the signal and background shown in Figure 4.5, we can construct a liner weight with the following structure:

$$\mathrm{W_{pt}} = \mathrm{jet_{pt}} \cdot \mathrm{pt_{WeightingSlope}} + \mathrm{pt_{WeightingOffset}} \tag{4.1}$$

where $\mathrm{jet_{pt}}$ depends on each jet, the $\mathrm{pt_{WeightingSlope}} = (0.18 \pm 0.03)\ \mathrm{GeV}^{-1}$ and $\mathrm{pt_{WeightingOffset}} = -6 \pm$ obtained from the fit in the range [0-80] GeV.

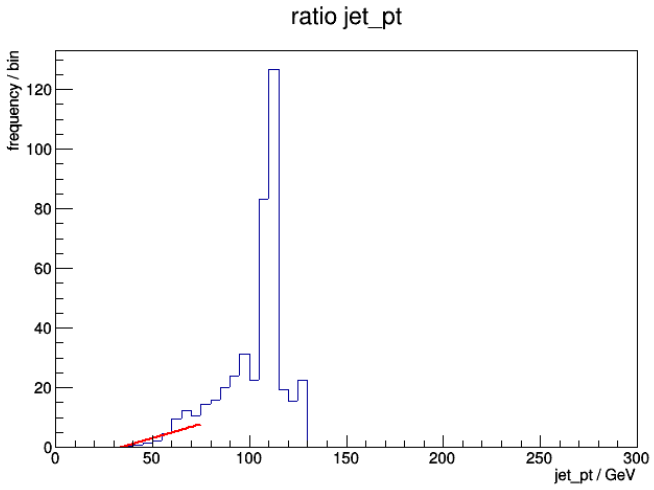


Figure 4.5: Ratio between the signal and background for the $\mathrm{Jet_{pt}}$ distribution.

4.2.2 ϕ coordinate

About ϕ weight, the idea is not only to consider pixel clusters satisfying $\Delta\phi$(jet, cluster) < 0.2 cm, but also, weight them. First of all, Figure 4.6 shows the signal and background distribution. In ad-

dition, we have to take into account the geometric difference between the barrels and the end-caps.

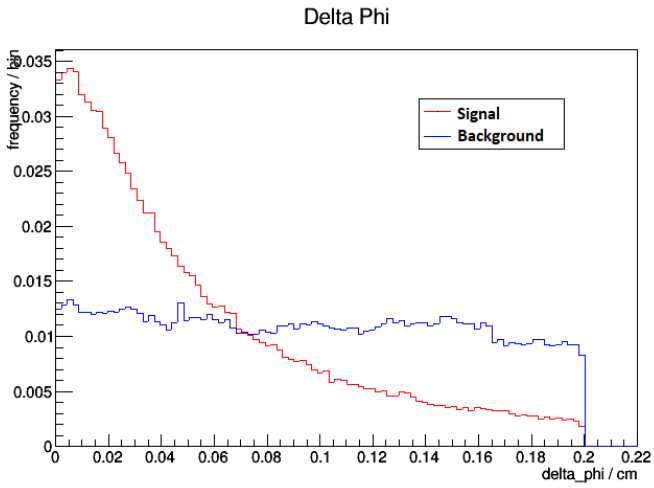
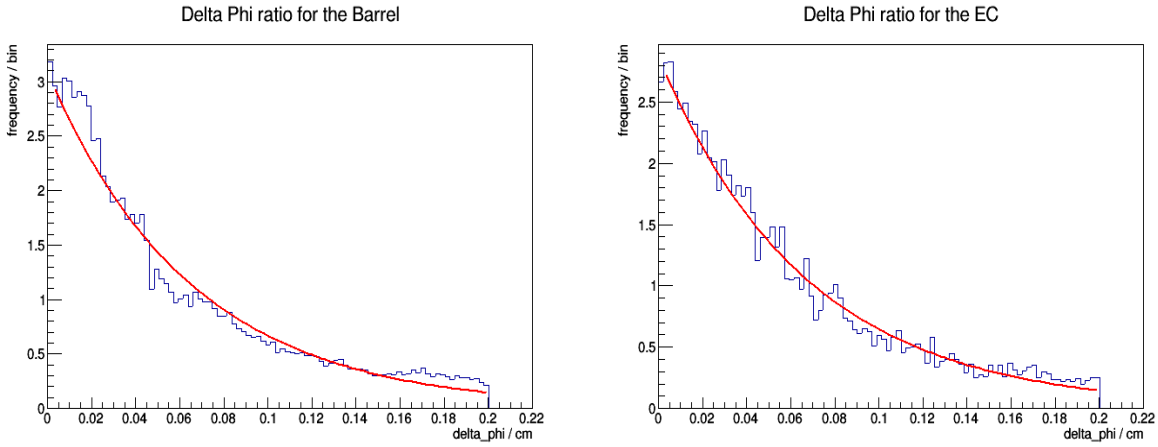


Figure 4.6: $\Delta\phi$(jet, cluster) distribution for the signal and background.

Figure 4.7 shows the ratio of signal and background for the barrel and the end-caps. From these plots we can extract the following exponential weights:

$$W_\phi = e^{\frac{-\Delta\phi}{m_{weight\Delta\phi}+k}} \tag{4.2}$$

where $\Delta\phi$ = $\Delta\phi$(jet,cluster) hence it depends on each jet and hit, $k = 0.17$, $m_{weight\Delta\phi,Barrel} = -1.53 \pm 0.09$ and $m_{weight\Delta\phi,EC} = -1.49 \pm 0.18$.



(a) (b)

Figure 4.7: Ratio between the signal and background for the $\Delta\phi$ distribution in the barrel and end-cap.

4.2.3 Cluster ρ coordinate

Figure 4.8 shows the distribution of the ρ-cluster coordinate. Again the idea is that: layers closest to the beam line (and to the interaction point) have the highest resolution. Thereby, pixel clusters with the highest ρ coordinate will have associated a low weight. Mathematically, we will use the following weight:

$$\mathrm{W}_{\rho} = 1 - \frac{\mathrm{Cluster}_{\rho}}{\mathrm{m}_{\mathrm{W}_{\rho}\mathrm{up}}} \tag{4.3}$$

where $\mathrm{Cluster}_{\rho}$ depends on each cluster and $\mathrm{m}_{\mathrm{W}_{\rho}\mathrm{up}} = 19.8$ (obtained empirically, probing different values and finding the one giving the best results).

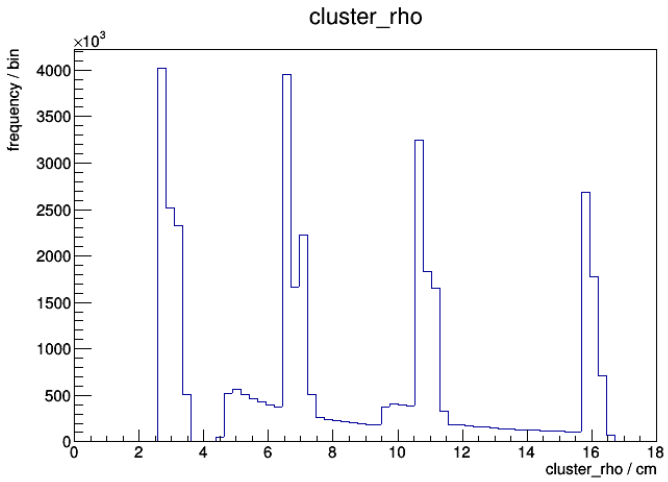


Figure 4.8: Cluster ρ distribution. Each of the four peaks observed correspond to what was detected by the 4 barrel layers.

4.2.4 Cluster width (size X)

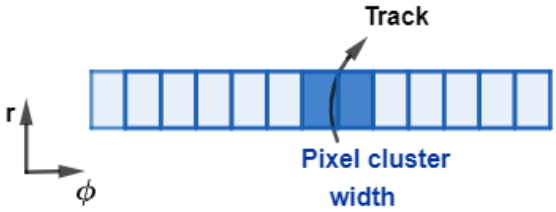


Figure 4.9: Shape of a pixel cluster along the r - ϕ projection.

The pixel cluster size X or cluster width is predominantly due to charge-sharing. It is defined as the number of pixels on the transverse plane (r - ϕ) illustrated in Figure 4.9.

Figure 4.10 shows the signal and background distribution for the cluster size x. We can directly construct a discrete width weight. We will increase the weight if sizeX is equal 2.

Thereby, we can associate the following values: If sizeX = 2, then $W_{sizeX} = 3$. Otherwise, $W_{sizeX} = 1$.

The reason why I chose it is due to the fact that the total weight will be a product of all the individual weighs. Introducing a factor of 1, will not produce any changes. Factor 3 was chosen in such a way that is proportional to factor 1 in Figure 4.10.

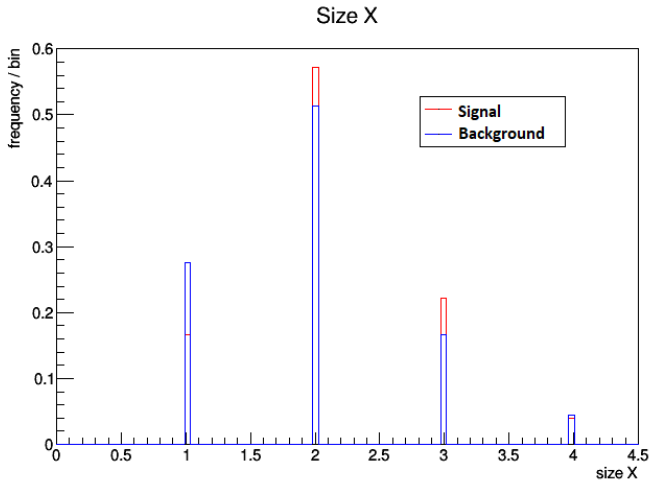


Figure 4.10: Size X distribution for the signal and background.

4.2.5 Cluster length (size Y)

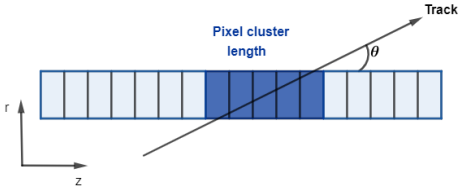


Figure 4.11: Shape of a pixel cluster along the r-z projection.

The pixel cluster size Y depends on the track impact angle. This weight has a slightly different idea to the others. We will study the difference between the cluster length and the expected one. The cluster length expected will be the mean number of pixels along the η direction. Figure 4.11 illustrates this idea. Therefore, the cluster length can be written as:

$$\text{cluster length} = \text{Jet}_{\eta}\text{Expected} = 1 + \frac{\text{pixel thickness}/\tan\theta}{\text{pixel length}} = 1 + \frac{1.9}{\tan\theta} \qquad (4.4)$$

where the pixel sensor thickness is 285 μm and the pixel sensor length is 150 μm.

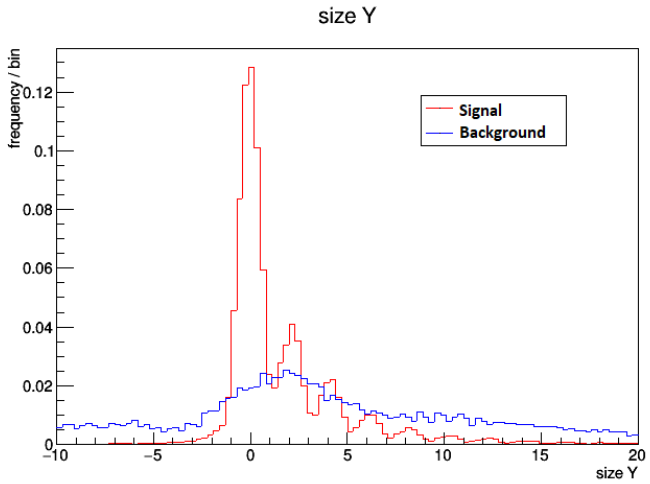


Figure 4.12: Size Y distribution for the signal and background.

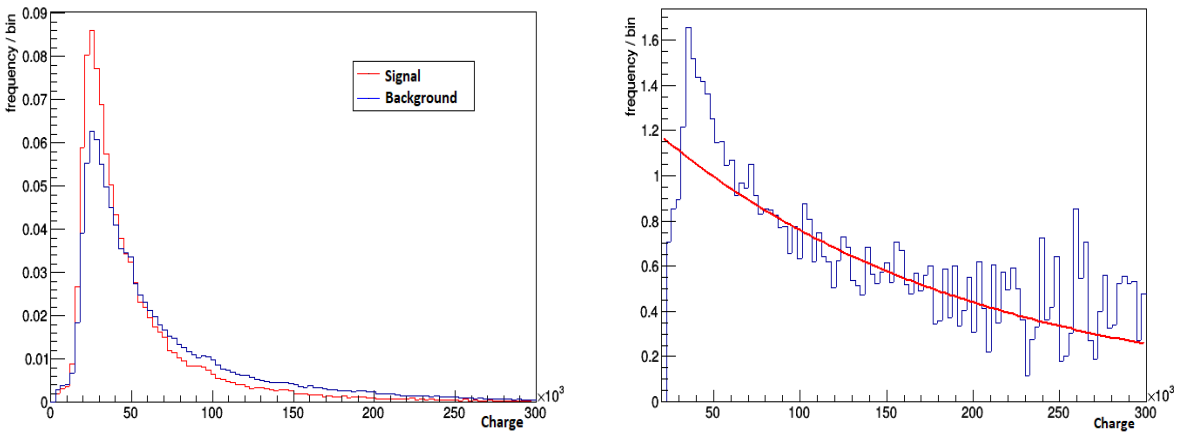
Again, Figure 4.12 gives us an idea about the signal and background distribution. We can define a discrete weight as in the case of cluster size X. If $|\mathrm{sizeY} - \mathrm{Jet}_{\eta}\mathrm{Expected}| < 0.5$, then $\mathrm{W}_{\mathrm{sizeY}} = 3$. Otherwise, $\mathrm{W}_{\mathrm{sizeY}} = 1$.

4.2.6 Cluster charge

About the cluster charge, comparing again the ratio of the signal and background plots in Figure 4.13, we got an exponential weight over the expectation value applied to the interval [10000,50000] of charge of equivalent electrons measured:

$$\mathrm{W}_{\mathrm{charge}} = \mathrm{e}^{-\mathrm{m}_{\mathrm{weightCharge}} * \mathrm{clusterCharge} + \mathrm{k}} \tag{4.5}$$

where $\mathrm{m}_{\mathrm{weightCharge}} = -(4.6 \pm 0.3) \cdot 10^{-5}$, $\mathrm{clusterCharge}$ depends on each cluster and k=0.167.



(a) (b)

Figure 4.13: (a) Signal and background distributions for the cluster charge. (b) Ratio between the signal and background distributions for the cluster charge.

4.2.7 Total weight

As a combination of the previous weights, a total cluster weight W_{total} is defined as:

$$\mathrm{W_{total}} = \mathrm{W}_{\phi} \cdot \mathrm{W_{sizeY}} \cdot \mathrm{W}_{\rho} \cdot \mathrm{W_{charge}} \cdot \mathrm{W_{sizeX}} \cdot \mathrm{W_{sizeX}} \cdot \mathrm{W_{pt}} \tag{4.6}$$

4.3 Peak-finder algorithm

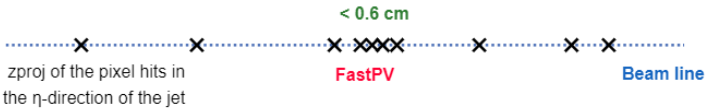


Figure 4.14: Sketch of the peak-finder algorithm. Clusters are projected back to the beam line along the jet η direction. The region with the highest weights and concentration of projected pixel clusters encloses the primary vertex.

The goal of the peak-finder algorithm developed in this thesis is to reconstruct the FastPV accurately and efficiently. The weights developed have been described above. Now, it is necessary to have an algorithm able to find the FastPV properly considering the projected z clusters distribution and these calculated weights.

The peak-finder algorithm is illustrated in Figure 4.14 and runs for all the events with the following structure:

- Looping over the projected z clusters.
- We defined a new set, which is appended with the highest weight-projected z cluster satisfying that the distance between them is smaller than 0.6 cm.
- The FastPV is defined as the weighted average of the projected z clusters in that set.

4.4 Vertex error

The vertex error is defined by the difference between the FastPV calculated and the MC-truth position. Figure 4.15 shows the error obtained using our peak-finder algorithm.

4.5 Bias feature. Correction factor

Due to the CMS geometry, more clusters are detected in the central region of the detector (pile-up). Thereby, we bias the FastPV reconstruction and this feature was observed in Figure 4.15. In

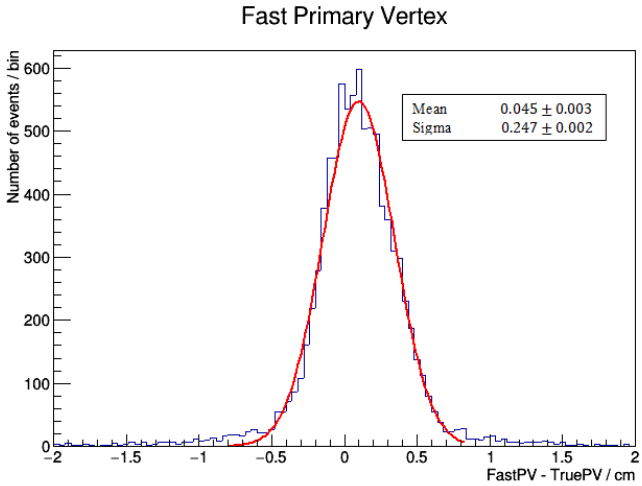


Figure 4.15: Difference between the FastPV and TruePV.

order to solve it, we need to add a correction factor. Figure 4.16 shows that our corrector factor will be: 0.97 ± 0.02. Figure 4.17 shows the fast primary vertex resolution once the correction have been applied: $0.97 \cdot z(\text{FastPV})$

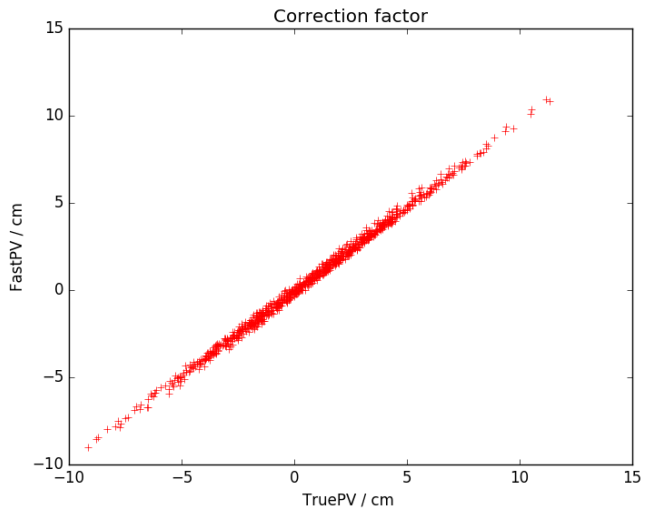


Figure 4.16: FastPV bias vs TruePV. From a linear fit, a correction factor of 0.97 ± 0.02 has been obtained.

4.6 Efficiency

First of all, we defined efficiency as the fraction of events having |FastPV - TruePV| < 1.5cm. The reason why we use exactly 1.5 cm is that in the track reconstruction, which comes after PV, we reconstruct tracks within 1.5 cm from FastPV. If |FastPV - TruePV| < 1.5 cm is not satisfied, we will not reconstruct the signal tracks.

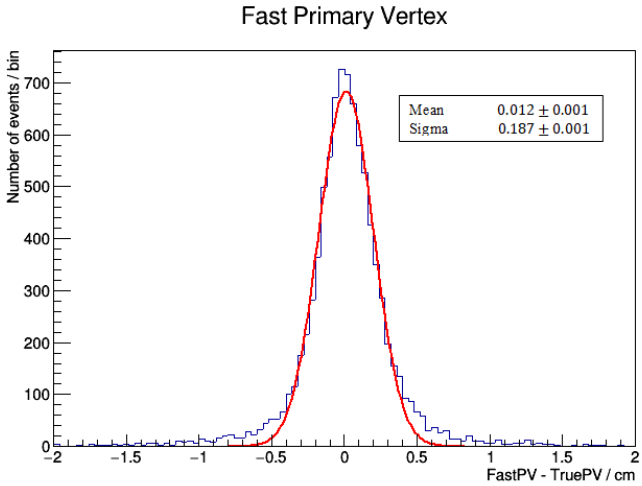


Figure 4.17: Fast Primary Vertex resolution.

Furthermore, we have to noticed that if FastPV reconstruction is wrong, it is completely wrong, |FastPV - TruePV| $\sim$ 10 cm. Events reconstructed accurately have |FastPV - TruePV| $\sim$ 1 mm.

The efficiency obtained without the weights implementation was 55 $\%$. With weights, it has improved up to 82$\%$.

Chapter 5

Conclusions

As a summary, we are working at the High Level Trigger, so we want reduce the time needed for the full event reconstruction. The motivation of the Fast Primary Vertex is to speed up tracking. We have seen how this algorithm gives a rough localization of the primary vertex.

The Fast Primary Vertex algorithm relies on the idea of tracks contained on a jet have almost the same direction of the jet. Thereby, we can project the pixel clusters to the beam line, along the jet direction. With this back projection, we will have a set of z projection over the beam line. From their distribution, we can guess where the primary vertex is. However, the efficiency was only of 55%.

In order to improve the efficiency, we applied different weights regarding the $\Delta\phi$(jet, cluster), cluster size X, cluster size Y, jet momentum and cluster charge. Figure 4.15 showed how the position of the primary vertex is much more accurate now.

Still, something else could be improved. To deal with the CMS geometry (pile-up) and avoid the bias in Figure 4.15, a correction in the central region factor was included. Finally, the efficiency was of 82% with pile-up 35.

Thereby, we can conclude that FastPV can be approximated in an easy and fast way from the projections of the cluster over the beam line with the jet-η direction. Their localization is improved if weights and the correction factor are implemented. A further study of the weights, working with more complex algorithm and having with more statistics could still make better the method.

Bibliography

[1] Mark Thomson. *Modern particle physics*. Cambridge University Press, New York, 2013.

[2] P Baudrenghien, T Bohl, T Linnecar, E Shaposhnikova, and J Tuckmantel. Nominal longitudinal parameters for the lhc beam in the cern sps. In *Particle Accelerator Conference, 2003. PAC 2003. Proceedings of the*, volume 5, pages 3050–3052. IEEE, 2003.

[3] Thomas Sven Pettersson and P. Lefèvre. The Large Hadron Collider: Conceptual design. 1995.

[4] N A Tahir, R Schmidt, M Brugger, R Assmann, A V Shutov, I V Lomonosov, A R Piriz, D H H Hoffmann, C Deutsch, and V E Fortov. The cern super proton synchrotron as a tool to study high energy density physics. *New Journal of Physics*, 10(7):073028, 2008.

[5] Lucio Rossi. Superconducting magnets for the lhc main lattice. *IEEE transactions on applied superconductivity*, 14(2):153–158, 2004.

[6] ALICE: Technical proposal for a large ion collider experiment at the CERN LHC. 1995.

[7] *LHCb reoptimized detector design and performance: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003.

[8] Monica Pepe Altarelli and Frederic Teubert. B physics at lhcb. In *Perspectives on LHC Physics*, pages 299–322. World Scientific, 2008.

[9] GL Bayatian, A Korablev, A Soha, O Sharif, M Chertok, W Mitaroff, F Pauss, V Genchev, M Wensveen, V Lemaitre, et al. Cms physics: Technical design report volume 1: Detector performance and software. Technical report, CMS-TDR-008-1, 2006.

[10] A. Airapetian et al. ATLAS: Detector and physics performance technical design report. Volume 1. 1999.

[11] Lucas Taylor. CMS detector design. 2011, 23 october. `http://CMS.web.cern.ch/news/CMS-detector-design`. Accessed: 2018, March 26.

[12] Lucas Taylor. Silicon strips. 2011, 23 october. http://CMS.web.cern.ch/news/silicon-strips. Accessed: 2018, March 30.

[13] S. Spannagel. Status of the CMS phase i pixel detector upgrade. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 831:71 – 75, 2016. Proceedings of the 10th International "Hiroshima" Symposium on the Development and Application of Semiconductor Tracking Detectors.

[14] Pippa S. Wells. The upgraded atlas and CMS detectors and their physics capabilities. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 373(2032), 2015.

[15] D. J. A. Cockerill. The CMS Electromagnetic Calorimeter at the LHC. In *Proceedings, 34th International Conference on High Energy Physics (ICHEP 2008): Philadelphia, Pennsylvania, July 30-August 5, 2008*, 2008.

[16] CMS collaboration et al. Performance and operation of the CMS electromagnetic calorimeter. *Journal of Instrumentation*, 5(03):T03010, 2010.

[17] Lucas Taylor. Ecal preshower. 2011, 23 November. http://CMS.web.cern.ch/news/ecal-preshower. Accessed: 2018, may 7.

[18]

[19] Lucas Taylor. Hadron calorimeter. 2011, 23 october. http://CMS.web.cern.ch/news/hadron-calorimeter. Accessed: 2018, March 31.

[20] Lucas Taylor. Muon detectors. 2011, 23 october. http://CMS.web.cern.ch/news/muon-detectors. Accessed: 2018, March 31.

[21] CMS collaboration et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10):P10009, 2014.

[22] Henning Kirschenmann. Jets at CMS and the determination of their energy scale. 2012, 9 july. http://CMS.web.cern.ch/news/CMS-detector-design. Accessed: 2018, April 23.

[23] Florian Beaudette. The CMS particle flow algorithm. *arXiv preprint arXiv:1401.8155*, 2014.

[24] http://erdit.eu/sites/erdit.web.cern.ch/files/pictures/HEP/multiVX. Accessed: 2018, April 23.

[25] CMS collaboration et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10):P10009, 2014.

[26] Thomas Speer, Kirill Prokofiev, R Frühwirth, Wolfgang Waltenberger, and Pascal Vanlaer. Vertex Fitting in the CMS Tracker. Technical Report CMS-NOTE-2006-032, CERN, Geneva, Feb 2006.

[27] Cristina Ferro. B-tagging at CMS. In *EPJ Web of Conferences*, volume 28, page 12055. EDP Sciences, 2012.

[28] `http://bartosik.pp.ua/hep`$sketches/btagging$. Accessed : 2018, April 23.

[29] CMS collaboration et al. The CMS trigger system. *arXiv preprint arXiv:1609.02366*, 2016.

[30] Vardan Khachatryan, AM Sirunyan, A Tumasyan, W Adam, E Asilar, T Bergauer, J Brandstetter, E Brondolin, M Dragicevic, J Eroe, et al. The CMS trigger system.

[31] Andrea Perrotta. Performance of the CMS high level trigger. *Journal of Physics: Conference Series*, 664(8):082044, 2015.

[32] Darin Acosta, CMS Collaboration, et al. CMS Trigger Improvements towards run ii. *Nuclear and Particle Physics Proceedings*, 273:1008–1013, 2016.

[33] Silvio Donato, Luigi Rolandi, and Andrea Rizzi. Search for the Standard Model Higgs boson decaying to b quarks with the CMS experiment, 2017.