SEARCH FOR PRODUCTION OF A HIGGS BOSON AND A SINGLE TOP QUARK IN MULTILEPTON FINAL STATES IN pp COLLISIONS AT $\sqrt{s} = 13$ TeV.

by

Jose Andres Monroy Montañez

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Doctor of Philosophy

Major: Physics and Astronomy

Under the Supervision of Kenneth Bloom and Aaron Dominguez

Lincoln, Nebraska

July, 2018

SEARCH FOR PRODUCTION OF A HIGGS BOSON AND A SINGLE TOP QUARK IN MULTILEPTON FINAL STATES IN pp COLLISIONS AT $\sqrt{s} = 13$ TeV.

Jose Andres Monroy Montañez, Ph.D.

University of Nebraska, 2018

Adviser: Kenneth Bloom and Aaron Dominguez

Table of Contents

List of Figures

List of Tables

Chapter 3

The CMS experiment at the LHC

3.1 Introduction

Located on the Swiss-French border, the European Council for Nuclear Research (CERN) is the largest scientific organization leading the particle physics research. About 13000 people in a broad range of fields including users, students, scientists, engineers, among others, contribute to the data taking and analysis, with the goal of unveiling the secrets of nature and revealing the fundamental structure of the universe. CERN is also the home of the Large Hadron Collider (LHC), the largest circular particle accelerator around the world, where protons (or heavy ions) traveling close to the speed of light, are made to collide. These collisions open a window to investigate how particles (and their constituents if they are composite) interact with each other, providing clues about the laws of nature. This chapter presents an overview of the LHC structure and operation. A detailed description of the CMS detector is offered, given that the data used in this thesis have been taken with this detector.

3.2 The LHC

With 27 km of circumference, the LHC is currently the largest and most powerful circular accelerator in the world. It is installed in the same tunnel where the Large Electron-Positron (LEP) collider was located, taking advantage of the existing infrastructure. The LHC is also the larger accelerator in the CERN's accelerator complex and is assisted by several successive accelerating stages before the particles are injected into the LHC ring where they reach their maximum energy (see Figure 3.1).

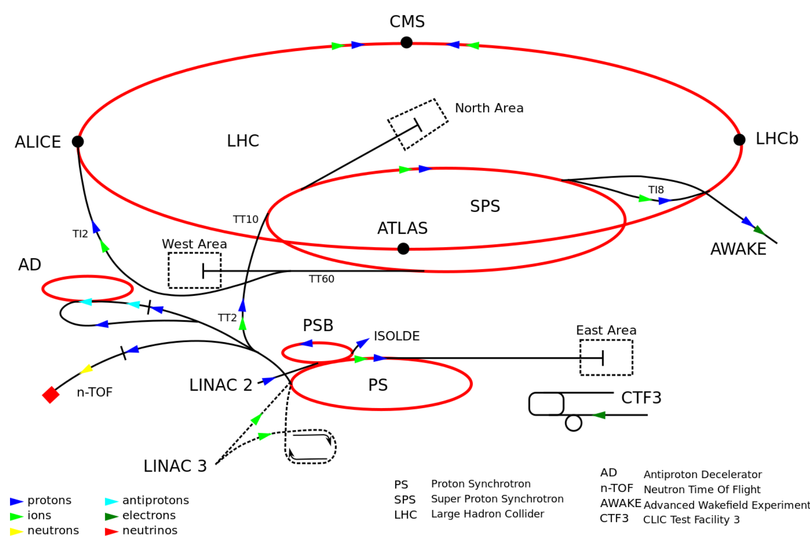

Figure 3.1: CERN accelerator complex. Blue arrows show the path followed by protons along the acceleration process [54].

LHC runs in three modes depending on the particles being accelerated

- Proton-Proton collisions (*pp*) for multiple physics experiments.
- Lead-Lead collisions (Pb-Pb) for heavy ion experiments.
- Proton-Lead collisions (p-Pb) for quark-gluon plasma experiments.

In this thesis only *pp* collisions will be considered.

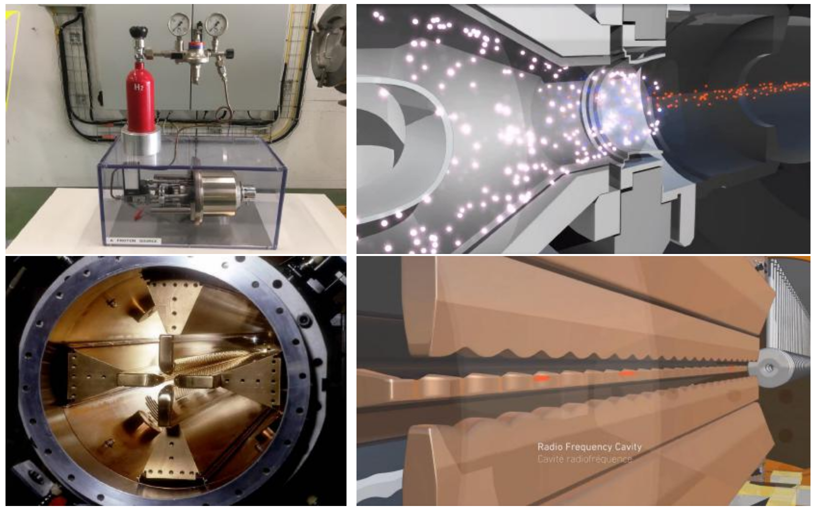

Figure 3.2: LHC protons source and the first acceleration stage. Top: the bottle contains hydrogen gas (white dots) which is injected into the metal cylinder to be broken down into electrons (blue dots) and protons (red dots); Bottom: the obtained protons are directed towards the radio frequency quadrupole which perform the first acceleration, focus the beam and create the bunches of protons. [58, 59]

Collection of protons starts with hydrogen atoms taken from a bottle, containing hydrogen gas, and injecting them in a metal cylinder; hydrogen atoms are broken down into electrons and protons by an intense electric field (see Figure 3.2 top). The resulting protons leave the metal cylinder towards a radio frequency quadrupole (RFQ) that focus the beam, accelerates the protons and creates the packets of protons called bunches. In the RFQ, an electric field is generated by a RF wave at a frequency that matches the resonance frequency of the cavity where the electrodes are contained. The beam of protons traveling on the RFQ axis experiences an alternating electric field gradient that generates the focusing forces.

In order to accelerate the protons, a longitudinal time-varying electric field component is added to the system; it is done by giving the electrodes a sinus-like profile as shown in Figure 3.2 bottom. By matching the speed and phase of the protons with the longitudinal electric field the bunching is performed; protons synchronized with the

RFQ (synchronous proton) do not feel an accelerating force, but those protons in the beam that have more (or less) energy than the synchronous proton (asynchronous protons) will feel a decelerating (accelerating) force; therefore, asynchronous protons will oscillate around the synchronous ones forming bunches of protons [56]. From the RFQ protons emerge with energy 750 keV in bunches of about 1.15×10^{11} protons [57].

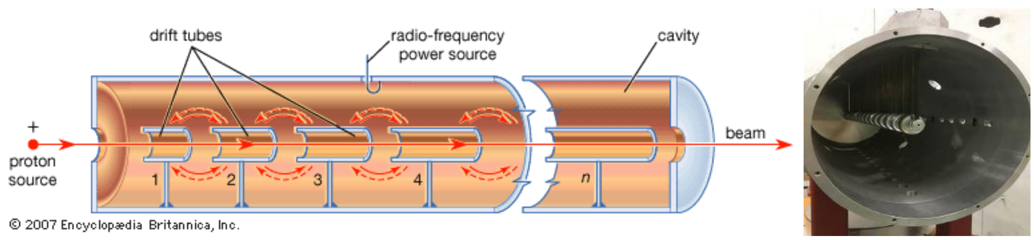

Figure 3.3: The LINAC2 accelerating system at CERN. Electric fields generated by radio frequency (RF) create acceleration and deceleration zones inside the cavity; deceleration zones are blocked by drift tubes where quadrupole magnets focus the proton beam. [60]

Proton bunches coming from the RFQ go to the linear accelerator 2 (LINAC2) where they are accelerated to reach 50 MeV energy. In the LINAC2 stage, acceleration is performed using electric fields generated by radio frequency which create zones of acceleration and deceleration as shown in Figure 3.3. In the deceleration zones, the electric field is blocked using drift tubes where protons are free to drift while quadrupole magnets focus the beam.

The beam coming from LINAC2 is injected into the proton synchrotron booster (PSB) to reach 1.4 GeV in energy. The next acceleration is provided at the proton synchrotron (PS) up to 26 GeV, followed by the injection into the super proton synchrotron (SPS) where protons are accelerated to 450 GeV. Finally, protons are injected into the LHC where they are accelerated to the target energy of 6.5 TeV. PSB, PS, SPS and LHC accelerate protons using the same RF acceleration technique described before.

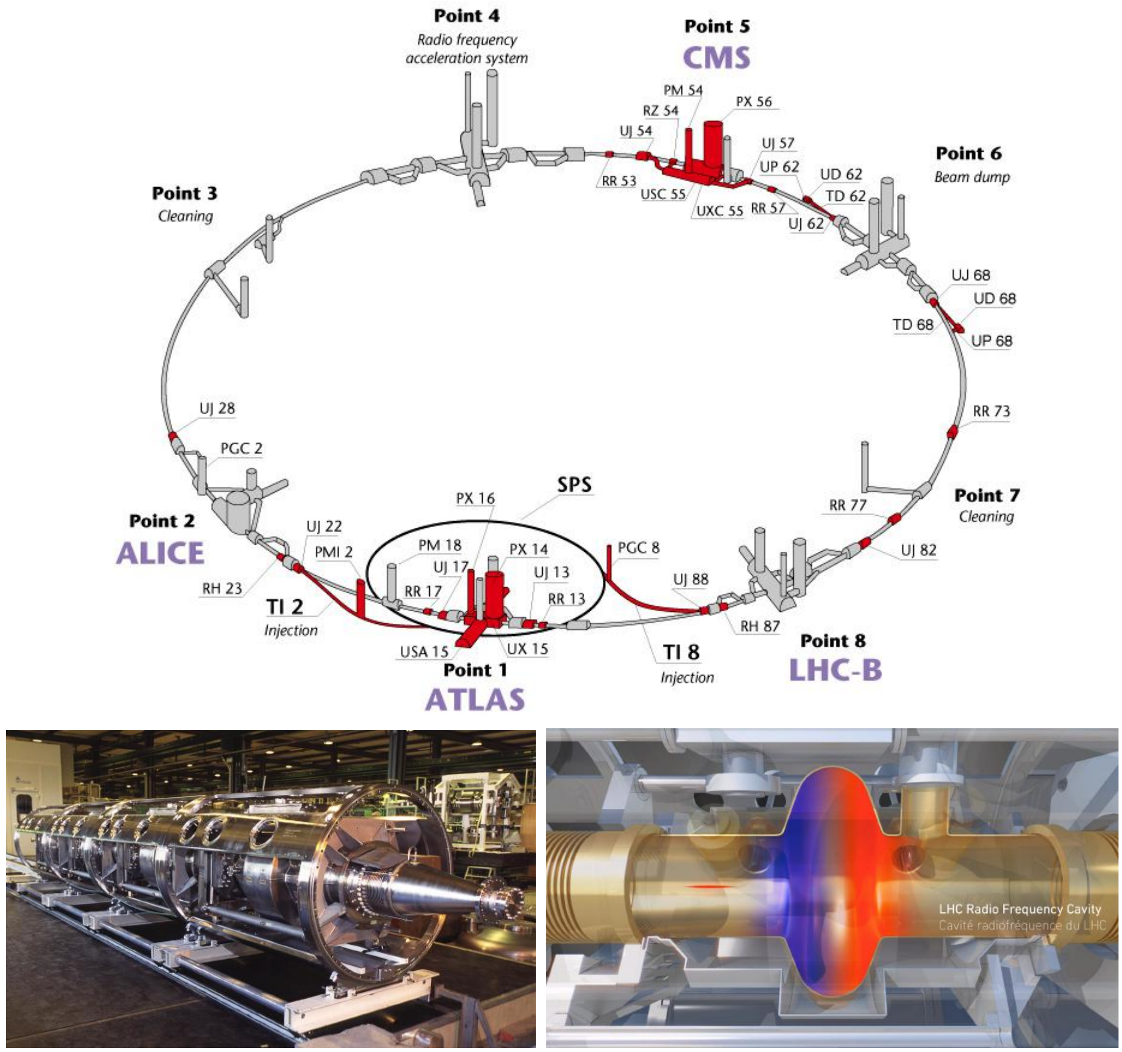

Figure 3.4: Top: LHC layout. The red zones indicate the infrastructure additions to the LEP installations, built to accommodate the ATLAS and CMS experiments which exceed the size of the former experiments located there [55]. Bottom: LHC RF cavities. A module accommodates 4 cavities that accelerate protons and preserve the bunch structure of the beam. [59, 61]

LHC has a system of 16 RF cavities located in the so-called point 4, as shown in Figure 3.4 top, tuned at a frequency of 400 MHz and the protons are carefully timed, so in addition to the acceleration effect the bunch structure of the beam is preserved. Bottom side of Figure 3.4 shows a picture of a RF module composed of 4 RF cavities working in a superconducting state at 4.5 K; also is showed a representation of the accelerating electric field that accelerates the protons in the bunch.

While protons are accelerated in one section of the LHC ring, where the RF cavities are located, in the rest of their path they have to be kept in the curved trajectory defined by the LHC ring. Technically, LHC is not a perfect circle; RF, injection, beam dumping, beam cleaning and sections before and after the experimental points where protons collide are all straight sections. In total, there are 8 arcs 2.45 Km long each and 8 straight sections 545 m long each. In order to curve the proton's trajectory in the arc sections, superconducting dipole magnets are used.

Inside the LHC ring, there are two proton beams traveling in opposite directions in two separated beam pipes; the beam pipes are kept at ultra-high vacuum ($\sim 10^{-9}$ Pa) to ensure that there are no particles that interact with the proton beams. The superconducting dipole magnets used in LHC are made of a NbTi alloy, capable of transporting currents of about 12000 A when cooled at a temperature below 2K using liquid helium (see Figure 3.5).

Protons in the arc sections of LHC feel a centripetal force exerted by the dipole magnets which is perpendicular to the beam trajectory; The magnitude of magnetic field needed can be found assuming that protons travel at $v \approx c$, using the standard values for proton mass and charge and the LHC radius, as

$$F_m = \frac{mv^2}{r} = qBv \quad \rightarrow B = 8.33T \tag{3.1}$$

which is about 100000 times the Earth's magnetic field. A representation of the magnetic field generated by the dipole magnets is shown on the bottom left side of Figure

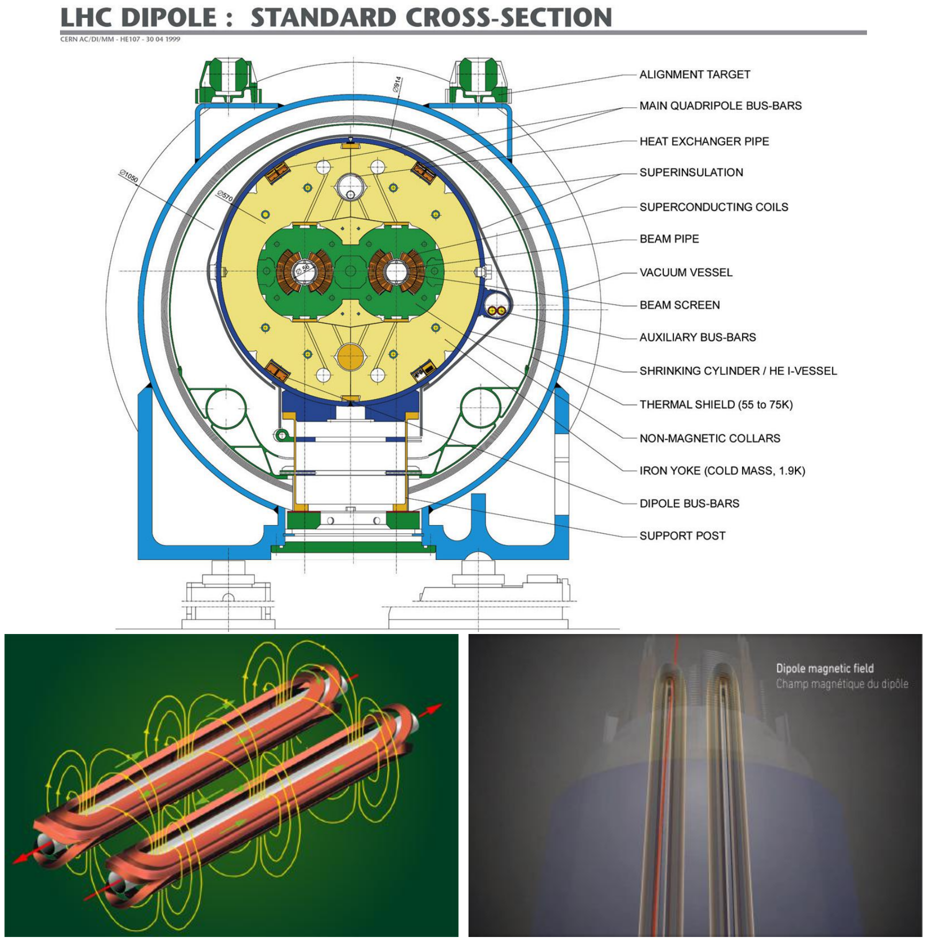

Figure 3.5: Top: LHC dipole magnet transverse view; cooling, shielding and mechanical support are indicated. Bottom left: Magnetic field generated by the dipole magnets; note that the direction of the field inside one beam pipe is opposite with respect to the other beam pipe which guarantee that both proton beams are curved in the same direction towards the center of the ring. The effect of the dipole magnetic field on the proton beam is represented on the bottom right side [59, 62, 63].

3.5. The bending effect of the magnetic field on the proton beam is shown on the bottom right side of Figure 3.5. Note that the dipole magnets are not curved; the arc section of the LHC ring is composed of straight dipole magnets of about 15 m. In total there are 1232 dipole magnets along the LHC ring.

In addition to bending the beam trajectory, the beam has to be focused so it stays inside the beam pipe. The focusing is performed by quadrupole magnets installed in a different straight section; in total 858 quadrupole magnets are installed along the LHC ring. Other effects like electromagnetic interaction among bunches, interaction with electron clouds from the beam pipe, the gravitational force on the protons, differences in energy among protons in the same bunch, among others, are corrected using sextupole and other magnetic multipoles.

The two proton beams inside the LHC ring are made of bunches with a cylindrical shape of about 7.5 cm long and about 1 mm in diameter; when bunches are close to the collision point (CP), the beam is focused up to a diameter of about 16 μm in order to maximize the number of collisions per unit area and per second, known as luminosity (L). Luminosity can be calculated using

$$L = fn\frac{N_1 N_2}{4\pi\sigma_x\sigma_y} \tag{3.2}$$

where f is the revolution frequency, n is the number of bunches per beam, N_1 and N_2 are the numbers of protons per bunch (1.5×10^{11}), σ_x and σ_y are the gaussian transverse sizes of the bunches. The expected luminosity is about

$$\begin{aligned}
f =& \frac{v}{2\pi r_{LHC}} \approx \frac{3\times 10^8 \mathrm{m/s}}{27\mathrm{km}} \approx 11.1\mathrm{kHz}, \\
n =& 2808 \\
N_1 =& N_2 = 1.5 \times 10^{11} \\
\sigma_x =& \sigma_y = 16\mu\mathrm{m}
\end{aligned}$$

$$L = 1.28 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1} = 1.28 \times 10^{-5}\mathrm{fb}^{-1}\mathrm{s}^{-1} \tag{3.3}$$

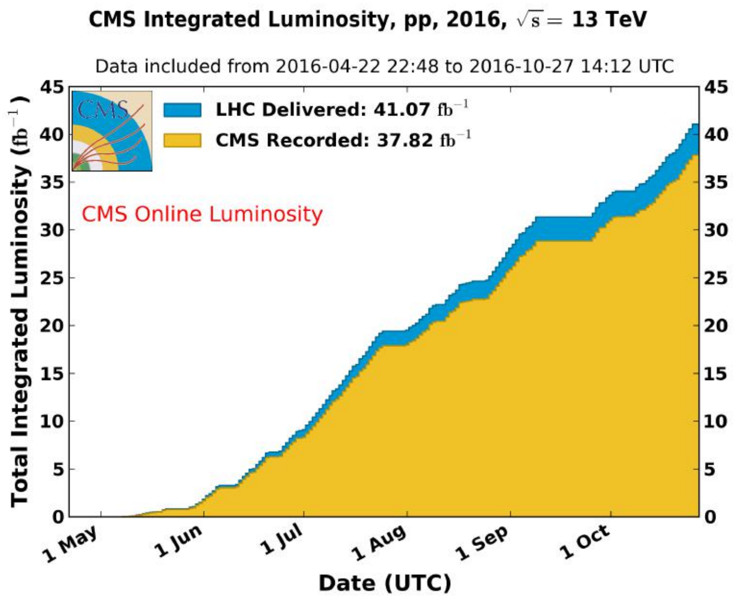

Figure 3.6: Integrated luminosity delivered by LHC and recorded by CMS during 2016. The difference between the delivered and the recorded luminosity is due to fails and issues occurred during the data taking in the CMS experiment [64].

Luminosity is a fundamental aspect of LHC given that the bigger luminosity the bigger number of collisions, which means that for processes with a very small cross section the number of expected occurrences is increased and so the chances of being detected. The integrated luminosity, i.e., the total luminosity, collected by the CMS experiment during 2016 is shown in Figure 3.6; the data analyzed in this thesis corresponds to an integrated luminosity of 35.9 fb^{-1} at a center of mass-energy $\sqrt{s} = 13$ TeV.

A way to increase L is increasing the number of bunches in the beam. Currently, the separation between two consecutive bunches in the beam is 7.5 m which corresponds to a time separation of 25 ns. In the full LHC ring the allowed number of bunches is $n = 27\mathrm{km}/7.5\mathrm{m} = 3600$; however, there are some gaps in the bunch pattern intended for preparing the dumping and injection of the beam, thus, the proton beams are composed of 2808 bunches.

Once the proton beams reach the desired energy, they are brought to cross each other producing proton-proton collisions. The bunch crossing happens in precise places where the four LHC experiments are located, as seen in the top of Figure 3.7. In 2008, the first set of collisions involved protons with $\sqrt{s} = 7$ TeV; the energy was increased to 8 TeV in 2012 and to 13 TeV in 2015.

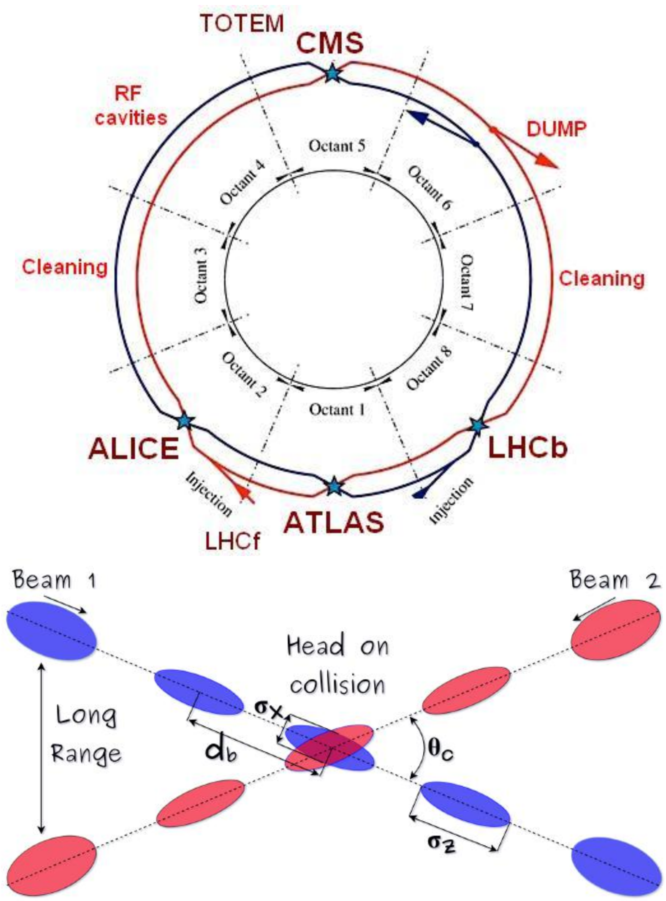

Figure 3.7: Top: LHC interaction points. Bunch crossing occurs where the LHC experiments are located [65]. Sections indicated as cleaning are dedicated to collimate the beam in order to protect the LHC ring from collisions with protons in very spreaded bunches. Bottom: bunch crossing scheme. Since the bunch crossing is not perfectly head-on, the luminosity is reduced in a factor of 17%; adapted from Reference [77].

CMS and ATLAS experiments, which are multi-purpose experiments, are enabled to explore physics in any of the collision modes. LHCb experiment is optimized to explore bottom quark physics, while ALICE is optimized for heavy ion collisions searches; TOTEM and LHCf are dedicated to forward physics studies; MoEDAL (not indicated in the Figure) is intended for monopoles or massive pseudo stable particles searches.

At the CP there are two interesting details that need to be addressed. The first one is that the bunch crossing does not occur head-on but at a small crossing angle θ_c (280 μrad in CMS and ATLAS) as shown in the bottom side of Figure 3.7, affecting the overlapping between bunches; the consequence is a reduction of about 17% in the luminosity (represented by a factor not included in eqn. 3.2). The second one is the occurrence of multiple pp collisions in the same bunch crossing; this effect is called pile-up (PU). A fairly simple estimation of the PU follows from estimating the probability of collision between two protons, one from each of the bunches in course of collision; it depends roughly on the ratio of proton size and the cross section of the bunch in the interaction point, i.e.,

$$P(pp-collision) \sim \frac{d^2_{proton}}{\sigma_x \sigma_y} = \frac{(1\text{fm})^2}{(16\mu\text{m})^2} \sim 4 \times 10^{-21} \tag{3.4}$$

however, there are $N = 1.15 \times 10^{11}$ protons in a bunch, thus the estimated number of collisions in a bunch crossing is

$$PU = N^2 * P(pp-collision) \sim 50 \ pp\text{-collision per bunch crossing,} \tag{3.5}$$

about 20 of those pp collisions are inelastic. Each collision generates a vertex, but only the most energetic is considered as a primary vertex; the rest are considered as PU vertices. A multiple *pp* collision event in a bunch crossing at CMS is showed in Figure 3.8. Unstable particles outgoing from the primary vertex will eventually decay; this decay vertex is known as a secondary vertex.

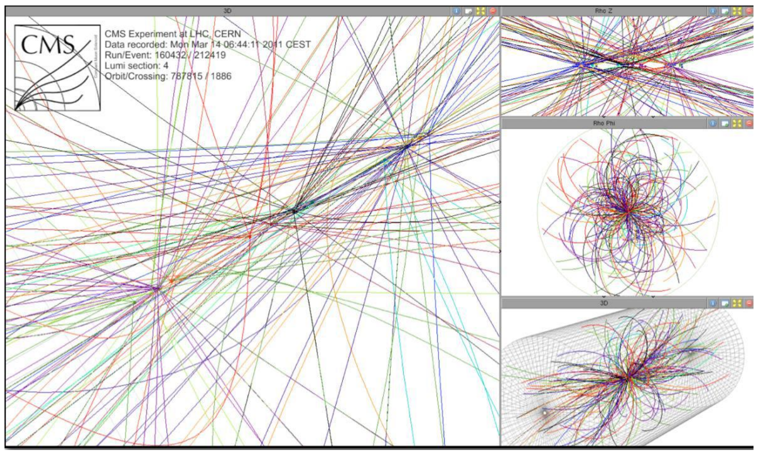

Figure 3.8: Multiple *pp* collision bunch crossing at CMS. Only the most energetic vertex is considered and the rest are cataloged as PU vertices [66].

Next section presents a description of the CMS detector which it is the detector used to collect the data used in this thesis.

3.3 The CMS experiment

CMS is a general-purpose detector designed to conduct research in a wide range of physics from the standard model to new physics like extra dimensions and dark matter. Located at the point 5 in the LHC layout as shown in Figure 3.4, CMS is composed of several detection systems distributed in a cylindrical structure; in total, CMS weights about 12500 tons in a very compact 21.6 m long and 14.6 m diameter cylinder. It was built in 15 separate sections at the ground level and lowered to the cavern individually to be assembled. A complete and detailed description of the CMS detector and its components is given in Reference [67] on which this section is based on.

Figure 3.9 shows the layout of the CMS detector. The design is driven by the requirements on the identification, momentum resolution and unambiguous charge determination of the muons; therefore, a large bending power is provided by the solenoid magnet made of superconducting cable capable to generate a 3.8 T magnetic field. The detection system is composed of (from the innermost to the outermost)

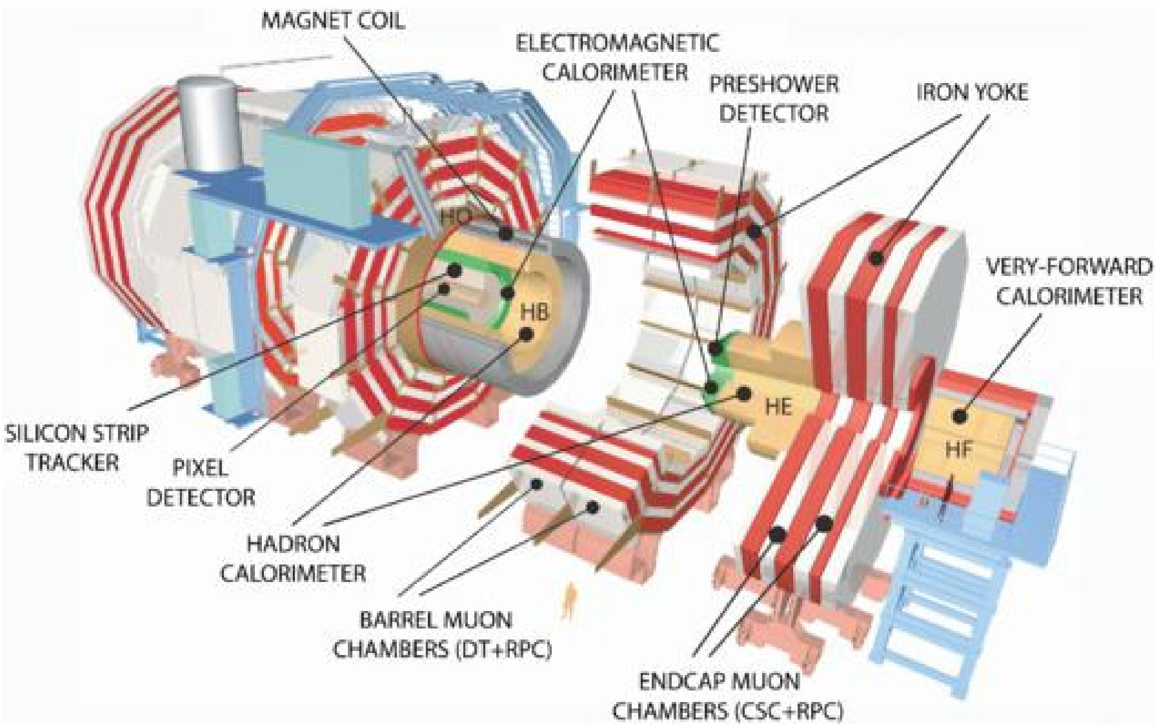

Figure 3.9: Layout of the CMS detector. The several subdetectors are indicated. The central region of the detector is referred as the Barrel section while the endcaps are referred as the forward sections. [68].

- Pixel detector.
- Silicon strip tracker.
- Preshower detector.

- Electromagnetic calorimeter.
- Hadronic calorimeter.
- Muon chambers (Barrel and endcap)

The central region of the detector is commonly referred as the barrel section while the endcaps are referred as the forward sections of the detector; thus, each subdetector is composed of a barrel section and a forward section.

3.3.1 Coordinate system

The coordinate system used by CMS is centered in the geometrical center of the detector which is the same as the CP as shown in Figure 3.10. The z-axis is parallel to the beam direction, while the Y-axis pointing vertically upward, and the X-axis pointing radially inward toward the center of the LHC.

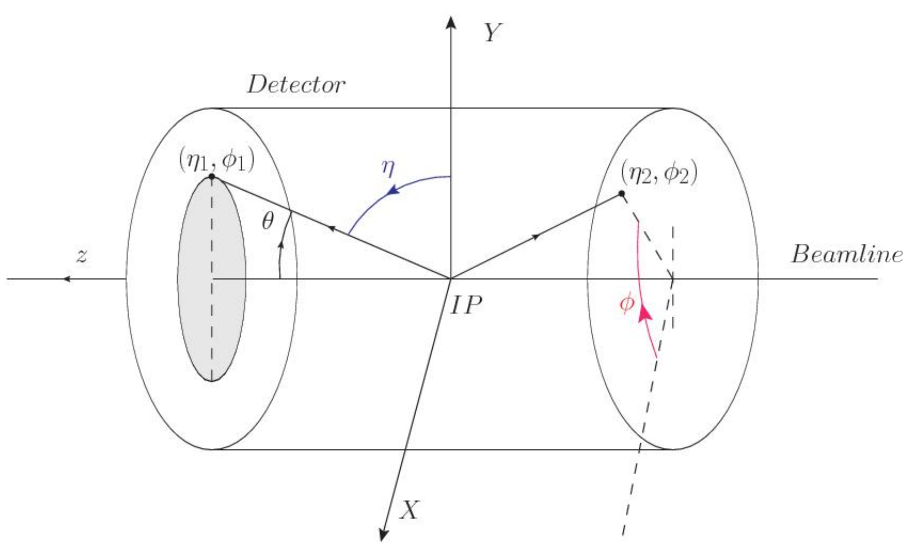

Figure 3.10: CMS detector coordinate system.

In addition to the common cartesian and cylindrical coordinate systems, two coordinates are of particular utility in particle physics: rapidity (y) and pseudorapidity

(η), defined in connection to the polar angle θ, energy and longitudinal momentum component (momentum along the z-axis) according to

$$y = \frac{1}{2} ln \frac{E + p_z}{E - p_z} \qquad \eta = -ln\left(tan \frac{\theta}{2}\right) \tag{3.6}$$

Rapidity is related to the angle between the XY-plane and the direction in which the products of a collision are emitted; it has the nice property that the difference between the rapidities of two particles is invariant with respect to Lorentz boosts along the z-axis. Thus, data analysis becomes more simple when based on rapidity; however, it is not simple to measure the rapidity of highly relativistic particles, as those produced after pp collisions. Under the highly relativistic motion approximation, y can be rewritten in terms of the polar angle, concluding that rapidity is approximately equal to the pseudorapidity defined above, i.e., $y \approx \eta$. Note that η is easier to measure that y given the direct relationship between the former and the polar angle. Angular distance between two objects in the detector (ΔR) is defined in terms of their coordinates (η_1, ϕ_1), (η_2, ϕ_2) as

$$\Delta R = \sqrt{(\Delta\eta)^2 - (\Delta\phi)^2} \tag{3.7}$$

3.3.2 Pixels detector

The CMS tracking system is designed to provide a precise measurement of the trajectory (*track*) followed by the charged particles created after the pp collisions; also, the precise reconstruction of the primary and secondary origins (*vertices*) is expected in an environment where, each 25 ns, the bunch crossing produce about 20 inelastic collisions and about 1000 particles. An increment in the luminosity is ongoing which implies that the PU will increase accordingly.

The pixel detector was replaced during the 2016-2017 extended year-end technical stop, due to the increasingly challenging operating conditions like the higher particle flow and more radiation harsh environment, among others. The new one is responding as expected, reinforcing its crucial role in the successful way to fulfill the new LHC physics objectives after the discovery of the Higgs boson. The last chapter of this thesis is dedicated to describe my contribution to the *Forward Pixel Phase 1 upgrade*.

The current pixel detector is composed of 1856 silicon pixel detector modules organized in four-barrel layers in the central region and three disks in the forward region; it is designed to record efficiently and with high precision, up to 10μm in the XY-plane and 20μm in the z-direction, the first four space-points (*hits*) near to the CP region (see Figure 3.11 left side) in the range $|\eta| \leq 2.5$. The first barrel layer is located at a radius of 30 mm from the beamline, while the fourth layer is located at a radius of 160 mm closer to the strip tracker inner barrel layer (see Section 3.3.3) in order to reduce the rate of fake tracks. The high granularity of the detector is represented in its about 123 Mpixels, each of size $100 \times 150\mu\mathrm{m}^2$, which is almost twice the channels of the old detector. The transverse momentum resolution of tracks can be measured with a resolution of 1-2% for muons of $p_T = 100$ GeV.

Some of the improvements with respect to the previous pixel detector include a higher average tracking efficiency and lower average fake rate as well as higher track impact parameter resolution which is fundamental in order to increase the efficiency in the identification of jets originating from b quarks (b-tagging). A significant source of improvement comes from the overall reduction in the material budget of the detector which results in fewer photon conversions and less multiple scattering from charged particles.

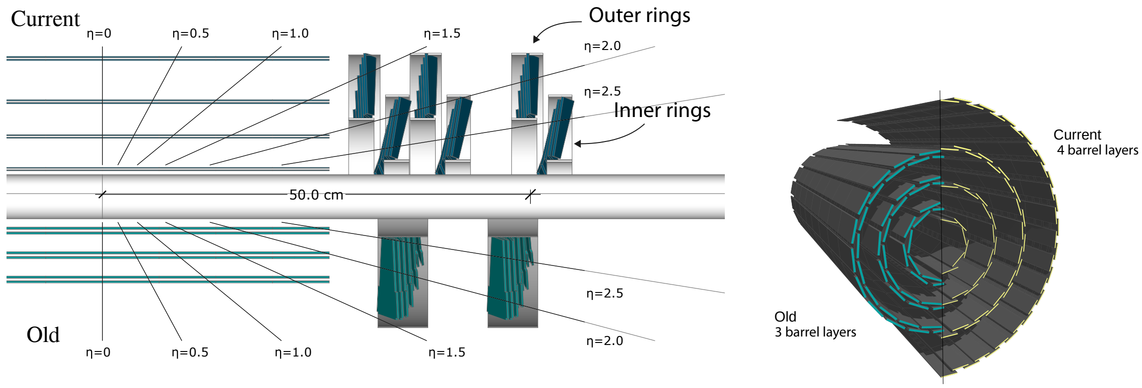

Figure 3.11: CMS pixel detector schematic view. Left: layout comparing the layers and disks in the old and current pixel detectors. Right: Transverse-oblique view comparing the pixel barrel layers in the two [70].

3.3.3 Silicon strip tracker

The silicon strip tracker (SST) is the second stage in the CMS tracking system. The top side of Figure 3.12 shows a schematic of the SST. The inner tracker region is composed of the tracker inner barrel (TIB) and the tracker inner disks (TID) covering the region $r < 55$ cm and $|z| < 118$ cm. The TIB is composed of 4 layers while the TID is composed of 3 disks at each end. The silicon sensors in the inner tracker are 320 μm thick, providing a resolution of about 13-38 μm in the $r\phi$ position measurement.

The modules indicated in blue in the schematic view of Figure 3.12 are two modules mounted back-to-back and rotated in the plane of the module by a *stereo* angle of 100 mrad; the hits from these two modules, known as *stereo hits*, are combined to provide a measurement of the second coordinate (z in the barrel and r on the disks) allowing the reconstruction of hit positions in 3-D.

The outer tracker region is composed of the tracker outer barrel (TOB) and the tracker endcaps (TEC). The six layers of the TOB offer coverage in the region $r > 55$

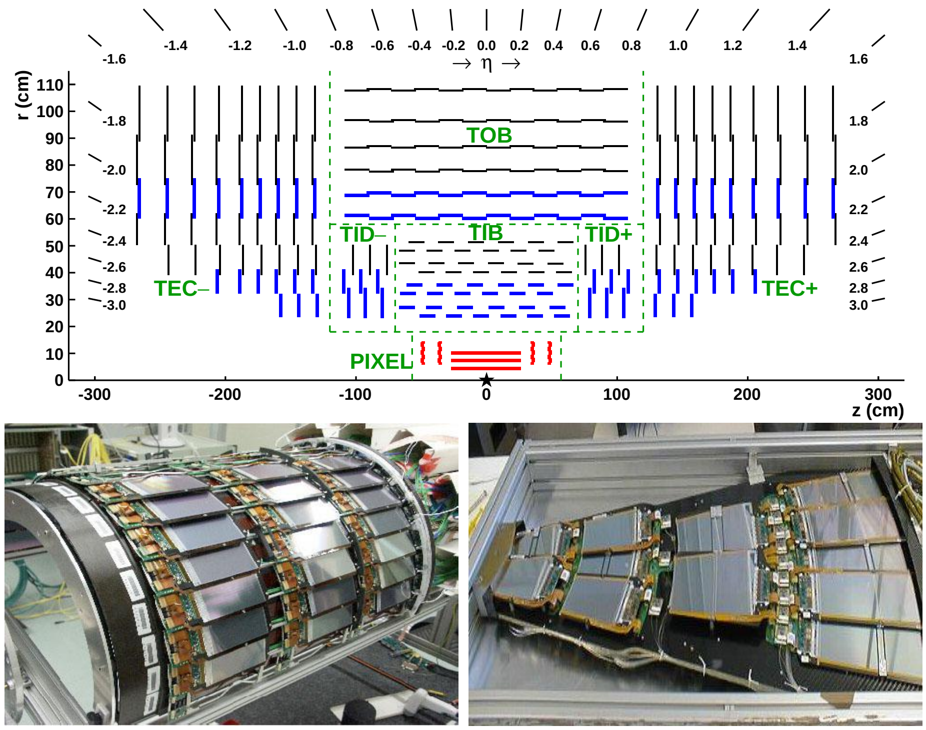

Figure 3.12: Top: CMS Silicon Strip Tracker (SST) schematic view. The SST is composed of the tracker inner barrel (TIB), the tracker inner disks (TID), the tracker outer barrel (TOB) and the tracker endcaps (TEC). Each part is made of silicon strip modules; the modules in blue represent two modules mounted back-to-back and rotated in the plane of the module by a stereo angle of 120 mrad in order to provide a 3-D reconstruction of the hit positions. Bottom: pictures of the TIB (left) and TEC (right) modules [71–73].

cm and $|z| < 118$ cm, while the 9 disks of the TEC cover the region $124 < |z| < 282$ cm. The resolution offered by the outer tracker is about 13-38 μm in the $r\phi$ position measurement. The inner four TEC disks use silicon sensors 320 μm thick; those in the TOB and the outer three TEC disks use silicon sensors of 500 μm thickness. The silicon strips run parallel to the z-axis and the distance between strips varies from 80 μm in the inner TIB layers to 183 μm in the inner TOB layers; in the endcaps the wedge-shaped sensors with radial strips, whose pitch range between 81 μm at small radii and 205 μm at large radii.

The whole SST has 15148 silicon modules, 9.3 million silicon strips and cover a total active area of about 198 m^2.

3.3.4 Electromagnetic calorimeter

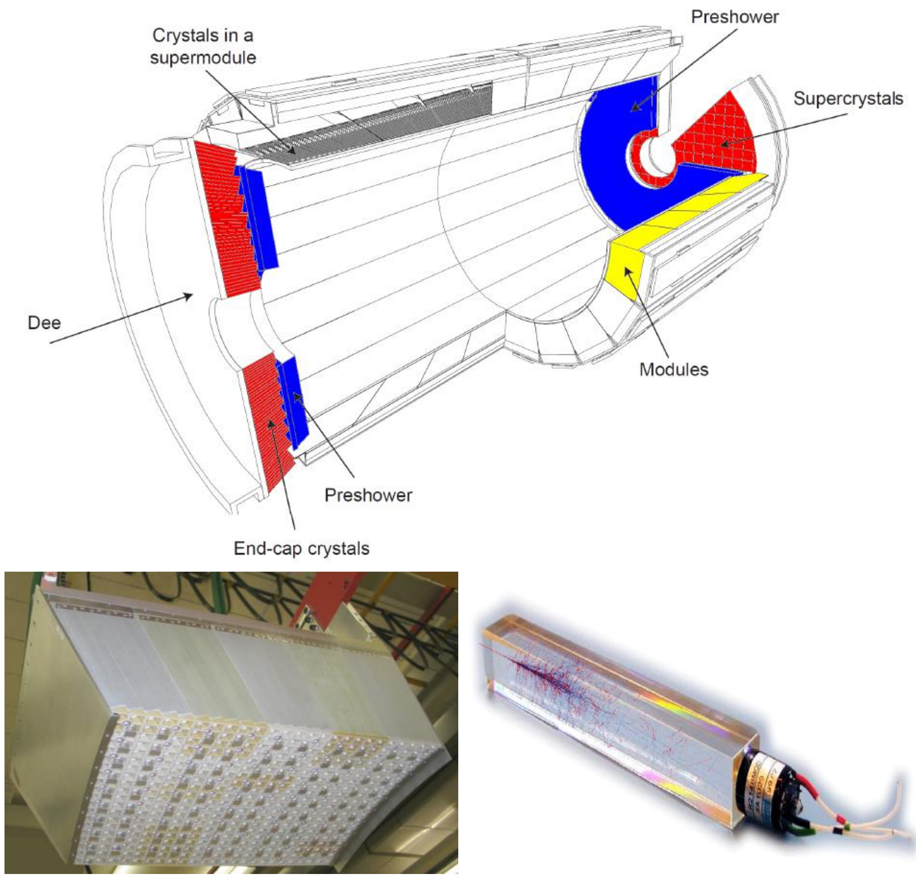

Figure 3.13: Top: CMS ECAL schematic view. Bottom: Module equipped with the crystals (left); ECAL crystal(right).

The CMS electromagnetic calorimeter (ECAL) is designed to measure the energy of electrons and photons. It is composed of 75848 lead tungstate crystals which have a short radiation length (0.89 cm) and fast response, since 80% of the light is emitted within 25 ns; however, they are combined with Avalanche photodiodes (APDs) as photodetectors given that crystals themselves have a low light yield (30γ/MeV). A schematic view of the ECAL is shown in Figure 3.13.

Energy is measured when electrons and photons are absorbed by the crystals which generates an electromagnetic *shower*, as seen in bottom right picture of the Figure 3.13; the shower is seen as a *cluster* of energy which depending on the amount of energy deposited can involve several crystals. The ECAL barrel (EB) covers the region $|\eta| < 1.479$, using crystals of depth of 23 cm and 2.2×2.2 cm^2 transverse section; the ECAL endcap (EE) covers the region $1.479 < |\eta| < 3.0$ using crystals of depth 22 cm and transverse section of 2.86×2.86 cm^2; the photodetectors used are vacuum phototriodes (VPTs). Each EE is divided in two structures called *Dees*.

In front of the EE, it is installed the preshower detector (ES) which covers the region $1.653 < |\eta| < 2.6$. The ES provides a precise measurement of the position of electromagnetic showers, which allows to distinguish electrons and photons signals from π^0 decay signals. The ES is composed of a layer of lead absorber followed by a layer of plastic scintillators

3.3.5 Hadronic calorimeter

Hadrons are not absorbed by the ECAL but by the hadron calorimeter (HCAL), which is made of a combination of alternating brass absorber layers and silicon photomultiplier(SiPM) layers; therefore, particles passing through the scintillator material produce showers, as in the ECAL, as a result of the inelastic scattering of the hadrons with the detector material. Since the particles are not absorbed in the scintillator, their energy is sampled; therefore the total energy is not measured but estimated from the energy clusters, which reduce the resolution of the detector. Brass was chosen

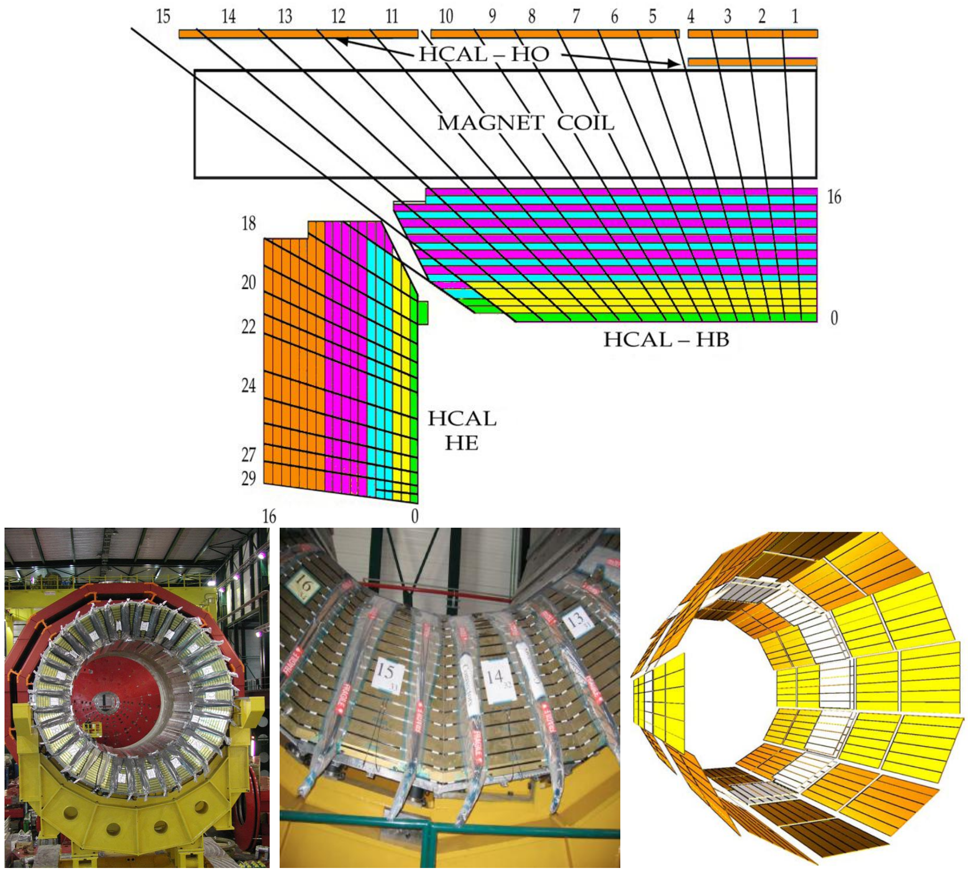

Figure 3.14: Top: CMS HCAL schematic view, the colors indicate the layers that are grouped into the same readout channels. Bottom: picture of a section of the HB; the absorber material is the golden region and scintillators are placed in between the absorber material (left and center). Schematic view of the HO (right). [74, 75]

as the absorber material due to its short interaction length ($\lambda_I = 16.42$cm) and its non-magnetivity. Figure 3.14 shows a schematic view of the CMS HCAL.

The HCAL is divided into four sections; the Hadron Barrel (HB), the Hadron Outer (HO), the Hadron Endcap (HE) and the Hadron Forward (HF) sections. The HB covers the region $0 < |\eta| < 1.4$, while the HE covers the region $1.3 < |\eta| < 3.0$. The HF, made of quartz fiber scintillator and steel as absorption material, covers the forward region $3.0 < |\eta| < 5.2$. Both the HB and HF are located inside the solenoid. The HO is placed outside the magnet as an additional layer of scintillators with the purpose of measure the energy tails of particles passing through the HB and the magnet (see Figure 3.14 top and bottom right). The upgrades made to the HCAL during the technical stop 2016-2017 consisted in the replacement of the photo transducer, in order to improve the efficiency.

3.3.6 Superconducting solenoid magnet

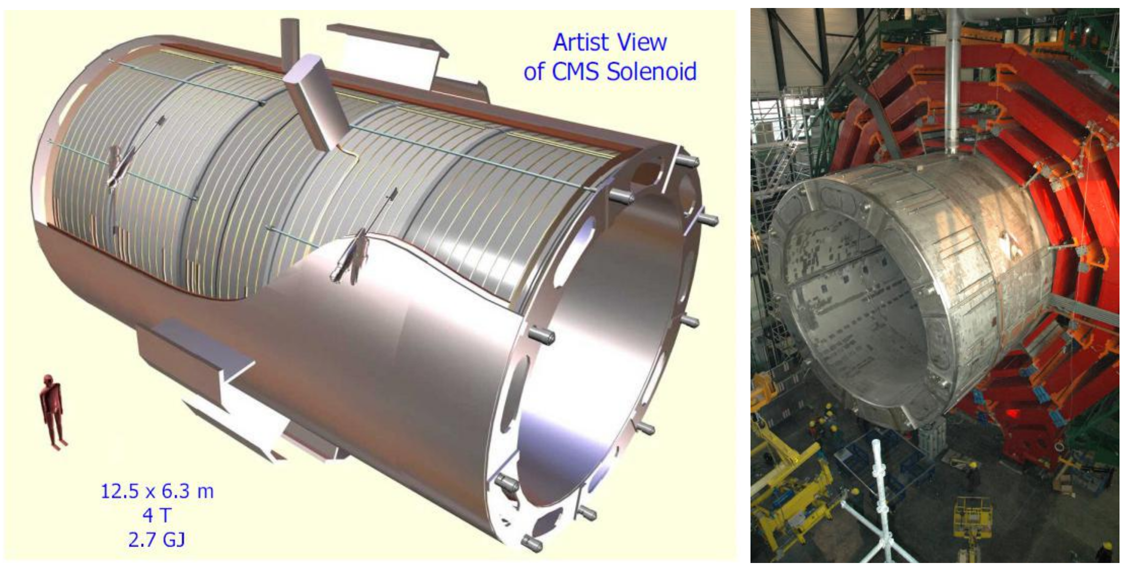

Figure 3.15: Artistic representation of the CMS solenoid magnet(left). The magnet is supported on an iron yoke (right) which also serves as the house of the muon detector and as mechanical support for the whole CMS detector [69].

The superconducting magnet installed in the CMS detector is designed to provide an intense and highly uniform magnetic field in the central part of the detector. In fact, the tracking system takes advantage of the bending power of the magnetic field to measure with precision the momentum of the particles that traverse it; the unambiguous determination of the sign for high momentum muons was a driven principle during the design of the magnet. The magnet has a diameter of 6.3 m, a length of 12.5 m and a cold mass of 220 t; the generated magnetic field reaches a strength of 3.8T.

Since it is made of Ni-Tb superconducting cable it has to operate at a temperature of 4.7 K by using a helium cryogenic system; the current circulating in the cables reaches 18800 A under normal running conditions. The left side of Figure 3.15 shows an artistic view of the CMS magnet, while the right side shows a transverse view of the cold mass where the winding structure is visible.

The yoke (see Figure 3.15), composed of 5 barrel wheels and 6 endcap disks made of iron, serves not only as the media for magnetic flux return but also provides the house for the muon detector system and structural stability to the full detector.

3.3.7 Muon system

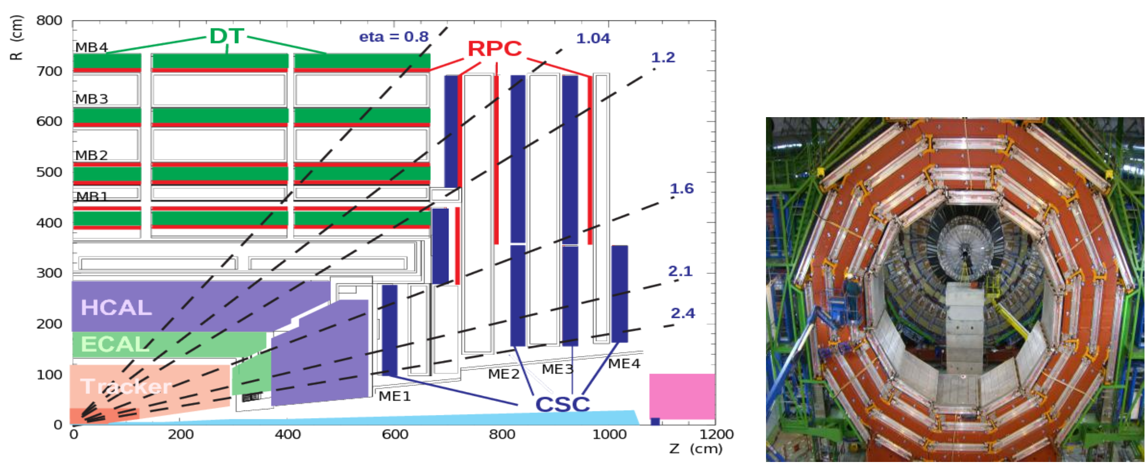

Figure 3.16: Left: CMS muon system schematic view; Right: one of the yoke rings with the muon DTs and RPCs installed; in the back it is possible to see the muon endcap [76].

Muons are the only charged particles able to pass through all the CMS detector due to their low ionization energy loss; thus, muons can be separated easily from the high amount of particles produced in a *pp* collision. Also, muons are expected to be produced in the decay of several new particles; therefore, a good detection of muons was on the leading principles when designing the CMS detector.

The CMS muon detection system (muon spectrometer) is embedded in the return yoke as seen in Figure 3.16. It is composed of three different detector types, the drift tube chambers (DT), Cathode strip chambers (CSC), and resistive plate chambers (RPC); DT are located in the central region $\eta < 1.2$ arranged in four layers of drift chambers filled with an Ar/CO_2 gas mixture.

The muon endcaps are made of CSCs covering the region $\eta < 2.4$ and filled with a mixture of $Ar/CO_2/CF_4$. The reason behind using a different detector type lies on the different conditions in the forward region like the higher muon rate and higher residual magnetic field compared to the central region.

The third type of detector used in the muon system is a set of four disks of RPCs working in avalanche mode. The RPCs provide good spatial and time resolutions. The track of $high - p_T$ muon candidates is built combining information from the tracking system and the signal from up to six RPCs and four DT chambers.
The muon tracks are reconstructed from the hits in the several layers of the muon system.

3.3.8 CMS trigger system

Under normal conditions, CMS expects pp collisions every 25 ns, i.e., an interaction rate of 40 MHz for which it is not possible to store the recorded data in full. In order to handle this high event rate data, an online event selection, known as triggering, is performed; triggering reduce the event rate to 100 Hz for storage and further offline analysis.

The trigger system starts with a reduction of the event rate to 100 kHz in the so-called *level 1 trigger (L1)*. L1 is based on dedicated programmable hardware like Field Programmable Gate Arrays (FPGAs) and Application Specific Integrated Circuits (ASICs), partly located in the detector itself; another portion is located in the CMS under-ground cavern. Hit patterns information from the muon chambers and the energy deposits in the calorimeter are used to decide if an event is accepted or rejected, according to selection requirements previously defined, which reflect the interesting physics processes. Figure 3.17 shows the L1 trigger architecture.

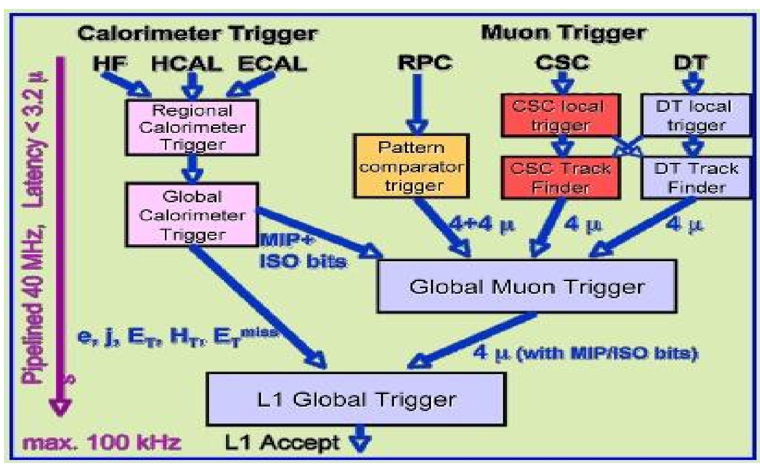

Figure 3.17: CMS Level-1 trigger architecture [77].

The second stage in the trigger system is called *high-level trigger (HLT)*; events accepted by L1 are passed to HLT in order to make an initial reconstruction of them. HLT is software based and runs on a dedicated server farm, using selection algorithms and high-level object definitions; the event rate at HLT is reduced to 100 Hz. The first HLT stage takes information from the muon detectors and the calorimeters to make the initial object reconstruction; in the next HLT stage, information from the pixel and strip detectors is used to do first fast-tracking and then full tracking online. This initial object reconstruction is used in further steps of the trigger system.

Events and preliminary reconstructed physics objects from HLT are sent to be fully reconstructed at the CERN computing center. Again, the pixel detector information provides high-quality seeds for the track reconstruction algorithm offline, primary vertex reconstruction, electron and photon identification, muon reconstruction, τ identification, and b-tagging. After full reconstruction, data sets are made available for offline analyses.

During the 2016-2017 technical stop, the L1 system was updated in order to improve the physics object identification by improving the algorithms and accounting for the increasing pile-up scenario.

3.3.9 CMS computing

After the data, coming from the experiment, are processed at several levels, they have to be stored and made available for further analysis; in order to cope all the tasks implied in the offline data processing, like transfer, simulation, reconstruction and reprocessing, among others, a big computing power is required. The CMS computing system is based on the distributed architecture concept, where users of the system and physical computer centers are distributed worldwide and interconnected by high-speed networks.
The worldwide LHC computing grid (WLCG) is the mechanism used to provide that distributed environment. WLCG is a tiered structure connecting computing centers around the world, which provides the necessary storage and computing facilities. The

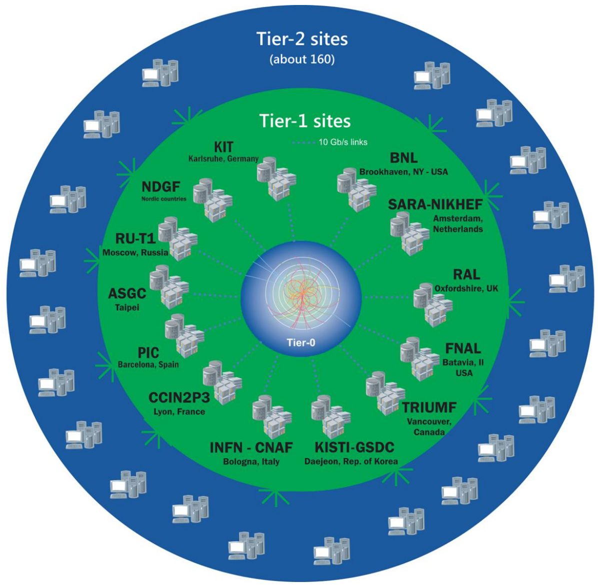

Figure 3.18: WLCG structure. The primary computer centers (Tier-0) are located at CERN (data center) and at the Wigner datacenter in Budapest. Tier-1 is composed of 13 centers and Tier-2 is composed of about 160 centers. [78].

primary computing centers of the WLCG are located at the CERN and the Wigner datacenter in Budapest and are known as Tier-0 as shown in Figure 3.18. The main responsibilities for each tier level are [78]

- **Tier-0**: initial reconstruction of recorded events and storage of the resulting datasets, the distribution of raw data to the Tier-1 centers.

- **Tier-1**: provide storage capacity, support for the Grid, safe-keeping of a proportional share of raw and reconstructed data, large-scale reprocessing and safe-keeping of corresponding output, generation of simulated events, distribution of data to Tier 2s, safe-keeping of a share of simulated data produced at these Tier 2s.

- **Tier-2**: store sufficient data and provide adequate computing power for specific analysis tasks, provide analysis requirements and proportional share of simulated event production and reconstruction.

Aside from the general computing strategy to manage the huge amount of data produced by experiments, CMS uses a framework to perform a variety of processing, selection and analysis tasks. The central concept of the CMS data model referred to as *event data model* (EDM) is the *Event*; therefore, an event is the unit that contains the information from a single bunch crossing as well as any data derived from that information like the reconstructed objects, the details under which additional data are derived.

Events are passed as the input to the *physics modules* that obtain information from them and create new one; for instance, *event data producers* add new data into the events, *analyzers* produce an information summary from an event set, *filters* perform selection and triggering.

CMS uses several event formats with different levels of detail and precision

- **Raw format**: events in this format contain the full recorded information from the detector as well as trigger decision and other metadata. An extended version of raw data is used to store information from the CMS Monte Carlo simulation tools. Raw data are stored permanently, occupying about 2MB/event
- **RECO format**: events in this format correspond to raw data that have been submitted to reconstruction algorithms like primary and secondary vertex reconstruction, particle ID, track-finding. RECO events contain physical objects and all the information used to reconstruct them; average size is about 0.5 MB/event.
- **AOD format**: Analysis Object Data (AOD) is the data format used in the physics analyses given that it contains the parameters describing the high-level

physics objects in addition to enough information to allow a kinematic refitting if needed. AOD events are filtered versions of the RECO events to which skimming or other kind processes have been applied. Requires about 100 kB/event.

- **Non-event data** are data needed to interpret and reconstruct events. Some of the non-event data used by CMS contains information about the detector contraction and condition data like calibrations, alignment, and detector status.

Figure 3.19 shows the data flow scheme between CMS detector and hardware tiers.

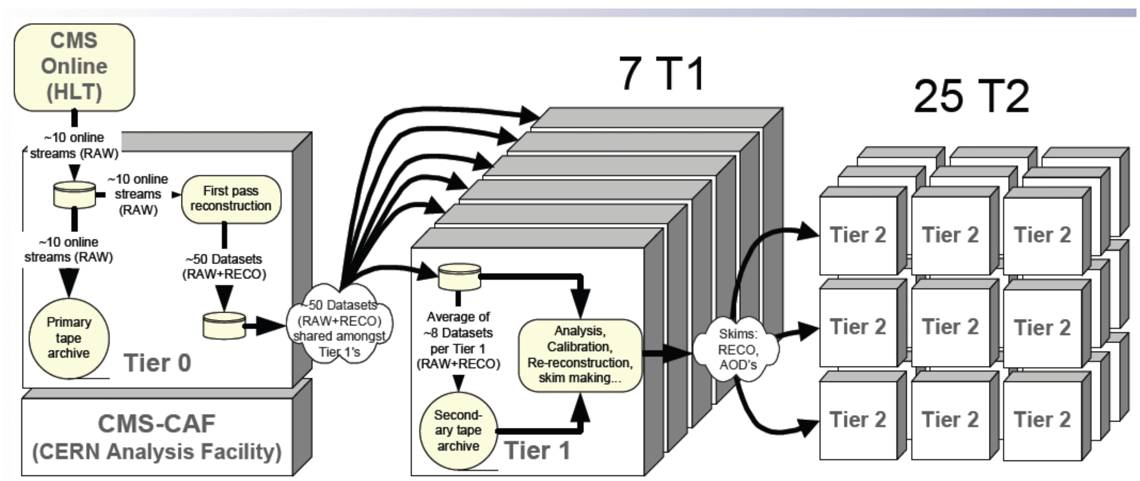

Figure 3.19: Data flow from CMS detector through hardware Tiers.

The whole collection of software built as a framework is referred to as *CMSSW*. This framework provides the services needed by the simulation, calibration and alignment, and reconstruction modules that process event data, so that physicists can perform analysis. The CMSSW event processing model is composed of one executable, called cmsRun, and several plug-in modules which contains all the tools (calibration, reconstruction algorithms) needed to process an event. The same executable is used for both detector and Monte Carlo data [79].

Chapter 4

Event generation, simulation and reconstruction

The process of analyzing data recorded by the CMS experiment involves several stages where the data are processed in order to interpret the information provided by all the detection systems; in those stages, the particles produced after the *pp* collision are identified by reconstructing their trajectories and measuring their features. In addition, the SM provides a set of predictions that have to be compared with the experimental results; however, in most of the cases, theoretical predictions are not directly comparable to experimental results due to the diverse source of uncertainties introduced by the experimental setup and theoretical approximations, among others.

The strategy to face these conditions consists in using statistical methods implemented in computational algorithms to produce numerical results that can be contrasted with the experimental results. These computational algorithms are commonly known as Monte Carlo (MC) methods and, in the case of particle physics, they are designed to apply the SM rules and produce predictions about the physical observables measured in the experiments. Since particle physics is governed by quantum mechanics principles, predictions are not allowed from single events; therefore, a high number of events are *generated* and predictions are produced in the form of statistical distributions for the observables. Effects of the detector presence are included in the predictions by introducing simulations of the detector itself.

This chapter presents a description of the event generation strategy and the tools used to perform the detector simulation and physics objects reconstruction. A comprehensive review of event generators for LHC physics can be found in Reference [80] on which this chapter is based.

4.1 Event generation

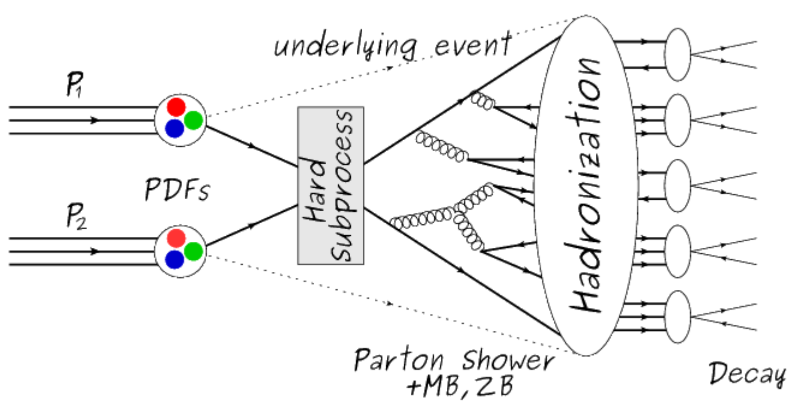

Figure 4.1: Event generation process. The actual interaction is generated in the hard subprocess. The parton shower describes the evolution of the partons from the hard subprocess.Modified from Reference [81].

The event generation is intended to create events that mimic the behavior of actual events produced in collisions; they obey a sequence of steps from the particles collision hard process to the decay process into the final state. Figure 4.1 shows a schematic view of the event generation process; the fact that the full process can be treated as several independent steps is motivated by the QCD factorization theorem.

Generation starts by taking into account the PDFs of the incoming particles. Event generators offer the option to chose from several PDF sets depending on the particular process under simulation[1]; in the following, pp collisions will be considered. The *hard subprocess* describes the actual interaction between partons from the incoming protons; it is represented by the matrix element connecting the initial and final states of the interaction. Normally, the matrix element can be written as a sum over Feynman diagrams and consider interferences between terms in the summation. During the generation of the hard subprocess, the production cross section is calculated.

The order to which the cross section is calculated depends on the order of the Feynman diagrams involved in the calculation; therefore, radiative corrections are included by considering a higher order Feynman diagrams where QCD radiation dominates. Currently, cross sections calculated to LO do not offer a satisfactory description of the processes, i.e., the results are only reliable for the shape of distributions; therefore, NLO calculations have to be performed with the implication that the computing time needed is highly increased.

The final parton content of the hard subprocess is subjected to the *parton shower* which generates the gluon radiation. Parton shower evolves the partons, i.e., glouns split into quark-antiquark pairs and quarks with enough energy radiate gluons giving rise to further parton multiplication, following the DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) equations. Showering continues until the energy scale is low enough to reach the non-perturbative limit.

[1] Tool in Reference [82] allows to plot different PDF sets under customizable conditions.

In the simulation of LHC processes that involve b quarks, like the single top quark or Higgs associated production, it is needed to consider that the b quark is heavier than the proton; hence, the QCD interaction description is made in two different schemes [83]

- four-flavor (4F) scheme. b quarks appear only in the final state because they are heavier than the proton and therefore they can be produced only from the splitting of a gluon into pairs or singly in association with a t quark in high energy-scale interactions; furthermore, during the simulation, the b-PDFs are set to zero. Calculations in this scheme are more complicated due to the presence of the second b quark but the full kinematics is considered already at LO and therefore the accuracy of the description is better.

- five-flavor (5F) scheme. b quarks are considered massless, therefore they can appear in both initial and final states since they can now be part of the proton; thus, during the simulation b-PDFs are not set to zero. In this scheme, calculations are simpler than in the 4F scheme and possible logarithmic divergences are absorbed by the PDFs through the DGLAP evolution.

In this thesis, the tHq events are generated using the 4F scheme in order to reduce uncertainties, while the tHW events are generated using the 5F scheme to eliminate LO interference with $t\bar{t}H$ process [48].

Partons involved in the pp collision are the focus of the simulation, however, the rest of the partons inside the incoming protons are also affected because the remnants are colored objects; also, multiple parton interactions can occur. The hadronization of the remnants and multiple parton interactions are known as *underlying event* and it has to be included in the simulation. In addition, multiple *pp* collisions in the same bunch crossing (pile-up mentioned in 3.2) occurs, actually in two forms

- *in-time PU* which refers to multiple *pp* collision in the bunch crossing but that are not considered as primary vertices.
- *Out-of-time PU* which refers to overlapping *pp* collisions from consecutive bunch crossings; this can occur due to the time-delays in the detection systems where information from one bunch crossing is assigned to the next or previous one.

While the underlying event effects are included in generation using generator-specific tools, PU effects are added to the generation by overlying Minimum-bias (MB) and Zero-bias (ZB) events to the generated events. MB events are inelastic events selected by using a loose trigger with as little bias as possible, therefore accepting a large fraction of the overall inelastic event; ZB events correspond to random events recorded by the detector when collisions are likely. MB models in-time PU and ZB models out-of-time PU.

The next step in the generation process is called *hadronization.* Since particles with a net color charge are not allowed to exits isolated, they have to recombine to form bound states. This is precisely the process by which the partons resulting from the parton shower arrange themselves as color singlets to form hadrons. At this step, the energy-scale is low and the strong coupling constant is large, therefore hadronization process is non-perturbative and the evolution of the partons is described using phenomenological models. Most of the baryons and mesons produced in the hadronization are unstable and hence they will decay in the detector.

The last step in the generation process corresponds to the decay of the unstable particles generated during hadronization; it is also simulated in the hadronization step, based on the known branching ratios.

4.2 Monte Carlo Event Generators.

The event generation described in the previous section has been implemented in several software packages for which a brief description is given.

- **PYTHIA 8**. It is a program designed to perform the generation of high energy physics events which describes the collisions between particles such as electrons and protons. Several theories and models are implemented in it, in order to describe physical aspects like hard and soft interaction, parton distributions, initial and final-state parton showers, multiple parton interactions, beam remnants, hadronization[2] and particle decay. Thanks to extensive testing, several optimized parametrizations, known as *tunings*, have been defined in order to improve the description of actual collisions to a high degree of precision; for analysis at $\sqrt{s} = 13$ TeV, the underline event CUETP8M1 tune is employed [85]. The calculation of the matrix element is performed at LO which is not enough for the current required level of precision; therefore, pythia is often used for parton shower, hadronization and decays, while other event generators are used to generate the matrix element at NLO.

- **MadGraph5_aMC@NLO**. MadGraph is a matrix element generator which calculates the amplitudes for all contributing Feynman diagrams of a given process but does not provide a parton shower while MC@NLO incorporates NLO

[2] based in the Lund string model [84]

QCD matrix elements consistently into a parton shower framework; thus, MadGraph5_aMC@NLO, as a merger of the two event generators MadGraph5 and aMC@NLO, is an event generator capable to calculate tree-level and NLO cross sections and perform the matching of those with the parton shower. It is one of the most frequently used matrix element generators; however, it has the particular feature of the presence of negative event weights which reduce the number of events used to reproduce the properties of the objects generated [86].

- **POWHEG**. It is an NLO matrix element generator where the hardest emission of color charged particles is generated in such a way that the negative event weights issue of MadGraph5_aMC@NLO is overcome; however, the method requires an interface with p_T-ordered parton shower or a parton shower generator where this highest emission can be vetoed in order to avoid double counting of this highest-energetic emission. PYTHIA is a commonly matched to POWHEG event generator [87].

Events resulting from the whole generation process are known as MC events.

4.3 CMS detector simulation.

After generation, MC events contain the physics of the collisions but they are not ready to be compared to the events recorded by the experiment since these recorded events correspond to the response of the detection systems to the interaction with the particles traversing them. The simulation of the CMS detector has to be applied on top of the event generation; it is simulated with a MC toolkit for the simulation of particles passing through matter called Geant4 which is also able to simulate the electronic signals that would be measured by all detectors inside CMS.

The simulation takes the generated particles contained in the MC events as input, makes them pass through the simulated geometry, and models physics processes that particles experience during their passage through matter. The full set of results from particle-matter interactions corresponds to the simulated hit which contains information about the energy loss, momentum and position. Particles of the input event are called *primary*, while the particles originating from GEANT4-modeled interactions of a primary particle with matter are called a *secondary*. Simulated hits are the input of subsequent modules that emulate the response of the detector readout system and triggers. The output from the emulated detection systems and triggers is known as digitization [88, 89].

The modeling of the CMS detector corresponds to the accurate modeling of the interaction among particles, the detector material, and the magnetic field. This simulation procedure includes the following standard steps

- Modeling of the Interaction Region.
- Modeling of the particle passage through the hierarchy of volumes that compose CMS detector and of the accompanying physics processes.
- Modeling of the effect of multiple interactions per beam crossing and/or the effect of events overlay (Pile-Up simulation).
- Modeling of the detector's electronics response, signal shape, noise, calibration constants (digitization).

In addition to the full simulation, i.e., a detailed detector simulation, a faster simulation (FastSim) have been developed, that may be used where much larger statistics are required. In FastSim, detector material effects are parametrized and included in the hits; those hits are used as input of the same higher-level algorithms[3] used to analyze the recorded events. In this way, comparisons between fast and full simulations can be performed [91].

After the full detector simulation, the output events can be directly compared to events actually recorded in the CMS detector. The collection of MC events that reproduces the expected physics for a given process isknown as MC sample.

4.4 Event reconstruction.

The CMS experiment use the *particle-flow event reconstruction algorithm (PF)* to do the reconstruction of particles produced in *pp* collisions. Next sections will present a basic description of the *Elements* used by PF (tracker tracks, energy clusters, and muon tracks), based in the References [92, 93] where more detailed descriptions can be found.

4.4.1 Particle-Flow Algorithm.

Each of the several sub detection systems of the CMS detector is dedicated to identify an specific type of particles, i.e., photons and electrons are absorbed by the ECAL and their reconstruction is based on ECAL information; hadrons are reconstructed from clusters in the HCAL while muons are reconstructed from hits in the muon

[3] track fitting, calorimeter clustering, b tagging, electron identification, jet reconstruction and calibration, trigger algorithms which will be considered in the next sections

chambers. PF is designed to correlate signals from all the detector layers (tracks and energy clusters) in order to reconstruct and identify each final state particle and its properties as sketched in Figure 4.2.

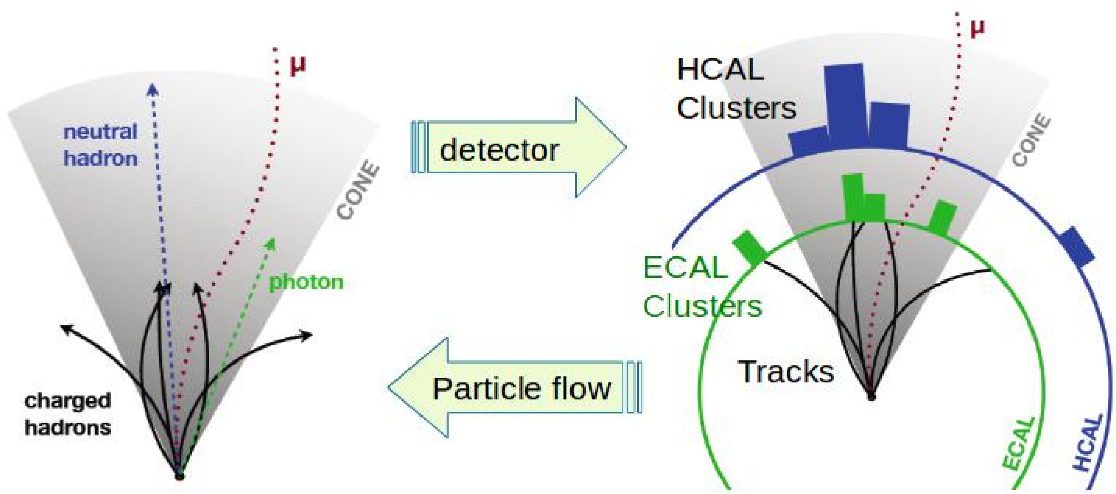

Figure 4.2: Particle flow algorithm. Information from the several CMS detection systems if provided as input to the algorithm which then combine it to identify and reconstruct all the particles in the final state and their properties. Reconstruction of simulated events is also performed by providing information from MC samples, detector and trigger simulation [94].

For instance, a charged hadron is identified by a geometrical connection, known as *link*, between one or more calorimeter clusters and a track in the tracker, provided there are no hits in the muon system; combining several measurements allows a better determination of the energy and charge sign of the charged hadron.

Charged-particle track reconstruction.

The strategy used by PF in order to reconstruct tracks is called *Iterative Tracking* which occurs in four steps

- Seed generation where initial track candidates are found by looking for a combination of hits in the pixel detector, strip tracker, and muon chambers. In total ten iterations are performed, each one with a different seeding requirement. Seeds are used to estimate the trajectory parameters and uncertainties at the time of the full track reconstruction. Seeds are also considered track candidates.

- Track finding using a tracking software known as Combinatorial Track Finder (CTF) [95]. The seed trajectories are extrapolated along the expected flight path of a charged particle, in agreement to the trajectory parameters obtained in the first step, in an attempt to find additional hits that can be assigned to the track candidates.

- Track-fitting where the found tracks are passed as input to a module which provides the best estimate of the parameters of each trajectory.

- Track selection where track candidates are submitted to a selection which discards those that fail a set of defined quality criteria.

Iterations differ in the seeding configuration and the final track selection as elaborated in References [92, 93]. In the first iteration, high p_T tracks and tracks produced near to the interaction region are identified and those hits are masked thereby reducing the combinatorial complexity. Next, iterations search for more complicated tracks, like low p_T tracks and tracks from b hadron decays, which tend to be displaced from the interaction region.

Vertex reconstruction.

During the track reconstruction, an extrapolation toward to the calorimeters is performed in order to match energy deposits; that extrapolation is performed also toward the beamline in order to find the origin of the track known as *vertex*. The vertex reconstruction is performed by selecting from the available reconstructed tracks, those that are consistent with being originated in the interaction region where pp collisions are produced. The selection involves a requirement on the number of tracker (pixel and strip) hits and the goodness of the track fit.

Selected tracks are clustered using a *deterministic annealing algorithm (DA)*[4]. A set of candidate vertices and their associated tracks, resulting from the DA, are then fitted with an *adaptive vertex fitter (AVF)* to produce the best estimate of the vertices locations.

The p_T of the tracks associated to a reconstructed vertex is added, squared and used to organize the vertices; the vertex with the highest squared sum is designated as the *primary vertex (PV)* while the rest are designated as PU vertices.

Calorimeter clustering.

After traversing the CMS tracker system, electrons, photons and hadrons deposit their energy in the ECAL and HCAL cells. The PF clustering algorithm aims to provide a high detection efficiency even for low-energy particles and an efficient distinction between close energy deposits. The clustering runs independently in the ECAL barrel and endcaps, HCAL barrel and endcaps, and the two preshower layers, following two steps

- cells with an energy larger than a given seed threshold and larger than the energy of the neighboring cells are identified as cluster seeds. The neighbor cells are those that either share a side with the cluster seed candidate, or the eight closest cells including cells that only share a corner with the seed candidate.

[4] DA algorithm and AVF are described in detail in References [97, 98]

- cells with at least a corner in common with a cell already in the cluster seed and with an energy above a cell threshold are grouped into topological clusters.

Clusters formed in this way are known as *particle-flow clusters*. With this clustering strategy, it is possible to detect and measure the energy and direction of photons and neutral hadrons as well as differentiate these neutral particles from the charged hadron energy deposits. In cases involving charged hadrons for which the track parameters are not determined accurately, for instance, low-quality and high-p_T tracks, clustering helps in the energy measurements.

Electron track reconstruction.

Although the charged-particle track reconstruction described above works for electrons, they lose a significant fraction of their energy via bremsstrahlung photon radiation before reaching the ECAL; thus, the reconstruction performance depends on the ability to measure also the radiated energy. The reconstruction strategy, in this case, requires information from the tracking system and from the ECAL. Bremsstrahlung photons are emitted at similar η values to that of the electron but at different values of ϕ; therefore, the radiated energy can be recovered by grouping ECAL clusters in a η window over a range of ϕ around the electron direction. The group is called ECAL supercluster.

Electron candidates from the track-seeding and ECAL super clustering are merged into a single collection which is submitted to a full electron tracking fit with a Gaussian-sum filter (GSF) [96]. The electron track and its associated ECAL supercluster form a *particle-flow electron*.

Muon track reconstruction.

Given that the CMS detector is equipped with a muon spectrometer capable to identify and measure the momentum of the muons traversing it, the muon reconstruction is not specific to PF; therefore, three different muon types are defined

- *Standalone muon.* A clustering on the DTs or CSCs hits is performed to form track segments; those segments are used as seeds for the reconstruction in the muon spectrometer. All DTs, CSCs, and RPCs hits along the muon trajectory are combined and fitted to form the full track. The fitting output is called a *standalone-muon track*.

- *Tracker muon.* Each track in the inner tracker with p_T larger than 0.5 GeV and a total momentum p larger than 2.5 GeV is extrapolated to the muon system. A *tracker muon track* corresponds to a extrapolated track that matches at least one muon segment.

- *Global muon.* When tracks in the inner tracker (inner tracks) and standalone-muon tracks are matched and turn out being compatibles, their hits are combined and fitted to form a *global-muon track*.

Global muons sharing the same inner track with tracker muons are merged into a single candidate. PF muon identification uses the muon energy deposits in ECAL, HCAL, and HO associated with the muon track to improve the muon identification.

Particle identification and reconstruction.

PF elements are connected by a linker algorithm that tests the connection between any pair of elements; if they are found to be linked, a geometrical distance that quantifies the quality of the link is assigned. Two elements may be linked indirectly through common elements. Linked elements form *PF blocks* and each PF block may contain elements originating in one or more particles. Links can be established between tracks, between calorimeter clusters, and between tracks and calorimeter clusters. The identification and reconstruction start with a PF block and proceed as follows

- Muons. An *isolated global muon* is identified by evaluating the presence of inner track and energy deposits close to the global muon track in the (η,ϕ) plane, i.e., in a particular point of the global muon track, inner tracks and energy deposits are sought within a radius of $\Delta R = 0.3$ (see eqn. 3.7) from the muon track; if they exit and the p_T of the found track added to the E_T of the found energy deposit does not exceed 10% of the muon p_T then the global muon is an isolated global muon. This isolation condition is stringent enough to reject hadrons misidentified as muons.

  Non-isolated global muons are identified using additional selection requirements on the number of track segments in the muon system and energy deposits along the muon track. Muons inside jets are identified with more stringent criteria in isolation and momentum as described in Reference [99]. The PF elements associated with an identified muon are masked from the PF block.

- Electrons are identified and reconstructed as described above plus some additional requirements on fourteen variables like the amount of energy radiated, the distance between the extrapolated track position at the ECAL and the position of the associated ECAL supercluster, among others, which are combined in an specialized multivariate analysis strategy that improves the electron identification. Tracks and clusters used to identify and reconstruct electrons are masked in the PF block.

- Isolated photons are identified from ECAL superclusters with E_T larger than 10 GeV, for which the energy deposited at a distance of 0.15, from the supercluster position on the (η,ϕ) plane, does not exceed 10% of the supercluster energy; note that this is an isolation requirement. In addition, there must not be links to tracks. Clusters involved in the identification and reconstruction are masked in the PF block.

- Bremsstrahlung photons and prompt photons tend to convert to electron-positron pairs inside the tracker, therefore, a dedicated finder algorithm is used to link tracks that seem to originate from a photon conversion; in case those two tracks are compatible with the direction of a bremsstrahlung photon, they are also linked to the original electron track. Photon conversion tracks are also masked in the PF block.

- The remaining elements in the PF block are used to identify hadrons. In the region $|\eta| \leq 2.5$, neutral hadrons are identified with HCAL clusters not linked to any track while photons from neutral pion decays are identified with ECAL clusters without links to tracks. In the region $|\eta| > 2.5$ ECAL clusters linked to HCAL clusters are identified with a charged or neutral hadron shower; ECAL clusters with no links are identified with photons. HCAL clusters not used yet, are linked to one or more unlinked tracks and to an unlinked ECAL in order to reconstruct charged-hadrons or a combination of photons and neutral hadrons according to certain conditions on the calibrated calorimetric energy.

- Charged-particle tracks may be liked together when they converge to a *secondary vertex (SV)* displaced from the interaction point where the PV and PU vertices are reconstructed; at least three tracks are needed in that case, of which at most one has to be an incoming track with hits in tracker region between a

PV and the SV.

The linker algorithm, as well as the whole PF algorithm, has been validated and commissioned; results from that validation are presented in the Reference [92].

Jet reconstruction.

Quarks and gluons may be produced in the *pp* collisions, therefore, their hadronization will be seen in the detector as a shower of hadrons and their decay products in the form of a *jet*. The anti-k_t algorithm [100] is used to perform the jet reconstruction by clustering those PF particles within a cone (see Figure 4.3); previously, isolated electrons, isolated muons, and charged particles associated with other interaction vertices are excluded from the clustering.
The anti-k_t algorithm proceeds in a sequential recombination of PF particles; the distance between particles i and j (d_{ij}) and the distance between particles and the beam are defined as

$$\begin{aligned} d_{ij} =& \min\left(\frac{1}{k_{ti}^2}, \frac{1}{k_{tj}^2}\right)\frac{\Delta_{ij}^2}{R^2} \\ d_{iB} =& \frac{1}{k_{ti}^2} \end{aligned} \tag{4.1}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, k_{ti}, y_i and ϕ_i are the transverse momentum, rapidity and azimuth of particle i respectively and R is the called jet radius. For all the remaining PF particles, after removing the isolated ones, d_{ij} and d_{iB} are calculated[5] and the smallest is identified; if it is a d_{ij}, particles i and j are replaced with

[5] Notice that this is a combinatorial calculation.

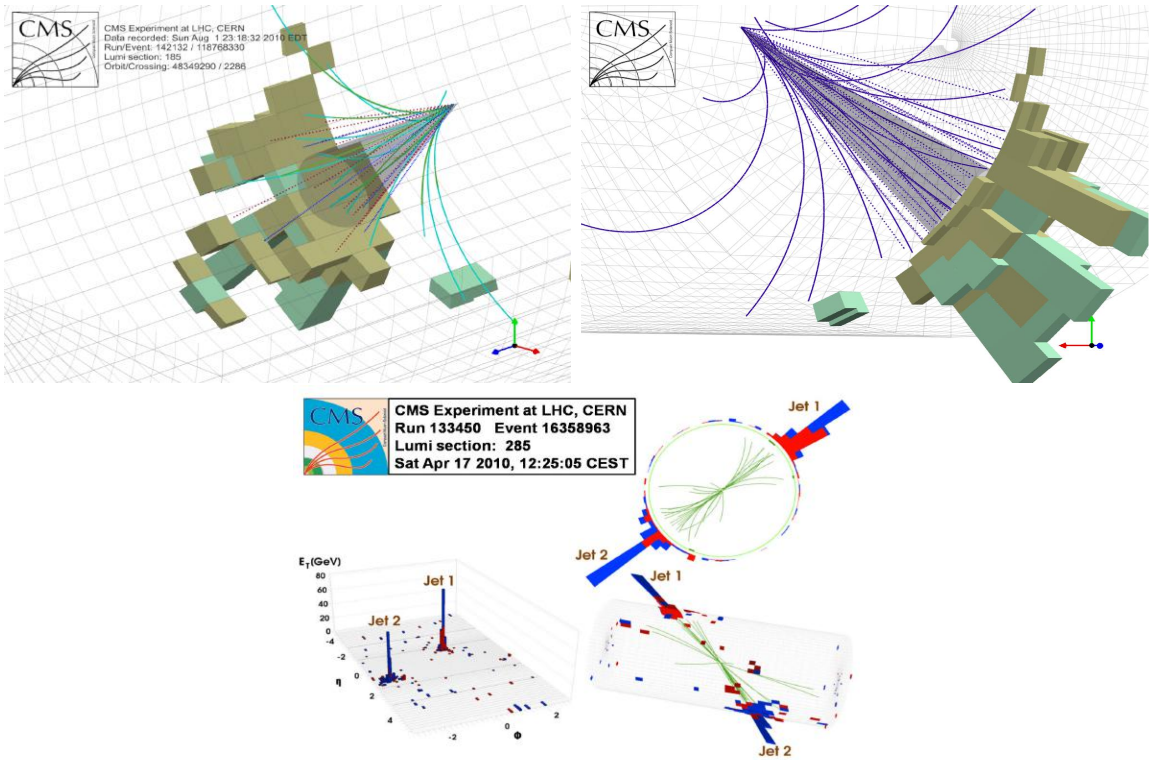

Figure 4.3: Jet reconstruction performed by the anti-k_t algorithm. Top: Two different views of a CMS recorded event are presented. Continuous lines correspond to tracks left by charged particles in the tracker while dotted lines are the imaginary paths followed by neutral particles. The green cubes represent the ECAL cells while the blue ones represent the HCAL cells; in both cases, the height of the cube represent the amount of energy deposited in the cells [101]. Bottom: Reconstruction of a recorded event with two jets [102].

a new object whose momentum is the vectorial sum of the combined particles. If the smallest distance is a d_{iB} the clustering process ends, the object i (which at this stage should be a combination of several PF particles) is declared as a *Particle-flow-jet* (PF jet) and all the associated PF particles are removed from the detector. The clustering process is repeated until no PF particles remain.

Even though jets can be reconstructed efficiently, there are some effects that are not included in the reconstruction and that lead to discrepancies between the reconstructed results and the predicted results; in order to overcome these discrepancies, a factorized model has been designed in the form of jet energy corrections (JEC) [103, 104]

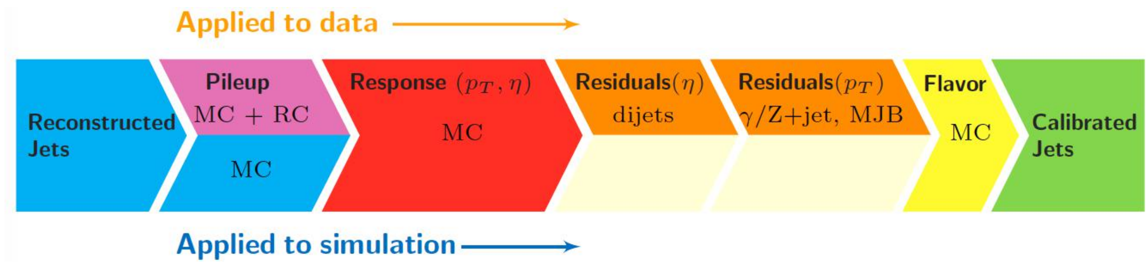

Figure 4.4: Jet energy correction diagram. Correction levels are applied sequentially in the indicated fixed order [104].

applied sequentially as shown in the diagram of Figure 4.4.

At each level, the jet four-momentum is multiplied by a scaling factor based on jet properties, i.e., η, flavor, etc.

- Level 1 correction removes the energy coming from pile-up. The scale factor is determined using a MC sample of QCD dijet (2 jets) events with and without pileup overlay; it is parametrized in terms of the offset energy density ρ, jet area A, jet η and jet p_T. Different corrections are applied to data and MC due to the detector simulation.

- MC-truth correction accounts for differences between the reconstructed jet energy and the MC particle-level energy. The correction is determined on a QCD dijet MC sample and is parametrized in terms of the jet p_T and η.

- Residuals correct remaining small differences within jet response in data and MC. The Residuals η-dependent correction compares jets of similar p_T in the barrel reference region. The Residuals p_T-dependent correct the jet absolute scale (JES vs p_T).

- Jet-flavor corrections are derived in the same way as MC-truth corrections but using QCD pure flavor samples.

b-tagging of jets.

A particular feature of the hadrons containing bottom quarks (b-hadrons) is that their lifetime is long enough to travel some distance before decaying, but it is not as long as those of light quark hadrons; therefore, when looking at the hadrons produced in *pp* collisions, b-hadrons decay typically inside the tracker rather than reaching the calorimeters as some light-hadrons do. As a result, a b-hadron decay gives rise to a displaced vertex (secondary vertex) with respect to the primary vertex as shown in Figure 4.5; the SV displacement is in the order of a few millimeters. A jet resulting from the decay of a b-hadron is called b jet; other jets are called light jets.

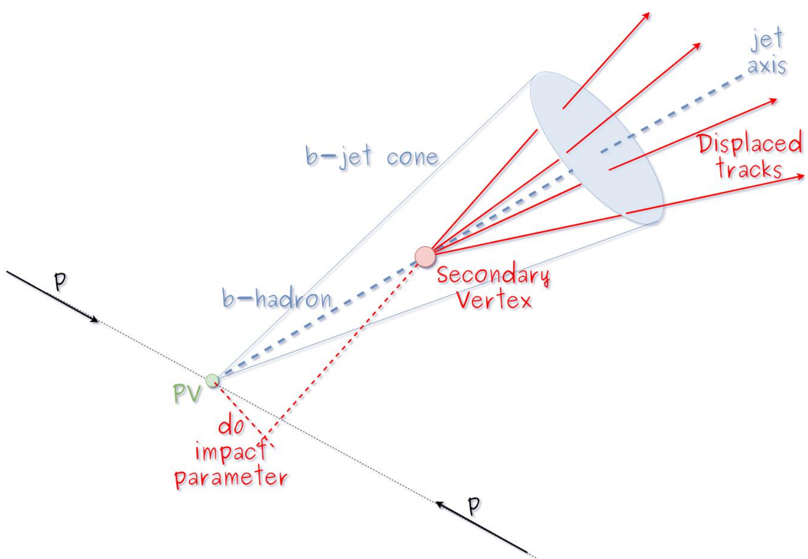

Figure 4.5: Secondary vertex in a b-hadron decay.

Several methods to identify b-jets (b-tagging) have been developed; the method used in this thesis is known as *Combined Secondary Vertex* algorithm in its second version (CSVv2) [105]. By using information of the impact parameter, the reconstructed secondary vertices, and the jet kinematics as input in a multivariate analysis that combines the discrimination power of each variable in one global discriminator variable, three working points (references): loose, medium and tight, are defined which quantify the probabilities of mistag jets from light quarks as jets from b quarks; 10, 1 and 0.1 % respectively. Although the mistagging probability decreases with the working point strength, the efficiency to correctly tag b-jets also decreases as 83, 69 and 49 % for the respective working point; therefore, a balance needs to be achieved according to the specific requirements of the analysis.

4.4.1.1 Missing transverse energy.

The fact that proton bunches carry momentum along the $z-$axis implies that for each event it is expected that the momentum in the transverse plane is balanced. Imbalances are quantified by the missing transverse energy (MET) and are attributed to several sources including particles escaping undetected through the beam pipe, neutrinos produced in weak interactions processes which do not interact with the detector and thus escaping without leaving a sign, or even undiscovered particles predicted by models beyond the SM.

The PF algorithm assigns the negative sum of the momenta of all reconstructed PF particles to the *particle-flow MET* according to

$$\vec{\not{E}}_T = -\sum_i \vec{p}_{T,i} \tag{4.2}$$

JEC are propagated to the calculation of the $\vec{\not{E}}_T$ as described in the Reference [106].

4.4.2 Event reconstruction examples

Figures 4.6-4.8 show the results of the reconstruction performed on 3 recorded events. Descriptions are taken directly from the source.

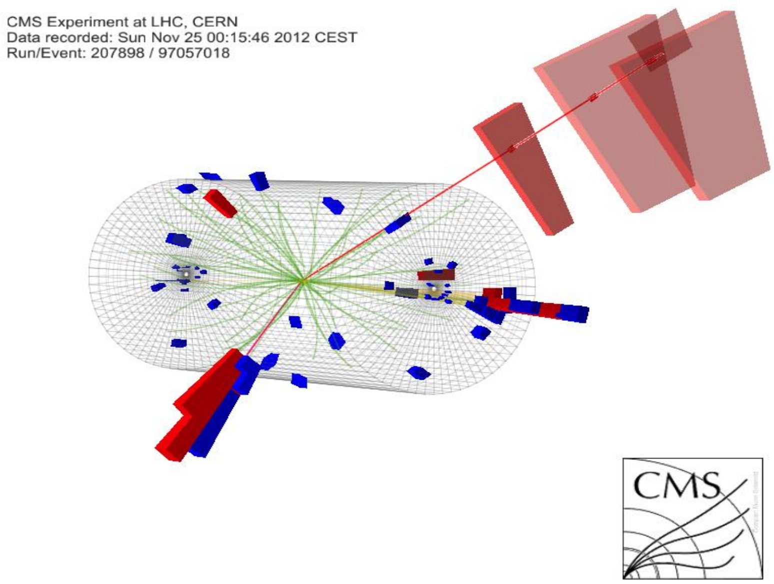

Figure 4.6: HIG-13-004 Event 1 reconstruction results; "HIG-13-004 Event 1: Event recorded with the CMS detector in 2012 at a proton-proton center-of-mass energy of 8 TeV. The event shows characteristics expected from the decay of the SM Higgs boson to a pair of τ leptons. Such an event is characterized by the production of two forward-going jets, seen here in opposite endcaps. One of the τ decays to a muon (red lines on the right) and neutrinos, while the other τ decays into a charged hadron and a neutrino." [107].

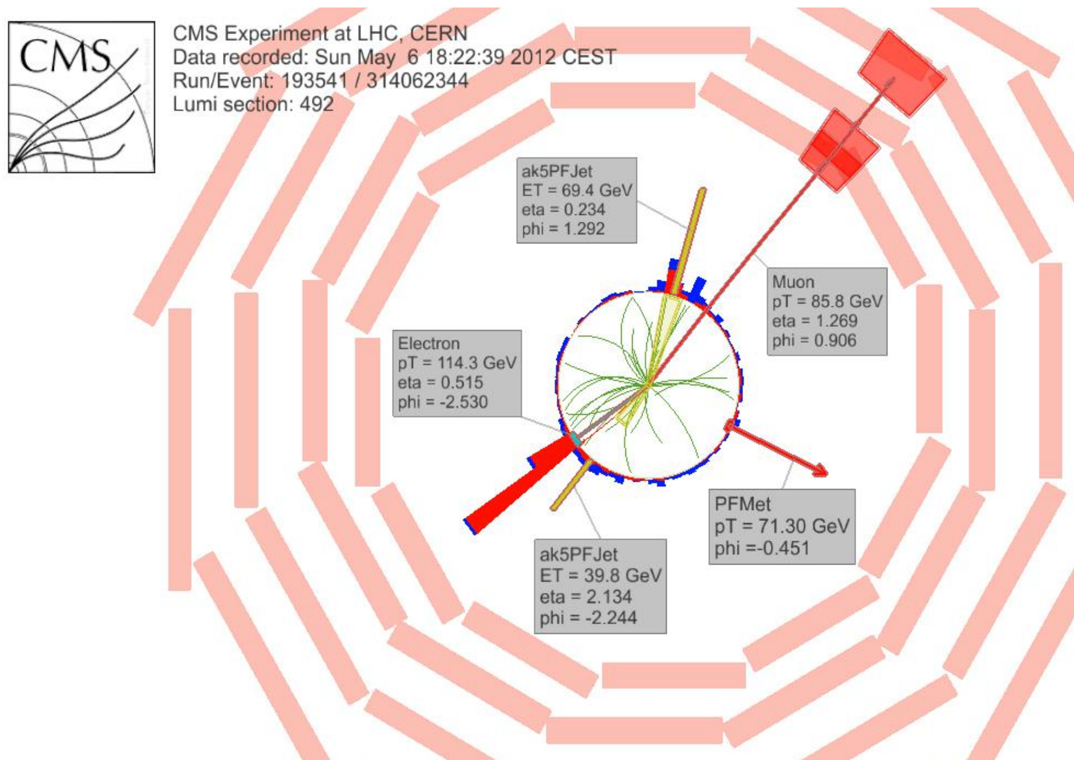

Figure 4.7: $e\mu$ event reconstruction results;"An $e\mu$ event candidate selected in 8 TeV data, as seen from the direction of the proton beams. The kinematics of the main objects used in the event selection are highlighted: two isolated leptons and two particle-flow jets. The reconstructed missing transverse energy is also displayed for reference" [108].

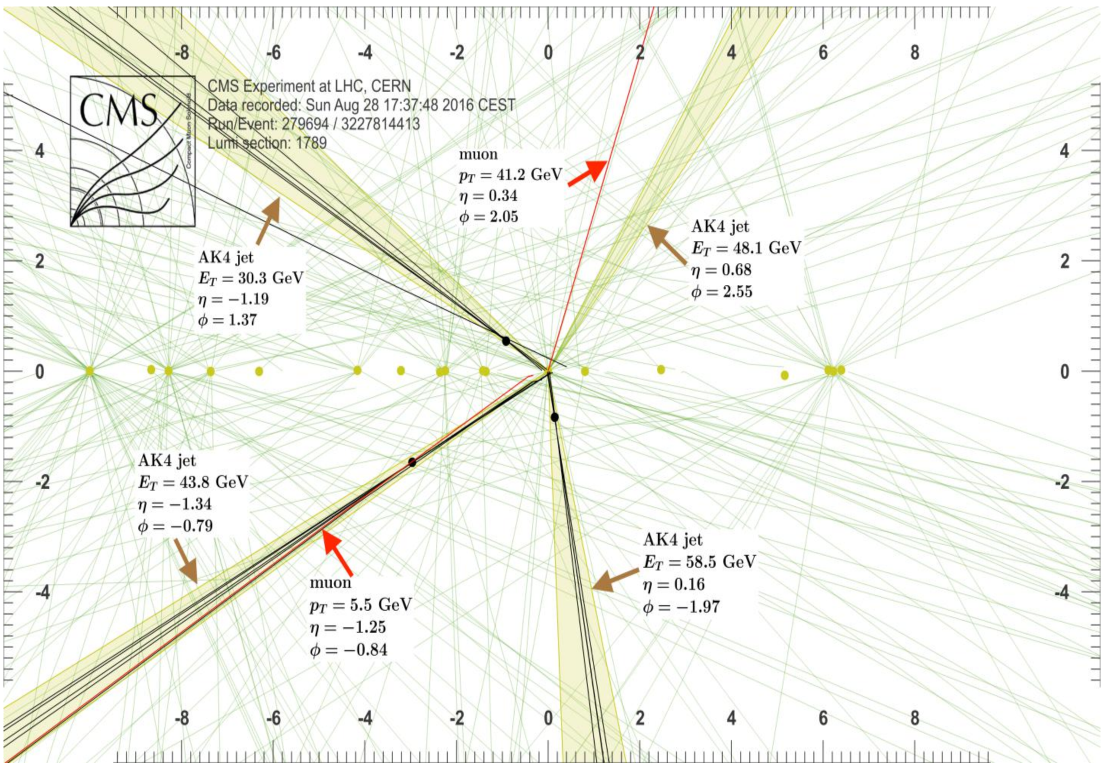

Figure 4.8: Recorded event reconstruction results;"Recorded event (ρ-z projection) with three jets with $p_T > 30$ GeV with one displaced muon track in 2016 data collected at 13 TeV. Each of the three jets has a displaced reconstructed vertex. The jet with $p_T(j) = 43.8$ GeV, $\eta(j) = -1.34$, $\phi(j) = -0.79$ contains muon with $p_T(\mu) = 5.5$ GeV, $\eta(\mu) = -1.25$, $\phi(\mu) = -0.84$. Event contains reconstructed isolated muon with $p_T(\mu) = 41.2$ GeV, $\eta(\mu) = 0.34$, $\phi(\mu) = 2.05$ and MET with $p_T = 72.5$ GeV, $\phi = -0.32$. Jet candidates for a b-jet from top quark leptonic and hadronic decays are tagged by CSVv2T algorithm. One of the other two jets is tagged by CharmT algorithm. Tracks with $p_T > 0.5$ GeV are shown. The number of reconstructed primary vertices is 18. Reconstructed m_T(W) is 101.8 GeV. Beam spot position correction is applied. Reconstructed primary vertices are shown in yellow color, while reconstructed displaced vertices and associated tracks are presented in black color. Dimensions are given in cm" [109].

References

[1] J. Schwinger. “Quantum Electrodynamics. I. A Covariant Formulation”. Physical Review. 74 (10): 1439-61, (1948).

[2] R. P. Feynman. “Space-Time Approach to Quantum Electrodynamics”. Physical Review. 76 (6): 769-89, (1949).

[3] S. Tomonaga. “On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields”. Progress of Theoretical Physics. 1 (2): 27-42, (1946).

[4] D.J. Griffiths, “Introduction to electrodynamics”. 4th ed. Pearson, (2013).

[5] F. Mandl, G. Shaw. “Quantum field theory.” Chichester: Wiley (2009).

[6] F. Halzen, and A.D. Martin, “Quarks and leptons: An introductory course in modern particle physics”. New York: Wiley, (1984) .

[7] File: Standard_Model_of_Elementary_Particle_dark.svg. (2017, June 12) Wikimedia Commons, the free media repository. Retrieved November 27, 2017 from https://www.collegiate-advanced-electricity.com/single-post/2017/04/10/The-Standard-Model-of-Particle-Physics.

[8] E. Noether, “Invariante Variationsprobleme”, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse, vol. 1918, pp. 235-257, (1918).

[9] C. Patrignani et al. (Particle Data Group), Chin. Phys. C, 40, 100001 (2016) and 2017 update.

[10] M. Goldhaber, L. Grodzins, A.W. Sunyar “Helicity of Neutrinos”, Phys. Rev. 109, 1015 (1958).

[11] Palanque-Delabrouille N et al. “Neutrino masses and cosmology with Lyman-alpha forest power spectrum”, JCAP 11 011 (2015).

[12] M. Gell-Mann. “A Schematic Model of Baryons and Mesons”. Physics Letters. 8 (3): 214-215 (1964).

[13] G. Zweig. “An SU(3) Model for Strong Interaction Symmetry and its Breaking” (PDF). CERN Report No.8182/TH.401 (1964).

[14] G. Zweig. “An SU(3) Model for Strong Interaction Symmetry and its Breaking: II” (PDF). CERN Report No.8419/TH.412(1964).

[15] M. Gell-Mann. “The Interpretation of the New Particles as Displaced Charged Multiplets”. Il Nuovo Cimento 4: 848. (1956).

[16] T. Nakano, K, Nishijima. “Charge Independence for V-particles”. Progress of Theoretical Physics 10 (5): 581-582. (1953).

[17] N. Cabibbo, “Unitary symmetry and leptonic decays” Physical Review Letters, vol. 10, no. 12, p. 531, (1963).

[18] M.Kobayashi, T.Maskawa, “CP-violation in the renormalizable theory of weak interaction,” Progress of Theoretical Physics, vol. 49, no. 2, pp. 652-657, (1973).

[19] File: Weak Decay (flipped).svg. (2017, June 12). Wikimedia Commons, the free media repository. Retrieved November 27, 2017 from

https://commons.wikimedia.org/w/index.php?title=File:Weak_Decay_(flipped).svg&oldid=247498592.

[20] Georgia Tech University. Coupling Constants for the Fundamental Forces(2005). Retrieved January 10, 2018, from http://hyperphysics.phy-astr.gsu.edu/hbase/Forces/couple.html#c2

[21] M. Strassler. (May 31, 2013).The Strengths of the Known Forces. Retrieved January 10, 2018, from https://profmattstrassler.com/articles-and-posts/particle-physics-basics/the-known-forces-of-nature/the-strength-of-the-known-forces/

[22] S.L. Glashow. “Partial symmetries of weak interactions”, Nucl. Phys. 22 579-588, (1961).

[23] A. Salam, J.C. Ward. “Electromagnetic and weak interactions”, Physics Letters 13 168-171, (1964).

[24] S. Weinberg, “A model of leptons”, Physical Review Letters, vol. 19, no. 21, p. 1264, (1967).

[25] M. Peskin, D. Schroeder, “An introduction to quantum field theory”. Perseus Books Publishing L.L.C., (1995).

[26] A. Pich. “The Standard Model of Electroweak Interactions” https://arxiv.org/abs/1201.0537

[27] F.Bellaiche. (2012, 2 September). “What’s this Higgs boson anyway?”. Retrieved from: https://www.quantum-bits.org/?p=233

[28] M. Endres et al. Nature 487, 454-458 (2012) doi:10.1038/nature11255

[29] F. Englert, R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. Physical Review Letters. 13 (9): 321-23.(1964) doi:10.1103/PhysRevLett.13.321

[30] P.Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. Physical Review Letters. 13 (16): 508-509,(1964). doi:10.1103/PhysRevLett.13.508.

[31] G.Guralnik, C.R. Hagen and T.W.B. Kibble. “Global Conservation Laws and Massless Particles”. Physical Review Letters. 13 (20): 585-587, (1964). doi:10.1103/PhysRevLett.13.585.

[32] CMS collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. Physics Letters B. 716 (1): 30-61 (2012). arXiv:1207.7235. doi:10.1016/j.physletb.2012.08.021

[33] ATLAS collaboration. “Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC”. Physics Letters B. 716 (1): 1-29 (2012). arXiv:1207.7214. doi:10.1016/j.physletb.2012.08.020.

[34] ATLAS collaboration; CMS collaboration (26 March 2015). “Combined Measurement of the Higgs Boson Mass in pp Collisions at âĹŽs=7 and 8 TeV with the ATLAS and CMS Experiments”. Physical Review Letters. 114 (19): 191803. arXiv:1503.07589. doi:10.1103/PhysRevLett.114.191803.

[35] LHC InternationalMasterclasses“When protons collide”. Retrieved from `http://atlas.physicsmasterclasses.org/en/zpath_protoncollisions.htm`

[36] CMS Collaboration, “SM Higgs Branching Ratios and Total Decay Widths (update in CERN Report4 2016)”. `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR` , last accessed on 17.12.2017.

[37] R.Grant V. "Determination of Higgs branching ratios in $H \to W^+W^- \to l\nu jj$ and $H \to ZZ \to l^+l^-jj$ channels". Physics Department, University of Tennessee (Dated: October 31, 2012). Retrieved from http://aesop.phys.utk.edu/ph611/2012/projects/Riley.pdf

[38] LHC Higgs Cross Section Working Group, Denner, A., Heinemeyer, S. et al. "Standard model Higgs-boson branching ratios with uncertainties". Eur. Phys. J. C (2011) 71: 1753. https://doi.org/10.1140/epjc/s10052-011-1753-8

[39] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, "Associated production of Higgs and single top at hadron colliders", Phys.Rev. D64 (2001) 094023, [hep-ph/0106293].

[40] S. Biswas, E. Gabrielli, F. Margaroli, and B. Mele, "Direct constraints on the top-Higgs coupling from the 8 TeV LHC data," Journal of High Energy Physics, vol. 07, p. 073, (2013).

[41] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, and A. Thamm, "Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson," Journal of High Energy Physics, vol. 05, p. 022, (2013).

[42] T.M. Tait and C.-P. Yuan, "Single top quark production as a window to physics beyond the standard model", Phys. Rev. D 63 (2000) 014018 [hep-ph/0007298].

[43] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, "Higgs production in association with a single top quark at the LHC," European Physical Journal C, vol. 75, p. 267, (2015).

[44] CMS Collaboration, "Modelling of the single top-quark production in association with the Higgs boson at 13 TeV." `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopHiggsGeneration13TeV`, last accessed on 16.01.2018.

[45] CMS Collaboration, "SM Higgs production cross sections at $\sqrt{s}$ = 13 TeV." `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV`, last accessed on 16.01.2018.

[46] S. Dawson, The effective W approximation, Nucl. Phys. B 249 (1985) 42.

[47] S. Biswas, E. Gabrielli and B. Mele, JHEP 1301 (2013) 088 [arXiv:1211.0499 [hep-ph]].

[48] F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, "tWH associated production at the LHC", European Physical Journal C, vol. 77, p. 34, (2017). arXiv:1607.05862

[49] LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs Cross Sections: 4.Deciphering the Nature of the Higgs Sector", arXiv:1610.07922.

[50] J. Ellis, D. S. Hwang, K. Sakurai, and M. Takeuchi."Disentangling Higgs-Top Couplings in Associated Production", JHEP 1404 (2014) 004, [arXiv:1312.5736].

[51] CMS Collaboration, V. Khachatryan et al., "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV," arXiv:1412.8662.

[52] ATLAS Collaboration, G. Aad et al., "Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fbâĹŠ1 of proton-proton collision data", ATLAS-CONF-2014-009.

[53] ATLAS and CMS Collaborations, “Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV,” (2016). CERN-EP-2016-100, ATLAS-HIGG-2015-07, CMS-HIG-15-002.

[54] File:Cern-accelerator-complex.svg. Wikimedia Commons, the free media repository. Retrieved January, 2018 from https://commons.wikimedia.org/wiki/File:Cern-accelerator-complex.svg

[55] J.L. Caron , “Layout of the LEP tunnel including future LHC infrastructures.”, (Nov, 1993). A C Collection. Legacy of AC. Pictures from 1992 to 2002. Retrieved from https://cds.cern.ch/record/841542

[56] M. Vretenar, “The radio-frequency quadrupole”. CERN Yellow Report CERN-2013-001, pp.207-223 DOI:10.5170/CERN-2013-001.207. arXiv:1303.6762

[57] L.Evans. P. Bryant (editors). “LHC Machine”. JINST 3 S08001 (2008).

[58] CERN Photographic Service.“Radio-frequency quadrupole, RFQ-1”, March 1983, CERN-AC-8303511. Retrieved from https://cds.cern.ch/record/615852.

[59] CERN Photographic Service “Animation of CERN’s accelerator network”, 14 October 2013. DOI: 10.17181/cds.1610170 Retrieved from https://videos.cern.ch/record/1610170

[60] C.Sutton. “Particle accelerator”.Encyclopedia Britannica. July 17, 2013. Retrieved from https://www.britannica.com/technology/particle-accelerator.

[61] L.Guiraud. “Installation of LHC cavity in vacuum tank.”. July 27 2000. CERN-AC-0007016. Retrieved from https://cds.cern.ch/record/41567.

[62] J.L. Caron, “Magnetic field induced by the LHC dipole’s superconducting coils”. March 1998. AC Collection. Legacy of AC. Pictures from 1992 to 2002. LHC-PHO-1998-325. Retrieved from https://cds.cern.ch/record/841511

[63] AC Team. “Diagram of an LHC dipole magnet”. June 1999. CERN-DI-9906025 retrieved from https://cds.cern.ch/record/40524.

[64] CMS Collaboration “Public CMS Luminosity Information”. https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2016 _proton_proton_13_TeV_collis, last accessed 24.01.2018

[65] J.L Caron. “LHC Layout” AC Collection. Legacy of AC. Pictures from 1992 to 2002. September 1997, LHC-PHO-1997-060. Retrieved from https://cds.cern.ch/record/841573.

[66] J.A. Coarasa. “The CMS Online Cluster:Setup, Operation and Maintenance of an Evolving Cluster”. ISGC 2012, 26 February - 2 March 2012, Academia Sinica, Taipei, Taiwan.

[67] CMS Collaboration. “The CMS experiment at the CERN LHC” JINST 3 S08004 (2008).

[68] CMS Collaboration. “CMS detector drawings 2012” CMS-PHO-GEN-2012-002. Retrieved from http://cds.cern.ch/record/1433717.

[69] R. Breedon. “View through the CMS detector during the cooldown of the solenoid on February 2006. CMS Collection”, February 2006, CMS-PHO-OREACH-2005-004, Retrieved from https://cds.cern.ch/record/930094.

[70] A. Dominguez et. al. “CMS Technical Design Report for the Pixel Detector Upgrade”, CERN-LHCC-2012-016. CMS-TDR-11.

[71] CMS Collaboration. “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” Journal of Instrumentation, vol. 9, no. 10, p. P10009,(2014).

[72] CMS Collaboration and M. Brice. “Images of the CMS Tracker Inner Barrel”, November 2008, CMS-PHO-TRACKER-2008-002. Retrieved from https://cds.cern.ch/record/1431467.

[73] M. Weber. “The CMS tracker”. 6th international conference on hyperons, charm and beauty hadrons Chicago, June 28-July 3 2004.

[74] CMS Collaboration. “Projected Performance of an Upgraded CMS Detector at the LHC and HL-LHC: Contribution to the Snowmass Process”. Jul 26, 2013. arXiv:1307.7135

[75] L. Veillet. “End assembly of HB with EB rails and rotation inside SX ”,January 2002. CMS-PHO-HCAL-2002-002. Retrieved from https://cds.cern.ch/record/42594.

[76] J. Puerta-Pelayo.“First DT+RPC chambers installation round in the UX5 cavern.”. January 2007, CMS-PHO-OREACH-2007-001. Retrieved from https://cds.cern.ch/record/1019185

[77] X. Cid Vidal and R. Cid Manzano. “CMS Global Muon Trigger” web site: Taking a closer look at LHC. Retrieved from https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.lhc_trigger

[78] WLCG Project Office, “Documents & Reference - Tiers - Structure,” (2014). http://wlcg.web.cern.ch/documents-reference , last accessed on 30.01.2018.

[79] CMS Collaboration. “CMSSW Application Framework”, https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework, last accesses 06.02.2018

[80] A. Buckleya, J. Butterworthb, S. Giesekec, et. al. “General-purpose event generators for LHC physics”. arXiv:1101.2599v1 [hep-ph] 13 Jan 2011

[81] A. Quadt. “Top Quark Physics at Hadron Colliders”. Advances in the Physics of Particles and Nuclei. Springer-Verlag Berlin Heidelberg. DOI: 10.1007/978-3-540-71060-8 (2007)

[82] DurhamHep Data Project, “The Durham HepData Project - PDF Plotter.” http://hepdata.cedar.ac.uk/pdf/pdf3.html , last accessed on 02.02.2018.

[83] F. Maltoni, G. Ridolfi, and M. Ubiali, “b-initiated processes at the LHC: a reappraisal,” Journal of High Energy Physics, vol. 07, p. 022, (2012).

[84] B. Andersson, G. Gustafson, G.Ingelman and T. Sjostrand, “Parton fragmentation and string dynamics”, Physics Reports, Vol. 97, No. 2-3, pp. 31-145, 1983.

[85] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements;” European Physical Journal C, vol. 76, no. 3, p. 155, (2016).

[86] J. Alwall et. al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” Journal of High Energy Physics, vol. 07, p. 079, (2014).

[87] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," Journal of High Energy Physics, vol. 11, p. 070, (2007).

[88] S. Agostinelli et al., "GEANT4: A Simulation toolkit," Nuclear Instruments and Methods in Physics, vol. A506, pp. 250âĂŞ303, (2003).

[89] J.Allison et.al.,"Recent developments in Geant4", Nuclear Instruments and Methods in Physics Research A 835 (2016) 186-225.

[90] CMS Collaboration "Full Simulation Offline Guide", `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideSimulation`, last accessed 04.02.2018

[91] A. Giammanco. "The Fast Simulation of the CMS Experiment" J. Phys.: Conf. Ser. 513 022012 (2014)

[92] A.M. Sirunyan et. al. "Particle-flow reconstruction and global event description with the CMS detector", JINST 12 P10003 (2017) `https://doi.org/10.1088/1748-0221/12/10/P10003`.

[93] The CMS Collaboration. " Description and performance of track and primary vertex reconstruction with the CMS tracker". JINST 9 P10009 (2014). doi:10.1088/1748-0221/9/10/P10009

[94] J. Incandela. "Status of the CMS SM Higgs Search" July 4, 2012. Pdf slides. Retrieved from `https://indico.cern.ch/event/197461/contributions/1478917/attachments/290954/406673/CMS_4July2012_Final.pdf`

[95] P. Billoir and S. Qian, "Simultaneous pattern recognition and track fitting by the Kalman filtering method", Nucl. Instrum. Meth. A 294 219. (1990).

[96] W. Adam, R. Fruhwirth, A. Strandlie and T. Todorov, “Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC”, eConf C 0303241 (2003) TULT009 [physics/0306087].

[97] K. Rose, “Deterministic Annealing for Clustering, Compression, Classification, Regression and related Optimisation Problems”, Proc. IEEE 86 (1998) 2210.

[98] R. Fruhwirth, W. Waltenberger and P. Vanlaer, “ Adaptive Vertex Fitting”, CMS Note 2007-008 (2007).

[99] CMS collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV ”, JINST 7 P10002 2012, [arXiv:1206.4071].

[100] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-k_t jet clustering algorithm,” Journal of High Energy Physics, vol. 04, p. 063, (2008).

[101] B. Dorney. “Anatomy of a Jet in CMS”. Quantum Diaries. June 1st, 2011. Retrieved from `https://www.quantumdiaries.org/2011/06/01/anatomy-of-a-jet-in-cms/`

[102] The CMS Collaboration.“Event Displays from the high-energy collisions at 7 TeV”, May 2010, CMS-PHO-EVENTS-2010-007, Retrieved from `https://cds.cern.ch/record/1429614`.

[103] The CMS collaboration. “Determination of jet energy calibration and transverse momentum resolution in CMS”. JINST 6 P11002 (2011). `http://dx.doi.org/10.1088/1748-0221/6/11/P11002`

[104] The CMS Collaboration, “Introduction to Jet Energy Corrections at CMS.”. `https://twiki.cern.ch/twiki/bin/view/CMS/IntroToJEC`, last accessed 10.02.2018.

[105] CMS Collaboration Collaboration. “Identification of b quark jets at the CMS Experiment in the LHC Run 2”. Tech. rep. CMS-PAS-BTV-15-001. Geneva: CERN, (2016). https://cds.cern.ch/record/2138504.

[106] CMS Collaboration Collaboration. “Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector”. Tech. rep. CMS-PAS-JME16-004. Geneva: CERN, 2016. https://cds.cern.ch/record/2205284.

[107] CMS Collaboration, “New CMS results at Moriond (Electroweak) 2013”, Retrieved from http://cms.web.cern.ch/sites/cms.web.cern.ch/files/styles/large/public/field/image/HIG13004_Event01_0.png?itok=LAwZzPHR

[108] CMS Collaboration, “New CMS results at Moriond (Electroweak) 2013”, Retrieved from http://cms.web.cern.ch/sites/cms.web.cern.ch/files/styles/large/public/field/image/TOP12035_Event01.png?itok=uMdnSqzC

[109] K. Skovpen. “Event displays highlighting the main properties of heavy flavour jets in the CMS Experiment”, Aug 2017, CMS-PHO-EVENTS-2017-006. Retrieved from https://cds.cern.ch/record/2280025.

[110] G. Cowan. “Topics in statistical data analysis for high-energy physics”. arXiv:1012.3589v1

[111] A. Hoecker et al., “TMVA-Toolkit for multivariate data analysis” arXiv:physics/0703039v5 (2009)

[112] L. Lista. “Statistical Methods for Data Analysis in Particle Physics”, 2nd ed. Springer International Publishing. (2017) https://dx.doi.org/10.1007/978-3-319-62840-0

[113] I. Antcheva et al., “ROOT-A C++ framework for petabyte data storage, statistical analysis and visualization ,” Computer Physics Communications, vol. 182, no. 6, pp. 1384âĂŞ1385, (2011).

[114] Y. Coadou. “Boosted decision trees”, ESIPAP, Archamps, 9 February 2016. Lecture. Retrieved from https://indico.cern.ch/event/472305/contributions/1982360/attachments/1224979/1792797/ESIPAP_MVA160208-BDT.pdf

[115] J.H. Friedman. “Greedy function approximation: A gradient boosting machine”. Ann. Statist. Volume 29, Number 5 (2001), 1189-1232. https://projecteuclid.org/download/pdf_1/euclid.aos/1013203451.

[116] F. James, M. Roos, “MINUIT: Function minimization and error analysis”. Cern Computer Centre Program Library, Geneve Long Write-up No. D506, 1989

[117] J. Neyman and E. S. Pearson, “On the problem of the most efficient tests of statistical hypotheses”. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character. Vol. 231 (1933), pp. 289-337

[118] B. Hespel, F. Maltoni, and E. Vryonidou, “Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM”, JHEP 06 (2015) 065, doi:10.1007/JHEP06(2015)065, arXiv:1503.01656.

[119] ATLAS Collaboration, "Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC", Phys. Lett. B726 (2013) 88âĂŞ119, doi:10.1016/j.physletb.2014.05.011,10.1016/j.physletb.2013.08.010, arXiv:1307.1427. [Erratum: Phys. Lett.B734,406(2014)].

[120] CMS Collaboration, "Search for the associated production of a Higgs boson with a single top quark in proton-proton collisions at $\sqrt{s} = 8$ TeV", JHEP 06 (2016) 177,doi:10.1007/JHEP06(2016)177, arXiv:1509.08159.

[121] B. Stieger, C. Jorda Lope et al., "Search for Associated Production of a Single Top Quark and a Higgs Boson in Leptonic Channels", CMS Analysis Note CMS AN-14-140, 2014.

[122] M. Peruzzi, C. Mueller, B. Stieger et al., "Search for ttH in multilepton final states at $\sqrt{s} = 13$ TeV", CMS Analysis Note CMS AN-16-211, 2016.

[123] CMS Collaboration, "Search for H to bbar in association with a single top quark as a test of Higgs boson couplings at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-16-019, 2016.

[124] CMS Collaboration, "Search for production of a Higgs boson and a single top quark in multilepton final states in proton collisions at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-17-005, 2016.

[125] B. WG, "BtagRecommendation80XReReco", February, 2017. `https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation80XReReco`.

[126] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, 2016.

[127] M. Peruzzi, F. Romeo, B. Stieger et al., "Search for ttH in multilepton final1 states at $\sqrt{s} = 13$ TeV", CMS Analysis Note CMS AN-17-029, 2017.

[128] CMS Collaboration, "Baseline muon selections for Run-II." `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonIdRun2`, last accessed on 24.02.2018.

[129] G. Petrucciani and C. Botta, "Two step prompt muon identification", January, 2015. `https://indico.cern.ch/event/368007/contribution/2/material/slides/0.pdf`.

[130] H. Brun and C. Ochando, "Updated Results on MVA eID with 13 TeV samples", October, 2014. `https://indico.cern.ch/event/298249/contribution/3/material/slides/0.pdf`.

[131] B. Maier, "SingleTopHiggProduction13TeV", February, 2016. https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopHiggsGeneration13TeV.