Search for Higgs Decay to Dark Matter and Trigger Studies

João Carlos Arnauth Pela
of Imperial College London

A dissertation submitted to Imperial College London
for the degree of Doctor of Philosophy

Abstract

Here the abstract of the thesis

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university. This dissertation does not exceed the word limit for the respective Degree Committee.

João Pela

Acknowledgements

TODO:

- Family
- Friends
- Work colleagues (include CMS collaboration)
- more

The work presented in this thesis was supported by the Portuguese Government through Fundação para a Ciência e a Tecnologia (FCT) in the form of my PhD grant with the reference SFRH/BD/77979/2011. I am thankful for their support which allowed me to attain higher education.

Preface

Thesis structure and so on...

"To my grand mother"

Contents

Chapter 1

Experimental Apparatus

1.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [1, 2] is currently the world's largest particle accelerator and is capable of producing the highest energy particle beams ever made by mankind. This machine has total perimeter of 26.7 km and was built at European Organization for Nuclear Research (CERN) in a circular tunnel, which previously housed the Large Electron Positron collider (LEP) [3], at an average depth of 100 m below ground under the Franco-Swiss border near Geneva, Switzerland. A diagram of the LHC tunnel and its experiments can be found at figure 1.1.

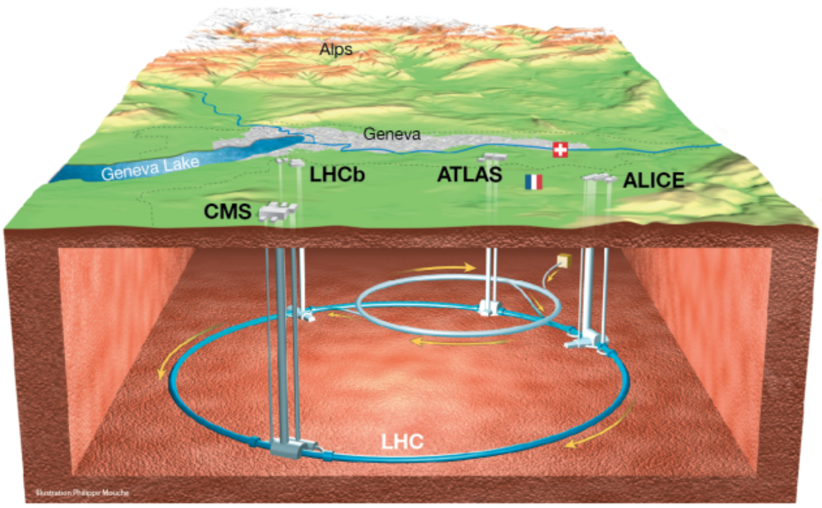


Figure 1.1: Underground diagram of the Geneva area showing the LHC and its experiments location [4].

The LHC is a synchrotron machine with the capability of accelerating two particles beams in opposite directions in two separated beam pipes. These beams only cross and are forced to collide in four points of the accelerator where particle detectors are installed to observe the products of such collisions. This experiments are: A Toroidal LHC ApparatuS (ATLAS) [5], Compact Muon Solenoid (CMS) [6], Large Hadron Collider beauty (LHCb) [7] and A Large Ion Collider Experiment (ALICE) [8].

The objective of the LHC program is to investigate physics at the TeV scale, more specifically to understand the electroweak symmetry breaking and if this phenomena could be explained by the Higgs mechanism. There are many Beyond the Standard Model (BSM) models that predict new physics at this energy regime making the LHC the perfect machine to investigate such phenomena. ATLAS and CMS are general-purpose detectors which aim to investigate a broad spectrum of physics. The LHCb detector is used to study processes that involve the decay of b-flavoured hadrons. The ALICE detector is optimised to look at heavy-ion collisions and to investigate the properties of extreme high density medium that is formed.

The LHC is only the last element of a complex accelerator chain which step-by-step increases the energy of the particles to eventually be collided [2]. Protons are initially obtained by stripping the electrons of hydrogen gas. This process happens at the begging of the Linear Particle Accelerator 2 (LINAC2) which then accelerates them up to the energy of 50 MeV. After this initial step proton are injected into the Proton Synchrotron Booster (PSB) and the energy ramps ups to 1.4 GeV. Particles are then passed to the Proton Synchrotron (PS) where the energy futher increases to 25 GeV. Subsequently they are the injected into the Super Proton Synchrotron (SPS) where the particle energy level reaches 450 GeV. Finally, protons pass to the LHC where they can be accelerated to a maximum energy of 7 TeV. A simplified diagram of the CERN accelerator chain can be found in figure 1.2. Normal operation of the LHC therefore depends on all the upstream accelerators availability. The typically turn around time, the time necessary to stop the accelerator from running and restart collisions, is around 2 hours. When stable beams are achieved, a single proton fill can be used to collide protons up to 24 hours, but it is common to restart more frequently to profit from the higher collision rates possible right at the beginning of a new fill.

Each beam pipe can be filled with proton or heavy ions. Three modes of operation have been tried: proton-proton, proton-lead ion and lead ion-lead ion. By changing the incoming particles we are changing the quantity of nucleons present at each interaction. The maximum design energy per proton is 7 TeV and is 2.76 TeV for each lead nucleon.

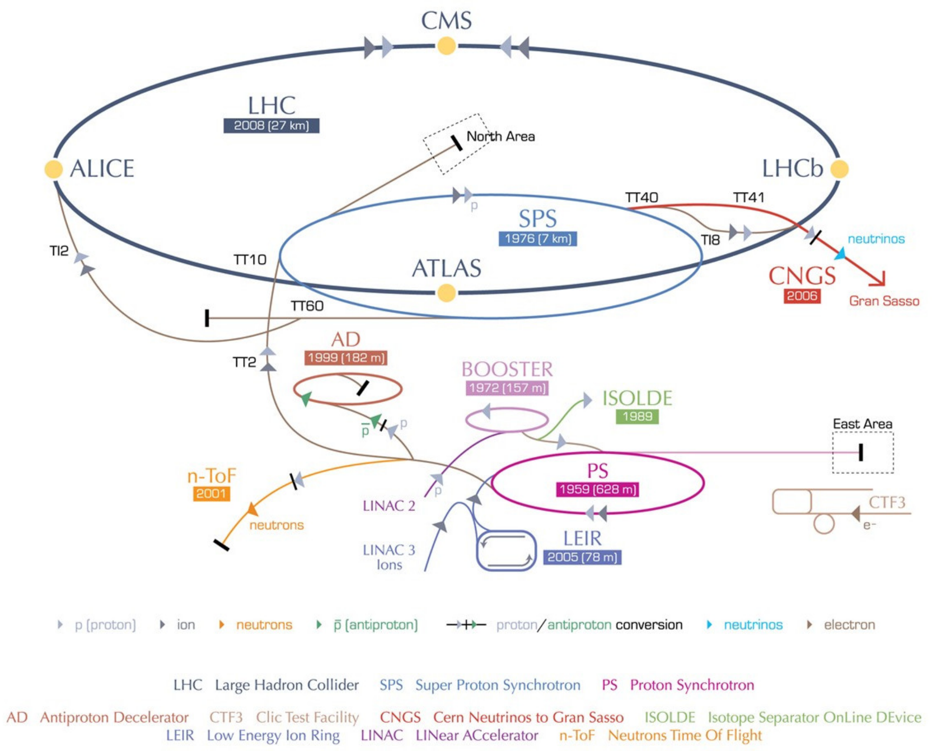


Figure 1.2: Diagram of the CERN accelerator complex [9].

The maximum design luminosity for proton-proton is $10^{34}\,\mathrm{cm^{-2}s^{-1}}$ and $10^{27}\,\mathrm{cm^{-2}s^{-1}}$ for lead ion-lead ion collisions.

Particles beams trajectory are curved by 1232 niobium-titanium superconducting dipole magnets each with a length of 14.3 m. They are cooled with superfluid helium to 1.9 K and produce the necessay magnetic field of 8.4 T. Eight Radio Frequency (RF) cavities located at the LHC point 4 are used to accelerate the beams. At each turn particle energy is increased to compensate for synchrotron radiation loss and increase the momentum. At nominal operation the LHC will steer 2808 bunches composed up to 10^{11} protons separated by 25 ns in each direction. Some of the key parameters of the LHC proton-proton and lead-lead operation can be found in table 1.1.

At the LHC we are looking for extremely rare processes. As is can be seen in figure 1.3 the production cross section of a Standard Model (SM) Higgs boson is more than 9 orders of magnitude smaller than the total proton-proton cross section.

		pp	**HI**	
Energy per nucleon	E	7	2.76	TeV
Dipole field at 7 TeV	B	8.33	8.33	T
Design Luminosity*	$\mathcal{L}$	10^{34}	10^{27}	$cm^{-2}s^{-1}$
Bunch separation		25	100	ns
No. of bunches	k_B	2808	592	
No. particles per bunch	N_p	1.15×10^{11}	7.0×10^{7}	
Collisions				
β-value at IP	β^*	0.55	0.5	m
RMS beam radius at IP	σ^*	16.7	15.9	µm
Luminosity lifetime	τ_L	15	6	h
Number of collisions/crossing	n_c	≈ 20	-	

* For heavy-ion (HI) operation the design luminosity for Pb-Pb collisions is given.

Table 1.1: The machine parameters relevant for the LHC detectors.[10]

To be able to record and study such rare processes we need to produce a significant number of collisions. For this purpose the LHC was designed to operate at high instantaneous luminosity, L. This quantity is defined as,

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma}{4\pi \epsilon_n \beta^*} F, \tag{1.1}$$

where N_b is the number of protons per bunch, n_b is the number of bunches, f_{rev} is the frequency of revolution, γ is the Lorentz factor, ϵ_n is the normalized emittance, β^* is the beta function at the collision point and F is the reduction factor due to the crossing angle.

1.1.1 Running and performance

Operation of the LHC has started when the first beams circulated in the machine in September 2008. Unfortunately, only a few days after a faulty weld between two dipole magnets caused a significant magnet quench which in turn damaged several dipoles and a simultaneous leak of a significant amount of helium happened. The event showed that

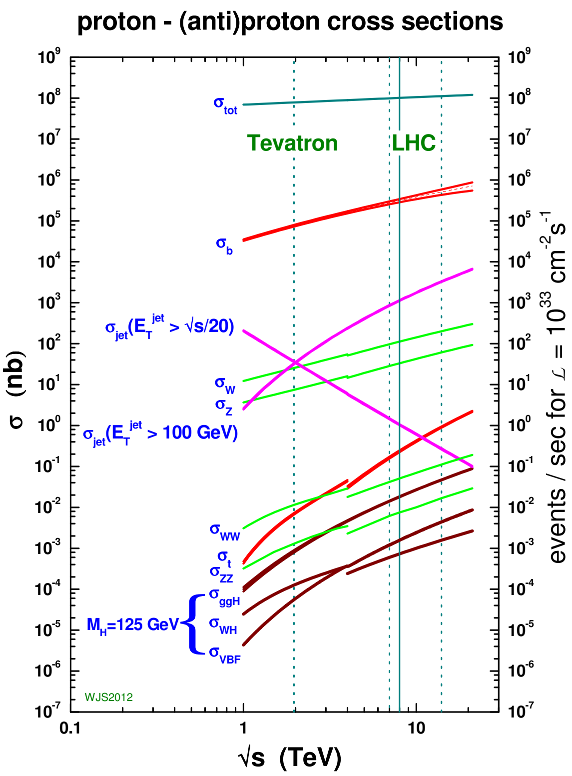


Figure 1.3: Cross sections for several processes for collisions of antiproton-proton and proton-proton as a function of the center of mass energy [6].

beyond the repair of the affected systems the accelerator needed a significant consolidation program to allow it to return to activity [11]. This consolidation program took over one year to finalise and to prevent further possible problems and allow better understanding of the machine while maximizing physics reach, it was decided to initially run the LHC at 7 TeV center-of-mass energy. First collisions happened at November 2009 just at the SPS injection energy of 450 GeV giving start to the LHC run I.

The collision energy was finally ramped up to 7 TeV with first collisions being observed during March 2010. Operation at this energy continued until the end of 2011, with a peak luminosity being achieved of $3.7 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$. The total integrated luminosity delivered to CMS was $6.1\,\text{fb}^{-1}$ with the total actually recorded being $5.6\,\text{fb}^{-1}$. During 2012 with the increase knowledge of the accelerator it was possible to increase the centre-of-mass energy further to 8 TeV and eventually reaching peak luminosity of $7.7 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-2}$ and delivering integrated luminosity of $23.3\,\text{fb}^{-1}$ to CMS of which $21.79, \text{fb}^{-1}\ \text{fb}^{-1}$ were recorded. Figure 1.4 shows the delivered luminosity in the period 2010-2013 over time.

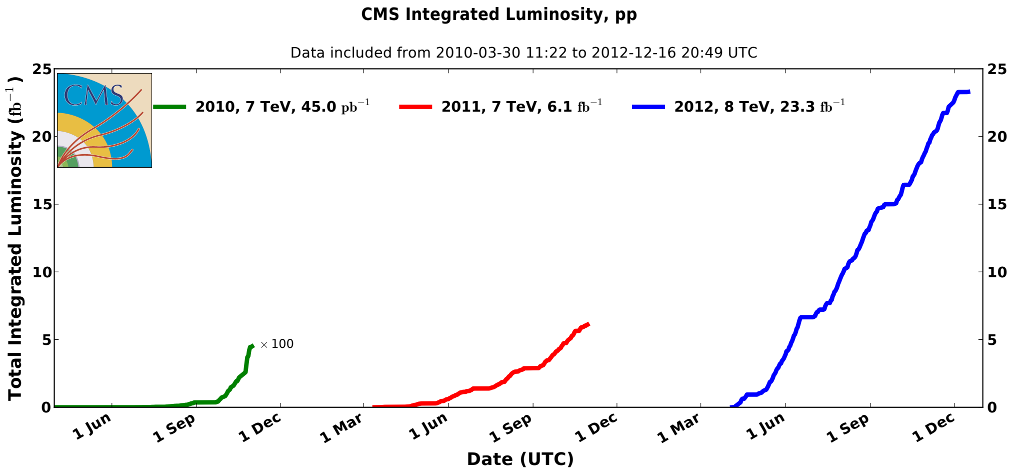


Figure 1.4: Cumulative luminosity versus day delivered to CMS during stable beams and for p-p collisions. This is shown for 2010 (green), 2011 (red) and 2012 (blue) data-taking [12].

For physics usage, data needs to undergo a certification process. In this process specialists from each CMS subsystem check that no problem has happened during data taking that would bias or invalidate the recorded events. For 2011 a total of 5.1 fb^{-1} and for 2012 a total 19.7 fb^{-1} were considered of good quality for physics.

In order to achieve high integrated luminosity LHC collides particle bunches up to 40 million times a second, and many interactions may happen simultaneously, this effect is called Pile-Up (PU). Figure 1.5 shows the distribution of the mean number of interaction per bunch crossing during 2012 at the CMS experiment.

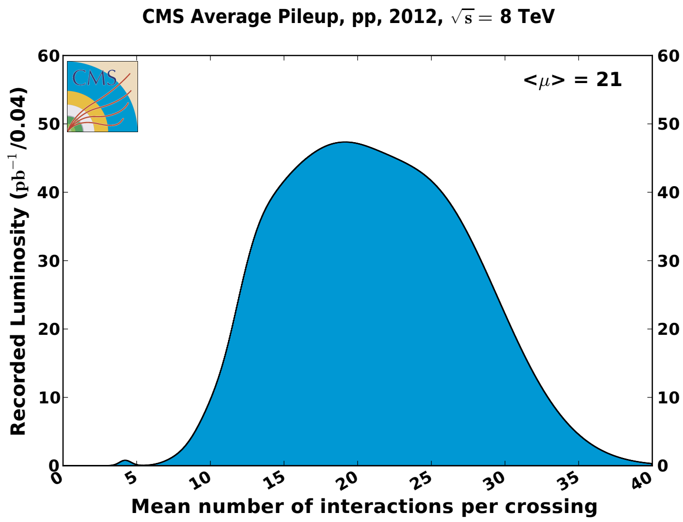


Figure 1.5: Mean number of interactions per bunch crossing at the CMS experiment during 2012 [13].

1.2 The Compact Muon Solenoid Experiment

The Compact Muon Solenoid (CMS) experiment is a general purpose experiment located at the LHC point 5, near the village of Cessy, France. It was designed to be a high performance detector studying collisions at its centre. It is composed of several subsystems in a classic onion shaped structure. A diagram of the experiment can be found in figure 1.6.

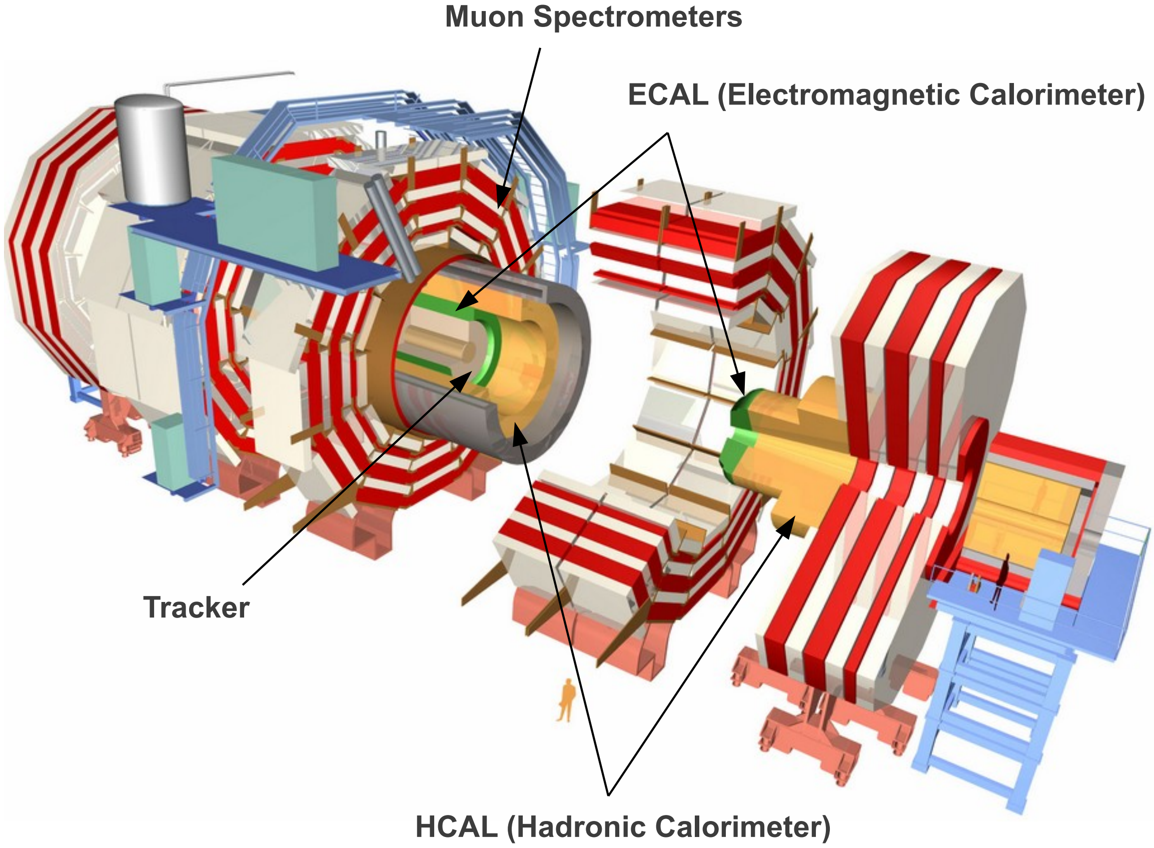


Figure 1.6: Diagram of CMS experiment showing the experiment in an open configuration and highlighting the position of its sub-detectors [14].

The main driving motivation for its design was to investigate the electroweak symmetry breaking and the Higgs mechanism at the design time was presumed to be the most likely explanation. Many alternative theories to the standard model predict new particles which could be observed at the TeV scale, CMS as a multi-purpose experiment is well suited to search for these new scenarios. If found, such new physics may allow us to understand some of the currently open questions in particle physics, like providing a particle candidates for dark dark matter. Further more, some of these possible new physics signals could point the way towards a grand unified theory. CMS is also capable of operating while the LHC is colliding heavy ions and has a rich program covering the study of matter at extreme temperatures, densities and parton momentum fraction (low-x).

The requirements imposed on the CMS design to meet its physics goals can be summarized in the following table [6, 15]:

- Good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution ($\approx 1\%$ at $100\,GeV$), and the ability to determine un-ambiguously the charge of muons with $p_{\mathrm{T}} < 1\,\mathrm{TeV}$.

- Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ's and b-jets, requiring pixel detectors close to the interaction region.

- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution ($\approx 1\%$ at 100 GeV), wide geometric coverage, π^0 rejection, and efficient photon and lepton isolation at high luminosities.

- Good missing-transverse-energy and dijet-mass resolution, requiring hadron calorimeters with with a large hermetic geometric coverage and with fine lateral segmentation.

The final detector design fulfils all these requirements. The experiment is compact compared to the other LHC experiments being 22 m long and 15 m in diameter. Although small, it is the heaviest of the four big detectors at 12500 tonnes. Its high density is a direct consequence of it producing the highest magnetic field at 4 T and therefore needing more material for it to be contain in its return yoke.

1.2.1 Geometry and conventions

The adopted coordinate system has it origin in the center of CMS where the nominal collision point is located, the y-axis points vertically upwards, and the x-axis points radially inward in the direction of the centre of the LHC. The z-axis points along the beam line towards the Jura mountains from the LHC point 5. The azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle θ is measured from the z-axis.

We define pseudorapidity as $\eta = -ln(tan(\theta/2))$. All transverse quantities, like the transverse momentum ($p_{\perp}$), are measured in the transverse plane of beam axis. The imbalance of energy is also measured in the x-y plane and is denoted as $E_{\perp}^{miss}$.

1.2.2 Inner tracking system

The inner tracking system is the closest detector to the beam axis and the interaction region [16, 17]. Its function is to measure the trajectory of all charged particles with momentum above 1 GeV being produced at each LHC collision. With the help of the strong magnetic field produced by the CMS magnet, particle trajectories are bent allowing for charge and momentum determination. With the resulting tracks is it then possible to determine the primary vertex as well as secondary vertexes like other lower energy proton-proton collision or displaced vertexes from the decay of long lived particles like B mesons.

Building a tracking system for an experiment at the LHC is very challenging. Such system at design luminosity will be hit by an average of 1000 particles at a rate approaching 40 MHz. It needs to be a fast, efficient, high granularity detector, radiation hard and as thin as possible to not deflect the incoming particles trajectory. At each layer the occupancy should be of the order of 1% or lower. This design requirements have lead to a tracker design entirely based on silicon detector technology.

The volume near the interaction point can be split according to the charged particle flux into three regions:

- $r < 10$ cm: highest particle flux, up to $\approx 10^8$cm^{-2}s^{-1} at $r \approx 4$cm, pixel detectors are used. The pixel size is $\approx 100 \times 150\,\mu\text{m}^2$, which translates into an occupancy of 10^{-4} per LHC bunch crossing.
- $20 < r < 55$ cm: particle flux decreases enough to use silicon micro-strips with a minimum cell size of 10 cm $\times$ 80 µm, leading to an occupancy of $\approx 2 - 3\%$ per LHC bunch crossing.
- $50 < r < 110$ cm: most outer region of the tracker, particle flux is low enough to use larger pitch silicon micro-strips. The maximum cell size is of 25 cm $\times$ 180 µm, and occupancy is of the order of $\approx 1\%$.

The CMS tracker final configuration is composed of a pixel detector with three barrel layers at radii between 4.4 cm and 10.2 cm and 2 disks on each side of the barrel. And a silicon strip tracker with 10 barrel detection layers extending up to 1.1 m with 3 plus 9 disks on each side of the barrel. A schematic of this detector module distribution can be found at figure 1.7. This detector has an acceptance covering up to pseudorapidity of $|\eta| < 2.5$ and has a total active area of about $200\,\mathrm{m}^2$ making the largest silicon tracker ever built.

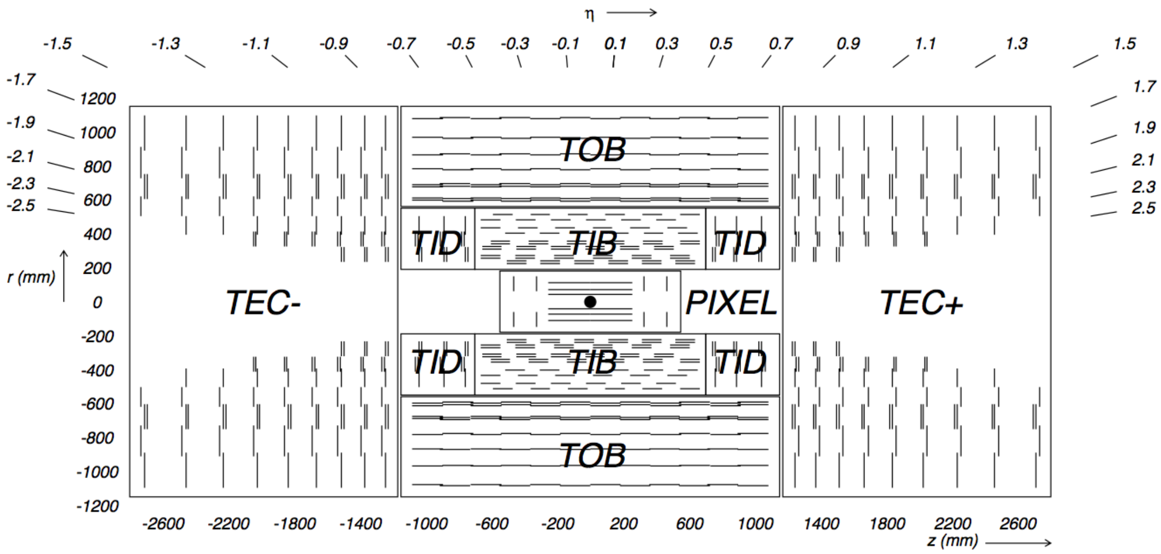


Figure 1.7: Schematic cross section of the CMS tracker [6]. Each line represent a detector module. Double lines represent dual surface back-to-back detector modules. The inner tracker components are shown components: Tracker Endcaps (TEC), Tracker Outer Barrel (TOB), Tracker Inner Disks (TID), Tracker Inner Barrel (TIB) and Pixels.

1.2.3 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) is the detector responsible for measuring the energy of electrons and photons [18, 19]. It is an hermetic energy measurement system comprised of 61200 lead tungstate ($PbWO_4$) crystals mounted in the barrel and 7324 crystals in each of the 2 endcaps and it has an acceptance up to $|\eta| < 3.0$.

Lead tungstate has a fairly high density ($8.28\,\mathrm{g/cm^3}$), has a short radiation length (0.89 cm) and a small Moliere redius (2.2 cm). The crystals also have a fast scintillation decay time emitting 80% of the light yield in 25 ns (the minimal bunch crossing time at the LHC). This characteristics make it a good choice for an electromagnetic calorimeter allowing a compact design with fine granularity. However, this crystals emit a fairly low light yield ($30\,\gamma/\mathrm{MeV}$) which requires the use of photo-detectors with intrinsic gain which will preform well inside a magnitic field. In the barrel region silicon Avalanche photo-diodes (APD) are used and Vacuum Photo-Triodes (VPT) are used in the endcaps. To guarantee good response from both crystals and APD it is necessary to have system thermal stability, with the goal being temperature variation of less than 0.1°C.

The barrel section, the ECAL Barrel (EB), has an inner radius of 129 cm and is composed of 36 identical "supermodules", each covers the barrel length and corresponding to a pseudo-rapidity interval of $0 < |\eta| < 1.479$. The crystals are quasi-projective (the axes are tilted at 3° with respect to the line from the nominal vertex position) and cover 0.0174 (i.e. 1°) in $\Delta\phi$ and $\Delta\eta$. The crystals have a front face cross-section of $\approx 22 \times 22\,\text{mm}^2$ and a length of 230 mm, corresponding to 25.8 X_0.

The endcap section, the ECAL Endcap (EE), is at a distance of 314 cm from the vertex and covering a pseudorapidity range of $1.479 < |\eta| < 3.0$, are each structured as 2 "Dees", consisting of semi-circular aluminium plates from which are cantilevered structural units of 5×5 crystals, known as "supercrystals". A diagram of the ECAL can be found on figure 1.8.

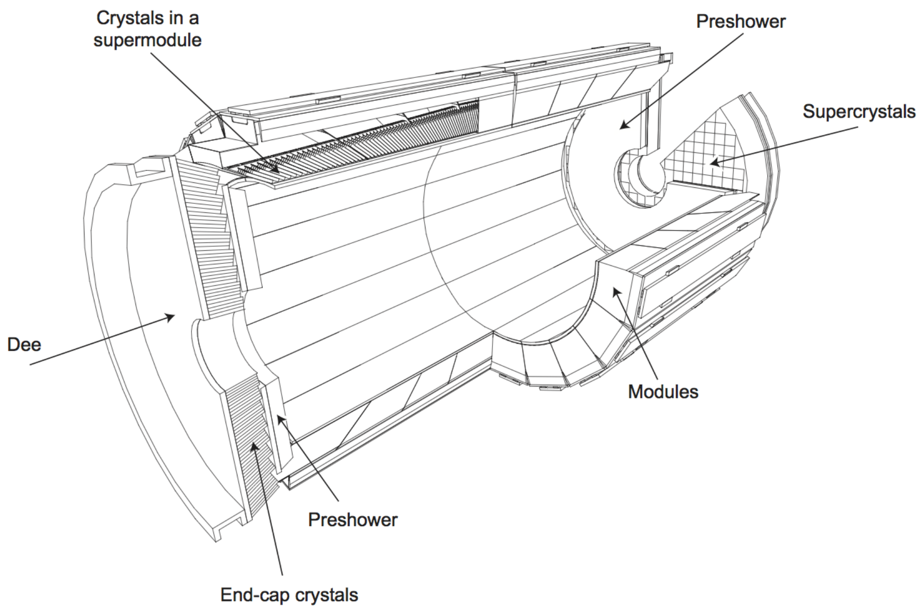


Figure 1.8: Diagram of the ECAL layout illustrating the positions of its components [6].

The energy resolution of the ECAL can be can be expressed as:

$$\frac{\sigma}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \tag{1.2}$$

Here E is the energy of the incoming particle, S is the stochastic term which quantifies the fluctuations in scintillation and lateral containment of the shower, N the noise term which relates with electronics and digitisation process and finally C is a constant term that quantifies the non-uniform longitudinal response and inter-calibration errors. These parameters have been measured to be $S = 0.028\,\mathrm{GeV}^{1/2}$, $N = 0.12\,\mathrm{GeV}$ and $C = 0.003$ with the help of an electron beam [20] and in the absence of magnetic field.

Preshower detector

The CMS Preshower is a detector locate in each endcap covering the fiducial region of $1.653 < |\eta| < 2.6$. Its mission is to identify neutral pions decay, help to identify electrons against minimum ionizing particles and improve electron and photon position determination.

This detector is sampling calorimeter composed by two layers of lead radiators each followed by silicon strip sensors. The lead layers have the function of forcing the incoming particles to initiate an electromagnetic shower. The first lead layer had $2X_0$ while the secong had $1X_0$, which results in 95% of the single incident photons starting their shower before hitting the first sensor [6]. The shape of the lead layers edge matches the ECAL crystal behind them to facilitate calculations at the Level 1 Trigger (L1T).

Each silicon sensors have an active area of $61 \times 61\,\mathrm{mm}$ and are $320\,\mu\mathrm{m}$ thick. The sensors are divided into 32 strips each with 1.9 mm. The preshower system has a total thickness of 20 cm and had 137000 individual read-out channels.

1.2.4 Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is a sampling calorimeter which is designed to measure the properties of hadron jets and indirectly neutrinos or other undiscovered particles that would result in apparent missing energy [21]. The design of the Hadronic Calorimeter (HCAL) was strongly influenced by the choice of the magnet parameters since most of the calorimetry is inside of the magnet. A diagram of the HCAL subsystems and their location inside CMS can be found in figure 1.9.

The HCAL Barrel (HB) covers the region up to $|\eta| < 1.3$ and is limited from the beam side by the ECAL at radius $r = 1.77\,\mathrm{m}$ and outwards by the magnet at radius $r = 2.95\,\mathrm{m}$. This is a strict limitation on the amount of absorber material to be used. This detector is

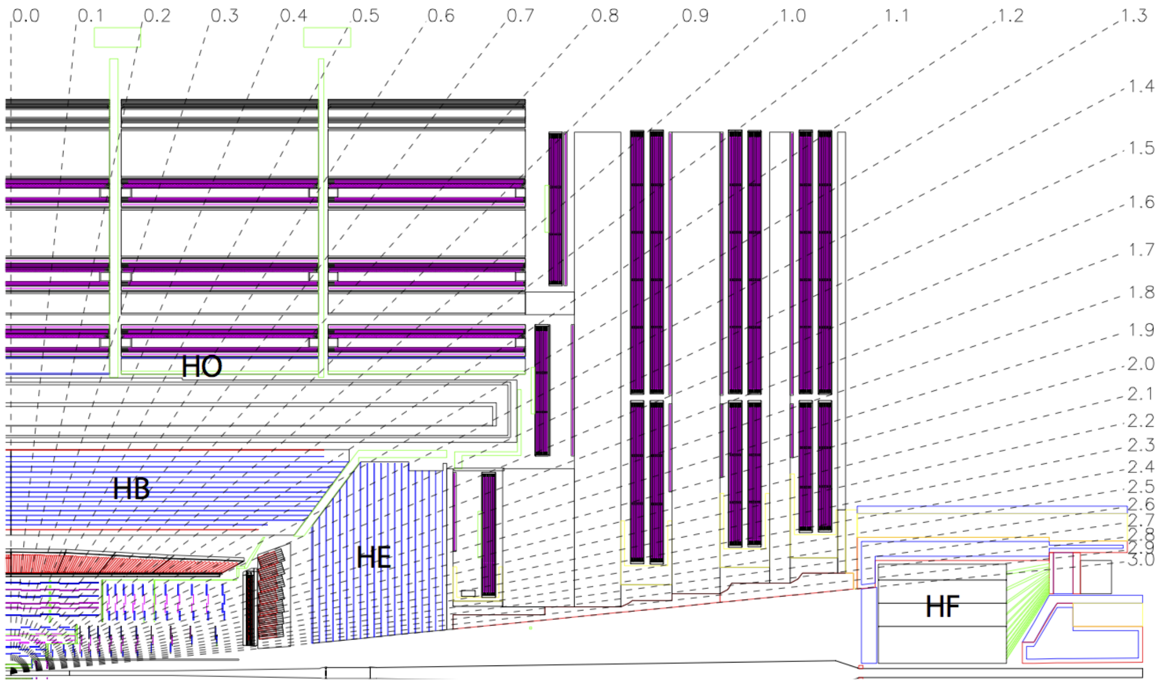


Figure 1.9: Longitudinal view of the CMS detector highlighting the location of the HCAL components: HCAL Barrel (HB), HCAL Endcap (HE) HCAL Outer (HO) and HCAL Forward (HF) [6].

composed of 36 identical azimuthal wedges split in two half-barrels. They are constructed of brass absorber plates alternated with plastic scintillator. Brass has a short interaction length ($X_0 = 16.42$ cm) and is non-magnetic. The detector is composed of 2304 towers with a segmentation of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ which corresponds to the same area of a 5×5 arrays of ECAL crystals.

To improve the measurement capability, an outer calorimeter, the HCAL Outer (HO), is placed outside of the magnet as a *tail catcher*. It increases the effective thickness of the hadronic calorimeter by over 10 interaction lengths. This detector covers the range $|\eta| < 1.26$, it is composed or iron absorber and scintillator and is subdivided into sectors that cover 30° azimuthal angle in each of the barrel wheels.

The HCAL Endcap (HE) covers the range of $1.3 < |\eta| < 3.0$. It is composed by 2034 towers with a 14 towers segmentation in η and 5° segmentation in ϕ. The 8 inner most towers the segmentation is 10° in ϕ, whilst the η segmentation increases in η from 0.09 to 0.35.

Additionally, to extend acceptance to $|\eta| < 5.2$ the HF is installed at 11.2 m from the interaction point providing excellent hermeticity for $E_{\perp}^{miss}$ measurement. Its steel absorber is 1.65 m deep and has quartz fibres running through it, parallel to the beam line. The energy measurement is made via Cerenkov light produced by the incoming particles inside the fibres. There are 13 tower in η with segmentation of $\approx \Delta\eta = 0.175$ except the lowest η tower with $\approx \Delta\eta = 0.1$ and highest η tower with $\approx \Delta\eta = 0.3$. The segmentation in ϕ is of $\Delta\phi = 10^{o}$ except in the highest η towers which is $\Delta\phi = 20^{o}$. There are a total of tower 900 per HF module.

Similarly to the ECAL the energy resolution HCAL was tested using a test beam of single charged pions [20], and it was obtained that:

$$\frac{\sigma}{E} = \frac{94.3\%}{\sqrt{E}} \oplus 8.4\%. \tag{1.3}$$

1.2.5 Solenoid Magnet

The design requirements for correct charge assignment and p_{T} determination for charge particles and specially muons drive the magnet parameters choice. For muons, unambiguously charge determination requires momentum resolution of $\Delta p/p \approx 10\%$ at $p = 1\mathrm{TeV}$. This requirements are specially difficult to obtain in the forward regions but with the correct length/radius ratio can be obtained with a modestly sized solenoid magnet but with large field [10, 22].

The choice of the CMS collaboration was to build a Niobium-Titanium (NbTi) superconducting solenoid magnet which has been design to operate at fields up to 4 T it has a diameter of 6 m and a length of 12.5 m at maximum field the stored energy reaches 2.7 GJ. Typically, the magnet is only run at 3.8 T in order to maximize its lifetime. To contain such an enormous magnetic flux a 10 kt return yoke envelopes the magnet with 5 wheels in the barrel region and 2 endcaps composed of 3 disks closing the sides [6]. A summary of the most important magnet parameters can be found at table 1.2.

1.2.6 Muon System

The muon detection is an important part of the mission of CMS [23]. Muons are fairly easy to detect when compared with other elementary particles and are only rarely produced in proton-proton collisions. To take the example of the SM Higgs boson, while the decay mode involving a pair of Z bosons is fairly unlikely compared with other decays

Parameter	Value
Field	4 T
Inner Bore	5.9 m
Length	12.9 m
Number of turns	2168
Current	19.5 kA
Stored Energy	2.7 GJ
Hoop Stress	64 atm

Table 1.2: Parameters of the CMS superconducting solenoid [10]

the Z bosons can decay into 4 mouns. This decay while rare does not have significant backgrounds making it a "golden channel" for discovery, which indeed was proven the case [24]. Many other models, like SUSY, use muon final states in their searches exactly for the same reason. The CMS muon system is composed of 3 types of gaseous detectors depending on they location and momentum reconstruction needs. A diagram of the disposition of this system inside CMS can be found on figure 1.10.

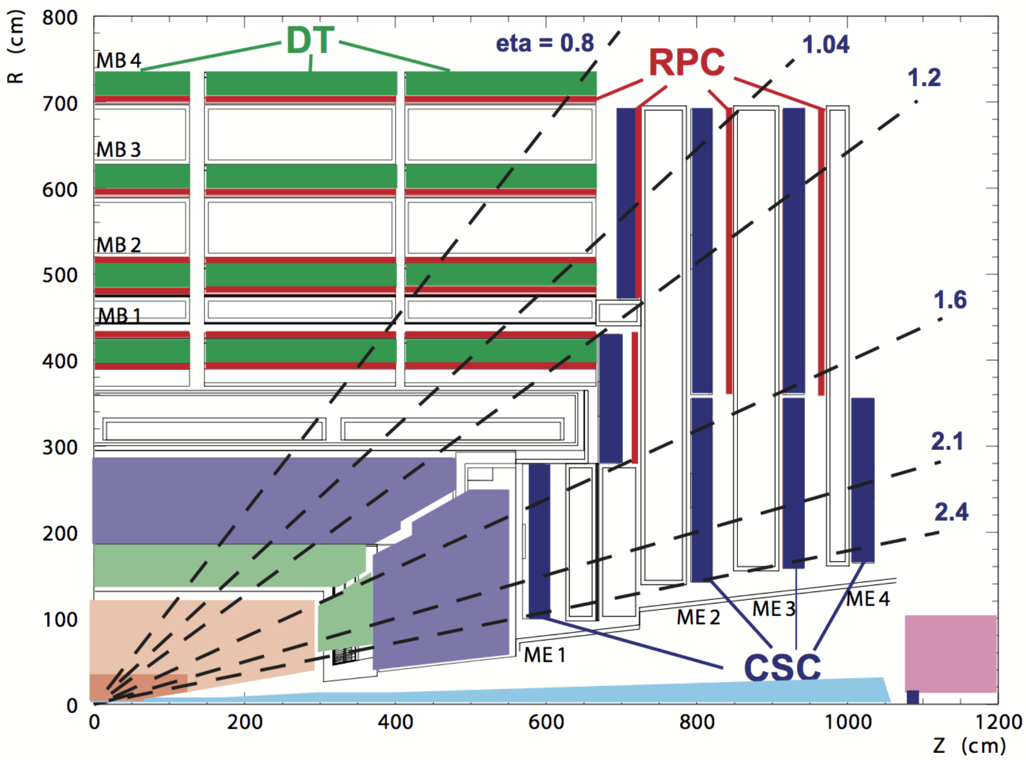


Figure 1.10: Diagram of the CMS muon systems. The location of each muon chamber for each subsystem is showed [10].

In the barrel and up to $|\eta| < 1.2$, Drift Tube (DT) are used. since the neutron background is small and the field is constant. This system is composed 250 chambers and is arranged in 4 concentric cylindrical layers which are installed inside of the return yoke. This chambers have a total of 172000 wires with a length of 2.4 m which are housed inside of tubes filled with a mixture of argon and carbon-dioxide. Each of the wheels of the barrel is split into 12 sectors covering 30^{o} azimuthal angle. The maximum gas ionization drift is of 2.0 cm and results in a single point resolution is $\approx$ 200 µm per wire. For each station each measured muon the ϕ resolution is better than 200 µm and direction resolution is $\approx$ 1 mrad.

In the endcaps Cathode Strip Chamber (CSC) are used in the region between $2.4 > |\eta| > 0.9$. Here, muon and background rates are high and the magnetic field is not uniform. This system has fast response and is radiation resistant. It is composed by 468 chambers arranged in 4 stations per side. Each chamber is trapezoidal in shape and made of 6 gas gaps and covers either 10^{o} or 20^{o} in ϕ. Each gap contains a plane of cathode strips and a plane of anode wires. For each chamber the spacial resolution is of the order of 200 µm and the angular resolution is $\approx$ 10 mrad in ϕ.

Finally the Resistive Plate Chamber (RPC) covers the $|\eta| < 1.6$ range. This system overlaps with the 2 other muon systems. It is very fast with an ionization event being much faster than the bunch crossing time. This fast response allows, in conjunction with a dedicated trigger system, to select the correct bunch crossing associated with the detection of a muon. In the barrel there 480 rectangular chambers arranged in 4 stations with 6 RPC layers (2 layers are present in the 2 stations closest to the beam pipe). In the endcaps there are 3 RPC disk shaped stations on each side, which are composed by trapezoidal shaped detectors.

The combined muon system offline momentum resolution is of the order of 9% for small values of η and p and for transverse momenta of up to 200 GeV. At higher energies of around 1 TeV the standalone momentum resolution is in the range of 15-40% depending on $|\eta|$. This values are limited by the muon multiple-scattering before arriving to the muon system. If we combine the tracker information into a global fit the resolution for lower p_{T} tracks improves an order of magnitude while at higher momenta (around 1 TeV) it is of about 5%, which is well inside the CMS design requirements.

1.2.7 Data Acquisition System

The CMS Data Acquisition (DAQ) system is designed to process, analyse and ultimately store the information collected by the detector [25]. The LHC produces bunch crossings at a rate of 40 MHz but we are only capable of storing between $10^2 - 10^3$ events per second. At design luminosity each collision will have an average of over 20 simultaneous collisions and produce a zero-suppressed data payload of around 1 MByte. To reduce the event rate to completely retrieve from the detector buffers a first level of trigger was developed. This hardware system reduces the amount of events to be processed to a maximum of 100kHz. Even with this event suppression the DAQ has to retrieve and move $\approx 100\,\text{GBytes}^{-1}$ from the detector to the surface. This data comes from approximately 650 data sources and has to be merged into a single event package. The information is then passed to a computer farm where a software filters serve as a second level of trigger. In this system the event rate is further reduced up top a factor of 1000 making the output rate compatible with what can be saved into permanent storage. A diagram of this system can be found on figure 1.11.

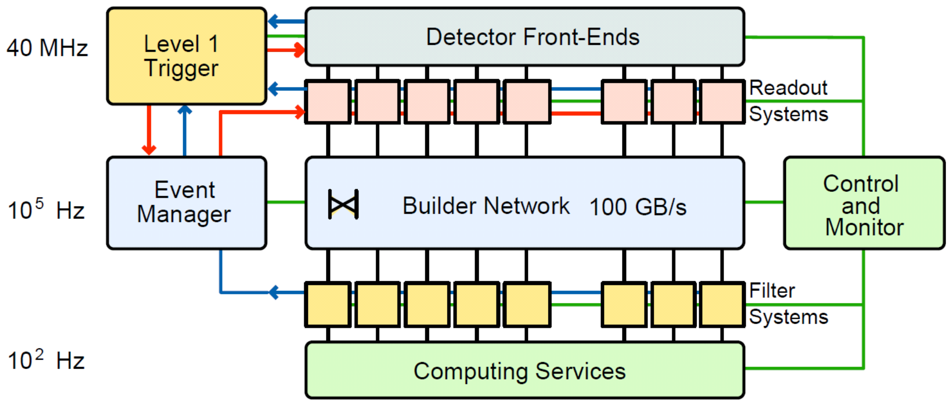


Figure 1.11: Diagram of the CMS DAQ system. Data flow is showed as the lines connecting each electronics or computing units [6].

1.2.8 Trigger System

As described on the previous section the CMS trigger system is responsible for selecting which collisions are recorded in real-time. We can only save $10^2 - 10^3$ events per second with the current systems. This implies that the trigger system needs to obtain a data reduction of a factor of $\mathcal{O}(10^6 - 10^7)$. This is achieved with a two level trigger system, the first is a dedicated hardware system named Level 1 Trigger (L1T) [26] and the second is a commercial computer system running dedicated software called the High Level Trigger (HLT) [25].

Initially, all data is stored for 128 bunch crossing which corresponds to 3.2 µs. This is the time we have to make a first decision to keep or discard an event. This is the task of the L1T which has the target to reduced the data to a maximum rate of 100 kHz. There isn't enough time to get all the information from the detector, so only a coarse version of the calorimetry and muon systems data, and some correlation between them is accessed. With this information the L1T produces a set of particle candidates and energy sums over which custom user defined algorithms can use to filter events. A diagram of the L1T trigger components and the data flow across the system is present on figure 1.12.

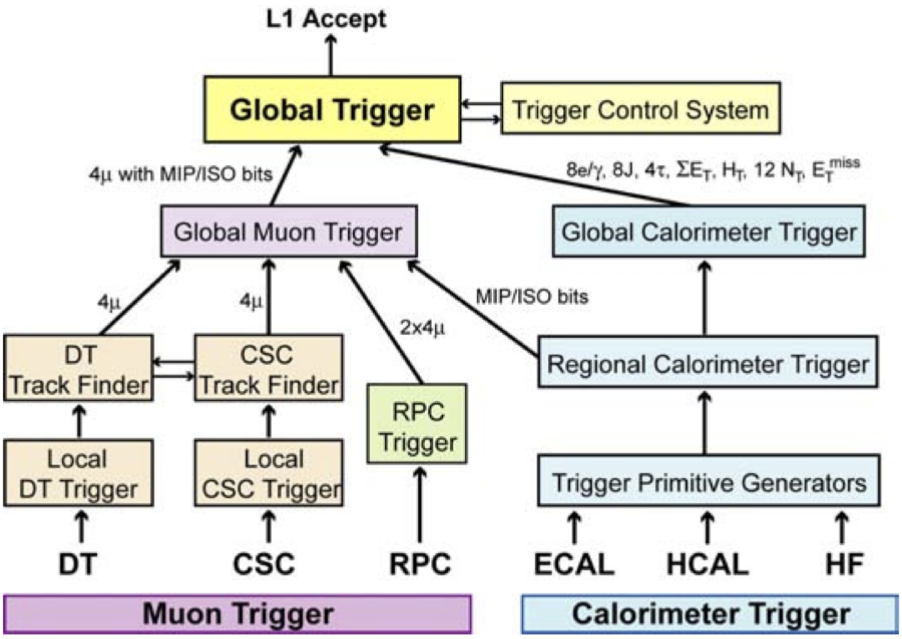


Figure 1.12: Diagram of the L1T system. The arrows indicate data flow and the number of particle candidates at each step is indicate [6].

The High Level Trigger (HLT) receives events accepted by the L1T and needs to perform further event reduction of $\mathcal{O}(10^3 - 10^2)$ to a final output rate of $\mathcal{O}(10^{2-3})$ Hz . This system is composed of standard computing hardware in the form of computing farm with ≈ 15 k Central Processing Unit (CPU). This system, using the additional latency created by the L1T event selection, is able to make use of the complete detector information including the tracker data. More sophisticated and precise algorithms are therefore possible which can be tailored to select any desired physical final state.

Event selection algorithms at both the L1T and HLT are frequently updated during data taking. The selection thresholds may be tuned in order to control the rate with the changes of LHC luminosity. Novel methods or strategies to identify particles more efficiently can be implemented, like PU subtraction or new calibrations. Analysis groups may also show interest in recording new event final final states for which new selection criteria may be developed. The set of algorithms used for data taking is normally referred as the *trigger menu*.

After events pass both levels of the trigger they are recorded into permanent storage. During 2012-13 operation, two output streams were saved. The *prompt data stream*, with a rate of approximately 300 Hz, was composed of high priority trigger paths which were immediately reconstructed. And the *parked data stream*, with an average rate of 600 Hz, was stored without reconstruction. This data waited until computing resources were free to go through reconstruction [27]. This process was finalised a few months after the LHC Run I was finished.

Even with such measures to reduce the data to be stored, each LHC experiment records several petabytes of data every year in addition to similarly sized amounts of simulated events.

1.2.9 Computing

The quantity of data produced by the LHC and the necessary processing capability is so big that it would be difficult to have all computing resources in a single place. For this reason a tiered system was developed, where all participating computing sites are connected and have specific roles and responsibilities in the data taking, processing and storing. This global computing system is know as the Grid [28].

The CERN Data Centre is the Tier 0 of this network, all data produced by the LHC experiments is handled by this facility. Only about 20% of the total capacity of the Grid is hosted here, but CERN Tier 0 has the very important mission of safe keeping all the raw data produced by the experiments. It also has the task of doing the first attempt of event reconstruction, which is the process identifying meaningful physics objects in data.

There are 7 CMS Tier 1 computer centres around the world. They are responsible to store a proportional amount of raw and reconstructed data. If any reprocessing of the data is needed, this centres are responsible for this task and storing the resulting output as well. Tier 1 centres also host simulated data and distribute it to Tier 2 centres.

Local research centres like universities of scientific laboratories are normally at the Tier 2 level. This centres have the responsibility of handling a proportional share of simulated data production and reconstruction. Currently there are over 150 Tier 2 centres around the world

Individual computers or local clusters without any formal engagement with the Grid structure, are considered to be the Tier 3 level of the Grid.

1.2.10 Level 1 Trigger: Stage I Upgrade

An extensive upgrade program for the L1T electronics was planed and is being executed in order to cope with the increase of luminosity and pile-up predicted for the period after Long Shutdown 1 (LS1) [29, 30]. The center-of-mass energy has almost double from 8 TeV to 13 TeV, instantaneous luminosity will also increase as will average pile-up. Also, the bunch separation has change from 50 ns to 25 ns making out-of-time pile-up a significant problem.

To ensure physics performance during 2015 and beyond only a partial upgrade was executed for the 2015 run which is known as the *Stage-1* upgrade. The main feature of this upgrade program is the replacement of the existing Global Calorimeter Trigger (GCT). Two key enhancements were possible:

- Event-by-event pile-up energy subtraction for jets reconstruction, e/γ isolation, τ isolation.
- Smaller feature size τ candidates, which will have significantly better energy estimation and background rejection.

The intermediate system will have significantly better performance than the now legacy system. The full 2016 calorimeter trigger system will additionally provide finer granularity which will lead to increased position and energy resolution.

Chapter 2

Event Reconstruction and simulation

This chapter describes how the Compact Muon Solenoid (CMS) detector produces physics objects from the information collected at each event. The Vector Boson Fusion (VBF) Higgs to invisible analysis uses almost all the physics objects reconstructed by the detector with making use of information from all the experiment sub-detectors. The following sections describe in detail each of these objects how they are reconstructed and what are the choices made to filter them. The last section describes how Monte Carlo (MC) methods are used to simulate physics processes and emulate the detector response.

2.1 Tracks

Reconstructing the trajectories of charged particles allows us to measure their momentum and determining their charge. This is possible by analysing the hit patterns in the inner tracking system. In CMS this reconstruction is made with the Combinatorial Track Finder (CTF) algorithm [31]. The relevant steps for track generation are described below:

- Seed generation is made with hits at the pixel detector. A track seeds can be made with two or three hits. In the first case a know vertex or the beam spot is used to constrain the seed momentum. The parameters of each seed are estimated using the assumption that the trajectory is a helix, but it takes into account hit errors and multiple scattering [32].
- The track seed is extrapolated through the tracker layers with a combinatorial Kalman filter [33]. For each additional layer, the best matching hit if any is added

and track parameters are recomputed. This procedure continues until the last layer is reached [31].

- Ambiguity resolution may be necessary since it is possible to have the same track being reconstructed from different seeds, or a seed may results in more than a single trajectory candidate. To resolve this possible double counting, when considering a pair of tracks with more than 50% of shared hits, we discard the one with the fewer amount of hits. In case of equal number of hits the one with lowest χ^2 is kept.
- After the track building and cleaning stages are done final refitting is performed. This procedure is aimed at removing possible bias by constrains at the seed forming stage. A standard Kalman filter and smoother are used.

The process of track finding is repeated up to six times where the hits for each successfully reconstructed track are removed for the next iteration. Using early Large Hadron Collider (LHC) data and a dataset of pions and muons it was possible to estimate that the tracking efficiency is $> 98\%$ for all track $p_{\mathrm{T}} > 500\,\mathrm{MeV}$ and $> 99\%$ for tracks with $p_{\mathrm{T}} > 2, \mathrm{GeV}$ [34].

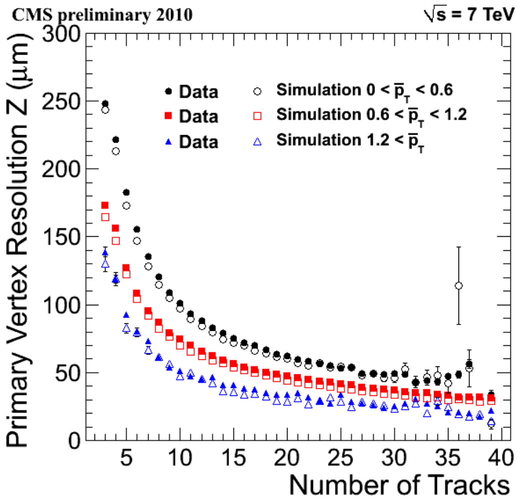
2.2 Vertex Reconstruction

The LHC can produce extreme collision intensities which are obtained partially by having multiple collisions happening at each bunch crossing. As it has been discussed in section 1.1.1 an average of 21 simultaneous collisions happened per bunch crossing in the CMS experiment during 2012. In this environment, it is crucial to identify the Primary Vertex (PV) and the particles that come from it. This information can then be used to reject particles coming from other additional collisions and to identify displaced vertices which can be the signature of long lived particles like b-mesons.

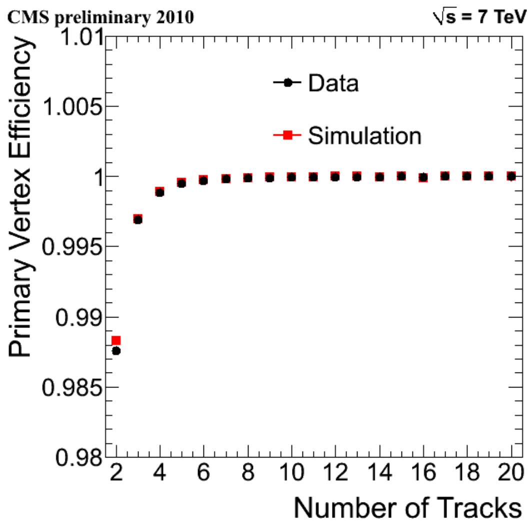
The individual tracks are reconstructed making use of the inner tracker. Each vertex is initially seeded by two tracks with separation in z less than $1\,\mathrm{cm}$. Then remaining track are clustered to the vertex seeds with the Deterministic Annealing (DA) algorithm [35]. After the clustering process is done, the position of each vertex is recomputed using the Adaptive Vertex Fitter (AVF) algorithm [36]. In this algorithm weights, w_i are assigned to each track according to how compatible they are with the fitted vertex position. Weight vary from 1 to 0, being that track assigned weights of close 1 are highly compatible with the vertex and close 0 would be given to low compatibility tracks. Then we can define the number of degrees of freedom of the new fit as:

$$n_{dof}(vertex) = 2 \sum_{i}^{tracks} w_i - 3 \tag{2.1}$$

This variable can be used to distinguish real proton-proton interactions from misclustered vertices, since it is correlated with the number of tracks compatible with that specific vertex [37]. The vertex position and resolution have been measured with LHC data and compared with simulation. The resulting plots can be found in figure 2.1 as a function of number of tracks.



(a)



(b)

Figure 2.1: (a) Primary vertex resolution in the z coordinate as function of the number of associated tracks. Results are give for three ranges of average track p_T. (b) Primary vertex efficiency as a function of the number of associated track [37]

The PV is defined as the vertex with highest sum of associated tracks p_T squared. In situations were no vertex can be reconstructed, like if there is a tracking failure, the beam spot position is assumed. Knowing precisely the interaction point allows to determine particle candidate quantities relative to it which allow for better object identification and pile-up control.

Most CMS analysis, including the ones presented in this thesis, require explicitly that a good vertex is reconstructed with the following characteristics:

- Real reconstructed vertex from tracks, not the beam spot.
- A minimum number of degrees of freedom: $n_{dof} > 4$.
- Collision must be near the interaction region. We require longitudinal distance to be $|z| <= 24\,\text{cm}$ (longitudinal impact parameter).
- Collision must be close to the beam line. Radial distance to beam line: $d_{xy} < 2\,\text{cm}$ (transverse impact parameter).

2.3 Particle Flow

The Particle Flow (PF) algorithm [38–40] is used in the CMS experiment with the objective of reconstructing every stable particle produced in the event. This is achieved by combining information from all CMS sub-detectors in order to identify electrons, photons, muons, charged hadrons and neutral hadrons and measure their direction, energy and type. The identified particles can in turn be used in jet clustering, determining the missing transverse energy, reconstructing and identifying taus, calculating particle isolation, identifying b-quark jets, etc.

The CMS experiment is very well suited for this approach since it is equipped with a high precision silicon tracker which is immersed in uniform axial magnetic field and its dual calorimeter design with high hermeticity and resolution. The tracker system allows very precise direction/momentum reconstruction for charged particles, down to transverse momentum as low as 150 MeV. The high granularity of the Electromagnetic Calorimeter (ECAL) allows for photons to be identified through deposit separation even inside high energy jets. In turn electrons can be reconstructed by combining their track and the energy deposits of the electron itself and its emissions, this algorithm will be explained further in section 2.5. The tracker information also allows to separate charged and neutral hadrons in close proximity, a task which is not possible with just the Hadronic Calorimeter (HCAL) due to its coarser granularity. We can determine the charged hadron momentum from the track information, and then, by removing its deposit from the calorimeter system we can determine the neutral hadron deposits. In areas outside the tracker and/or ECAL coverage, measurements are more coarse since we have less information available.

The clustering is performed separately in the ECAL and HCAL algorithm. We start by identifying *seed clusters* which are local maxima of calorimeter cell energy deposits. We add neighbouring cell into *topological clusters* if their energy deposit is bigger than two standard deviations of the electronics noise. This value was determined to be 80 MeV for the ECAL barrel, up to 300 MeV for the ECAL endcap and 800 MeV for the HCAL. The energy of each cell may be shared between multiple clusters.

Tracks and clusters are PF elements that need to be linked together to reconstruct the particle they came from and also to avoid double counting. We pair elements based on a metric of distance between elements and if compatible we merge them into *blocks* which can interpreted as particle candidates. As an example, a pair of a track and energy cluster on the calorimeter system would be linked if you could extrapolate the track to the cluster volume.

2.4 Isolation

To reduce the probability of misidentification of a lepton coming from Quantum Chromodynamics (QCD) jets as opposed to one coming from the hard scattering we can require isolation [41, 42]. We compute the isolation by summing the transverse momenta of all particles inside a cone around the selected lepton. In this sum we include all charged particles, neutral hadrons and photons. But we do not want to include the Pile-Up (PU) contribution to this sum so we only include the charged candidates with an impact parameter smaller than 0.1 cm. Different methods are used for each particle to estimate and subtract the neutral component of the PU depending on Particle Object Group (POG) recommendations.

Normally, for physics analysis we defined the more meaningful *relative isolation* as $I_{rel} = I/p_{\mathrm{T}}^{lepton}$. By using, a quantity that is relative to the the candidate p_{T} and not an absolute cut we avoid wrongly accepting low energy candidates or rejecting an high energy candidates. In the next sections the steps taken to calculate this quantity for each particle candidate are explained.

2.5 Electrons

In the CMS experiment electrons are reconstructed by matching energy clusters in the ECAL with tracks coming from the inner tracking system. Unfortunately, electrons can loose and disperse significant amounts of energy until they reach the ECAL. While they transverse the inner tracker they may emit photons through bremsstrahlung and in turn these photon can convert to e^+e^- pairs. About 35% of the electron radiate at least 70% of their energy in this way [43]. This spread of energy is mostly in ϕ due to the applied magnetic field [44]. Dedicated algorithms were developed to combine the the ECAL energy deposits, into a so called *supercluster*, of the initial electron and its emissions.

Different algorithms are used in the barrel and endcaps regions. In the barrel region we explore the simple $\eta - \phi$ geometry with the *hybrid clustering algorithm*. The procedure starts by identifying *seed crystals* with $E_T > 1\,\mathrm{GeV}$. A domino shaped cluster is formed around this seed in the η direction of 3×1 or 5×1 crystals centred at the seed. Additional dominoes are added in both ϕ direction in an attempt to collect the bremsstrahlung emissions up to $\Delta\phi \approx 0.3\,\mathrm{rad}$. Any domino with energy below $100\,\mathrm{MeV}$ is disregarded. The resulting additional sub-clusters must have its own seed with $E_T > 350\,\mathrm{MeV}$ and they are all combined to form the final *supercluster*.

In the encaps the *Multi-*5×5 *algorithm* is used. In this region of the detector the geometry is more complex and does not follow a simple $\eta - \phi$ symmetry. The seeds for the this clustering procedure are the crystals which are local maxima over their four direct neighbours and have a deposit of $E_T > 0.18\,\mathrm{GeV}$. Then, and starting with the seeds with highest E_T, we collect the energy around them into clusters of 5×5 crystals. We then search for similar seeds and form clusters that can overlap within $\Delta\eta < 0.07$ and $\Delta\phi < 0.3\,\mathrm{rad}$ of the initial seed. Those clusters are then combined into a single *supercluster* which needs to have at least $E_T > 1\,\mathrm{GeV}$. The *supercluster* is then extrapolated to the ECAL preshower by clustering the energy within $\Delta\eta < 0.15$ and $\Delta\phi < 0.45$ around the most energetic cluster and adding it to the *supercluster* itself [41].

In order to reconstruct the electron track we need to take into account the bremsstrahlung emissions. The CTF algorithm is not appropriate for this purpose so a different track-finding algorithm had to be developed. For high p_T electrons we use the ECAL supercluster energy deposit weighted mean impact point as a seed. If we combine this information with the determined E_T we can define two $\eta - \phi$ search regions in the pixel detector depending on the charge hypothesis. If we find two compatible hits, the electron trajectory is updated. From this point normal track building is performed but instead of a Kalman filter algorithm we use a Gaussian Sum Filter (GSF) algorithm [45]. This method performs better in the presence of non-Gaussian losses like the one coming from the bremsstrahlung emissions.

The typical background to real electrons are collimated hadronic jets, like from π^0 and $\pi^\pm$ overlap or from $\pi^\pm$ showers [43]. There are many useful variables that may be used to reduce such background and are often used in *electron identification* criteria:

- $\Delta\eta_{in}$ and $\Delta\phi_{in}$, are the distance between the track direction at the vertex and extrapolated to the ECAL and supercluster.
- $\sigma_{i\eta i\eta}$ is the energy-weighted η width of the cluster. For real prompt electrons this is normally small since this quantity is not significantly affected by the magnetic field.
- H/E is the ration of hadronic to electromagnetic energy in the region of the seed cluster.

Distributions of these variables for simulated electrons and jets can be found in figure 2.2.

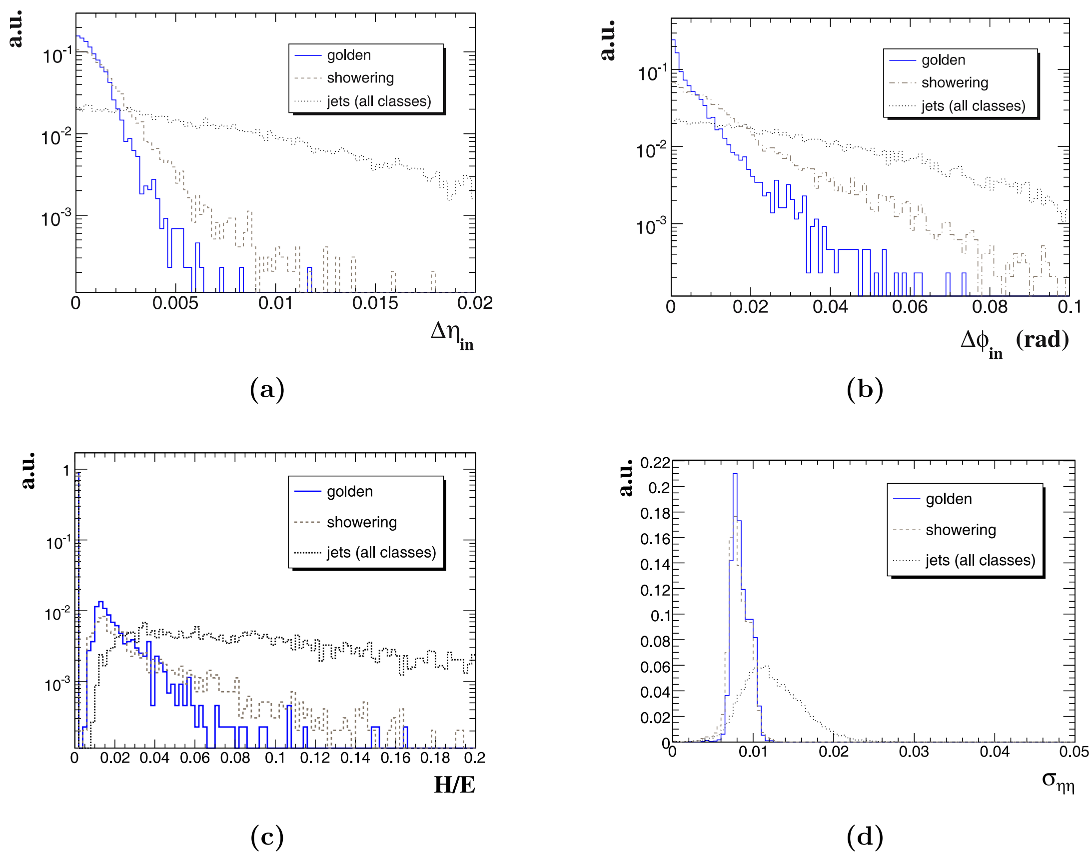


Figure 2.2: Distributions for (a) $\Delta\eta_{in}$, (b) $\Delta\phi_{in}$, (c) H/E and (d) $\sigma_{i\eta i\eta}$. Here *golden electrons* are those who emit minimal bremmstrahlung photons, *showering* are electrons that lose a large faction of their energy in emissions and *jets* are the typical distributions for hadronic jets. [43]

2.5.1 Isolation

For electrons we calculate isolation with the *effective area corrected isolation* method over a cone of $\Delta R < 0.3$ around the electron. For the neutral PU subtraction we uses a look-up table of effective areas according to electron $|eta|$ which is multiplied by the estimated neutral PU energy density by unit of effective area. The definition for this isolation can be found in equation 2.2.

$$I = \sum^{\substack{\text{charged}\\\text{non-pileup}}} p_{\mathrm{T}} + \max\left(0, \sum^{\text{neutral}} p_{\mathrm{T}} + \sum^{\text{photon}} p_{\mathrm{T}} - \rho(\text{lepton}) \times \text{Eff. Area}(\text{lepton})\right) \quad (2.2)$$

2.5.2 Veto electrons

We define *veto electrons* as an electron candidate with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.4$ which passes the CMS Electron/Gamma POG [46] requirements of the cut based electron Identification (ID) *veto electron* working point. A summary of these conditions can be found in table 2.1.

Variable	Barel	Endcap
$\|\Delta\eta(track, supercluster)\|$	< 0.007	< 0.1
$\|\Delta\phi(track, supercluster)\|$	< 0.8	< 0.7
$\sigma(i\eta, i\eta)$	< 0.01	< 0.03
H/E	< 0.15	-
$\|d_0(vertex)\|$	< 0.04	< 0.04
$\|d_Z(vertex)\|$	< 0.2	< 0.2
$\frac{PF_{isolation}}{p_\perp}$ for $\Delta R_{cone} = 0.3$	< 0.15	< 0.15

Table 2.1: Details of the CMS Electron-Gamma POG recommendations for a *veto electron*. Here barrel is defined as $|\eta_{supercluster}| <= 1.479$ and endcap is $1.479 < |\eta_{supercluster}| < 2.5$.

2.5.3 Tight electrons

We also define *tight electrons* as an electron candidate with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$ which passes the CMS Electron/Gamma POG requirements of the cut based electron ID *tight electron* working point. This working point is similar to the 2011 very tight WP70 working point. A summary of these conditions can be found in table 2.2.

Variable $p_{\mathrm{T}} > 20 (p_T <= 20)$	Barel	Endcap
$\lvert\Delta\eta(track, supercluster)\rvert$	< 0.004	< 0.005
$\lvert\Delta\phi(track, supercluster)\rvert$	< 0.3	< 0.2
$\sigma(i\eta, i\eta)$	< 0.01	< 0.03
H/E	< 0.12	< 0.10
$\lvert d_0(vertex)\rvert$	$< 0.02\,\mathrm{cm}$	
$\lvert d_Z(vertex)\rvert$	$< 0.1\,\mathrm{cm}$	
$\lvert\frac{1}{E} - \frac{1}{p}\rvert$	< 0.05	
$\frac{PF_{isolation}}{p_\perp}$ for $\Delta R_{cone} = 0.3$	< 0.10	$< 0.10(0.07)$
Conversion rejection: vertex fit probability	$< 1 \times 10^6$	
Conversion rejection: missing hits	$= 0$	

Table 2.2: Details of the CMS Electron-Gamma POG recommendations for a *tight electron*. Here barrel is defined as $|\eta_{supercluster}| <= 1.479$ and endcap is $1.479 < |\eta_{supercluster}| < 2.5$.

2.6 Muons

Muon track reconstruction starts independently at the inner-tracker (*tracker track*) and in the muon systems (*standalone muon track*) [42]. Then this information can be combined into a single muon track in two possible ways.

Global Muon reconstruction is an *outside-in algorithm.* We starts by finding tracker track match for each standalone muon track. This is done by propagating the match candidate pair to a common surface and comparing track parameters. For each matched pair, a *global-muon fit* is performed using all hits from the two tracks using a Kalman-filter algorithm [33]. For muons of $p_{\mathrm{T}} \gtrsim 200\,\mathrm{GeV}/c$, it has been showed that a *global-muon fit* improves the momentum resolution compared to a *tracker-only fit* [10, 47].

Tracker Muon reconstruction is an *inside-out algorithm.* In this method we start by selecting all tracker tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ and $p > 2.5\,\mathrm{GeV}$. We extrapolate those tracks to the muon system while taking into account the magnetic field, energy loss and scattering. If we find a match with at least one muon segment in the muon system (track stub in the Drift Tube (DT) or Cathode Strip Chamber (CSC)) this tracker track now becomes a Tracker Muon.

Tracker muon reconstructions is more efficient than the global muon reconstruction at low momenta at $p \lesssim 5\,\text{GeV}$. This difference is due to tracker muons reconstruction only requiring one segment on the muon system. While global muon reconstruction is more efficient for higher energies where the muons are more likely to pass several muon stations.

Muons can be also be classified as prompt of non-prompt. The prompt muons are the ones produced directly in the hard process like the decays of vector bosons or quarkonia particle decays. On the other hand, non-prompt muons typically come from in-flight decays of light hadrons, from taus or heavy quark decays.

When reconstructing global muons, its unlikely to find non-prompt muons but we may have hadronic activity "punching-through" the calorimeter system and appearing in the muon system. To reduce this types of background we can use different muon identification criteria.

Studies with the CMS detector have been performed to asses muon reconstruction efficiency [42]. Muon was defined as candidate with global fit track using tracker and muon chamber hits with a χ^2 per degree of freedom of less than 10. This fit must include at least one segment in the muon chamber, track segments in at least 2 muon stations, use more than 10 hits in the inner tracker of which at least one in the a pixel layer and finally a small transverse impact paramenter $|d_{xy}| < 2\,\text{mm}$. The efficiency for such a criteria has been measured both in data and Monte Carlo using $J/\psi \to \mu^+\mu^-$ and $Z \to \mu^+\mu^-$ and for $p_\text{T} > 10\,\text{GeV}$ it plateaus at 96-99%.

2.6.1 Isolation

For muons we use the *combined isolation* over a cone of $\Delta R < 0.4$ around the muon. For neutral PU subtraction we use the charged PU component inside the cone and multiply it by a factor of 0.5 which is determined from simulation. The definition for this isolation can be found in equation 2.3.

$$I = \sum^{\substack{\text{charged}\\\text{non-pileup}}} p_{\mathrm{T}} + \max\left(0, \sum^{\text{neutral}} p_{\mathrm{T}} + \sum^{\text{photon}} p_{\mathrm{T}} - \frac{1}{2}\sum^{\substack{\text{charged}\\\text{pileup}}} p_{\mathrm{T}}\right) \tag{2.3}$$

2.6.2 Loose Muons

We can define *loose muon* using the cut based definitions recommend by the CMS Muon POG [48] with the same name, where we require the muon candidate to be a PF muon which is also a tracker or global muon. We exclude only standalone muons which are only $\approx 0.01\%$ of the PF muons. Additionally we require the muon candidate to have $p_{\mathrm{T}} > 10\,\mathrm{GeV}$, $|\eta| < 2.1$ and relative combined isolation < 0.2.

2.6.3 Tight Muons

We can also define *tight muon* as a muon candidate with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, $|\eta| < 2.1$ passing relative combined isolation < 0.12. Additionally, we require compatibility of being produced at the primary vertex by requiring $d_{xy} < 0.045\,\mathrm{cm}$ and $d_z < 0.2\,\mathrm{cm}$. We also require the muon to pass the CMS Muon POG recommended cut based *tight muon* identification criteria that requires the candidate to be a PF muon which is also a global muon. Where the the global track fit has at least one muon chamber hit and $\chi^2/ndof < 10$. The presence of muon segments in at least two chambers, at least five tracker layers with hits and at least one pixel hit.

2.7 Jets

When we collide hadrons the most probable hard processes will be the scattering quarks and gluons. However, these do not reach our detectors. They quickly hadronize and fragment generating a collimated spray of particles which is commonly referrer to as a jet. To determine the properties of this outgoing quarks and gluons we need therefore to look at the characteristics of their associated jets. To achieve this goal we need to combine the measured jet remnants in a way that preserves the physical properties of the original parton.

2.7.1 Jet Clustering

Jet clustering algorithms are sets of rules that allows us to combine particle candidates into a jets [49]. These algorithms normally are controlled by parameters that define how close particles need to be in order to be associated into a jet and a way to combine their momentum. However, a jet definition should be robust and provide consistent measurements about the parton. There are two major families of problems that may affect a jet algorithms. These problems appear when the number of jets in an event changes by adding a soft collinear gluon emissions (collinear safety) or by parton splitting (infrared safety).

In CMS we use a sequential recombination algorithm known as anti-k_T [50] which is both infrared and collinear safe. This algorithm starts be determining a measurement of distance between every pair of objects d_{ij} and a the distance of each object to the beamline d_{iB}. The definition of this distances can be found in equations 2.4 and 2.5 respectively.

$$d_{ij} = \min({p_{\mathrm{T}}}_i^{2p}, {p_{\mathrm{T}}}_j^{2p})\frac{\Delta R_{ij}^2}{R^2} \tag{2.4}$$

$$d_{iB} = {p_{\mathrm{T}}}_i^{2p} \tag{2.5}$$

Where ΔR is the separation the $\eta - \phi$ plane and R is the maximum diameter for the jet. The parameter p determines the type of algorithm. When p equal to 1 its a k_T algorithm, 0 for the Cambridge/Aachen algorithm and -1 for the anti-k_T.

After determining all the d_{ij} and d_{iB} we determine which is the minimum distance. If it is a d_{ij} we combine those two object and recalculate all the distances. If the minimum is a d_{iB} we declare i to be a *final state jet*, and remove it from the list of particles and recalculate all distances the again. The procedure continues until there are no more objects remaining.

The anti-k_T algorithm tend to cluster particles around a the hardest particle in a region which normally leads to a cone like jet area in the $\eta - \phi$ plane. In the VBF Higgs to invisible analysis the clustering is made over PF particle candidates using the implementation in the FASTJET software package [51]. The CMS recommended cone diameter size for 2012-13 analysis is of 0.5 while for 2015 is of 0.4.

2.7.2 Particle Flow Jet Identification

The CMS Jet-MET POG has defined criteria to reject fake, badly reconstructed, and noisy PF jets while keeping 98-99% real jets. All the presented analysis in this thesis we have used the recommended PF jet ID in the loose working point. In this working point all jets are required to have at least two constituents, and both neutral hadron fraction and a neutral Electromagnetic (EM) fraction to be below 99%. Additionally for jets inside the tracker acceptance with $|\eta| < 2.4$ we require the charged multiplicity and charged hadron fraction to be bigger than zero, and the charged EM fraction to be less than 99%.

2.7.3 Pileup Jet Identification

To identify if a PF jet has come from PU or from the primary vertex we make of a Boosted Decision Tree (BDT). This machine learning algorithm was trained with information about the trajectory of the tracks associated with the jet, the jet shape, and object multiplicity. In the presented analyses we have used the recommended loose working point of the *full BDT method*. This method was applied to each jet which would only be accepted if the BDT output score would pass the cuts defined in table 2.3 depending on jet p_{T} and η.

Jet p_T	Jet $\|\eta\|$	BDT_{score}
$20 < p_T \leq 30$	$\|\eta\| < 2.5$	> -0.80
$20 < p_T \leq 30$	$2.50 \leq \|\eta\| < 2.75$	> -0.85
$20 < p_T \leq 30$	$2.75 \leq \|\eta\| < 3.00$	> -0.84
$20 < p_T \leq 30$	$3.00 \leq \|\eta\| < 5.00$	> -0.85
$30 < p_T$	$\|\eta\| < 2.5$	> -0.80
$30 < p_T$	$2.50 \leq \|\eta\| < 2.75$	> -0.74
$30 < p_T$	$2.75 \leq \|\eta\| < 3.00$	> -0.68
$30 < p_T$	$3.00 \leq \|\eta\| < 5.00$	> -0.77

Table 2.3: Table of the minimum values of *full BDT method* score for a PF jet to be accepted as coming from the PV using a loose working point. Required minimum values have been binned in jet p_{T} and η.

2.7.4 Lepton cleaning

To avoid having leptons being miss reconstructed as jets we filter out all jets which are located at $\Delta R < 0.5$ to any veto electron or loose muons.

2.7.5 Jet Energy Corrections

When reconstructing a jet the clustered energy often does not match the parton energy that gave it origin. There are many reason for this effect like non-linearity of the calorimeters response, detector noise, overlap with problematic detector areas, additional energy from PU, miss calibration, etc. To fix this problem corrections are determined and applied to each jet in order to in average have an energy measurements that is equal to the original hadron. This corrections can be factorized into components as it is represented in equation 2.6 [52].

$$P^{\mu}_{\text{corr}} = C_{\text{offset}}({p_{\text{T}}}^{\text{raw}}, \eta) \cdot C_{\text{rel}}({p_{\text{T}}}^{\text{off}}, \eta) \cdot C_{\text{abs}}({p_{\text{T}}}^{\text{rel}}, \eta) \cdot P^{\mu}_{\text{raw}} \tag{2.6}$$

The C_{offset} term accounts for and subtracts the contribution of PU and noise in the detector measurements. Its value is determined by taking into account the specific event p_{T}-density expressed in the ρ variable, and the individual jet area A [53]. The event ρ is calculated as the median p_{T}-density of all jets present in the event. Since the median is taken it will not be affected by the presence of hard jets. Unfortunately, the Underlying Event (UE) activity has similar characteristic to the PU and should not be subtracted. To avoid this effect the correction takes the form of $\rho - \langle\rho\rangle_{UE} \cdot A$, where $\langle\rho\rangle_{UE}$ is the average expected UE contribution.

The C_{rel} term is applied to make the energy response flat as a function of η. It is applied to the offset corrected transverse momentum ${p_{\text{T}}}^{off}$. To determine its value the p_{T}-balancing method is used [52]. In this method we select a reference jet located in the central region where energy measurement is expected to be flat and a probe jet at any value of η. We can calculate the average of balance quantity as $(p_{\text{T}}^{\text{probe}} - p_{\text{T}}^{\text{reference}})/p_{\text{T}}^{\text{average}}$ which is used to determine to the correction to response in bins of jet η and dijet average p_{T}.

The C_{abs} term is intended to make the make the response uniform in p_{T}. It is applies to the η corrected transverse momentum ${p_{\text{T}}}^{rel}$ and is calculated using the Missing

Transverse Energy Projection Fraction (MPF) method [54]. In this method we use the good experimental resolution for leptons and photons in processes like $\gamma +$ jets and $Z +$ jets to infer on the properties of the recoil jets. Since this processes should not have Missing Transverse Energy (MET), if observed it can be used to calibrate the jet response for the jets present in the event.

The total uncertainty on the jet energy scale is obtained by summing in quadrature the estimated uncertainties of each one of the correction terms. The total uncertainty is in the range of $\approx 3-5\%$ depending on p_{T} and η [52].

2.8 Hadronic Taus

Taus can decay leptonically and hadronically. In leptonic decays the tau decays directly to an electron and two additional neutrinos. Therefore it is very difficult to identify such decays experimentally. On the other hand an hadronic tau decay produces a characteristic signature of a narrow jet containing an odd number of charged particles and additional neutral hadrons as well as a tau neutrino. In all the analysis presented in this thesis when referring to a tau we refer an hadronically decaying tau. The most probable decay modes have one or three charged π mesons and are summarized in table 2.4.

Decay Channel	Resonance	Mass [MeV]	Branching Fraction [%]
$\tau^{\pm} \to \pi^{\pm}\nu_{\tau}$			11.6
$\tau^{\pm} \to \pi^{\pm}\pi^{0}\nu_{\tau}$	ρ	770	26.0
$\tau^{\pm} \to \pi^{\pm}\pi^{0}\pi^{0}\nu_{\tau}$	a_1	1260	10.8
$\tau^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm}\nu_{\tau}$	a_1	1260	9.8
$\tau^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm}\pi^{0}\nu_{\tau}$			4.8
Other hadronic modes			1.7
Total			64.7

Table 2.4: Summary of the hadronic tau decay modes, with the branching fractions and intermediate resonances listed where relevant [55].

Reconstruction of hadronic tau neutrinos with PF is done by identifying the specific decay mode visible products. The approach is at the core of the Hadron Plus Stips (HPS) algorithm [56, 57]. It combines reconstructed charged hadrons with strips of clustered photons which are interpreted as π_0. The reconstructed system is constrained by the tau mass and intermediate resonances which results in a highly collimated jet when compared with a typical QCD jet.

2.8.1 Hadron Plus Stips Algorithm

The HPS algorithm utilizes PF candidates to reconstruct charged pions and photons resulting from neutral pions decay. These photon can convert in into electron-positron pairs in the tracker material. This factors are taken into consideration, as well as deflection in the magnetic field. We also attempt to find the intermediate resonances listed in table 2.4 has a handle to determine the tau decay channel. A tau neutrino is present on all decays and cannot be directly measured, this results in a smearing of the measured tau mass when considering only the visible products.

We seed the algorithm with PF anti-k_T jets with $R = 0.5$ where $p_{\mathrm{T}} > 14\,\mathrm{GeV}$ and $|\eta| < 2.5$. To search for the π^0 decay products we try to identify strips by clustering PF electrons and photons with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$. We start from the most energetic electromagnetic particle inside the jet area and make that the centre of our candidate strip. We look for other electromagnetic objects within a window of $\Delta\eta = 0.05$ and $\Delta\phi = 0.20$ of the centre of the strip. If an object is found it gets associated with the strip and its four-momentum gets recalculated. We repeat the procedure until we cannot find any more unassociated EM objects inside the strip area. If the final strip object has a mass compatible with a π^0, in the interval between $50 - 200\,\mathrm{MeV}$, and has $p_{\mathrm{T}} > 2.5\,\mathrm{GeV}$ it is kept. We then start the next strip clustering with the highest p_{T} electron or gamma not already belonging to a strip.

The charged pion candidates are required to have $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ and its track pass $d_z < 0.4\,\mathrm{cm}$ and $d_{xy} < 0.03\,\mathrm{cm}$ to the vertex associated with the highest p_{T} track in the jet, which is assumed to be the τ production vertex.

The following topologies are taken into account by the HPS algorithm:

1. *single hadron*: tries to identify tau decays into $\pi^{\pm}\nu_{\tau}$ or $\pi^{\pm}\pi^0\nu_{\tau}$ where the netral pion decay cannot be identified as a strips.

2. *One hadron + one strip*: tries to identify tau decays into $\pi^{\pm}\pi^0\nu_{\tau}$ where the π^0 decay photons are close together. In this case we are selecting the $\rho(770)$ intermediate resonance. The mass of the reconstructed τ_{had} is required to be in the interval $0.4 < m_{\tau_{had}} < 1.3\,\mathrm{GeV}$ for $p_{\mathrm{T}}^{\tau_{had}} < 200\,\mathrm{GeV}$. The upper limit in the mass window

can go up to 2.1 GeV for candidates with $p_{\mathrm{T}}^{\tau_{had}} > 200$ GeV to account for resolution effects.

3. *One hadron + two strip*: tries to identify tau decays into $\pi^{\pm}\pi^{0}\pi^{0}\nu_{\tau}$. In this case we are selecting the $a_1(1260)$ intermediate resonance. The mass of the reconstructed τ_{had} is required to be in the interval $0.4 < m_{\tau_{had}} < 1.2$ GeV for $p_{\mathrm{T}}^{\tau_{had}} < 200$ GeV. The upper limit in the mass window can go up to 2.0 GeV if the $p_{\mathrm{T}}^{\tau_{had}}$ increases above 200 GeV.

4. *Three hadrons*: tries to identify tau decays into $\pi^{\pm}\pi^{\mp}\pi^{\pm}\nu_{\tau}$. The hadrons are required to have mass in the interval $0.8 - 1.5$ GeV since we assume the $a_1(1260)$ intermediate resonance. Total charged is required to be one.

There is no dedicated search for $\pi^{\pm}\pi^{\mp}\pi^{\pm}\pi^{0}\nu_{\tau}$ or higher pion multiplicity decay modes. These topologies are reconstructed with the currently defined criteria.

All selected hadrons and strips are required to be inside of cone of $\Delta R < 2.8\,\mathrm{GeV}/p_{\mathrm{T}}^{\tau_{had}}$. The cone size is is constrained to the interval $\Delta R = 0.05 - 0.10$.

2.8.2 Isolation and Discriminants

Isolation for taus is calculates in a similar way to electrons ans muons. The isolation variable is defined by summing the p_{T} of all PF hadron and photon candidates in a cone of $\Delta R < 0.5$ around the tau axis. Here the charged hadron tracks are required to have $d_z < 2$ cm to the tau production vertex. We can subtract the contribution to isolation coming from PU estimating its density in a cone of $\Delta R < 0.8$ around the tau and considering track with $d_z > 2$ cm. All tau constituents are ignores in this sums. Working points have been defined for loose, medium and tight isolation [57].

Electrons can be reconstructed as taus when they make isolated deposits in the calorimeter or emit enough energy via bremsstrahlung to form a strip. A BDT has been trained with a set of variables similar to the ones used in electron identification to exclude such miss reconstructions. Similarly to isolation three working points have been defined [56, 57].

Muons are less likely to be reconstructed as a tau. We can exclude such tau candidates by requiring that the track of the leading charged hadron is not also a tracker muon. This discriminator also has three possible working points [56, 57].

We can now define an hadronic tau candidate as a PF tau candidates with $p_{\mathrm{T}} > 20$ GeV, $|\eta| < 2.3$ and $d_z < 0.2$ cm to the primary vertex. We require that the candidate passes decay-mode identification, tight isolation and finally tight discriminators against electrons and muons.

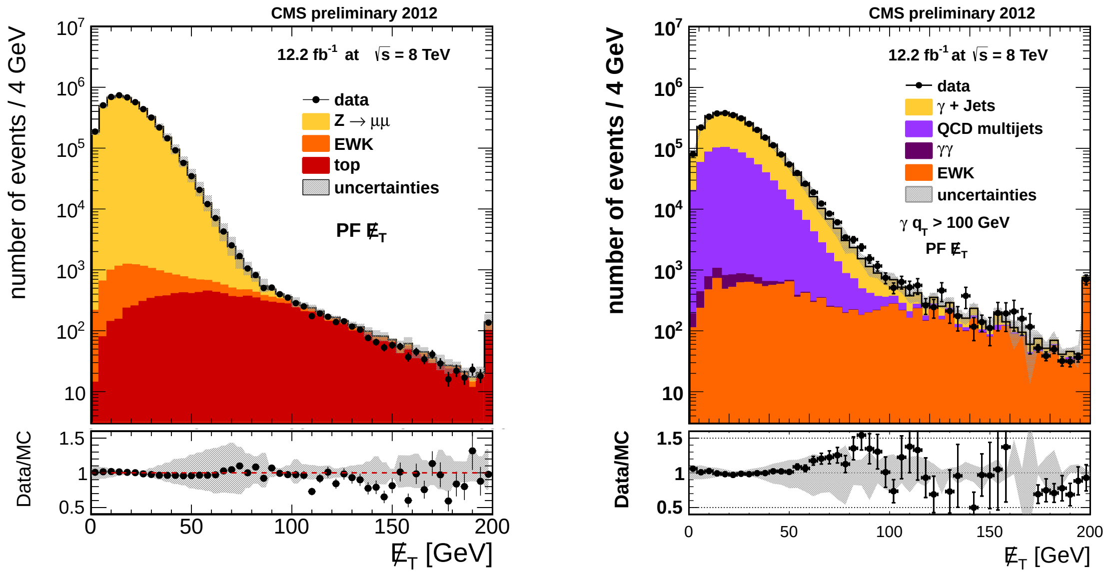
2.9 Missing Transverse Energy

The Standard Model describes neutrinos as particles which only interact via the weak force. They can pass through our detectors without interacting and therefore not allowing any direct measurement. Many new models describe additional particles that would also be able to escape detection by leaving very small or no energy deposits in our experiments. The appearance of such particles can only be inferred through the measurement of an imbalance of transverse momentum of all detected particles. These effect can be quantified as the negative sum off all visible particle candidates transverse momentum in an event.

The magnitude of that vector is referred to as Missing Transverse Energy (MET). Particle flow methodology provides a complete list of objects candidates in the event with excellent resolution achieved by combining all available information. Making it well suited to be the input for MET calculation. Although CMS has an excellent individual particle resolution the calculation of MET is affected by the combined resolution of the measurement of all particles in the event. Figure 2.3 shows the distributions of PF MET for both data and simulation for event selections of $Z \to \mu\mu$ and $\gamma + jets$ processes at $\sqrt{s} = 8\,\mathrm{TeV}$.

Both photons and muons energy measurements have good resolution in the CMS experiment and these processes do not involve real MET. The observed distribution in both plots are predominantly shaped by the energy resolution of jet energy measurement.

During data taking issues with the detector or data acquisition can happen creating anomalously high MET and rendering this events unusable. The groups responsible for each part of the detector and individual physics objects, check the data after it was taken to find if such problems have occurred. After this problems are identified they produce software event filters for analysts to be able to remove this problematic events. The CMS JET-MET POG compiled a list of the recommended filters for analysis using 2012-13 data to remove events affected by energy deposits from beam halo, noise in HCAL readout electronics, particles directly hitting the ECAL photodiodes, track reconstruction problems and finally ECAL and HCAL miss timed laser calibration sequence. This filters have been used in both prompt and parked VBF Higgs to invisible analyses.



(a) (b)

Figure 2.3: Distributions of the particle flow $E_{\mathrm{T}}^{\mathrm{miss}}$ in a) $Z \to \gamma\gamma$ and b) γ+jets events in $\sqrt{s} = 8\,\mathrm{TeV}$ data and simulation. The uncertainty in the muon, photon, jet and neutral hadron energy responses is showed by the shaded band [58].

There are many factors that affect MET response and resolution. These include zero suppression thresholds which dictate the minimum energy a calorimeter cell will report, dead or non-instrumented regions of the detector and reconstruction inefficiencies. Techniques have been developed to correct both response and resolution when using PF MET [59]. These corrections include accounting for the bias in response due to using incorrect energy scale of the jets, and reducing the impact of pileup on the resolution [58].

In the VBF Higgs to invisible analysis MET calculated without including muons is used in the offline analysis and trigger. This choice allows to investigate the irreducible background of $Z \to \nu\nu$ by using $Z \to \mu\mu$ as a proxy. Muons are vetoes in the signal region and required in the control regions to recover the usual MET value.

2.10 Monte Carlo Simulation

Monte Carlo (MC) methods are a class of computer algorithms that rely on random sampling to obtain numerical results. This type of methods are especially useful in problems with many coupled degrees of freedom where it is difficult to perform analytical calculations. In particle physics these methods are often used to simulate physics processes, their interaction with detectors and the obtained response.

To simulate one event in the CMS experiment we first start by the physics process itself. It can be split into two sub-processes: hard scattering and hadronization. There are many purpose built software programs that will perform each of these steps. An illustration of how the simulation of proton-proton collision is done with MC programs can be found in figure 2.4. A review of the available generators for LHC physics can be found in reference [60].

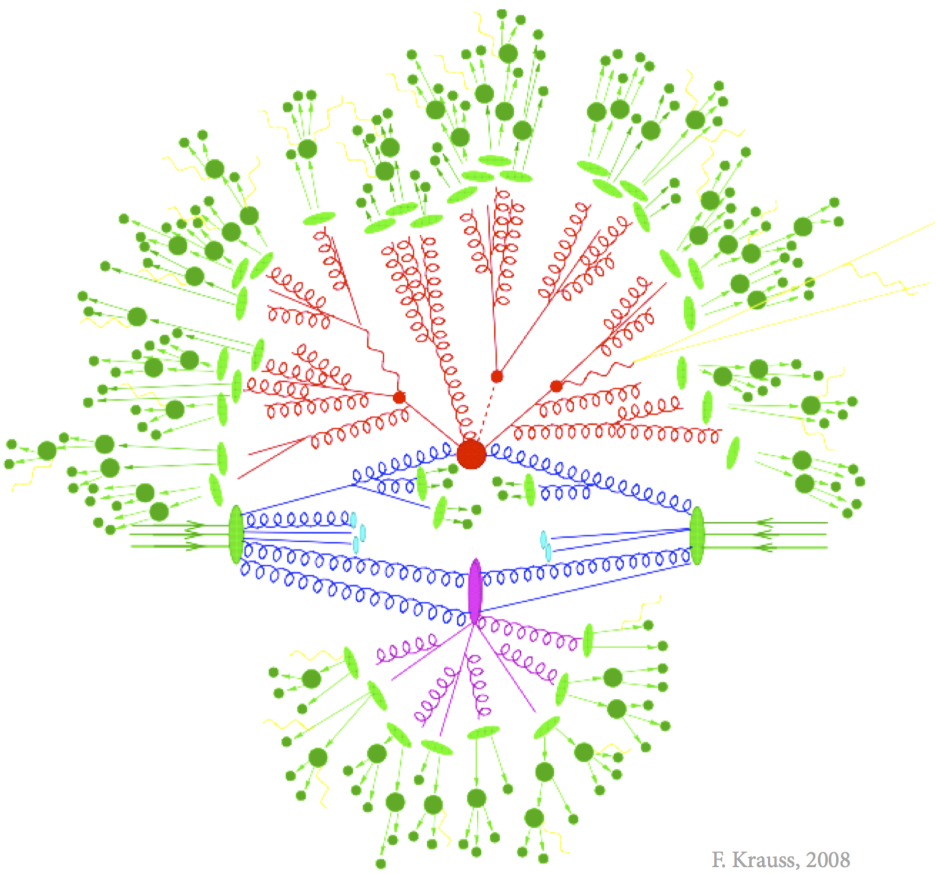


Figure 2.4: Illustration of a proton-proton collision as implemented in some MC event generators [61]. Sub-processes are represented, the hard-scattering in the center of the diagram, the parton showering in red, hadronization in green. We can also observe the UE interaction and its showering in purple.

General purpose particle physics event generators like PYTHIA 8 [62, 63], HERWIG++ [64] and SHERPA [65] are able to do both hard scattering and hadronization steps for a wide variety of physics processes. Typically these programs are restricted to $2 \to 2$ and $2 \to 1$ hard processes calculated at Leading Order (LO).

There are many other dedicated matrix-element generators, like MADGRAPH 5 [66], ALPGEN [67] and also SHERPA that focus on the hard process simulation. These programs provide $2 \rightarrow X$ hard scattering where a higher number of partons in the final state is possible. Some generators have also implemented Next to Leading Order (NLO) calculations, which provide better kinematics discription and lower uncertainties. Two examples of such generators are aMC@NLO [68] and POWHEG [69, 70]. The simulated parton level events then need to be passed to a general purpose event generators for hadronization.

Overlapping of the phase-space description of matrix-element and showering programs needs to be avoided when simulating multi-jets events. This problem rises from software like PYTHIA or HERWIG describing parton radiation as a Markov Chain process based on Sudakov form factors. This approach is only formally correct in the limit of soft and collinear emissions. On the other hand Matrix Element (ME) programs like MadGraph work well for the hard scattering high energy limit but diverges when the partons become soft or collinear.

There are a few jet-parton matching schemes developed to account for this overlap [71]. Showering can be vetoed and the event reweighed accordingly, like in the CKKW scheme [72–74] or events can be rejected altogether like in the MLM scheme [75]. Depending on the generator used for the showering, different schemes are implemented and care must be taken in the definition of the matching parameters.

Most event generators can be finely tuned so important aspects of the simulation can be adjusted to experimental conditions. As an example, in the CMS experiment PYTHIA is used with the Z2 tune, which was produced using measurements made using minimum bias data at the Tevatron and at the LHC [76].

After the physics event is simulated, the interaction with the detector and the corresponding electronics response is estimated using a precise model of the experiment. In the CMS experiment GEANT 4 [77, 78] software is used for this task which also relies heavily on MC methods.

Bibliography

[1] O. S. Bruning, P. Collier, P. Lebrun et al., “LHC Design Report Vol.1: The LHC Main Ring”,.

[2] L. Evans and P. Bryant, “LHC Machine”, *JINST* **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.

[3] LEP Injector Study Group Collaboration, “LEP Design Report, Volume 1: The LEP Injector Chain”,.

[4] P. Mouche, “CERN Document Server: Overall view of the LHC”, *http://cds.cern.ch/record/1708847* (Accessed Dec, 2015).

[5] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *Journal of Instrumentation* **3** (2008), no. 08, S08003.

[6] CMS Collaboration, “The CMS experiment at the CERN LHC”, *Journal of Instrumentation* **3** (2008), no. 08, S08004.

[7] LHCb Collaboration, “The LHCb Detector at the LHC”, *Journal of Instrumentation* **3** (2008), no. 08, S08005.

[8] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *Journal of Instrumentation* **3** (2008), no. 08, S08002.

[9] C. Lefèvre, “The CERN accelerator complex”, *https://cds.cern.ch/record/1260465* (Accessed Dec, 2015).

[10] CMS Collaboration, “CMS Physics: Technical Design Report Volume 1: Detector Performance and Software”, technical report, Geneva, (2006). There is an error on cover due to a technical problem for some items.

[11] M. Bajko, F. Bertinelli, N. Catalan Lasheras et al., “Report of the task force on the incident of 19th September 2008 at the LHC”, *LHC Project Report* **LHC-PROJECT-REPORT-1168** (2009).

[12] CMS Collaboration, "Multi-year Collisions Plots", *https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults* (Accessed Dec, 2015).

[13] CMS Collaboration, "Average Pileup in 2012", *https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults* (Accessed Dec, 2015).

[14] CMS Collaboration, "CMS Compact Muon Solenoid", *http://public.web.cern.ch/public/Objects/LHC/CMSnc.jpg* (Accessed Nov, 2010).

[15] CMS Collaboration, "CMS, the Compact Muon Solenoid: Technical proposal",.

[16] CMS Collaboration, C. Collaboration, "CMS, tracker technical design report". 1998.

[17] CMS Collaboration, "Addendum to the CMS tracker TDR",.

[18] CMS Collaboration, "The CMS electromagnetic calorimeter project : Technical Design Report",.

[19] CMS Collaboration, "Changes to CMS ECAL electronics : addendum to the Technical Design Report",.

[20] USCMS, ECAL/HCAL Collaboration, "The CMS barrel calorimeter response to particle beams from 2-GeV/c to 350-GeV/c", *Eur. Phys. J.* **C60** (2009) 359–373, doi:10.1140/epjc/s10052-009-0959-5,10.1140/epjc/s10052-009-1024-0. [Erratum: Eur. Phys. J.C61,353(2009)].

[21] CMS Collaboration, "The CMS hadron calorimeter project : Technical Design Report",.

[22] CMS Collaboration, "The CMS magnet project : Technical Design Report",.

[23] CMS Collaboration, "The CMS muon project : Technical Design Report",.

[24] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* **B716** (2013) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[25] CMS Collaboration, "CMS The TriDAS Project : Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger",.

[26] CMS Collaboration, "CMS TriDAS project : Technical Design Report, Volume 1: The Trigger Systems",.

[27] CMS Collaboration Collaboration, "Data Parking and Data Scouting at the CMS Experiment",.

[28] CMS Collaboration, "CMS computing: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 2005. Submitted on 31 May 2005.

[29] CMS Collaboration, "CMS Technical Design Report for the Level-1 Trigger Upgrade",.

[30] CMS Collaboration, "Technical proposal for the upgrade of the CMS detector through 2020", Technical Report CERN-LHCC-2011-006. LHCC-P-004, CERN, Geneva, (Jun, 2011).

[31] W. Adam, B. Mangano, T. Speer et al., "Track Reconstruction in the CMS tracker", Technical Report CMS-NOTE-2006-041, CERN, Geneva, (Dec, 2006).

[32] S. Cucciarelli, M. Konecki, D. Kotlinski et al., "Track reconstruction, primary vertex finding and seed generation with the Pixel Detector", Technical Report CMS-NOTE-2006-026, CERN, Geneva, (Jan, 2006).

[33] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth.* **A262** (1987) 444–450, doi:10.1016/0168-9002(87)90887-4.

[34] CMS Collaboration, "Measurement of Tracking Efficiency",.

[35] K. Rose, "Deterministic Annealing for Clustering, Compression, Classification, Regression, and related Optimization Problems", *Proceedings of the IEEE* **86** (1998) 2210.

[36] W. Waltenberger, R. Frühwirth, and P. Vanlaer, "Adaptive vertex fitting", *Journal of Physics G: Nuclear and Particle Physics* **34** (2007), no. 12, N343.

[37] CMS Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions",.

[38] CMS Collaboration Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", Technical Report CMS-PAS-PFT-10-001, 2010.

[39] CMS Collaboration Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", Technical Report CMS-PAS-PFT-09-001, CERN, 2009. Geneva, (Apr, 2009).

[40] CMS Collaboration Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", Technical Report CMS-PAS-PFT-10-002, CERN, Geneva, (2010).

[41] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[42] T. C. collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s}$= 7 TeV", *Journal of Instrumentation* **7** (2012), no. 10, P10002.

[43] S. Baffioni, C. Charlot, F. Ferri et al., "Electron reconstruction in CMS", Technical Report CMS-NOTE-2006-040, CERN, Geneva, (Feb, 2006).

[44] E. Meschi, T. Monteiro, C. Seez et al., "Electron Reconstruction in the CMS Electromagnetic Calorimeter", Technical Report CMS-NOTE-2001-034, CERN, Geneva, (Jun, 2001).

[45] W. Adam, R. Frühwirth, A. Strandlie et al., "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC", *Journal of Physics G: Nuclear and Particle Physics* **31** (2005), no. 9, N9.

[46] CMS Collaboration, "Electron Reconstruction and Identification at $\sqrt{s} = 7$ TeV",.

[47] C. Collaboration, "Performance of CMS muon reconstruction in cosmic-ray events", *Journal of Instrumentation* **5** (2010), no. 03, T03022.

[48] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV",.

[49] G. P. Salam, "Towards Jetography", *Eur. Phys. J.* **C67** (2010) 637–686, doi:10.1140/epjc/s10052-010-1314-6, arXiv:0906.1833.

[50] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[51] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J.* **C72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[52] T. C. collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *Journal of Instrumentation* **6** (2011), no. 11, P11002.

[53] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett.* **B659** (2008) 119–126, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[54] CDF Collaboration, "The Dijet angular distribution in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **69** (1992) 2896–2900, doi:10.1103/PhysRevLett.69.2896.

[55] Particle Data Group Collaboration, "Review of Particle Physics*", *Phys. Rev. D* **86** (Jul, 2012) 010001, doi:10.1103/PhysRevD.86.010001.

[56] CMS Collaboration, "Performance of tau-lepton reconstruction and identification in CMS", *JINST* **7** (2012) P01001, doi:10.1088/1748-0221/7/01/P01001, arXiv:1109.6034.

[57] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS", arXiv:1510.07488.

[58] "Performance of Missing Transverse Momentum Reconstruction Algorithms in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV with the CMS Detector",.

[59] T. C. collaboration, "Missing transverse energy performance of the CMS detector", *Journal of Instrumentation* **6** (2011), no. 09, P09001.

[60] A. Buckley et al., "General-purpose event generators for LHC physics", *Phys. Rept.* **504** (2011) 145–233, doi:10.1016/j.physrep.2011.03.005, arXiv:1101.2599.

[61] F. Krauss, "Phenomenology at collider experiments", *http://www.ippp.dur.ac.uk/ krauss/Lectures/MonteCarlos/Pheno_5_08.pdf* (Accessed May, 2014).

[62] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[63] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[64] M. Bahr et al., "Herwig++ Physics and Manual", *Eur. Phys. J.* **C58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.

[65] T. Gleisberg, S. Hoeche, F. Krauss et al., "Event generation with SHERPA 1.1", *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007, arXiv:0811.4622.

[66] J. Alwall, M. Herquet, F. Maltoni et al., "MadGraph 5 : Going Beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[67] M. L. Mangano, M. Moretti, F. Piccinini et al., "ALPGEN, a generator for hard multiparton processes in hadronic collisions", *JHEP* **07** (2003) 001, doi:10.1088/1126-6708/2003/07/001, arXiv:hep-ph/0206293.

[68] J. Alwall, R. Frederix, S. Frixione et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[69] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[70] S. Alioli, K. Hamilton, P. Nason et al., "Jet pair production in POWHEG", *JHEP* **04** (2011) 081, doi:10.1007/JHEP04(2011)081, arXiv:1012.3380.

[71] S. Hoeche, F. Krauss, N. Lavesson et al., "Matching parton showers and matrix elements", in *HERA and the LHC: A Workshop on the implications of HERA for LHC physics: Proceedings Part A*. 2006. arXiv:hep-ph/0602031.

[72] S. Catani, F. Krauss, R. Kuhn et al., "QCD matrix elements + parton showers", *JHEP* **11** (2001) 063, doi:10.1088/1126-6708/2001/11/063, arXiv:hep-ph/0109231.

[73] F. Krauss, "Matrix elements and parton showers in hadronic interactions", *JHEP* **08** (2002) 015, doi:10.1088/1126-6708/2002/08/015, arXiv:hep-ph/0205283.

[74] S. Mrenna and P. Richardson, "Matching matrix elements and parton showers with HERWIG and PYTHIA", *JHEP* **05** (2004) 040, doi:10.1088/1126-6708/2004/05/040, arXiv:hep-ph/0312274.

[75] M. L. Mangano, M. Moretti, F. Piccinini et al., "Matching matrix elements and shower evolution for top-quark production in hadronic collisions", *JHEP* **01** (2007) 013, doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.

[76] CMS Collaboration, "Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and Comparison with $\sqrt{s} = 0.9$ TeV", *JHEP* **09** (2011) 109, doi:10.1007/JHEP09(2011)109, arXiv:1107.0330.

[77] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[78] J. Allison et al., "Geant4 developments and applications", *IEEE Trans. Nucl. Sci.* **53** (2006) 270, doi:10.1109/TNS.2006.869826.

List of Figures

List of Tables

Acronyms

ALICE A Large Ion Collider Experiment. 3

APD Avalanche photo-diodes. 12

ATLAS A Toroidal LHC ApparatuS. 3

AVF Adaptive Vertex Fitter. 23

BDT Boosted Decision Tree. 34, 38, 53

BSM Beyond the Standard Model. 3

CERN European Organization for Nuclear Research. 2–4, 20, 51

CMS Compact Muon Solenoid. 3, 6–10, 13, 15–18, 20, 22–26, 29–34, 39, 41, 42, 51, 53

CPU Central Processing Unit. 19

CSC Cathode Strip Chamber. 17, 31

CTF Combinatorial Track Finder. 22, 27

DA Deterministic Annealing. 23

DAQ Data Acquisition. 18, 51

DT Drift Tube. 17, 31

EB ECAL Barrel. 12

ECAL Electromagnetic Calorimeter. 11–15, 25–28, 39

EE ECAL Endcap. 12

EM Electromagnetic. 34, 37

FCT Fundação para a Ciência e a Tecnologia. vi

GCT Global Calorimeter Trigger. 21

GSF Gaussian Sum Filter. 27

HB HCAL Barrel. 13, 14, 51

HCAL Hadronic Calorimeter. 13–15, 25, 26, 39, 51

HE HCAL Endcap. 14, 51

HF HCAL Forward. 14, 15, 51

HLT High Level Trigger. 19, 20

HO HCAL Outer. 14, 51

HPS Hadron Plus Stips. 36, 37

ID Identification. 29, 34

L1T Level 1 Trigger. 13, 18–21, 51

LEP Large Electron Positron collider. 2

LHC Large Hadron Collider. 2–11, 18, 20, 23, 24, 41, 42, 51

LHCb Large Hadron Collider beauty. 3

LINAC2 Linear Particle Accelerator 2. 3

LO Leading Order. 41

LS1 Long Shutdown 1. 21

MC Monte Carlo. 22, 40–42

ME Matrix Element. 42

MET Missing Transverse Energy. 36, 39, 40

MPF Missing Transverse Energy Projection Fraction. 35

NLO Next to Leading Order. 42

PF Particle Flow. 25, 26, 32–34, 36–40, 53

POG Particle Object Group. 26, 29, 30, 32, 34, 39, 53

PS Proton Synchrotron. 3

PSB Proton Synchrotron Booster. 3

PU Pile-Up. 7, 20, 26, 29, 31, 34, 35, 38

PV Primary Vertex. 23, 24, 34, 53

QCD Quantum Chromodynamics. 26, 37

RF Radio Frequency. 4

RPC Resistive Plate Chamber. 17

SM Standard Model. 4, 15

SPS Super Proton Synchrotron. 3, 6

TEC Tracker Endcaps. 11, 51

TIB Tracker Inner Barrel. 11, 51

TID Tracker Inner Disks. 11, 51

TOB Tracker Outer Barrel. 11, 51

UE Underlying Event. 35, 41

VBF Vector Boson Fusion. 22, 33, 39, 40

VPT Vacuum Photo-Triodes. 12