

CHAPTER 3

The LHC and the CMS experiment

The CERN accelerator complex is a sequence of machines that produce and accelerate collections of 10^7 protons, called “bunches”, to nearly light speed. In the Large Hadron Collider (LHC) the bunches collide at specific interaction points (IP), where the four main experiments are located: ALICE, ATLAS, CMS, and LHCb. This section will start with a discussion of the LHC machine and then describe the CMS detector.

3.1 The Large Hadron Collider

3.1.1 The history of the LHC

The story of the LHC begins in 1977, when the CERN director general Sir John Adams suggested that the tunnel of the future Large Electron-Positron Collider (LEP) could be reused to accommodate a future hadron collider of more than 3 TeV center-of-mass (COM) energies [56]. At the 1984 ECFA-CERN workshop on a “Large Hadron Collider in the LEP Tunnel” [57], the physics goals of the LHC were stated: confirming BEH mechanism, searching for the Higgs Boson, and exploring of the origin of the masses of W and Z bosons. The parameters of the proposed LHC were very ambitious: the center-of-mass collision energy of 10 to 20 TeV, and a target instantaneous

luminosity of $10^{33-34} \frac{1}{\text{cm}^2 \text{s}}$.

The Large Hadron Collider (LHC) is the most powerful (in terms of the COM energies) particle accelerator that has ever been built. It is located at the border of France and Switzerland at a depth from 50 to 175 m underground. The LHC ring is 26.7 km in circumference and is the final stage in a sequence of accelerators. In the following section, there is a discussion of the whole sequence of accelerators.

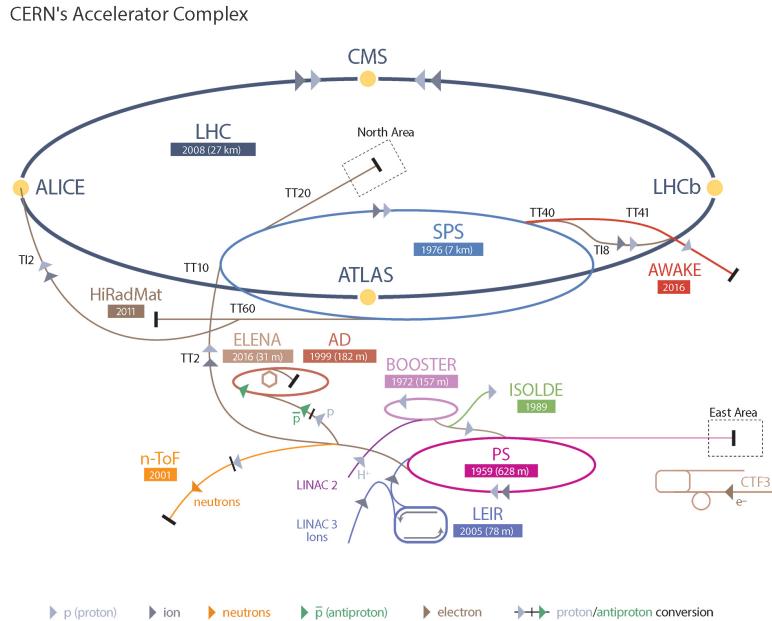


Figure 3.1: Schematic layout of the LHC.

3.1.2 The layout of the LHC

It is a complex process to start proton-proton collisions in the LHC at 13 TeV and, therefore, the process consists of several stages (see Fig. 3.1). The process begins with a bottle of hydrogen. The hydrogen atoms from the bottle are fed into the source chamber of the Linear Accelerator (Linac). In the chamber, the hydrogen is heated up to the plasma state until the electrons are stripped off of the hydrogen atoms. Electrons are then removed and the remaining protons are directed to the

first acceleration stage which increases the energy of the protons to 50 MeV. After the Linac, the beam of protons is injected into the Proton Synchrotron Booster (PSB). The PSB contains four rings, which accelerate protons to 1.4 GeV. The third stage is the Proton Synchrotron (PS), which further splits the incoming beam into 72 bunches separated by 7.5 m. The energy of the protons is increased to 25 GeV. After that, the protons are sent to the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV and split into more bunches. The SPS then fills the LHC ring with two beams, each consisting of 2808 bunches of protons with nearly 10^{11} protons in total. It takes SPS about $O(10)$ minutes to fill each LHC ring with bunches. In the LHC, the two beams are circulating in opposite directions in two separate beam pipes. During standard data collecting (“data taking”), the beams circulate for $O(10)$ hours. This number has been found experimentally and leads to the highest possible amount of data per year.

3.1.3 LHC infrastructure

The equipment of the LHC tunnel serves several purposes, the main objective being to keep the colliding beams on a circular orbit, which requires a complex synchronization of bending dipole magnets, cooling systems, accelerating radio frequency cavities, and vacuum insulation systems.

3.1.3.1 Magnets

Most of the LHC’s circumference is covered by 1232 superconducting magnets placed evenly around the tunnel to approximate a circular orbit. These are dipole magnets (see Fig. 3.2) that bend the beam and keep it on the circular orbit, hence why they are commonly called “Main Bends” (MB). The proven technology for magnets has

existed since Tevatron and relied on NbTi superconductors. This technology also satisfied the LHC cost and performance requirements, thus, the decision was made to reuse the same choice of the alloy for the LHC superconducting dipole magnets that steer the proton beams. The dipoles need to produce the magnetic field of 8.3T. Each dipole is 16.5 *m* long (with ancillaries) and 570 *mm* in diameter and is placed inside of the dipole cryostat, called the “Helium bath”.

The cryostat is a long cylindrical tube 914 *mm* in diameter, made of low-carbon steel, and is where the dipole mass is cooled down to 1.9 *K*. Even though the inner structure of such cryostats is very complex and includes both beam pipes, two sets of coils for two beam pipes, vacuum pipes etc., this compound object is normally called a dipole magnet. The name “dipole” is reserved for MBs as each beam pipe in the magnet consists of two “poles” that provide a vertical magnetic field similarly to a simple dipole system of magnets.

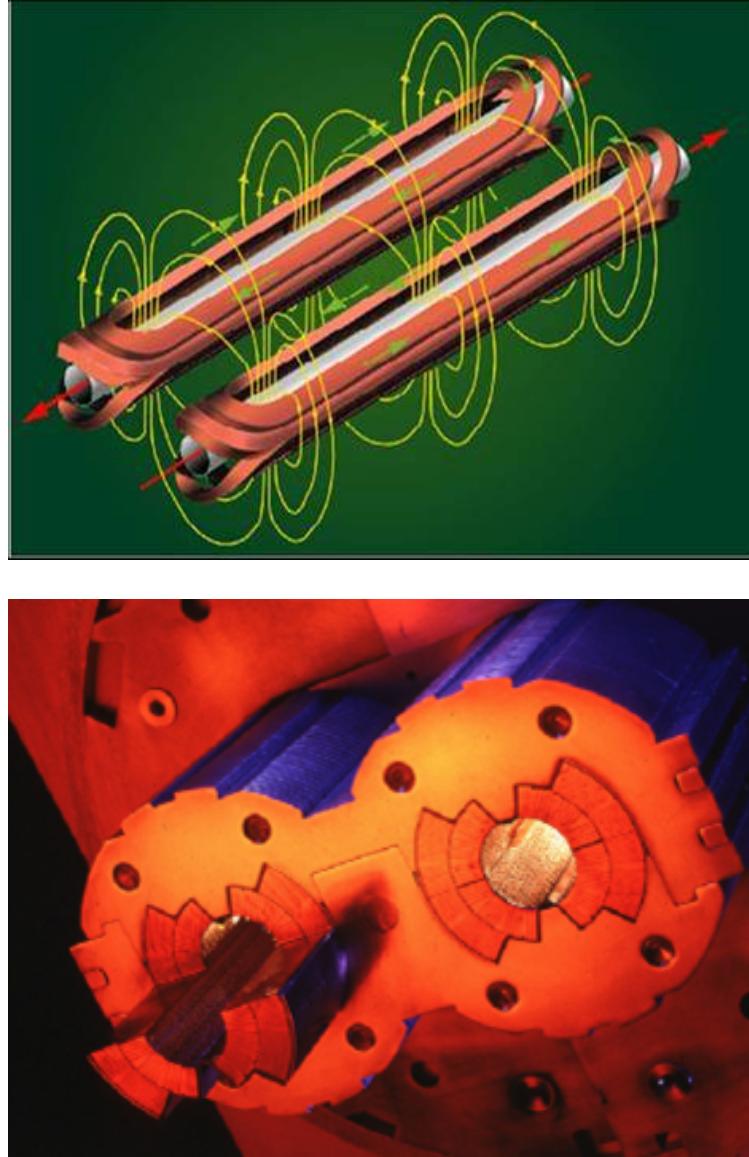


Figure 3.2: LHC dipole magnets. Top: two dipole coils and magnetic field lines. Bottom: two beam pipes with the coils inside of the dipole magnet. Each dipole magnet contains two magnetic configurations to steer the two proton beams travelling in opposite directions.

A dipole magnet must be curved to help a chain of dipoles complete 360 degrees. The curvature is 5.1 *mrad* per dipole, which is equivalent to a sagitta (explained later in this chapter) of about 9 mm, corresponding to a radius of curvature of 2812.36 m.

Other important magnets are quadrupoles. They are used to ensure proper beam

dynamics. In total, 392 quadrupole magnets ranging from 5 to 7 metres in length are used to squeeze the beam in transverse directions (with respect to the z axis) and to keep it narrow during the run duration. Additional special quadrupole magnets (SQM) are installed right before the IPs to focus the beams even more. These increase the density of protons in the beam and guarantees the maximum luminosity (explained in [3.1.5](#)). In addition, SQMs help to decrease the chance of the parasitic collisions when bunches from the same beam or bunches outside of the IP center interact (see Fig. [3.3](#)). To further correct the beam path or orbit, about 5000 higher order correcting magnets are used, which are evenly spaced around the circular trajectory of the LHC.

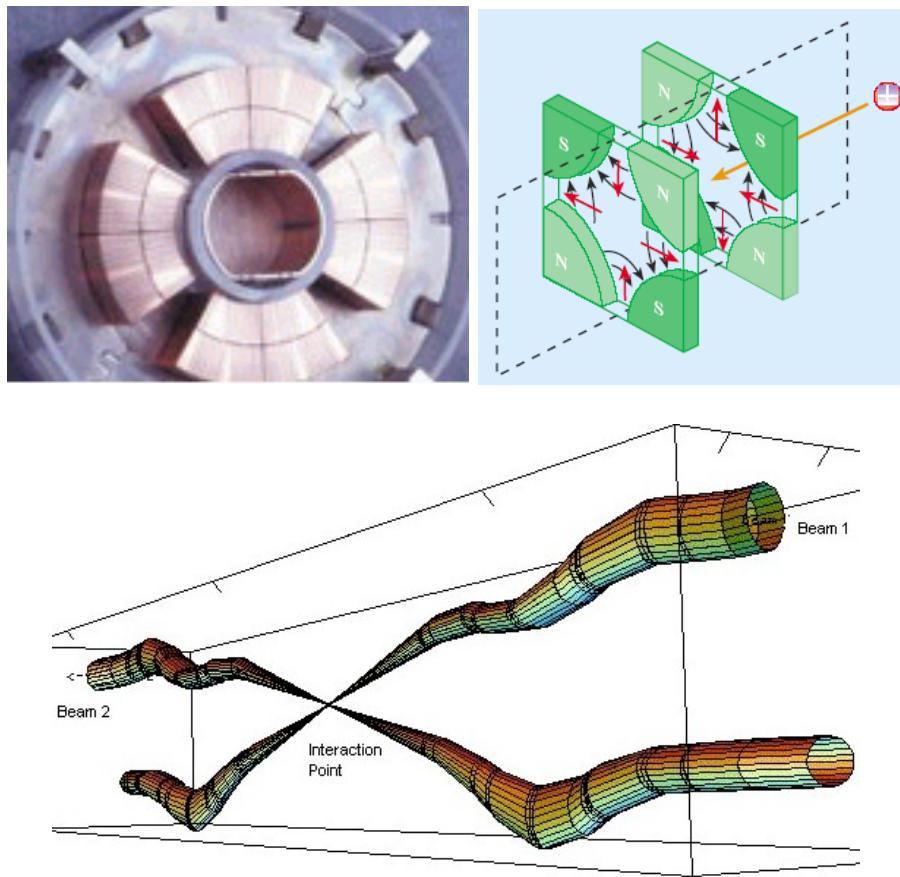


Figure 3.3: LHC quadrupoles. Top left: the coil of the quadrupole magnet. Top right: schematic view of the magnetic fields in the quadrupole. Bottom: two beams and the IP.

To power the LHC, 1612 electrical circuits are used. These circuits are mostly needed to power the dipole and quadrupole magnets, which is done in eight evenly-spaced locations of the LHC. A total of 3286 current leads are needed to connect all the circuits and power cables. More than a thousand of the leads operate between 600 A and 13 kA (see Fig. 3.4). The other leads operate in the range 60 to 120 A.



Figure 3.4: 13 kA high-temperature superconducting current lead.

3.1.3.2 Cooling System

To ensure that the dipoles are in the superconducting state, they have to be cooled to 1.9 K using superfluid helium-4. The cooling (cryogenic) system is needed to keep the superconducting LHC magnets at the appropriate temperature. The choice of cooling gas depends on the magnet type and location. This dictates the required range of temperatures, which differs from system to system by 75 K. The cryogenic system uses a layered design with the temperature becoming progressively colder moving

from outside the dipoles to closer towards the beam pipe.

The “coldest” part of the cryogenic system is designed for the inner part of the dipoles. This system (see Fig. 3.5) must cool down 37 Mkg of the LHC magnets within 15 days to the required temperatures, which is done through a system of pipes that transports and directs the flow of the superfluid helium. The cryogenic system must also be able to deal with the fast increases of the pressure flow and flow surges, as it is crucial for the LHC operation to keep dipoles constantly cooled and at the superconducting state.

As the cost to cool the LHC equipment to 1.8-1.9 K temperatures is high, several temperature levels are employed (see Fig. 3.5):

- 50 to 75 K for the thermal shielding used in the dipoles,
- 20 to 300 K for upper (“warm”) sections of the high-temperature superconducting current leads,
- 4.6 to 20 K for lower temperature interception,
- 4.5 K for radio frequency cavities and lower (“cold”) sections of the high-temperature superconducting current leads,
- 4 K for the transportation system that directs the 1.8 K helium to dipoles,
- 1.9 K for helium in the superfluid state to cool magnet masses.

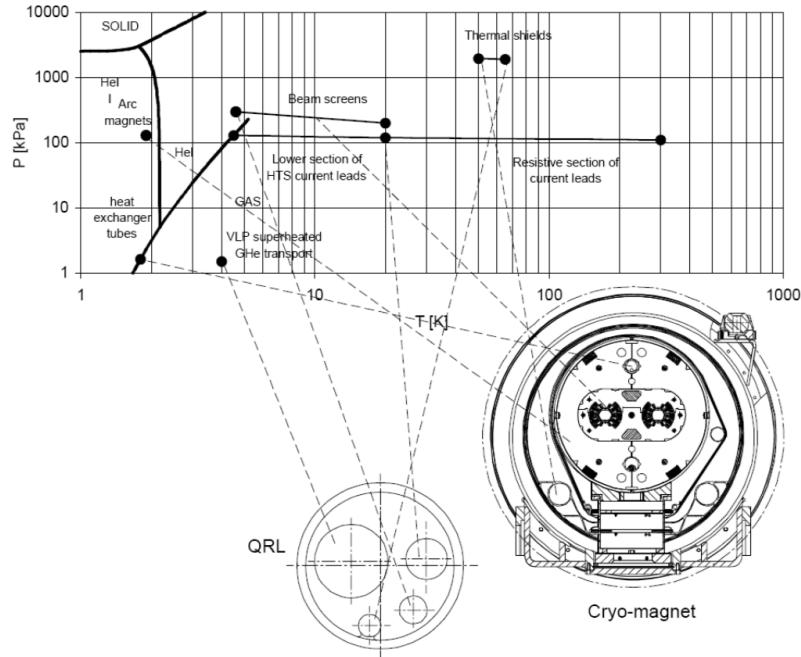


Figure 3.5: LHC cryogenic states and the temperature scale.

3.1.3.3 Radio Frequency Cavities

Proton bunches need to be ramped to 6.5 TeV energies. To achieve this 13 TeV COM energy, eight superconducting radio-frequency cavities (RFC) are used per beam. They are located in front of the IPs of four experiments. Electromagnetic waves of 400 MHz with a peak field strength of 5.5 MV/m adjust the speed of protons in bunches. Each RFC (see Fig. 3.6) increases the energy of protons by 60 keV per revolution and it takes $O(20)$ minutes to reach 6.5 TeV beam energy. The RFC frequencies are gradually increased by 1 kHz to match the acceleration of the protons in the bunch as they gain more energy. When the ramp is completed, the RFCs are used to compensate for small energy losses due to the synchrotron radiation (7 keV per revolution).

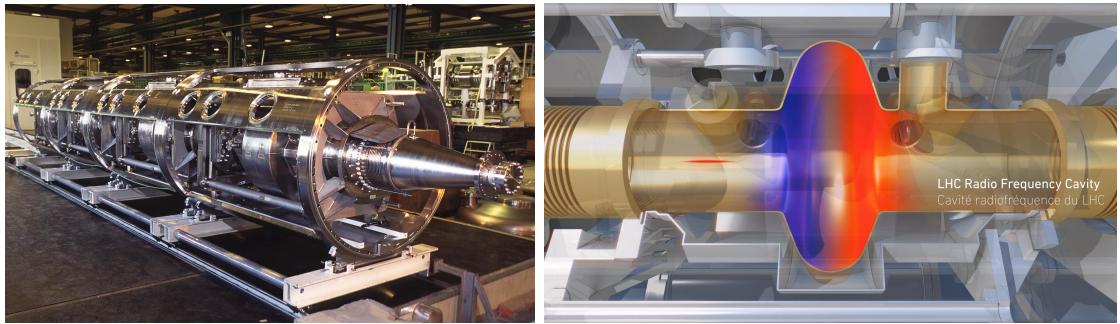


Figure 3.6: LHC RF cavities. Left: a cryomodule with four RF cavities. Right: a schematic drawing of a single RF cavity. The colour field is used to denote positive (red) and negative (blue) polarities. A narrow beam traversing the cavity is coming from the left and is shown in red.

3.1.3.4 Vacuum System

The work of the LHC depends on three vacuum systems [58]. Without them, the dipoles would not achieve the superconducting state, the beams would not be able to circulate, and no stable collisions would occur. With a total of 104 kilometres of vacuum pipes, the LHC owns the largest vacuum system in the world. The main types of vacuum systems are:

- insulation vacuum for cryomagnets,
- insulation vacuum for the helium distribution line,
- beam vacuum.

The insulation vacuum is needed to ensure the operations at both the low temperatures of the magnets and the room temperatures in the tunnel. The insulation vacuum of 10^{-6} mbar is used for a total of 15000 cubic metres. In building this vacuum system, the LHC used 250,000 welded joints and 18,000 vacuum seals.

The vacuum for the helium distribution lines is needed to protect the flow of the helium-4 from the heat. This helium flow is used to cool down the dipole mass.

Cryogenic distribution lines (QRL) of 3.3 km each are connected to eight cryogenic plants that pump the helium-4 into the LHC. The vacuum in these systems is at $10^{-7} - 10^{-10}$ mbar level.

For the beam pipes, the LHC uses ultra-high vacuum of 10^{-10} mbar at a cryogenic temperature of 5 K. The vacuum is getting progressively closer to 10^{-11} mbar near the IPs, because these locations are where collisions take place and any additional gas is highly undesirable. This vacuum is the emptiest space in the Solar System. This ultra-high vacuum is needed to reduce the beam degradation due to the beam-gas interactions in the pipe and parasitic collisions of the proton bunches with the collimators near the IPs.

The vacuum systems are affected by the heat produced from the synchrotron radiation that is emitted by the proton beams when they are bent. To reduce the amount of this heat and to narrow down the beam size in the transverse direction when the beam widens, the LHC uses “beam screens”, which operate at between 5 and 20 K.

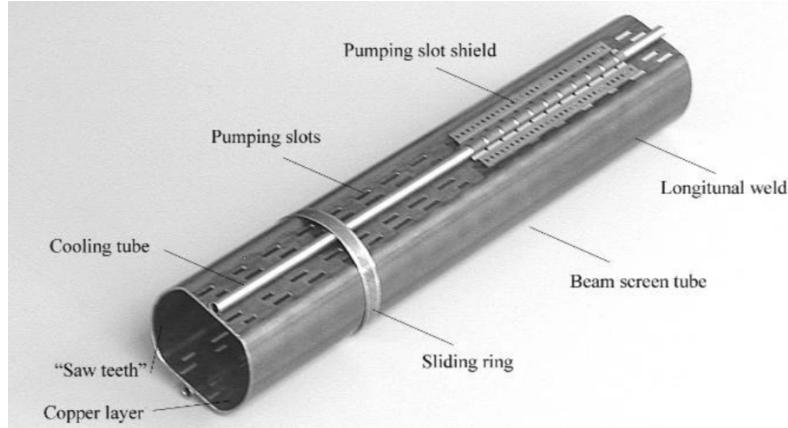


Figure 3.7: Beam screen.

The beam screens are necessary to reduce the number of protons scattering on the residual gas of the beam pipes, which could lead to a magnet quench and even

interrupt the machine operation.

The table below summarises the main heat sources that degrade the vacuum quality in the beam pipe, where the vacuum must exist at 1.9 K:

- synchrotron radiation (0.2 W/m per beam),
- energy loss by nuclear scattering (30 mW/m per beam),
- image currents (0.2 W/m per beam),
- electron cloud related effects (vary).

Having discussed the LHC collider, the following section explores the Compact Muon Solenoid (CMS) detector, which was used to collect the data analysed in this thesis.

3.1.4 LHC operations

The first LHC budget plan was finalised in 1996, with the final cost being approved just a few years later. The first proton beam entered the LHC ring in 2008. However, an incident delayed the original LHC plans. It was caused by the mechanical damage of the tunnel equipment due to the release of the helium. Thus, the real data taking period (called LHC Run 1) began only in 2010, and lasted for two years, with 7-8 TeV COM energies used. The recorded dataset contained enough Higgs bosons to claim a discovery of this rarely produced particle. After this achievement, the LHC was closed in 2012 for the first long shutdown (LS1). During this time, necessary upgrades of the main detectors and the LHC were performed. This was an unavoidable and essential step to prepare the LHC for the more challenging environment of COM energies increased to 13 TeV.

If denote the area of 10^{-28} m^2 as barn (b), then the femtobarn (fb) equals to 10^{-43} m^2 . When HEP physicists talk about inverse femtobarns, they mean collisions per femtobarn of beam cross-sectional area. Thus, the inverse femtobarn unit can be thought of as the effectiveness of the particle accelerator. In terms of these new units, the LHC can theoretically produce $80 - 120 \text{ fb}^{-1}$ (inverse femtobarns) of data a year. In practice, these numbers were lower because the LHC operated at the revolution frequency below the nominal, used fewer proton bunches in the beam, etc. These resulted in lower than the expected instantaneous luminosity, an important term in collider physics that will be explained in the next section.

The LHC Run 2 started in 2015 and the CMS collected 4.2 fb^{-1} of data that year. Over the course of the 2016 data collection, an integrated luminosity of 35.9 fb^{-1} was recorded. This luminosity is the amount of data that was collected by the CMS detector and later approved by the CMS physics coordination for the use in the physics analyses. Note that the CMS collects as much data as possible, but not all data have good quality for the offline physics analyses, so the amount of the approved data is always less than the collected one. The data set of proton-proton collisions collected in 2016 at 13 TeV COM energy will be used in this thesis to analyse double Higgs boson decays. Together with the 2017 and 2018 data taking, almost 150 fb^{-1} have been delivered and recorded by the CMS detector during the whole Run 2 period of four years (2015-2018).

At the moment of writing this thesis, the LHC has entered LS2. Data taking will resume in 2020 and proton-proton collisions will continue for three years with the integrated luminosity expected to be equal to nearly 300 fb^{-1} . This will conclude the LHC Phase 1 program.

The new upgraded LHC, the High-Luminosity LHC (LHC) or the Phase 2, will start operations in 2026 and run until 2035. The COM energy will be increased to 14

TeV (may happen during Run 3) and physicists expect to record an unprecedented dataset of 3000 fb^{-1} .

3.1.5 Luminosity

The instantaneous luminosity \mathcal{L} is the parameter which relates the cross section σ of the process to the number of events N_{events} produced per unit time during the interaction: $N_{events}/dt = \mathcal{L}\sigma$. Luminosity is the parameter controlled by the machine and can be written as:

$$\mathcal{L} = \frac{N^2 n_b f_{rev}}{4\pi\sigma_x\sigma_y}$$

where N_b is the number of particles in the colliding bunch, n_b is the number of colliding bunches in the beam, f_{rev} is the revolution frequency of the beam, a σ_x and σ_y are the standard deviations of the beam density profile (BDP) in the transverse plane, where it is assumed that the BDP of both beams can be described by a Gaussian distribution.

To maximise the amount of collected data, the luminosity parameter should be as high as possible. It is worth noting that the luminosity is not constant and decays with time due to the degradation of the initial circulating beams. The theoretical decay time (the time to reach $1/e$ level) is approximately 29 h. In practice, taking into account the decrease of protons in the bunch due to collisions, contributions from the intrabeam scattering, scattering on the residual gas, etc., the real luminosity lifetime is about 15 h.

A useful variation of the luminosity parameter is the total integrated luminosity. This is the number normally quoted for the dataset collected over the period T:

$$L = \int_0^T \mathcal{L} dt.$$

In collider physics the “beam dump” is a process of burning off exhausted low

luminosity beams by intentionally directing them towards a target made of concrete and steel. The time from the start of the collisions to the beam dump is usually called the “run” or “fill”.

The amount of data delivered by the LHC can be calculated for a single run period of $O(10)$ h. Performing the integration, this is obtained:

$$L = \mathcal{L}_0 \tau_{\mathcal{L}} \left[1 - e^{-\frac{\tau_{run}}{\tau_{\mathcal{L}}}} \right],$$

where \mathcal{L}_0 is the initial peak instantaneous luminosity at the start of the run, τ_{run} is the total duration of a run, and $\tau_{\mathcal{L}}$ is the luminosity lifetime. The optimum run time is 12 hours. Between the runs, the LHC center needs to dump the old beams, fill the rings with the new beams, and increase (“ramp”) the energy of new beams to 13 TeV. After that a new run can be started. This restarting process normally takes two to six hours.

3.2 The CMS experiment

The CMS detector (later the CMS for brevity) is a multi-purpose particle detector built to study a variety of complex particle interactions produced by the LHC. The CMS is located in the underground cavern at the Point 5, which is one of the four main IPs of the LHC. The CMS detector, with additional computing infrastructure, is able to detect the produced particles, to measure their main physics parameters, and to send the related data to computing data centers for persistent storage.

The CMS detector has a cylindrical shape and consists of a central cylinder (the “barrel”) and two forward (the “endcaps”) sections (see Fig. 3.8). The CMS detector is the heaviest detector ever built with the mass of nearly 12500 tons. The mass is explained by the amount of the used superconducting metal, which serves as the magnet, as well as by all the iron on the outside used for the magnetic flux return

and shielding. The CMS is 21.6 m long and 14.6 m high. The CMS has an onion-like structure of concentric layers made up of detectors around the IP. Additionally, the outer part has a large superconducting solenoid to produce a homogeneous magnetic field of 3.8 T inside the detector.

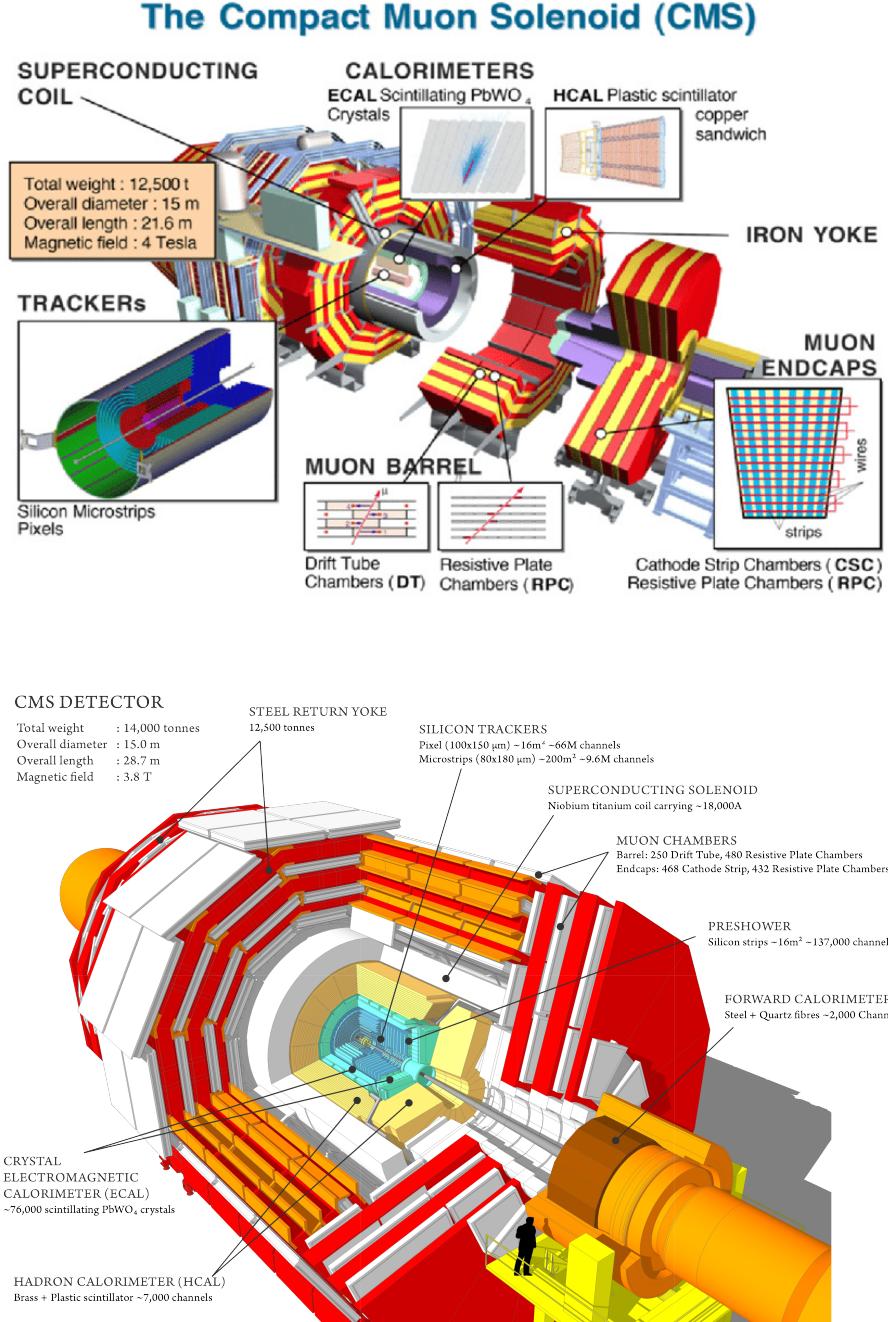


Figure 3.8: CMS experiment with the main sub-detectors.

All sub-detectors can be categorized into trackers and calorimeters [59]. As the particle passes through the material of the tracker, it leaves a path of the emerging particle, called a “track”. Trackers focus on the direction and the track curvature of

the charged particles. Tracking information allows the determination of the particle's momentum.

There are two trackers in the CMS detector: an inner tracking system that encloses the IP and an outer tracking system that is located outside of the solenoid magnet. The first system contains the pixel and the strip trackers. The second tracking system, embedded within the steel yoke of the magnet, is dedicated to the detection of muons, and is usually called a muon tracker or a muon system.

The magnet yoke is made of five barrel wheels. Such an arrangement saves the CMS some space and also is used for the magnetic flux return. Additionally, it serves as a support for the embedded muon system, which is located outside of the ECAL and HCAL systems (described later). Prompt muons from the heavy particle decays are energetic enough to traverse the ECAL and leave the detector because they are minimum ionizing particles. The muon system-magnet yoke structure provides a return field of the magnet of about 2 T and is used to measure the momentum of muons. The complex magnetic field causes the muons trajectories to be bent in opposite directions in the inner tracker in contrast to the outer tracker. This important feature of the CMS detector is depicted in the CMS logo (see Fig. 3.9).

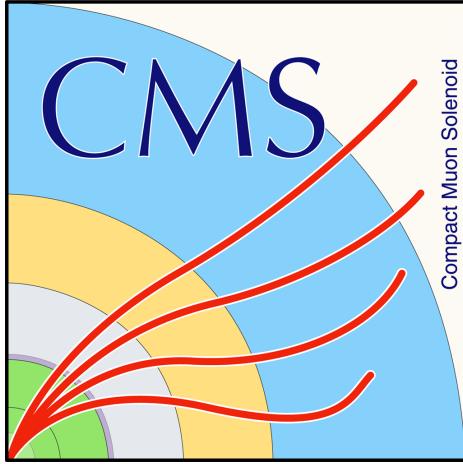


Figure 3.9: The logo of the CMS experiment that is showing curved trajectories of the emerging muons.

The CMS has two calorimeters: the electromagnetic and the hadronic. They both rely on high density materials either to sample or to contain almost all the energy of the incoming particles with their secondary interaction products. However, these two systems focus on two different sets of particles. As will be discussed later, the electromagnetic calorimeter (ECAL) is dedicated to measuring the energy of photons and electrons, while the hadronic calorimeter is targeting the measurement of the energy of hadrons.

The rate of the incoming data at the LHC is 40 MHz, which is related to the 25 ns bunch spacing. This corresponds to almost 70 TB produced every second. It is impossible to store that much data, and, most importantly, most of the information in this data is not pertinent for future physics analyses (the Physics program of the CMS is well defined and specific). To reduce the data rate, the CMS uses a highly efficient system of triggers. The first one, the Level-1 (L1) trigger, reduces the readout rate from 40 MHz to 100 kHz. The subsequent High-Level Trigger (HLT) further decreases the rate to 1 kHz. With the help of the trigger system, the original 70 TB per second rate is transformed into manageable 1 GB per second that is stored for offline analysis.

use.

3.2.1 The CMS coordinate system

The CMS uses a right-handed Cartesian coordinate system to define the axes of the colliding beams (see Fig. 3.10). The center is located at the IP and the x axis points to the center of the LHC ring. The y axis points upwards, and the z axis points along the proton beam direction. This is defined for one of the proton beam directions, and is opposite for the other one. As the CMS detector has a cylindrical shape, the polar system is used in the x-y plane: a standard set of the azimuthal angle φ and the radial coordinate r . A polar angle θ is defined in the r-z plane and an angular variable η (called pseudorapidity), which is widely used in this thesis, is defined as $\eta = \ln \tan(\theta/2) = \ln(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z})$. Another useful quantity in the collider physics - the rapidity - is given by $y = 1/2 \ln(\frac{E + p_z}{E - p_z})$. Rapidity is a function of the energy E and longitudinal momentum p_z of the particle (the projection of \vec{p} on the z axis). Note that η converges to y when the mass is negligible and the particle travels with the speed close to the speed of light. Most angular variables that are used currently in modern high-energy physics (HEP) at the LHC are defined in terms of η and φ . For example, a relative distance ΔR in $\eta - \phi$ plane between two particles is given by: $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}$, with $\Delta\eta$ and $\Delta\varphi$ being the absolute values of the relative differences in η 's and φ 's of two particles.

Another useful quantity is the projection of the momentum of a particle on the transverse plane, which is called “transverse momentum” p_T . This variation of the momentum is independent of the z axis, and, thus, of the Lorentz boost. Similarly, the transverse energy of a particle is defined as $E_T = \sqrt{m^2 + p_T^2}$.

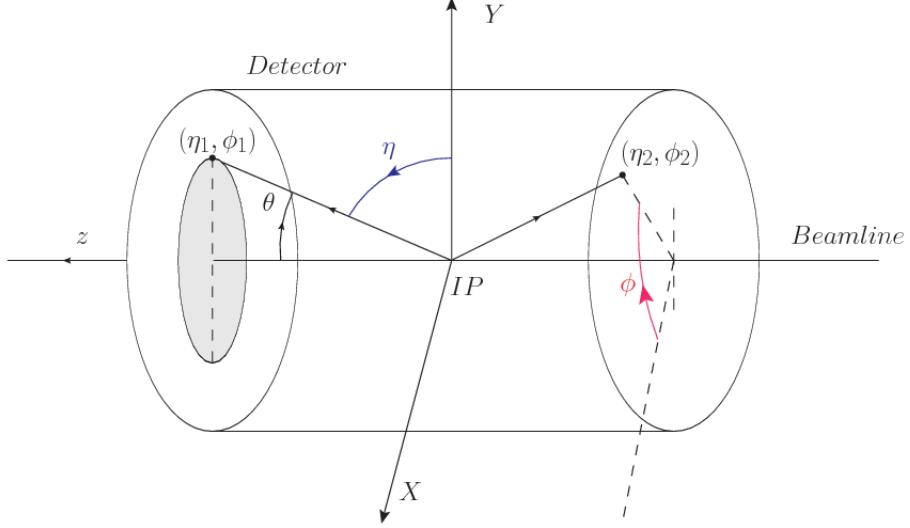


Figure 3.10: Coordinate system of the CMS detector [35]. Two particles (1 and 2) are shown with the corresponding angular variables ($\Delta\eta_1, \Delta\varphi_1$) for the first and ($\Delta\eta_2, \Delta\varphi_2$) for the second particle respectively.

3.2.2 The Inner Tracker

The inner tracker [60] (see Fig. [3.11]) is the closest subdetector to the IP. Using the tracker, the experiment measures the trajectories of charged particles and reconstructs decay vertices. Because this system is constantly under the radiation coming from the interactions with the particle flux of nearly 100 MHz/cm at $r = 4$ cm, the design of the tracker is focused on two main requirements: high granularity for precise determination of the vertices and tracks, and robustness against the radiation-hard environment with the operational time of at least ten years. As a solution to both challenges, the CMS relies on the silicon technology that provides the tracker with the large surface of thin but highly granular active detectors. The tracking system has a diameter of 2.4 m and a length of 5.4 m covering the detector space of $|\eta| < 2.5$.

The innermost part of the tracker - the pixel detector (“pixel”)- consists of three layers in the barrel at the radii of 4.4 cm, 7.3 cm, and 10.2 cm respectively. The

pixel also has two active disks in forward regions. They are positioned 34.5 and 46.6 cm away from the IP. The pixel is made of 1440 modules which contain 66 million pixel cells. Each cell is 100 by 150 μm with 285 μm thickness, which allows the determination of “hit” positions (the passage of the particle through the pixel cells) in two directions: $z\text{-}\varphi$ in the barrel and $r\text{-}\varphi$ in the endcaps.

The spatial resolution of each pixel is about 10 μm in the $r\text{-}\varphi$ plane and 20 μm along the z direction. The spatial information that comes from the tracker is used to determine the main interaction point of the hard scattering (the primary vertex) and also additional interaction vertices (the pileup). The tracker also helps to reconstruct the displaced vertices (the secondary vertices) of the particles that decay relatively quickly, e.g., b-jets, which will be discussed later in this chapter.

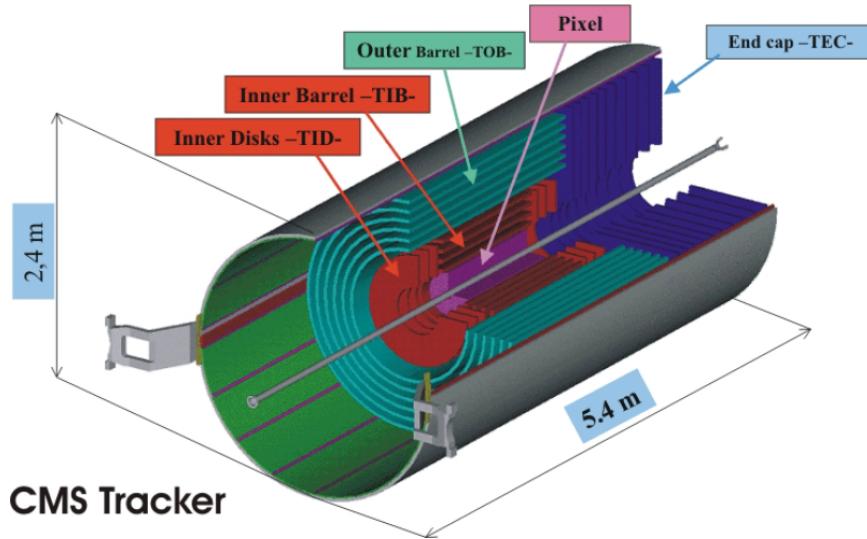


Figure 3.11: The inner tracker. The pixel and strip detectors are shown.

The outer part of the inner tracker is the strip tracker. It contains several subsystems and is made of almost 9.3 million strips arranged in different configurations in 15148 modules. The first subsystem is the tracker inner barrel (TIB), which consists of the four barrel layers of strip modules. The second subsystem is the tracker inner

disks (TIDs), which is made of three disks of strip modules. At a radius of about 60 cm, the tracker outer barrel (TOB) starts. TOB is made of six layers of strips. Finally, to cover high η regions, the tracker endcaps (TECs) are used, which are made of two sets of nine disks of strips.

Each strip is O(20) cm long. Its thickness varies from 320 μm for TIB and TID, to 320 μm - 500 μm for TOB and TEC, respectively. Also the width changes from 80 μm - 141 μm for TIB and TID, to 97 μm - 184 μm for TOB and TEC, correspondingly. The resolution on the single point in the radial direction is 20 - 50 μm , and in the z direction it varies from 200 to 500 μm , depending on the value of r .

All subsystems of the inner tracker have to be cooled down to about -20°C. This requirement is needed to minimize the damage to the tracker caused by the radiation from the collisions and to reduce overheating of the electronics. The material of the inner tracker has 0.4 to 1.8 radiation lengths (X_0), which corresponds to 0.1 to 0.5 nuclear interaction lengths (λ_i). Numbers vary with the η . The inner tracker is surrounded by the ECAL described in the following subsection.

3.2.3 The ECAL

The inner tracker and the ECAL provide the detector with complementary measurements. The tracker focuses on the direction and the momentum of the particle and identifies only charged particles. The ECAL [61] (see Fig. 3.12), on the other hand, determines the energy of the particles and detects all particles that interact electromagnetically, including photons and neutral pions. However, the ECAL is primarily designed to make precise measurements of the energy of electrons and photons by selecting the proper type of crystals and their length to contain enough X_0 's to stop incoming electrons and photons.

The ECAL is a highly granular detector that relies on the lead tungstate crystal (PbWO_4) technology. Electrons and photon passing through the crystal interact with its material and their energy is converted into the produced electromagnetic shower. The scintillation light, produced by PbWO_4 scintillators, is further read out by the electronics, more details in [62]. The PbWO_4 crystals have a high density ($8.28\text{g}/\text{cm}^3$), a small radiation length ($X_0 = 0.89\text{ cm}$), a short Moliere radius ($R = 2.2\text{ cm}$), and a fast response (80% of its scintillation light is produced within 25 ns). These characteristics make PbWO_4 crystals ideal candidates for the ECAL, because they guarantee an excellent containment of the electromagnetic shower within the crystals.

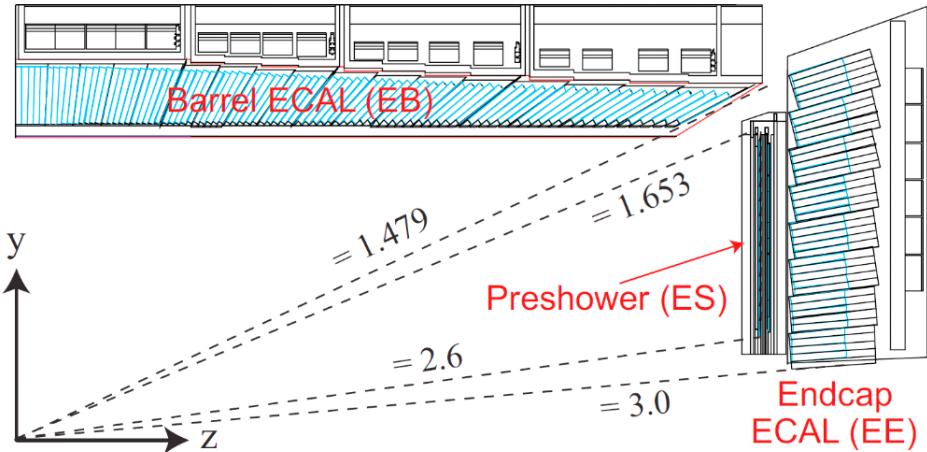


Figure 3.12: The ECAL and the Preshower detectors.

The ECAL has a barrel part (EB), covering the $|\eta| < 1.479$, and two endcaps (EE) covering $1.479 < |\eta| < 3.0$. In the barrel, the ECAL is made of 61 200 crystals. Each crystal is 22 by 22 mm with a length of 23 cm. The endcaps of the ECAL have 7324 crystals. There, each crystal is 28.62 by 28.62 mm with a length of 22 cm. The crystals' layout follows a quasi-geometric projection where the axes of the crystals are slightly tilted to ensure that particle trajectories are never aligned with intercrystal

cracks. This layout is optimized for the best particle shower containment with respect to the position of the interaction point.

The resolution of the ECAL is a function of the energy of the incident particle E and can be decomposed into three terms. The first term is a stochastic term that is inversely proportional to the square root of the number n of scintillation photons produced in the interaction. In the main formula n is replaced by E, since n is proportional to E. The second term is a “noise” term that describes the noise in the detector. The dependence of the resolution on the energy component in the noise term is inverse. The third term is related to detector imperfections and is represented by a constant C. The final dependence of the ECAL energy resolution σ on the particle energy E is given by:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (3.1)$$

From dedicated calibration studies, the parameters in the formula above are found to be equal to $S = 2.8\%$, $N = 12\%$, and $C = 0.3\%$. As a “standard” procedure, the CMS often optimizes the performance of the subdetectors for 45 GeV electrons, since they correspond to a classical Drell-Yan decay of Z boson to two electrons. In this case, a typical energy resolution for 45 GeV electrons is about 2% in EB and 2-5% for EE. The constant terms dominate the resolution near the Z peak (91 GeV). The ECAL is operated at a temperature of 18°C and the “active thickness” of the ECAL material corresponds to 25 X_0 .

An additional subdetector, called the Preshower, is installed right in front of the EE and covers $1.653 < |\eta| < 2.6$. The Preshower is designed to improve the discrimination of photons from diphoton decays of neutral pions $\pi^0 \rightarrow \gamma\gamma$. This is a sampling calorimeter in which the material that produces the particle shower is distinct from

the material that measures the deposited energy. Typically, the two materials alternate. The Preshower has two lead layers which launch the electromagnetic showers. This samples the energy of the particles traversing the Preshower material. After these layers, 2 mm-wide silicon strips are placed, and these measure the deposited energy and transverse profile of the shower shape initiated by the lead layers. The thickness of the Preshower material corresponds to $3 X_0$.

3.2.4 The HCAL

Hadrons normally go through the ECAL layers without being stopped since ECAL crystals are optimised to contain mostly electron and photon showers. To absorb these particles, the HCAL [63] (see Fig. 3.13) is placed around the ECAL. The HCAL focuses on particles that hadronize. This is the process of the formation of hadrons out of quarks and gluons. The HCAL detects charged and neutral hadrons such as pions, kaons, protons, and neutrons. Hadrons also produce collimated streams of secondary particles (jets) and these jets are identified by the HCAL. Additionally, the HCAL is used to measure the transverse energy of neutrinos indirectly, through the momentum imbalance technique, which will be discussed later in this chapter.

The HCAL is split into the HCAL barrel (HB) and the HCAL endcap (HE) sections. They cover $|\eta| < 1.3$ and $1.3 < |\eta| < 3.0$ respectively. The HB and HE are sampling calorimeters. They are made of a brass absorber and of active plastic scintillating tiles. The brass plates in HB have thickness of 56.5 mm and in HE the thickness is increased to 79 mm. The absorber material corresponds to $5.82 \lambda_I$ at $\eta = 0$ and almost $10 \lambda_I$ at $0 < |\eta| < 1.3$.

The gaps in the absorber of the HCAL are filled with an active medium of 70000 plastic scintillator tiles. The scintillation light is guided by wavelength shifting fibers

(WLSs) to hybrid photodiodes (HPDs). The scintillator is quite fast with the 68 % of the light being produced within 25 ns.

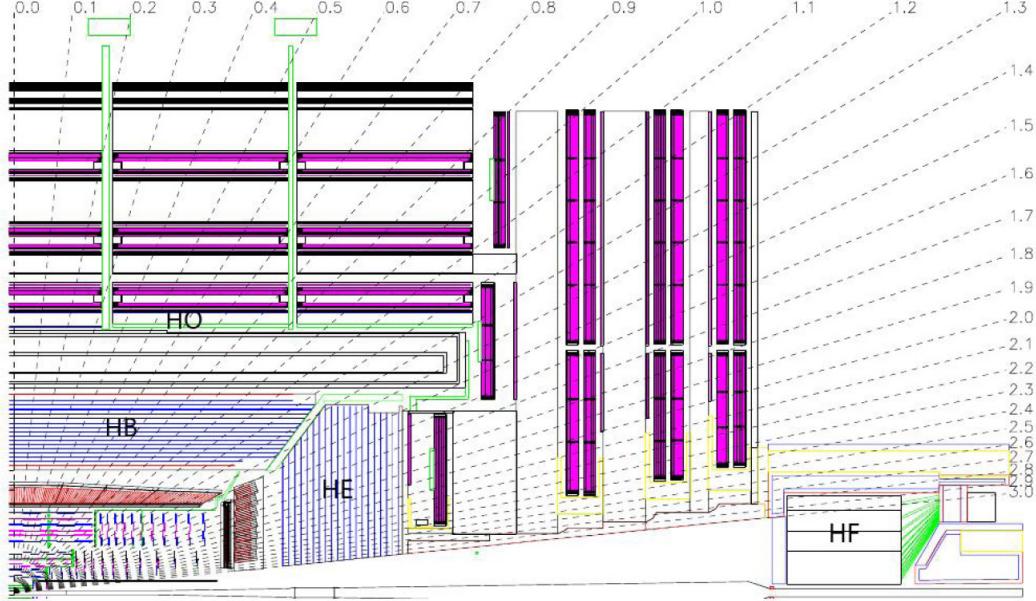


Figure 3.13: The HCAL with the η coverage map.

The CMS also has an outer calorimeter (HO) placed above the HB outside the solenoid. The HO is called a tail catcher system and increases the total calorimeter thickness to $11.8 \lambda_I$ in the barrel, with the magnet coil working as an extra absorption layer. The HO consists of five rings of scintillator tiles. A supplementary 19.5 cm-thick iron plate and a second layer of sensitive material are placed around $\eta = 0$ to enhance the absorber depth in the HO.

In the forward directions, two forward calorimeters (HF) extend the coverage to $|\eta| = 5.2$. The HF is composed of steel absorbers and quartz fibres that produce Cherenkov light when the particle in the material travels faster than the light in that medium. The light is then collected by photomultiplier tubes (PMTs).

Since the HCAL is located between the ECAL and the internal surface of the solenoid, the space allocated for the HCAL has not been enough for the HCAL to

fully absorb the hadronic showers. This imperfect containment of the hadronic shower limits the performance of the HCAL. Comparing with the formula [3.1], from the calibration using the single pions the values are given by: by $S = 115 \%$, $N = 52 \%$, and $C = 5.5 \%$ [64]. This underperformance of the HCAL is perhaps the greatest weakness of the entire detector, which is mitigated by the strengths of the other subdetectors. This motivated the CMS to employ the particle flow algorithm (described in the next chapter), which uses the information from all subdetectors.

3.2.5 The Superconducting Solenoid

The NbTi superconducting solenoid (see Fig. [3.14]) of 6 m in diameter is the core of the CMS experiment. The magnet operates at a temperature of 4.5K. The bulk of the CMS detector weight (90 %) comes from the magnet steel return yoke and structural supports, which together weigh 12500 tons.

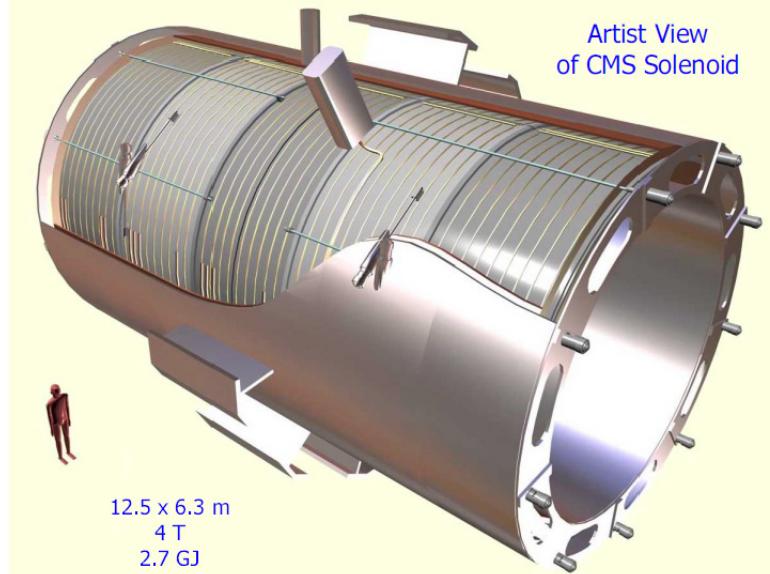


Figure 3.14: The CMS superconducting solenoid. The person on the left is shown to emphasise the size of the magnet.

The solenoid is a central part in the CMS detector design. The idea was to have

a uniform magnetic field capable of bending the trajectories of charged particles as they traverse the detector. When a low energy particle is produced ($O(0.5)$ GeV), it curls back and has a helical path that is fully contained within the detector. However, when a highly energised particle is produced, the trajectory is seen as a “straight” incomplete arc. Both situations lead to an imperfect measurement of momentum.

The primary physics quantities measured by the tracking system are presumed to be Gaussian distributed; but there are exceptions, e.g., momentum of the particle is not Gaussian distributed. However, the sagitta, which is inversely proportional to momentum, is Gaussian distributed. When a particle in the magnetic field passes through the material of the detector, the path deviates from the ideal circular line due to multiple scattering effects. The sagitta term is used to quantify the depth of the circular arc and is equal to the distance from the center of the arc to the center of its base. Since the sagitta is following a Gaussian distribution, it may be approximated by simpler expressions in many calculations of the momentum resolution. Hence, the sagitta is widely used in particle physics.

The magnetic field strength B and the length of the track L are dictated by the design of the detector. Because the momentum resolution is given by $\sigma_p/p^2 \approx \sigma_x/BL^2$ (see [59]) and improves linearly with magnetic field B , the CMS decided to invest much of the detector space and budget in the magnet, since the tracking material is more expensive. For a track of the length of $O(1)$ m in the magnetic field of $O(3)$ T, the sagitta is equal to 1 mm, which can be measured with the precision of $O(10)$ μm .

3.2.6 The Muon Tracker

Many physics analyses in the CMS rely on precise measurements of the muons in the detector. Although muons are detected by the inner tracker, that information cannot

be used by the trigger (which will be discussed in the following subsection). Therefore, the CMS has an outer tracker or muon tracker [65] (see Fig. 3.15) located outside the calorimeters and the solenoid. Typical high energy muons that are produced in collisions at the LHC traverse the detector material with the minimal energy losses since their EM interaction cross section is much smaller than for electrons. To measure the energy of muons, the CMS uses the muon tracker, which relies on various gaseous detector technologies. The muon tracker is inserted into the gaps of the flux-return yoke. Tracks in the muon system are used to reconstruct standalone muons (see section 4) and, in combination with the inner tracker, to reconstruct the global muons, more in section 4.

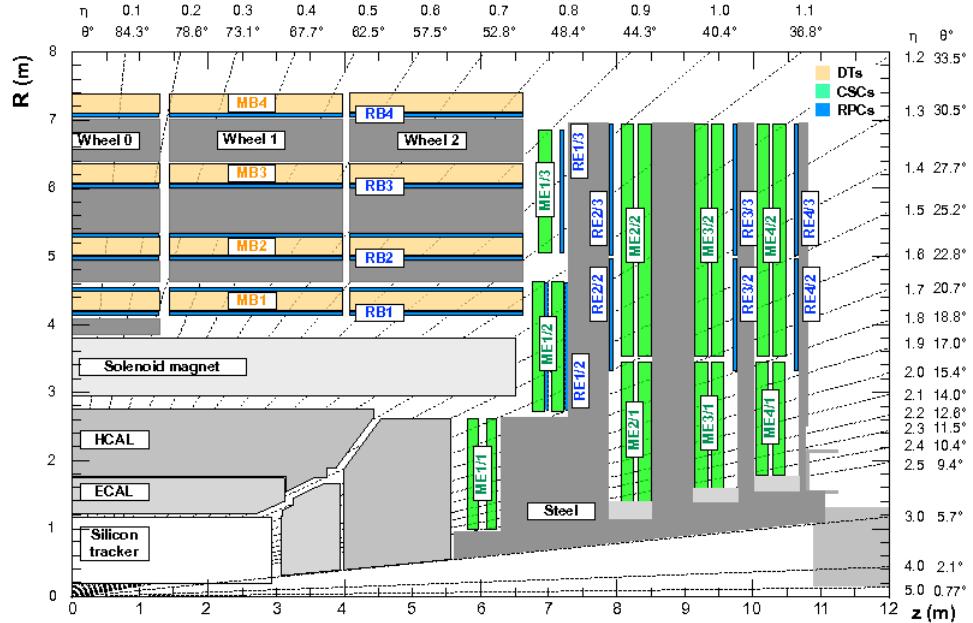


Figure 3.15: The CMS muon tracker. DT, CSC, and RPC detectors are shown in yellow, green, and blue respectively.

CMS muon system has three subdetectors: the drift tubes (DTs), the cathode strip chamber detectors (CSCs), and the resistive plate chambers (RPCs).

In the barrel region, the CMS is equipped with the DT system, consisting of 250

drift tubes arranged into five barrel sections (“wheels”). The working elements of the DT system - cylindrical cells with the rectangular base of 4.2 by 1.3 cm - are tubes with an anode wire in a mix of argon and CO₂ gases. DT cells are 2.4 m long and are organised in three groups of four elements (three “super-layers”). When the muon passes through super-layers, it ionizes the gas in the cells and the released electrons start moving to anodes. In simple terms, the muon position and direction can be determined from the time it takes for electrons to reach the anodes. DT resolution of single-cell hit positions ranges from 200 μm in the r- φ plane to 200-600 μm for forward directions.

CSCs are used in the forward direction to cover the region of $0.9 < |\eta| < 2.4$. The CSCs are multi-wire chambers made of cells that have a trapezoidal shape. The chambers contain radial copper cathode strips and, perpendicular to those, gold-plated tungsten anode wires. Each cell is filled with the mix of argon, CO₂, and CF₄ gases. The strip cells have a single-layer resolution of 300-900 μm . A CSC provides a spatial resolution of 40 -150 μm .

To improve the performance of DTs and CSCs, RPCs are used and cover the barrel and endcaps in the range of $|\eta| < 1.9$. RPCs are double-gap chambers consisting of two resistive 2 mm in thickness Bakelite layers separated by a 2 mm layers filled with a mix of C₂H₂F₄, isobutane C₄H₁₀, and SF₆ gases. RPCs operate in avalanche mode, producing an avalanche when the muon traverses the gas of the cell. RPCs have a spatial resolution of 0.8 - 1.2 cm, which is not as good as the ones provided by other muon subsystems, but RPCs have an advantage in terms of an excellent time resolution - just 3 ns. The barrel and the endcaps contain in total 10 RPC stations.

3.2.7 The Triggers and DAQ

The CMS trigger [66] is a system responsible for selecting events of interest and storing them for the offline analysis. The trigger has two stages: the L1 trigger (see Fig. 3.16), which reduces the event rate from 40 MHz to 100 kHz, and the HLT, which further decreases the rate to nearly 1 kHz. The L1 trigger consists of custom hardware that processes a part of the information from calorimeters and the outer tracker systems. The HLT is a part of the detector readout system (DRS) and uses the full detector information for event reconstruction. The HLT is a computing farm consisting of 22000 CPU cores that produce a decision on whether to save or to skip the event with an average time of about 220 μ s. The DRS is integrated in the higher level data acquisition (DAQ) system [67]. The events selected by the HLT, which is very similar to offline event selection, are collected and sent by the DAQ to the tapes of the main CERN computing centre (Tier-0) for the persistent storage.

The L1 and HLT systems have differences and similarities. They operate at different time scales and the volumes of data they are processing are completely different. However, the goals of these systems are similar: to identify and reconstruct physics objects and combine their properties to produce an acceptance/rejection decision for each event.

3.2.7.1 The L1 Trigger

The L1 system [62] contains a so-called menu of 500 algorithms, called “seeds”, designed to identify useful physics events. These seeds include trigger criteria varying from basic single-object identification to complicated selections that require certain topological conditions to be met. Each seed has a set of assigned “prescale” factors f that reduce the rate of events accepted by a particular trigger algorithm from 100% to

$100/f\%$. One of the purposes to introduce prescale factors is that the luminosity level decreases during the run period and they adjust the trigger rate to keep it constant during the data taking time.

Since the processing time of the L1 system is very important for the whole CMS operation, the L1 is built using FPGAs and ASICs custom hardware. L1 produces decisions within $3.8 \mu s$. Data from all the calorimeters are first processed by the L1 regional calorimeter trigger (RCT) and then by a more selective global calorimeter trigger (GCT).

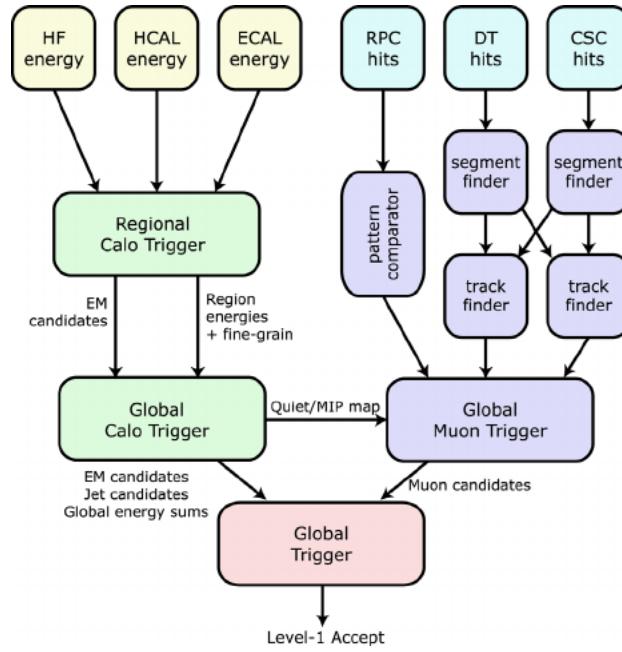


Figure 3.16: The CMS L1 trigger layout.

The RCT receives the information about energy deposits from all calorimeters and covers the range $|\eta| < 5$. The RCT processes this information in parallel and produces e/γ candidates as an output. The information from the inner tracker is not available, therefore, L1 identifies both electrons or photons, but cannot distinguish them. L1 also detects jets, taus, missing transverse energy (MET or E_T or \not{p}_T), and muons. L1 RCT is responsible for determining the first estimates of several main

parameters of interest: p_T , isolation (described later in this chapter), etc. The CMS has a limited resolution on any of these quantities at L1 and, therefore, the HLT is needed (discussed in the next subsection).

First stage reconstruction uses particle hits in the muon detectors and analyzes them using track finder algorithms. All three muon detectors of the CMS are used by the L1 muon trigger. Using DT and CSC systems, track segments from the hit information are identified. The pattern recognition algorithms are applied to these segments to reconstruct muon candidates and measure their momenta.

More complex, but slower, algorithms then reuse hits for more precise particle identification using a global muon trigger (GMT). The hits from the RPCs are used directly by the pattern comparator trigger (PACT) which reconstructs muon candidates at the high radii. Then several regional track finder algorithms sort the identified muon candidates and send this information to the GMT. Each candidate contains p_T and angular information. The GMT then combines the muon information from different subsystems to avoid duplicating the candidates. The GMT also performs more precise quality checks and may discard a portion of the input candidates it receives.

Finally, the information from the GCT and GMT is combined by a global trigger (GT). The GCT sorts the created e/γ candidates, identifies jets, and calculates \cancel{E}_T . The final decision of the GT is to store or to skip the event. If the event satisfies the acceptance requirements and is going to be retained, the L1 accept signal (LAS) is generated and propagated by the trigger control and distribution system (TCDS) to all subdetectors. The information from this signal is then refined by the HLT.

The GT is the final step of the CMS L1 trigger system and implements a menu of triggers. The output of this system is used as an input to the HLT algorithms.

3.2.7.2 The High-Level Trigger

The selection done by the HLT mimics the offline analysis. For all reconstructed objects in the event, including electrons, muons, and jets, the identification criteria are applied to select only events of interest (defined by the physics program of the CMS). Each offline analysis defines what these events are in a different way, but to name a few, almost all analyses need true prompt leptons, well-reconstructed jets, or some other commonly used objects.

The HLT computing farm has an event filter farm, which consists of filter-builder units (FBU). In the FBU the parts of the events and information from different detector subsystems is combined to produce complete events. Then the filter unit unfolds the raw detector data into an experiment specific data structure and performs the event reconstruction and trigger filtering.

The whole event processing procedure of the HLT is centered around the HLT path. The HLT path is a set of algorithmic instructions that in a sequential manner reconstructs physics objects and performs the object selection. The instructions include kinematic and isolation requirements on the objects. The complexity of the steps in the path sequence increases and the quality of the physics objects (the probability to have a correct label) improves too (with respect to the L1). After this step is completed, selected events are sent to another software processing farm. In this storage-manager farm, the date is archived, stored locally on disk, and later sent to the CMS Tier-0 for offline use.

Most data enters the queue for processing and is ready to be sent to Tier-0 very soon. In some cases, the CMS prefers not to process immediately the special data, also known as the parked data. These data are collected and kept until the run is finished. In these situations, the data is sent for parking and the CMS tape is

used. This mostly includes physics analyses such as vector boson fusion or parton distribution studies with Drell-Yan processes.

The output of the HLT is limited by capacities of the Tier-0. This includes the bandwidth of the data transfer, how quickly we want to process and store the data, and, to a degree, the amount of the available tape. All these factors complicate the work of the DAQ, because in addition to physics data streams, the calibration streams also need to be processed and stored. These streams, though, use information only from few subdetectors.

3.2.7.3 The DAQ system

The DAQ systems in modern high energy physics are responsible for many tasks. The challenges are well known: high data rates and volumes, limited tape space, and limited transfer bandwidth. CMS DAQ is based on homogeneous architecture, scales well with the different beam energy regimes and data rates, and has a stable performance in a variety of operating conditions.

To illustrate an example of a complex computing task that is elegantly solved by the DAQ system, discussing the aforementioned FBUs system of the HLT in more detail is needed. The FBU relies on a single multi-core machine in which the communication with other units is done via shared memory. The data from the full detector is used for the filtering process. Complicated offline-like reconstruction algorithms are then used for the full precision event selection. With more CPU cores available for Run 2, the per-event time budget to run HLT-style reconstruction is increased to a “comfortable” 175 ms per event, which is a long enough time to run most of the CMS reconstruction algorithms.

The current CMS DAQ was developed to address these core requirements:

- The data from one or several data transfer lines are available for other lines,
- the event building is done in parallel profiting from multiple processing units,
- almost real-time process monitoring.

Proper design patterns are used in the software for the DAQ, which decouple the user interface from the implementation. The design also allows for the remote control. The software system can be run on a number of different operating systems and hardware platforms. Finally, the memory management tools of the underlying system are not linked directly to the applications, it is done using a dedicated abstract addressing scheme.

3.2.8 The CMS design

We discussed all CMS subdetectors and DAQ. Now, one can summarize in one list the requirements on the design of the CMS detector to successfully complete its physics program. We refer to the CMS Technical Design Report [62]:

- good muon momentum resolution over the momentum scale covering almost a TeV range with the focus on dimuon resolution (mostly $Z \rightarrow \mu\mu$ and $H \rightarrow \mu\mu$) at the O(100) GeV. The capability to determine correctly the charge of the highly energetic muon all the way up to 1 TeV,
- good momentum resolution of all charged particles in the inner tracker,
- good diphoton mass resolution with the focus on the $H \rightarrow \gamma\gamma$ discovery channel. Also, the ability to reject $\pi^0 \rightarrow \gamma\gamma$, which is one of the main background processes to many physics analyses. This requirement mostly concern the performance of the ECAL,

- good resolution of the missing transverse energy (discussed in the next section) and of the mass of the two-jet system. This task depends heavily on the performance of the HCAL.

Since in this dissertation we study boosted Higgs and Z bosons, as well as leptons from $Z(\ell\ell)$ decays and b jets from the $H \rightarrow b\bar{b}$ decays, all the CMS subdetectors are equally important for the double Higgs boson measurement.