

1 **Using Machine Learning Techniques for Data Quality**
2 **Monitoring at CMS Experiment**

3 by

4 Guillermo A. Fidalgo Rodríguez

5 A thesis presented for the degree of

6 BACHELLOR'S OF SCIENCE??

7 in

8 Physics

9 UNIVERSITY OF PUERTO RICO
10 MAYAGÜEZ CAMPUS

11 2018

12 Approved by:

13 _____
14 Sudhir Malik, Ph.D.

15 President, Graduate Committee
Date

16 _____
17 Héctor Méndez, Ph.D.

18 Member, Graduate Committee
Date

19 _____
20 Samuel Santana Colón, Ph.D.
21 Member, Graduate Committee

Date

22 _____
23 Rafael A. Ramos, Ph.D.
24 Chairperson of the Department

Date

²⁵ **Abstract**

²⁶ The Data Quality Monitoring (DQM) of CMS is a key asset to deliver high-quality
²⁷ data for physics analysis and it is used both in the online and offline environment. The cur-
²⁸ rent paradigm of the quality assessment is labor intensive and it is based on the scrutiny of
²⁹ a large number of histograms by detector experts comparing them with a reference. This
³⁰ project aims at applying recent progress in Machine Learning techniques to the automa-
³¹ tion of the DQM scrutiny. In particular the use of convolutional neural networks to spot
³² problems in the acquired data is presented with particular attention to semi-supervised
³³ models (e.g. autoencoders) to define a classification strategy that doesn't assume previ-
³⁴ous knowledge of failure modes. Real data from the hadron calorimeter of CMS are used
³⁵ to demonstrate the effectiveness of the proposed approach.

³⁶ *Keywords:* [DQM, online, offline, Machine Learning]

³⁷ **Acknowledgments**

³⁸ I wish to thank United States State Department and University of Michigan for pro-
³⁹ viding the opportunity to work abroad at CERN during the 2018 Winter Semester. I also
⁴⁰ wish to thank CERN staff, CMS Experiment , Texas Tech University, and the University
⁴¹ of Puerto Rico at Mayagüez, with special thanks to Dr. Federico de Guio for his local
⁴² mentorship and Dr. Nural Akchurin, and Dr. Sudhir Malik for their guidance. A very
⁴³ special thanks to Dr. Jean Krisch for accepting me for this great research opportunity and
⁴⁴ Dr. Steven Goldfarb for being a wonderful host and overall local guidance at CERN.

⁴⁵ List of Figures

46	Figure 2.1: CMS Detector	4
47	Figure 2.2: The trajectory of a particle traveling through the layers of the	
48	detector leaving behind a it's signature footprint	5
49	Figure 3.1: The CERN Accelerator Complex [1].	9
50	Figure 3.2: The CMS Detector Layout [2].	11
51	Figure 3.3: Overview of the CMS Tracker Layout [3].	12
52	Figure 3.4: Geometrical layout of the CMS Electromagnetic Calorimeter [4].	14
53	Figure 3.5: Layout of the CMS ECAL illustrating its various components [5].	16
54	Figure 3.6: Geometrical layout of the HCAL showing the locations of the	
55	hadron barrel (HB), endcap (HE), outer (HO) and forward (HF)	
56	calorimeters [6].	17
57	Figure 3.7: Quarter-view of the CMS muon system (left) and a muon in the	
58	transverse plane leaves a curved trajectory across the four muon	
59	detector stations (right).	19
60	Figure 3.8: Transverse slice of the CMS detector, showing the individual de-	
61	tector subsystems and particle signatures in each. The particle	
62	type can be inferred by combining the detector response in the	
63	different subdetectors [7].	21
64	Figure 3.9: Phase-2 Pixel Detector Layout[ref].	24

65 **Contents**

66	Abstract	i
67	Acknowledgments	ii
68	List of Figures	iii
69	1 Introduction	1
70	2 The CMS Experiment	3
71	3 Data Collection and Data Quality Monitoring	7
72	3.1 What is Data Collection for CMS?	7
73	3.2 What is Data Quality Monitoring	8
74	3.3 The CMS Detector	10
75	3.3.1 Silicon Tracking System	12
76	3.3.2 Electromagnetic Calorimeter	14
77	3.3.3 Hadron Calorimeter	16
78	3.3.4 Magnet	18
79	3.3.5 Muon Detector	18
80	3.3.6 Trigger and Data Acquisition	20
81	3.3.7 Event Reconstruction	21
82	3.3.8 Future Upgrade of Pixel Detector	22
83	A Appendix Title	25
84	B References	26

⁸⁵ **Chapter 1**

⁸⁶ **Introduction**

⁸⁷ The work for this thesis was performed at CERN on CMS Experiment. CERN stands
⁸⁸ for European Organization for Nuclear Research. It was founded in 1954 and is located
⁸⁹ at the Franco-Swiss border near Geneva. At CERN, physicists and engineers are probing
⁹⁰ the fundamental structure of the universe. They use the world's largest and most complex
⁹¹ scientific instruments to study the basic constituents of matter – the fundamental parti-
⁹² cles. The instruments used at CERN are purpose-built particle accelerators and detectors.
⁹³ Accelerators boost beams of particles to high energies before the beams are made to col-
⁹⁴ lide with each other or with stationary targets. Detectors observe and record the results
⁹⁵ of these collisions. The accelerator at CERN is called the Large Hadron Collider (LHC),
⁹⁶ the largest machine ever built by humans and it collides particles (protons) at close to the
⁹⁷ speed of light. The process gives the physicists clues about how the particles interact, and
⁹⁸ provides insights into the fundamental laws of nature. Seven experiments at the LHC use
⁹⁹ detectors to analyze particles produced by proton-proton collisions. The biggest of these
¹⁰⁰ experiments, ATLAS and CMS, use general-purpose detectors designed to study the fun-
¹⁰¹ damental nature of matter and fundamental forces and to look for new physics or evidence
¹⁰² of particles that are beyond the Standard Model. Having two independently designed de-
¹⁰³ tectors is vital for cross-confirmation of any new discoveries made. The other two major
¹⁰⁴ detectors ALICE and LHCb, respectively, study a state of matter that was present just
¹⁰⁵ moments after the Big Bang and preponderance of matter than antimatter. Each experi-

106 ment does important research that is key to understanding the universe that surrounds and
107 makes us.

108

109 [Chapter 2](#) presents a basic description of the Large Hadron Collider and CMS Detector

110

111 ?? gives a brief motivation

112

113 ?? is dedicated to a study optimizing

114

115 ?? ptimated.

116

117 ?? details an improvarger production cross-section than Z+jets process used before.

118

119 The conclusions and results of each chapter are presented in the corresponding chap-
120 ter.

121

122 This thesis work has been presented at several internal meetings of the CMS Experi-
123 ment and at the following international meetings and conferences:

124 1. **Andrés Abreu** gave a talk “*Estimation of the Z Invisible Background for Searches*

125 *for Supersymmetry in the All-Hadronic Channel*” at “APS April 2018: American
126 Physical Society April Meeting 2018, 14-17 Apr 2018”, Columbus, OH

127 2. **Andrés Abreu** gave a talk “*Phase-2 Pixel upgrade simulations*” at the “USLUA

128 Annual meeting: 2017 US LHC Users Association Meeting, 1-3 Nov 2017”, Fer-
129 milab, Batavia, IL

¹³⁰ **Chapter 2**

¹³¹ **The CMS Experiment**

¹³² The Compact Muon Solenoid (CMS) detector is a general purpose particle detector
¹³³ designed to investigate various physical phenomena concerning the SM and beyond it,
¹³⁴ such as Supersymmetry, Extra Dimensions and Dark Matter. As its name implies, the
¹³⁵ detector is a solenoid which is constructed around a superconducting magnet capable of
¹³⁶ producing a magnetic field of 3.8 T. The magnetic coil is 13m long with an inner diameter
¹³⁷ of 6m, making it the largest superconducting magnet ever constructed. The CMS detector
¹³⁸ itself is 21m long with a diameter of 15m and it has a weight of approximately 14,000
¹³⁹ tons. The CMS experiment is one of the largest scientific collaborations in the history
¹⁴⁰ of mankind with over 4,000 participants from 42 countries and 182 institutions. CMS is
¹⁴¹ located at one of these points and it essentially acts as a giant super highspeed camera
¹⁴² that makes 3D images of the collisions that are produced at a rate of 40 MHz (40 million
¹⁴³ times per second). The detector has an onion-like structure to capture all the particles that
¹⁴⁴ are produced in these high energy collisions most of them being unstable and decaying
¹⁴⁵ further to stable particles that are detected. CMS detector was designed with the following
¹⁴⁶ features (as shown in [Figure 3.2](#)) :

- ¹⁴⁷ 1. A **magnet** with large bending power and high performance muon detector for good
¹⁴⁸ muon identification and momentum resolution over a wide range of momenta and
¹⁴⁹ angles.

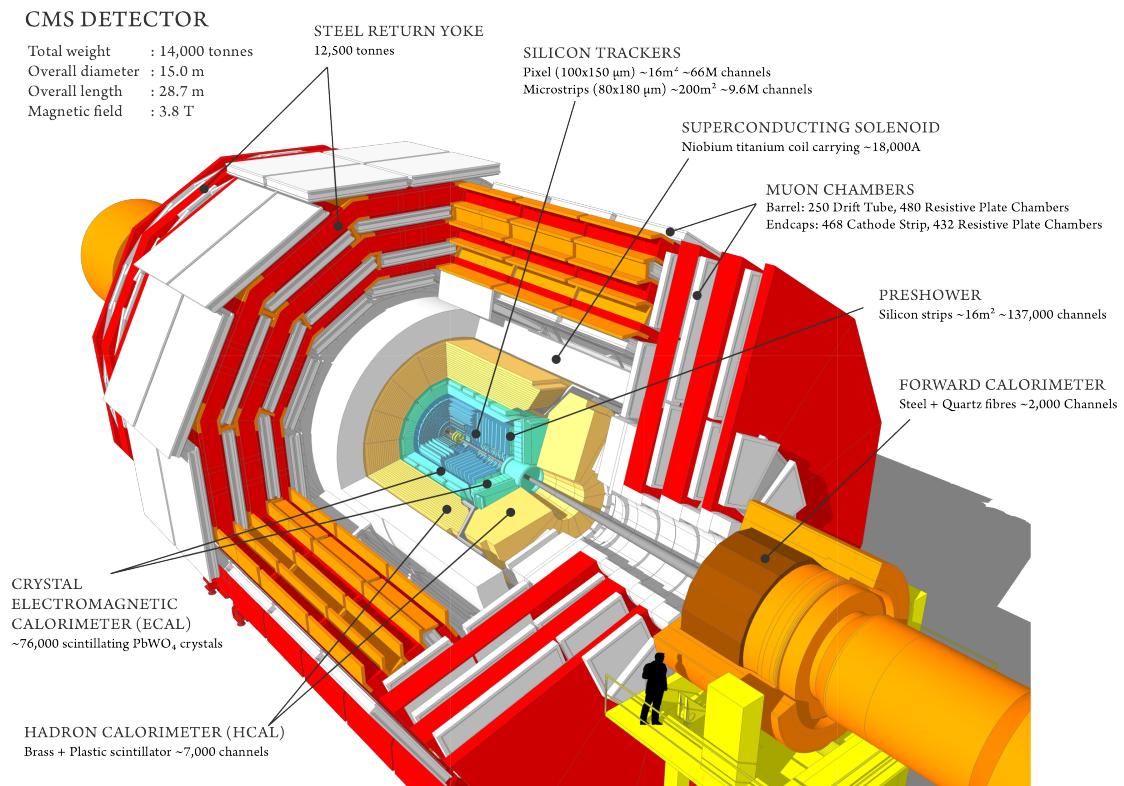


Figure 2.1: CMS Detector

- 150 2. An **inner tracking system** capable of high reconstruction efficiency and momen-
151 tum resolution requiring **pixel detectors** close to the interaction region.
- 152 3. An **electromagnetic calorimeter** able to provide good electromagnetic energy res-
153 olution and a high isolation efficiency for photons and leptons.
- 154 4. A **hadron calorimeter** capable of providing precise missing-transverse-energy and
155 dijet-mass resolution.

156 A property from these particles that is exploited is their charge. Normally, particles
157 produced in collisions travel in a straight line, but in the presence of a magnetic field,
158 their paths are skewed and curved. Except the muon system, the rest of the subdetectors
159 lie inside a 3.8 Tesla magnetic field . Due to the magnetic field the trajectory of charged
160 particle produced in the collisions gets curved (as shown in [Figure 2.2](#)) and one can
161 calculate the particle's momentum and know the type of charge on the particle. The
162 Tracking devices are responsible for drawing the trajectory of the particles by using a
163 computer program that reconstructs the path by using electrical signals that are left by

the particle as they move. The Calorimeters measure the energy of particles that pass through them by absorbing their energy with the intent of stopping them. The particle identification detectors work by detecting radiation emitted by charged particles and using this information they can measure the speed, momentum, and mass of a particle. After the information is put together to make the “snapshot” of the collision one looks for results that do not fit the current theories or models in order to look for new physics.

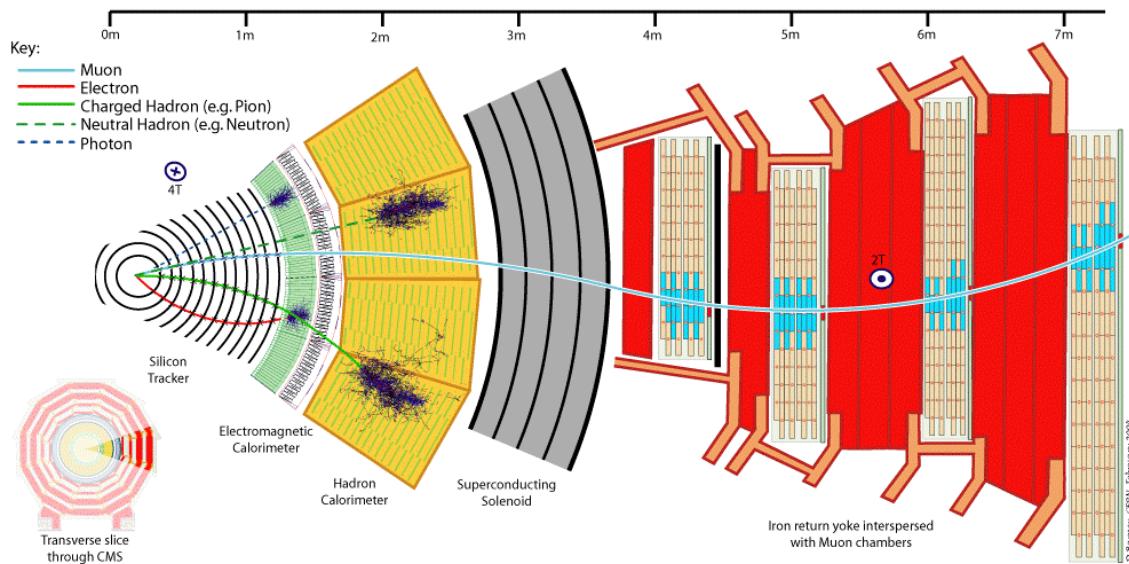


Figure 2.2: The trajectory of a particle traveling through the layers of the detector leaving behind a its signature footprint

The project focusses specifically on data collected from one of the Calorimeters, - the Hadron Calorimeter (HCAL). The HCAL, as its name indicates, is designed to detect and measure the energy of hadrons or, particles that are composed of quarks and gluons, like protons and neutrons. Additionally, it provides an indirect measurement of the presence of non-interacting, uncharged particles such as neutrinos (missing energy) . Measuring these particles is important as they can tell us if new particles such as the Higgs boson or supersymmetric particles (much heavier versions of the standard particles we know) have been formed. The layers of the HCAL are structured in a staggered fashion to prevent any gaps that a particle might pass through undetected. There are two main parts: the barrel and the end caps. There are 36 barrel wedges that form the last layer of the detector inside the magnet coil, there is another layer outside this, and on the endcaps, there are another 36 wedges to detect particles that come out at shallow angles with respect to the beam

182 line.

¹⁸³ Chapter 3

¹⁸⁴ Data Collection and Data Quality

¹⁸⁵ Monitoring

¹⁸⁶ 3.1 What is Data Collection for CMS?

¹⁸⁷ During data taking there are millions of collisions occurring in the center of the detector every second. The data per event is around one million bytes (1 MB), that is produced at a rate of about 600 million events per second [8], that's about 600 MB/s. Keeping in mind that only certain events are considered "interesting" for analysis, the task of deciding what events to consider out of all the data collected is a two-stage process. First, the events are filtered down to 100 thousand events per second for digital reconstruction and then more specialized algorithms filter the data even more to around 100 200 events per second that are found interesting. For CMS there is a Data Acquisition System that records the raw data to what's called a High-Level Trigger farm which is a room full of servers that are dedicated to processing and classify this raw data quickly. The data then gets sent to what's known as the Tier-0 farm where the full processing and the first reconstruction of the data are done. [9]

¹⁹⁹ 3.2 What is Data Quality Monitoring

²⁰⁰ To operate a sophisticated and complex apparatus as CMS, a quick online feedback on
²⁰¹ the quality of the data recorded is needed to avoid taking low quality data and to guarantee
²⁰² a good baseline for the offline analysis. Collecting a good data sets from the collisions
²⁰³ is an important step towards search for new physics as deluge of new data poses an extra
²⁰⁴ challenge of processing and storage. This all makes it all the more important to design
²⁰⁵ algorithms and special software to control the quality of the data. This is where the Data
²⁰⁶ Quality Monitoring (DQM) plays a critical in the maintainability of the experiment, the
²⁰⁷ operation efficiency and performs a reliable data certification. The high-level goal of
²⁰⁸ the system is to discover and pinpoint errors, problems occurring in detector hardware
²⁰⁹ or reconstruction software, early, with sufficient accuracy and clarity to maintain good
²¹⁰ detector and operation efficiency. The DQM workflow consists of 2 types: **Online** and
²¹¹ **Offline**.

²¹² The Online DQM consists of receiving data taken from the event and trigger his-
²¹³ tograms to produce results in the form of monitoring elements like histogram references
²¹⁴ and quality reports. This live monitoring of each detector's status during data taking gives
²¹⁵ the online crew the possibility to identify problems with extremely low latency, mini-
²¹⁶ mizing the amount of data that would otherwise be unsuitable for physics analysis. The
²¹⁷ scrutiny of the Online DQM is a 24/7 job that consists of people or shifters that work at the
²¹⁸ CMS control center constantly monitoring the hundreds of different plots and histograms
²¹⁹ produced by the DQM software. This consumes a lot of manpower and is strenuous work.
²²⁰ The Offline DQM is more focused on the full statistics over the entire run of the experi-
²²¹ ment and works more on the data certification. In the offline environment, the system is
²²² used to review the results of the final data reconstruction on a run-by-run basis, serving as
²²³ the basis for certified data used across the CMS collaboration in all physics analyses. In
²²⁴ addition, the DQM framework is an integral part of the prompt calibration loop. This is
²²⁵ a specialized workflow run before the data are reconstructed to compute and validate the
²²⁶ most up-to-date set of conditions and calibrations subsequently used during the prompt

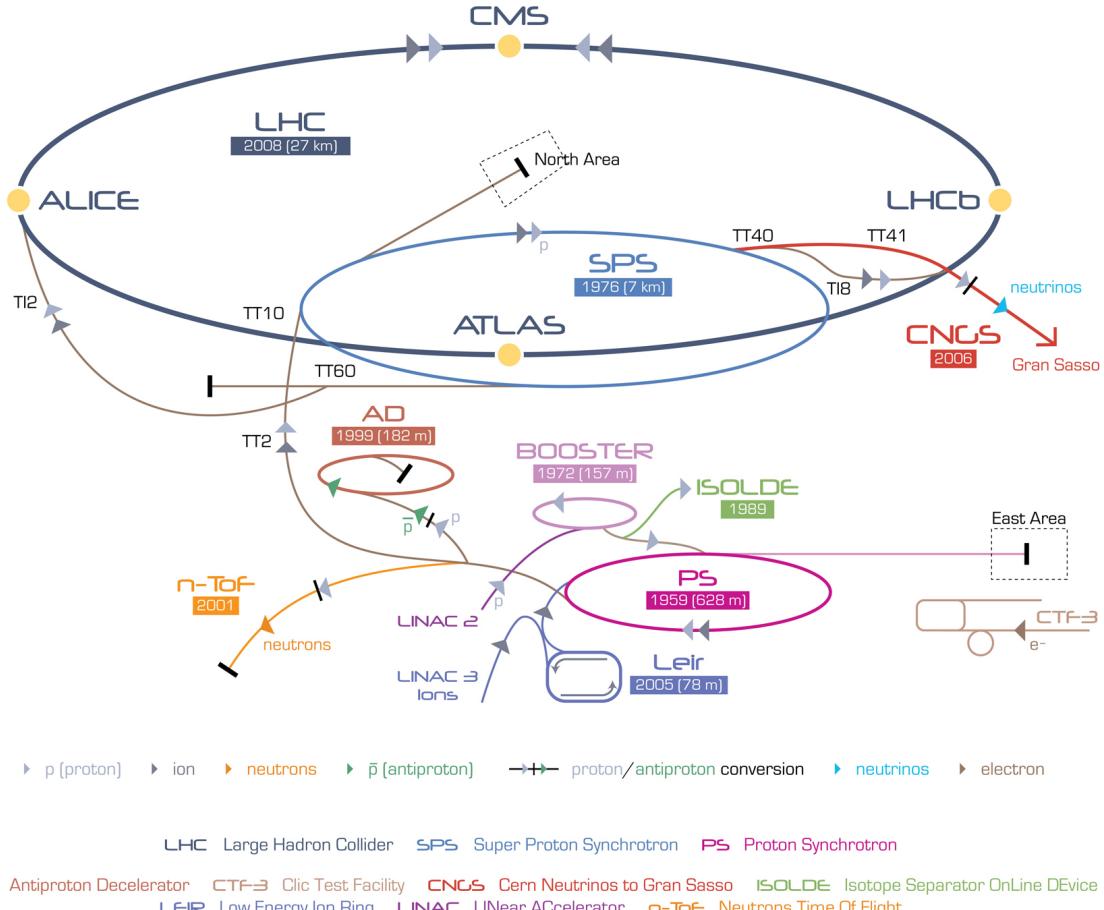


Figure 3.1: The CERN Accelerator Complex [1].

reconstruction. This project aims to minimize the DQM scrutiny by eye and automate the process so that there is a more efficient process to monitor the detector and the quality of the data by implementing Machine Learning techniques.

The four main detectors comprising the LHC machine are CMS, ATLAS [10], LHCb [11] and ALICE [12]. Both CMS and ATLAS are general purpose detectors whose initial designs had the detection of the SM Higgs boson, with its wide range of decay modes, in mind. Both detectors managed to accomplish this goal when a 126 GeV scalar boson consistent with the SM Higgs was independently verified by both experiments in July of 2012. Furthermore, the designs for CMS and ATLAS allow for the search of many other additional phenomena in BSM physics such as Supersymmetry, Dark Matter [13], Dark Sector [14], etc. On the other hand, the LHCb and ALICE detectors focus on more particular kinds of searches. The main motivation for the LHCb experiment, where the b stands

239 for beauty, concerns itself with the measurement of CP violation parameters in b-hadron
 240 interactions and studies cover a wide range of aspects of Heavy Flavor Electroweak and
 241 QCD physics. Meanwhile, the ALICE experiment focuses on the study of heavy ion (Pb-
 242 Pb) nuclei collisions at a centre-of-mass energy of 2.76 TeV in order to better understand
 243 the physics behind strongly interacting matter at extreme energy densities.

244 3.3 The CMS Detector

245 The Compact Muon Solenoid (CMS) detector is a general purpose particle detector
 246 designed to investigate various physical phenomena concerning the SM and beyond it,
 247 such as Supersymmetry, Extra Dimensions and Dark Matter. As its name implies, the
 248 detector is a solenoid which is constructed around a superconducting magnet capable of
 249 producing a magnetic field of 3.8 T. The magnetic coil is 13m long with an inner diameter
 250 of 6m, making it the largest superconducting magnet ever constructed. The CMS detector
 251 itself is 21m long with a diameter of 15m and it has a weight of approximately 14,000
 252 tonnes. The CMS experiment is one of the largest scientific collaborations in the history
 253 of mankind with over 4,000 participants from 42 countries and 182 institutions.

254

255 In order to meet the many needs of the SM and BSM searches, and the goals of the
 256 LHC physics program, the CMS detector was designed with the following features:

257

- 258 • A magnet with large bending power and high performance muon detector for good
 259 muon identification and momentum resolution over a wide range of momenta and
 260 angles.

261

- 262 • An inner tracking system capable of high reconstruction efficiency and momentum
 263 resolution requiring pixel detectors close to the interaction region.

264

- 265 • An electromagnetic calorimeter able to provide good electromagnetic energy reso-
 266 lution and a high isolation efficiency for photons and leptons.

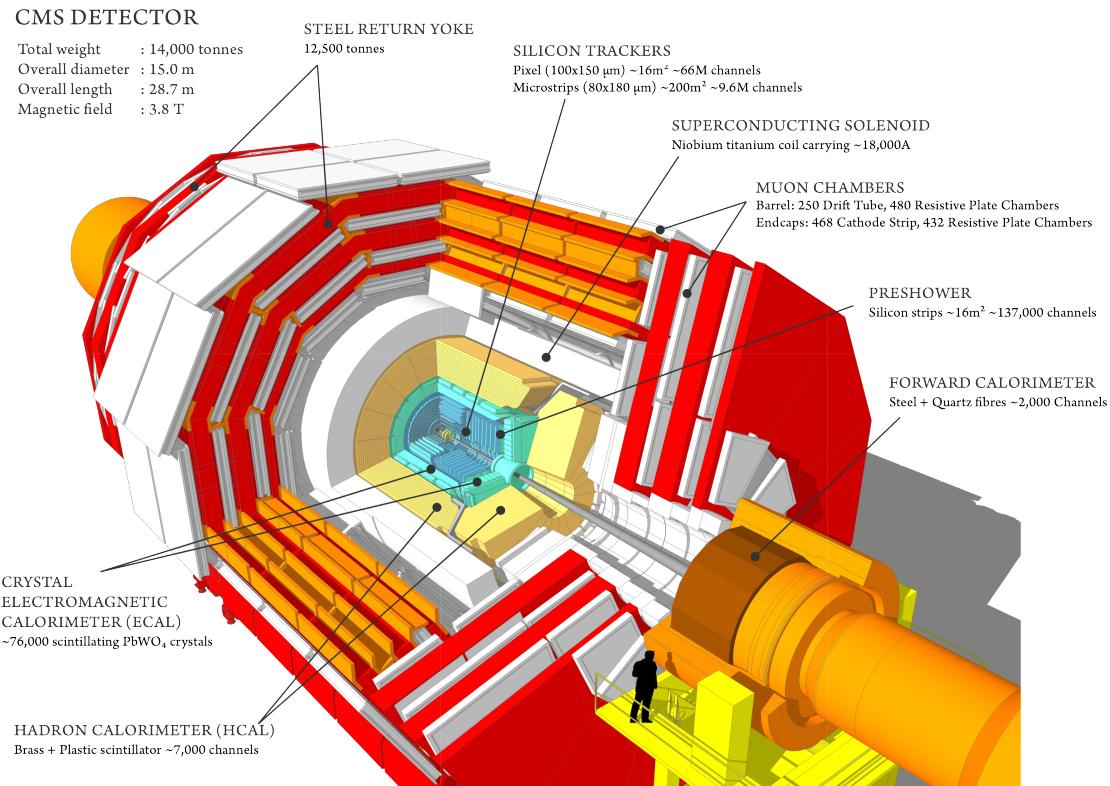


Figure 3.2: The CMS Detector Layout [2].

267

- 268 • A hadron calorimeter capable of providing precise missing-transverse-energy (p_T^{miss})
269 and dijet-mass resolution.

270

271 A general layout of the CMS detector and all its constituent sub-detectors can be seen
272 in [Figure 3.2](#). The configuration of the CMS sub-detectors follow a cylindrical layer
273 pattern that is symmetrical about the interaction region and consists of a central barrel
274 with endcaps on both ends.

275 The coordinate system for the CMS detector design uses a right-hand rule conven-
276 tion centered around the ideal interaction point to describe the positions of objects in the
277 experiment. The z-axis is defined along the direction of the LHC beam, with the x-axis
278 pointing towards the center of the LHC ring. In terms of polar coordinates then, r is the
279 radial distance from the center of the pipe, the polar angle θ is measured against the z-axis
280 and the azimuthal angle ϕ is measured with respect to the x-axis. However, the pseudo-
281 rapidity η is generally preferred over the polar angle θ . The pseudorapidity is defined

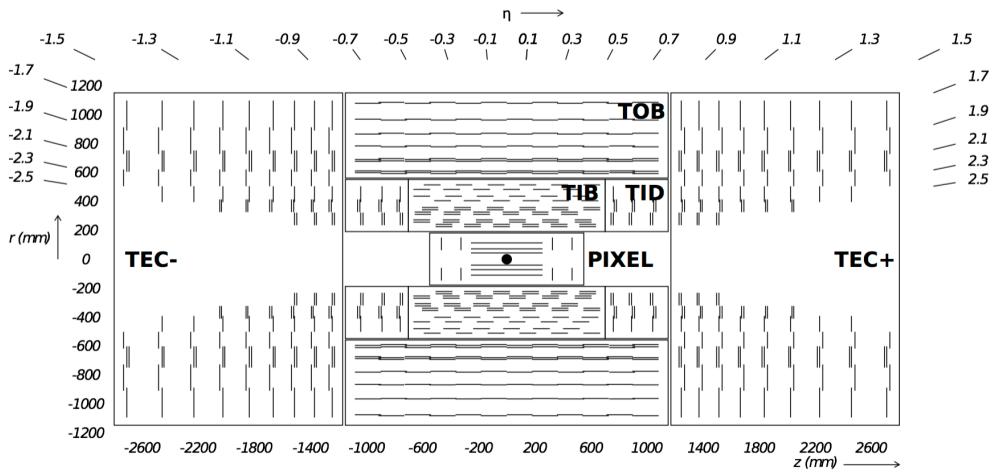


Figure 3.3: Overview of the CMS Tracker Layout [3].

282 as:

283
$$\eta = -\ln \tan \frac{\theta}{2}.$$

284 **3.3.1 Silicon Tracking System**

285 The CMS tracking system was designed with the goal of obtaining precise and effi-
 286 cient measurements for the trajectories of charged particles resulting from proton-proton
 287 collisions at the LHC. In addition, it allows for the precise measurement of secondary ver-
 288 tices and impact parameters needed to efficiently identify the heavy flavours produced in
 289 many interesting physics channels. Due to the LHC's design Luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$,
 290 the currently installed CMS phase-1 tracker is expected to handle an average of 1000 par-
 291 ticles from over 20 overlapping proton-proton interactions per bunch crossing, every 25
 292 ns. This required a detector technology capable of achieving a high granularity and fast
 293 response as well as being tolerant to the radiation produced from the intense particle flux.
 294 These considerations lead to a tracker design composed entirely of silicon detector tech-
 295 nology which features an active silicon area of about 200 m^2 , making it the largest silicon
 296 tracker ever built [5].

297

298 The CMS tracker is built in a cylindrical manner around the interaction point and has
 299 a diameter of 2.5 m and a length of 5.8 m. It is comprised of a pixel detector with three

300 barrel layers, positioned at a distance between 4.4 cm and 10.2 cm from the interaction
301 region, and a silicon strip tracker with 10 barrel detection layers extending to a radius of
302 1.1 m. Each system is made complete by endcaps at opposite sides of the barrel, consist-
303 ing of 2 discs for the pixel detector and 3 plus 9 discs for the strip tracker, extending the
304 acceptance region of the tracker up to a pseudorapidity $|\eta| < 2.5$.

305

306 The pixel detector is the part of the CMS tracking system closest to the interaction
307 region and consists of 3 barrel layers (BPix) and 2 endcap discs (FPix). It is responsible
308 for providing precise tracking points in $r\text{-}\phi$ and z , a feature that is required for the small
309 impact parameter resolution, needed for good secondary vertex reconstruction. The de-
310 tector contains 1440 modules covering an area of approximately 1 m^2 making up a total
311 of 66 million pixels. The sensors were designed using an n-on-n concept from $320 \mu\text{m}$
312 thick silicon wafers and are fabricated with read-out chips (ROCs) that are bump-bonded
313 to the sensor in standard $0.25 \mu\text{m}$ CMOS technology. Each of the pixels has a pitch size
314 of $100 \times 150 \mu\text{m}^2$, which corresponds to an occupancy of about 10^{-4} per bunch crossing.

315

316 The silicon strip tracker is built surrounding the pixel detector and consists of three
317 large subsystems. The Tracker Inner Barrel and Disks (TIB/TID) extend to a radius of
318 55 cm and are composed of four barrel layers, completed by three disks at each end. The
319 TIB/TID employs the use of $320 \mu\text{m}$ thick silicon micro-strip sensors in order to deliver
320 up to 4 $r\text{-}\phi$ measurements on a trajectory. Surrounding the TIB/TID is the Tracker Outer
321 Barrel (TOB), which consists of 6 barrel layers and has an outer radius of 116 cm. The
322 TOB extends symmetrically in z between ± 118 cm and provides an additional 6 $r\text{-}\phi$
323 measurements for a trajectory. Beyond the TOB's z range lie the Tracker Endcaps (TEC+
324 and TEC-, where the sign indicates their position respect to z). Each TEC consists of 9
325 discs with up to 7 rings of silicon micro-strip detectors, providing up to 9 additional ϕ
326 measurements per trajectory. The CMS silicon strip tracker has a total active silicon area
327 of 198 m^2 and is composed of 15,148 sensor modules.

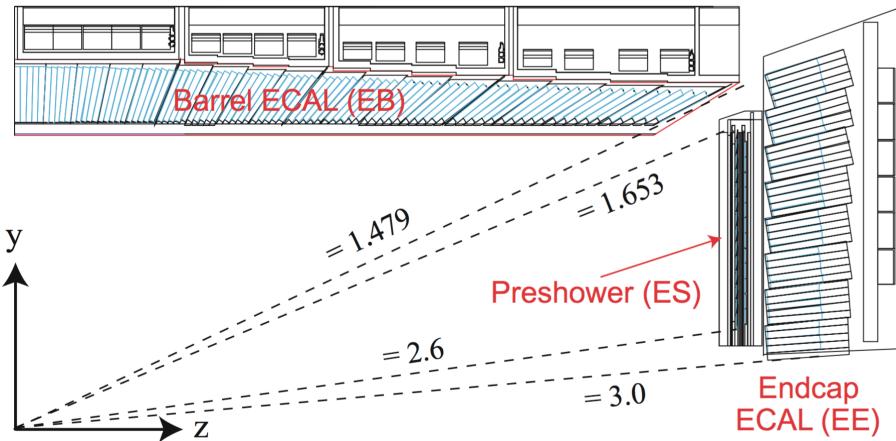


Figure 3.4: Geometrical layout of the CMS Electromagnetic Calorimeter [4].

3.3.2 Electromagnetic Calorimeter

The CMS Electromagnetic Calorimeter (ECAL) is a hermetic homogeneous calorimeter whose function is to measure the energy of particles that interact via the electromagnetic force. With the use of 75,848 scintillator crystals, it is capable of providing good energy resolution within the requirements of the ambitious LHC program. In particular, the ECAL's design was optimized to search for diphoton events resulting from Higgs boson decays ($H \rightarrow \gamma\gamma$).

335

The ECAL is composed of two main sub-systems – the barrel calorimeter (EB) and the endcap calorimeter (EE) – and is completed by a preshower calorimeter (ES), as shown in Figure 3.5. It covers a solid angle up to a pseudorapidity of $|\eta| < 3$, with the EB extending in the range of $|\eta| < 1.479$ and the EE covering a range of $1.479 < |\eta| < 3$. Both subsystems are composed of lead tungstate (PbWO₄) scintillator crystals which provide a fast response time and high radiation tolerance, both crucial requirements for optimal performance at LHC operating conditions. In addition, the properties of the PbWO₄ crystals (high density, short radiation length and small Moliére radius) led to the design of a compact calorimeter with fine granularity.

345

The barrel part of the ECAL (EB) is composed of specially designed avalanche pho-

347 todiodes (APD). It consists of 61,200 crystals, forming a total volume of 8.14 m^3 and
348 weighing about 67.4 t. The crystals that form the EB are 230 mm long with a cross-
349 section of $22 \times 22 \text{ mm}^2$ at the front face and $26 \times 26 \text{ mm}^2$ at the back. They are organized
350 in pairs within thin-walled alveolar structures called submodules. The submodules are
351 assembled into different types of modules that differ by their location with respect to η
352 and contain 400 or 500 crystals each. Theses modules are then arranged in sets of four
353 modules, called supermodules, and contain 1700 crystals each. Thus, the EB is composed
354 of two half-barrels, each consisting of 18 supermodules.

355

356 In contrast, the photodetectors used in the endcap section of the ECAL are vacuum
357 phototriodes (VPT) [15]. Each of the endcaps contain 7,324 crystals which, in total, oc-
358 cupy a volume of 2.90 m^3 and weigh about 24.0 t. The crystals in the EE are 220 mm in
359 length with a cross-section of $28.62 \times 28.62 \text{ mm}^2$ for the front face and $30 \times 30 \text{ mm}^2$ in the
360 rear. They are all identical in shape and are arranged in mechanical units of 5×5 crystals,
361 called supercrystals (SC), which consist of carbon-fibre alveola structures. Each of the
362 endcaps are divided into two semi-circular structures, called *Dees*, which hold a total of
363 3,662 crystals.

364

365 The ES preshower is located before the EE detector and spans a pseudorapidity range
366 of $1.653 < |\eta| < 2.6$. It's main purpose is to identify neutral pions in the endcaps as
367 well as to improve the determination of electrons and photons with high granularity. The
368 preshower consists of two layers and has a total thickness of 20 cm. The first layer is
369 conformed by lead radiators which initiate electromagnetic showers from incoming pho-
370 tons and electrons. Meanwhile, the second layer is composed of silicon strips which are
371 capable of measuring the deposited energy and the transverse shower profiles.

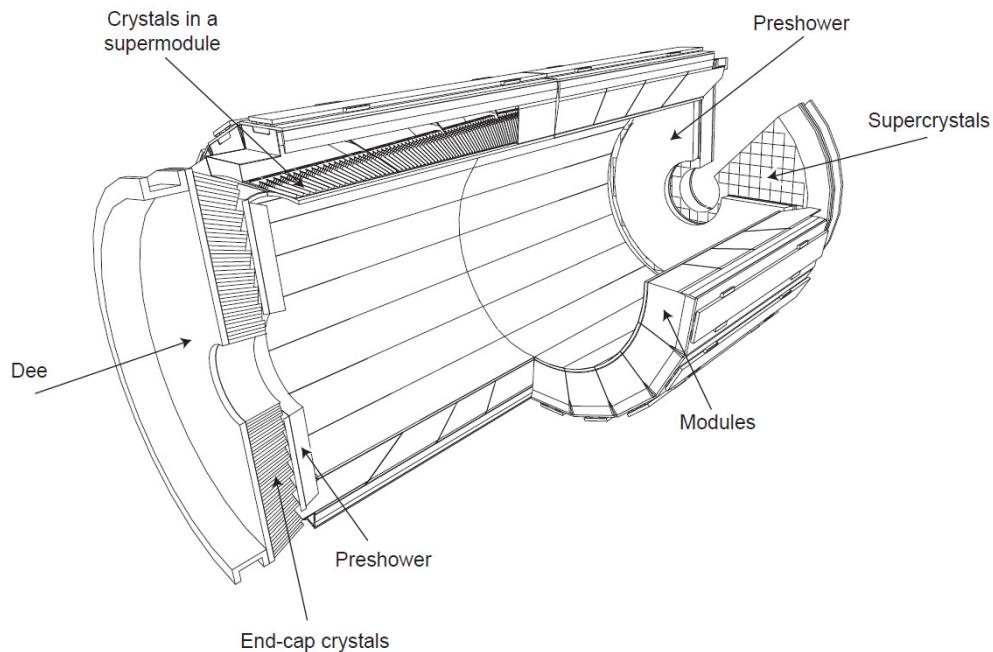


Figure 3.5: Layout of the CMS ECAL illustrating its various components [5].

3.3.3 Hadron Calorimeter

The CMS Hadron calorimeter (HCAL) conforms the next layer of the CMS detector. It is a sampling calorimeter that consists of alternating layers of massive absorbing brass plates and plastic scintillator tiles and is of particular importance for the measurement of hadron jet energy and p_T^{miss} . The HCAL detector is located in between the outer extent of the ECAL ($R = 1.77$ m) and the inner extent of the magnet coil ($R = 2.95$ m). Similar to the other CMS subsystems, it's composed of a barrel part (HB) and an endcap part (HE). In addition, it features a tail-catching outer calorimeter (HO), located outside the magnet, and a forward calorimeter (HF) in the very forward region near the beam line. A layout of the HCAL system can be seen in Figure 3.6.

382

The barrel component of the HCAL is a sampling calorimeter which covers the pseudorapidity range $|\eta| < 1.3$. It consists of both the HB and the HO detectors. The reason behind separating the barrel detector into the HB and HO is due to the limited amount of space available for the barrel detector. The HB is located within the superconducting magnet coil and is supplemented by the HO in between the outer solenoid coil and the

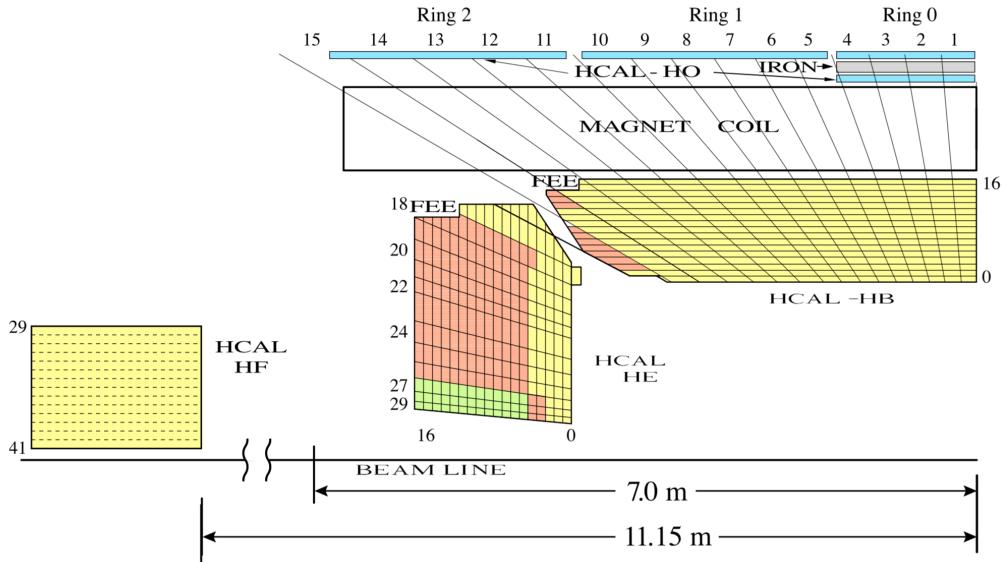


Figure 3.6: Geometrical layout of the HCAL showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters [6].

388 muon chambers. Therefore, the HO acts as a tail-catcher in order to improve the jet en-
 389 ergy measurements and p_T^{miss} , with the solenoid in between acting as absorber material.
 390 The HB consists of two half-barrel sections, identified as HB+ and HB- due to their ge-
 391 ometrical location, which are composed of 36 identical azimuthal wedges. The wedges,
 392 which are constructed out of flat brass absorber plates, are aligned parallel to the beam
 393 axis and are segmented into four azimuthal angle (ϕ) sections.

394

395 The HE covers a significant amount of the pseudorapidity in the range of $1.3 < |\eta| <$
 396 3, a region containing about 34% of the particles produced in the final state. Due to
 397 the high luminosity of the LHC, the HE is required to have a high radiation tolerance at
 398 $|\eta| \simeq 3$, as well as being capable of handling high counting rates. Similar to the HB, the
 399 HE is also composed of brass absorber plates and scintillator plates which are read out
 400 by wavelength shifting fibers. The light captured by the scintillators merges within the
 401 wavelength shifting fibers and then it's read out by hybrid photo-diodes. The scintillators
 402 are partitioned in towers with an area of $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$.

403

404 3.3.4 Magnet

405 The CMS superconducting solenoid is one of the driving features of the detector de-
 406 sign. It is capable of providing a magnetic field with a 3.8T magnitude, which allows
 407 for the large bending power needed for precise particle transverse momentum (p_T) mea-
 408 surements. The magnet is made of four layers of stabilized reinforced Niobium-Titanium
 409 (NbTi) and has a cold mass of 220 t. The solenoid consists of a 13 m long coil with an
 410 internal diameter of 6 m, which houses both the tracking and calorimetric system. This
 411 design allows for particles to be measured prior to crossing the magnetic coil which sig-
 412 nificantly improves the energy resolution.

413 3.3.5 Muon Detector

414 As implied by the detector's name, precise and robust muon measurements have been
 415 a central theme of the CMS experiment since the early stages of its design. The detector
 416 design takes into account that muons behave as minimum ionizing particles (MIPs) [16]
 417 and can therefore manage to traverse the tracker and calorimeters with minimal energy
 418 loss. Furthermore, due to their relatively long lifetime they can be efficiently identified
 419 by a dedicated system at the outer region of the detector. Consequently, the CMS muon
 420 systems comprise the outermost layer of the detector, which are integrated into the mag-
 421 net return yoke that surrounds the solenoid.

422

423 The muon system is capable of three main functions: muon identification, muon p_T
 424 measurement and triggering. It is composed of three different types of detectors, all of
 425 which make use of gaseous chamber technology. This choice of detector provides a cost
 426 efficient way of covering most of the full solid angle, featuring a total of 25,000 m² of de-
 427 tection plates. As a consequence of the shape of the solenoid magnet, the muon detector
 428 was designed to have a cylindrical barrel section as well as two planar endcap regions.

429

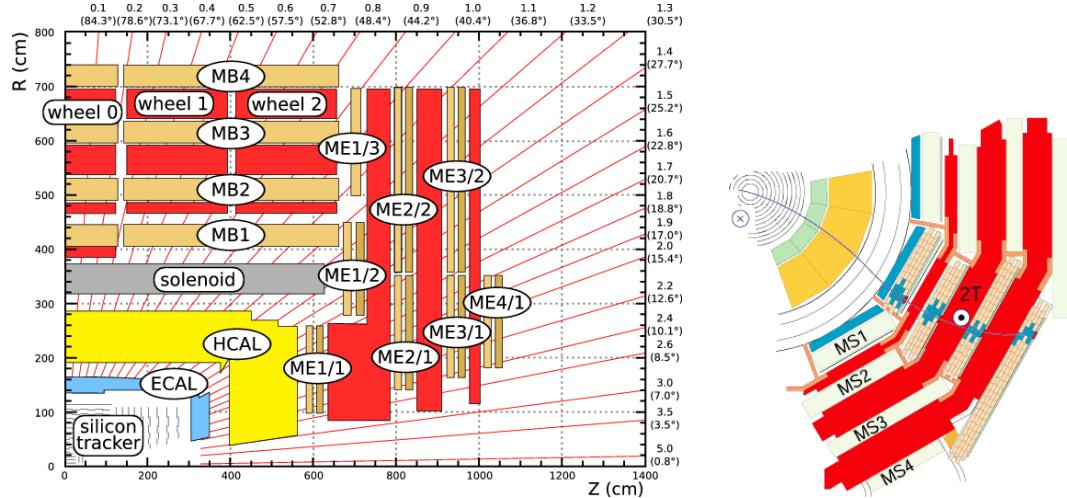


Figure 3.7: The left diagram shows a quarter-view of CMS with both the muon barrel (MB) and endcap (ME) stations [17]. The right diagram shows a muon in the transverse plane leaves a curved trajectory across the four muon detector stations [18].

430 For the barrel region of the muon detector, four layers of drift tube (DT) modules are
 431 used, covering a pseudorapidity of up to $|\eta| < 1.2$. These four layers, called “stations”,
 432 are arranged in cylindrical concentric layers around the beam line, where the first three
 433 layers have 60 DTs each and the outer cylinder has 70. Each of the DT stations contain
 434 12 individual gas filled tubes, all of which have a 4cm diameter and a center electrode.
 435 The use of DTs as tracking detectors for the barrel muon system is possible because of
 436 the low expected muon rate and the relatively low strength of the local magnetic field.

437

438 The endcap regions of the muon system are subject to a higher muon rate, and cover
 439 the range $0.9 < |\eta| < 2.4$ where the magnetic field is stronger and less homogeneous.
 440 Considering these conditions, cathode strip chambers (CSCs) are employed, which fea-
 441 ture high granularity, fast response time and adequate radiation hardness. CSCs consist of
 442 arrays of positively-charged “anode” wires crossed with negatively-charged copper “cath-
 443 ode” strips within a gas volume. They are trapezoidal in shape and can cover either 10°
 444 or 20° in ϕ . Furthermore, CSCs have the advantage of featuring both precision muon
 445 measurement and muon trigger in a single device.

446

447 The third type of detector used in the CMS muon system are called resistive plate
 448 chambers (RPCs) and can be found in both the barrel and endcap regions. They consist

449 of gaseous parallel-plate detectors capable of providing precise timing information and
 450 adequate spatial resolution. Because of their excellent time resolution, RPCs provide the
 451 capability of tagging the time of an ionizing event between 2 consecutive LHC bunch
 452 crossings (BX) in a much shorter time (~ 1 ns) than the interval between the BXs (25
 453 ns). For this reason, an RPC-based dedicated muon trigger device can be implemented to
 454 unambiguously identify the relevant BX to which a muon track is associated with, despite
 455 the high rate of events and background expected at the LHC.

456 3.3.6 Trigger and Data Acquisition

457 Due to the vast volume of data originating from the proton-proton collisions (deliv-
 458 ered by the LHC at a rate of 40 MHz), a method of eliminating the majority of the un-
 459 interesting/unwanted events was a requisite for the CMS detector design. This event rate
 460 reduction is achieved by the implementation of the so-called trigger system, which man-
 461 ages to select the potentially interesting interactions and reduce the rate from a staggering
 462 40 TBs^{-1} to a manageable value of just a few hundred Hz.

463

464 The CMS trigger system is implemented using a two-stage rate reduction, which com-
 465 bines both a hardware and software phase. The combination of both of these triggers is
 466 designed to reduce the rate by a factor of $\sim 10^6$. The first stage used in the rate reduction
 467 is purely hardware based and it is called the Level 1 (L1) Trigger [19], which consists of
 468 both Field Programmable Gate Array (FPGA) and Application Specific Integrated Circuit
 469 (ASIC) technology. During this initial stage, the rate is reduced to about 100 kHz with
 470 a latency of $3.2 \mu\text{s}$. This time interval constrains the trigger decision, allowing only for
 471 data from the calorimeters and muon system to be processed. Trigger primitives (TP)
 472 from these subsystems are processed through a series of steps before the combined event
 473 information is evaluated by the global trigger (GT) where the final decision, whether or
 474 not to accept the event, is made.

475

Figure 3.8: Transverse slice of the CMS detector, showing the individual detector subsystems and particle signatures in each. The particle type can be inferred by combining the detector response in the different subdetectors [7].

476 The second stage, which implements offline-quality reconstruction algorithms in its
477 event selection, is referred to as the High-Level Trigger (HLT) [20]. The event selection
478 process for the HLT requires that physics objects for each event, such as electrons, muons
479 and jets, are reconstructed and undergo predetermined identification criteria.

480 3.3.7 Event Reconstruction

481 In order to reconstruct the events the particle flow (PF) algorithm [21] is used. This
482 algorithm gathers information from all the CMS sub-detectors to reconstruct charged and
483 neutral hadrons, photons, muons, and electrons. It relies on an efficient and pure track
484 reconstruction, a clustering algorithm able to distinguish overlapping tracks originated
485 from different vertices, and an efficient link procedure to associate each particle deposit
486 in the sub-detectors. Once all the deposits of a particle are associated, it can be correctly
487 identified and its four-momentum optimally determined from the combined information
488 of the sub-detectors. The resulting list of particles are then used to reconstruct higher level
489 objects such as jets, taus, missing transverse energy, and to compute charged lepton and
490 photon isolation, etc [22]. The CMS experiment is provided with millions of collisions
491 per second, which need to be triggered, detected, stored and analyzed in a collaboration of
492 several thousand physicists. This huge amount of data and the complexity of the detector
493 require a flexible data model that serves all the needs of the collaboration. The data for-
494 mat is optimized for performance and flexibility of the reconstruction for the end user's
495 analysis.

496

497 Event information from each step in the simulation and reconstruction chain is logi-
498 cally grouped into what is called a data tier [23]. From the physicist's point of view the
499 most important data tiers are RECO, which contains all reconstructed objects and hits,
500 and AOD (a subset of RECO). The AOD will contain a copy of all the high-level physics

objects (such as muons, electrons, taus, etc.) and enough information about the event to support all the typical usage patterns of a physics analysis. It also contains a summary of the RECO information sufficient to support typical analysis actions such as track refitting with improved alignment or kinematic constraints, re-evaluation of energy and/or position of ECAL clusters based on analysis-specific corrections. The format of each data tier is ROOT [24]. ROOT is a framework for data processing developed at CERN with the sole purpose of aiding high energy physics research. The various AOD datasets are stored worldwide at various data tier centers. From the AOD's the analysis groups create data structures called NTuples containing only the high-level physics objects needed for their particular analysis.

3.3.8 Future Upgrade of Pixel Detector

After the first LHC shutdown called LS1 (2013-2014), and the installation of the Phase-1 Pixel Detector [25] in early 2017, among other things, the LHC is planning another series of upgrades during two major shutdowns, called LS2 and LS3, currently planned for 2019-2020 and \sim 2024, respectively. LS2 would result in a further increase of the luminosity beyond the original design value, to over $2 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. With the LS3 upgrade of the LHC (called HL-LHC [26]) the luminosity is expected to reach up to $7.5 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. Correspondingly, the CMS collaboration has planned a series of further upgrades [27, 28] that will ensure the capabilities of the CMS detector to match to the HL-LHC running conditions, while taking the opportunity to improve the performance and repair any problems uncovered during the data-taking periods. The UPRM group will continue its involvement in the Phase-1 Pixel Detector operations and Phase-2 Pixel Upgrade design.

524

The HL-LHC conditions of instantaneous peak luminosities of up to $7.5 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and an integrated luminosity of the order of 300 fb^{-1} would result in 1 MeV neutron equivalent fluence of $2.3 \cdot 10^{16} \text{ neq/cm}^2$ and a total ionizing dose (TID) of 12MGy (1.2

528 Grad) at the center of CMS, where its innermost component, the Phase-2 Pixel Detector
529 will be installed. The detector should be able to withstand the above radiation dose, han-
530 dle projected hit rates of $3\text{GHz}/\text{cm}^2$ at the lowest radius, be able to separate and identify
531 particles in extremely dense collision debris, deal with a pileup of 140-200 collisions per
532 bunch crossing and have a high impact parameter resolution. This translates into requiring
533 a detector design that is highly granular, has thinner sensors and smaller pixels, and faster,
534 radiation hard electronics compared to its Phase-1 counterpart. The selection of interest-
535 ing physics events at the Level-1 (L1) trigger and inefficiency of selection algorithms in
536 high pileup conditions further require the Tracker to be included in this trigger stage, help-
537 ing reduce the event rate from 40 MHz rate to 7.5 kHz. The physics goals also require
538 an increase in Pixel Detector coverage to $|\eta| = 4.0$ which improves the p_T^{miss} resolution
539 and particle-flow event reconstruction by providing p_T measurements and trajectories for
540 charged particles entering the calorimeters. The p_T^{miss} resolution is an essential perfor-
541 mance parameter for many BSM physics searches including SUSY and extra dimension
542 models where particles escape undetected from the detector space. The smaller pixel size
543 will further improve b-tagging as well as hadronic reconstruction and track reconstruction
544 efficiencies within boosted jets, which can be produced from new heavy objects decaying
545 to Higgs, Z bosons, or top quarks – all heavy probes that can be exploited for new physics
546 searches. Improving the b-tagging capabilities directly affects our analysis due to the im-
547 portance of properly identifying top quark decays.

548

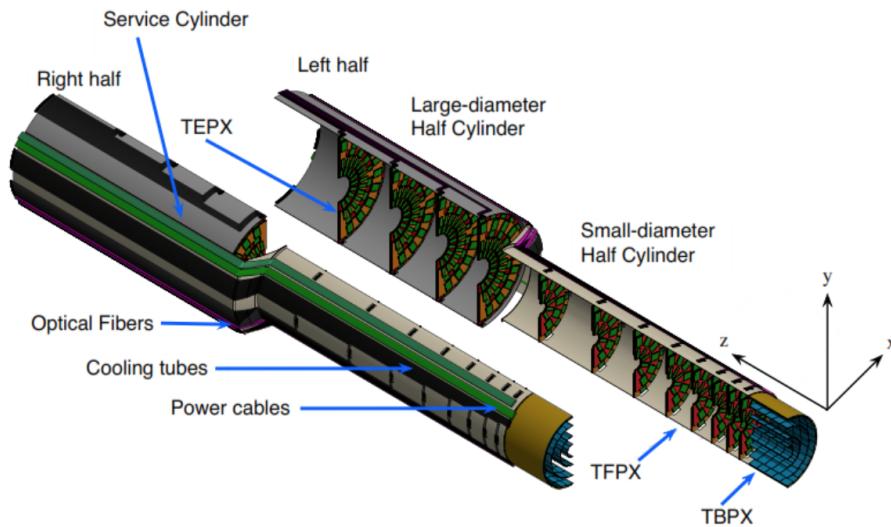


Figure 3.9: Phase-2 Pixel Detector Layout[ref].

549 The Phase-2 Pixel Detector [29, 30] baseline design comprises a barrel part with 4
 550 layers of Tracker Barrel Pixel Detector (TBPX), 8 small double-discs per side of Tracker
 551 Forward Pixel Detector (TFPX) and 4 large double-discs per side of Tracker Endcap Pixel
 552 Detector (TEPX). This forward and end part is referred to as 8l4s (8 TFPX and 4 TEPX).
 553 In the TBPX the pixel modules are arranged in “ladders”. In each layer, neighboring ladders
 554 are mounted staggered in radius, so that $r\text{-}\phi$ overlap between the ladders is achieved.
 555 The modules on a ladder do not overlap in z . A projective gap at $\eta = 0$ is avoided by
 556 mounting an odd number of modules along z , and by splitting the barrel mechanics in z
 557 into slightly asymmetric halves. In TFPX and TEPX the modules are arranged in concen-
 558 tric rings. Each double-disc is physically made of two discs, which facilitates to mount
 559 modules onto four planes, with overlaps in r as well as $r\text{-}\phi$. Each disc is split into two
 560 halves, and these D-shaped structures are referred to as “dees”. The TEPX will provide
 561 the required luminosity measurement capability, by an appropriate implementation of the
 562 readout architecture. In total, the pixel detector will have an active surface of approxi-
 563 mately 4.9 m^2 . **Figure 3.9** shows the layout of the Phase-2 Pixel Detector.

⁵⁶⁴ **Appendix A**

⁵⁶⁵ **Appendix Title**

566 **Appendix B**

567 **References**

- 568 [1] C. Lefevre, “The CERN accelerator complex. Complexe des accelerateurs du
569 CERN.” Available on the [CERN document server](#), Dec 2008.
- 570 [2] T. Sakuma and T. McCauley, “Detector and Event Visualization with SketchUp
571 at the CMS Experiment,” *J. Phys. Conf. Ser.*, vol. 513, p. 022032, 2014.
572 [doi:10.1088/1742-6596/513/2/022032](https://doi.org/10.1088/1742-6596/513/2/022032).
- 573 [3] S. Spannagel, *CMS Pixel Detector Upgrade and Top Quark Pole Mass Determina-*
574 *tion*, vol. 27. Springer, 2017. [doi:10.1007/978-3-319-58880-3](https://doi.org/10.1007/978-3-319-58880-3).
- 575 [4] CMS Collaboration, “CMS Physics Technical Design Report Volume 1: Detec-
576 tor Performance and Software,” Technical Design Report CERN-LHCC-2006-001,
577 CERN (2006).
- 578 [5] CMS Collaboration, “The CMS Experiment at the CERN LHC,” *JINST*, vol. 3,
579 p. S08004, 2008. [doi:10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- 580 [6] CMS Collaboration, “Performance of the CMS Hadron Calorimeter with Cos-
581 mic Ray Muons and LHC Beam Data,” *JINST*, vol. 5, p. T03012, 2010.
582 [doi:10.1088/1748-0221/5/03/T03012](https://doi.org/10.1088/1748-0221/5/03/T03012).
- 583 [7] J. Nielsen, “Fundamentals of LHC Experiments,” in *Proceedings, Theoretical Ad-*
584 *vanced Study Institute in Elementary Particle Physics (TASI 2010). String Theory*

- 585 *and Its Applications: From meV to the Planck Scale: Boulder, Colorado, USA, June*
586 *1-25, 2010*, pp. 127–152, 2011. [arXiv:1106.2516](https://arxiv.org/abs/1106.2516).
- 587 [8] CERN, “Processing what to record?,” 2018.
- 588 [9] CMS, “The CMS Computing Project,” tech. rep., CERN, 2005.
- 589 [10] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Col-
590 lider,” *JINST*, vol. 3, p. S08003, 2008. [doi:10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- 591 [11] LHCb Collaboration, “The LHCb Detector at the LHC,” *JINST*, vol. 3, p. S08005,
592 2008. [doi:10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005).
- 593 [12] K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC,” *JINST*, vol. 3,
594 p. S08002, 2008. [doi:10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002).
- 595 [13] K. Garrett and G. Duda, “Dark Matter: A Primer,” *Adv. Astron.*, vol. 2011,
596 p. 968283, 2011. [doi:10.1155/2011/968283](https://doi.org/10.1155/2011/968283).
- 597 [14] R. C. G. Landim, *Dark Sector Cosmology*. PhD thesis, Sao Paulo U., 2017.
598 [arXiv:1702.06024](https://arxiv.org/abs/1702.06024).
- 599 [15] K. W. Bell *et al.*, “Vacuum phototriodes for the CMS electromagnetic
600 calorimeter endcap,” *IEEE Trans. Nucl. Sci.*, vol. 51, pp. 2284–2287, 2004.
601 [doi:10.1109/TNS.2004.836053](https://doi.org/10.1109/TNS.2004.836053).
- 602 [16] H. Bichsel, D. E. Groom, and S. R. Klein, “Passage of particles through matter,”
603 2004.
- 604 [17] CMS Collaboration, “Alignment of the CMS Muon System with Cosmic-Ray
605 and Beam-Halo Muons,” *JINST*, vol. 5, p. T03020, 2010. [doi:10.1088/1748-0221/5/03/T03020](https://doi.org/10.1088/1748-0221/5/03/T03020).
- 607 [18] “Muon Detectors.” <http://cms.web.cern.ch/news/muon-detectors>.
608 Accessed: 2018-04-08.

- 609 [19] A. Tapper and D. Acosta, “CMS Technical Design Report for the Level-1 Trigger
610 Upgrade,” 2013.
- 611 [20] W. Adam *et al.*, “The CMS high level trigger,” *Eur. Phys. J.*, vol. C46, pp. 605–667,
612 2006. [doi:10.1140/epjc/s2006-02495-8](https://doi.org/10.1140/epjc/s2006-02495-8).
- 613 [21] F. Beaudette, *The CMS Particle Flow Algorithm*, 2013.
- 614 [22] A. K. Nayak, “Reconstruction of physics objects in the CMS detector,” *PoS*,
615 vol. CHARGED2012, p. 010, 2012.
- 616 [23] CMS Collaboration, “CMS: The computing project. Technical design report,” Technical
617 Design Report CERN-LHCC-2005-023, CERN (2005).
- 618 [24] R. Brun and F. Rademakers, “ROOT: An object oriented data analysis frame-
619 work,” *Nucl. Instrum. Meth.*, vol. A389, pp. 81–86, 1997. [doi:10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X).
- 621 [25] M. Lipinski, “The Phase-1 upgrade of the CMS pixel detector,” *JINST*, vol. 12,
622 no. 07, p. C07009, 2017. [doi:10.1088/1748-0221/12/07/C07009](https://doi.org/10.1088/1748-0221/12/07/C07009).
- 623 [26] G. Apollinari, I. Bejar Alonso, O. Bruning, M. Lamont, and L. Rossi, “High-
624 Luminosity Large Hadron Collider (HL-LHC) : Preliminary Design Report,” 2015.
625 [doi:10.5170/CERN-2015-005](https://doi.org/10.5170/CERN-2015-005).
- 626 [27] D. Contardo, M. Klute, J. Mans, L. Silvestris, and J. Butler, “Technical Proposal for
627 the Phase-II Upgrade of the CMS Detector,” 2015. [CMS-TDR-15-02](https://cds.cern.ch/record/1590004).
- 628 [28] J. Pilot, “Future of the CMS Muon System: Upgrades and Aging,” *PoS*,
629 vol. ICHEP2016, p. 281, 2016. [doi:10.22323/1.282.0281](https://doi.org/10.22323/1.282.0281).
- 630 [29] K. Klein, “The Phase-2 Upgrade of the CMS Tracker,” 2017. [CMS-TDR-014](https://cds.cern.ch/record/2200000).
- 631 [30] S. Mersi, “Phase-2 Upgrade of the CMS Tracker,” *Nucl. Part. Phys. Proc.*, vol. 273-
632 275, pp. 1034–1041, 2016. [doi:10.1016/j.nuclphysbps.2015.09.162](https://doi.org/10.1016/j.nuclphysbps.2015.09.162).