



VRIJE  
UNIVERSITEIT  
BRUSSEL

Vrije Universiteit Brussel  
Faculteit Wetenschappen en  
Bio-ingenieurswetenschappen  
Departement Natuurkunde

1

2

3

4

5

6

7

8

9

10

11

12

---

# Search for Dark Matter in the Monojet and Trackless Jets Final States with the CMS Detector at the LHC

---

Isabelle De Bruyn

Promotor

Prof. Dr. Steven Lowette

Proefschrift ingediend met het oog op het behalen van de  
academische graad van Doctor in de Wetenschappen

September 2017



## Acknowledgements



# Table of Contents

1

2	<b>Acknowledgements</b>	i
3	<b>1 Introduction</b>	1
4	<b>2 Dark Matter Scenarios Beyond the Standard Model</b>	3
5	2.1 The Standard Model of Particle Physics . . . . .	3
6	2.1.1 Elementary particles and their interactions . . . . .	3
7	2.1.2 Unanswered questions of the Standard Model . . . . .	3
8	2.2 Dark matter . . . . .	3
9	2.2.1 Observational evidence . . . . .	3
10	2.2.2 Detection of dark matter . . . . .	3
11	2.2.3 Dark matter models . . . . .	3
12	2.3 Strongly Interacting Massive Particles . . . . .	3
13	2.3.1 Motivation . . . . .	3
14	2.3.2 Experimental constraints . . . . .	3
15	<b>3 The LHC and the CMS Detector</b>	5
16	3.1 The Large Hadron Collider at CERN . . . . .	5
17	3.1.1 The LHC injector chain . . . . .	5
18	3.1.2 The Large Hadron Collider . . . . .	6
19	3.1.3 The experiments at the LHC . . . . .	7
20	3.2 The CMS detector . . . . .	8
21	3.2.1 The tracker . . . . .	9
22	3.2.1.1 The pixel tracker . . . . .	9
23	3.2.1.2 The strip tracker . . . . .	10
24	3.2.2 The electromagnetic calorimeter . . . . .	10
25	3.2.3 The hadronic calorimeter . . . . .	11
26	3.2.4 The muon system . . . . .	13
27	3.2.5 Trigger and data acquisition . . . . .	13
28	3.2.6 CMS performance in Run 2 . . . . .	15
29	3.2.6.1 Pre-amplifier saturation in the APV25 chip . . . . .	15
30	<b>4 Event Simulation and Reconstruction</b>	19
31	4.1 Event generation . . . . .	19
32	4.2 Detector simulation . . . . .	20
33	4.3 Event reconstruction . . . . .	21
34	4.3.1 Track reconstruction . . . . .	21
35	4.3.2 Electron and isolated photon reconstruction . . . . .	22
36	4.3.3 Muon reconstruction . . . . .	23
37	4.3.4 Particle flow . . . . .	23
38	4.3.5 Jet reconstruction . . . . .	24
39	4.3.6 Missing transverse energy reconstruction . . . . .	25
40	4.4 Simulation of the SIMP signal . . . . .	26

1	<b>5 The Monojet Analysis</b>	<b>29</b>
2	5.1 Introduction . . . . .	29
3	5.2 Event selection . . . . .	29
4	5.3 Background estimation . . . . .	29
5	5.4 Results . . . . .	29
6	5.5 Improvement going from the 2015 to 2016 analysis . . . . .	29
7	5.6 Interpretation . . . . .	29
8	<b>6 Search for SIMPs using Trackless Jets</b>	<b>31</b>
9	6.1 Introduction . . . . .	31
10	6.2 Event selection . . . . .	31
11	6.3 Background estimation . . . . .	31
12	6.4 Results . . . . .	31
13	6.5 SIMP model interpretation . . . . .	31
14	<b>7 Conclusion &amp; Outlook</b>	<b>33</b>

# 1

## Introduction

1

2



# 2

1

2

3

## Dark Matter Scenarios Beyond the Standard Model

### 2.1 The Standard Model of Particle Physics

#### 2.1.1 Elementary particles and their interactions

#### 2.1.2 Unanswered questions of the Standard Model

### 2.2 Dark matter

#### 2.2.1 Observational evidence

#### 2.2.2 Detection of dark matter

#### 2.2.3 Dark matter models

### 2.3 Strongly Interacting Massive Particles

#### 2.3.1 Motivation

#### 2.3.2 Experimental constraints



# 3

1

2

## The LHC and the CMS Detector

3 In order to investigate the currently unsolved mysteries of particle physics, such as the existence of dark  
4 matter, many experiments can be conducted, among other things at particle colliders. The largest particle  
5 accelerator in the world is the Large Hadron Collider (LHC), located at the European Organization for  
6 Nuclear Research (CERN) in Geneva, Switzerland. At this accelerator, protons are being accelerated at  
7 energies up to 6.5 TeV, giving rise to a record center-of-mass energy of 13 TeV in the proton collisions.  
8 Using data from the collisions generated at the interaction points along the accelerator ring, the Stan-  
9 dard Model can be tested in many ways and searches for particles beyond the Standard Model can be  
10 performed.

11 In Section 3.1 more details are given about the LHC and the 4 main experiments situated at the inter-  
12 action points. In particular, the general-purpose Compact Muon Solenoid (CMS) detector is described in  
13 Section 3.2.

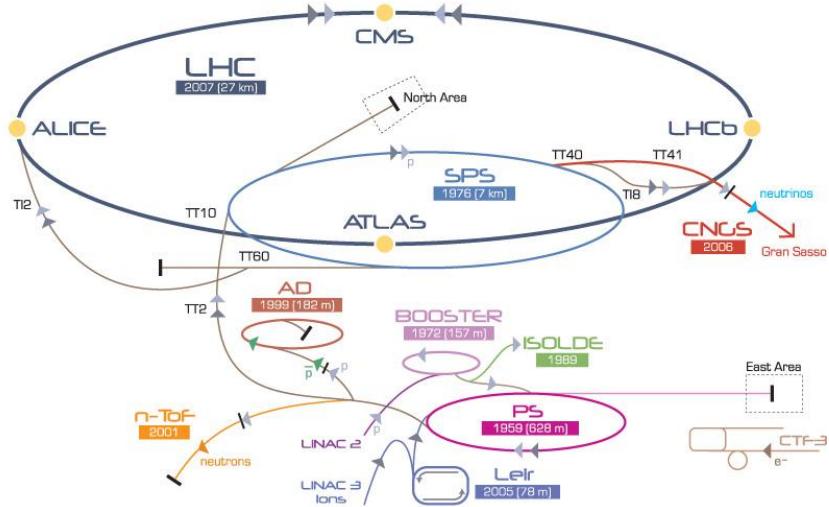
### 14 3.1 The Large Hadron Collider at CERN

15 The LHC was built in the already existing Large Electron Positron (LEP) collider tunnel, which was  
16 excavated in the 1980's and has a circumference of 27.6 km. Contrary to the LEP collider, the LHC  
17 accelerates particles of the same charge, namely protons or lead ions. Much higher luminosities can  
18 therefore be reached, since only particles are used and the generation of anti-particles is not needed. This  
19 was the limiting factor at the Tevatron, where protons and anti-protons were used. Additionally, at the  
20 probed energies the colliding particles are not the protons or ions, but their constituents, which carry  
21 a varying fraction of the total momentum. This makes the LHC an ideal instrument for exploration at  
22 higher energies, as the collisions naturally cover a wide energy range.

#### 23 3.1.1 The LHC injector chain

24 The protons (or lead ions) can not directly be injected in the LHC, but need to be accelerated gradually  
25 in several pre-accelerators, as illustrated in Figure 3.1. For the proton beams, the LHC injection chain  
26 starts at a bottle of hydrogen, where protons are stripped from the hydrogen atoms and accelerated up  
27 to 50 MeV by a linear accelerator (LINAC2). The protons are then transferred to a chain of circular  
28 accelerators, starting with the Proton Synchrotron Booster (PSB) which accelerates them to an energy of  
29 1.4 GeV. Next, the protons go through the Proton Synchrotron (PS) and are delivered to the Super Proton  
30 Synchrotron (SPS) at an energy of 26 GeV. Finally, the protons are injected in the LHC in opposite  
31 direction with an energy of 450 GeV.

The lead ions are first accelerated in a different linear accelerator, LINAC3, before being injected in the Low Energy Ion Ring (LEIR) at an energy of 4.5 MeV per nucleon. Here the ions are accelerated to an energy of 72 MeV per nucleon, and they then follow the same path as the protons through the PS, where they are accelerated to 5.9 GeV and stripped from the last of their electrons, and the SPS, where they are accelerated to 177 GeV. The record center-of-mass energy for heavy ion collisions at the LHC so far has been 5.02 TeV and 8.16 TeV, for lead-lead (Pb-Pb) and proton-lead (p-Pb) collisions in 2015 and 2016, respectively.



*Figure 3.1: Schematic view of the various linear and circular accelerators of the CERN accelerator complex, including the LHC injection chain.*

### 3.1.2 The Large Hadron Collider

The most relevant specifications for a particle physics accelerator are the maximum energy and the luminosity that can be reached. High energy is necessary in order to be able to create new heavy particles, which are for example predicted in many theories beyond the Standard Model.

The protons are kept on the correct orbit by the 1232 LHC dipole magnets. These magnets are cooled down to 1.9 K with liquid Helium and supplied with a current of 12 kA to reach the design field of 8.33 T. This limits the maximum beam momentum of the accelerator to

$$p = B/\rho = 8.33 \text{ T}/2804 \text{ m} = 7 \text{ TeV}/c, \quad (3.1)$$

with  $\rho$  the bending radius of the tunnel. The protons are accelerated up to the desired energy by radio-frequency (RF) cavities, which produce an oscillating electric field.

A high event rate or luminosity  $\mathcal{L}$  is equally important, to obtain a sufficiently high number of collisions. For a process with cross section  $\sigma$ , this rate is

$$\frac{dN}{dt} = \mathcal{L}\sigma. \quad (3.2)$$

In order to achieve the high design luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$  in the LHC, the protons are concentrated in bunches with a 25 ns spacing, which are focused by strong quadrupole magnets around the interaction regions. Additionally, 225 ns gaps are present as well between the bunch trains, corresponding to the rise time of the injection kicker magnets. One gap of 3  $\mu\text{s}$  is necessary as well to allow clean beam dumps. These requirements limit the number of bunches to a maximum of 2808.

After almost 25 years of design and construction, the LHC was completed in 2008 and the commissioning of the machine started. However, during a powering test on 19 September of the same year an

1 electric arc developed inside a bus bar which led to a large release of helium and a pressure wave that  
 2 caused extensive mechanical damage to the affected LHC sector. This incident delayed the first collisions,  
 3 with one bunch per beam and at a beam energy of 900 GeV, until late 2009. During 2010 and  
 4 2011 a center-of-mass energy of 7 TeV was used for the collisions, which was then increased to 8 TeV in  
 5 2012. The instantaneous luminosity was also increased, starting from  $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$  in 2010 to more  
 6 than  $6 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  in 2012. During the 3 years of data-taking in Run 1, data corresponding to an  
 7 integrated luminosity of  $45.0 \text{ pb}^{-1}$ ,  $6.1 \text{ fb}^{-1}$ , and  $23.3 \text{ fb}^{-1}$ , respectively, were delivered. After Run 1, a  
 8 long shutdown (LS1) of 2 years followed, which was used to correct the problems that were discovered  
 9 in the aftermath of the incident at the startup in 2008, and to upgrade and consolidate the experiments  
 10 located on the LHC ring.

11 In 2015, the LHC restarted operations with Run 2, at an even higher center-of-mass energy of 13 TeV.  
 12 During 2016 the design luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$  was exceeded and a total of  $41 \text{ fb}^{-1}$  of data were  
 13 delivered. A comparison of the delivered integrated luminosity per year is shown in Figure 3.2.

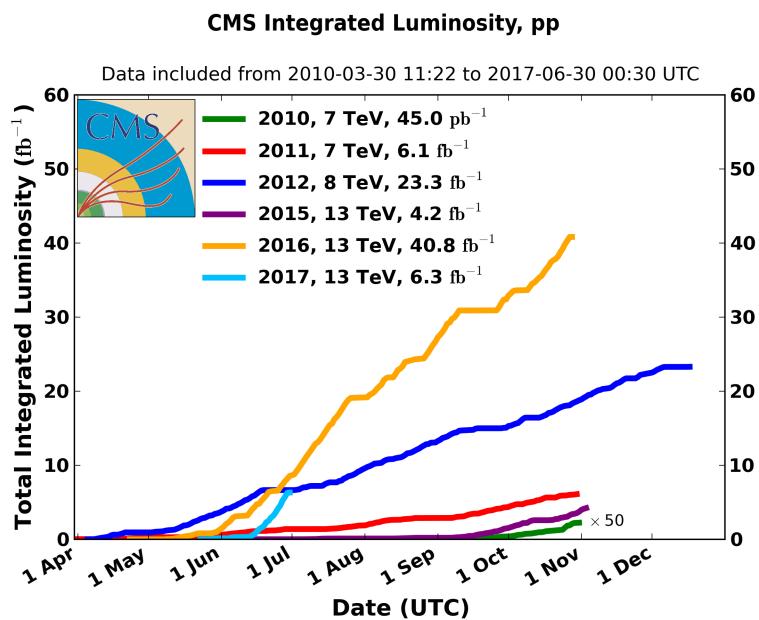


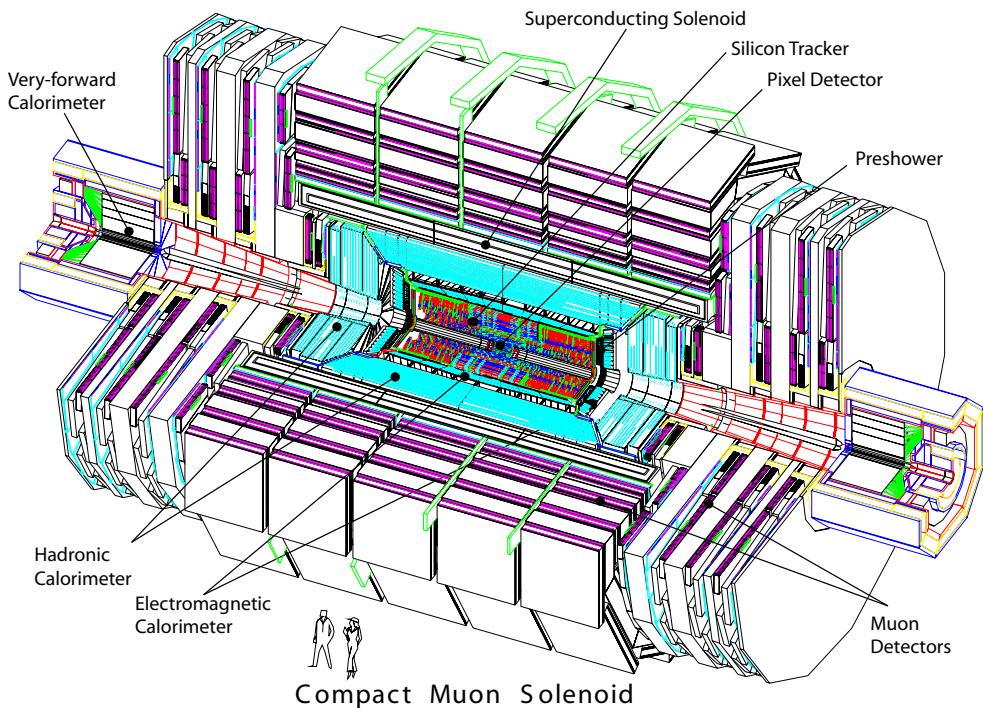
Figure 3.2: Overview of the integrated luminosity delivered to the CMS detector during Run 1 (2010 to 2012) and Run 2 (2015 to 2017).

### 3.1.3 The experiments at the LHC

15 There are four interaction points (IPs) where the proton or lead ion beams of the LHC can collide, and  
 16 around each of these points large particle detectors were built in order to record the generated colli-  
 17 sions. The ATLAS and CMS detectors, located at IP1 and IP5, are both high luminosity general-purpose  
 18 detectors and consist of several layers surrounding the IP in an onion-like structure to avoid particles  
 19 escaping detection. These detectors can cover a wide range of high energy physics, from precision mea-  
 20 surements of the Standard Model to searches beyond the Standard Model. At IP2 the ALICE detector  
 21 is specialized in heavy ion collisions with low instantaneous luminosities, around  $10^{27} \text{ cm}^{-2}\text{s}^{-1}$ . With  
 22 this detector information is gathered about the quark-gluon plasma, a state of matter that exists at ex-  
 23 tremely high temperatures and densities where quarks and gluons are no longer confined in hadrons. The  
 24 fourth main detector, LHCb, is located at IP8 and requires instantaneous luminosities of the order of a  
 25 few  $10^{32} \text{ cm}^{-2}\text{s}^{-1}$ . Using this detector  $b$  quarks are being studied, focusing among other things on the  
 26 matter-antimatter asymmetry in the universe.

## 3.2 The CMS detector

The searches described in this thesis were conducted using data collected with the CMS detector, a general-purpose detector located on the LHC ring. It consists of the typical components of a particle physics detector, namely a tracker, an electromagnetic calorimeter (ECAL), a hadronic calorimeter (HCAL), a solenoidal magnet, and muon detectors. One peculiar aspect is however that both calorimeters are situated inside the superconducting magnet. This design was chosen in order to improve the energy resolution by reducing the amount of material in front of the calorimeters. The overall detector has a length of 21.6 m, a diameter of 14.6 m and a total weight of 12500 t. A schematic overview is shown in Figure 3.3.



*Figure 3.3: The CMS detector, consisting of the pixel and strip tracker, the electromagnetic calorimeter (ECAL) with preshower, the hadronic calorimeter (HCAL) with its forward component, and the muon systems.*

The CMS coordinate system places the origin at the nominal collision point. The  $x$  axis is perpendicular to the beam and points towards the center of the LHC ring, the  $y$  axis is vertical and pointing upwards, and the  $z$  axis is defined anti-clockwise along the beam direction. The azimuthal angle  $\phi$  is then defined in the  $xy$  plane, relative the the  $x$  axis and the polar angle  $\theta$  is measured with respect to the  $z$  axis. In general, the polar angle is converted into the pseudorapidity

$$\eta = -\ln \left( \tan \frac{\theta}{2} \right) \quad (3.3)$$

for convenience, since differences in pseudorapidity are invariant under Lorentz boosts along the  $z$  axis. A pseudorapidity of 0 corresponds to the direction perpendicular to the beam ( $\theta = \pi/2$ ), and an infinite pseudorapidity corresponds to the direction parallel to the beam ( $\theta = 0$ ).

Due to the conservation of momentum before and after the collision, the momenta of the particles in the final state of a collision should be balanced in the transverse plane. Another variable that is therefore often used in particle physics is the transverse momentum of a particle, defined as

$$p_T = p \cdot \sin \theta. \quad (3.4)$$

### 3.2.1 The tracker

The innermost part of the CMS detector, closest to the IP, is the tracking system, which is designed to provide a precise measurement of the trajectories of charged particles. This all-silicon detector is divided into a pixel and a strip detector, with a layout as shown in Figure 3.4. The inner part, consisting of pixel modules, provides very precise 3D hits, which are important for vertex reconstruction and track seeding. This allows to have a precise measurement of secondary vertices and track impact parameters, necessary for the efficient identification of e.g. heavy flavor particles. As the hit occupancy is lower in the outer part of the detector, a larger cell size can be afforded, and silicon strips are used instead of pixels. This strip detector provides a large lever arm and a link to the calorimeters and the muon system. The tracker covers a pseudorapidity range  $|\eta| < 2.5$ .

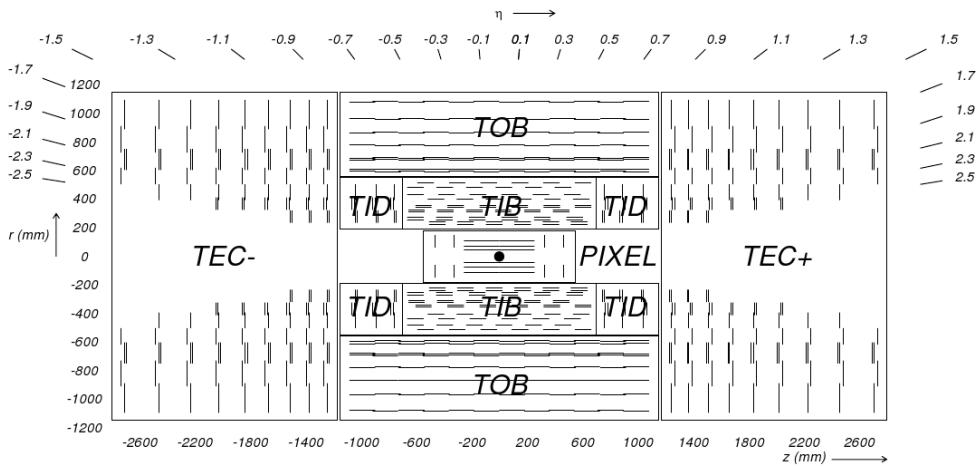


Figure 3.4: A transverse view of the pixel and strip tracker detectors.

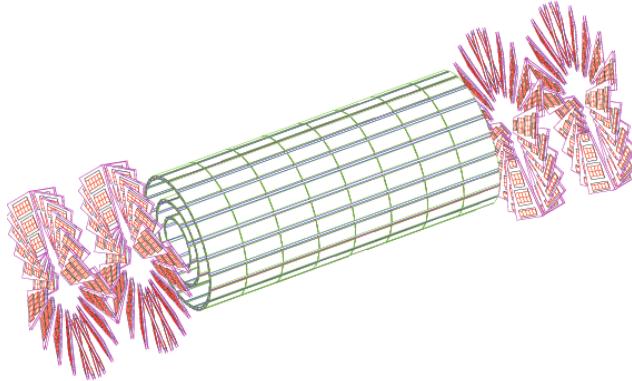
#### 3.2.1.1 The pixel tracker

The pixel tracker was replaced during the extended technical stop in 2016 and 2017 [1], as a part of the CMS Phase 1 upgrades. As the data used for this thesis were recorded before that, only the so-called Phase 0 detector is described here.

For the pixel modules n+ pixels on n-substrate are used, allowing the sensors to also work in under-depletion after type inversion. The 1440 modules are arranged in several cylindrical layers and disks, as illustrated in Figure 3.4. The barrel, consisting of 3 pixel layers surrounding the beam pipe at radii of 4.4, 7.3 and 10.2 cm, is complemented by the forward pixel detector, composed of 2 endcap disks on each side extending from 6 to 15 cm in radius. The barrel and the forward parts contain respectively 48 million and 18 million pixels with a size of  $100 \times 150 \mu\text{m}^2$ , covering a total area of  $1.06 \text{ m}^2$ .

In the barrel, the magnetic field of CMS is perpendicular to the drift of the electrons to the collecting pixels, which results in a Lorentz drift. This drift leads to a spread of the charge over several pixels. Since the read-out of the modules is analog, an improved spatial resolution can therefore be achieved with charge interpolation. In the forward pixel detector the drift of the electrons would be parallel to the magnetic field so in order to profit from the Lorentz angle, the modules are tilted by  $20^\circ$  in a turbine-like arrangement, as can be seen in Figure 3.5. A spatial resolution of  $10 \mu\text{m}$  ( $30 \mu\text{m}$ ) can be achieved in the local directions  $x$  ( $y$ ) of the module, respectively. In the barrel  $x$  is the longitudinal direction perpendicular to the beam and  $y$  is the longitudinal direction parallel to the beam.

The signals from the pixel sensors are read out by custom read-out chips (ROCs), which amplify and store the signals, and already apply zero-suppression on-detector. The data rate from the detector to the Front End Drivers (FEDs) is therefore not constant for every event. Additionally, if there are too many hits on a pixel module for a given event, they can not all be stored on the finite buffer of the ROC. Consequently, as the instantaneous luminosity increases the pixel modules start to show a “dynamic



*Figure 3.5: A 3D view of the barrel and forward pixel detector.*

inefficiency” which is most pronounced in the first layer, closest to the beampipe. This was one of the main motivations for the Phase 1 upgrade of the pixel detector.

### 3.2.1.2 The strip tracker

The outer part of the tracker consists of 15 148 strip modules, which are distributed among multiple barrel layers and endcap disks and make up a total active area of  $198 \text{ m}^2$ . The inner part is composed of 4 Tracker Inner Barrel (TIB) layers with 3 Tracker Inner Disks (TID) on each side. Surrounding these are 6 Tracker Outer Barrel (TOB) layers and the 2 Tracker EndCaps (TEC), which are composed of 9 disks. This geometric arrangement is shown in Figure 3.4, with double lines to indicate back-to-back modules. These so-called double-sided modules are mounted with a stereo angle of 100 mrad to improve the 3D point resolution by providing a measurement of the  $z$  and  $r$  coordinate in the barrel and disks, respectively. The choice of strip pitches is driven by the two particle separation capability and two-hit resolution, and ranges from  $80 \mu\text{m}$  to  $205 \mu\text{m}$ . The length of the strips varies from 63 mm to 117 mm, minimizing the occupancy and noise levels.

In the TOB and the 3 outermost rings of the TEC two silicon sensors are daisy chained, while single sensors are used in the inner part. This is done to limit the number of read-out channels, since the area that had to be instrumented is larger in the outer region. The larger cell size can be afforded due to the lower occupancy in the outer part. However, the noise of the sensors also increases with strip length, so thicker silicon sensors,  $500 \mu\text{m}$  compared to  $320 \mu\text{m}$  in the inner part, are used in order to collect more signal per traversing particle.

The strip sensors are single sided p-on-n type silicon. The signals from the sensors are amplified, shaped, and stored by 4 or 6 custom APV25 chips per module. When the trigger has made a positive decision, the analog signals from two APV25 chips are multiplexed and sent to the FED boards in the service cavern via optical fibers, where they are converted to digital signals. The FEDs then perform pedestal and common mode subtraction as well as cluster finding. Additionally, the data is sparsified in these off-detector electronics, before being sent to the CMS central data acquisition (DAQ). Due to charge sharing, this analog read-out scheme also results in an improved spatial resolution of 15 to  $40 \mu\text{m}$ , depending on the position of the modules and the strip pitch.

### 3.2.2 The electromagnetic calorimeter

Surrounding the tracker, the CMS electromagnetic calorimeter (ECAL) is designed to measure the energy of photons and electrons. It is composed of 75 848 lead tungstate ( $\text{PbWO}_4$ ) crystals arranged in a cylindrical barrel and 2 endcaps. The barrel crystals measure  $22 \times 22 \text{ mm}^2$  at the front face of crystal, and  $26 \times 26 \text{ mm}^2$  at the rear face, which corresponds to approximately  $0.0174 \times 0.0174$  in  $\eta\text{-}\phi$ . The length of the crystal is 230 mm, corresponding to 25.8 radiation lengths. In the endcaps, the crystals have a rear face cross section of  $30 \times 30 \text{ mm}^2$ , front face cross section of  $28.62 \times 28.62 \text{ mm}^2$ , and a length of 220 mm, corresponding to 24.7 radiation lengths.

1 The high density material was chosen due to its short radiation length and small Moli  re radius,  
 2 resulting in a small spread of the electromagnetic shower generated by an incoming photon or electron.  
 3 This allows for a fine granularity, a better shower separation, and a compact calorimeter. Additionally,  
 4 this scintillating material has a fast response, as about 80% of the light is emitted during the first 25 ns.  
 5 The scintillation light is collected by photodetectors, digitized, and read out.

6 The layout of the ECAL is shown in Figure 3.6, with the barrel (EB) extending up to  $|\eta| < 1.470$   
 7 and the endcaps (EE) on each side covering the range  $1.479 < |\eta| < 3.0$ . A preshower detector (ES)  
 8 is positioned in front of the endcap crystals, covering the pseudorapidity range between  $|\eta| = 1.653$   
 9 and  $|\eta| = 2.6$ . This detector consists of a layer of lead which initiates an electromagnetic shower from  
 10 incoming photons or electrons, and a layer of silicon sensors which measures the deposited energy. The  
 11 main goal of this 20 cm thick detector is to discriminate between photons and neutral pions.

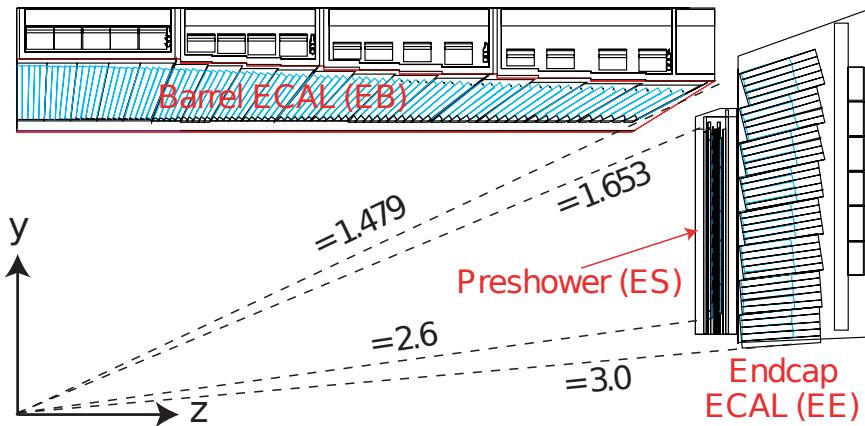


Figure 3.6: A transverse view parallel to the beamline showing one quarter of the ECAL, with its barrel (EB), endcap (EE), and preshower (ES) detectors.

12 The energy resolution of calorimeters can be parametrized by the following stochastic ( $S$ ), noise ( $N$ ),  
 13 and constant ( $C$ ) terms:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (3.5)$$

14 The stochastic term represents contributions from the shower containment, the number of photoelectrons  
 15 and the fluctuations in the gain process. The noise term takes into account all noise components, such as  
 16 electronics and digitization noise. Finally, the constant term characterizes among others energy leakage  
 17 from the back of the calorimeter crystals and non-uniformities of the longitudinal light collection. The  
 18 latter term dominates the energy resolution for high-energy electron and photon showers. Figure 3.7  
 19 shows the energy dependence of this resolution for incident electrons as measured in a beam test, as well  
 20 as the determined stochastic, noise, and constant terms obtained by fitting equation 3.5 to the data.

21 A more recent measurement of the energy resolution was performed using electrons from Z boson  
 22 decays in collision data. In the central region, up to  $|\eta| < 0.8$ , it was measured to be better than 2%.  
 23 Outside of this region, in the more forward direction, the energy resolution is 2-5% [2]. The reconstruction  
 24 of the electrons and photons will be discussed in Section 4.3.2.

### 25 3.2.3 The hadronic calorimeter

26 The hadronic calorimeter (HCAL) surrounds the ECAL with the aim to measure the energy of charged  
 27 and neutral hadrons. The missing transverse energy can then be inferred from this measurement together  
 28 with the measured energy in the ECAL, in order to identify neutrinos or exotic particles. The HCAL  
 29 consists of brass absorber plates interleaved with plastic scintillator tiles.

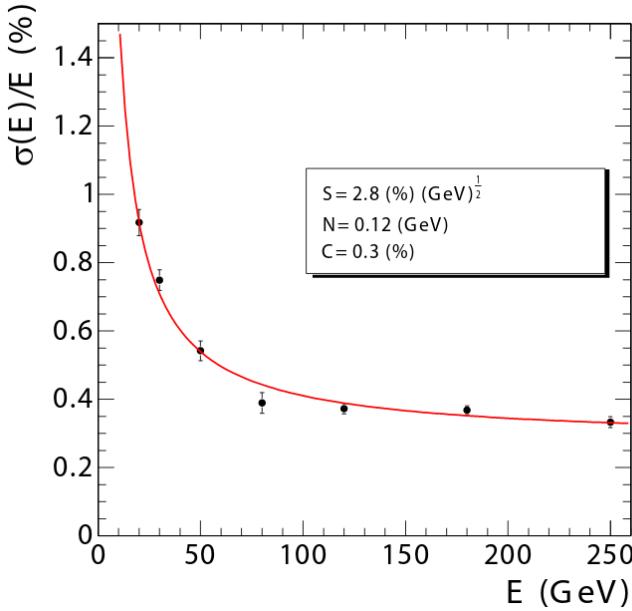


Figure 3.7: The ECAL energy resolution as a function of the electron energy, measured from a beam test. The stochastic ( $S$ ), noise ( $N$ ), and constant ( $C$ ) are given as well.

Figure 3.8 shows a longitudinal quarter view of the different HCAL components. A cylindrical barrel (HB) covers the region up to  $|\eta| < 1.4$  and is complemented by endcaps (HE) on each side, extending the pseudorapidity range to  $|\eta| < 3.0$ . In the central region, the stopping power of the ECAL and HCAL barrel is not sufficient to contain the entire hadron showers. The HCAL was therefore extended outside the solenoid with an outer calorimeter (HO), which uses the magnet coil as absorber and consists of scintillators. Two layers are positioned at  $\eta = 0$ , where the absorber depth is minimal, and only 1 layer is used for the 2 rings on each side of the central ring. Finally, a forward calorimeter (HF) is positioned at 11.2 m from the IP covering  $3.0 < |\eta| < 5.2$ . Unlike the other HCAL components, this detector consists of iron and quartz fibers. Cherenkov-based, radiation-hard technology, since it is exposed to very large particle fluxes.

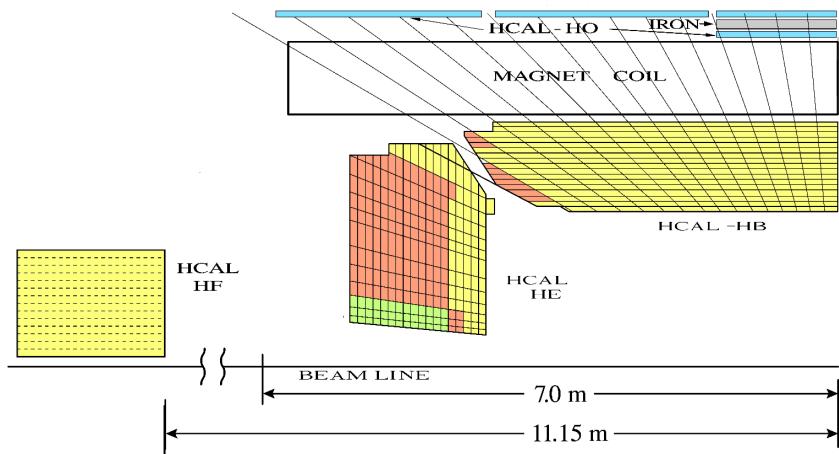


Figure 3.8: A quarter view of the hadronic calorimeter (HCAL), parallel to the beamline. The barrel (HB), endcap (HE), outer (HO), and forward (HF) detectors are indicated.

The optical signals from the scintillators in the HB and HE are converted to electrical signals by multichannel hybrid photodiodes, while silicon photomultipliers (SiPMs) are used in the HO. In the HF,

the Cherenkov light emitted in the quartz fibers is detected by standard photomultiplier tubes (PMTs), since the magnetic field is much smaller in this region.

The expected transverse energy resolution for jets is shown in Figure 3.9 for various pseudorapidity regions: barrel jets ( $|\eta| < 1.4$ ), endcap jets ( $1.4 < |\eta| < 3.0$ ), and very forward jets ( $3.0 < |\eta| < 5.0$ ). Details about the reconstruction of jets from calorimeter and tracking information will be given in Section 4.3.5.

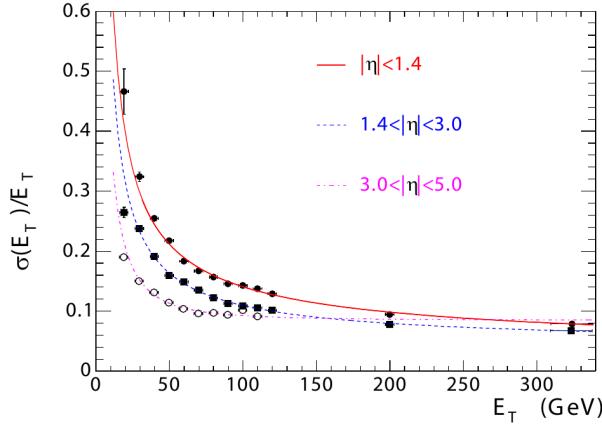


Figure 3.9: The jet transverse energy resolution as a function of the jet transverse energy, for barrel jets ( $|\eta| < 1.4$ ), endcap jets ( $1.4 < |\eta| < 3.0$ ), and very forward jets ( $3.0 < |\eta| < 5.0$ ).

### 3.2.4 The muon system

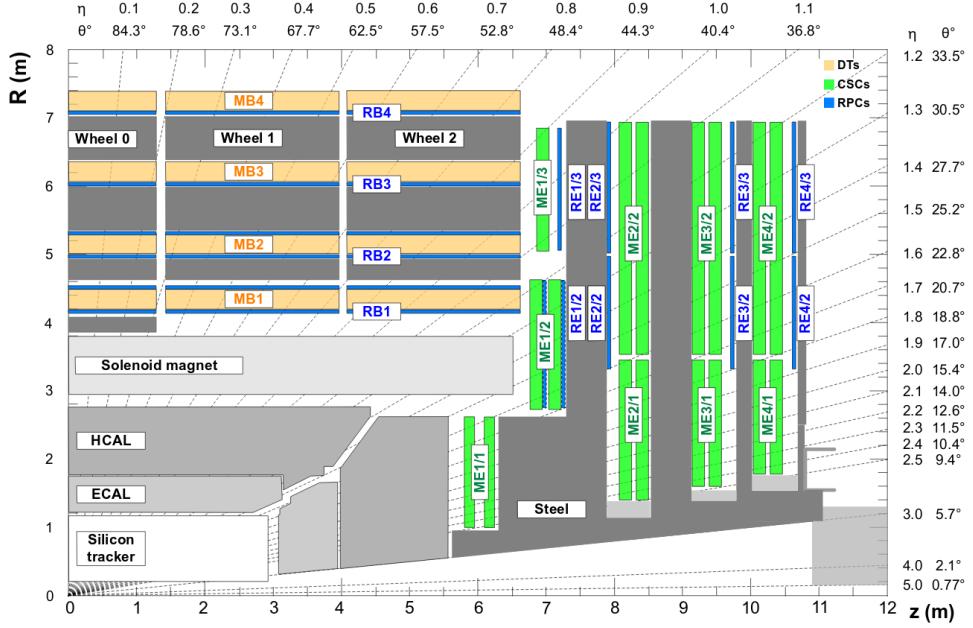
The outermost detector, located entirely on the outside of the solenoid, is a dedicated muon detection system. The purpose of this subsystem is muon identification, momentum measurement, and triggering. As illustrated in Figure 3.10, the layers of muon chambers are embedded in the iron yoke constraining the magnetic field lines. The strong magnetic field completely saturates the return yoke with a field of about 2 T, in opposite direction with respect to the field inside the magnet.

Three different types of gaseous detectors are used. In the barrel, 4 layers of Drift Tubes (DT) are installed, covering the pseudorapidity range up to  $|\eta| < 1.2$ . Due to the higher flux and the larger and non-uniform magnetic field at larger pseudorapidities, Cathode Strip Chambers (CSC) are used in the endcap region ( $0.9 < |\eta| < 2.4$ ). The DTs are designed for the low muon rates that are expected in the barrel and thus have a slower response time than the CSCs. Resistive Plate Chambers (RPCs) complement the DT and CSC systems in the pseudorapidity region up to  $|\eta| < 1.8$ . They provide a fast response, with a good time resolution but a worse spatial resolution than the DTs or CSCs. The RPCs are therefore very well suited to trigger on muons.

The offline reconstruction efficiency of simulated events containing one muon is typically between 95% and 99%, except for the regions between 2 DT wheels ( $|\eta| = 0.25$  and  $|\eta| = 0.8$ ) and the transition region between the DTs and CSCs ( $|\eta| = 1.2$ ), where the efficiency drops to 92%. The reconstruction of muons using the information from the tracker and the muon detectors will be detailed in Section 4.3.3. For low pseudorapidities and small momenta, the offline momentum resolution of the standalone muon system is about 9%. At momenta around 1 TeV, the resolution varies from 15% to 40%, depending on the pseudorapidity. As demonstrated in Figure 3.11, performing a global momentum fit using the tracker as well improves the resolution by an order of magnitude at low muon momenta. At high momenta the resolution of the full system is about 5%.

### 3.2.5 Trigger and data acquisition

Collisions are provided by the LHC at high interaction rates, with an interval of 25 ns between bunch crossings. This corresponds to a frequency of 40 MHz. Additionally, multiple collisions occur at the



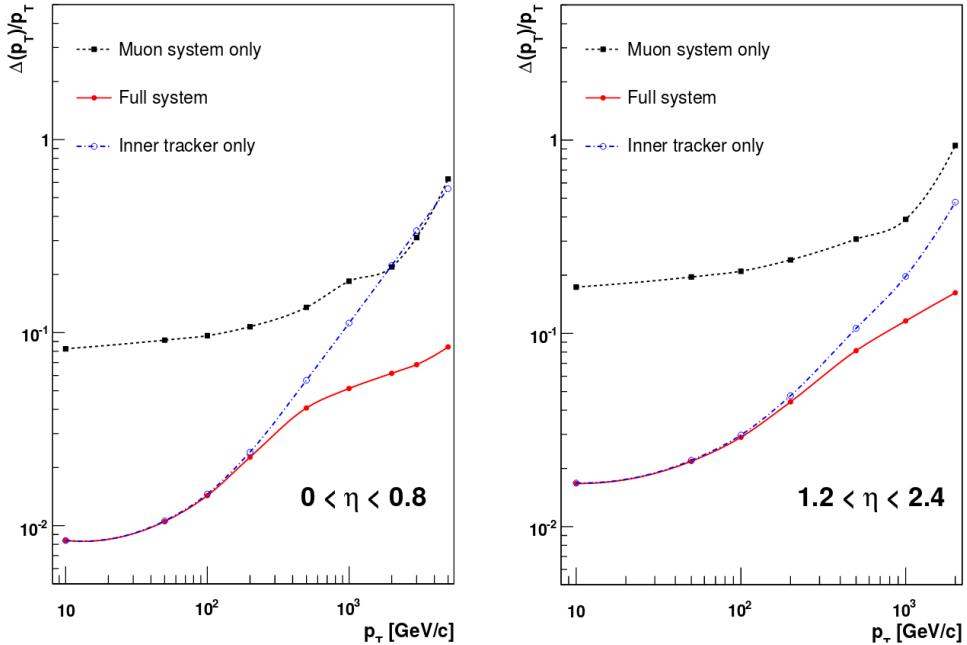
*Figure 3.10: A transverse view of one quarter of CMS showing the position of the 3 types of muon detectors. The Drift Tubes (DT) are located in the barrel, the Cathode Strip Chambers (CSC) in the endcaps, and the Resistive Plate Chambers (RPC) in both regions up to  $|\eta| < 1.8$ .*

same time, depending on the luminosity. Since it is impossible to store and process the large amount of data produced in the collisions at this high rate, a severe rate reduction is needed. This rate reduction is performed by the trigger system, which decides whether to store or reject an event. Since this decision process is constrained in time, the computing time is optimized by rejecting uninteresting events as quickly as possible. The rate is reduced to 1 kHz in two steps by the Level-1 (L1) Trigger and the High-Level Trigger (HLT).

The L1 Trigger decision is based on information from the calorimeters and muon systems, following the structure illustrated in Figure 3.12. At the lowest level, the Local Triggers are based on energy deposits in calorimeter towers and track segments or hit patterns in the muon system. Regional triggers, indicated as Calo Trigger Layer 1 and Muon Track-Finder Layer in the figure, then combine this information and use pattern logic to determine trigger objects such as jet or muon candidates in separated spatial regions. The candidates are ranked based on their energy or momentum and quality, reflecting the level of confidence assigned to the L1 parameter measurements. Finally, the Calo Trigger Layer 2 and the Global Muon Trigger (GMT) determine the highest-rank calorimeter and muon objects across the whole detector and transfer them to the Global Trigger, which makes the final decision to accept or reject an event. Following this procedure, the L1 Trigger thresholds are tuned to reduce the event rate to 100 kHz. The L1 Trigger is composed of custom electronics located partially on the detectors, and partially in the underground service cavern. The L1 decision needs to be made and distributed to the detector front-end electronics within 3.8  $\mu$ s [4].

The readout of the data proceeds as illustrated in Figure 3.13. When an event is accepted by the L1 Trigger, the data from about 740 FEDs is read out by the Readout Units (RUs). For so-called *legacy* systems, i.e. systems which are using VME-based hardware from the initial installation, the FEDs are read out by custom Front-End-Readout-Link (FRL) cards, while for systems that changed their readout architecture from the VME standard to the newer  $\mu$ TCA standard during or after LS1 they are read out via the newer Front-End-Readout-Optical-Link (FEROL) cards. The event fragments are then sent over the event-builder switch to the Builder Units (BUs), which assemble the events. Next, the events are distributed to the Filter Units (FUs) by a large switch network.

The HLT software system is implemented in this filter farm, which uses more than 15000 CPU cores for the final event selection. In this second step, the HLT reduces the event rate further to 1 kHz. The



*Figure 3.11: The muon transverse momentum resolution as a function of transverse momentum for low (left) and (high) pseudorapidities. The resolution is shown for the muon system and the tracker separately, and for the full system. [3]*

- 1 complete read-out data, including information from the pixel and strip tracker, are available for this step.
- 2 New objects can therefore be reconstructed such as e.g. tau leptons and b-jets, as is done in the offline
- 3 software, but speed-optimized.

#### 4 **3.2.6 CMS performance in Run 2**

- 5 The number of collisions recorded at the experiments will differ from the amount delivered by the LHC.
- 6 Data loss can be caused by e.g. problems with a particular subdetector, the trigger rate, the data acqui-
- 7 sition, or the infrastructure. During Run 2, CMS achieved a data taking efficiency of 89% and 92% in
- 8 2015 and 2016, respectively. The comparison between the delivered and recorded cumulative integrated
- 9 luminosity in 2016 is shown in Figure 3.14. Subsequently, the recorded data is certified by the offline
- 10 Data Quality Monitoring (DQM), to ensure that the data are suited for physics analysis.

##### 11 **3.2.6.1 Pre-amplifier saturation in the APV25 chip**

- 12 During Run 2, the instantaneous luminosity delivered by the LHC increased continuously, and even ex-
- 13 ceeded the design luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$  in 2016. As the luminosity increased, a dynamic in-
- 14 efficiency appeared in the strip tracker, which was most noticeable in the first layer of the TOB. The
- 15 symptoms were a change in the signal-to-noise ratio and loss of hits. As can be seen from Figure 3.15,
- 16 the most probable value (MPV) of the signal-to-noise ratio is shifted towards lower values and the low
- 17 tail increased as well. The loss of hits is clearly visible in Figure 3.16, showing the change in number
- 18 of hits per track for increasing instantaneous luminosities. The run periods indicated in the plot refer
- 19 to a subset of the data taken over the course of the year. Run period boundaries are typically defined
- 20 by changes in the LHC running conditions, changes to the detector configuration or calibration, or other
- 21 parameters. The number of hits decreases for later run periods such as D and F, as the instantaneous
- 22 luminosity increases. This loss of hits results in less and shorter tracks.

23 The origin of this inefficiency was eventually tracked down to saturation effects in the pre-amplifier of  
24 the APV25 chip. The pre-amplifier decay time changes significantly with temperature. As the operating  
25 temperature of the strip tracker was lowered from  $+4^\circ\text{C}$  to  $-15^\circ\text{C}$  coolant temperature during LS1, the

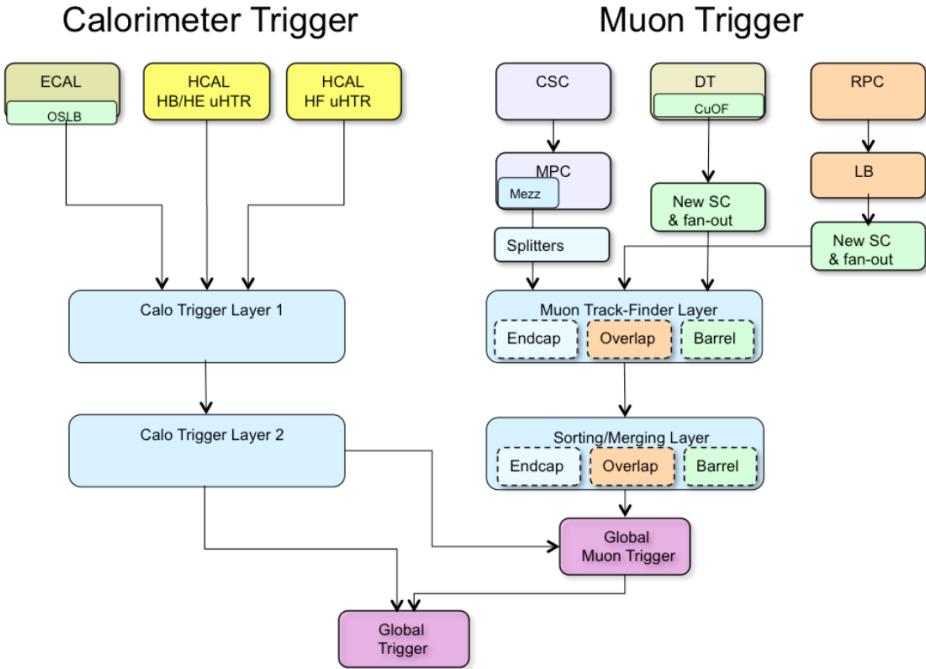


Figure 3.12: Schematic overview of the L1 Trigger. [4]

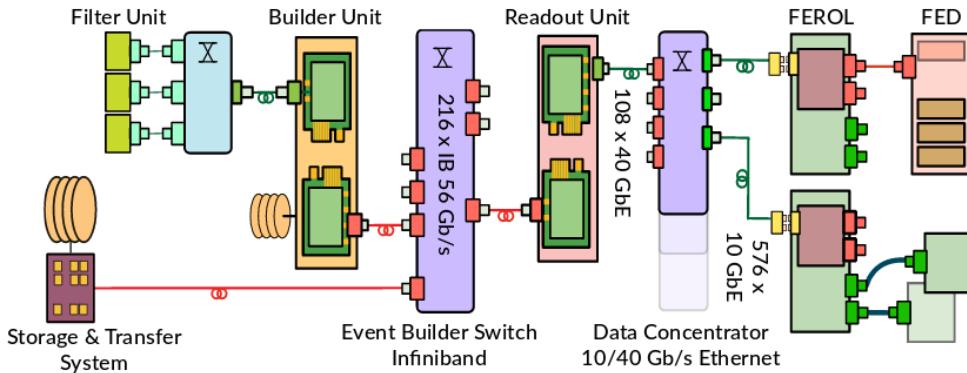


Figure 3.13: Schematic of the data acquisition (DAQ) system. [5]

<sup>1</sup> decay time was no longer sufficient to cope with the high luminosities. The dynamic inefficiency was  
<sup>2</sup> cured in August 2016 by changing the pre-amplifier drain speed. This lead among others to the recovery  
<sup>3</sup> of the muon efficiency, which showed a large drop for the highest luminosities before the change and an  
<sup>4</sup> essentially flat behavior afterwards, as demonstrated in Figure 3.17.

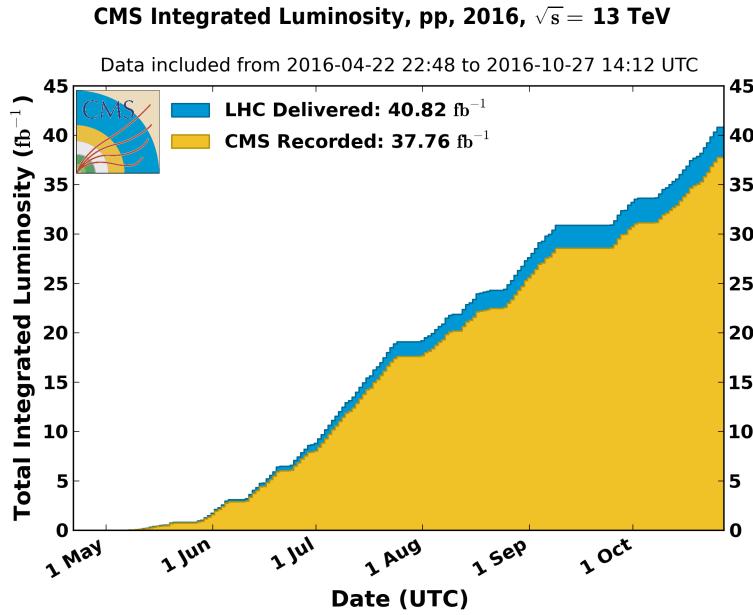


Figure 3.14: The cumulative distribution of the instantaneous luminosity delivered by the LHC (blue) and recorded by CMS (yellow) in 2016.

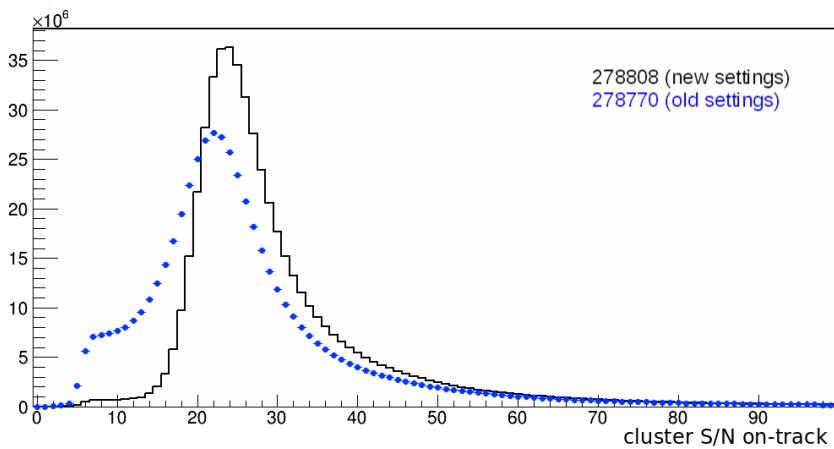


Figure 3.15: The signal-to-noise ratio for clusters on reconstructed tracks in the first layer of the TOB for a run before (blue) and after (black) the change of pre-amplifier drain speed.

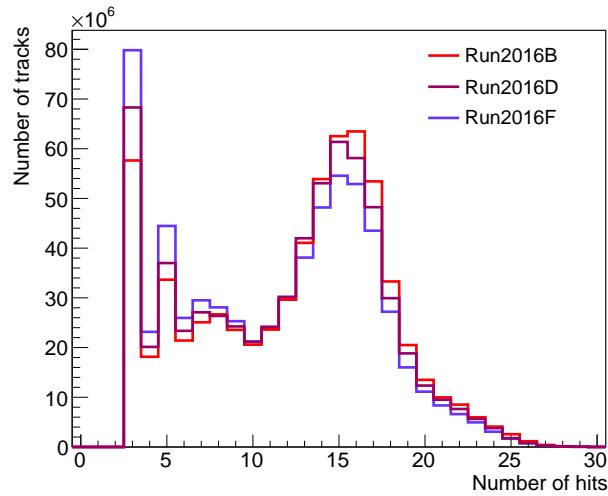


Figure 3.16: The number of hits per track for run periods B, D, and F, showing the effect of the increasing instantaneous luminosity.

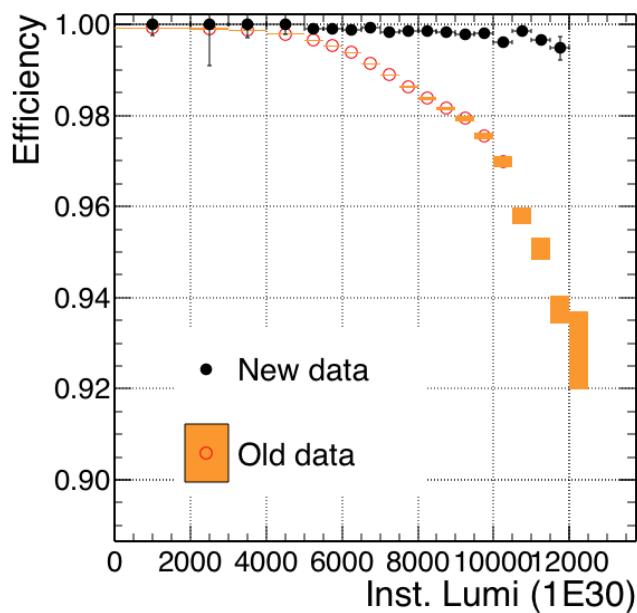


Figure 3.17: The muon efficiency as a function of the instantaneous luminosity for before (orange) and after (black) the change of pre-amplifier drain speed which cured the dynamic inefficiency.

# 4

1  
2

## Event Simulation and Reconstruction

3 In order to use the recorded data, the obtained signals coming from various parts of the detector must be  
4 reconstructed to be able to identify the particles in the event. Additionally, to compare the experimental  
5 results with theory, events are generated and the resulting signals in the detector are simulated, as detailed  
6 in Sections 4.1 and 4.2, respectively. The event reconstruction is detailed in Section 4.3. Finally, some  
7 details about the simulation of strongly interacting massive particles (SIMPs) are given in Section 4.4.

### 8 4.1 Event generation

9 The event structure at the LHC is complicated by the composite nature of protons, as well as the attainable  
10 high momentum transfers. A number of aspects must therefore be taken into account when generating  
11 events, such as parton distribution functions (PDFs), hard scattering, the parton shower, hadronization,  
12 and additional activity in the event.

13 Two partons, meaning the quark or gluon constituents of the colliding protons, will interact with a cer-  
14 tain probability for a given momentum transfer. This is parametrized by the PDFs  $f(x, Q^2)$ , were  $x$  is the  
15 momentum fraction of the partons and  $Q^2$  is the momentum transfer scale. Experimentally determined  
16 PDFs are available from various groups, including e.g. CTEQ [6], MRST/MSTW [7], and NNPDF [8].  
17 An example of such PDFs obtained by the NNPDF group is shown in Figure 4.1. The PDFs are then  
18 used to calculate the matrix element of the hard scattering, which is the process of interest where the  
19 two colliding partons create high-energetic final state particles. This is done using an event generator,  
20 such as MADGRAPH5\_aMC@NLO [9] and POWHEG [10]. With MADGRAPH5\_aMC@NLO the ma-  
21 trix element can be calculated analytically at tree-level or leading order (LO), and since the addition of  
22 aMC@NLO at next-to-leading order (NLO) as well. This generator was used to generate most of the  
23 background processes for the Monojet analysis detailed in Chapter 5 and for the SIMP signal used in  
24 Chapter 6. POWHEG is able to generate events using NLO computations, but only for a relatively lim-  
25 ited number of physics processes. In this thesis, background processes from single-top production were  
26 generated with this program. Since NLO calculations are more time-consuming, one can also scale a LO  
27 cross section to the NLO level by using a so-called k-factor, defined as the ratio of the NLO and LO cross  
28 sections. However, these k-factors often need to be determined as a function of the relevant kinematic  
29 variables as they depend on the kinematic phase space and the probed energy scale.

30 Since the colliding partons have a color charge, the hard scattering will be accompanied by a cascade  
31 of radiation from QCD processes. This radiation can originate from the incoming partons, which is  
32 referred to as initial state radiation (ISR), or the outgoing partons in the final state, the so-called final state  
33 radiation (FSR). The perturbative evolution of the cascade can be modeled using the DGLAP (Dokshitzer-  
34 Gribov-Lipatov-Altarelli-Parisi) equations [11–13]. These equations describe the time evolution of the

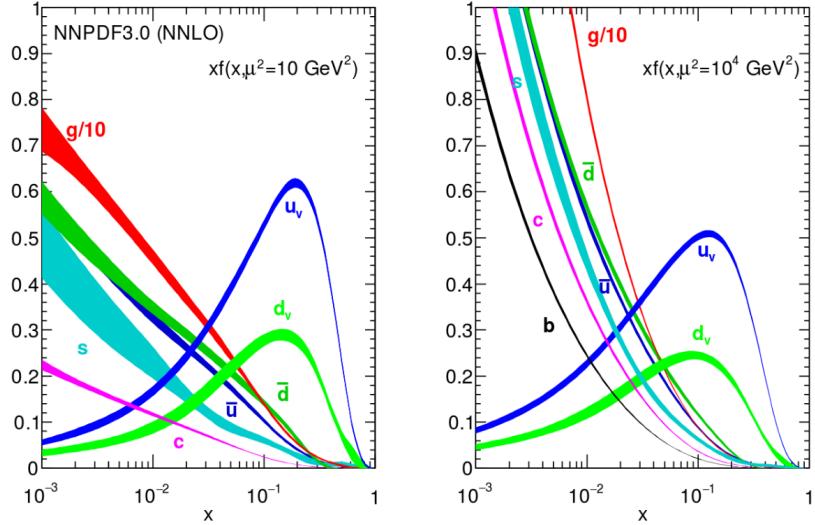


Figure 4.1: The parton distribution functions times the momentum fraction  $x$  at energy scales  $10 \text{ GeV}^2$  (left) and  $10 000 \text{ GeV}^2$  (right), obtained in NNLO NNPDF3.0 global analysis [8].

1 probability of a ‘mother’ parton to split into ‘daughter’ partons at an energy scale  $Q^2$ . The momentum  
 2 of the mother is then divided among the daughter partons, which can in turn split into other partons at  
 3 a lower  $Q^2$  scale. The cascade continues down to an energy scale  $\Lambda_{QCD}$  where the strong coupling  
 4 constant becomes unity.

5 The next step after the showering is the hadronization of the colored particles produced in the parton  
 6 shower, transforming them into color-neutral hadrons. Since this happens at low energy scales where the  
 7 perturbative approach of QCD is not valid, phenomenological models have to be used. For most of the  
 8 processes considered in this thesis, the showering and hadronization is done with PYTHIA 8 [14], using a  
 9 standard set of parameters which were tuned to reproduce the experimental data.

10 In addition to ISR and FSR, also beam remnants and multiple parton interactions give rise to ad-  
 11 ditional activity in the event, referred to as the underlying event. After the partons participating in the  
 12 hard scattering are extracted, the remainder of the protons have a non-zero color charge. The creation of  
 13 additional hadrons during the hadronization is therefore possible. Multiple parton interactions represent  
 14 additional interactions which can take place between other incoming partons. Finally, additional colli-  
 15 sions between other protons in the same bunch crossing or from a previous bunch crossing, respectively  
 16 referred to as in-time and out-of-time pileup, add extra activity in the event.

## 17 4.2 Detector simulation

18 After being generated, the collision events are passed on to the CMS detector simulation, which is based  
 19 on the GEANT 4 [15] simulation toolkit. This toolkit provides a description of the interaction between  
 20 particles and the detector material, including effects such as bremsstrahlung of charged particles, photon  
 21 conversions, energy loss of charged particles by ionization, and the showering of electrons, photons and  
 22 hadrons in the calorimeters due to interaction with the material. The CMS simulation package contains  
 23 the geometry of the detector with all the sensitive layers designed to detect the traversing particles, as well  
 24 as the dead material regions consisting of e.g. support structures, cables and cooling pipes. A precise map  
 25 of the magnetic field is also included in order to simulate the curvature of the charged particles correctly.

26 Next, the electronic response produced by the hits in the active detector material is simulated, resulting  
 27 in an event content similar to the output of the real detector. At this point the effect of pileup is also  
 28 included by adding detector hits of generated proton-proton interactions on top of the hits resulting from  
 29 the main interaction. Most of the simulated event samples used in this thesis are processed using this  
 30 detector simulation. However, the interaction of new particles that can arise from specific theory models  
 31 is not always readily described in GEANT. This is the case for the signal samples used in the analysis

<sup>1</sup> described in Chapter 6, so an additional step was needed in order to simulate strongly interacting massive  
<sup>2</sup> particles (SIMPs) in the CMS detector.

### <sup>3</sup> 4.3 Event reconstruction

<sup>4</sup> Once the detector response has been simulated, the obtained events can be reconstructed. The same  
<sup>5</sup> method is applied for these simulated events and for data coming from the detector. First, the reconstruc-  
<sup>6</sup> tion of tracks is performed, with a specific track reconstruction for electrons and muons. Furthermore,  
<sup>7</sup> the calorimeter deposits, generated by electrons, photons, and hadrons, are grouped into clusters. Addi-  
<sup>8</sup> tionally, the reconstruction is further improved by using the so-called particle flow (PF) algorithm. This  
<sup>9</sup> algorithm greatly improves the performance for jet and hadronic  $\tau$  decay reconstruction, missing trans-  
<sup>10</sup> verse energy momentum determination, as well as electron and muon identification. Finally, the obtained  
<sup>11</sup> particle flow (PF) candidates are clustered into jets, and the missing transverse energy can be derived.

#### <sup>12</sup> 4.3.1 Track reconstruction

<sup>13</sup> The tracks of charged particles going through the CMS tracker are reconstructed with an iterative tracking  
<sup>14</sup> approach. This is used to cope with the high occupancy and consequently high combinatorics. Addition-  
<sup>15</sup> ally, the first iterations search for tracks with less possible combinations, such as tracks with many pixel  
<sup>16</sup> hits or a high momentum. After every iteration, the hits associated with the found track are removed to  
<sup>17</sup> reduce the combinatorics. Each iteration consists of four steps:

- <sup>18</sup> 1. **Seed generation.** In this first step hits are combined into seeds for the subsequent track finding. In  
<sup>19</sup> the initial iterations pixel triplets are used, then pixel pairs, in order to take gaps or non-working  
<sup>20</sup> modules into account. Next, mixed pixel/strip triplets are taken, and finally strip-only seeds are  
<sup>21</sup> used. These additional iterations improve the acceptance in  $p_T$  and in displacement with respect to  
<sup>22</sup> the primary vertex.
- <sup>23</sup> 2. **Track finding.** The seeds are used as starting point for a Kalman filter algorithm. This method  
<sup>24</sup> extrapolates the seed trajectory outward to the next layer, taking into account potential energy  
<sup>25</sup> loss and multiple scattering. If compatible hits are found in the next layer, the parameters of the  
<sup>26</sup> trajectory are updated. This process continues until the outermost layer of the tracking system.  
<sup>27</sup> Using this method, a given seed can generate multiple tracks, or different tracks can share hits. A  
<sup>28</sup> trajectory cleaner therefore determines the fraction of hits the tracks have in common and discards  
<sup>29</sup> the track with the lowest number of hits when there are too many shared hits. If both tracks have  
<sup>30</sup> the same number of hits, the track with the largest  $\chi^2$  value is removed.
- <sup>31</sup> 3. **Track fitting.** The track parameters are then refitted using a Kalman filter and smoother, taking all  
<sup>32</sup> hits determined in the track finding step into account.
- <sup>33</sup> 4. **Track selection.** Finally, the tracks are selected based on quality requirements, such as the number  
<sup>34</sup> of layers that have hits, the  $\chi^2/\text{dof}$ , and the distance to a primary vertex. This greatly reduces the  
<sup>35</sup> fraction of reconstructed tracks that are fake.

<sup>36</sup> The performance of the track reconstruction is excellent, and a high track-finding efficiency is ob-  
<sup>37</sup> tained [16] while keeping the rate of fake tracks negligible. The highest tracking efficiency is obtained  
<sup>38</sup> for muons, which traverse the full detector volume and have an improved momentum resolution due to  
<sup>39</sup> tracking information from the muon detectors giving a long lever arm. For isolated muons with  $p_T$  be-  
<sup>40</sup> tween 1 and 100 GeV the tracking efficiency is higher than 99% for the entire  $\eta$  coverage of the tracker,  
<sup>41</sup> as can be seen from the left plot in Figure 4.2. The  $p_T$  resolution is about 2-3% for a muon with  $p_T =$   
<sup>42</sup> 100 GeV up to  $|\eta| < 1.6$ , but worsens for higher pseudorapidities. Different types of particles interact  
<sup>43</sup> differently with the detector material. Charged hadrons, for example, are also subject to elastic and in-  
<sup>44</sup> elastic nuclear interactions and have a tracking efficiency of 80-95% depending on pseudorapidity and  
<sup>45</sup> transverse momentum, as shown in the right plot of Figure 4.2.

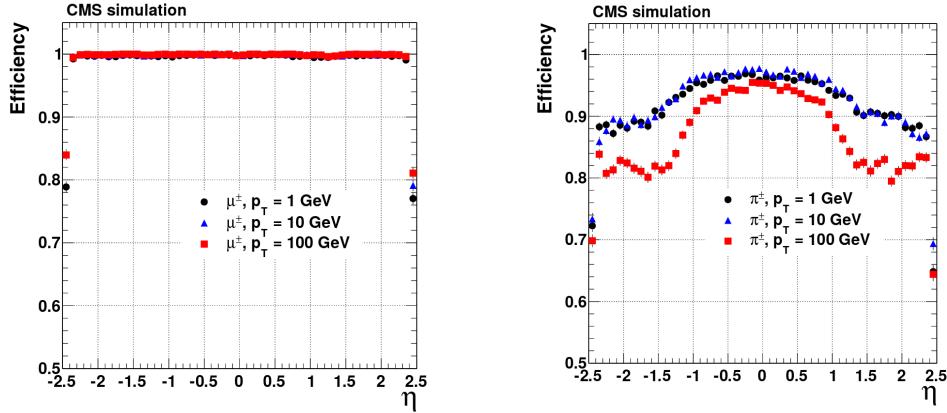


Figure 4.2: The muon efficiency (left) and pion efficiency (right) as a function of pseudorapidity, for multiple transverse momenta. [16]

Finally, the primary vertex is reconstructed from the tracks. Since the collisions happen between bunches of protons, multiple protons will be colliding at the same time. The extra collisions, next to the potentially interesting collision, are referred to as pile-up interactions. The particles generated in these collisions are all detected simultaneously and form a challenge to disentangle them from the particles coming from the to be studied interaction.

The reconstruction is done in 2 steps: first the tracks that appear to originate from the same interaction vertex are clustered, then a fitting procedure computes the vertex parameters and assigns a weight to each associated track, reflecting the probability that it corresponds to the considered vertex. Figure 4.3 shows the reconstruction efficiency and the resolution of the primary vertex. The more tracks, the better the vertex is constrained and thus the better the resolution.

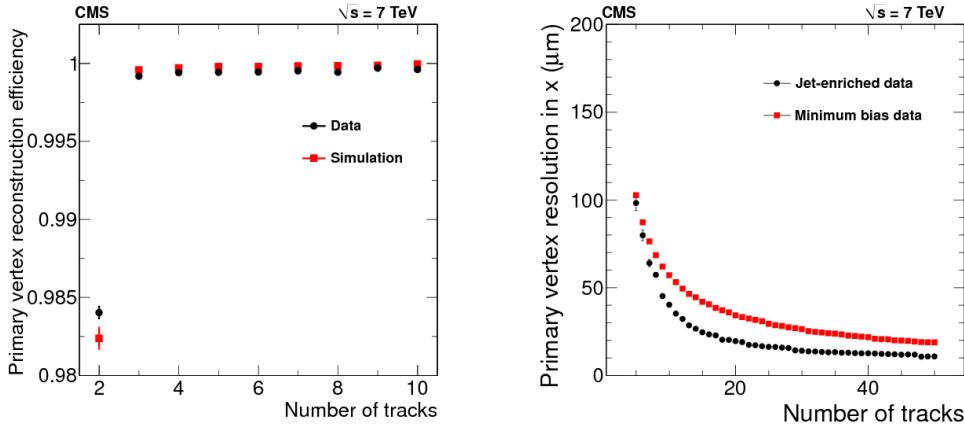


Figure 4.3: The primary vertex reconstruction efficiency (left) and resolution (right) as a function of the number of tracks associated to it. [16]

### 4.3.2 Electron and isolated photon reconstruction

Electrons are reconstructed using information from both the tracker and the calorimeters. Due to the large amount of material present in the tracker, electrons will emit bremsstrahlung photons, and photons will often convert into  $e^+e^-$  pairs, which can again radiate bremsstrahlung photons. The electron and photon reconstruction is therefore very similar.

For electrons, a Gaussian-sum filter (GSF) [17] candidate is taken as starting point. This GSF candidate is obtained using 2 different methods to reconstruct the electron track from the hits in the tracker, which should gather all radiated energy from the electron. First, the ECAL-based approach is used,

grouping ECAL clusters into superclusters. These superclusters collect the energy of the electron and the bremsstrahlung photons in a small  $\eta$  window and a large  $\phi$  window, taking the bending of the electron track in the magnetic field into account. The supercluster energy and position is then used to estimate the position of the corresponding hits in the tracker layers. Subsequently, the tracker-based approach is used to reconstruct electrons missed by the ECAL-based method. In this case, all the tracks from the iterative tracking with transverse momentum larger than 2 GeV are used. Next, the specific electron tracking is performed, using a GSF fit, which is more adapted to electrons than the Kalman filter used in the iterative tracking, as it describes the energy loss in each tracker layer. The electron seeds obtained with both methods are merged and used as input for the full electron tracking, which is performed with twelve GSF components. The obtained electron tracks are then linked to ECAL clusters by the PF algorithm, as described in Section 4.3.4. In the case of isolated photons, a candidate is seeded from an ECAL supercluster with transverse energy larger than 10 GeV which is not linked to a GSF track.

The total energy of the accumulated ECAL clusters is corrected for the energy that was lost in the process of reconstruction, using analytical functions of the energy and pseudorapidity. The applied corrections can be as large as 25%, at low transverse momentum and at  $|\eta| = 1.5$ , where the material density in the tracker is largest. The energy of the electron is then obtained from a combination of the corrected energy and the momentum of the GSF track, while the direction of the electron is taken from the GSF track. For photons, the corrected energy and the direction of the supercluster are used.

Additionally, the electron and photon candidates must satisfy identification criteria to be retained. In the case of electrons a boosted decision tree is used, combining fourteen variables including the amount of energy radiated and the ratio between the energies gathered in HCAL and ECAL, while for photons the candidates must be isolated from other tracks and calorimeter clusters, and the energy distribution in the ECAL and the ratio between the HCAL and ECAL energies must be compatible with the expectation from a photon shower.

### 4.3.3 Muon reconstruction

Muon tracking is performed using 2 complementary approaches. The first method starts from standalone muons which are reconstructed from hits in the muon detectors using pattern recognition. The standalone muons are then matched to tracks in the tracker, and the hits are combined to form a global muon track. This global muon fit improves the momentum resolution compared to the tracker-only fit at muon momenta larger than 200 GeV.

For momenta below 10 GeV, muons often fail the global muon conditions which require the muon to penetrate through more than one muon detector plane, due to the large multiple scattering in the return yoke. In this case, tracker muon reconstruction is more efficient since it only requires one muon segment. Each track in the tracker with a transverse momentum larger than 0.5 GeV and a total momentum larger than 2.5 GeV is therefore extrapolated to the muon system and if at least one matching track segment is found, it is retained as muon candidate.

Within the geometrical acceptance of the muon system about 99% of the muons are reconstructed, either as global muon or as tracker muon and frequently as both. Global and tracker muons that share the same track inside the tracker are merged into a single candidate. Muons that are only reconstructed as standalone muons have a worse momentum resolution compared to the global and tracker muons.

Charged hadrons can be misreconstructed as muons if e.g. a part of the hadron shower reaches the muon system. In order to improve the muon identification, the PF muon identification algorithm described in Section 4.3.4 also matches energy deposits in the ECAL and HCAL with the muon track.

### 4.3.4 Particle flow

The particle flow (PF) algorithm reconstructs so-called particle flow candidates by combining information from all different CMS subdetectors, linking different elements, such as tracks in the tracker, calorimeter clusters, and muon tracks. The obtained collection of particle candidates is subsequently used to reconstruct jets and to determine the missing transverse energy.

In a first step, the PF algorithm identifies charged particle tracks, as defined in Section 4.3.1, and Sections 4.3.2 and 4.3.3 for electron and muon tracks, respectively. At the same time, the calorimeter clusters are reconstructed with a clustering algorithm designed specifically for the PF event reconstruction. In this algorithm, cluster seeds are first identified as local energy maxima with respect to the four or eight closest cells, if the energy deposited in the cell is above a given seed threshold. The clusters are then formed by accumulating neighboring cells with an energy above a given cell threshold, suppressing noise.

The PF elements in the different subdetectors are then connected by a link algorithm which avoids any double counting. The link algorithm produces blocks of associated elements, quantifying the quality of the link by defining a geometrical distance between the elements. When an element is linked to multiple other elements, only the link with the shortest distance is kept. More precisely, a link between a track in the tracker and a calorimeter cluster is made by extrapolating it from the last hit in the tracker to the calorimeters. The distance between the position of the extrapolated track and the cluster in the  $(\eta, \phi)$  plane is then used to define the link distance. At the interaction points between the track and the tracker layers, tangents to the GSF tracks are extrapolated to the ECAL in order to collect the energy of photons radiated by electron bremsstrahlung. A dedicated conversion finder was also developed to identify bremsstrahlung and prompt photon conversions into  $e^+e^-$  pairs. Links between calorimeter clusters are established outside of the tracker acceptance, or between the preshower and ECAL clusters in the preshower acceptance. In this case the link distance is also defined as the distance between the position of the clusters. Charged particle tracks can also be linked by a common secondary vertex. Finally, the PF muon identification algorithm associates the muon tracks to the muon energy deposits in the ECAL and HCAL, to improve the muon identification performance.

In a next step, the PF blocks are classified as muons, electrons, or isolated photons. The corresponding elements are then excluded from further consideration. Once electrons, muons, and isolated photons have been identified, the remaining elements are identified as charged hadrons, neutral hadrons, or photons produced in jets. Within the tracker acceptance, the ECAL clusters not linked to any track are classified as photons, while the clusters in the HCAL without a matched track are labeled as neutral hadrons. Outside of the tracker acceptance, charged and neutral hadrons can not be distinguished. ECAL clusters linked to an HCAL cluster are then assumed to arise from the same hadron shower, and the estimated energy for these particles is the sum of the energy deposited in the ECAL and the HCAL. The remaining clusters are then linked to one or several tracks in order to reconstruct the charged hadrons.

### 4.3.5 Jet reconstruction

Jets are reconstructed with the anti- $k_T$  algorithm [18], which clusters either the particles reconstructed by the PF algorithm (PF jets) or the energy deposits in the calorimeters (Calo jets). This procedure takes into account the transverse momentum  $p_T$ , also called  $k_T$ , of the particles and the distance between particles, defined as

$$\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}. \quad (4.1)$$

The strategy consists of the following steps:

1. For every pair of particles  $i$  and  $j$ , a distance  $d_{ij}$  defined as

$$d_{ij} = \min \left( \frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2} \right) \frac{\Delta R_{ij}^2}{R^2} \quad (4.2)$$

is calculated.

2. For every particle  $i$ , a distance  $d_{iB}$  to the beam pipe is calculated with

$$d_{iB} = 1/p_{Ti}^2. \quad (4.3)$$

3. The minimum of  $d_{ij}$  and  $d_{iB}$  is then determined.

4. If it is  $d_{ij}$ , particles  $i$  and  $j$  are recombined into a new particle by adding the four-momenta of the particles. If it is  $d_{iB}$ , particle  $i$  is declared to be a jet and it is removed from the list of particles.

1       5. This is repeated until no particles remain.

2       In this clustering algorithm, the parameter  $R$  determines what is called a jet. If a particle  $i$  has  
3       no other particles within a distance  $R$ ,  $d_{iB}$  will be smaller than  $d_{ij}$  and the particle will become a jet.  
4       A consequence of this is that an arbitrarily soft particle can become a jet, and therefore a minimum  
5       transverse momentum for a jet to be of interest is defined.

6       The anti- $k_T$  algorithm favors clustering around hard particles, and the jets then grows outward from  
7       this seed. However, since it still involves a combination of energy and angle in the distance measure, this  
8       is a collinear-safe growth, meaning that the jet will not change when one of the particles of the jet is split  
9       collinearly. This algorithm is also infrared-safe, i.e. the same set of jets is obtained when soft particles  
10      are emitted. and gives rise to circular jets.

11      A reliable determination of the jet energy is however not straightforward, since many effects can dis-  
12      tort the energy estimation, such as the calorimeter response, the limited particle reconstruction efficiency,  
13      the underlying event, the pileup, and the charged particles bending out of the jet cone due to the strong  
14      magnetic field. The pileup is mitigated by applying charged hadron subtraction (CHS), which consists of  
15      removing charged hadrons associated with vertices other than the primary vertex from the list of PF can-  
16      didates. Additionally, the jet energy is corrected using a factorized approach, as illustrated in Figure 4.4,  
17      with the following steps:

- 18      • **Pileup correction (L1).** The first level of jet energy corrections is applied event-by-event and jet-  
19      by-jet, and is determined from simulation. It is dependent on the pseudorapidity and transverse  
20      momentum of the jet, the average  $p_T$  density in the event, and the effective jet area. This effective  
21      area is determined by injecting a large number of very soft particles in the event before the jet  
22      clustering. The spread of the soft particles in each jet then defines the jet area.
- 23      • **Relative  $\eta$  and absolute  $p_T$  corrections (L2L3).** This correction is also obtained from simulations  
24      and corrects for the non-uniform response of the calorimeters in  $\eta$  and  $p_T$ .
- 25      • **Residual  $\eta$  and  $p_T$  corrections (L2L3Residual).** Since the L2L3 correction is derived from simu-  
26      lation, additional residual corrections are needed in order to correct for the remaining small differ-  
27      ences between the jet response in data and simulation. These corrections are typically of the order  
28      of a few percent.

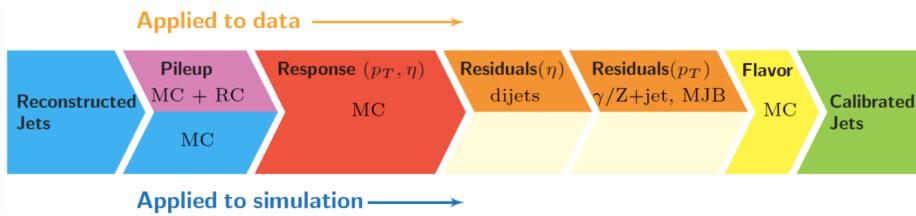


Figure 4.4: Graphical overview of the factorized approach used at CMS to apply jet energy corrections.

29      Finally, a set of identification criteria are applied on the PF jets. A jet is required to consist of at least  
30      two particles. For jets in the region  $|\eta| < 2.7$ , the fraction of energy coming from ether neutral hadrons  
31      or photons should not exceed 99%. Additionally, for jets restricted to the tracker acceptance ( $|\eta| < 2.4$ ),  
32      there should at least be some energy deposited in the HCAL, the jet should contain 1 or more charged  
33      constituent, and the fraction of energy corresponding to electrons or photons should not exceed 99%.

#### 34      4.3.6 Missing transverse energy reconstruction

35      While most particles produced in the collisions can be reconstructed from the hits and energy deposits  
36      in the detector, some collision products might not leave energy deposits in tracker, calorimeters or muon  
37      system. This makes an accurate reconstruction of this type of particles rather challenging. Another  
38      method is therefore used, based on indirect observations. As the detector is hermetically closed such that  
39      all other particles in the event can be detected, the missing transverse energy can be determined. This

1 energy then corresponds to all undetected particles in the event, and can be calculated from the vectorial  
 2 sum of the transverse momenta of all the observed final state particles:

$$\vec{E}_T^{miss} = - \sum \vec{p}_T, \quad (4.4)$$

3 where the sum runs over all reconstructed PF particles.

4 A notable example of particles leaving no hits or energy deposits behind are neutrinos, as they are  
 5 neutral and weakly interacting and will therefore traverse the entire detector. Other hypothetical neutral  
 6 weakly interacting particles, which are being searched for in many physics analyses, would escape the  
 7 detector without producing hits as well.

## 8 4.4 Simulation of the SIMP signal

9 For the generation of the SIMP signal, the model Lagrangian given in equation ?? is implemented in  
 10 FEYNRULES 2.0 [19]. Next, the matrix element is calculated at LO and events are generated using  
 11 MADGRAPH 5. The subsequent parton shower and hadronization is done with PYTHIA 8, using tune  
 12 CUEP8M1. Next, the events are simulated in the CMS detector using GEANT. However, the SIMPs  
 13 are not included in the simulation, as their interaction with matter is not implemented in GEANT. The  
 14 SIMPs were instead incorporated by adding an additional step to the standard reconstruction described in  
 15 Section 4.3. In this additional step the SIMPs are directly converted to neutral PF candidates and merged  
 16 with the rest of the PF candidates. The standard pileup corrections, jet clustering, and charged hadron  
 17 subtraction are then applied in order to obtain the resulting jets, denoted here as P2PF jets.

18 In order to validate this method, a second sample was produced using neutrons instead of SIMPs and  
 19 applying the same additional step. In this case the the neutrons will also be correctly reconstructed by the  
 20 standard reconstruction. The reconstructed PF candidates that are matched to the generated neutrons were  
 21 therefore removed before injecting the converted generated neutrons to the collection of PF candidates.  
 22 This sample is then used to evaluate the difference with a standard neutron sample where the full GEANT  
 23 simulation is done. Neutrons were chosen because of their resemblance to the SIMPs as single neutral  
 24 particles generating a hadronic shower.

25 The two leading generator-level jets (GEN) can then be compared to the uncorrected P2PF jets from  
 26 the custom sample and the PF jets from the standard sample. The transverse momentum of these jets is  
 27 compared in Figure 4.5. This shows that the jet energy resolution (JER) is not described properly, since  
 28 the additional step directly converts generated particles to PF candidates without taking into account any  
 29 other effects. In order to produce a more realistic simulation, the new PF candidates are therefore smeared  
 30 with JER distributions derived using the uncorrected PF jets matching the neutrons in the standard neutron  
 31 sample in bins of  $\eta$  and  $p_T$ . An example of this resolution is shown in Figure 4.6 for central neutrons  
 32 with low and high transverse momentum.

33 After applying this smearing, the P2PF jets are processed with the standard sequence of charged  
 34 hadron subtraction, jet clustering, L1FastJet, and L2/L3 corrections. The comparison of the corrected  
 35 P2PF jets and the standard corrected PF jets is shown in Figure 4.7 for the neutron sample, validating  
 36 that the jet transverse momentum is now correctly smeared. The JER distributions are also compared in  
 37 Figure 4.8 and fitted with a Crystal Ball function, showing compatible parameters. This demonstrates  
 38 that the procedure, where the JER distributions derived from a neutron sample are used to smear the PF  
 39 candidates from generator-level SIMPs, can sufficiently accurately simulate SIMPs in a realistic detector.

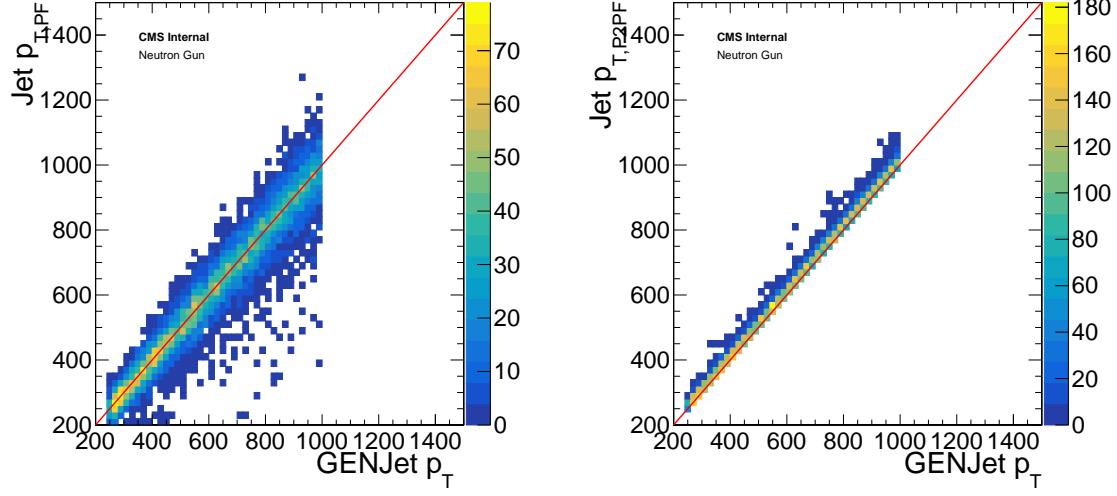


Figure 4.5: Comparison of the transverse momentum of the generator-level jets to the PF jets (left) and P2PF jets (right) without jet energy resolution smearing, using a neutron sample.

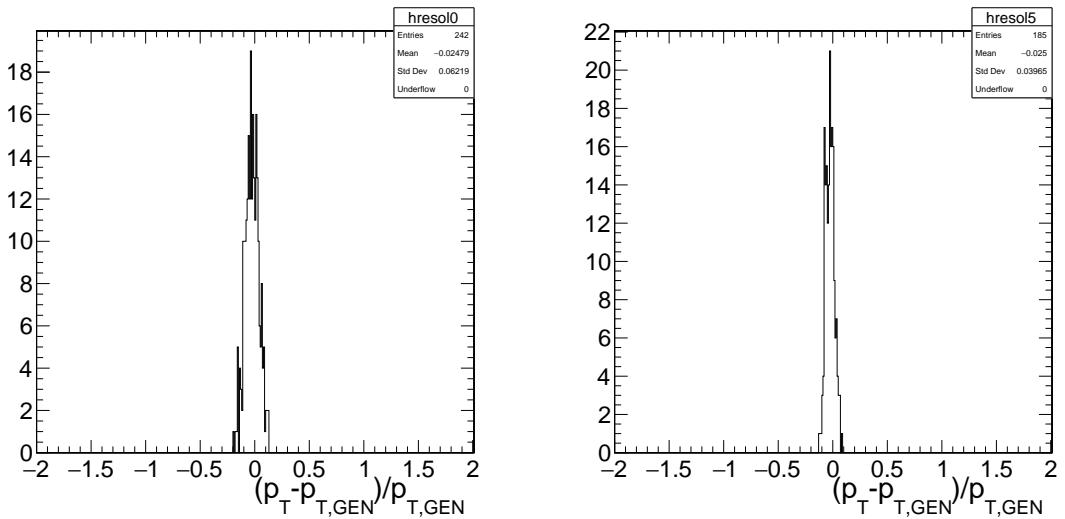


Figure 4.6: The jet energy resolution of neutrons with  $0 < |\eta| < 0.5$  and  $200 \text{ GeV} < p_T < 300 \text{ GeV}$  (left) or  $700 \text{ GeV} < p_T < 800 \text{ GeV}$  (right).

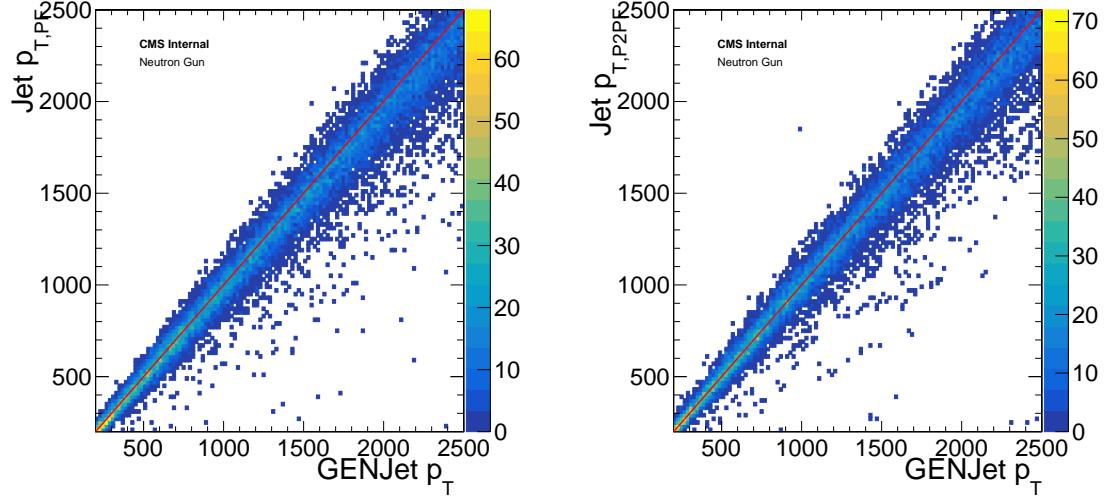


Figure 4.7: Comparison of the transverse momentum of the generator-level jets to the PF jets (left) and P2PF jets (right) in the region  $0 < |\eta| < 0.5$  with jet energy resolution smearing, using a neutron sample.

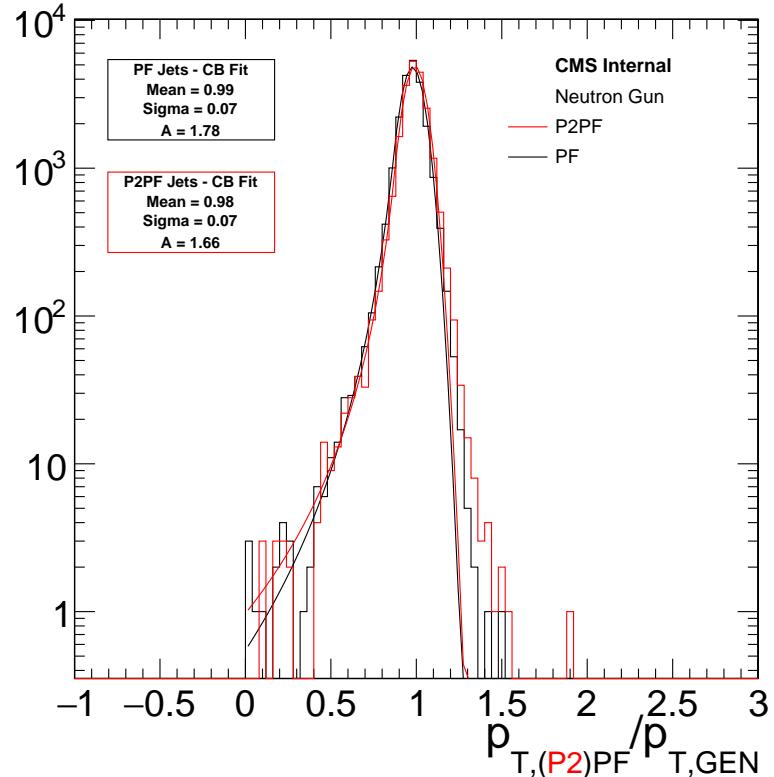


Figure 4.8: The jet energy resolution of the corrected P2PF jets (red) and PF jets (black), fitted with a Crystal Ball function.

# 5

## The Monojet Analysis

<sup>1</sup> **5.1 Introduction**

<sup>2</sup> **5.2 Event selection**

<sup>3</sup> **5.3 Background estimation**

<sup>4</sup> **5.4 Results**

<sup>5</sup> **5.5 Improvement going from the 2015 to 2016 analysis**

<sup>6</sup> **5.6 Interpretation**



# 6

1

2

## Search for SIMPs using Trackless Jets

3 **6.1 Introduction**

4 **6.2 Event selection**

5 **6.3 Background estimation**

6 **6.4 Results**

7 **6.5 SIMP model interpretation**



# 7

1

2

## Conclusion & Outlook



## References

1

- 2 [1] David Aaron Matzner Dominguez, D. Abbaneo, K. Arndt, N. Bacchetta, A. Ball, E. Bartz, W. Bertl,  
3 G. M. Bilei, G. Bolla, H. W. K. Cheung, et al. *CMS Technical Design Report for the Pixel Detector  
4 Upgrade*. 2012.
- 5 [2] Serguei Chatrchyan et al. *Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter  
6 in pp Collisions at  $\sqrt{s} = 7$  TeV*. JINST, 8:P09009, 2013. [JINST8,9009(2013)].
- 7 [3] S. Chatrchyan et al. *The CMS Experiment at the CERN LHC*. JINST, 3:S08004, 2008.
- 8 [4] A. Tapper and Darin Acosta. *CMS Technical Design Report for the Level-1 Trigger Upgrade*. 2013.
- 9 [5] Jean-marc Olivier Andre, Ulf Behrens, James Branson, Philipp Maximilian Brummer, Olivier  
10 Chaze, Sergio Cittolin, Cristian Contescu, Benjamin Gordon Craigs, Georgiana Lavinia Dar-  
11 lea, Christian Deldicque, Zeynep Demiragli, Marc Dobson, Nicolas Doualot, Samim Erhan,  
12 Jonathan Richard Fulcher, Dominique Gigi, Maciej Szymon Gladki, Frank Glege, Guillermo  
13 Gomez Ceballos, Jeroen Guido Hegeman, Andre Georg Holzner, Mindaugas Janulis, Raul  
14 Jimenez Estupinan, Lorenzo Masetti, Franciscus Meijers, Emilio Meschi, Remigius Mommsen,  
15 Srecko Morovic, Vivian O'Dell, Luciano Orsini, Christoph Maria Ernst Paus, Petia Petrova, Marco  
16 Pieri, Attila Racz, Thomas Reis, Hannes Sakulin, Christoph Schwick, Dainius Simelevicius, Petr  
17 Zejdl, Hamburg, California, San Diego, California, Chicago, Illinois, Massachusetts Institute  
18 of Technology, Cambridge, Massachusetts, also at Vilnius University, Vilnius, Lithuania, and also  
19 at. *Performance of the CMS Event Builder*. Technical Report CMS-CR-2017-034, CERN, Geneva,  
20 Feb 2017.
- 21 [6] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, and W. K. Tung. *New generation  
22 of parton distributions with uncertainties from global QCD analysis*. JHEP, 07:012, 2002.
- 23 [7] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. *Parton distributions for the LHC*. Eur. Phys.  
24 J., C63:189–285, 2009.
- 25 [8] Richard D. Ball et al. *Parton distributions for the LHC Run II*. JHEP, 04:040, 2015.
- 26 [9] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Tor-  
27 rielli, and M. Zaro. *The automated computation of tree-level and next-to-leading order differential  
28 cross sections, and their matching to parton shower simulations*. JHEP, 07:079, 2014.
- 29 [10] Stefano Frixione, Paolo Nason, and Carlo Oleari. *Matching NLO QCD computations with Parton  
30 Shower simulations: the POWHEG method*. JHEP, 11:070, 2007.
- 31 [11] V. N. Gribov and L. N. Lipatov. *Deep inelastic e p scattering in perturbation theory*. Sov. J. Nucl.  
32 Phys., 15:438–450, 1972. [Yad. Fiz. 15, 781 (1972)].
- 33 [12] Yuri L. Dokshitzer. *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e-  
34 Annihilation by Perturbation Theory in Quantum Chromodynamics*. Sov. Phys. JETP, 46:641–653,  
35 1977. [Zh. Eksp. Teor. Fiz. 73, 1216 (1977)].
- 36 [13] Guido Altarelli and G. Parisi. *Asymptotic Freedom in Parton Language*. Nucl. Phys., B126:298–  
37 318, 1977.

- 1 [14] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. *PYTHIA 6.4 physics and manual*. JHEP,  
2 05:026, 2006.
- 3 [15] John Allison et al. *Geant4 developments and applications*. IEEE Trans. Nucl. Sci., 53:270, 2006.
- 4 [16] Serguei Chatrchyan et al. *Description and performance of track and primary-vertex reconstruction  
5 with the CMS tracker*. JINST, 9(10):P10009, 2014.
- 6 [17] A. Strandlie and R. Fruhwirth. *Discrimination between different types of material in track recon-  
7 struction with a Gaussian-sum filter*. IEEE Trans. Nucl. Sci., 53:3842–3849, 2006.
- 8 [18] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. *The anti-  $k_t$  jet clustering algorithm*. Journal  
9 of High Energy Physics, 2008(04):063, 2008.
- 10 [19] Adam Alloul, Neil D. Christensen, Céline Degrande, Claude Duhr, and Benjamin Fuks. *FeynRules  
11 2.0 - A complete toolbox for tree-level phenomenology*. Comput.Phys.Commun., 185:2250–2300,  
12 2014.

# List of Acronyms

1

2

3

## A

4

5

6 ATLAS A Toroidal LHC ApparatuS

7

8

## B

9

10 BU Builder Unit

11

12

13

## C

14

15 CERN European Organization for Nuclear Research  
16 CHS charged hadron subtraction  
17 CMS Compact Muon Solenoid  
18 CSC Cathode Strip Chambers

19

20

21

## D

22

23 DAQ data acquisition  
24 DQM Data Quality Monitoring  
25 DT Drift Tubes

26

27

28

## E

29

30 ECAL electromagnetic calorimeter

31

32

33

## F

34

35 FED Front End Driver  
36 FSR final state radiation  
37 FU Filter Unit

1

2

**G**

3

5 GSF

Gaussian-sum filter

6

7

**H**

8

10 HCAL

hadronic calorimeter

11 HLT

High-Level Trigger

12

13

**I**

14

16 IP

interaction point

17 ISR

initial state radiation

18

19

**J**

21

22 JER

jet energy resolution

23

24

**L**

25

27

28

29

30

31

32

L1

Level-1

Low Energy Ion Ring

LEP

Large Electron Positron

LHC

Large Hadron Collider

LO

leading order

33

34

**N**

35

36 NLO

next-to-leading order

37

38

**P**

40

41 PDF

parton distribution function

42 PF

particle flow

43 PS

Proton Synchrotron

44 PSB

Proton Synchrotron Booster

1

2

3 **Q**

4

5 QCD                    Quantum Chromodynamics

6

7

8 **R**

9

10 RF                    radio-frequency

11 ROC                  read-out chip

12 RPC                  Resistive Plate Chambers

13 RU                   Readout Unit

14

15

16 **S**

17

18 SIMP                strongly interacting massive particle

19 SPS                  Super Proton Synchrotron

20

21

22 **T**

23

24 TEC                  Tracker EndCaps

25 TIB                  Tracker Inner Barrel

26 TID                  Tracker Inner Disks

27 TOB                  Tracker Outer Barrel

