

# Contents

<b>1</b>	<b>Experimental setup: Collider, detector and algorithms</b>	<b>1</b>
<b>1.1.</b>	<b>The Large Hadron Collider . . . . .</b>	<b>3</b>
<b>1.2.</b>	<b>The CMS detector . . . . .</b>	<b>5</b>
1.2.1	The tracking system . . . . .	6
1.2.2	The electromagnetic calorimeter . . . . .	8
1.2.3	The hadronic calorimeter . . . . .	9
1.2.4	The muon system . . . . .	11
1.2.5	The trigger system . . . . .	12
<b>1.3.</b>	<b>Event reconstruction and particle identification . . . . .</b>	<b>12</b>
1.3.1	The particle-flow algorithm . . . . .	13
1.3.2	Object reconstruction . . . . .	14
1.3.2.1	Reconstruction of primary vertices . . . . .	14
1.3.2.2	Reconstruction of tracks . . . . .	15
1.3.2.3	Reconstruction of jets . . . . .	17
1.3.2.4	Reconstruction of photons . . . . .	20
1.3.2.5	Reconstruction of muons . . . . .	21
1.3.2.6	Reconstruction of electrons . . . . .	22
1.3.2.7	Reconstruction of taus . . . . .	23
1.3.2.8	Reconstruction of missing transverse energy . . . . .	24
1.3.3	Event cleaning . . . . .	24
<b>1.4.</b>	<b>Simulation of events . . . . .</b>	<b>25</b>



## **Part 1**

**Experimental setup: Collider, detector and algorithms**



## 1.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [1, 2] is a particle accelerator installed in the former LEP [3] tunnel at CERN<sup>1</sup>. It is 26.7 km in circumference and consists of two separate rings, which are, in periods of operation, inhabited by two counter-circulating beams. At the interaction points of the two beams, either proton-proton collisions or heavy ion collisions take place. In this thesis, only  $pp$ -collision data from the year 2012 is analysed. Thus, all machine values cited in the following chapters and paragraphs refer to the setup for  $pp$  collisions in 2012 if not stated otherwise.

The beams are separated into bunches which rotate with a bunch spacing of 50 ns corresponding to a collision frequency of 20 MHz. Before the bunches are actually filled into the LHC ring they are pre-accelerated in other accelerators, which are in the order they are actually passed by the protons: Linac2, Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), Super Proton Synchrotron (SPS). The beam energy is 450 GeV when the beams enter the LHC ring. The injector chain and the LHC ring with its experiments are visualised in Fig 1.1.

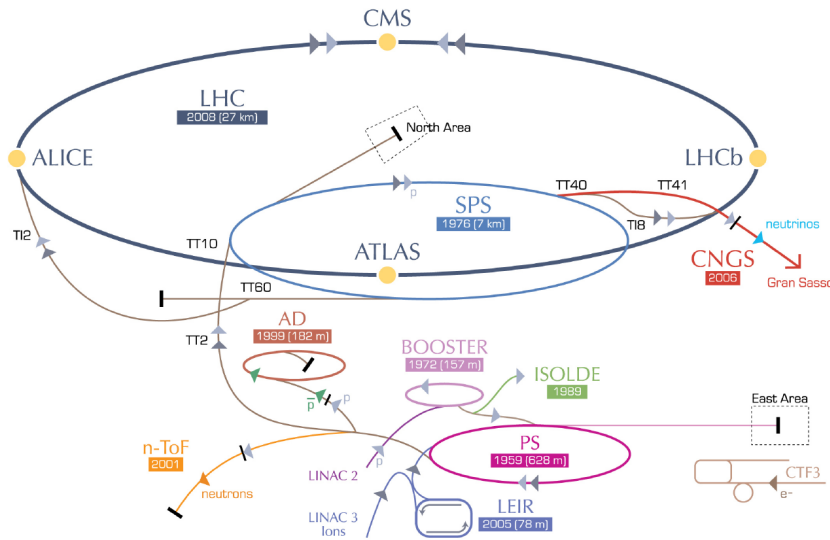


Figure 1.1: Visualisation of the LHC with its experiments and the injector chain. Taken from [4].

<sup>1</sup>European Organisation for Nuclear Research

In the LHC, the beams are kept on their circular path with the help of a magnetic dipole field of 4.76 T. Further quadrupole and sextupole magnets squeeze and focus the bunches resulting in a bunch spread of roughly 8 cm length and a Gaussian shape radius of  $20 \mu\text{m}$  RMS at the interaction point. The number of protons contained in each bunch is of the order  $10^{11}$ . The LHC hosts four main particle physics experiments: the CMS, ATLAS, LHCb and ALICE experiments. CMS [5,6] and ATLAS [7–9] are so-called “general purpose experiments”, that are used for a variety of different physics analyses. In contrast, the LHCb [10] and ALICE [11] experiments are designed with an emphasis on CP-violation measurements and heavy ion collisions, respectively.

The number of expected events  $N$  for a given process can be expressed in terms of the corresponding cross section  $\sigma$  times the integrated luminosity  $L$

$$N = L \cdot \sigma, \quad (1.1)$$

with an integrated luminosity of  $L = \int \mathcal{L} dt$ , where  $\mathcal{L}$  is the instantaneous luminosity. The instantaneous luminosity  $\mathcal{L}$  depends on several machine parameters, such as the collision frequency  $f$ , the number of particles in the bunches  $n_1$  and  $n_2$ , the spread in the transverse plane of the bunches  $\sigma_x$  and  $\sigma_y$ , and a geometrical correction parameter  $F$  due to the crossing angle of the two bunches at the interaction point:

$$\mathcal{L} = \frac{f n_1 n_2}{4\pi \sigma_x \sigma_y} \cdot F. \quad (1.2)$$

In 2012, the peak luminosity was  $7.7 \cdot 10^{33} \frac{1}{\text{cm}^2 \text{s}}$ . The total integrated luminosity of  $pp$  collisions over time recorded at the CMS experiment is shown in Fig. 1.2. Events recorded

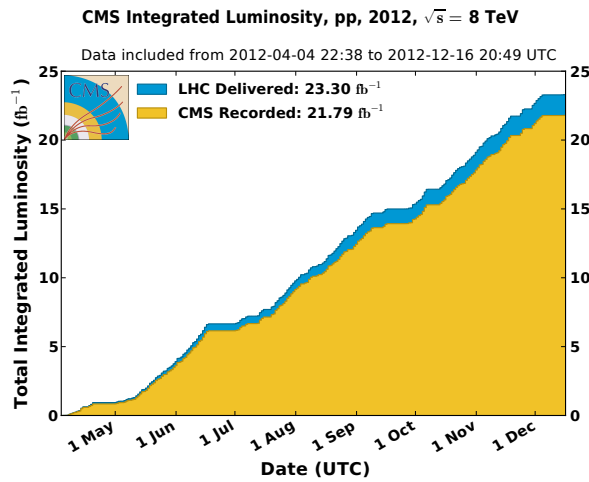


Figure 1.2: Integrated luminosity delivered by LHC (blue) and recorded by CMS (orange) in the year 2012. Taken from [12].

at the CMS experiment are grouped together within “luminosity blocks”. They refer to short intervals of data taking ( $\sim 2\text{--}3$  min) in which the instantaneous luminosity is stable. In addition, luminosity blocks are comprised to so-called “runs”. They correspond to periods of data taking over which the overall conditions remain almost stable.

## 1.2 The CMS detector

The Compact Muon Solenoid (CMS) detector [5,6] is a general purpose detector, designed to explore particle physics phenomena up to the multi-TeV scale. The detector concept is an onion-like structure of different layers, each one made up of a different type of detector. The CMS detector measures 21.6 m in length and 14.6 m in diameter with a total weight of 12 500 tons. In Fig. 1.3, a perspective view of the CMS detector is depicted.

The coordinate system used at CMS consists of the pseudorapidity  $\eta = -\ln[\tan(\theta/2)]$  and the azimuthal angle  $\phi$ . The advantage of the pseudorapidity  $\eta$  is the Lorentz invariance with respect to the  $z$ -axis (beam axis). The angle  $\phi$  covers the direction in the  $x - y$  plane (orthogonal to the beam axis). Since the CMS experiment utilises proton-proton collisions, where the inelastic scattering happens on parton level, the initial total momentum of the

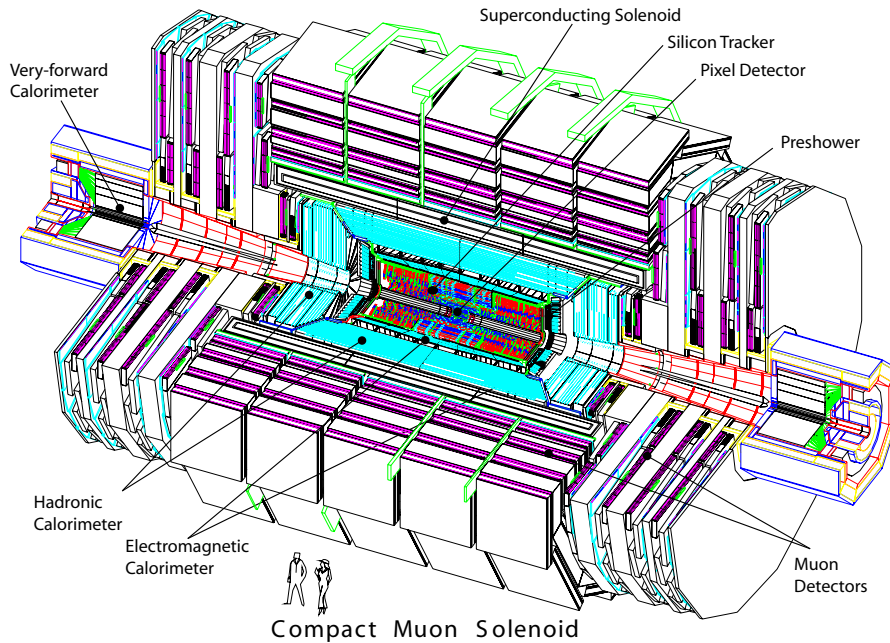


Figure 1.3: A perspective view of the CMS detector. Taken from [5]

corresponding partons are unknown. Thus, instead of exploiting the conservation of total momentum, only the conservation of transverse momentum is usually used. This can be done because the partons have no (or only an negligible amount) of initial momentum in the transverse plane. Thus the initial state is characterised by  $\sum p_T = 0$ , with  $p_T = \sqrt{p_x^2 + p_y^2}$ .

In order to bend the particle trajectories and thus to measure the momentum of charged particles a superconducting solenoid is incorporated between the calorimeter system and the muon system (see Fig. 1.3) providing a uniform axial magnet field of 3.8 T. Iron yokes contained within the muon system ensure the return of the magnetic flux.

In the following, the various detector components of the CMS detector from the inside to the outside as well as the trigger system will be explained.

### 1.2.1 The tracking system

The tracking detector [13–15] is the innermost detector at CMS. It is a silicon semiconductor detector and is included for the tasks of vertex and track reconstruction by the measurement of particles' energy losses. A schematic sketch of the tracker at CMS is depicted in Fig. 1.4. The tracking system is divided into two parts: the innermost tracker is a silicon pixel detector surrounded by a silicon strip detector. Since the silicon pixel tracker is closest to the interaction point and has thus to deal with much higher particle fluxes, its granularity is higher compared to the silicon strip detector.

Both parts will be explained in detail in the upcoming sections. Because, the variable

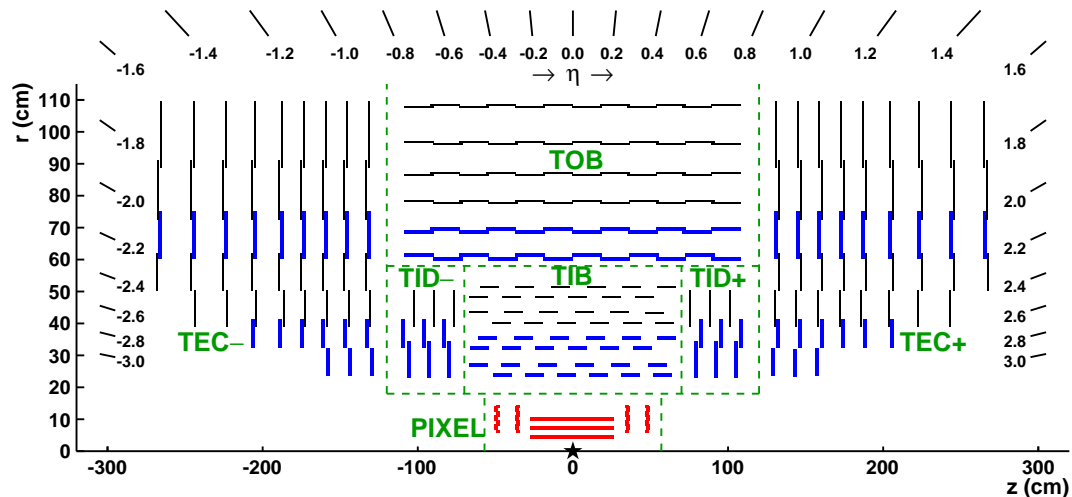


Figure 1.4: Schematic sketch of the silicon tracker at CMS in the  $r - z$  plane including the silicon pixel detector (PIXEL) as well as the different components of the silicon strip detector: tracker inner barrel (TIB), tracker outer barrel (TOB), tracker endcap (TEC), and tracker inner disk (TID). Taken from [16].



$dE/dx$ , the energy loss per path length, is used in this thesis for the search for highly ionising, short tracks (Part ??), the last part of this section is dedicated to a short description of how the energy of a traversing particle is measured with the silicon sensors. Furthermore, since a calibration of the silicon pixel detector was performed within this PhD thesis (see Section ??), an emphasis will be put on the pixel detector.

### The silicon pixel tracker

The silicon pixel detector consists of three different cylindrical layers in the barrel region at radii of 4.4 cm, 7.3 cm and 10.2 cm and two discs in the endcaps at  $z$ -distances of 34.5 cm and 46.5 cm. It is made up of 1440 modules in total (barrel + endcaps), each module comprising 8 or 16 read-out-chips (ROCs). The read-out-chips are bump bonded [17] to a pixel system of  $52 \times 80$  pixels, which are read out in double columns (see [17] on detailed information of the readout electronics). A visualisation of a part of a pixel module is shown in Fig. 1.5. In total, there are 65 million pixels comprised in the pixel detector. The large number of pixels and their small size ensure a low occupancy close to the vertex of around  $0.002 - 0.02\%$  [16] and a high hit efficiency of around 99% [19].

The silicon pixel detector is very important for the reconstruction of primary and secondary vertices as well as the reconstruction of particle tracks. Therefore, a high spatial resolution is needed. This is achieved by the small size of the pixels ( $150 \times 100 \mu\text{m}^2$ ) and the exploitation of the spread of the energy deposition across several pixels, called charge sharing (in average the energy is deposited across 3-5 pixels [20]). Exploiting the energy spread across pixels, a spatial resolution in the barrel region of  $9.4 \mu\text{m}$  in the  $r - \phi$  plane and - dependent on the incident angle of a track - a hit resolution between  $20 - 45 \mu\text{m}$  in the  $z$ -direction is achieved [16]. The spatial resolution of the primary vertex depends on the number of tracks taken into account for the reconstruction of the primary ver-

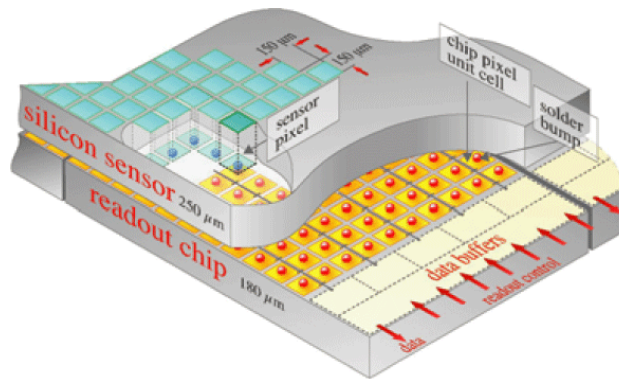


Figure 1.5: Schematic sketch of a part of a silicon pixel tracker module including the silicon sensors and the read-out-chip (ROC). Taken from [18].

tex. For more than 50 tracks originating from the primary vertex the spatial resolution is around  $10 - 12 \mu\text{m}$  for each of the three spatial dimensions [16]. The reconstruction efficiency of primary vertices is close to 100% if more than two tracks are used for the vertex reconstruction [16].

### The silicon strip tracker

The silicon strip tracker is the next-to innermost detector of the CMS detector and ranges up to a radius of 1.1 m. The barrel region consists of a tracker inner barrel (TIB) and a tracker outer barrel (TOB). The TIB has four layers with two layers equipped with so-called “stereo” modules that can measure the hit position additionally in the  $r - z$  plane. The silicon sensors in the TIB are of  $320 \mu\text{m}$  thickness with a strip pitch varying between  $80 - 120 \mu\text{m}$ . The TOB has six different layers (two layers of stereo modules) with silicon sensors of  $500 \mu\text{m}$  thickness and strip pitches between 120 and  $180 \mu\text{m}$ .

The endcaps are subdivided into a tracker endcap (TEC) and a tracker inner disk (TID). They ensure a coverage of a pseudorapidity up to  $|\eta| = 2.5$ . In each TEC, 9 disks between a  $z$ -position of  $120 \text{ cm} < z < 280 \text{ cm}$  are contained. Each of the TID comprises three disks between  $60 \text{ cm} < z < 110 \text{ cm}$ .

In the barrel region, a single-point resolution between  $23 - 52 \mu\text{m}$  in the  $r - \phi$  plane and  $230 - 530 \mu\text{m}$  in the  $z$ -direction is achieved.

### Energy measurements in the tracking system

A charged particle traversing the above mentioned silicon detectors produces electron-hole pairs in the semiconducting material along its trajectory, thus losing energy due to ionisation. For silicon, the mean energy to create an electron-hole pair is  $3.61 \text{ eV}$  at  $-10^\circ\text{C}$ . Minimally ionising particles produce an average of 22 000 electron-hole pairs in silicon sensors [17]. Electrons that are subject to a hard collision with the incoming particle (so-called delta-rays), produce further ionisation and can thus lead to much higher energy deposits in the silicon. Because of the applied bias voltage at the sensors (for the creation of a depletion zone), the released electrons (holes) travel to the n-contacts (p-contacts), thereby inducing a current proportional the released energy, which is measured by the readout electronics. A more detailed description of the energy measurement in silicon sensors can be found in [17].

### 1.2.2 The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) [13,21] encloses the tracking system and starts at a radius of 129 cm in the barrel region. It measures the energy and position of a traversing particle. It is divided into a barrel part and two endcaps, which are at a distance of 314 cm

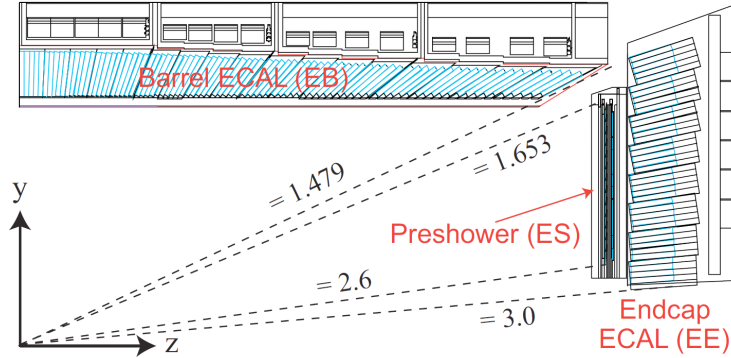


Figure 1.6: A quarter section of the ECAL in the  $r - z$  plane. Taken from [13].

from the vertex. Figure 1.6 depicts a schematic sketch of the electromagnetic calorimeter system in the transverse plane. It can be seen, that the ECAL barrel (EB) covers a pseudorapidity region up to  $|\eta| = 1.479$ . The ECAL endcaps (EE) start at  $|\eta| = 1.653$  and reach up to  $|\eta| = 3.0$ . In front of the endcaps, a preshower detector ( $1.653 < |\eta| < 2.6$ ) is installed with the main task to identify neutral pions. It additionally improves the position measurement of electrons and photons.

The EB and EE consist of lead tungstate ( $\text{PbWO}_4$ ) scintillating crystals, 61200 in the barrel region and 7324 in each of the endcaps. Their advantage is the short radiation length ( $X_0 = 0.89 \text{ cm}$ ) and a small Molière radius (2.2 cm). Thus, particles deposit their energy on rather short distances and a compact design is possible. Since they lack a fine longitudinal segmentation, the information on the shower development is limited. To detect the rather low light yield ( $30 \gamma / \text{MeV}$ ) of a traversing particle, silicon avalanche photodiodes (APDs) are used in the barrel region and vacuum phototriodes (VPTs) in the endcaps. For information on the readout electronics, the reader is referred to [13].

The resolution of an energy measurement in the calorimeter can be expressed by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \quad (1.3)$$

with  $S$  referring to the stochastic term,  $N$  to the noise term, and  $C$  to a constant term. For the ECAL at CMS the parameters of Eq. (1.3) are measured to  $S = 3.63 \sqrt{\text{GeV}}$ ,  $N = 0.124 \text{ GeV}$ , and  $C = 0.26$  [13]. These numbers lead to an energy resolution of around 0.4% for an electron with  $E \approx 200 \text{ GeV}$  and around 0.6% for an electron with  $E \approx 50 \text{ GeV}$ .

### 1.2.3 The hadronic calorimeter

The hadronic calorimeter (HCAL) [13,22] of the CMS detector is splitted into four different detector modules: the hadron barrel (HB), the hadron outer (HO), the hadron endcap

(HE) and the hadron forward (HF) calorimeters. A schematic sketch is depicted in Fig. 1.7.

The HCAL is dedicated to measuring the energy of hadrons as well as improving the estimate of the missing energy in the event because of its large pseudorapidity coverage. The latter one is achieved by the high pseudorapidity coverage ( $|\eta| < 5.0$ ) that assures the detection of most of the visible particles.

The HCAL is a so-called sampling calorimeter which consists of brass absorber material, initiating the hadronic shower, as well as active plastic scintillators. The emitted photons are read out with wavelength-shifting (WLS) fibres which are embedded into the scintillators. These in turn are connected to fibres that transfer the light to the readout system.

The hadron barrel (HB) covers the pseudorapidity range between  $-1.4 < \eta < 1.4$ . It is composed of 17 layers of absorber material (15 brass and 2 steel layers) interleaved with scintillator layers. The scintillator layers are divided into 2304 towers with a size of  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ .

The hadron outer (HO) covers a pseudorapidity range up to  $|\eta| = 1.26$ . It is located between the solenoid and the barrel detector of the muon system. The HO is dedicated to measuring the energy of the shower leakage of hadrons. Its thickness corresponds to over ten interaction lengths.

The hadron endcap (HE) extends the pseudorapidity coverage of the HCAL up to  $|\eta| = 3.0$  and starts at  $|\eta| = 1.3$ . It consists of 2304 towers in total, which vary in size between  $5 - 10^\circ$  in the  $\phi$  direction and  $0.087 - 0.35$  in  $\eta$  direction.

Finally, the hadron forward (HF) calorimeter extends the pseudorapidity range up to  $|\eta| = 5.0$ , starting from  $|\eta| = 3.0$ . It is build out of steel plates with  $1 \text{ mm}^2$  grooves contain-

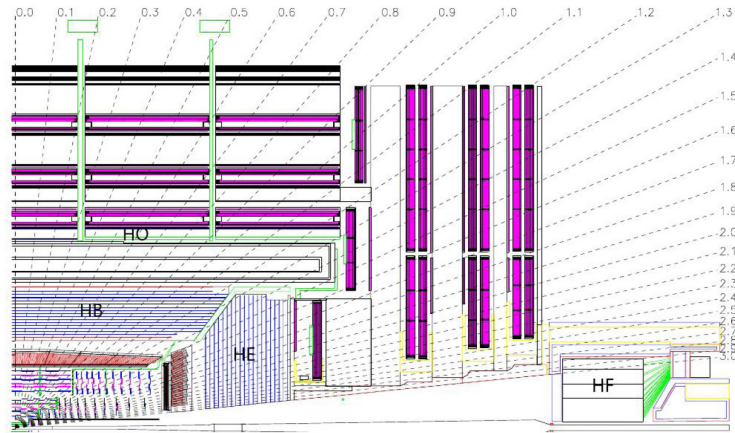


Figure 1.7: A quarter section of the CMS detector including the HCAL in a transverse view. Taken from [5].

ing quartz fibres. The emitted light by the quartz fibres is transferred to photomultipliers. The HF is divided into 13 towers where almost all towers have a spread of  $\Delta\eta \approx 0.175$  in  $\eta$  direction and  $\sim 10^\circ$  in  $\phi$  direction.

### 1.2.4 The muon system

The muon system [13, 23] is the outermost component of the CMS detector. It comprises three different types of gaseous detectors, mounted into the iron return yokes: drift tube (DT) chambers in the barrel region ( $|\eta| < 1.2$ ), cathode strip chambers (CSC) in the endcap region ( $1.04 < |\eta| < 2.4$ ) and resistive plate chambers (RPC) in the barrel as well as the endcap region ( $|\eta| < 1.6$ ) (see Fig. 1.8 for a schematic sketch of the muon system). In the barrel part of the muon system, four layers (so-called stations) of drift-tube chambers are assembled inside the iron return yoke layers at radii of 4.0, 4.9, 5.9 and 7.0 m from the beam axis. The position of a muon traversing these layers can be measured with a precision of  $\approx 100 \mu\text{m}$  in radial direction and  $\approx 1 \text{ mrad}$  in  $\phi$  direction.

The four endcap disks are made up of 468 cathode strip chambers in total. By measuring the centre-of-gravity, they achieve a spatial resolution of  $\approx 100 - 200 \mu\text{m}$  and an angular resolution of  $\approx 10 \text{ mrad}$  in  $\phi$  direction. They are designed in order to cope with a high particle flux of about  $1 \text{ kHz}/\text{cm}^2$  and a non-uniform magnetic field. As signals can be transferred very fast, they are used for the level-1 trigger system (see Section 1.2.5).

Finally, the resistive plate chambers cover the barrel as well as the endcap region up to a pseudorapidity of  $\eta = 1.6$ . They provide a fast response with a good time resolution enabling the exact identification of the respective bunch-crossing. It is used for the level-1

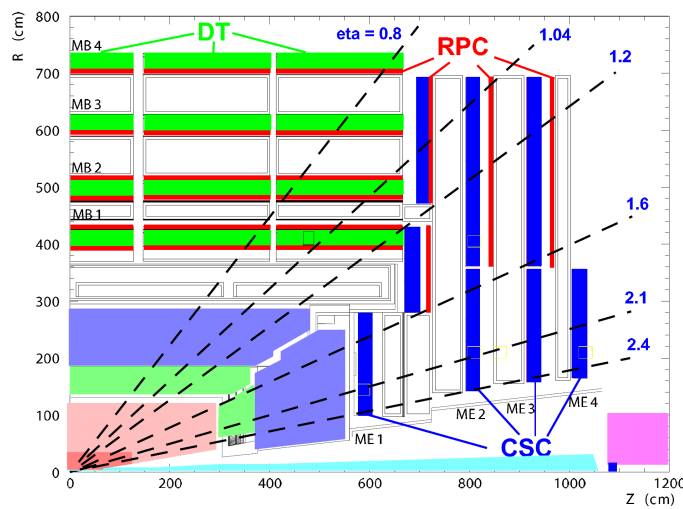


Figure 1.8: A quarter section of the CMS detector in the transverse plane with a detailed view on the muon system. Taken from [13].

trigger system as well.

### 1.2.5 The trigger system

Because of the impossibility of storing each event occurring at the CMS experiment (cf. bunch crossing rate of 20 MHz), a multistage trigger system [13] is used to achieve a drastic reduction of recorded events by nearly six orders of magnitude. It comprises two main parts: a so-called level-1 (L1) trigger system and a high-level trigger (HLT) system.

The L1 triggers need to provide a very fast decision ( $3.2\,\mu\text{s}$ , where around  $1\,\mu\text{s}$  is allocated to the actual L1 trigger calculations) whether an event shall be recorded or not. During this time, the recorded event data is held in buffers located close to the single detector components. Information from the muon system and the calorimeters is used for the L1-trigger decisions. Objects used for these decisions are so-called “trigger primitive” objects: photons, electrons, muons, jets above certain  $E_T$  and  $p_T$  thresholds and global variables like missing transverse energy,  $\cancel{E}_T$ . The design value of the number of events per second that pass this trigger stage is 100 kHz.

After a time of  $3.2\,\mu\text{s}$ , the stored data in the buffers close to the single detector components are transferred to the front-end readout buffers in case the event passed the L1-trigger requirements. By partial event reconstruction and the use of various detector signals (calorimeter, muon information followed by pixel information and full event reconstruction), higher event objects can be used to check whether HLT-trigger requirements are fulfilled. On HLT level, the decision time amounts to 50 ms and a reduction from 100 kHz to 100 Hz of event recording is achieved.

## 1.3 Event reconstruction and particle identification

A crucial ingredient of data analysis at the CMS experiment is the translation of the signal measurements in the various sub-detector components into physical objects, like muons, electrons, etc.. For this translation, i. e. for particle identification, information from all detector components is used. This is known as the particle-flow event reconstruction algorithm [24]. In Fig. 1.9, a slice through the CMS detector is shown with the signatures of different particles indicated as coloured lines.

In the next section an introduction to this algorithm is given, followed by the definition and classification of physics objects at the CMS experiment.

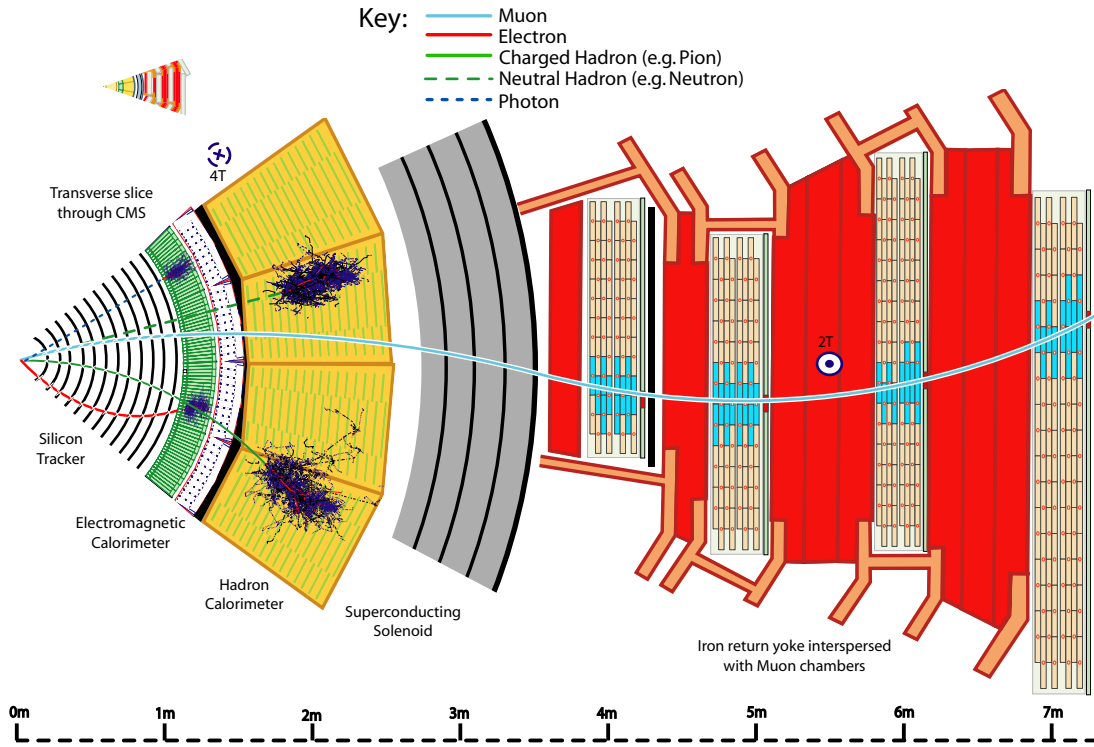


Figure 1.9: A radial slice through the CMS detector with several particle signatures indicated as coloured lines. Taken from [25].

### 1.3.1 The particle-flow algorithm

The particle-flow (PF) event description [24] aims at optimising particle identification and reconstruction by the usage of all sub-detector components of the CMS detector. There are three main building bricks used for global event description: reconstructed charged-particle tracks, calorimeter clusters, and muon tracks. The main requirements for these building bricks are a high reconstruction efficiency and a small fake rate. Therefore, a special emphasis was put on developing a very efficient tracking algorithm (see Section 1.3.2) and a well performing calorimeter clustering algorithm [24].

The particle-flow algorithm proceeds as follows for each event (the description is based on [24]):

1. For each pair of detected building bricks, a “link distance” is calculated in order to quantify the quality of their link, i.e. whether the two building bricks stem from the same particle.
2. “Blocks” are produced from the building bricks that are linked together (with a typical number of one, two or three building bricks contained in a block) using

different algorithms for different sets of building bricks (see [24] for detailed information about these algorithms). Blocks comprise either charged-particle tracks and calorimeter clusters, or several calorimeter clusters, or a charged-particle track in the tracker and a muon track in the muon system. The latter is called a global muon.

3. For each block the following steps are performed:
  - a) Each global muon where the  $p_T$  measured in both sub-detectors is compatible with the  $p_T$  measurement in the tracker is defined as a particle-flow muon and the track in both sub-detectors is removed from the block.
  - b) Next, electrons are reconstructed and identified using tracker hits and ECAL clusters. For an identified particle-flow electron, the corresponding tracker hits and the ECAL clusters (including energy deposits from Bremsstrahlung photons) are removed from the block.
  - c) Tighter track quality criteria are applied.
  - d) The compatibility of the remaining ECAL and HCAL energy deposits to the transverse momentum of the reconstructed tracks within a block is checked. This allows for the identification of particle-flow charged hadrons with a momentum estimate using tracker and calorimeter information. If the energy deposits in the ECAL or HCAL are much larger than the corresponding track  $p_T$ , the signature is identified as a particle-flow photon or particle-flow neutral hadron, respectively. All ECAL and HCAL clusters used for the identification as well as the reconstructed tracks are removed from the block.
  - e) Finally, the remaining ECAL and HCAL clusters (which are all not linked to any other building block) are identified as particle-flow photons or particle-flow neutral hadrons, respectively.

For the final identification of particles and other objects further reconstruction criteria are applied (see Section 1.3.2).

## 1.3.2 Object reconstruction

In this section, an overview of the required criteria for the identification of particles and other physics objects is given.

### 1.3.2.1 Reconstruction of primary vertices

The reconstruction of primary vertices (PVs) is important in order to determine the location of the interaction point and to get an estimate of the corresponding uncertainty. At the CMS experiment, the reconstruction of vertices proceeds as follows [16]: First, a



track selection is performed which depends on the number of hits in the tracker, on the significance of the transverse distance to the beam spot ( $|d_0^{\text{BS}}|/\delta d_0$ ), and the normalised  $\chi^2$  of the track fit (see the subsequent Section 1.3.2.2 for more information on the track reconstruction algorithms). All selected tracks are afterwards clustered to several vertices based on their point of closest approach in  $z$ -direction to the beam spot. The clustering first identifies candidate vertices using a so-called deterministic annealing (DA) algorithm [26]. Subsequently, the candidate vertices are evaluated with the adaptive vertex fitter [27], which estimates the location of the vertices as well as performance parameters of the fit. A weight is assigned to each track that is close to one in case for a good compatibility and close to zero for a bad compatibility with the vertex. The track with the largest sum of all squared track transverse momenta,  $\sum p_{\text{T}}^2$ , is referred to as the primary vertex. All other vertices during a bunch-crossing are so-called pileup interactions.

Further quality criteria of the primary vertex are required in the search for highly ionising, short tracks (Part ??) and the measurement of the jet transverse-momentum resolution (Part ??). For this purpose, the “number of degrees of freedom” of the primary vertex is introduced:

$$n_{\text{dof}} = -3 + 2 \sum_{\text{tracks}} w_i. \quad (1.4)$$

The requirement of a well reconstructed primary vertex is  $n_{\text{dof}} > 4$ . Furthermore, the vertex is required to be within 24 cm in  $z$ - and 2 cm in  $r$ -direction with respect to the nominal interaction point.

### 1.3.2.2 Reconstruction of tracks

The reconstruction of tracks aims at linking several hits in the tracking system to one reconstructed track that matches the original trajectory of the particle with a high probability. With the track reconstruction, an estimate of the particle momentum as well as its position can be achieved. Track reconstruction is challenging because the large number of hits<sup>1</sup>, especially in the layers close to the interaction vertex, leads to a high combinatorial complexity. In the following, an overview of the tracking algorithm used at CMS is given. It is based on [16] and the reader is referred to this reference for more information on the reconstruction of tracks at CMS.

The tracking software used at CMS is usually referred to as the Combinatorial Track Finder (CTF). It is based on the so-called combinatorial Kalman filter [28–30] which in turn is based on the Kalman filter [31]. The Kalman filter is an algorithm that allows for the estimation of parameters of interest based on a set of observations that are subject

---

<sup>1</sup>At design luminosity around 1000 charged particles are expected to traverse the tracker in each bunch-crossing [16]

to noise or other inaccuracies. It is mathematically equivalent to a global least square minimisation for linear models with Gaussian noise.

The basic idea of the tracking algorithm at CMS is to avoid applying the combinatorial Kalman filter on all hits in one step by using an iterative procedure (called iterative tracking). A reduction of complexity can be achieved by first reconstructing tracks that are easy to identify because of e.g. a relatively high  $p_T$ . These tracks are removed afterwards and the remaining tracker hits are subject to further reconstruction iterations. The following iterations are performed (these steps refer to the setting from May to August in 2011 but are in their basic structure retained in the year 2012):

- Iteration 0: Tracks near the  $pp$ -interaction point that have three pixel hits and a  $p_T > 0.8$  GeV are reconstructed.
- Iteration 1: Tracks with only two pixel hits and  $p_T > 0.8$  GeV are reconstructed.
- Iteration 2: Low  $p_T$  tracks from the  $pp$ -interaction point are reconstructed.
- Iteration 3-5: Reconstruction of tracks that are not originating from the primary vertex or that were not found by previous iterations.

Within these iterations, the reconstruction is subdivided into four different steps:

- Seed generation: Only 2-3 hits are used to define track candidates.
- Extrapolation: Based on the expected flight path, additional hits are assigned to the candidate track using a combinatorial Kalman filter.
- Track fitting: With the usage of the Kalman filter and a smoother, the trajectory is fitted in order to estimate the track parameters.
- Setting of quality flags: Quality flags are assigned to all tracks and tracks that fail certain quality criteria are discarded.

A special task when reconstruction tracks consists of the suppression of fake tracks, i.e. tracks that are not associated with a charged particle. Therefore, the requirement of fulfilling certain quality criteria (fourth step) is crucial to substantially reduce the contamination of fake tracks. For this purpose, requirements on the following variables are imposed (the values of the parameters  $\alpha$  and  $\beta$  vary across iterations):

- A minimum number of layers in which the track has an associated hit.
- A minimum number of layers in which the track has an associated 3D hit<sup>2</sup>.

---

<sup>2</sup>A 3D hit refers to a measurement that provides 3D position information, such as a pixel hit or a strip hit in a stereo module.

- A maximum number of layers in which the track has no associated hit.
- A high quality of the trajectory fit:  $\chi^2/\text{ndof} < \alpha_0 N_{\text{layers}}$ .
- A low transverse impact parameter significance:  $|d_0^{\text{BS}}|/\delta d_0 < (\alpha_3 N_{\text{layers}})^\beta$ ,  
 $|d_0^{\text{BS}}|/\sigma_{d_0}(p_T) < (\alpha_1 N_{\text{layers}})^\beta$ .
- A low longitudinal impact parameter significance:  $|z_0^{\text{PV}}|/\delta z_0 < (\alpha_4 N_{\text{layers}})^\beta$ ,  
 $|z_0^{\text{PV}}|/\sigma_{z_0}(p_T, \eta) < (\alpha_2 N_{\text{layers}})^\beta$ .

The variable  $d_0^{\text{BS}}$  is the distance of the track to the centre of the beam spot in the transverse plane to the beam line and  $z_0^{\text{PV}}$  is the distance along the beam line to the closest pixel. The variables  $\delta d_0$  and  $\delta z_0$  are the uncertainties on the distance (impact) parameters and  $\sigma$  is the standard deviation corresponding to the length of the beam spot in z-direction.

### 1.3.2.3 Reconstruction of jets

In Section 1.3.1, the reconstruction of particle-flow neutral and charged hadrons was already introduced. Charged and neutral hadrons are usually formed in the hadronisation process of final state quarks and gluons. Since the interest at particle colliders concerns the sum of all particles arising from one final state gluon or quark, which is referred to as a jet, the reconstruction of jets is a very important task at particle colliders. Foremost, the clustering of the particle-flow candidates to a jet is a crucial ingredient for the reconstruction of jets. Furthermore, the calibration of the measured jet transverse momentum as well as the subtraction of energy due to pileup interactions is important to ensure a good quality  $p_T$  measurement of jets at CMS. All three subjects are explained in detail in the following sections.

#### Clustering of jets

The clustering algorithm used at the CMS experiment belongs to the class of sequential recombination jet algorithms. In sequential recombination algorithms, a distance measure between two particles - or more generally objects - is used to cluster them into jet cones. For this purposes, two quantities are defined

$$d_{ij} = \min \left( p_{T,i}^{2k}, p_{T,j}^{2k} \right) \frac{\Delta_{ij}^2}{R^2}, \quad (1.5)$$

$$d_{iB} = p_{T,i}^{2k},$$

with  $\Delta_{ij}$  referring to the distance between particles  $i$  and  $j$  in the  $\eta - \phi$  plane and  $R$  denoting the radius parameter of the algorithm which determines the cluster scale. The quantity  $k$  can be any integer and specifies the behaviour of the algorithm, as will be seen later. The algorithm is an iterative procedure comprising the following steps: First, the

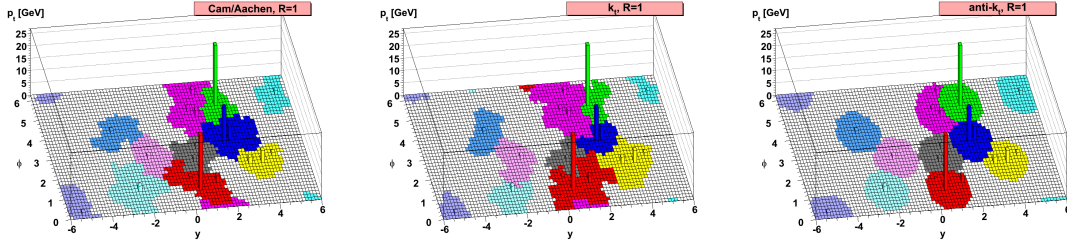


Figure 1.10: Parton-level event in the  $\eta - \phi$  plane with  $\sim 1000$  soft particles added. Three different algorithms are used to cluster the particles into jet cones (coloured areas): the Cambridge/Aachen algorithm (left), the  $k_T$  algorithm (middle), and the anti- $k_T$  algorithm (right). Taken from [36].

quantities  $d_{iB}$  and  $d_{ij}$  are calculated for every single and every pair of “objects” in the event, respectively. An object can be a particle or a set of already clustered particles. If the smallest quantity is a distance between a pair of objects, these objects are clustered together. If the smallest quantity is  $d_{iB}$ , the corresponding object is considered a jet and removed from the object list. If all objects are removed from the input list of the algorithm, the clustering comes to an end.

As noted before, the integer  $k$  in Eq. (1.5) defines different types of algorithms. For the choice of  $k = 0$ , the algorithm is referred to as Cambridge/Aachen algorithm [32, 33]. It doesn’t take the momentum of particles into account for the clustering but is only based on spatial information. The  $k_T$  algorithm [34, 35] uses the configuration with  $k = 1$ , thus starting the clustering with low momentum particles. At CMS, the anti- $k_T$  algorithm is used with  $k = -1$ . It begins with clustering high momentum particles and leads to very regular shapes of the jet cones. The behaviour of the three different clustering algorithm is illustrated in Fig. 1.10. One single parton-level event was subject to the three different jet clustering algorithms. Roughly 1000 soft particles are added to the event in order to test the adaptiveness of the jet algorithms to soft particles. It can be seen, that the Cambridge/Aachen and the  $k_T$  algorithm are both more irregular in shape, while jets clustered with the anti- $k_T$  method result in circular shapes. Furthermore, jets clustered with sequential recombination jet algorithms are infrared and collinear safe.

Throughout this thesis jets clustered with the anti- $k_T$  algorithm with a radius parameter of  $R = 0.5$  are used.

### Jet energy calibration

In order to account for possible mismeasurements of the jet transverse momentum, all jets at CMS are subject to a calibration procedure. This calibration aims at eliminating discrepancies between the measured jet  $p_T$  and the true jet  $p_T$  because of pileup interac-

tions (level-1 corrections) as well as discrepancies depending on jet  $p_T$  and  $\eta$  (level-2 and level-3 corrections). The calibration factors of the level-1, level-2 and level-3 corrections are determined fully in simulation. To ensure also a complete calibration of jets in data, additional correction factors are determined which are only applied on the jet  $p_T$  measured in data (level-2 and level-3 residual corrections). These calibration factors are determined with the help of reference objects in data, such as photons. An overview of the applied correction factors in simulation and data is shown in Fig. 1.11.

In the following, a more detailed description of the determination of the factorised calibration factors will be given. For an extensive description of the calibration approach, the reader is referred to [38]

**L1 pileup offset corrections:** In order to eliminate a mismeasurement of the jet  $p_T$  because of pileup interactions, correction factors ( $C_{\text{offset}}$ ) are determined that are supposed to restore the true jet  $p_T$ . This correction is determined in bins of jet  $p_T$ , jet  $\eta$ , the jet area  $A$ , and the event energy density  $\rho$ . The correction factors are determined with simulated samples and are applied to measured data as well as simulated events.

**L2 relative + L3 absolute corrections:** These two corrections are devoted to correct the measured jet  $p_T$  in order to ensure a uniform response ( $p_T^{\text{reco. jet}}/p_T^{\text{true jet}}$ ) vs.  $\eta$  (L2) ( $C_{\text{rel}}$ ) and  $p_T$  (L3) ( $C_{\text{abs}}$ ). They are also determined in simulated samples and applied to simulated and measured events.

**L2L3 residual corrections:** Finally, correction factors ( $C_{\text{res}}$ ) are determined in order to remove any residual bias in data. They are estimated in data with the help of

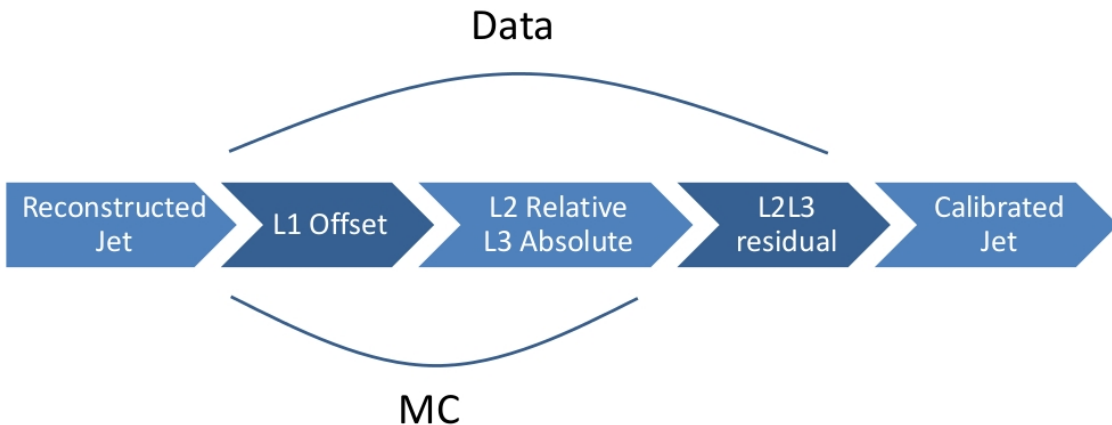


Figure 1.11: Visualisation of the various steps of the jet energy calibration at the CMS experiment. The lines embrace the applied calibration steps for simulation (MC) and data. Taken from [37].

QCD-multijet events to ensure a uniform response vs.  $\eta$  (L2) and with  $\gamma/Z$ + jets events in order to ensure a uniform response vs.  $p_T$  (L3). These factors are not applied to simulation.

The derived correction factors are applied on the measured transverse momentum in a factorised approach according to the following equation:

$$p_T^{\text{cor}} = C_{\text{res}}(p_T'', \eta) \cdot C_{\text{rel}}(\eta) \cdot C_{\text{abs}}(p_T') \cdot C_{\text{offset}}(p_T^{\text{raw}}, \eta) \cdot p_T^{\text{raw}}, \quad (1.6)$$

where  $p_T^{\text{cor}}$  is the corrected jet  $p_T$  and  $p_T^{\text{raw}}$  is the fully uncorrected jet  $p_T$ . By the application of the various correction factors, the jet transverse momentum is calibrated in a factorised manner.

Throughout this thesis, all jets are calibrated by the application of the jet transverse momentum correction factors, referred to as jet energy corrections.

### Charged hadron subtraction

Particles originating from pileup vertices (additional vertices during a bunch crossing) can be clustered into jets leading to higher energies of the reconstructed jet. Therefore a procedure called charged hadron subtraction (CHS) is applied. It removes reconstructed charged particles that do not originate from the primary vertex (see [39] for more details).

In this thesis, all jets are subject to charged hadron subtraction.

#### 1.3.2.4 Reconstruction of photons

The following description of the photon reconstruction is based on [40], where also a more detailed explanation of the photon reconstruction algorithms can be found.

Photons are reconstructed from energy deposits in the ECAL. The clustering algorithms of the ECAL energy deposits do not differentiate between electrons and photons. Thus, they are the same as used for electron identification (Section 1.3.2.6). So-called “super-clusters” are formed from a “seed crystal” which has an energy deposit greater than all of its neighbouring crystals and above a certain threshold. In the barrel region a so-called “hybrid” algorithm is used which proceeds by adding in  $\phi$  direction fixed arrays of  $5 \times 1$  crystals in  $\eta \times \phi$  if their energy deposits are larger than a certain minimal energy. In the endcap region, the so-called “multi  $5 \times 5$ ” clustering algorithm is used. It proceeds by adding fixed arrays of  $5 \times 5$  crystals if their energy exceeds a certain threshold. These clustering procedures collect energy from radiating electrons as well as converted photons.

After the clustering the measured energy in a supercluster is subject to an energy correction procedure [40]. The energy estimate is finally based on the variable  $R_9$ . It is the ratio of the energy deposit within a  $3 \times 3$  array of crystals around the seed crystal

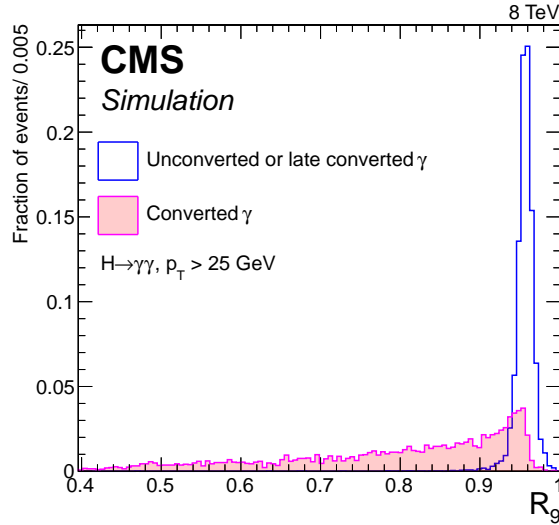


Figure 1.12: Normalised distribution of the  $R_9$  variable for converted and unconverted (or late converted) photons. Taken from [40]

divided by the total energy in the supercluster. This variable nicely discriminates between converted and unconverted photons as can be seen in Fig. 1.12. The energy spread is wider for converted photons than for unconverted photons. Thus, the energy is estimated as the sum of energy deposits in the superclusters for converted photons ( $R_9 < 0.94$  in the barrel and  $R_9 < 0.94$  in the endcap region) and as the sum of energy deposits in the  $5 \times 5$  crystal matrix for the unconverted photons.

### 1.3.2.5 Reconstruction of muons

There are three different muon definitions at CMS [41]: global muons, tracker muons, and standalone muons. They all have in common that they require energy deposits in the muon system. The reconstruction of each of the three muon types is explained in the following.

**Standalone muons:** For the reconstruction of standalone muons, all reconstructed segments in the muon system are utilised. Similar to the track reconstruction in the tracking system, Kalman filter techniques [31] are exploited to reconstruct muon trajectories in the muon chambers. A compatibility to the interaction point is imposed to reconstruct only muons produced at the LHC (no cosmic muons). Further details about the reconstruction of standalone muons can be found in [13, 42]

**Tracker muons:** To reconstruct so-called tracker muons, all tracker tracks with a  $p_T > 0.5$  GeV and  $p > 2.5$  GeV are extrapolated to the muon system. If at least one muon

segment is matched to a reconstructed track fulfilling certain quality criteria, the trajectory is considered as a tracker muon (see [41] for more detailed information).

**Global muons:** For the reconstruction of global muons an outside-in approach is utilised. For each reconstructed standalone muon, the compatibility to the reconstructed tracks in the tracking system is checked. If compatible, a global muon track is reconstructed using Kalman filter techniques. For high- $p_T$  muons, the momentum resolution can be increased this way compared to the momentum estimated using tracker information only [41].

### 1.3.2.6 Reconstruction of electrons

The reconstruction of electrons at the CMS experiment is based on a mixture of the particle-flow algorithm explained in Section 1.3.1 and a standalone approach [43]. Thus, it is a very complex procedure and only the basic approach shall be explained here. The reader is referred to [44] for a complete description of the reconstruction procedure.

Electrons are identified using clustered energy deposits in the ECAL and reconstructed tracks in the tracker system. Both inputs are tested on compatibility and finally used to estimate the electron properties.

The difficulty of the electron reconstruction lies in the possibly large energy losses due to bremsstrahlung. This can change the direction of the electron significantly and lead to a reduced efficiency of the standard track reconstruction as well as a reduced efficiency of the association of ECAL clusters to the reconstructed track in the inner tracking system. Therefore, an optimised track reconstruction for electrons is additionally performed in order to account for direction changes due to the radiation of photons. Because the dedicated electron track reconstruction can be very time consuming, the seeding of tracker hits relies already on ECAL information to reduce the number of candidate tracks (ECAL-based seeding).

Therefore, the electron reconstruction proceeds in detail as follows. First, ECAL clusters, so-called superclusters, are formed using a “hybrid” algorithm in the barrel region and a “multi- $5 \times 5$ ” algorithm in the endcaps. It is the same clustering used for photons (see Section 1.3.2.4 where the more information about these algorithms can be found). Superclusters are used, since due to bremsstrahlung, the energy deposits by electrons are usually spread out over several crystals. Additionally, so-called PF clusters are formed that add all neighbouring crystals to the above mentioned superclusters. Afterwards, the standard track reconstruction is performed which is able to reconstruct electron tracks when bremsstrahlung is negligible. Very short reconstructed tracks or tracks with a bad  $\chi^2$  are refitted using a so-called Gaussian sum filter (GSF) [45] to account for electrons with high radiative energy losses. As mentioned above, these tracks use ECAL-based seeding to reduce the number of candidate tracks. To identify these refitted tracks as candidate



electron tracks, a multivariate analysis is performed using tracker and ECAL information. All ECAL PF clusters tangent to the electron track are added in order to capture the energy of bremsstrahlung photons. Finally, track-cluster association criteria are applied in order to reduce misidentification while preserving a high reconstruction efficiency. The association criteria rely on spatial distances between the superclusters and the tracker track for ECAL-seeded tracks (GSF tracks) and rely on multivariate techniques for the tracker-seeded (standard track reconstruction) electrons.

### 1.3.2.7 Reconstruction of taus

The tau reconstruction at CMS refers to the reconstruction of hadronically decaying taus (a fraction of around 65% of taus decays hadronically). Leptonically decaying taus can be identified through the reconstruction of the muon or the electron in the final state. In the following a tau refers therefore to hadronically decaying taus only. At CMS, the main algorithm used to identify and reconstruct taus is the so-called hadron plus strip algorithm (HPS) [46, 47], which is also used in this thesis. The description of the tau reconstruction algorithm within this thesis follows [46].

The HPS algorithm consists of two separate steps. First, reconstructed PF particles are combined in a way that they are compatible with any of the tau decay modes and the four-momentum of this combined object is estimated. Second, multivariate techniques are used in order to discriminate the candidate taus against decays from gluon or quark jets and from muons and electrons. The HPS algorithm starts from reconstructed jets with a radius parameter of  $R = 0.5$  and  $p_T > 14 \text{ GeV}$  and  $|\eta| < 2.5$ .

The successful identification of a tau requires the reconstruction of a neutral pion. A neutral pion is present in most of the hadronic decays and decays into two photons with a probability of  $\sim 99\%$ . The photons in turn convert into electron pairs with a high probability inside the tracker volume. Therefore, photons and electrons are clustered into strips in the  $\eta - \phi$  plane in order to identify possible  $\pi^0$ . The initial position of the strip equals the position of the seed electron or photon, i.e. the photon or electron with the highest  $p_T$  not yet clustered into any strip. The next-highest  $p_T$  electron/photon within a certain distance ( $\eta \times \phi = 0.05 \times 0.20$ ) to the original strip is merged into the strip. Afterwards, the strip position is recalculated and the step is repeated. The clustering stops if no more electrons/photons are found.

After the reconstruction of the neutral pion, multiple decay mode hypotheses are tested by combining up to two clustered strips with one or three charged PF hadrons. For all hypotheses, several restrictions on the invariant mass of the hadronically decaying tau are imposed. Finally, the hypothesis fulfilling the invariant mass requirement and - in case this happens for more than one hypothesis - the one with the highest  $p_T$  tau is chosen.

Furthermore, for all taus in this thesis, the discriminators used for the reduction of jet,

electron and muon contamination are required to be zero (see [46] for further details on the discriminators).

Finally, a loose isolation criterion is imposed, which is calculated by the pileup corrected scalar sum of the transverse momentum of all charged particles and photons with a  $p_T$  larger than 0.5 GeV in a cone of  $R = 0.5$  around the tau.

### 1.3.2.8 Reconstruction of missing transverse energy

The particle-flow missing transverse energy, PF  $\cancel{E}_T$ , in an event is defined as the negative vectorial sum of all particle-flow particles'  $p_T$ . Thus - neglecting momentum mismeasurements - it refers to the transverse momentum of all only weakly interacting particles in the event. In order to avoid negative effects from e.g. tracker inefficiencies, non-linear calorimeter responses, etc.,  $\cancel{E}_T$  can be measured more precisely by correcting the transverse momentum of all jets contained in an event with the jet energy corrections (see Section 1.3.2.3).

For a more detailed explanation of  $\cancel{E}_T$  as well as for information on the performance of the reconstruction, the reader is referred to [48]

### 1.3.3 Event cleaning

The two analyses described in this thesis are subject to an event cleaning procedure as recommended by the CMS collaboration [49]. It refers to the application of several filters to remove events with detector signals not caused by particles from the hard interaction. In the following, the filters are listed and a short description will be given.

**HBHE noise filter:** This filter rejects anomalous HCAL noise in the barrel (HB) and in the endcap (HE). The noise in the HCAL is mainly related to the hybrid photodiodes (HPD) and readout boxes (RBX). It rejects events considering pulse shape information and anomalous high hit multiplicities in an HPD or RBX.

**CSC beam halo filter:** Particles contained in the beam halo can induce signals in the detector. Therefore, events with a muon moving parallel to the beam are rejected.

**HCAL laser event filter:** This filter provides a list of events, where the laser used for the HCAL calibration fired when a bunch crossing occurred. Detector signals from the laser can be misinterpreted as particle signals and, therefore, these events are removed.

**ECAL dead cell filter:** Events with jets pointing towards a non-working ECAL cell are rejected. This filter relies on information of the surrounding crystals as well as on “trigger-primitive” information.

**EE bad supercrystal filter:** In the ECAL endcap, two supercrystals were not working properly during the data taking period in 2012. Therefore, events where at least one of the two crystals detected unnaturally high energy deposits are removed.

**Tracking failure filter:** Sometimes, unnaturally high energy deposits are recorded in the calorimeters without corresponding tracks in the inner tracking system. The reconstruction of tracks can fail, because of too many tracker seeds or because the hard interaction was too far away from the interaction point. This filter rejects therefore events, where less than 10% of the energy in the event was measured by the tracker.

**Scraping filter (data only):** When there are at least ten tracks in the event, at least 25% of them need to be flagged as high-purity tracks as defined in [50].

**Tracking POG filter:** Events with aborted track reconstruction and/or strip tracker noise are rejected.

## 1.4 Simulation of events

A very important input for physics analyses at particle colliders is the use of simulated collision events. They serve as a comparison to measured data and can give an estimation of the expected number of background and signal events for a given selection. Furthermore, they are used to study event kinematics for Standard Model (SM) and beyond-SM processes.

The generation of simulated samples is done with the help of Monte Carlo (MC) methods (see e. g. [51]). MC techniques are used to randomly sample points according to a specified probability distribution. The underlying probability distributions for the event generation are the transition probabilities from the initial into all possible final states.

Besides the simulation of the hard interaction of the event, also the subsequent showering and hadronisation as well as the detector response needs to be simulated. In the following sections, a short introduction into the various steps of event simulation is given. For more information, the reader is referred to [52].

The first step of the generation of  $pp$ -collision events consists of the simulation of the hard process, i. e. the primary interaction of two partons of the two colliding protons. Which types of partons are involved in the primary interaction is specified by so-called

parton distribution functions (PDFs). They describe the probability of a parton (up-quarks, down-quarks, gluons, sea quarks) to carry a certain energy fraction of the proton and thus define the probability for each parton to take part in the hard interaction process.

After the simulation of the hard interaction between the selected partons, the showering and hadronisation of the particles is simulated. Throughout this thesis, the matrix element generator MADGRAPH [53] plus the event generator PYTHIA [54] or the event generator PYTHIA only are used. Since MADGRAPH only provides the generation of matrix elements, the subsequent showering and hadronisation is done by PYTHIA. For this task, particles produced in the hard interaction are matched to the particles produced in the showers in order to avoid a double-counting of emissions in multijet processes. Since MADGRAPH is responsible for the simulation of well separated partons whereas PYTHIA is designed for simulating the soft and collinear showers, the event is rejected if a parton from the matrix element generator cannot be matched within a certain distance  $\Delta R$  to any of the simulated showers (MLM matching scheme [55]). The showers are clustered with the  $k_T$  method (see Section 1.3.2.3). Additionally, events are rejected if there are more showers than partons in the event.

The formation of colour neutral particles, called hadronisation, and the subsequent decay of unstable particles is simulated with PYTHIA for all samples used in this thesis. A visualisation of the various steps performed for the generation of events is shown in Fig. ??.

In order to account for signals caused by the interaction of the non-primary partons of the interacting protons - called the underlying event (UE) - various models (tunes) are exploited. They rely on the measured signals in data in proton collisions without any hard interaction. Detector signals caused by additional interactions (pileup interactions) during a bunch-crossing, are also mixed into every simulated sample.

Finally, the interaction of the traversing particles with the detector material as well as the response and the readout electronics is simulated with GEANT4 [56,57] throughout this thesis.

## Bibliography

- [1] L. Evans and P. Bryant, “LHC Machine”, *JINST* **3** (2008) S08001.  
doi:10.1088/1748-0221/3/08/S08001.
- [2] O. S. Bruning, P. Collier, P. Lebrun et al., “LHC Design Report Vol.1: The LHC Main Ring”, *Technical Design Report* **CERN-2004-003** (2004).
- [3] “LEP Design Report: Vol.2. The LEP Main Ring”,.
- [4] CERN, “LHC: the guide”, *CERN Brochure* **CERN-Brochure-2009-003-Eng** (2009).
- [5] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004. doi:10.1088/1748-0221/3/08/S08004.
- [6] CMS Collaboration, “CMS physics: Technical design report”, *Technical Design Report* **CERN-LHCC-2006-001, CMS-TDR-008-1** (2006).
- [7] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* **3** (2008) S08003. doi:10.1088/1748-0221/3/08/S08003.
- [8] ATLAS Collaboration, “ATLAS detector and physics performance: Technical Design Report. Volume 1”, *Technical Design Report* **CERN-LHCC-99-14, ATLAS-TDR-14** (1999).
- [9] ATLAS Collaboration, “ATLAS detector and physics performance: Technical Design Report. Volume 2”, *Technical Design Report* **CERN-LHCC-99-15, ATLAS-TDR-15** (1999).
- [10] LHCb Collaboration, “The LHCb Detector at the LHC”, *JINST* **3** (2008) S08005. doi:10.1088/1748-0221/3/08/S08005.
- [11] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST* **3** (2008) S08002. doi:10.1088/1748-0221/3/08/S08002.
- [12] CMS Collaboration, “CMS Luminosity - Public Results”, *Public CMS Wiki* (2015).  
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>, Topic revision: r111.

- [13] CMS Collaboration, “CMS Physics: Technical Design Report Volume 1: Detector Performance and Software”, *Technical Design Report* **CERN-LHCC-2006-001,CMS-TDR-8-1** (2006).
- [14] CMS Collaboration, “The CMS tracker system project: Technical Design Report”, *Technical Design Report* **CERN-LHCC-98-006,CMS-TDR-5** (1997).
- [15] CMS Collaboration, “The CMS tracker: addendum to the Technical Design Report”, *Technical Design Report* **CERN-LHCC-2000-016** (2000).
- [16] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014), no. 10, P10009, [arXiv:1405.6569](https://arxiv.org/abs/1405.6569). doi:10.1088/1748-0221/9/10/P10009.
- [17] J. A. Sibille, “Radiation Hard Hybrid Pixel Detectors, and a  $b\bar{b}$  Cross-Section Measurement at the CMS Experiment”. PhD thesis, Hamburg University, 2013. <http://cds.cern.ch/record/1569832>.
- [18] L. Taylor, “Silicon Pixels”. <http://cms.web.cern.ch/news/silicon-pixels>. Accessed: 2016-02-12.
- [19] CMS Collaboration, “Offline calibrations and performance of the CMS pixel detector”, *Nucl. Instrum. Meth.* **A650** (2011) 25–29. doi:10.1016/j.nima.2010.11.188.
- [20] CMS Collaboration, “Pixel Performance plots for 2012 data and simulation”, *Public CMS Wiki*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PixelOfflinePlots2013>, Topic revision: r22.
- [21] CMS Collaboration, “The CMS electromagnetic calorimeter project: Technical Design Report”, *Technical Design Report* **CERN-LHCC-97-033,CMS-TDR-4** (1997).
- [22] CMS Collaboration, “The CMS hadron calorimeter project: Technical Design Report”, *Technical Design Report* **CERN-LHCC-97-031,CMS-TDR-2** (1997).
- [23] CMS Collaboration, “The CMS muon project: Technical Design Report”, *Technical Design Report* **CERN-LHCC-97-032,CMS-TDR-3** (1997).
- [24] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-001** (2009).

- [25] J. Nielsen, “Fundamentals of LHC Experiments”, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2010). String Theory and Its Applications: From meV to the Planck Scale*, pp. 127–152. 2011. [arXiv:1106.2516](#).
- [26] K. Rose, “Deterministic Annealing for Clustering, Compression, Classification, Regression, and Related Optimization Problems”, in *Proceedings of the IEEE*, pp. 2210–2239. 1998.
- [27] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, “Adaptive vertex fitting”, *J. Phys.* **G34** (2007) N343. [doi:10.1088/0954-3899/34/12/N01](#).
- [28] P. Billoir, “Progressive track recognition with a Kalman-like fitting procedure”, *Comput. Phys. Commun.* **57** (1989) 390–394. [doi:10.1016/0010-4655\(89\)90249-X](#).
- [29] P. Billoir and S. Qian, “Simultaneous pattern recognition and track fitting by the Kalman filtering method”, *Nucl. Instrum. Meth.* **A294** (1990) 219–228. [doi:10.1016/0168-9002\(90\)91835-Y](#).
- [30] R. Mankel, “A concurrent track evolution algorithm for pattern recognition in the HERA-B main tracking system”, *Nucl. Instrum. Meth.* **A395** (1997) 169–184. [doi:10.1016/S0168-9002\(97\)00705-5](#).
- [31] R. Fruhwirth, “Application of Kalman filtering to track and vertex fitting”, *Nucl. Instrum. Meth.* **A262** (1987) 444–450. [doi:10.1016/0168-9002\(87\)90887-4](#).
- [32] Y. L. Dokshitzer, G. D. Leder, S. Moretti et al., “Better jet clustering algorithms”, *JHEP* **08** (1997) 001, [arXiv:hep-ph/9707323](#). [doi:10.1088/1126-6708/1997/08/001](#).
- [33] M. Wobisch and T. Wengler, “Hadronization corrections to jet cross-sections in deep inelastic scattering”, in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*. 1998. [arXiv:hep-ph/9907280](#).
- [34] S. Catani, Y. L. Dokshitzer, M. H. Seymour et al., “Longitudinally-invariant  $k_{\perp}$ -clustering algorithms for hadron-hadron collisions”, *Nucl. Phys.* **B406** (1993) 187–224. [doi:10.1016/0550-3213\(93\)90166-M](#).
- [35] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions”, *Phys. Rev.* **D48** (1993) 3160–3166, [arXiv:hep-ph/9305266](#). [doi:10.1103/PhysRevD.48.3160](#).

- [36] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_t$  jet clustering algorithm”, *JHEP* **04** (2008) 063, [arXiv:0802.1189](#). doi:10.1088/1126-6708/2008/04/063.
- [37] K. Goebel, “Probing supersymmetry based on precise jet measurements at the CMS experiment”. PhD thesis, Hamburg University, 2015.  
<http://www-library.desy.de/cgi-bin/showprep.pl?thesis15-003>.
- [38] CMS Collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS”, *JINST* **6** (2011) P11002, [arXiv:1107.4277](#). doi:10.1088/1748-0221/6/11/P11002.
- [39] CMS Collaboration, “Pileup measurement and mitigation techniques in CMS”, *J. Phys. Conf. Ser.* **404** (2012) 012045. doi:10.1088/1742-6596/404/1/012045.
- [40] CMS Collaboration, “Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *JINST* **10** (2015), no. 08, P08010, [arXiv:1502.02702](#). doi:10.1088/1748-0221/10/08/P08010.
- [41] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV”, *JINST* **7** (2012) P10002, [arXiv:1206.4071](#). doi:10.1088/1748-0221/7/10/P10002.
- [42] CMS Collaboration, “Performance of CMS muon reconstruction in cosmic-ray events”, *JINST* **5** (2010) T03022, [arXiv:0911.4994](#). doi:10.1088/1748-0221/5/03/T03022.
- [43] S. Baffioni, C. Charlot, F. Ferri et al., “Electron reconstruction in CMS”, *Eur. Phys. J. C* **49** (2007) 1099–1116. doi:10.1140/epjc/s10052-006-0175-5.
- [44] CMS Collaboration, “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *JINST* **10** (2015), no. 06, P06005, [arXiv:1502.02701](#). doi:10.1088/1748-0221/10/06/P06005.
- [45] W. Adam, R. Fruhwirth, A. Strandlie et al., “Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC”, *eConf* **C0303241** (2003) TULT009, [arXiv:physics/0306087](#). [J. Phys.G31,N9(2005)]. doi:10.1088/0954-3899/31/9/N01.
- [46] CMS Collaboration, “Reconstruction and identification of  $\tau$  lepton decays to hadrons and  $\nu_\tau$  at CMS”, *JINST* **11** (2016), no. 01, P01019, [arXiv:1510.07488](#). doi:10.1088/1748-0221/11/01/P01019.
- [47] CMS Collaboration, “Performance of  $\tau$ -lepton reconstruction and identification in CMS”, *JINST* **7** (2012) P01001, [arXiv:1109.6034](#). doi:10.1088/1748-0221/7/01/P01001.



- 
- [48] CMS Collaboration, “Performance of the CMS missing transverse momentum reconstruction in pp data at  $\sqrt{s} = 8$  TeV”, *JINST* **10** (2015), no. 02, P02006, [arXiv:1411.0511](#). doi:10.1088/1748-0221/10/02/P02006.
- [49] CMS Collaboration, “MET Optional Filters”, *Internal CMS Wiki* (2015). <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters>, Topic revision: r66.
- [50] CMS Collaboration, “Tracking and Vertexing Results from First Collisions”, *CMS Physics Analysis Summary CMS-PAS-TRK-10-001* (2010).
- [51] S. Weinzierl, “Introduction to Monte Carlo methods”, [arXiv:hep-ph/0006269](#).
- [52] M. Schröder, “Quality of jet measurements and impact on a search for new physics at CMS”. PhD thesis, Hamburg University, 2012. <http://www-library.desy.de/cgi-bin/showprep.pl?thesis12-042>.
- [53] J. Alwall, R. Frederix, S. Frixione et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079, [arXiv:1405.0301](#). doi:10.1007/JHEP07(2014)079.
- [54] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175](#). doi:10.1088/1126-6708/2006/05/026.
- [55] M. L. Mangano, M. Moretti, F. Piccinini et al., “Matching matrix elements and shower evolution for top-quark production in hadronic collisions”, *JHEP* **01** (2007) 013, [arXiv:hep-ph/0611129](#). doi:10.1088/1126-6708/2007/01/013.
- [56] GEANT4 Collaboration, “GEANT4: A Simulation toolkit”, *Nucl. Instrum. Meth.* **A506** (2003) 250–303. doi:10.1016/S0168-9002(03)01368-8.
- [57] J. Allison et al., “Geant4 developments and applications”, *IEEE Trans. Nucl. Sci.* **53** (2006) 270. doi:10.1109/TNS.2006.869826.