Search for Higgs Decay to Dark Matter and Trigger Studies

João Carlos Arnauth Pela
of Imperial College London

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

Abstract

Here the abstract of the thesis

Declaration

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

João Pela

Acknowledgements

TODO:

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

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

Preface

Thesis structure and so on...

"To my grand mother"

Contents

Chapter 1

Theory

[1]

[2]

TODO:

- Global status

1.1 Standard Model of Particle Physics

TODO:

- Very brief summary of the Standard Model.

The Standard Model (SM) of particle physics is the currently accepted model for describing the physics of elementary particles.

Leptons (J=1/2)				
Generation	Particle Name	Symbol	Mass (GeV/c^2)	Q/e
1^{st}	Electron	e	0.000511	1
	Electron Neutrino	ν_e	$< 3 \times 10^-9$	0
2^{nd}	Muon	μ	0.106	1
	Muon Neutrino	ν_μ	$< 1.9 \times 10^-4$	0
3^{rd}	Tau	τ	1.777	1
	Tau Neutrino	ν_τ	$< 1.82 \times 10^-2$	0

Table 1.1: List of leptons and their fundamental properties

Quarks (J=1/2)				
Generation	Particle Name	Symbol	Mass (GeV/c^2)	Q/e
1^{st}	Up	u	$1.5 - 3.3 \times 10^{-3}$	-2/3
	Down	d	$3.5 - 6 \times 10^{-3}$	1/3
2^{nd}	Charm	c	1.16-1.34	-2/3
	Strange	s	$70 - 130 \times 10^{-3}$	1/3
3^{rd}	Top	t	169-173	-2/3
	Bottom	b	$4.13 - 4.37$	1/3

Table 1.2: List of quarks and their fundamental properties

Bosons			
Particle Name	Mass (GeV/c^2)	Q/e	Spin
Photon (γ)	0	0	1
$W^{\pm}$	80.4	∓ 1	1
Z^0	91.2	0	1
Gloun (g)	0	0	1
Higgs (H^0)	> 114	0	0

Table 1.3: List of bosons and their fundamental properties

1.2 Higgs Mechanism

Summary of the Higgs Mechanism. Should include

- Motivations
- Explanation of the mechanism itself
- Consequences
- Possible decays

1.3 Higgs Invisible decays

TODO:

- Explain what are SM Higgs invisible decays.
- Go over the possibility of BSM invisible decays.

Chapter 2

Experimental Apparatus

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC)[3] is currently the world's largest particle accelerator and is capable to produce the highest energy particle beams ever made by mankind. This gigantic machine with a total perimeter of 26.7 km was built at European Organization for Nuclear Research (CERN) in a circular tunnel, where previously the Large Electron Positron collider (LEP)[4] was installed, at an average depth of 100 m below ground under the Franco-Swiss border near Geneva, Switzerland. A diagram of the LHC tunnel and its experiments can be found at figure 2.1.

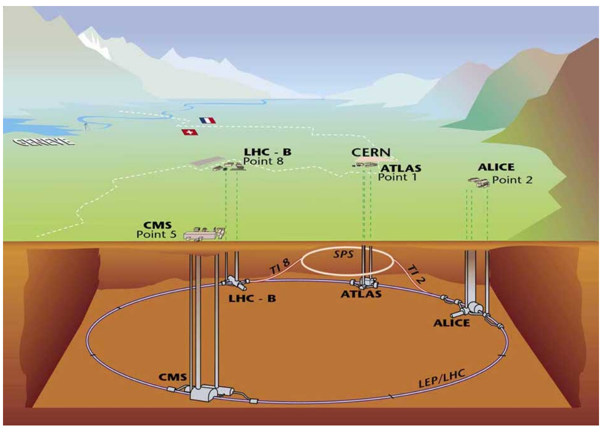


Figure 2.1: Underground diagram of the Geneva area showing the LHC and its experiments location.

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

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

The LHC is only the last element of a complex accelerator chain which step-by-step increases the energy of the particles to eventually be collided. Protons are initially obtained by stripping the electrons of hydrogen gas. This process happens at the begging of the Linear Particle Accelerator 2 (LINAC2) which then accelerates them up to the energy of 50 MeV. After this initial step proton are injected into the Proton Synchrotron Booster (PSB) and the energy ramps ups to 1.4 GeV. Particles are then passed to the Proton Synchrotron (PS) where the energy futher increases to 25 GeV. Subsequently they are the injected into the Super Proton Synchrotron (SPS) where the particle energy level reaches 450 GeV. Finally, protons pass to the LHC where they can be accelerated to a maximum energy of 7 TeV. A simplified diagram of the CERN accelerator chain can be found in figure 2.2. Normal operation of the LHC therefore depends on all the upstream accelerators availability. The typically turn around time, the time necessary to stop the accelerator from running and restart collisions, is around 2 hours. When stable beams are achieved, a single proton fill can be used to collide protons up to 24 hours, but it is common to restart more frequently to take profit of the higher collision rates possible right at the beginning of a new fill.

The LHC as its name indicates can collide hadrons, more specifically proton or heavy ions. Three modes of operation have been tried according to the particles used: proton-proton, proton-lead and lead-lead. By changing the incoming particles we are changing the quantity of nucleons present at each interaction. The maximum design energy per proton is 7 TeV and for each lead nucleon 2.76 TeV. The maximum design luminosity for proton-proton is of $10^{34}\,\mathrm{cm^{-2}s^{-1}}$ and for lead-lead is of $10^{27}\,\mathrm{cm^{-2}s^{-1}}$.

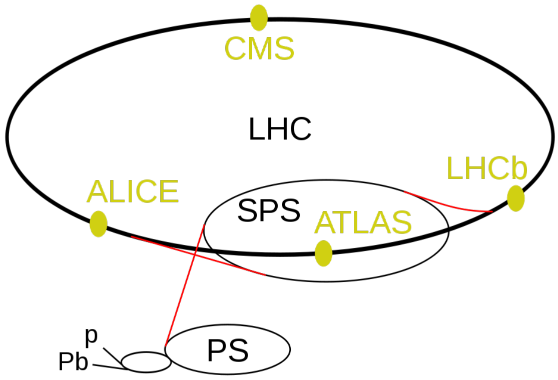


Figure 2.2: Diagram of the CERN accelerator complex.

Particles beams trajectory are curved by 1232 niobium-titanium superconducting dipole magnets each with a length of 14.3 m, they are cooled with superfluid helium to 1.9 K to be able to produce the necessay magnetic field of 8.4 T. To accelerate the beam eight Radio Frequency (RF) cavities located at the LHC point 4 are used. At each turn particle energy is increased to compensate for synchrotron radiation loss and increase the momentum. At nominal operation the LHC will steer 2808 bunches separated by 25 ns in each direction each bunch is composed up to 10^{11} protons. Some of the key parameters of the LHC proton-proton and lead-lead operation can be found in table 2.1.

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

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

Table 2.1: The machine parameters relevant for the LHC detectors.[9]

At the LHC we are looking for extremely rare processes as is can be seen in figure 2.3 the production cross section of a Standard Model (SM) Higgs boson is more than 9 orders of magnitude smaller than the total proton-proton cross section.

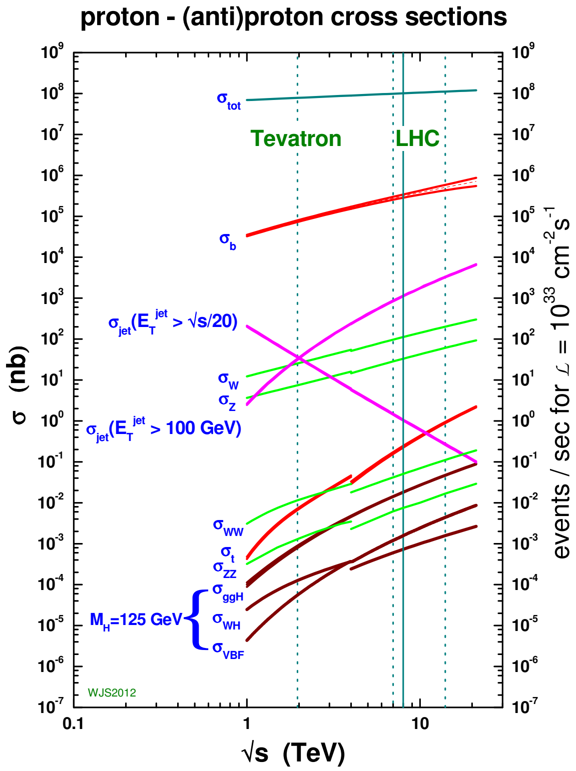


Figure 2.3: Cross sections for several processes for collisions of antiproton-proton and proton-proton as a function of the center of mass energy[6].

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

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma}{4\pi \epsilon_n \beta^*} F, \tag{2.1}$$

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

2.1.1 Running and performance

The LHC has started its operation with the first circulation beams in September 2008. Unfortunately, only a few days after a faulty weld between two dipole magnets caused a significant magnet quench which in turn damaged several dipoles and a simultaneous leak of a significant amount of helium happened. The event showed that beyond the repair of the affected systems the accelerator needed a significant consolidation program to allow it to return to activity[10]. This consolidation program took over one year to finalise and to prevent further possible problems and allow better understanding of the machine while maximizing physics reach, it was decided to initially run the LHC at 7 TeV center-of-mass energy. First collisions happened at November 2009 just at the SPS injection energy of 450 GeV giving start to the LHC run I.

The collision energy was finally ramped up to 7 TeV with first collisions being observed during March 2010. Operation at this energy continued until the end of 2011, with the peak luminosity being achieved of $3.7 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The total amount of of integrated luminosity delivered to CMS was $6.1\,\mathrm{fb}^{-1}$ with the total actually recorded being $5.6\,\mathrm{fb}^{-1}$. During 2012 with the increase knowledge of the accelerator it was possible to increase the centre-of-mass energy further to 8 TeV and eventually reaching peak luminosity of $7.7 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-2}$ and delivering $23.3\,\mathrm{fb}^{-1}$ to CMS of which $21.79, \mathrm{fb}^{-1}\ \mathrm{fb}^{-1}$ were recorded. Figure 2.4 shows the delivered luminosity in the period 2010-2013 over time.

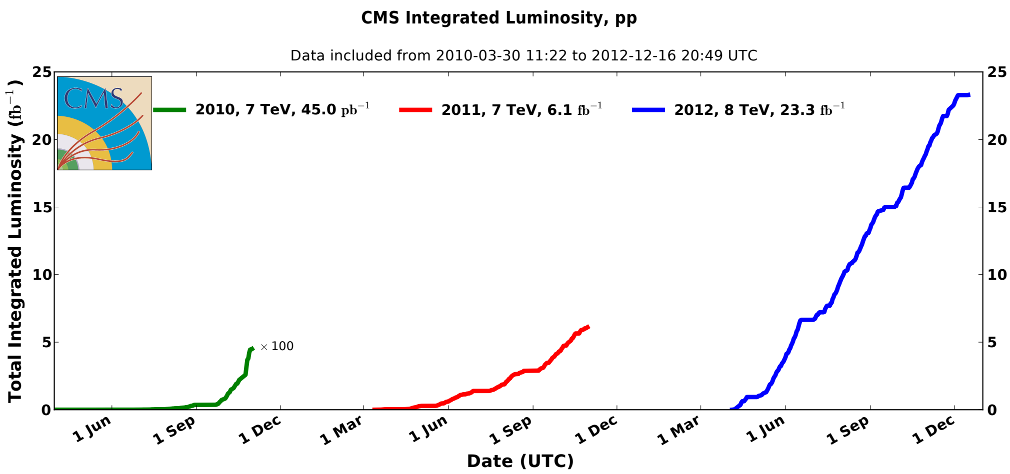


Figure 2.4: Cumulative luminosity versus day delivered to CMS during stable beams and for p-p collisions. This is shown for 2010 (green), 2011 (red) and 2012 (blue) data-taking.

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

In order to achieve high integrated luminosity LHC collides particle bunches 40 millions times a second, and many interactions may happen simultaneously, this effect is called Pile-Up (PU). A figure of the distribution of the mean number of interaction per bunch crossing during 2012 at the CMS experiment can be found in figure 2.5.

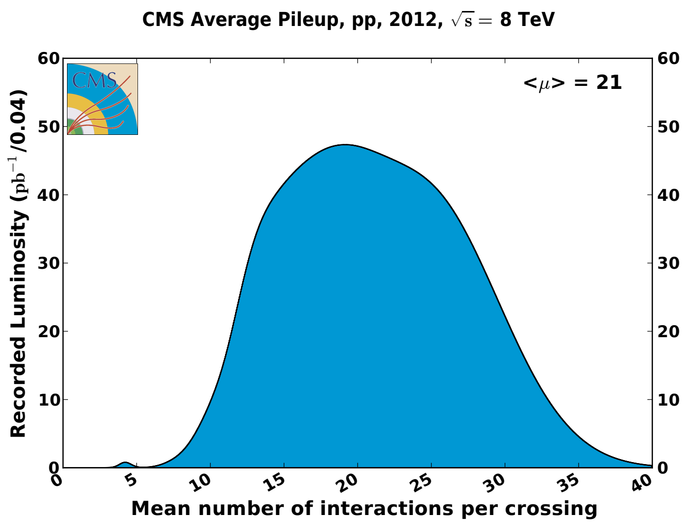


Figure 2.5: Mean number of interactions per bunch crossing at the CMS experiment during 2012.

2.2 The Compact Muon Solenoid Experiment

The experiment is a general purpose experiment located at the LHC point 5, near the village of Cessy, France. It was designed to be a high performance detector studying collisions at its centre. It is composed of several subsystems in a classic onion shaped structure. A diagram of the experiment can be found in figure 2.6.

The main driving motivation for its design is to investigate the electroweak symmetry breaking for which the Higgs mechanism at the design time was presumed to be the most likely explanation. Many other alternative theories to the standard model predict new particles which could be observed at the TeV scale, CMS as a multi-purpose experiment is well suited to search for this new scenarios. If found, such new physics may allow us to understand some of the currently open questions in particle physics, like providing a particle candidates for dark dark matter. Further more, some of this possible new

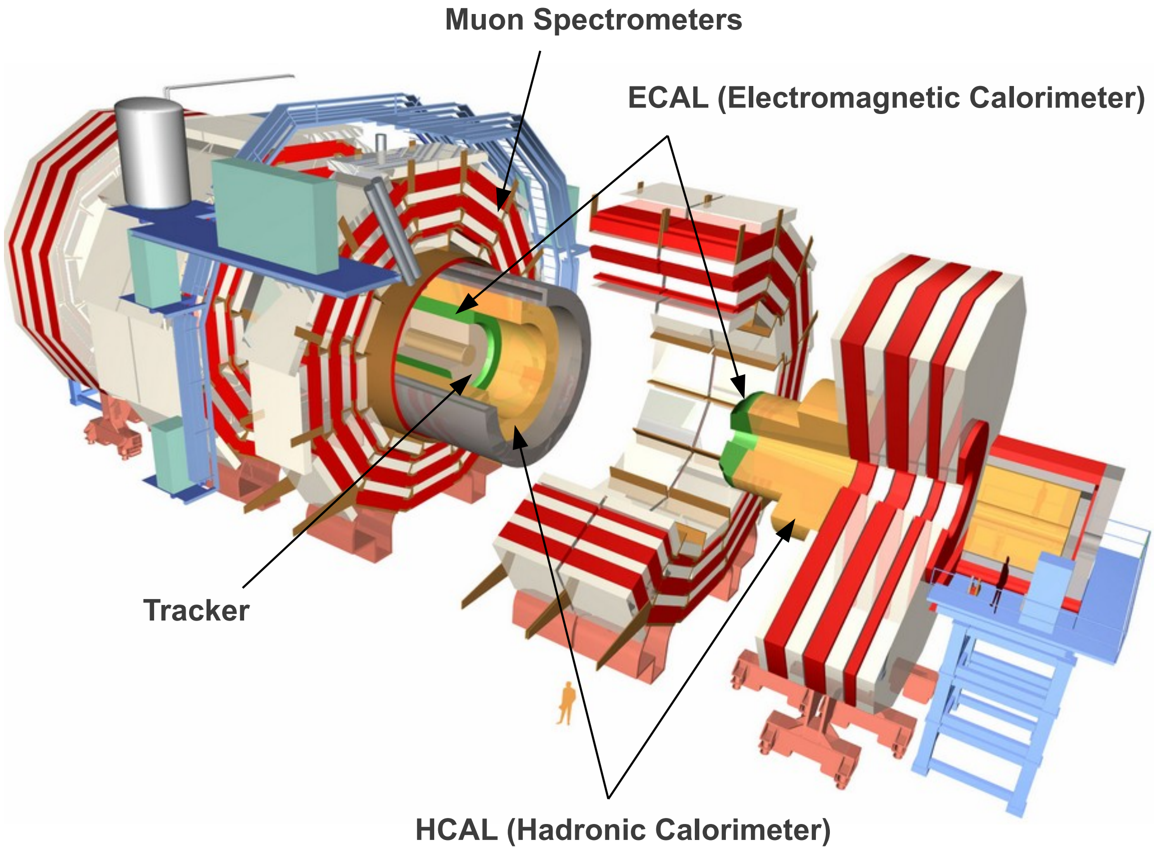


Figure 2.6: Diagram of Compact Muon Solenoid (CMS) experiment showing the experiment in an open configuration and highlighting the position of its sub-detectors. [11]

physics signals could point the way towards a grand unified theory. CMS is also capable of operating while the LHC is colliding heavy ions and has a rich program covering the study of Quantum Chromodynamics (QCD) matter at extreme temperatures, density and parton momentum fraction (low-x).

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

- Good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution ($\approx 1\%$ at $100\,GeV$), and the ability to determine un-ambiguously the charge of muons with $p_\mathrm{T} < 1\,\mathrm{TeV}$.
- Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ's and b-jets, requiring pixel detectors close to the interaction region.

- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution ($\approx 1\%$ at 100 GeV), wide geometric coverage, π^0 rejection, and efficient photon and lepton isolation at high luminosities.
- Good missing-transverse-energy and dijet-mass resolution, requiring hadron calorimeters with with a large hermetic geometric coverage and with fine lateral segmentation.

The final detector design fulfils all this requirements. The experiment is compact compared with the other LHC experiments being 22 m long and 15 m in diameter. Although small, it is the heaviest of the four big detectors at 12500 t. Its high density is a direct consequence of it producing the highest magnetic field at 4 T and therefore needing more material for it to be contain in its return yoke. On the next section we will go in detail over the features and technologies used.

2.2.1 Geometry and conventions

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

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

2.2.2 Inner tracking system

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

Building a tracking system for an experiment at the LHC is very challenging. At design luminosity an average of 1000 particles will hit such system at a rate approaching 40 MHz, leading to high hit density at high rate. It is therefore desirable to have a fast, efficient and high granularity detector. Where at each layer the occupancy should be at or below 1%. On the other hand each layer should be as thin as possible in order to not change the incoming particles trajectory or make them lose too much energy. The detector should also be radiation hard and survive for a period of at least 10 years due to its importance and location. This design requirements have lead to a tracker design entirely based on silicon detector technology.

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

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

The CMS tracker final configuration is composed of a pixel detector with three barrel layers at radii between 4.4 cm and 10.2 cm and 2 disks on each side of the barrel. And a silicon strip tracker with 10 barrel detection layers extending up to 1.1 m with 3 plus 9 disks on each side of the barrel. A schematic of this detector module distribution can be found at figure 2.7. This detector has an acceptance covering up to pseudorapidity of $|\eta| < 2.5$ and has a total active area of about $200\,\mathrm{m}^2$ making the largest silicon tracker ever built.

2.2.3 Electromagnetic Calorimeter

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

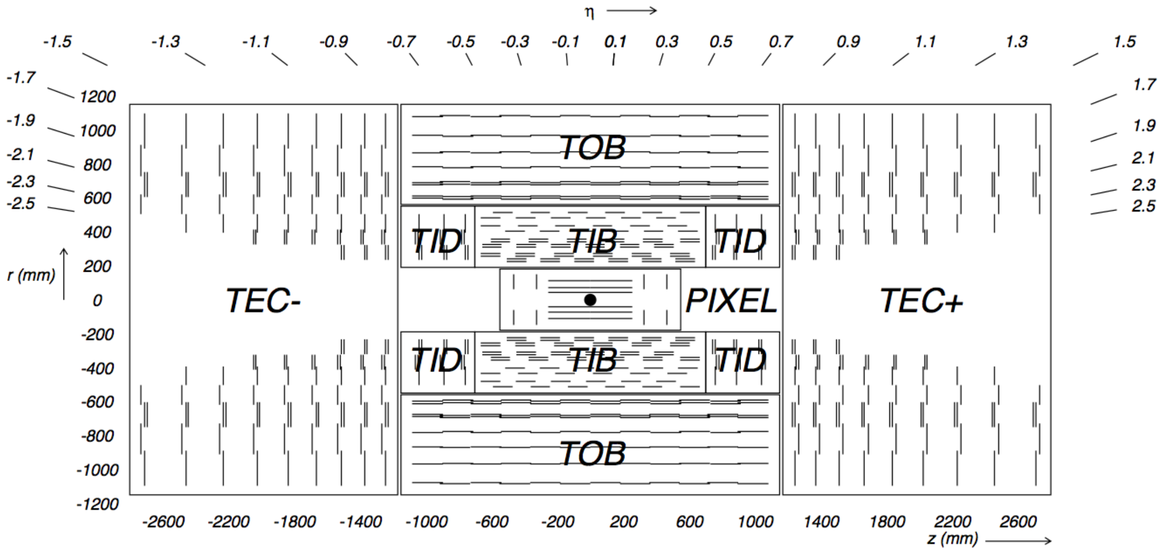


Figure 2.7: Schematic cross section of the CMS tracker. Each line represent a detector module. Double lines represent dual surface back-to-back detector modules.

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

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

The endcap section, the ECAL Endcap (EE), is at a distance of 314 cm from the vertex and covering a pseudorapidity range of $1.479 < |\eta| < 3.0$, are each structured as 2 "Dees", consisting of semi-circular aluminium plates from which are cantilevered structural units of 5×5 crystals, known as "supercrystals". A diagram of the ECAL can be found on figure 2.8.

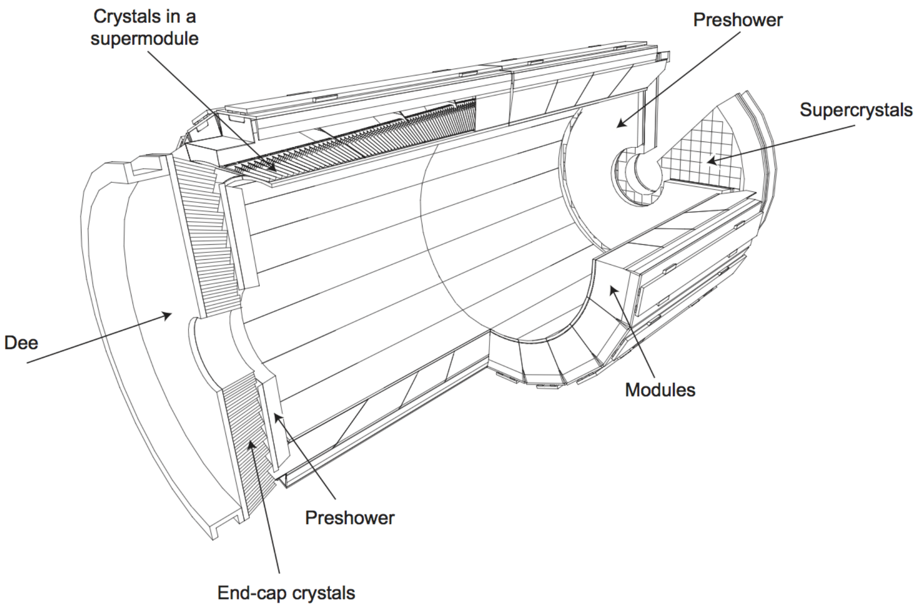


Figure 2.8: Diagram of the ECAL layout illustrating the positions of its components.

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

$$\frac{\sigma}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \tag{2.2}$$

Here E is the energy of the incoming particle, S is the stochastic term which quantifies the fluctuations in scintillation and lateral containment of the shower, N the noise term which relates with electronics and digitisation process and finally C is a constant term that quantifies the non-uniform longitudinal response and inter-calibration errors. These parameters have been measured to be $S = 0.028\,\mathrm{GeV}^{1/2}$, $N = 0.12\,\mathrm{GeV}$ and $C = 0.003$ with the help of an electron beam[12] and in the absence of magnetic field.

2.2.4 Hadronic Calorimeter

The is a sampling calorimeter which is designed to measure the properties of hadron jets and indirectly neutrinos or other undiscovered particles that would result in apparent missing energy[13]. The design of the Hadronic Calorimeter (HCAL) was strongly influenced by the choice of the magnet parameters since most of the calorimetry is inside of the magnet. A diagram of the HCAL subsystems and their location inside CMS can be found in figure 2.9.

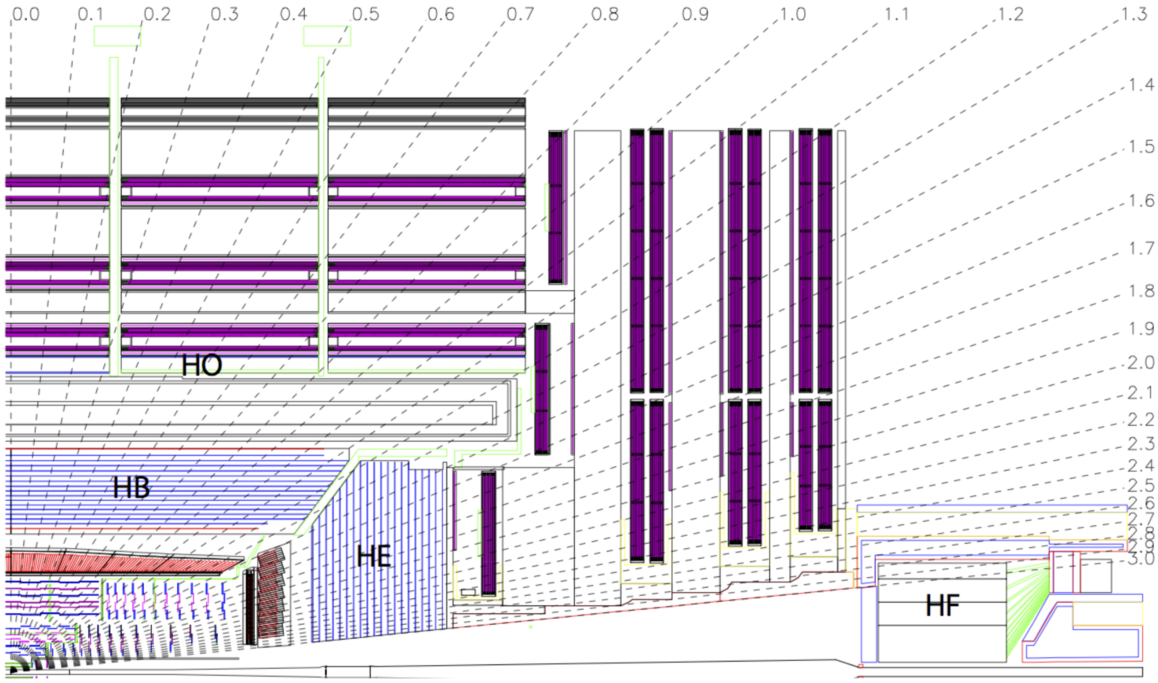


Figure 2.9: Longitudinal view of the CMS detector highlighting the location of the HCAL components: HCAL Barrel (HB), HCAL Endcap (HE) HCAL Outer (HO) and HCAL Forward (HF).

The covers the region up to $|\eta| < 1.3$ and is limited from the beam side by the ECAL at radius $r = 1.77\,\mathrm{m}$ and outwards by the magnet at radius $r = 2.95\,\mathrm{m}$. This is a strict limitation on the amount of absorber material to be used. This detector is composed of 36 identical azimuthal wedges split in two half-barrels. They are constructed of brass absorber plates alternated with plastic scintillator. Brass has a short interaction length ($X_0 = 16.42\,\mathrm{cm}$) and is non-magnetic. The detector is composed of 2304 towers with a segmentation of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ which corresponds to the same area of a 5×5 arrays of ECAL crystals.

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

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

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

Similarly to the ECAL the energy resolution HCAL was tested using a test beam of single charged pions[12], and it was obtained that:

$$\frac{\sigma}{E} = \frac{94.3\%}{\sqrt{E}} \oplus 8.4\%. \tag{2.3}$$

2.2.5 Solenoid Magnet

The design requirements for correct charge assignment and p_{T} determination for charge particles and specially muons drive the magnet parameters choice. For muons, unambiguously charge determination requires momentum resolution of $\Delta p/p \approx 10\%$ at $p = 1\mathrm{TeV}$. This requirements are specially difficult to obtain in the forward regions but with the correct length/radius ratio can be obtained with a modestly sized solenoid magnet but with large field[14].

The choice of the CMS collaboration was to build a Niobium-Titanium (NbTi) superconducting solenoid magnet which has been design to operate at fields up to 4 T it has a diameter of 6 m and a length of 12.5 m at maximum field the stored energy reaches 2.7 GJ. Typically, the magnet is only run at 3.8 T in order to maximize its lifetime. To contain such an enormous magnetic flux a 10 kt return yoke envelopes the magnet with 5 wheels in the barrel region and 2 endcaps composed of 3 disks closing the sides[6]. A summary of the most important magnet parameters can be found at table 2.2.

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

Table 2.2: Parameters of the CMS superconducting solenoid

2.2.6 Muon System

The muon detection is an important part of the mission of CMS as the middle name of the experiment indicates. Muons are fairly easy to detect when compared with other elementary particles and are only rarely produced in proton-proton collisions. Lets take the example of the SM Higgs boson, while the decay mode involving a pair of Z bosons is fairly unlikely compared with other decays the Z bosons can decay into 4 mouns. This decay while rare does not have significant backgrounds making it a "golden channel" for discovery, which indeed was proven the case[15]. Many other models, like SUSY, use muon final states in their searches exactly for the same reason. The CMS muon system is composed of 3 types of gaseous detectors depending on they location and momentum reconstruction needs. A diagram of the disposition of this system inside CMS can be found on figure 2.10.

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

In the endcaps Cathode Strip Chamber (CSC) are used in the region between $2.4 > |\eta| > 0.9$. Here, muon and background rates are high and the magnetic field is not uniform. This system has fast response and is radiation resistant. It is composed by 468 chambers

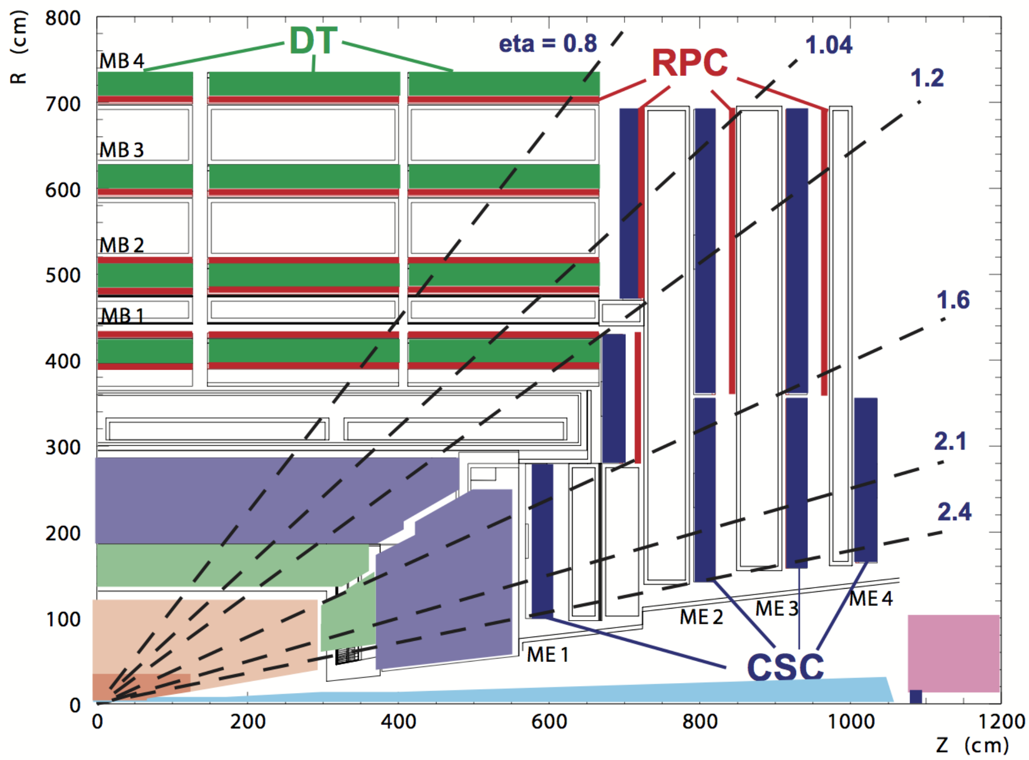


Figure 2.10: Diagram of the CMS muon systems. The location of each muon chamber for each subsystem is showed.

arranged in 4 stations per side. Each chamber is trapezoidal in shape and made of 6 gas gaps and covers either 10° or 20° in ϕ. Each gap contains a plane of cathode strips and a plane of anode wires. For each chamber the spacial resolution is of the order of 200 µm and the angular resolution is $\approx$ 10 mrad in ϕ.

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

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

2.2.7 Data Acquisition System

The CMS Data Acquisition (DAQ) system is designed to process, analyse and ultimately store the information collected by the detector. The LHC produces bunch crossings at a rate of 40 MHz but we are only capable of storing between $10^2 - 10^3$ events per second. At design luminosity each collision will have an average of 20 simultaneous collisions and produce a zero-suppressed data payload of around 1 MByte. To reduce the initial event rate a first level of trigger was designed in order to reduce the amount of events to be processed to a maximum of 100kHz. Even with this event suppression the DAQ has to handle $\approx$ 100GBytes^{-1} which come from approximately 650 data sources. The information is collected and passed to a computer farm where a software filters serve as a second level of trigger, this system is known as the . In this system the event rate is further reduced by a factor of 1000 making the output rate compatible with what can be saved into permanent storage. A diagram of this system can be found on figure 2.11.

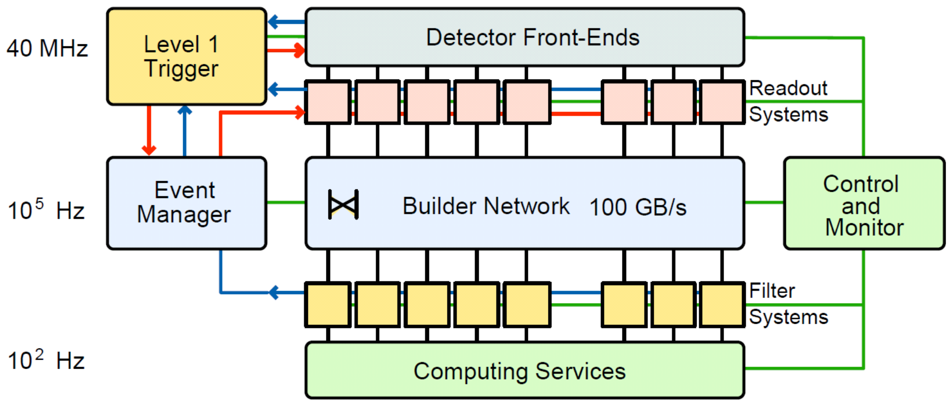


Figure 2.11: Diagram of the CMS DAQ system. Data flow is showed as the lines connecting each electronics or computing units.

2.2.8 Trigger System

At the nominal LHC running conditions the luminosity will be of the order of 10^{34} cm^{-2}s^{-1} where 20 simultaneous collisions will happen at a rate of 40 MHz spaced only by 25 nm.

This means we have every second 10^9 interaction. We can only save $10^2 - 10^3$ events per second with our computer systems. This implies that our trigger system needs to obtain a data reduction of a factor of $10^6 - 10^7$. This is achieved with a two level trigger system, the first is a dedicated hardware system named and the second is a commercial computer system running dedicated software called the [16].

Initially, all data is stored for 128 bunch crossing which corresponds to 3.2 µs. This is the time we have to make a first decision to keep or discard each event. This is the task of the Level 1 Trigger (L1T) which has the target to reduced the data to a maximum rate of 100 kHz. In the allocated time there isn't enough time to get all the information from the detector, so only a coarse version of the calorimetry and muon systems data and some correlation between them is accessed. With this information the L1T produces a set of particle candidates and energy sums over which custom user defined algorithm can filter the events. A diagram of the L1T trigger components and the data flow across the system is present on figure 2.12.

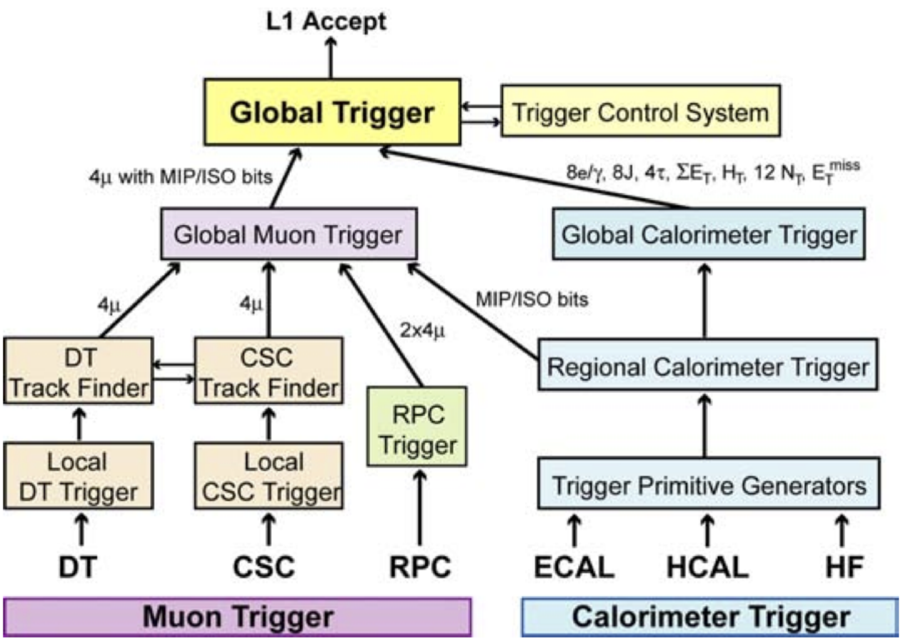


Figure 2.12: Diagram of the L1T system. The arrows indicate data flow and the number of particle candidates at each step is indicate.

The High Level Trigger (HLT) receives events from the L1T and needs to preform further event reduction of $\mathcal{O}(10^5)$ to a final output rate of $\mathcal{O}(10^{2-3})$ Hz . This system is composed of standard computing hardware in the form of computing farm with several thousands of Central Processing Unit (CPU). This system, using the additional latency created by the

L1T event selection, is able to make use of the complete detector information including the tracker data. More sophisticated and precise algorithms are therefore possible which can be tailored to select any desired physical final state.

Both event selection algorithms at the L1T and HLT are constantly updated during data taking. Many times it is necessary to update the selection thresholds in order to control the rate with the increase of LHC luminosity, but such changes can also be due to the development of novel methods or strategies to identify particles more efficiently or the interest in recording new event final final states. The set of algorithms used for data taking is normally referred as the *trigger menu*.

After events passing both levels of the trigger they are recorded into permanent storage. During 2012-13 operation two output streams were saved. The *prompt data stream*, with a rate of approximately 300 Hz, was composed of high priority trigger paths which were immediately reconstructed. And the *parked data stream*, with a rate up to 1 kHz, was stored without reconstruction. This data waited until computing resources would be free to go through reconstruction[17]. This process was finalised a few months after the LHC run I was finished.

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

2.2.9 Computing

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

The CERN Data Centre is the Tier 0 of this network, know as the Grid. All data produced by the LHC passes through it. Only about 20% of the total capacity of the Grid is hosted here, but CERN has the very important mission of safe keeping all the raw data produced by the LHC experiments. It also has the task of doing the first attempt at reconstructing this data into meaningful physics objects.

There are 13 Tier 1 computer centres around the world. They are responsible to store a proportional amount of raw and reconstructed data among them. If any reprocessing of the data is needed, this centres are responsible for this task and storing the resulting output as well. Tier 1 centres also host simulated data and distributing it to affiliated Tier 2 centres.

Local research centres like universities of scientific laboratories are normally at the Tier 2 level. They should have enough computing power and storage space for the analysis in which those centres are involved. This centres will have the responsibility of handling a proportional share of simulated data production and reconstruction. Currently there are over 150 tier 2 centres around the world

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

2.2.10 Level 1 Trigger: Stage I Upgrade

An extensive upgrade program for the L1T electronics was planed and is being executed in order to cope with the increase of luminosity and pile-up predicted for the period after Long Shutdown 1 (LS1)[18]. It is expected that center-of-mass energy will almost double from 8 TeV to 13 TeV, instantaneous luminosity will also increase as will average pile-up. Also, the bunch separation with change from 50 ns to 25 ns making out-of-time pile-up a significant problem.

To ensure physics performance during 2015 only a partial upgrade is planned which is known as the *Stage-1* upgrade. The main feature of this upgrade is the replacement of the existing Global Calorimeter Trigger (GCT). Two key enhancements will be possible from the upgrade:

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

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

Chapter 3

Technical work

3.1 Level 1 Trigger Data Quality Monitoring System

3.1.1 Online Monitoring

3.1.2 Offline Monitoring

3.1.3 Release Validation

3.2 Implemented Tests

3.2.1 Rates Monitoring

3.2.2 Synchronization Monitoring

3.2.3 Occupancy Monitoring

3.2.4 Status Summary Display

3.3 Certification

3.4 Proposed Future Upgrades

Chapter 4

Event Reconstruction and Physics Objects

This chapter describes how the Compact Muon Solenoid (CMS) detector produces physics objects from the information collected at each event. The Higgs to invisible analysis uses almost all the physics objects produced by the detector and for this uses information from all the experiment sub-detectors. The following sections detail for each of these objects how they are reconstructed and what are the choices made to filter them.

4.1 Tracks

Reconstructing the trajectories of charged particles allows us to measure their momentum and determining their charge. This is possible by analysing the hit patterns in the inner tracking system. In CMS this reconstruction is made with the Combinatorial Track Finder (CTF) algorithm [19]. The relevant steps for track generation are described below:

- Seed generation is made with hits at the pixel detector. A track seeds can be made with two or three hits. In the first case a know vertex or the beam spot is used to constrain the seed momentum. The parameters of each seed are estimated using the assumption that the trajectory is a helix, but it takes into account hit errors and multiple scattering [20].
- The track seed is extrapolated through the tracker layers with on a combinatorial Kalman filter [21] . For each additional layer, the best matching hit if any is added and track parameters are recomputed. This procedure continues until the last layer is reached [19].

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

The process of track finding is repeated up to six times where the hits for each successfully reconstructed track are removed for the next iteration. Using early Large Hadron Collider (LHC) data and a dataset of pions and muons it was possible to estimate that the tracking efficiency is $> 98\%$ for all track $p_{\mathrm{T}} > 500\,\mathrm{MeV}$ and $> 99\%$ for tracks with $p_{\mathrm{T}} > 2,\mathrm{GeV}$ [22].

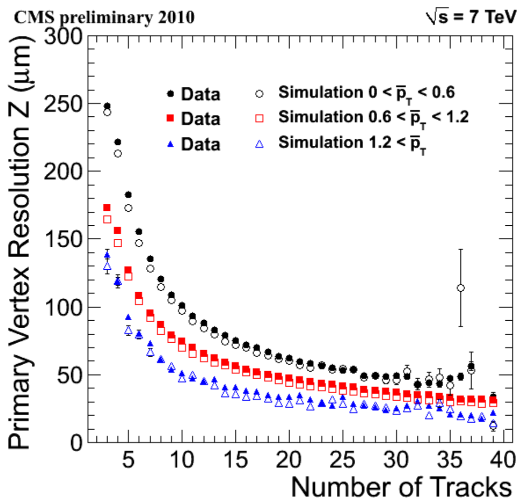
4.2 Vertex Reconstruction

The LHC can produce extreme collision intensities which are obtained partially by having multiple collisions happening at each bunch crossing. As it has been discussed in section 2.1.1 an average of 21 simultaneous collisions happened per bunch crossing in the CMS experiment during 2012. In this environment, it is crucial to identify the Primary Vertex (PV) and the particles that come from it. This information can then be used to reject particles coming from other additional collisions and to identify displaced vertices which can be the signature of long lived particles like b-mesons.

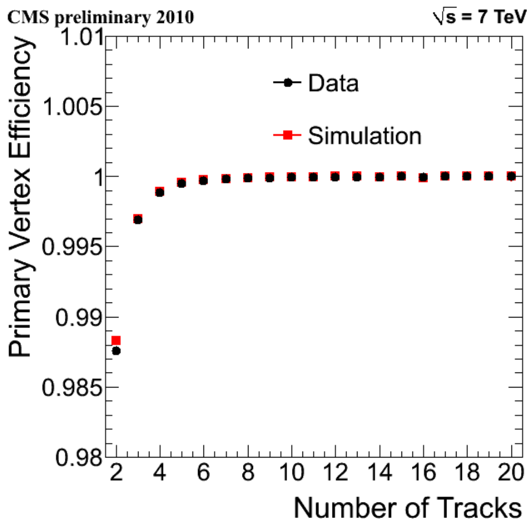
The individual tracks are reconstructed making use of the inner tracker. Each vertex is initially seeded by two tracks with separation in z less than 1 cm. Then remaining track are clustered to seed vertex with the Deterministic Annealing (DA) algorithm [23]. After the clustering process is done, the position of each vertex is recomputed using the adaptive vertex fitter algorithm [24]. In this algorithm weights, w_i are assigned to each track according to how compatible they are with the fitted vertex position. Weight vary from 1 to 0, being that track assigned weights of close 1 are highly compatible with the vertex and close 0 would be given to low compatibility tracks. Then we can define the number of degrees of freedom of the new fit as:

$$n_{dof}(vertex) = 2\sum_{i}^{tracks} w_i - 3 \tag{4.1}$$

This variable can be used to distinguish real proton-proton interactions from misclustered vertices, since it is correlated with the number of tracks compatible with that specific vertex [25]. The vertex position and resolution have been measured with LHC data and compared with simulation. The resulting plots can be found in figure 4.1 as a function of number of tracks.



(a)



(b)

Figure 4.1: (a) Primary vertex resolution in the z coordinate a function of the number of associated tracks. Results are give for three ranges of average track p_{T}. (b) Primary vertex efficiency as a function of the number of associated track [25]

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

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

- We require a real reconstructed vertex from tracks, not the beam spot.
- A minimum number of degrees of freedom: $n_{dof} > 4$.

- Collision must be near the interaction point. We require longitudinal distance to be $|z| <= 24\,\mathrm{cm}$.
- We required the collision be be close to the beam line. Radial distance to beam line: $d_{xy} < 2\,\mathrm{cm}$.

4.3 Particle Flow

The Particle Flow (PF) algorithm [26–28] is used in the CMS experiment with the objective of reconstructing every stable particle produced in the event. This is achieved by combining information from all CMS sub-detectors in order to identify electrons, photons, muons, charged hadrons and neutral hadrons and measure their direction, energy and type. The identified particles can in turn be used in jet clustering, to determine the missing transverse energy, to reconstruct and identify taus, to calculate particle isolation, for identify b-quark jets, etc.

The CMS experiment is very well suited for this approach since we can is equipped with a high precision silicon tracker which is immersed in uniform axial magnetic field and its dual calorimeter design with high hermeticity and resolution. The tracker system allows Very precise direction/momentum reconstruction for charged particles down to transverse momentum as low as 150 MeV. The high granularity of the Electromagnetic Calorimeter (ECAL) allows for photons to be identified through deposit separation even inside high energy jets. In turn electrons can be reconstructed by combining their track and the energy deposits of the electron itself and its emissions, this algorithm will be explained further in section 4.4. The tracker information also allows to separate charged and neutral hadrons in close proximity, a task which is not possible with just the Hadronic Calorimeter (HCAL) due to its coarser granularity. We can determine the charged hadron momentum from the track information, and then, by removing its deposit from the calorimeter system we can determine the neutral hadron deposits. In areas outside the tracker and/or ECAL coverage measurements are more coarse Since we less information available.

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

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

4.4 Electrons

In the CMS experiment electrons are reconstructed by matching energy clusters in the ECAL with tracks coming from the inner tracking system. Unfortunately, electrons can loose and disperse significant amounts of energy until they reach the ECAL. While they transverse the inner tracker they may emit photons through bremsstrahlung and in turn this photon can convert to e^+e^- pairs. About 35% of the electron radiate at least 70% of their energy in this way [29]. This spread of energy is mostly in ϕ due to the applied magnetic field [30]. Dedicated algorithms were developed to combine the the ECAL energy deposits, into a so called *super-clustering algorithm*, of the initial electron and its emissions.

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

In the encaps the "Multi-5×5" is used. In the region of the detector the geometry is more complex and does not follow a simple $\eta - \phi$ symmetry. We start by selecting for seeds the crystals which are local maxima over their four direct neighbours and have a deposit of $E_\perp > 0.18\,\mathrm{GeV}$. Then, and starting with the seeds with highest $E_\perp$, we collect the energy around them into clusters of 5×5 crystals. We then search for similar seeds and form clusters that can overlap within $\Delta\eta < 0.07$ and $\Delta\phi < 0.3\,\mathrm{rad}$ of the initial seed. Those clusters are then combined into a single *supercluster* which needs to have at least $E_{\perp} > 1\,\text{GeV}$. The *supercluster* is then extrapolated to the ECAL preshower by clustering the energy within $\Delta\eta < 0.15$ and $\Delta\phi < 0.45$ around the most energetic cluster and adding it to the *supercluster* itself [31].

In order to reconstruct the electron track we need to take into account the bremsstrahlung emissions. The CTF algorithm is not appropriate for this purpose so a different track-finding algorithm had to be developed. For high p_{T} electrons we use the ECAL supercluster energy deposit weighted mean impact point as a seed. If we combine this information with the determined $E_{\perp}$ we can define two $\eta - \phi$ search regions in the pixel detector depending on the charge hypothesis. If we find two compatible hits, the electron trajectory is updated. From this point normal track building is performed but instead of a Kalman filter algorithm we use a Gaussian Sum Filter (GSF) algorithm [32]. This method performs better in the presence of non-Gaussian losses like the one coming from the bremsstrahlung emissions.

The typical background to real electrons are collimated hadronic jets, like from π^0 and $\pi^{\pm}$ overlap or from $\pi^{\pm}$ showers [29]. There are many useful variables that may be used to reduce such background and are often used in *electron identification* criteria:

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

Distributions of the variables for simulated electrons and jets can be found in figure 4.2.

4.5 Muons

Muon track reconstruction starts independently at the inner-tracker (*tracker track* and in the muon systems (*standalone-muon track*) [33]. Then this information can be combined into a single muon track in two possible ways.

Global Muon reconstruction is an *outside-in algorithm.* We starts by finding tracker track match for each stand-alone muon track. This is done by propagating the match candidate

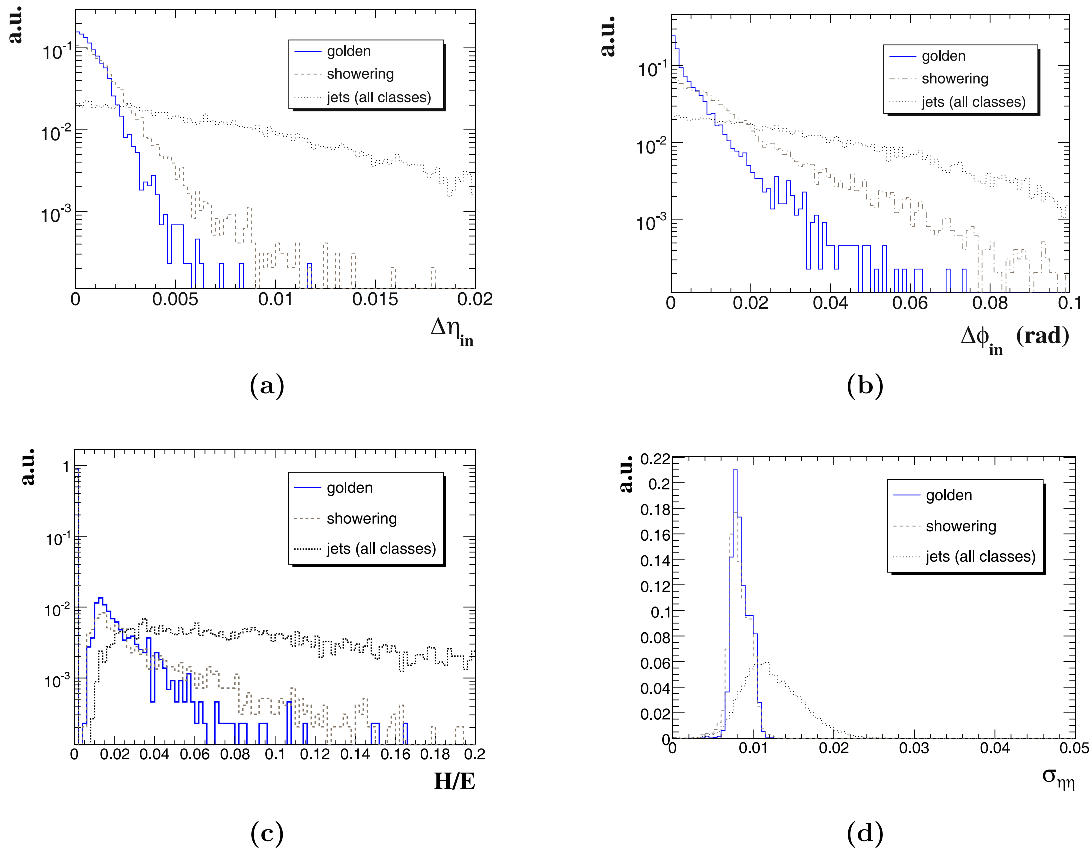


Figure 4.2: Distributions for (a) $\Delta\eta_{in}$, (b) $\Delta\phi_{in}$, (c) H/E and (d) $\sigma_{i\eta i\eta}$. Here *golden electrons* are those who emit minimal bremmstrahlung photons, *showering* are electrons that lose a large faction of their energy in emissions and *jets* are the typical distributions for hadronic jets.

pair to a common surface and comparing track parameters. For each matched pair, a *global-muon fit* is performed using all hits from the two tracks using a Kalman-filter algorithm [21]. For muons of $p_{\mathrm{T}} \gtrsim 200\,\mathrm{GeV}/c$, it has been showed that a *global-muon fit* improves the momentum resolution compared to a *tracker-only fit* [9, 34].

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

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

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

When reconstructing global muons, its unlikely to find non-prompt muons but we may have hadronic activity "punching-through" the calorimeter system and appearing in the muon system. To reduce this types of background we can use different muon identification criteria. We can define a "tight muon" as global fit track using tracker and muon chamber hits with a χ^2 per degree of freedom os less than 10. This fit must include at least one segment in the muon chamber, track segments in at least 2 muon stations, use more than 10 hits in the inner tracker of which at least one in the a pixel layer and finally a small transverse impact paramenter $|d_{xy}| < 2\,\mathrm{mm}$. The efficiency for such a criteria has been measured both in data and Monte Carlo using $J/\psi \to \mu^+\mu^-$ and $Z \to \mu^+\mu^-$ and for $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ it plataus at 96-99% [33].

4.6 Lepton Isolation

To reduce the the probability of misidentification of a lepton coming from Quantum Chromodynamics (QCD) jets as one coming from the hard scattering we can require isolation [31, 33]. We compute the isolation by summing the transverse momenta of all particles inside a cone around the selected lepton. In this sum we include all charged particles, neutral hadrons and photons. But we do not want to include the Pile-Up (PU) contribution to this sum so we only include the charged candidates with an impact parameter smaller than 0.1 c m. Different methods can be used to subtract the neutral component of the PU.

Normally, for physics analysis we defined the more meaningful *relative isolation* as $I_{rel} = I/p_{\mathrm{T}}^{lepton}$.

4.7 Electrons Isolation

For electrons we calculate the *effective area corrected isolation* over a cone of $\Delta R < 0.3$ around the electron. For the neutral PU subtraction we uses a look-up table of effective areas according to electron $|eta|$ which is multiplied by the estimated neutral PU energy density by unit of effective area. The definition for this isolation can be found in equation 4.2.

$$I = \sum^{\substack{\text{charged}\\ \text{non-pileup}}} p_{\mathrm{T}} + \max\left(0, \sum^{\text{neutral}} p_{\mathrm{T}} + \sum^{\text{photon}} p_{\mathrm{T}} - \rho(\text{lepton}) \times \text{Eff. Area}(\text{lepton})\right) \tag{4.2}$$

4.8 Muons Isolation

For muons we use the *combined isolation* over a cone of $\Delta R < 0.4$ around the muon. For neutral PU subtraction we use the determined charged PU component inside the cone and multiply it by a factor of 0.5 which is determined from simulation. The definition for this isolation can be found in equation 4.3.

$$I = \sum^{\substack{\text{charged}\\\text{non-pileup}}} p_{\mathrm{T}} + \max\left(0, \sum^{\text{neutral}} p_{\mathrm{T}} + \sum^{\text{photon}} p_{\mathrm{T}} - \frac{1}{2}\sum^{\substack{\text{charged}\\\text{pileup}}} p_{\mathrm{T}}\right) \tag{4.3}$$

4.9 Jets

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

4.9.1 Jet Clustering

Jet clustering algorithms are sets of rules that allows us to combine a particle candidates into a jets [35]. These algorithms normally are controlled by parameters that define how close particles need to be in order to be associated into a jet and a way to combine their momentums. However jet definition should be robust and provide consistent measurements about the parton. There are two major problems that may affect a jet algorithm. This problems appear when the number of jets in an event changes by adding a soft collinear gluon emissions (collinear safety) or by parton splitting (infrared safety).

In CMS we used a sequential recombination algorithm known as *anti-*k_T [36] which is both infrared and collinear safe.

$$d_{ij} = \min({p_{\mathrm{T}}}_i^{2p}, {p_{\mathrm{T}}}_j^{2p})\frac{\Delta R_{ij}^2}{R^2} \tag{4.4}$$

$$d_{iB} = {p_{\mathrm{T}}}_i^{2p} \tag{4.5}$$

[37]

4.9.2 Jet Energy Corrections

$$P^{\mu}_{\mathrm{corr}} = C_{\mathrm{offset}}({p_{\mathrm{T}}}^{\mathrm{raw}}, \eta) \cdot C_{\mathrm{rel}}({p_{\mathrm{T}}}'', \eta) \cdot C_{\mathrm{abs}}({p_{\mathrm{T}}}', \eta) \cdot P^{\mu}_{\mathrm{raw}} \, , \tag{4.6}$$

[38]

[39]

4.10 Taus

4.11 Missing Transverse Energy

[40]

[41]

Chapter 5

Prompt Data Analysis

Chapter 6

Parked Data Analysis

This chapter describes the analysis performed over the Compact Muon Solenoid (CMS) Run I parked data collected over 2012 and 2013. This data was collected and stored without reconstruction and only became fully available a few months after data taking was finished. The advantage of this dataset is the possibility to use lower threshold triggers which can collect more signal but also more backgrounds. To take full advantage of this data the analysis had to be redesigned and extended with new control regions.

6.1 Data and MC samples

6.1.1 Data

In this analysis we used the full certified data with collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ from 2012-13 data acquisition (Run I), using golden JSON file `Cert_190456-208686_8TeV_22Jan2013ReReco_Collisi` was used in this analysis. It amounts to an integrated luminosity of $19.2 \pm 0.5\,\mathrm{fb}^{-1}$. A summary of the dataset names and their integrated luminosity can be found in table 6.1.

The difference between certified and analysed datasets is due to out analysis trigger not being active for the first few runs of of Run 2012B.

6.2 Event quality filters

During data recording issues may happen with the detector or data acquisition which may render some of the events unusable. The groups responsible for each part of the

Dataset	$\int$ *Luminosity* $[pb^{-1}]$
/MET/Run2012A-22Jan2013-v1/AOD	889
/VBF1Parked/Run2012B-22Jan2013-v1/AOD	3871
/VBF1Parked/Run2012C-22Jan2013-v1/AOD	7152
/VBF1Parked/Run2012D-22Jan2013-v1/AOD	7317
Total analysed	19229
Total certified luminosity	19789

Table 6.1: Relevant parked datasets from Run I and their total analysed integrated luminosity. Total analysed and certified also showed.

detector and physics object check the data after it was taken and if they find such problems occurred. This groups produce software event filters analysts to be able to remove this problematic events. This issues cover from know detector problems to miss firing of calibration sequences or even failure to reconstruct physics objects.

The Jet-MET Particle Object Group (POG) recommends the usage of the following filters which are used in this analysis.

- CSCTightHaloFilter
- HBHENoiseFilter
- EcalDeadCellTriggerPrimitiveFilter
- trackingFailureFilter
- eeBadScFilter
- ECAL Laser filter
- HCAL Laser filter

In turn the JetMET group recommend the usage of the following Tracking POG Filter:

- logErrorTooManyClusters
- manystripclus53X
- toomanystripclus53X

6.3 Electrons

6.3.1 Veto electrons

The "veto electrons" are defined by the base requirements of the cut based electron ID veto working point of the EGamma POG with some additional cuts on top.

Requirements of the cut based electron ID veto working point:

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

Table 6.2: Details of the CMS Electron-Gamma Particle Object Group (POG) recommendations for a *veto electron*. Here barrel is defined as $|\eta_{supercluster}| <= 1.479$ and endcap is $1.479 < |\eta_{supercluster}| < 2.5$.

Additional requirements for this analysis

- $p_\perp > 10\,\mathrm{GeV}$
- $|\eta| < 2.4$

6.3.2 Tight electrons

The "tight electrons" are defined by using the base requirements of the cut based electron ID tight working point (similar to 2011 very tight WP70) of the EGamma POG with some additional cuts on top.

Requirements of the cut based electron ID tight working point (similar to 2011 very tight WP70) and can be found in table 6.3

Additional requirements for this analysis:

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

Table 6.3: Details of the CMS Electron-Gamma POG recommendations for a *tight electron*. Here barrel is defined as $\lvert\eta_{supercluster}\rvert <= 1.479$ and endcap is $1.479 < \lvert\eta_{supercluster}\rvert < 2.5$.

- $p_\perp > 20\,\mathrm{GeV}$
- $\lvert\eta\rvert < 2.4$

6.4 Muons

6.4.1 Loose Muons

Requirements:

- $p_\mathrm{T} > 10\,\mathrm{GeV}$
- $\lvert\eta\rvert < 2.1$
- Relative Combined Isolation < 0.2

POG Loose ID:

- Particle-Flow muon ID
- Global or Tracker Muon

6.4.2 Tight Muons

Requirements:

- $p_{\mathrm{T}} > 20\,\mathrm{GeV}$
- $|\eta| < 2.1$
- Relative Combined Isolation < 0.12
- $d_{xy} < 0.045$
- $d_z < 0.2$

POG Tight ID:

- Global Muon
- Particle-Flow muon id
- $\chi^2/ndof < 10$
- At least one muon chamber hit included in the global-muon track fit
- Muon segments in at least two muon stations
- $d_{xy} < 2$ mm w.r.t. the primary vertex
- $d_z < 5$ mm
- Number of pixel hits > 0
- Number of tracker layers with hits > 5

6.5 Taus

Requirements

- $p_{\mathrm{T}} > 20\,\mathrm{GeV}$
- $|\eta| < 2.3$
- $d_z < 0.2\,\mathrm{cm}$

Descriminants:

- decayModeFinding

- byTightCombinedIsolationDeltaBetaCorr3Hits
- againstMuonTight
- againstElectronTight

6.6 The Cross Check Analysis

Chapter 7

Run II Preparation

After the successful completion of the the first data taking period, the Large Hadron Collider (LHC) Run I, the accelerator and detectors went through a two year long technical shut-down which was designated the Long Shutdown 1 (LS1). During the period the accelerator completed a consolidation and improvement program to allow a ramp up of the beams energy up to the design value of 7 TeV per beam in proton-proton mode. At the same time the experiments also performed maintenance, repair and improvement programs.

Data analysis continued during this period of no data taking using the datasets already available or the newly reconstructed parked data. After this final work over 8 TeV data was completed most Compact Muon Solenoid (CMS) physics analysis started their preparation for the LHC Run II, where higher collision energies, even higher values of Pile-Up (PU) and more recorded integrated luminosity are expected. Following this global effort the CMS vector Boson Fusion (VBF) Higgs to invisible analysis also started its own preparation work.

The first step is always the definition of a trigger condition for data taking. The effort made to create and study such an adequate set of trigger for the use of this analysis during run II is documented in section 7.1. Additionally, work was made to study and propose the creation of a dedicated Quantum Chromodynamics (QCD) Monte Carlo (MC) sample with signal like characteristics expanding on the one already created for Run I. This study can be found in section **??**.

7.1 Run II trigger studies

The first step of any CMS physics analysis is to define which trigger to use for data taking. The Trigger Studies Group (TSG) develops generic usage trigger conditions, know as trigger paths, which can be used by any analysis. Typically this conditions cover all possible single objects (single electron, single jets, etc), multiple objects (double electron, triple muon, etc), cross triggers, (single electron + sigle muon, etc). In some cases, like for our analysis, it with better to define a custom condition to obtain maximum physics content. The following reasons drove the decision to create a set dedicated trigger paths.

- Maximize signal collection efficiency by selecting our signal topology with reduced trigger cuts while compared with generic triggers;
- Use again a trigger condition with $MET_{\text{no muon}}$ instead of MET to study Electroweak (EWK) Z irreducible background;
- Create a new dedicate pre-scaled trigger path with reduced thresholds with objective of reducing systematics;

For the proposal of our triggers it was decided to produce numbers for conservative and aggressive scenarios in terms of available High Level Trigger (HLT) bandwidth. For the signal trigger path rate 1.5 Hz and 5.0 Hz were considered. While for the systematics paths rate 0.1 Hz and 0.1 Hz were considered.

7.1.1 Methodology

7.1.2 Signal path

7.1.3 Systematics path

7.2 Additional L1 trigger

7.3 Run II QCD Monte Carlo samples

Simulating and reconstructing quantities of QCD events comparable to the ones produced at the LHC experiments is impractical. At every second of LHC physics operation several millions of bunch crossings happen, each one able to create several simultaneous collisions.

With the currently available hardware it takes in excess of one minute to fully simulate one of such bunch crossings. Further more, most of this events have very low energy collisions and are unlikely to be picked up by any physics analysis.

This constraints lead to QCD events being simulated in $p_\perp$ hats, where the first collision outgoing particles summed $p_\perp$ generated within a predefined rangee. Then several other collisions are added to the event as PU. This additional collisions are generated without any constraints in $p_\perp$.

This bin method allows the user to have QCD hard scattering samples with increasing energies and study the influence of each one of them in their own analysis. As a practical example we do not need to look over millions of QCD events to find high energy jets. We can just start from the higher QCD $p_\perp$ hats add lower ones until the contributing to our selection is negligible. On the other hand, analysis like the CMS VBF Higgs to invisible analysis, search for event topologies with low energy jets and/or Missing Transverse Energy (MET). In this cases available inclusive QCD samples will not have enough statistics to provide insight into this backgrounds behaviour.

During the preparation of the Run I the VBF Higgs to invisible analysis privately produced a set of QCD samples with VBF like jets and real MET. This samples allowed to understand the mechanisms that create real MET in QCD and how those could be mitigated.

In the preparation for Run II it was considered once again to be useful to have similar samples remade and possibly extended. It was identified that not only real MET is significant but also fake MET coming from detector miss-measurement. The QCD background is currently the only background we do not have any MC event sample. If such a sample could be produced it could allow the analysis to evolve to a shape based analysis or to use machine learning techniques, since we would have signal and all backgrounds simulations.

7.3.1 Monte Carlo sample simulation

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

To simulate one event on the CMS experiment we first start by simulating the physics process itself. We can split this into two sub-processes: hard scattering and hadronization. There are many purpose developed software programs that will perform each one of this steps of even both.

General purpose particle physics event generators like Pythia8 [42, 43], Herwig++ [44] or Sherpa [45] are able to do both hard scattering and hadronization steps. Typically these programs are restricted to 2 by 2 hard processes which are calculated at Leading Order (LO).

There are many other event matrix-element generators, like MadGraph5_aMC@NLO [46] and Alpgen [47] that focus on the hard process simulation programs providing events from 2 by X hard scattering with some implementing Next to Leading Order (NLO) calculations. This parton level events then need to be passed to a one of the general purpose event generators for further hadronization.

We need to avoid overlapping in the phase-space description of matrix-element and showering programs when simulating multi-jets events. The overlap comes from software like Pythia or Herwig describing parton radiation as a Markov Chain process based on Sudakov form factors. This approach is only formally correct in the limit of soft and collinear emissions. On the other hand Matrix Element (ME) programs like MadGraph works well for the hard scattering but diverges when the partons become soft or collinear.

There are a few jet-parton matching schemes developed to account for this overlap. Showering can be vetoed and the event reweighed according like in the CKKW scheme [48–50] or events can be rejected altogether like in the MLM scheme [51]. Depending on the generator used for the showering, different schemes are implemented and care must be taken in the definition of the matching parameters.

After the physics event is simulated, the interaction with the detector and the corresponding electronics response is estimated using a precise model of the experiment. In the CMS experiment we use GEANT4 [52, 53] software for this task which also relies heavily on Monte Carlo (MC) methods.

When the detector response is obtained we can proceed with the same event reconstruction algorithms already described in chapter 4.

7.3.2 Goals and first attempt

Building on the knowledge gained from the samples produced during Run I we can defined the goals for this new samples. Cuts at generator level involving MET should be avoided in order to not filter out events where the MET comes from miss-measurement. Variables that my bias $\Delta\phi(jet - jet)$ distribution should also avoided since the Run I analysis uses inverted cuts in this variable to perform data-driven QCD estimation. All cuts should be below the event selections used during Run I and if possible around or even below the Run II trigger conditions. The sample to be simulated should be equivalent to at least $1\,\mathrm{fb}^1$ of data but of a size comparable with the current official QCD Inclusive sample. This last requirement is to ensure that the computing resources necessary for making such sample do not go above what currently is used to produce similar purpose samples.

The first attempt to produce a proposal for the production of this QCD VBF-like samples was based on filtering events produced by Pythia 8. The filtering of this events was made by first clustering the generator particles in anti-k_T jets with $\Delta R < 0.4$ where muons were ignored. Only the events where at least one dijet with VBF characteristics would kept. Unfortunately, this approach lead to a very large number of event being generated (hard scattering and hadronization) and clustered only do be discarded. The computing time was considered too large to be feasible considering the physics case by the CMS team responsible for official sample production. However, is was recommended to take a different approach by using a ME generator, like MadGraph and cut already at the parton level, before any hadronization or clustering. After this initial event selection a second layer of cuts could be applied after hadronization to ensure the actual outgoing jets would pass out criteria. Furthermore, using a ME generator should provide a more accurate description of multi-jet events while the two steps approach should allow a significant reduction of the necessary computing time.

7.3.3 MadGraph parton level simulation

The MadGraph event generator was selected produce the parton level simulation. With this generator it is possible to make events from the interaction of two proton partons and obtain a final state with any number of partons. Each additional parton on the final state comes at the cost of an exponential increase of the possible diagrams, which in turn means more time is necessary to create events. I was chosen to only produce final states with 2, 3 and 4 outgoing partons. This generator has been used to create similar QCD samples used by some CMS Super Symmetry (SUSY) analyses.

The outgoing partons are defined to be a gluon or a quark (u, d, c, s or b). We do not allow diagrams with top quarks since they do not hadronize and lead to event topologies which are already accounted for in our analysis. The outgoing partons will be the seed of our final state jets.

A custom parton level filter was implemented inside the MadGraph code to select events with VBF characteristics. To pass the filter the event must have at least one outgoing di-parton with invariant mass of 800 GeV, where both parton are inside the detector volume with $|\eta| < 5.0$ and have more than $p_T > 30\,\text{GeV}$. The distributions of this variables for events passing this cuts can be found in figure 7.1.

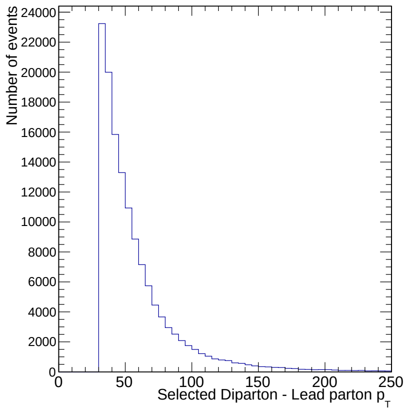
The estimated cross section for this processes and selection is $1.029 \times 10^7 \pm 1.614 \times 10^4$ pb and we request the production of 1.2^{10} events. That corresponds to an equivalent luminosity of just over $1.1\,\text{fb}^{-1}$ of equivalent integrated luminosity.

7.3.4 Hadronization with Pythia 8

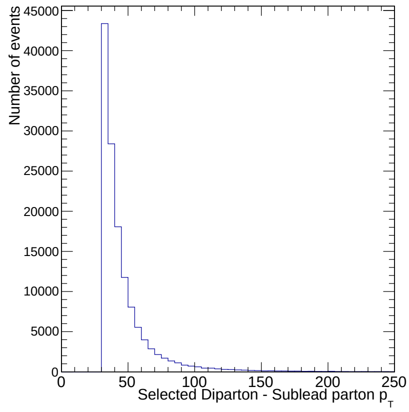
The parton level events that have passed the initial filter now have to be hadronized. Similarly to other samples produced in the CMS we have chosen Pythia 8 for this task. As described in section 7.3.1 when using a ME generator with a shower generator we need to filter the overlapping phase-space. As for recommendation of the CMS generator group we used the MLM scheme with the same parameters used for the production of previous official samples. The results of the hadronization process are summarized in table 7.1.

	Events			Cross Section [pb]	
Process	Tried	Passed	accepted [%]	Before	After
$pp \to jj$	231789	54291	23.4 ± 1.01	$1.675 \times 10^6 \pm 4.536 \times 10^3$	$3.924 \times 10^5 \pm 1.817 \times 10^3$
$pp \to jjj$	502287	36250	7.2 ± 0.03	$3.622 \times 10^6 \pm 9.809 \times 10^3$	$2.614 \times 10^5 \pm 1.500 \times 10^3$
$pp \to jjjj$	692600	44299	6.4 ± 0.03	$4.972 \times 10^6 \pm 1.346 \times 10^4$	$3.180 \times 10^5 \pm 1.697 \times 10^3$
Total	1426676	134840	9.45 ± 0.03	$1.027 \times 10^7 \pm 1.727 \times 10^4$	$9.718 \times 10^5 \pm 2.903 \times 10^3$

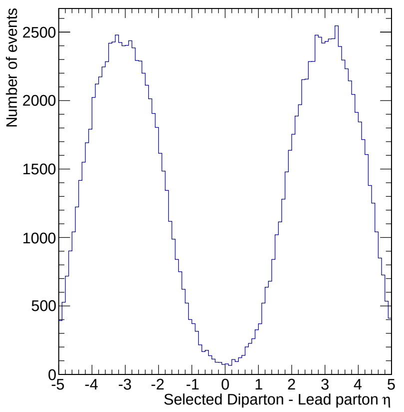
Table 7.1: Summary of the results of the Hadronization with Pythia 8 of 1.4M MadGraph events passing the parton level filter.



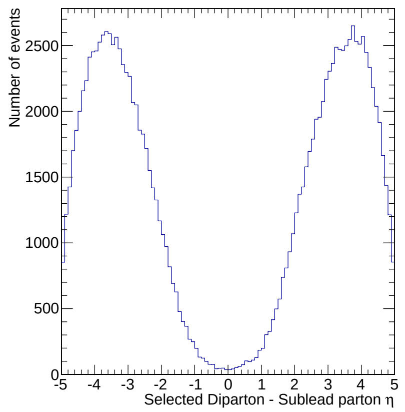
(a)



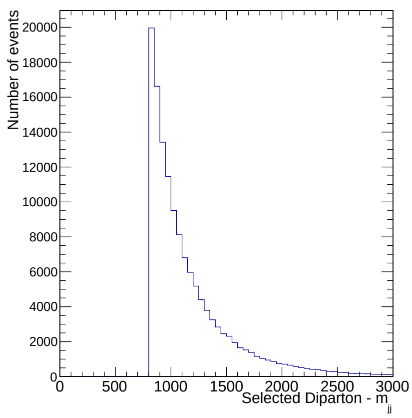
(b)



(c)



(d)



(e)

Figure 7.1: Parton p_{T}, η and di-parton $m_j j$ distributions for the leading di-parton passing cuts: parton $p_T > 30\,\mathrm{GeV}$ and $|\eta| < 5.0$, di-parton $m_{jj} > 800\,\mathrm{GeV}$

Efficiency of the post hadronization event matching has been estimated of $9.45\% \pm 0.03\%$, leading to an sample cross section of $9.718 \times 10^5 \pm 2.903 \times 10^3$ pb. The lower matching efficiency in the 3 and 4 jets final states is due to the absence of a restriction on minimum jet p_T on any additional jets to the dijet passing the parton level cuts. This jets, if low enough on energy will hardly be clusters into a jet and therefore cannot be match to its seed parton.

7.3.5 Generator level cuts

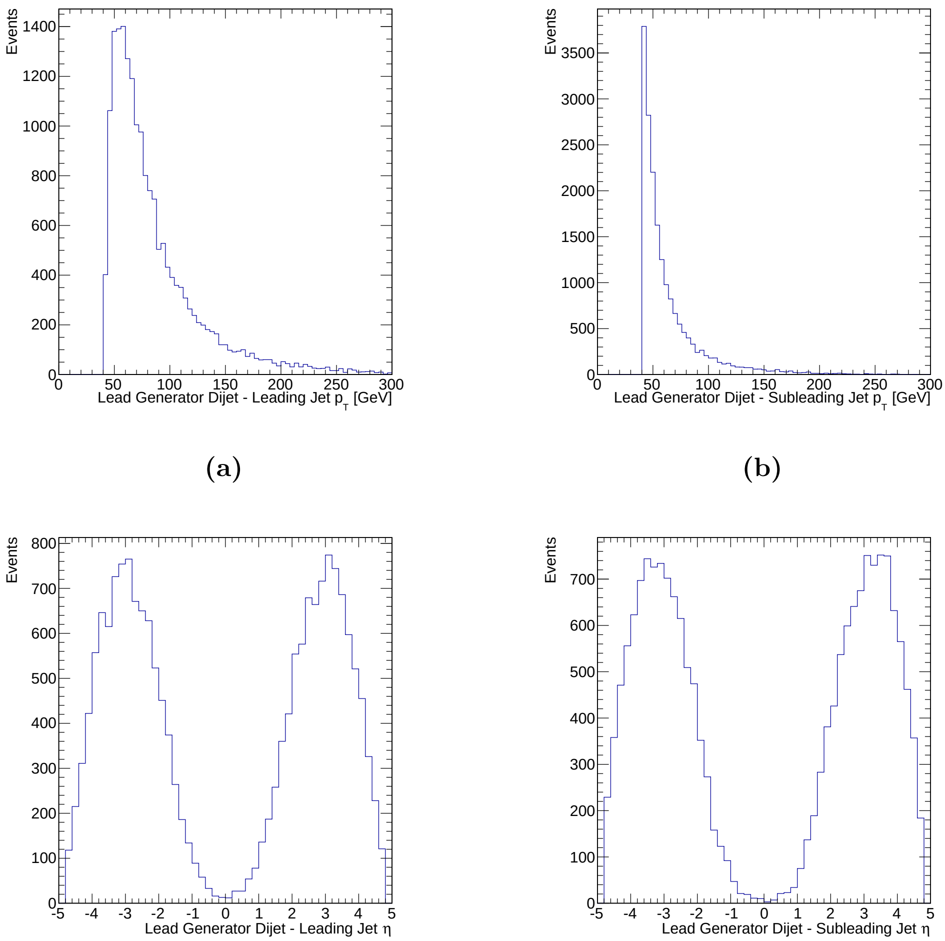
After hadronization we cluster the outgoing stable particles with the anti-k_T algorithm with $\Delta R < 0.4$ while ignoring muons. The reason to ignore muons is that CMS muon detector coverage only goes up to $|\eta| < 2.4$ so all muons outside this region will not be seen by the experiment and therefore will not be clustered into jets. Most of our signal like events will have at least one jet in the region $|\eta| > 2.4$.

We start by making an initial selection of the events with at least one generator level dijet passing $\Delta\eta > 3.0$, $m_{jj} > 1000$ GeV where both jets pass $p_\mathrm{T} > 40$ GeV and $|\eta| < 4.8$. The events passing this cuts are split into two sub-samples. Sub-sample A will have the events where the selected dijet passes $\Delta\phi <= 2.15$ and sub-sample B where at least on dijet passing all initial conditions and an inverted $\Delta\phi$ cut. Plots over all the relevant variable before the $\Delta\phi$ cut and for the leading dijet passing the cuts can be found in figure 7.2 and 7.3.

All the distributions show the expected features of the generator level filter cuts. As expected the peak of the $\Delta\phi$ distribution is at π when the 2 jets are back to back, but a tail of events is visible down to zero. A similar shape was observed in the Run I analysis before applying a cut on $min(\Delta\phi(jets, MET))$ in QCD dominated regions.

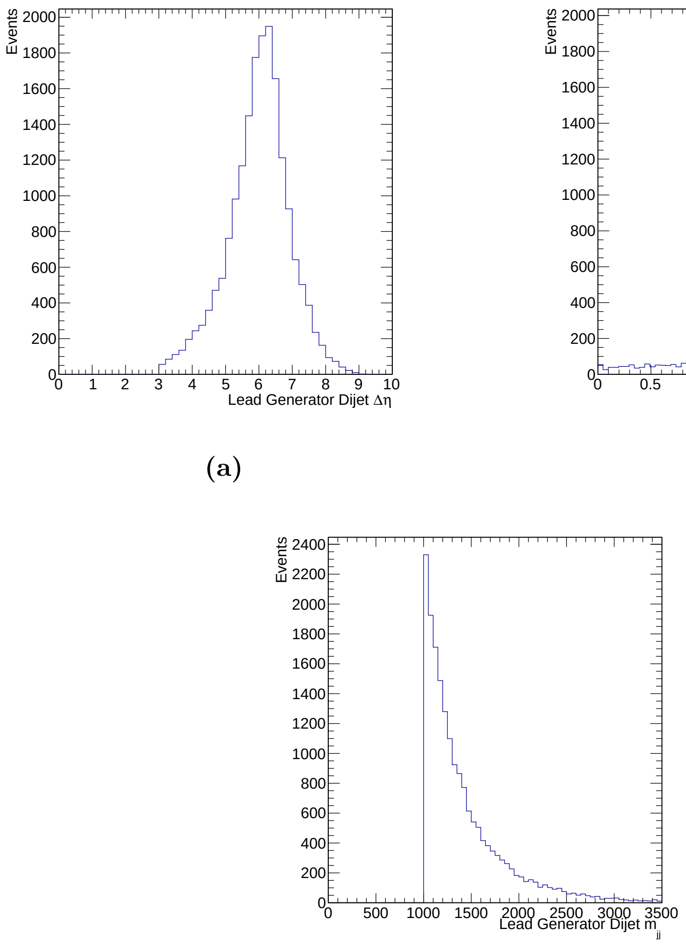
Sub-sample A will be produced by running over all the events produced up to the hadronization step. Its estimated filter efficiency is of $2.938 \times 10^{-1} \pm 4.67^{-4}$ which would lead to a sample since of approximately 29 million events and corresponding to an equivalent luminosity of over $1.1\,\mathrm{fb}^{-1}$.

Sub-sample B will result from running over only 10% of the events available at the hadronization step. This filter has an efficiency of $1.125^{-1} \pm 9.13^{-4}$ and would lead to sample of about 14 million events, corresponding to an equivalent luminosity of over 110 pb. If additional computing resources would become available this sample could be



(a) **(b)** **(c)** **(d)**

Figure 7.2: Relevant distributions for the two jets comprising the the leading dijet passing a generator filter requiring at least one dijet with $\Delta\eta > 3.0$ and $m_{jj} > 1000\,\mathrm{GeV}$ where the jets have $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ and $|\eta| < 4.8$



(a) (b)

(c)

Figure 7.3: Relevant distributions for the leading dijet passing a generator filter requiring at least one dijet with $\Delta\eta > 3.0$ and $m_{jj} > 1000\,\mathrm{GeV}$ where the jets have $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ and $|\eta| < 4.8$

expanded up to 100% of the base sample to a total of 141 million events and equivalent luminosity over 1.1 fb^{-1}.

It is necessary to determine the overlap between this two sub-samples. From table 7.2 we can see that a significant amount of events on each sub-sample have additional jets, passing all required jet conditions.

	$\Delta\phi$ cut		
N_{Jets}	no cut	< 2.15	$\gtrsim 2.15$
2	63.83 ± 0.59	15.80 ± 0.63	73.38 ± 0.70
3	23.53 ± 0.36	50.21 ± 1.13	17.59 ± 0.34
4	9.43 ± 0.23	24.34 ± 0.78	6.70 ± 0.21
5	2.42 ± 0.11	7.14 ± 0.42	1.70 ± 0.11
+6	0.79 ± 0.07	2.50 ± 0.25	0.63 ± 0.06

Table 7.2: Table showing the percentage of generator AK4 jets passing cuts $p_T > 40\,\text{GeV}$ and $|\eta| < 4.8$ for events with at least one dijet with $\Delta\eta < 3.0$ and $m_{jj} < 1000\,\text{GeV}$ and according to an additional dijet $\Delta\phi$ cut.

This additional jets lead to additional combinations that may pass the criteria of the opposite sub-sample. As it can be seen in table 7.3 in as much as 5% of the events in the $\Delta\phi <= 2.15$ sub-sample there is a second combination of two jets that would pass the criteria to be in that sub-sample.

	$\Delta\phi$ cut		
N_{Dijets}	no cut	< 2.15	$\gtrsim 2.15$
1	93.53 ± 0.71	94.29 ± 1.54	97.51 ± 0.80
2	5.84 ± 0.18	5.35 ± 0.37	2.39 ± 0.13
3	0.44 ± 0.05	0.30 ± 0.09	0.07 ± 0.02
+4	0.19 ± 0.03	0.05 ± 0.04	0.03 ± 0.01

Table 7.3: Table showing the percentage of generator AK4 dijets passing cuts $p_T^{jet} > 40\,\text{GeV}$, $|\eta|^{jet} < 4.8$, $\Delta\eta < 3.0$ and $m_{jj} < 1000\,\text{GeV}$ and according to an additional dijet $\Delta\phi$ cut.

The overlap between the two sub-samples has been determined to be 3.95% $\pm$ 0.14% of the events passing the initial selection. Since this number is relevant, and to to avoid event double counting, events with combinations that would pass both sub-sample definitions should be vetoed in one of the samples.

7.3.6 Migration study

On concern when making cuts at steps below event reconstruction is the possibility of cutting events that may pass analysis event selections. This migration of events needs to be taken into account while defining the cuts at parton and generator levels. If we take the relevant variables of signal region selection used during the 2012-13 parked data analysis, we selected a dijet with $\Delta\eta > 3.6$ and dijet $m_{jj} > 1200\,\text{GeV}$ where the lead jet $p_\text{T} > 50\,\text{GeV}$ and sub-lead jet $p_\text{T} > 45\,\text{GeV}$ and both have $|\eta| < 4.7$ (condition to guarantee the used AK5 jets are fully contained in the detector). It is unlikely that the Run II offline selection would be able to cut below jet $p_\text{T} > 50\,\text{GeV}$.

In order to study migration a second MC sample with lower parton cuts was generated. We also used MadGraph as an event generator with all the same parameters with the only difference being the dijet cuts. We now select events with at least a pair of outgoing partons with invariant mass of 600 GeV, where both parton are inside the detector volume with $|\eta| < 5.0$ and have more than $p_T > 10\,\text{GeV}$. Hadronization was then performed with the same procedure described in the previous section.

In order to compare generator jets to the partons that created them we need to match them. For each parton we select all generator jet which are located at $\Delta R < 0.4$ and from those we select the generator jet with less difference of p_T to our parton. This procedure attempts to account for situation where more than one jet is within the matching distance but the best match in p_T is not the closest one in ΔR. Using this procedure we can find a match for the di-parton passing the imposed cuts for 73.24% of the events and we also find that the matched generator jet is not the closest one in ΔR for 3.45% of the partons. A table of the matching efficiency for discriminated by physics process can be found in table 7.4.

Since we are simulating partons with fairly low p_T, two jets with $p_\text{T} > 10\,\text{GeV}$ and up to two more with no restriction on energy. It is not a surprise that in significant amount of events we cannot match all partons to generator jets. This is due to the spread of energy over a larger area then the jet algorithm can cluster and due to the default AK4 minimum p_T necessary to form a jet of 3 GeV. A set of plots of the of the relevant variable are plotted in figure 7.4. Here we take the selected di-parton and plot each variable against the matched dijet.

On plots 7.4 a), b) and f) we can see two populations. In the parton p_T plots they are along the diagonal and and along the line of generator jet p_T equal to zero and in the

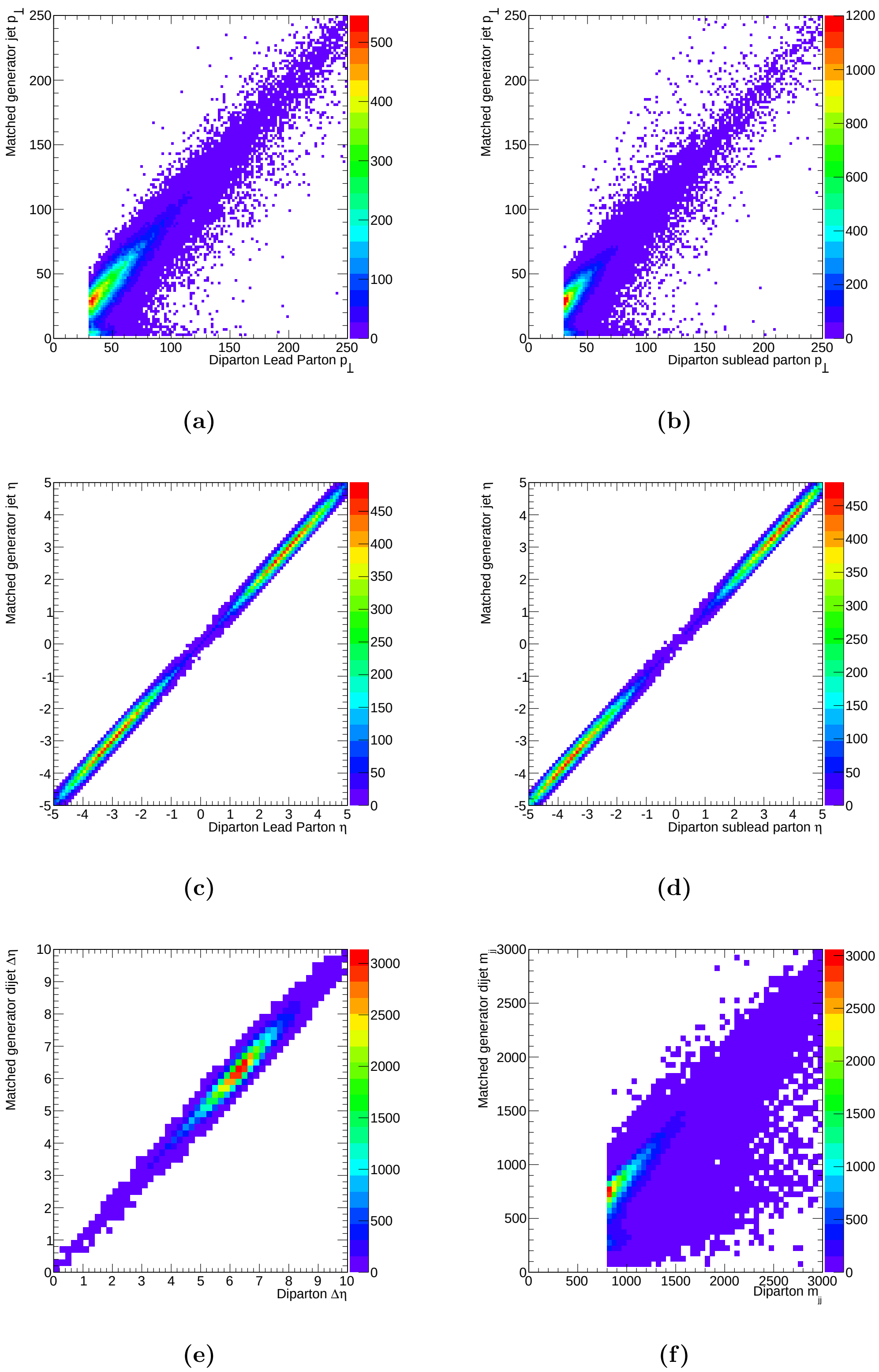


Figure 7.4: Plots for relevant variables of the selected di-parton against its matched dijet. On plot a) lead parton p_{T} b) sub-leading parton p_{T} c) lead parton η d) sub-lead parton η e) di-parton $\Delta\eta$ d) di-parton m_{jj}

n_{match}	Process			
	jj	jjj	jjjj	Total
0	22.04% ±0.22%	2.18% ±0.09%	0.14% ±0.03%	11.62% ±0.11%
1	38.60% ±0.30%	17.82% ±0.25%	3.02% ±0.13%	25.27% ±0.17%
2	39.35% ±0.30%	42.35% ±0.39%	16.91% ±0.32%	35.99% ±0.20%
3		37.65% ±0.37%	41.88% ±0.50%	19.83% ±0.15%
4			38.05% ±0.47%	7.29% ±0.09%

Table 7.4: Table showing the percentage of partons successfully matched to a generator AK4 jets. Numbers obtained for a total of 88282 events over all 3 possible hard scattering processes and for events with at least on di-parton with $m_{jj} > 600\, GeV$ where each parton has $p_{\mathrm{T}} < 10\,\mathrm{GeV}$ and $|\eta| < 5.0$

m_{jj} plot along the diagonal and along the line of $y = x/2$. This is due to the fact that at parton level the partons are perfectly matched in energy and momentum but if they are matched to only one correct generator jet and second jet with p_{T} close to zero, the system will have half the energy of the correctly assigned events.

We can now calculate the event migrations from events that did not pass will not pass the parton event selection but could have passed the generator level selection. This effect can be from jet dispersion or overlap, and clustering artefacts. Let's first consider the migrations on each variable separately, lead jet p_{T} (eq. 7.1), sub-lead jet p_{T} (eq. 7.2) and dijet m_{jj} (eq. 7.3).

$$\frac{p_T^{Parton} < 30 \wedge p_T^{GenJet} \geq 40}{p_T^{GenJet} \geq 40} = 0.27\% \pm 0.04\% \tag{7.1}$$

$$\frac{p_T^{Parton} < 30 \wedge p_T^{GenJet} \geq 40}{p_T^{GenJet} \geq 40} = 0.56\% \pm 0.08\% \tag{7.2}$$

$$\frac{m_{jj}^{Parton} < 800 \wedge m_{jj}^{GenJet} \geq 1000}{m_{jj}^{GenJet} \geq 800} = 0.13\% \pm 0.04\% \tag{7.3}$$

Now we can consider the migrations of events over all variables simultaneously in equation 7.4.

$$\frac{(p_T^{GenJet} > 40 \wedge m_{jj}^{GenJet} > 1000) \wedge (p_T^{Parton} < 30 \vee m_{jj}^{Parton} < 800)}{p_T^{GenJet} > 40 \cup m_{jj}^{GenJet} > 1000} = 0.23\% \pm 0.13\% \quad (7.4)$$

We can that that the global migrations of events from below the selected parton level cuts to above the selected generator cuts are of $0.23\% \pm 0.13\%$ of the total number events passing the generator filter. This is an acceptable value which should not bias in any relevant way the physics usage of this sample.

7.3.7 Summary

The production on new QCD MC event sample with VBF characteristics was studied and all objectives were achieved. We propose the use as MadGraph as the event generator, configured to produce proton-proton to two, three or four outgoing partons where this partons can be gluons or quarks except the top quark. At the stage we filter the events only accepting those that have at least one di-parton with $m_{jj} > 800\,\text{GeV}$ where each parton has at least 30 GeV and is contained inside the detector acceptance of $|\eta| < 5.0|$. This process has a cross section of $1.029 \times 10^7 \pm 1.614 \times 10^4$ pb.

We proceed with event hadronization using Pythia 8 event generator with MLM jet matching scheme as traditionally done in the CMS experiment. We estimate this step to have an efficiency of $9.45\% \pm 0.03\%$ which leads cross section of $9.718^5 \pm 2.903^3$ pb. From those events, we only keep the ones containing at least one generator dijet passing $\Delta\eta > 3.0$, $m_{jj} > 1000\,\text{GeV}$ where both jets pass $p_{\text{T}} > 40\,\text{GeV}$ and $|\eta| < 4.8$. We split those event into 2 sub-samples according if the dijet passing all cuts is below (sub-sample A) of above $\Delta\phi = 2.15$ (sub-sample B). The filter efficiency for sub-sample A is $2.938 \times 10^{-1} \pm 4.67^{-4}$ and it is aimed to have 1 fb of equivalent integrated luminosity. Sub-ample B filter efficiency is $1.125^{-1} \pm 9.13^{-4}$ and will have $0.1 - 1.0$ fb of equivalent integrated luminosity depending on available resources. The overlap between the two sub-samples has been estimated to be of $3.95\% \pm 0.14\%$ thus requiring care in combining them.

Migrations from events below the parton level cuts to about the generator level cuts have been determined to be $0.23\% \pm 0.13\%$ of the total number of events passing the generator filters.

The MadGraph code for event generation have been approved by the CMS MC production team. The additional code necessary for the generator level filtering has been queued for integration in the experiment software. Final approval of this sample production is under way.

Chapter 8

Conclusions

Summary of relevant results and their impact on Particle Physics

Bibliography

[1] P. W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons", *Phys. Rev.* **145** (May, 1966) 1156–1163, doi:10.1103/PhysRev.145.1156.

[2] D. Griffiths, "Introduction to Elementary Particles". Wiley, 2nd edition, 2008.

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

[4] LEP Injector Study Group Collaboration, "LEP Design Report, Volume 1: The LEP Injector Chain",.

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

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

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

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

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

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

[11] CERN, "CMS Compact Muon Solenoid",. http://public.web.cern.ch/public/Objects/LHC/CMSnc.jpg.

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

[13] CMS Collaboration, “CMS, the Compact Muon Solenoid: Technical proposal”,.

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

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

[16] CMS Collaboration, “CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems”,.

[17] CMS Collaboration Collaboration, “Data Parking and Data Scouting at the CMS Experiment”,.

[18] CMS Collaboration, “CMS Technical Design Report for the Level-1 Trigger Upgrade”,.

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

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

[21] 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.

[22] CMS Collaboration, “Measurement of Tracking Efficiency”,.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[41] CMS Collaboration Collaboration, "MET performance in 8 TeV data", Technical Report CMS-PAS-JME-12-002, CERN, Geneva, (2013).

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

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

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

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

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

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

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

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

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

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

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

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

List of Figures

List of Tables

Acronyms

GCT Global Calorimeter Trigger. 22

GSF Gaussian Sum Filter. 29

HB HCAL Barrel. 15, 65

HCAL Hadronic Calorimeter. 14–16, 27, 28, 65

HE HCAL Endcap. 15, 65

HF HCAL Forward. 15, 16, 65

HLT High Level Trigger. 20, 21, 43

HO HCAL Outer. 15, 65

L1T Level 1 Trigger. 20–22, 65

LEP Large Electron Positron collider. 4

LHC Large Hadron Collider. 4–13, 19, 21, 25, 26, 42, 43, 65

LHCb Large Hadron Collider beauty. 5

LINAC2 Linear Particle Accelerator 2. 5

LO Leading Order. 45

LS1 Long Shutdown 1. 22, 42

MC Monte Carlo. 42, 44, 45, 53, 56, 57

ME Matrix Element. 45–47

MET Missing Transverse Energy. 44, 46

NLO Next to Leading Order. 45

PF Particle Flow. 27, 28

POG Particle Object Group. 38, 39, 67

PS Proton Synchrotron. 5

PSB Proton Synchrotron Booster. 5

PU Pile-Up. 9, 32, 42, 44

PV Primary Vertex. 25, 26

QCD Quantum Chromodynamics. 10, 32, 42–44, 46, 47, 56

RF Radio Frequency. 6

RPC Resistive Plate Chamber. 18

SM Standard Model. 2, 7, 17

SPS Super Proton Synchrotron. 5, 8

SUSY Super Symmetry. 47

TSG Trigger Studies Group. 43

VBF vector Boson Fusion. 42, 44, 46, 47

VPT Vacuum Photo-Triodes. 13