SEARCH FOR CHARGINOS, NEUTRALINOS AND SLEPTONS IN OPPOSITE-SIGN
DI-LEPTON FINAL STATES AT CENTER-OF-MASS ENERGY OF 13 TEV WITH
THE CMS DETECTOR.

By

LEONORA VESTERBACKA OLSSON

A DISSERTATION Doctor of Sciences
DOCTOR OF PHILOSOPHY

ETH Zurich

2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

9

Abstract of Dissertation
Doctor of Philosophy

SEARCH FOR CHARGINOS, NEUTRALINOS AND SLEPTONS IN OPPOSITE-SIGN DI-LEPTON FINAL STATES AT CENTER-OF-MASS ENERGY OF 13 TEV WITH THE CMS DETECTOR.

By

Leonora Vesterbacka Olsson

March 2019

Chair: Professor Rainer Wallny
Major: Physics

This document presents two searches for physics beyond the Standard Model (SM), each using $35.9\mathrm{fb}^{-1}$ of proton–proton collision data collected with the CMS detector at a center-of-mass energy of 13 TeV, at the CERN Large Hadron Collider (LHC). The two searches for new phenomena is targeting electroweak production of Supersymmetric (SUSY) particles, so called Charginos, Neutralinos and sleptons, in a production mode that results in two leptons of opposite-sign and same-flavor, large missing transverse momentum, p_T^{miss}.

The document contains a summary of the theoretical framework that make up the SM and SUSY, along with a comprehensive description of the CMS experiment at the LHC accelerator complex. The two searches presented in this thesis both target the production of electroweak SUSY particles, but are divided into two types, according to the production mode. The search for Charginos and Neutralinos result in final states where two or more jets resulting from hadronization are produced, while the search for the direct production of sleptons is chrarcterized by the fact that no hadronization is expected, and thus results in a final state without any jets. The search strategies thus differ slightly, and the two strategies are presented, along with a description of the SM background processes that govern these final states.

Since no excess of collision data is observed with respect to the predicted SM backgrounds in neither of the searches, a statistical interpretation of the results yielding

12

upper limits in the production cross sections on the SUSY particles, is performed. These limits greatly extend the limits set using 8 TeV collision data during the LHC Run 1.

Concluding remarks commenting on the current absence of evidence for physics beyond the SM (BSM) are given, and an outlook highlighting the unprecedented instantaneous luminosity expected at the LHC, and the window of opportunity for searches for BSM physics that it presents.

CHAPTER 1

THE STANDARD MODEL OF PARTICLE PHYSICS

CHAPTER 2

SUPERSYMMETRY

CHAPTER 3

EXPERIMENTAL APPARATUS

This thesis uses data recorded with the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) near Geneva, Switzerland. This chapter provides a short overview of CERN and its accelerators, the LHC, as well as a short description of the main hardware components of the CMS experiment.

3.1   The Large Hadron Collider

The Large Hadron Collider (LHC) is the worlds most powerful particle accelerator[1]. It is a circular superconducting accelerator that measures 27 km in circumference and is located 100 m underground in the same tunnel that hosted the Large Electron Positron (LEP) collider[11]. The ring is divided into 8 sectors separated by eight access points to the tunnel, named interaction points (IPs). The purpose of the LHC is to accelerate protons or heavy ions up to center-of-mass energies of 14 TeV for protons and 2.76 TeV per nucleon for lead ions. The two rings in the LHC accelerate protons or heavy ions in opposite directions, to be collided at four of the eight IPs where huge particle detectors are located. The multi-purpose high luminosity experiments along the LHC ring are the CMS (Compact Muon Solenoid) [12] and ATLAS (A Toroidal LHC ApparatuS) [13] detectors located at IP5 and IP1, respectively. Both experiments collect data from

proton-proton (pp), proton-ion and ion-ion collisions. The dedicated heavy ion collision detector ALICE (A Large Ion Collider Experiment)[14] is located at IP2, that collect the same data as ATLAS and CMS. The LHCb[15] experiment is dedicated to low luminosity B-physics and collect data from pp collisions at IP8.
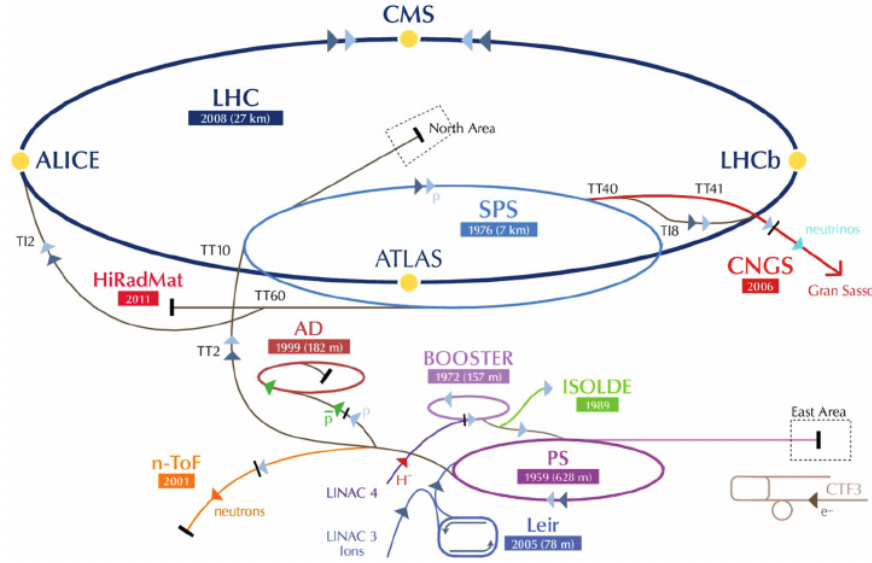
Figure 3-1. Sketch of the various accelerators and experiments hosted at CERN.

3.1.1   The accelerator chain

In order to get the protons to the center-of-mass energies quoted above, a long chain of circular and linear accelerators are needed. The protons start by being part of atoms in Hydrogen gas. Upon passing through an electric field, protons are separated from their electron and injected into the Linear Accelerator 2 (Linac 2). The Linac 2 accelerates the protons up to 50 MeV by passing them through alternating positive and negative cylindical conductors charged by radiofrequency cavities. After the Linac 2, the protons are injected into a circular collider, the Proton Synchrotron (PS) Booster, where four superimposed synchrotron rings accelerate the protons to an energy of 1.4 GeV. The protons in the Booster are kept in a circular path by magnetic dipole fields. Following the PS Booster is the PS which is a circular accelerator with a circumference of 628 meters. The protons

are accelerated by conventional magnets to an energy of 25 GeV, and they are kept on the circular path by 100 bending dipole magnets. The Super Proton Synchrotron (SPS) is the next accelerator for the protons and the last step before injection to the LHC. The SPS make use of the same technique as the PS but the larger circumference enables for the larger output energy of 450 GeV. Up until now, the protons have been travelling in "bunches", which are packets of hundred billion protons, that are separated in time by 25ns(forming so called bunch trains). Upon injection into the LHC, the proton bunch trains are split in to two trains that enter the LHC in two opposing directions. The goal of the LHC accelerator is to get these protons from 450 GeV to the maximum design energy of 7 TeV. This is achieved by letting the proton beams traverse radiofrequency (RF) cavities that are cooled down to 4.5 K using liquid Helium (up until this stage all accelerators have been operating in room temperature). The RF cavities provide a high frequency alternating electric field of 400.8 MHz. There are eight single-cell cavities per beam that produce the nominal voltage of 16 MV during storage that results in an energy gain per particle per turn is 485 keV. In addition to the RF cavities there are 1232 dipole magnets keeping the protons on the circular path. The dipole magnets measure 15 meters in length and weighing 30 tons each and are made of a niobium-titanium alloy[16]. The magnets are superconductors that can be operated at a temperature of 1.9 K and can reach a magnetic field of 8.33 T. In addition to the dipole magnets for bending the path of the protons, variuos other magnets such as decapole, sextupole and quadropole magnets for controlling or correcting the path. The accelerator chain is illustrated in Figure 3-1. Between 2010-2011 and during 2012, the LHC collided protons at a center-of-mass beam energy of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV respectively, while the energy increased to $\sqrt{s} = 13$ TeV in 2015 and 2016.

3.1.2   Beam parameters

The number of events (N_{event}) that can be generated in a collision is dictated by an interplay between theoretical predictions and engineering capabilities, summarizes as

$$N_{process} = \mathcal{L}_{int}\sigma_{process} \tag{3–1}$$

where σ_{event} is the cross section of a particular process and $\mathcal{L}_{int}$ is the so-called integrated luminosity that is defined through the *instantaneous* luminosity. The instantaneous luminosity depends only on the beam parameters and is defined as:

$$L = \frac{N_b^2 n_b f_{rev}\gamma_r}{4\pi\epsilon_n\beta^*}F \tag{3–2}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam f_{rev} the revolution frequency of each bunch. Further, the γ_r is the relativistic gamma factor, ϵ_n is the normalized beam emittance, β^* is the $\beta-$function at the collision point and F a geometrical factor inversely proportional to the crossing angle of the two beams at the IP. The peak LHC instantaneous luminosity is $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}s^{-1}$, and is reached for the beam parameters summarized in Table 3-1 By integrating the instantaneous luminosity L (that has the unit of $\mathrm{cm}^{-2}s^{-1}$) over time, the result is the integrated luminosity, $\mathcal{L}_{int}$ (which has the unit cm^{-2}). By integrating the instantaneous luminosity L (that has the unit of

Table 3-1. Beam parameters for beams in the LHC at injection and collision energy [1].

Parameter	Injection	Collision
Beam energy [GeV]	450	7000
Relativistic gamma factor (γ_r)	479.6	7461
Beam emittance (ϵ_n) [μ rad]	3.5	3.75
Half crossing angle [μrad]	±160	±142.5
$\beta-$function (β^*) [m]	18	0.55
Revolution frequency (f_{rev}) [Hz]	11245	
Number of bunches (n_b)	2808	
Particles per bunch (N_b)	1.15×10^{11}	

$cm^{-2}s^{-1}$) over time, the result is the so called integrated luminosity, denoted by $\mathcal{L}_{int}$. The integrated luminosity has the unit cm^{-2} which is the inverse of the unit of cross section, which makes the resulting number of events in a process in Eq. 3–1 unitless. The peak instantaneous stable luminosity reached by the LHC during 2016 was $1.527 \times 10^{34}\, cm^2 s^1$ and the total integrated luminosity greatly surpassed the predictions and expectations.

3.2   The CMS experiment

The CMS experiment is located $100\,m$ under ground at interaction point number five ("Point 5") in Cessy, France, along the LHC ring. A full description of the CMS experiment can be found in [12] and a sketch is shown in Figure 3-2. This multi-purpose
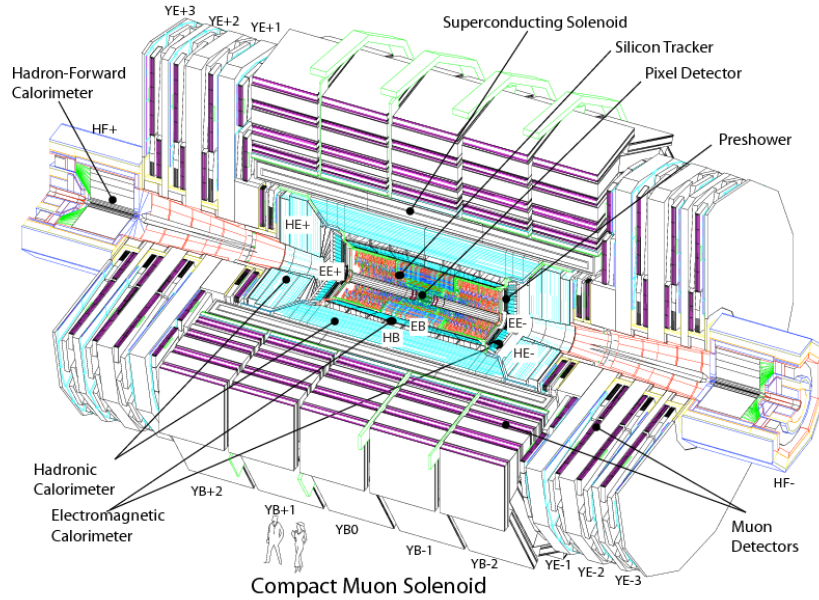
Figure 3-2. General view of the CMS detector. The major detector components are indicated, together with the acronyms for the various CMS construction modules.[2]

apparatus is built around the LHC beam line with circular layers of subdetectors with increasing radii, resulting in a cylindrical shape with a diameter of $16\,\mathrm{m}$ and length of $22\,\mathrm{m}$. The $z-$axis is defined to be the direction along the LHC beamline inside CMS, with the positive direction pointing towards the Jura mountains from Point 5. The cylindrical shape lends itself well to the use of cylindrical coordinates, where the the azimuthal angle ϕ is defined in the transverse $x - y$ plane perpendicular to the beam line and the polar angle θ is measured from the $z-$axis, as shown in Figure 3-3. For relativistic particles
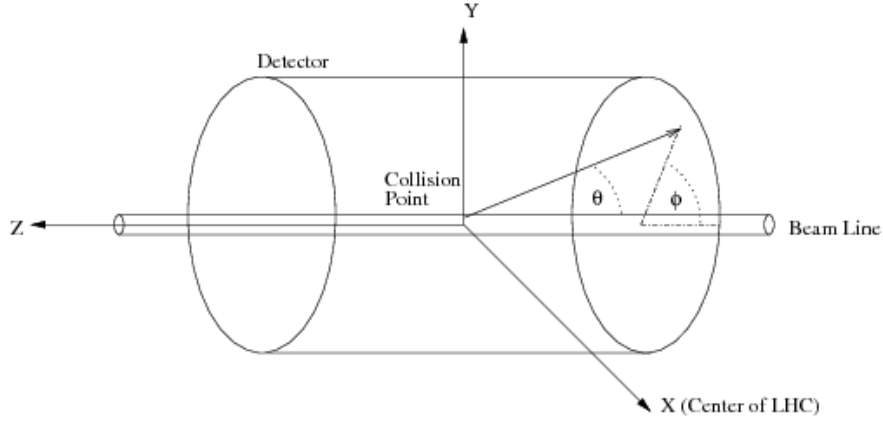
Figure 3-3. Illustration of the CMS coordinate system.[3]

it is common practice to describe the trajectory of the particles using the so called *pseudorapidity* η rather than θ, where η is defined as

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] = \frac{1}{2}\ln\left(\frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z}\right) \qquad (3\text{--}3)$$

where $\mathbf{p}$ is the three-momentum and p_z the momentum along the $z-$axis. This quantity will be heavily used in the rest of the section to define the coverage of the various subdetectors, and beyond this chapter to define the invariant angular distance between particles as $(\Delta R)^2 = (\Delta \eta)^2 + (\Delta \phi)^2$. Additionally, the transverse component of variables is more commonly used in relativsitc collisions, e.g. transverse momentum p_{T}, transverse energy E_{T} and missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$. In 2016, the CMS experiment recorded data from pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ corresponding to $37.8\,\mathrm{fb}^{-1}$ of integrated luminosity,

as seen in Figure 3-4, of which $35.9\,\mathrm{fb}^{-1}$ was declared to be good for analysis and used throughout this thesis.

CMS Integrated Luminosity, pp, 2016, $\sqrt{s} = $ 13 TeV

Data included from 2016-04-22 22:48 to 2016-10-27 14:12 UTC
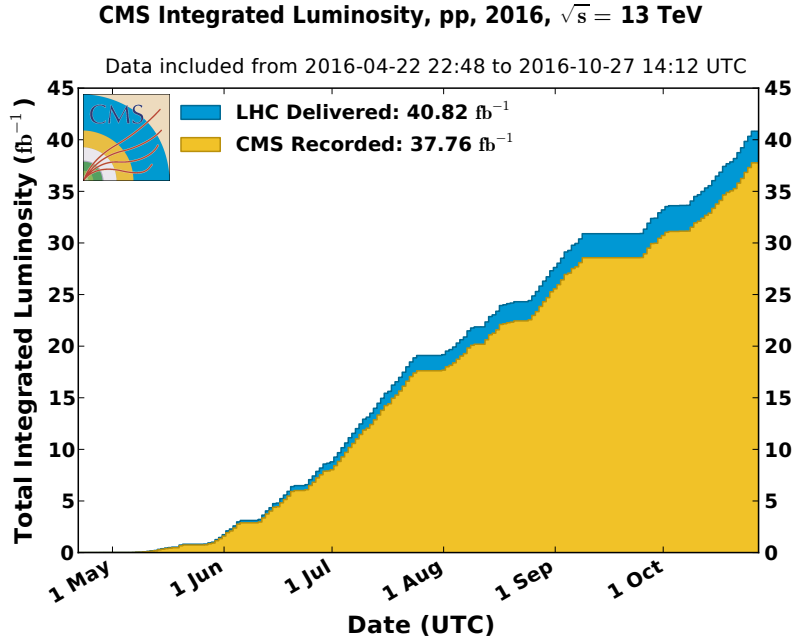
Figure 3-4. Illustration of the CMS coordinate system.[4]

In the following subsections, the magnet and the various subdetectors are introduced.

3.2.1   The Magnet

The word "Solenoid" in Compact Muon Solenoid refers to the superconducting solenoid magnet that the experiment is built around. A strong magnetic field is a key feature at an experiment of this magnitude as it is essential for momentum resolution of charged particles in the tracker. The solenoid measures a length of $12.9\,\mathrm{m}$ and six meters in inner diameter. The large inner radius allows for the tracker and calorimeters to be fully contained within the solenoid. This design reduces the material between the calorimeters and the IP which is desireable for precise momentum measurements. The solenoid is composed of a conducting material, a NbTi alloy, subjected to a current of $19\,\mathrm{kA}$ that results in a magnetic field of $3.8\,\mathrm{T}$. The magnet system includes a return yoke to keep the

magnetic field lines homogenous with the distance from the IP, and total weight including the solenoid is 11 000 tons. [17]

3.2.2 The Tracker

The CMS inner tracker measures a length of 5.8 m and a diameter of 2.5 m, and the overview that will be presented in the following is visualized in Figure 3-5 and follows the description in [5].
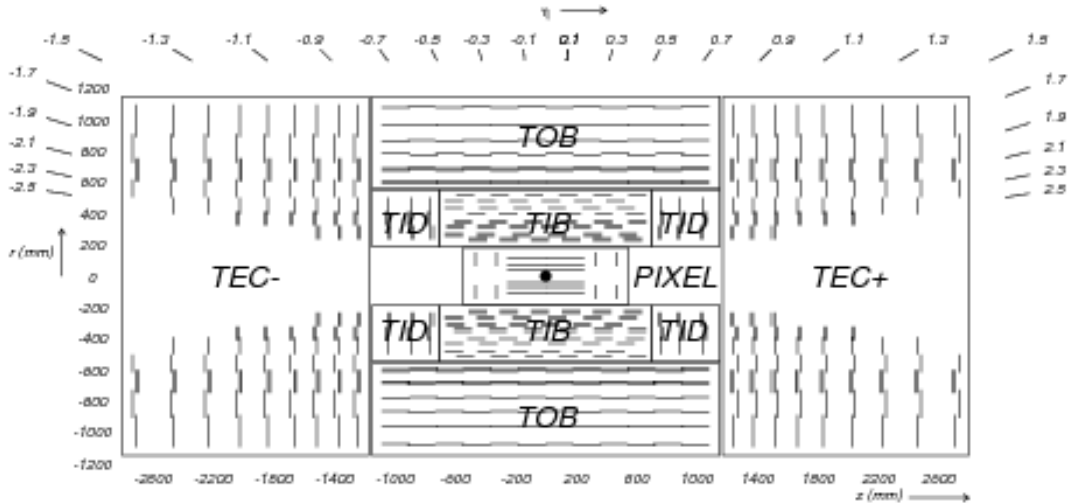
Figure 3-5. Schematic cross section through the CMS tracker in the r-z plane. Each line-element represents a detector module. Closely spaced double line-elements indicate back-to-back silicon strip modules, in which one module is rotated through a 'stereo' angle, so as to permit reconstruction of the hit positions in 3-D. Within a given layer, each module is shifted slightly in r or z with respect to its neighbouring modules, which allows them to overlap, thereby avoiding gaps in the acceptance [5].

In order to achieve a high efficiency in tagging b-jets, it is crucial that the tracker layers are as close to the LHC beamline as possible. Closest to the beamline is the smallest subdetector of the tracker, the silicon pixel tracker, consisting of a barrel (BPix) of three layers of radii 4.4, 7.3, 10.2 cm, and two pairs of endcap disks (FPix) at a distance of $|z| = 34.5$ cm and $|z| = 46.5$ cm away from the interaction point. These layers provide a

three dimensional position measurement with a transverse coordinate position cresolution of $10\,\mu$m and a longitudinal coordinate position resolution of $20-40\,\mu$m. The total number of pixels in the BPix and FPix amounts to 66 million and cover in total an area of $1\,m^2$. Outside of the pixel tracker are four subsystems of 9.3 million silicon micro strips. The tracker inner barrel (TIB) is composed of four layers and cover a region of radius $20\,cm$ to $55\,cm$. The pitch[1] is $80\,\mu$m for the first two layers of the TIB and $120\,\mu$m for the next two layers, and they are all oriented parallel to the beamline[18]. The tracker inner disks (TID) complements the TIB with three disks on each side that extends to coverage in the $|z|$ direction to $\pm118\,cm$. The TID modules are built using radially placed sensors forming a wedge shape with pitch ranging from 81 to $158\,\mu$m. Outside of the TIB, the tracker outer barrel (TOB) consisting of six barrel layers cover the region beyond $55\,cm$ in radius and the same $|z|$ coverage as the TID along the beamline. The TOB has thicker strip sensors of $500\,\mu$m and the first four layers use strips with a pitch of $183\,\mu$m and the last two layers use strips with a pitch of $122\,\mu$m. Finally, the last tracker subsystem is the tracker endcap (TEC) that consist of nine disks on each side of the TIDs and TOB, with a total coverage of the region $124 < |z| < 282\,cm$. The TOB and TEC have a resolution of the position measurement ranging from $18-47\,\mu$m. Putting all of this together, the inner tracker fully covers the region $|\eta| < 2.4$ and guarantees in total 9 hit measurements. The performance is quantified in terms of the resolution of the p_T for single muons, which results in 0.65 to 1.5% at $\eta = 0$ and 1 to 2% at $|\eta| = 1.6$ for muons of 10 and $100\,GeV$.

3.2.3   The ECAL

The choice of a high resolution CMS electromagnetic calorimter (ECAL) is motivated by the physics potential that was forseen for the discovery of the Higgs boson in 2012,

---

[1] The pitch is the distance between the p+ implants in the n type $Si0_2$ bulk.

as photons and electrons are key ingredients in at least three decay channels[19] This motivation lead to the choice of the design of a hermetic, homogeneous, fine grained scintillating calorimeter consisting of 75,848 lead tungstate ($PbWO_4$) crystals. The crystals are arranged in a central barrel region ($|\eta| \leq 1.48$) have a crystal length of $23\,\text{cm}$ and a front face of $2.2 \times 2.2\,\text{cm}^2$. The crystals arranged in the two ECAL endcap regions (up to $|\eta| = 3.0$) measure $22\,\text{cm}$ in length and a front face of $2.86 \times 2.86\,\text{cm}^2$. In the barrel, the crystals are organized into 36 supermodules, each containing 1,700 crystals. In the endcaps, the crystals are organized into dees each containing 3662 crystals. The layout of the supermodules and dees is presented in Figure 3-6.
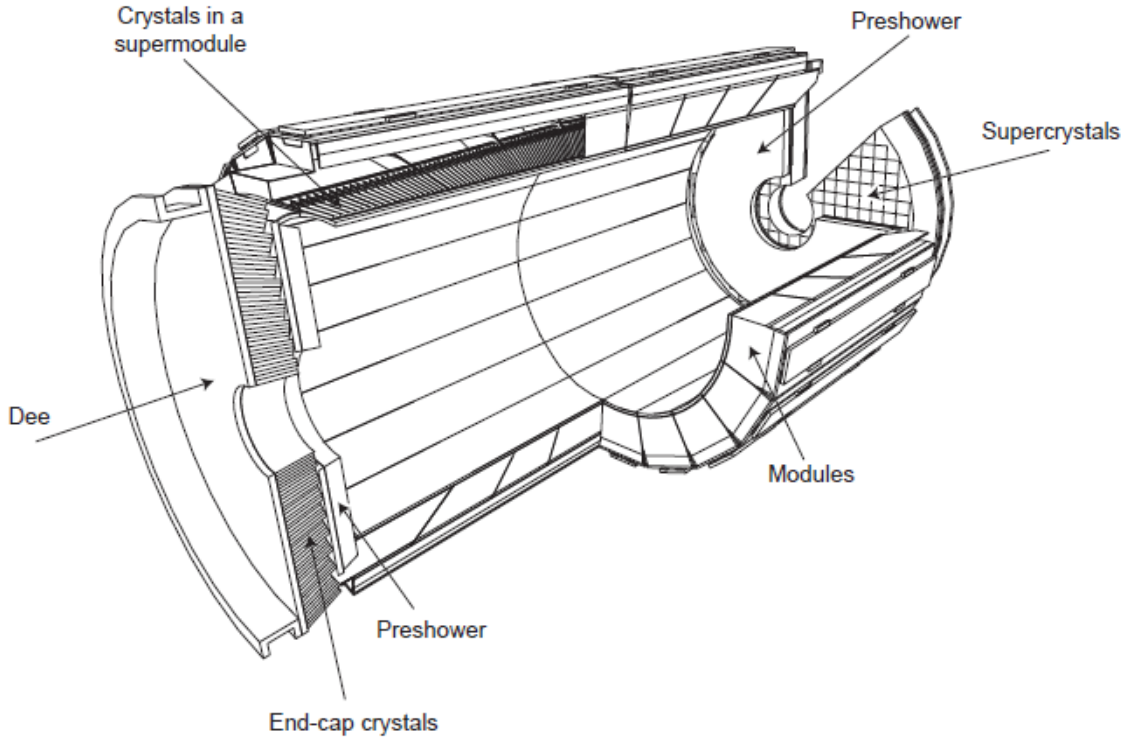
Figure 3-6. Layout of the CMS ECAL, showing the barrel supermodules, the two endcaps and the preshower detectors. The ECAL barrel coverage is up to $|\eta| = 1.48$, the endcaps extend the coverage to $|\eta| = 3.0$ and the preshower detector fiducial area is approximately $1.65 \leq |\eta| \leq 2.6$.[6].

In addition to the barrel and endcaps, the ECAl also consists of a preshower detector which is based on lead absorber and silicon strip sensors. The preshower covers a region of $1.65 \leq |\eta| \leq 2.6$ and is motivated by the ability to improve the differentiation of the

$\pi^0 \to \gamma\gamma$ process from $h \to \gamma\gamma$. The lead tungstate scintillating crystals are of high density $(\rho = 8.28g/\,\mathrm{cm}^3)$ and have the nice feature of short radiation length (X_0) and small Molière radius (R_M), where these quantities are related according to

$$R_M = 0.0265 X_0 (Z + 1.2) \tag{3-4}$$

where Z is the atomic number. The energy absorbed in the crystals by the incoming electrons and photons is emitted as light, 80% of the timw within 25ns. The energy resolution obtained with this design has been quantified at beam tests and results in

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\% \tag{3-5}$$

where the order of the terms are stochastic, noise and constant term respectively[6]. The light produced in the scintillating crystals is collected by avalanche photodiodes (APDs) in the barrel and by vacuum phototriodes (VPTs) in the endcaps. The APDs are sensitive to changes in temperature according to $-2.3\%^\circ C$. In the very front end (VFE) cards the signals from the APDs are pre-amplified and shaped by an ASIC multi-gain pre-amplifier chip. Trigger towers are formed by 5×5 crystals, and the trigger primitives are generated from the summed amplitudes of these 25 crystals in the front end (FE) cards.

3.2.4   The HCAL

The hadronic calorimeter is located outside of the ECAL while still being, mostly, contained within the solenoid. In contrast to the ECAL, the HCAL has the purpose to identify a variety of particles, mainly hadronic jets but also help in identifying electrons, photons and muons with information from the ECAL and muon systems. The ECAL is designed to fully contain the electromagnetic shower induced by electrons or photons, and the HCAL, being the next layer, is designed to contain hadronic showers and fully absorb hadrons before the solenoid. As much larger depth of detector material is required compared to that of the ECAL, hadronic calorimetry is considered much more challenging.

Additionally, the HCAL energy resolution is also worse compared to that of the ECAL due to intrinsic fluctuations[2] The HCAL layout consist of fours sections as shown in Figure 3-7.
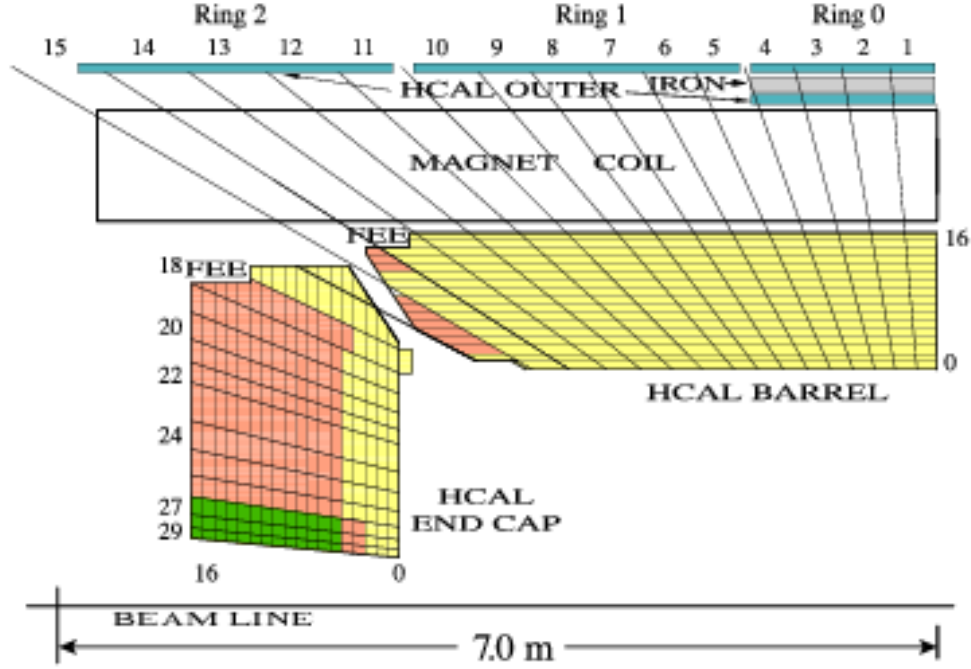
Figure 3-7. The HCAL tower segmentation for one-fourth of the HB, HO, and HE detectors. The numbers on top and on the left refer to the tower numbers. The numbers on the right and on the bottom (0-16) indicate the scintillator layers numbers inserted into slots in the absorber. The shading represents independent longitudinal readouts in the HB/HE overlap and the small angle regions.[7].

The barrel (HB) covers the $0 \leq |\eta| \leq 1.4$ range and the endcaps (HE) cover the $1.4 \leq |\eta| \leq 3$ range, and both are sampling calorimeters made by layers of brass absorber alternated with plastic scintillator. As the depth of the HB is limited to what can be contained within the solenoid, the thickness at $\eta = 0$ only covers 5.8 nuclear interaction

---

[2] Intrinsic fluctuations are a result of the significant incoming energy fraction being invisible, since it is employed in processes like nuclear break-up.

lengths (λ_I) which is increased to $10\lambda_I$ at $|\eta| = 1.2$[3] . Only about 95% of hadrons above $p_T \geq 100\,\text{GeV}$ are fully contained in the HCAL, leading to a small fraction of hadrons making it through the solenoid. For this reason, an Outer HCAL (HO) is located outside of the solenoid to recover these hadrons, and uses the same plastic scintillator as HB and HE but the magnet material as the absorber. The HB, HE and HO sections all use wave-length shifting fibers to extract the scintillating light that is guided to hybrid photodiodes (HPDs). The forward part of the HCAL (HF) covers the range $3 \leq |\eta| \leq 5$ and is located $11.1\,\text{m}$ from the interaction point. The purpose of the HF is to improve the measurement of p_T^{miss} and to identify very forward jets, as those produced in vector boson fusion (VBF). The region covered by the HF is subject to the largest particle flux with radiation doses reaching $100\,\text{Mrad/year}$. For this reason, a different construction choice is made for this region, exploiting Cherenkov calorimetry technique. The HF is composed of steel absorber interspersed with quartz core and acrylic clad fibers in the longitudinal direction[20]. The fibers collect the Cherenkov light produced be the showers in the absorbers, and the light is sent to a photomultiplier tube (PMT).

3.2.5   The muon system

The final layer of the CMS detector is dedicated to the measurement of muons, and is the origin of the "Muon" in Compact Muon Solenoid. The objective of this subsystem is to identify muons, trigger on muons and measure the momentum and charge of muons. A sketch of the experimental setup is shown in Figure 3-8 and the following descriptions follow closely that in[8]. A combination of three detector techniques is employed that are motivated by the various expected experimental implications. Similarly to the subdetector

---

[3] The nuclear interaction length is the mean-free path that an incident hadron can travel in a medium before it is fully absorbed due to nuclear interaction. It is defined as $\frac{1}{\lambda_I} = \sigma_{inel}\frac{N_A \cdot \rho}{A}$, where σ_{inel} is the inelastic cross-section, ρ the density and A the atomic mass.
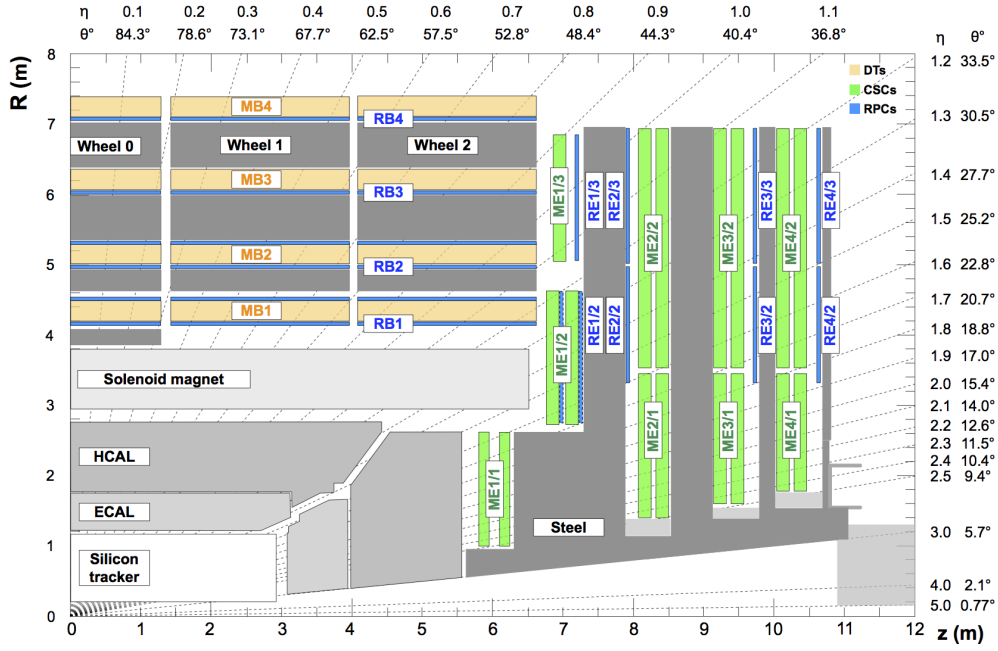
Figure 3-8. Cross section of a quadrant of the CMS detector with the axis parallel to the beam (z) running horizontally and the radius (R) increasing upward. The interaction point is at the lower left corner. The locations of the various muon stations and the steel flux-return disks (dark areas) are shown. The DTs are labeled MB ("Muon Barrel") and the CSCs are labeled ME ("Muon Endcap"). RPCs are mounted in both the barrel and endcaps of CMS, where they are labeled RB and RE, respectively.[8].

systems described in the previous sections, the muon systems consists of a barrel and two endcaps, where the barrel is interleaved with layers of the steel flux-return yoke. The detector technique employed in both the barrel and endcaps is gas ionization particle detectors and each of the modules are commonly referred to as chambers. As the muon system is located outside of the solenoid, the effective magnetic field is diminished to below $0.4\,\mathrm{T}$ between the yoke segments in the barrel. This results in non-uniform magnetic field strengths and a reversion of the muon trajectory. In addition to the relatively low magnetic field in the barrel, the expected rate of muons is low, thus making the use of drift tubes (DT) a suitable choice. The barrel DT chambers cover a region of $|\eta| \leq 1.2$ and is split in 12 segments around ϕ. The DT chambers consist of drift cells. Each cell contains a gold plated stainless steel anode wire operating at $3.6\,\mathrm{kV}$, surrounded by a gas mixture of 85% of Argon and 15% CO_2. This results in a drift time of $400ns$. Cathode plates on the sides of the cells operate at $\pm 1.8\,\mathrm{kV}$. An incident muon will release electrons in the gas volume that will drift to the anode and produce avalanches in the region close to the wire where the field strength is increased. Four layers of parallel cells form a super layer (SL) and a chamber consists of three SLs where one layer is oriented orthogonally to the other two in order to enable precise measurement in both the $r - \phi$ and the $r - z$ directions. The muon system endcaps cover a region $(0.9 \leq |\eta| \leq 2.4)$ where the expected rate of muons and neutron background is much higher than that in the barrel. As this requires a faster response, Cathode Strip Chambers (CSCs) that are multiwire proportional counters are used, motivated by the shorter drift path than that of the DTs. Additionally, the CSCs can tolerate the higher magnetic field that the muon endcap regions is subjected to. Each endcap consist of four rings of chambers mounted on the face of the endcap steel disks. A CSC consist of 6 layers, and the cathode strips are aligned radially outwards while the anode wires run perpendicularly to the strips, allowing for a position measurement in $r - \phi$. Each layer has 80 cathode strips. The chambers are filled with a gas mixture of 50% CO_2, 40% Argon and 10% CF_4, and the anode wires operate

at $2.9\,\mathrm{kV}$ to $3.6\,\mathrm{kV}$ depending on the ring. Interspersed in the barrel DT layers and endcap CSC layers are resistive plate chambers (RPCs). The purpose of these layers is to provide fast and independent triggering at lower p_{T} thresholds in the region $|\eta| \leq 1.6$. The RPCs are double gap chambers, where each gap consist of resistive Bakelite plates separated by a gas filled gap volume of the same thickness. When a charged particle crosses the RPC, the gas in the gap volumes is ionized and an avalanche is created by the high electrid fiels (due to the application of a $9.6\,\mathrm{kV}$ voltage) and an image is induced that is picked up by the readout strips.

CHAPTER 4

EVENT RECONSTRUCTION AND SIMULATION

The basis of any LHC data analysis relies on the concept of an 'event'. An event refers to a succesful collision of two protons that results in the full readout of the CMS detector, and subsequently the combination of subdetector information that forms physics objects. The LHC delivers proton bunches at a rate of 40MHzbut only a small fraction of the protons in the bunches result in collisions of interest to the CMS physics program. Additionally, if all collisions would lead to a full readout of the detector, there would not be enough bandwidth to readout the information, nor space to store it for offline analysis. With these limitations in mind, the CMS has developed a two-tier system to select events of physics interest, the first level (L1) trigger and the high level trigger (HLT). The trigger system is presented in the following section. Furthermore, the way in which the subdetector information is combined to form physics objects, namely the Particle Flow technique, is presented in this section. Apart from only analyzing the recorded events that pass the two-tier trigger system, simulated events of the CMS detector are used to validate the performance of the physics object reconstruction, for the estimation of SM background processes and for visualizing how a potential BSM signal could be found. The last part of this section is presenting how the physics processes are generated and how their interaction with the detector material is simulated.

4.1 Trigger

During 2016 proton-proton run of the LHC, the LHC delivered proton bunches with a time separation of 25ns, and a peak luminosity reaching the unprecedented value of $10^{-34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. At this peak luminosity, the pp interaction rate exceeds 1GHz, as the mean of the number of interaction per bunch is 25, which is a rate impossible to readout with the technology to date. A two tier triggering system helps differentiate which of the 1GHzcontain interesting events, while discarding the rest. This enables a reduction of the rate down to 400Hz, which is feasible for offline storage.

4.1.1 L1 trigger

The L1 trigger is a hardware system that uses input from the calorimeters and the muon detectors to make a decision to keep the event or not [21]. Tracking information is not included at this step, as it is too time consuming. Instead, information from the
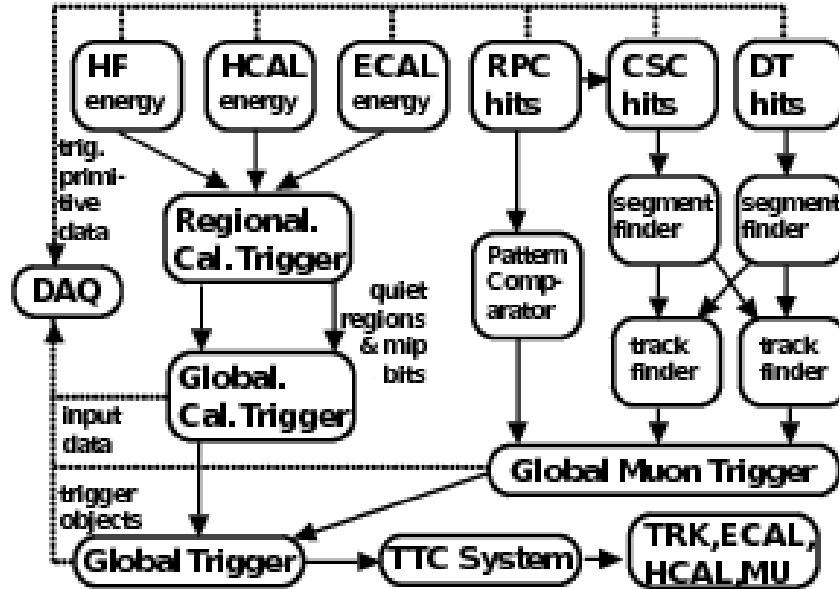
Figure 4-1. Chart showing the organization of the various components of any L1-accept. The calorimetry and muons systems work in parallel and are combined into a global trigger.

calorimeters and the muon detectors is combined into a global trigger, as illustrated in Figure 4-1. The L1 calorimeter trigger uses inputs like transverse energy and quality flags of the ECAL, HCAL and HF, in the form of Trigger Primitives (TP) from coarsely grouped trigger towers of the calorimeters. This information is provided to the regional calorimeter trigger (RCT), where it is combined to form $e\gamma$ candidates. The next step is the global calorimeter trigger (GCT), where jets are formed using the sum of transverse energy E_T, and the information on the pseudorapidity and τ veto from the RCT is used to label them as central, forward and τ jets. A crude calculation of the $p_\mathrm{T}^\mathrm{miss}$ can also be performed at L1, by summing the x and y components of the transverse energy in quadrature and rotating the vector by 180°. The objects returned by the GCT, isolated and non-isolated $e\gamma$ candidates, central, forward and τ jets and $p_\mathrm{T}^\mathrm{miss}$, are passed on to the global trigger (GT) and the information is used, together with that of the muon triggers, to decide if an event is kept or not. At this time, there has also been a collection of information gathered from the muon system. Again, no information from the tracker is used at L1, but muons can still be more or less efficiently identified by the muon system. Using the CSC and DT track finders, the tracks of muons can be identified and their p_T. This information, together with muon trigger candidate hits in the RPC, is sent to the global muon trigger (GMT). The information from the GCT and GMT is combined and a decision is made whether to keep the event or not, so called L1-accept. This trigger system has now brought down the rate of 1GHzto 100kHz, and the L1-accept is passed to all subdetectors that are read out and passed on the HLT. A way to further reduce the rates is to scale them down. Some processes with large cross sections, such as QCD, produces single photon events at rates higher than managable, especially at low photon p_T. For this reason, the single photon triggers are *prescaled*, meaning only a fraction of the events are recorded, and the fraction is evolving with the luminosity during data-taking. At analysis level, the events recorded with the prescaled triggers are scaled up according to whatever

value they were prescaled with. In this thesis, these kinds of prescaled triggers are used when collecting the single photon sample used for the $p_\mathrm{T}^\mathrm{miss}$ performance studies.

4.1.2 The HLT

While the L1 is completely hardware based and process the information from the detector underground in the experimental cavern, the HLT is both software and hardware based and located in computer farms on the ground level, running on 13,000 CPU cores. At this level, a more thorough object reconstruction is performed with the L1 information. The so called HLT path is a set of algorithms executed in a sequence of steps. As the tracking is more computing expensive, the first steps is to make a requirement on information from the calorimeters and the muon detectors, before performing the track reconstruction. The basic idea is to enable triggering on high quality objects that eventually can be reconstucted offline, while keeping the rates to a minimum. For this reason, variables such as the p_T and isolation of an object is used in different combination. The rate for triggering on a low p_T muon would be very high if no other requirements are imposed, but if one further requires a well isolated muon, this reduces the rate and possible misidentification of the muon. Conversely, as the production rate of higher p_T muons is lower, one can afford to only impose a p_T requirement and still keep the rate low. As will be seen later in the thesis, this is the reason for the use of the different HLT paths such as `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL` (involving isolation requirement on both muons) and `HLT_Mu30_TkMu11` (involving higher p_T requirements and no isolation requirements), to ensure triggering on all possible events with interesting physics. Further, HLT paths can involve objects like the Calo $p_\mathrm{T}^\mathrm{miss}$ (computed using only calorimeter deposits), PF $p_\mathrm{T}^\mathrm{miss}$ (computed using only PF jets), H_T (scalar sum of all jet p_T above a threshold) and $\not{H}_\mathrm{T}$ (the missing H_T). The $p_\mathrm{T}^\mathrm{miss}$ triggers are most sensitive to triggering on anomalous events where the large $p_\mathrm{T}^\mathrm{miss}$ is originating from noise, beam halo or other sources, which will be discussed in 8. In order to keep low rates for these triggers, noise

cleaning algorithms are applied at the HLT, and energy deposits associated to beam halo or HB/HE noise is removed from the energy sum and the calorimeter based $p_{\mathrm{T}}^{\mathrm{miss}}$ is recomputed. The noise cleaning algorithms are fully efficient in Run 2 and reduce the rate by a factor of 2.5. Additionally, there is even the possibility to get a better estimation of the $p_{\mathrm{T}}^{\mathrm{miss}}$ at HLT, by propagating the JECs to the jets and in the computation of the PF $p_{\mathrm{T}}^{\mathrm{miss}}$. This results in an improved efficiency and a further rate reduction. This thesis contains a detailed study on the performance of the $p_{\mathrm{T}}^{\mathrm{miss}}$ in Run 2, and related to this, Figure 4-2 is showing the efficiencies of the different $p_{\mathrm{T}}^{\mathrm{miss}}$ algorithms at L1 and at HLT compared to the offline $p_{\mathrm{T}}^{\mathrm{miss}}$. The HLT reduces the rate from the L1 output of 100kHzto
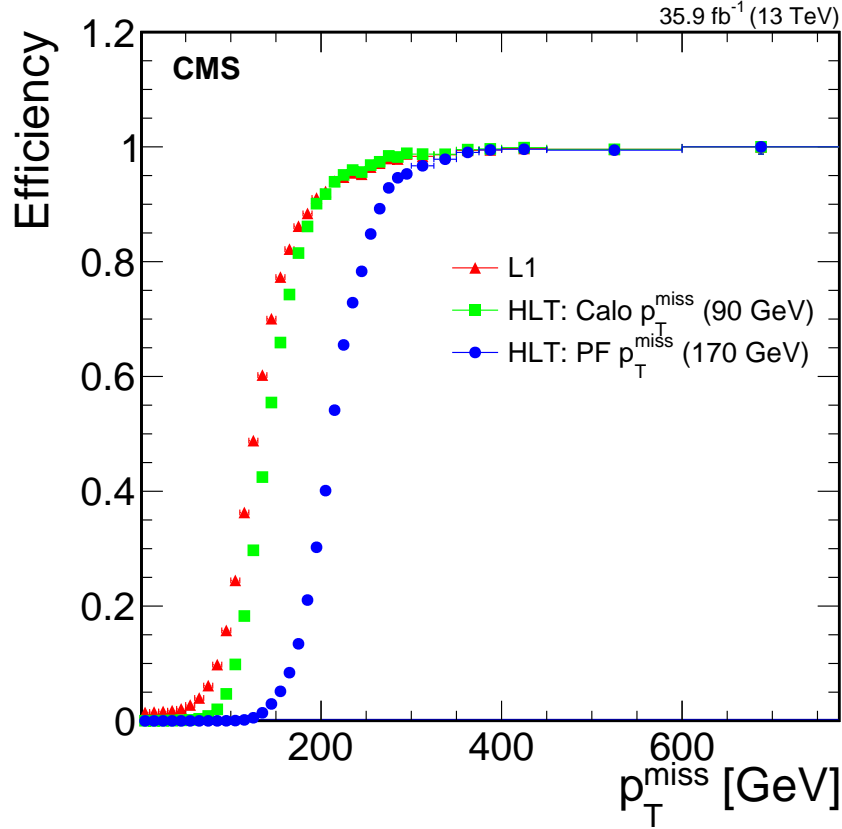
Figure 4-2. The $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency measured in the single-electron sample. The efficiency of each reconstruction algorithm, namely the L1, the calorimeter and the PF based $p_{\mathrm{T}}^{\mathrm{miss}}$ algorithms, is shown separately. The numbers in parenthesis correspond to the online $p_{\mathrm{T}}^{\mathrm{miss}}$ thresholds.

1kHz. The data that now have passed the L1 and HLT will be sent of the CERN Tier-0,

described in the subsequent section. As the following reconstruction of the data is very computing intense and time consuming, there is the possibility to only use the cores to process some of the data immediately, while some is left for later reconstruction. This concept of data-parking has been used during both Run 1 and 2 of the LHC, and is at the time of writing used to park some of the b-physics data collected during the end of the Run 2.

4.2 Data reconstruction

The computing of the LHC experiments is in large parts made in a four tiered computing system, with decending order of importance. The Tier-0 is the CERN Data center, located onsite in Geneva, and is the first line of tiers that the data collected by the detectors at the LHC experiment is processed at. The Tier-1 consists of 13 computing centers connected to the Grid located all over the world, and share the data reconstruction with the Tier-0, along with providing storage. The next computing tiers are the Tier-2 and Tier-3, which consists of computing resources at universities and institutes where analysis work is for example performed and stored. Once the CMS HLT has decided what events to keep, they are sent to the Tier-0 at CERN, where the reconstruction of the event is initialized. The CMS software, CMSSW, is a centrally maintained code base that reconstructs objects out of calorimeter deposits, hits in the muon chambers and tracks in the pixel and silicon trackers, and stores them in *Event Data Model* format (EDM).

4.3 Simulated events

Any search or measurement at the LHC are relying on simulated events, so called Monte Carlo (MC) generated events. Physical processes are simulated in a chain, starting from the foundations dictated by the theoretical framework as described in Chapter1, followed by the decay, radiation and hadronization of the particles produced, and finally,

the simulation of the interaction of the generated process with the detector material. Events are generated from both known SM processes to predict some backgrounds in the searches, and in order to validate the data-driven backgrounds. As the SUSY signals analyzed in this thesis are mulitple, and include numerous assumptions on the masses of the SUSY particles, the full chain of generation for all signal scenarious would be too computationally heavy. For this reason, a faster simulation package is used for the SUSY signals on the level of the detector simulation step, that greatly reduced the computational time by paramtrizing the interactions. The gain in the reduction of the computational time is traded off with a slight decrease in precision, but is taken into account in the statistical analysis. The subsequent chapters contain the description of the three levels of simulation, and the software packages used.

4.3.1   Hard parton scattering

This section relies on the theoretical framework presented in Chapter 1. At hadron colliders, as opposed to lepton colliders, the energy available in the collisions is distributed over the three valence quarks of the hadrons, the sea quarks and the gluons. In order to properly model the possible interactions, parton distribution functions (PDFs) are used, that dictates the probability of finding a parton within a proton, with a certain energy fraction x of the proton. Further, there are several ways of calculating the matrix element (ME). The square of the MEs are the transition probabilities of the gluons or (anti-)quarks in to the physical process of interest. The different packages utilizes different order of the strong coupling constant α_s, and are choice of package depends on the requested precision. These packages are known as Monte Carlo (MC) generators, which utilizes the MC random sampling technique.

4.3.1.1 Parton distribution functions

Parton distribution functions explain the probability to find a parton in a proton with some given energy fraction. Due to the non-perturbative nature of partons, i.e. that partons can not be freely observed, the parton distribution functions can not be computed from first principles, however, they can determined by fitting observables to experimental data. One way of determining the PDFs is through the so-called NNPDF3.0, that uses Neural Networks (NN) to model datasets collected by the ATLAS, CMS, HERA-II and LHCb experiments[22]. The NNPDP3.0 tool is developed to be used for searches and measurements during the LHC Run 2, and is used for all generated samples used in this thesis. To demonstrate the behaviour of the parton distributions, the NNPDF3.1 NNLO are displayed in Figure 4-3. N.B that the NNPDF3.1 is a successor of NNPDF3.0, and not used for the MC in this thesis, but lends itself well for a demonstration. As can be seen in Figure 4-3, the red curve, corresponding to the gluon PDF scaled down by a factor of 10, is dominating at the low energies, for both resolution scales. The corresponding interpretation is that the LHC pp collisions are dominated by low energy gluon induced processes, with a sub-dominant contribution of valence quarks at higher energy fractions. This feature is explaining the necessity of the very high collision rate of 40MHz, as most of the interactions are between low-energy carrying gluons, which do not have enough energy to creat "interesting" physics events.

4.3.1.2 Hard scattering processes

The generation of a physics process relies on the calculation of the matrix elements, which states the transition probabilites of the partons involved in the interaction to particles. The MEs can be calculated in many different ways, and each calculation corresponds to a MC generator, with the most significant difference being the order of α_s used. Leading order (LO) corresponds to the first order in α_s. In this thesis, a generator that uses LO ME calculation is Madgraph [23], which includes contributions with higher
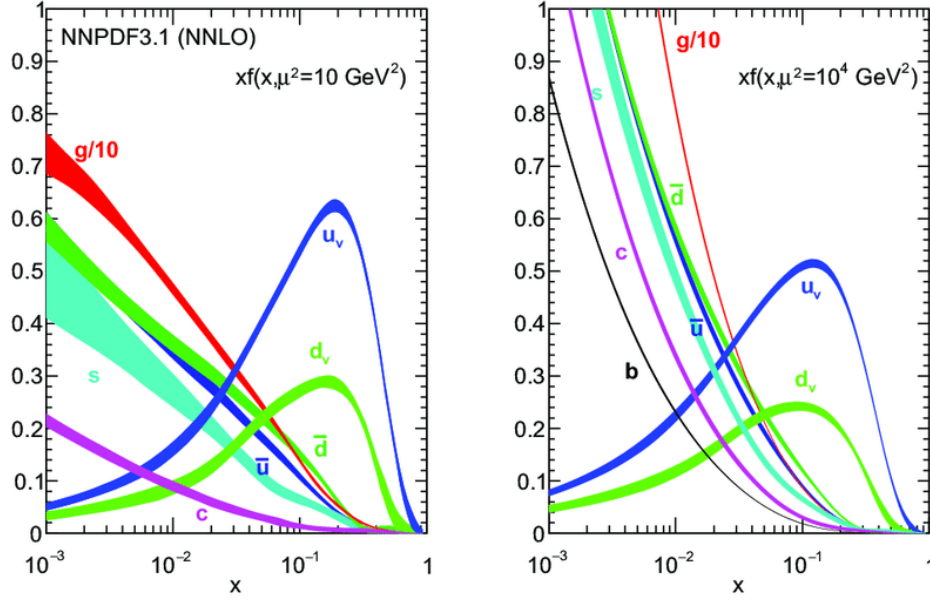
Figure 4-3. The NNPDF3.1 NNLO PDFs, evaluated at two resolution scales; $\mu^2 = 10\,\mathrm{GeV}^2$ (left) and $\mu^2 = 10^4\,\mathrm{GeV}^2$ (right)[9].

jet multiplicities via MLM [24] multileg matching and is used to simulate e.g. single γ events. Higher order calculations, such as next-to LO (NLO) and next-to-next-to LO (NNLO) include higher orders in α_s and are therefore more accurately describing the physical process. Madgraph can also be combined with a NLO generator, MC@NLO, into MADGRAPH5_aMC@NLO, and is used when generating e.g. the WZ $\to 3l1\nu$ process. A full NLO generator is POWHEG [25, 26], that is used to simulate e.g. the $t\bar{t} \to 2l2\nu$ process. The complete list of samples used in this thesis will be presented in the analysis section.

4.3.1.3 Multiparton interactions and parton showers

Multiparton interactions (MPI) are a result of the hadron collisions busy nature, where partons other than those involved in the hard scatter event interact. Initial state radiation (ISR) and final state radiation (FSR) are, as the names suggest, an emission of a γ, either before or after the main interaction in the event. Apart from both types of radiation are adding a photon to the event, ISR and FSR result in slight differences.

To illustrate this, we assume a Drell-Yan production, $Z \rightarrow e^+e^-$, with ISR and FSR. In the case of the ISR, the photon is radiated by the partons before the production of the Z boson, with a result of carrying off some of the available energy in the collision. Conversely, in FSR, the photon is radiated off of the electron or positron decayed from the Z boson, carrying off some of the energy of that lepton. So if one would reconstruct the invariant mass of the leptons, it would not completely form the Z boson mass of 91 GeV. Instead, one would have to construct the invariant mass of the leptons and the photon in order to recover the Z boson mass. The simulation of both the MPI and ISR anf FSR is done with PYTHIA [27].

4.3.1.4   Hadronization

Colored quarks and gluons produced in the hard scatter process can not exist freely due to color confinement. The mechanism of transforming the colored partons to colorless particles is known as hadronization. The hadronization step is challenging, as is all QCD calculations, and a phenomenological approach is taken in generators such as PYTHIA 8.2 [27]. The Lund string model [28] is the basis of the phenomenological approach, where the QCD field lines can be interpreted as being compressed into a tube like structure, a string. The fragmentation in this model can be viewed as starting in a middle and spreading outwards by repeated breaks of these strings, forming new quark anti-quark pairs.

4.3.1.5   Decay

Finally, PYTHIA is also taking care of the production of resonances and the decay of these unstable particles into stable particles or partons. All resonances are decayed sequentially as part of the hard process, and so the total cross-section as calculated by PYTHIA is dependent upon the available decay channels of the resonance, with the effect that not including a decay channel will decrease the cross-section accordingly. Conversely,

particle decays are performed after hadronisation, and changing the decay channels of a particle will not affect the total cross-section.

4.3.2   Detector simulation

Now that the simulation of the physics processes in hadron collisions have been described, one final step remains in order to be able to compare the data to the simulation, namely how the produced particles interact with the matter in the detector. GEANT4 is a simulation toolkit used to describe the interaction of particles with matter and can simulate everything from tracking of particles bending in a magnetic field to the response of detector components. In GEANT4, a detailed model of CMS is implemented that is taking care of the ionization, multiple scattering, and nuclear interactions and outputs a data format that is similar to that of the collision data. It is this step of the MC simulation that is most computationally heavy, and for this reason, a faster, simpler package is used to simulate the many SUSY scenarios, with the trade off of less precision. Worth noting is that if there would be any hints of SUSY in a particlular scenario, a full simulation would be performed.

4.4   Datasets

Once the collision data and the simulated data has been introduced, the datasets analyzed in this thesis is presented, along with a short introduction of the data formats and weighting of the MC samples.

4.4.1   Dataformats

The reconstruction of the collision data is processed at the Tier-0 and is stored in RECO format, a format containing much of the detector information and is thus associated with very large event sizes of 1.2MB. In order to make the storage and

analysis as easy as possible, a set of data tiers are processed that keeps the event sizes to a minimum. The first step in the data tier format processing is the Analysis Object Data (AOD) which has discarded much of the RAW detector information not needed for analysis, and decreased the event size greatly to around 300kB. The last step is the MINIAOD format, that was introduced for the Run II of the LHC. The MINIAOD contains all high level physics objects, all high level corrections for jets and $p_\mathrm{T}^\mathrm{miss}$, all particles returned by Particle Flow algorithm, all MC truth information for simulation, and all trigger information. By only saving the above information needed for mainstream analyses, the event sizes kept at 30–50kB. In order to cope with the huge amounts of data collected during Run II, and the various year specific set of MC samples, a further reduction in the event size is needed, and to this end the so called nanoAOD is developed, that manages to keep the event size at 1kB. The MINIAOD data and MC samples are stored on the Tier-0 and Tier-1, and analysis specific frameworks are used to create subsets of samples containing the particlular datasets needed. As the analyses presented in this thesis are based on different dileptonic datasets, the analysis specific framework used is aimed at picking data collected with dileptonic triggers and the various SM processes containing leptons, and store this in a format called Trees. These trees are used on analysis level for plotting, counting, fitting and statistical analysis.

4.4.2 Event weights

In order to be able to compare the simulted data to the real collision data, each MC sample need to be reweighted according to a so called event weight. The weight takes the form of

$$w = \frac{\sigma \cdot \mathcal{L}}{N} \tag{4–1}$$

where σ is the cross section of that process, $\mathcal{L}$ is the integrated luminosity of the collision data the simulation is compared to, and N is the number of generated events. This weight is then multiplied to each event. From the above equation, one can see that for a process

with large cross section, the more simulated events are needed to keep the weight to a desirable small value. The motivation to keep the weight as small as possible, is the statistical error on the MC sample, which goes as the square root of the sum of squared event weights. So the larger the N, the smaller the w and thus the smaller the statistical error.

4.4.3 Pileup

As described in Chapter 3, the LHC delivers more than one pp interaction per bunch crossing in order to increase chances for successful interactions. The motivation for this originates from the PDFs discussed previously in this chapter, and the fact that most pp collisions at the LHC produced not so interesting low mass QCD interactions with little physical interest. The number of interactions per bunch crossing is described by a Poissonian distribution, and has a mean of 27 interactions during the 2016 datataking, and is depicted in Figure 4-4[4]. This multiple pp interactions per bunch crossing,
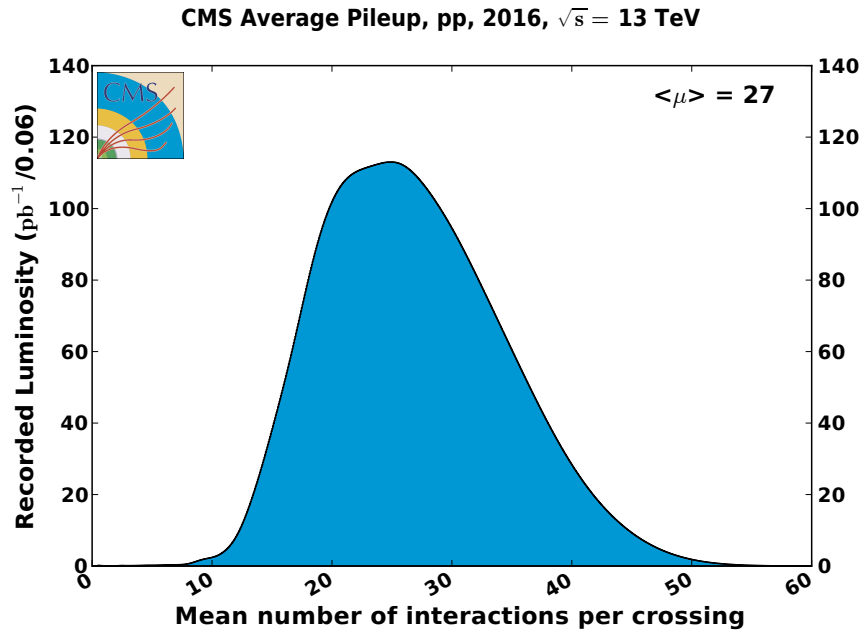
Figure 4-4. Mean number of interactions per bunch crossing for the 2016 pp run at 13 TeV. The cross section is taken to be 80 mb.

known as pileup, will be greatly discussed in the chapter on the performance of the p_T^{miss} algorithms. The reason for this is the challenges that pileup poses on the experiment in efficiently distinguishing the primary vertex from pileup vertices, and results in inefficiencies in the energy assignment of various physics objects. Specifically, the pileup has a significant effect on the reconstruction on the p_T^{miss}, which will be greatly discussed in the chapter on the performance of the p_T^{miss} algorithms. As the multiple pp collisions per bunch crossing depends on the beam conditions, and can vary over time between LHC fills, the pileup distribution varies significantly over time. For this reason, it is difficult to properly model the pileup in simulation, which results in the necessity of a correction of the modelled number of vertices in simulation to match the distribution in data. This is known as "pileup reweighting" and is used for all simulated samples. The two searches presented in this thesis use a pileup reweighting technique that reweights the total number of true interaction in the event, whereas the p_T^{miss} performance study reweights using the number of reconstructed vertices. The slight mismodelling of the pileup in the simulated signal samples is taken into account as a systematic uncertainty in the signal interpretation, which will be further discussed later.

CHAPTER 5

LEPTONIC SUSY SEARCHES

After having introduced the CMS detector, its abilities of reconstructing a wide variety of SM particles and the data processing, it is time to present what all this can be used for. The theoretical framework presented in Chapter 1 lays the foundation for the search for strong and electroweak production of SUSY particles that this thesis is presenting. Two opposite sign same flavor leptons provide a powerful search tool for stron and electroweak SUSY. This particular final state can appear in many SUSY scenarios, either through the production of an on-shell Z-boson or through the direct or intermediate production of sleptons. The strength of this final state does not only lie in the ability to target many SUSY scenarios, the SM backgrounds are relatively small and very well understood, which makes them ideal for various data-driven background estimation techniques. To put this into perspective, there are inclusive searches for strong SUSY prodcution in all hadronic final states that are completely swamped by QCD and W+jets processes. On the other side of the spectrum, there are searches trageting electroweak SUSY production in multilepton final states that instead are low in SM backgrounds, but the backgrounds from charge misidentification of leptons and the jets faking leptons poses a major challenge. The searches in this thesis is the middle ground between these two extreme final state scenarios with the strength of targeting a whole range of SUSY production modes. This thesis is based on two papers. The first paper [29] is a general search for SUSY in strong and electroweak production modes. The second paper [30]

presents a search targeting one particular SUSY production mode, namely direct selectron and smuon production. As all work performed in big collaborations such as the CMS experiment, no paper can be published independently. Instead, the work of many people at different institutes is required, and all authors of the CMS experiment contribute to the collection of the data, validation fo reconstruction algorithms, derivation of corrections, developing of analysis methods, to name a few. With this in mind, it is appropriate to mention that the first paper presented in this thesis relies on a collaboration of researchers at ETH Zürich, RWTH Aachen, University of California at San Diego and University of Oviedo. I will highlight my contributions to the first paper, and cite the work of my collaborators when necessary. Since my main contribution to the first paper was in the search for chargino-neutralino production ($\widetilde{\chi}_1^{\pm}$-$\widetilde{\chi}_2^0$) and higgsino production, I will dedicate a large fraction of this thesis to a description of this search, and refer to it in the following as the "Electroweak search".

5.1 SUSY searches with opposite sign same flavor leptons

The strong SUSY search, targeting gluino ($\widetilde{g}$) or sbottom ($\widetilde{b}$) production will also be presented in this thesis, refered to in the following as "Strong search". The second paper that this thesis is based on, is a search for directly produced selectrons ($\widetilde{e}_L$, $\widetilde{e}_R$) and smuons ($\widetilde{\mu}_L$, $\widetilde{\mu}_R$), refered to in the following as "Slepton search". The slepton search has many similarities to the electroweak search, but has some main differences that needs other SM background prediction techniques. This chapter will contain a brief overview of the different SUSY production modes in the first section, followed by a description of the SM background processes common for the three searches. The order of the SUSY scenarios presented below is dictated by the cross section of the processes. The strong SUSY searches is the oldest search presented in this thesis, as it has a higher production

cross section associated to it that made it relevant already in the Run 1 of the LHC. The
Electroweak search has a lower cross section associated to it, and many of the scenarios
probed are for the first time presented in this paper. Lastly, the direct slepton search has
the lowest cross section associated to it, making it the latest production mode covered by
this final state.

5.1.1 Strong SUSY search

The strong SUSY search using opposite sign same flavor leptons can target two
production modes. The first one is Gauge mediated SUSY breaking (GMSB) [31–33],
a model that assume strong production of a pair of gluinos ($\widetilde{g}$) that each decays into
a pair of quarks (u, d, s, c, or b) and the lightest neutralino, $\widetilde{\chi}^0_1$, shown on the left of
Figure 5-1. The leptons are a result of the decay of the on-shell Z-boson at the end of the
decay chain, and the large $p_\mathrm{T}^\mathrm{miss}$ is due to the gravitino ($\widetilde{G}$). The second strong SUSY
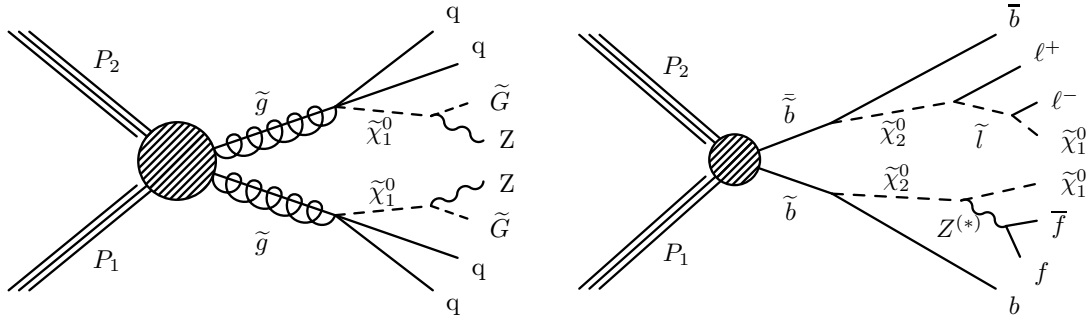
Figure 5-1. Diagrams for strong SUSY production. with decays containing at least one
dilepton pair stemming from a Z decay are shown. The gluino GMSB model
targeted by the strong on-Z search is shown on the left, that contains a one
dilepton pair stemming from an on-shell Z boson decay. On the right is a
diagram showing a model in which bottom squarks are pair produced with
subsequent decays that contain at least one dilepton pair. This model features
a characteristic edge shape in the $m_{\ell\ell}$ spectrum given approximately by the
mass difference of the $\widetilde{\chi}^0_2$ and $\widetilde{\chi}^0_1$.

production model that can be targeted using opposite sign same flavor leptons is the
direct pair production of $\widetilde{b}$. In this model, the $\widetilde{b}$ quarks decay to a bottom quark and $\widetilde{\chi}^0_2$.

Two assumptions are made for the decay of the $\widetilde{\chi}_2^0$. In the one case (upper half of the Feynman diagram in the right of Figure 5-1), the $\widetilde{\chi}_2^0$ decay to a $\widetilde{\ell}$ and a lepton of the same flavor, where the $\widetilde{\ell}$ decays to a lepton of the same flavor but opposite charge, and the $\widetilde{\chi}_1^0$ LSP. The sequential decay of the $\widetilde{\ell}$ (that in this work considered to be either a $\widetilde{e}$ or $\widetilde{\mu}$) results in the particular feature that the invariant mass of the two leptons has an edge shape with an endpoint dictated by the mass difference of the $\widetilde{\chi}_2^0$ and the $\widetilde{\chi}_1^0$. Further, the $\widetilde{b}$ and $\widetilde{\chi}_2^0$ masses are free parameters, the $\widetilde{\chi}_1^0$ is set to $100\,\text{GeV}$ and the sleptons are assumed to be degenrate with the mass being average of the $\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^0$ masses. The other decay mode of the $\widetilde{\chi}_2^0$ considered it visualized in the lower part of the Feynman diagram on the right of Figure 5-1. In this model, the $\widetilde{\chi}_2^0$ decays to a Z boson and the $\widetilde{\chi}_1^0$ LSP. The Z boson can be on-shell or off-shell depending on the mass difference between the neutralinos, and can decay to any fermion pair allowed by the SM but in this work only the decay to an electron or muon pair is considered. The branching fractions of the two $\widetilde{\chi}_2^0$ decay modes is considered to be 50% each, and the model is interpreted in context of the exclusion of $\widetilde{b}$ and $\widetilde{\chi}_2^0$.

5.1.2   Electroweak SUSY search

The SUSY models considered in the first paper assume electroweak production (EW) and Gauge mediated SUSY breaking (GMSB), and are presented in Figure 5-2. The $\widetilde{\chi}_1^\pm$-$\widetilde{\chi}_2^0$ production is depicted in the upper part of Figure 5-2. In this model, the $\widetilde{\chi}_1^\pm$ is set to decay to a W boson and a $\widetilde{\chi}_1^0$, which is the LSP, while the next-to-lightest neutralino, $\widetilde{\chi}_2^0$, decays to a Z boson and $\widetilde{\chi}_1^0$. The production cross sections for this model are computed in a limit of mass-degenerate wino $\widetilde{\chi}_1^\pm$ and $\widetilde{\chi}_2^0$, and light bino $\widetilde{\chi}_1^0$. All the other SUSY particles are assumed to be heavy and decoupled. The remaining two models considered in the first paper assume the production of $\widetilde{\chi}_1^0$-$\widetilde{\chi}_1^0$ pairs in GMSB. For bino- or wino-like neutralinos, the neutralino pair production cross section is very small, and thus a specific GMSB model is considered, with mass-degenerate higgsinos $\widetilde{\chi}_1^\pm$, $\widetilde{\chi}_2^0$, and $\widetilde{\chi}_1^0$ as

the next-to-lightest SUSY particles and a massless gravitino as the LSP [31–33]. In the production of any two of these, $\widetilde{\chi}_1^\pm$ or $\widetilde{\chi}_2^0$ decays immediately to $\widetilde{\chi}_1^0$ and low-momentum particles that do not impact the analysis, effectively yielding pair production of $\widetilde{\chi}_1^0 \widetilde{\chi}_1^0$. Intermediate production of either $\widetilde{\chi}_1^\pm$ or $\widetilde{\chi}_2^0$ is therefore not explicitly shown in the lower two diagrams of Fig. 5-2 representing these models. In the first model (lower left of Figure 5-2), the only allowed decay of the lightest neutralino is to a Z boson and a massless gravitino. In the other model (lower two diagrams of Figure 5-2), the lightest neutralino is allowed to decay to a gravitino and either a Z boson or an SM-like Higgs boson, with a 50% branching fraction to each decay channel. The cross sections for higgsino pair production are computed in a limit of mass-degenerate higgsino states $\widetilde{\chi}_2^0$, $\widetilde{\chi}_1^\pm$, and $\widetilde{\chi}_1^0$. Again, all the other SUSY particles are assumed to be heavy and decoupled. Following the convention of real mixing matrices and signed neutralino masses [34], the sign of the mass of $\widetilde{\chi}_1^0$ ($\widetilde{\chi}_2^0$) are set to $+1$ (-1). The lightest two neutralino states are defined as symmetric (anti-symmetric) combinations of higgsino states by setting the product of the elements N_{i3} and N_{i4} of the neutralino mixing matrix N to $+0.5$ (-0.5) for $i = 1$ (2). The elements U_{12} and V_{12} of the chargino mixing matrices U and V are set to 1. Common to these models is that they produce an on-shell Z boson giving the OCSF leptons, in association with a SM boson, and two LSPs resuting in large $p_\mathrm{T}^\mathrm{miss}$. The SM boson is either a W, Z or a H, and their decay mode to jets (or b-tagged jets in the case of the H) is targeted in the search.

5.1.3 Slepton search

SUSY models predict charged sleptons ($\widetilde{e}_L$, $\widetilde{\mu}_L$, $\widetilde{\tau}_L$, $\widetilde{e}_R$, $\widetilde{\mu}_R$, $\widetilde{\tau}_R$), the superpartners of the charged left-handed and right-handed SM leptons. At sufficiently heavy slepton masses, the sleptons undergo a two-body decay into one of the heavier neutralinos or a chargino, while direct decays to a neutralino LSP are favored for light slepton masses. The second paper that this thesis is based on is on a search for directly produced selectrons
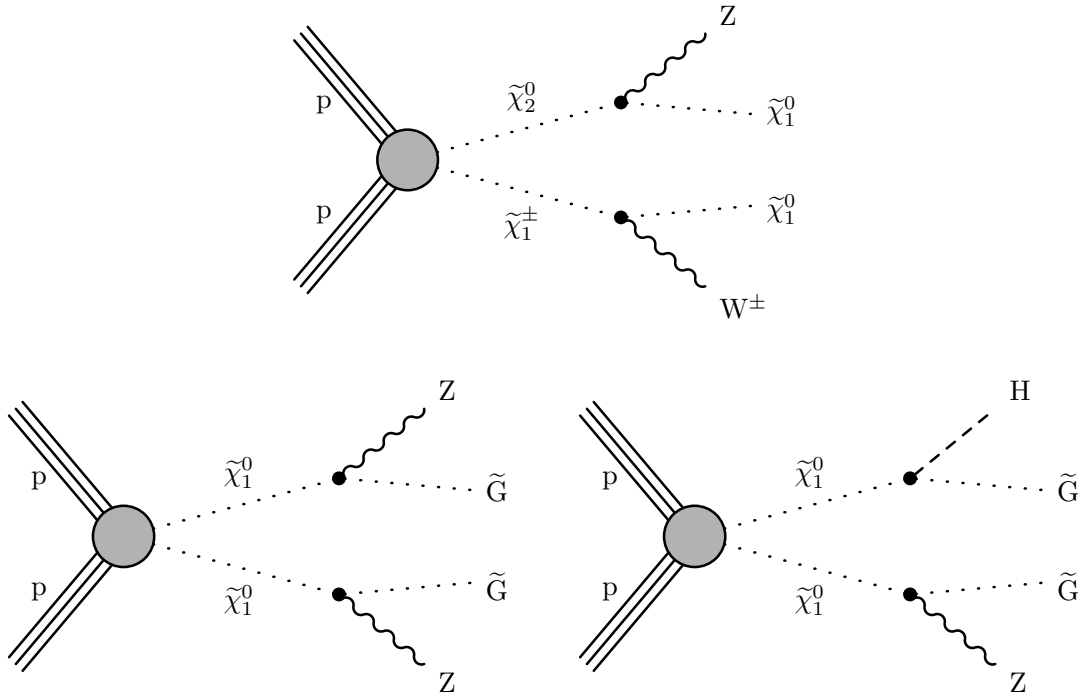
Figure 5-2. (Upper) Diagram corresponding to the chargino-neutralino production with the $\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^0$ decaying into vector bosons and the LSP. (Lower) Diagrams corresponding to the neutralino-neutralino model of where the neutralinos are allowed to decay to a gravitino and a Z boson (left) and where the neutralinos are allowed to decay to a gravitino and a Z boson or a Higgs boson, with a 50% branching fraction to each decay channel (right).

and smuons, under the assumption of direct decays $\widetilde{\ell} \to \ell\widetilde{\chi}_1^0$ with 100% branching ratio, as sketched in Fig. 5-3.
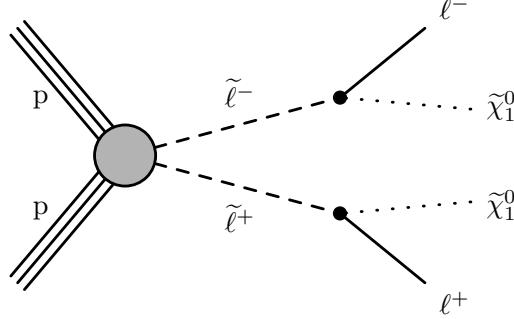
Figure 5-3. Diagram corresponding to the slepton model with two selectrons (smuons) directly produced and decay into electrons (muons) and a LSP

5.2   Background processes

The final states resulting from directly produced selectrons (smuons) are a pair of electrons (muons) and large p_T^{miss} from the LSPs. The main differences in the final state compared to the $\widetilde{\chi}_1^\pm$-$\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^0$-$\widetilde{\chi}_1^0$ production, are that absence of jets and the fact that the lepton pair is not compatible to the Z boson mass. These two distinctions make for a different search strategy, where no jets are required, and the contribution from Drell–Yan can be greatly suppressed through a veto. Below are the background processes listed for the search for $\widetilde{\chi}_1^\pm$-$\widetilde{\chi}_2^0$ and $\widetilde{\chi}_1^0$-$\widetilde{\chi}_1^0$ production and for direct slepton production, and a short description on how these processes are estimated.

5.2.1   Top related processes

Leptonically decaying top anti-top pair production provides a major background in this search, as the leptons are of opposite charge. Additionally, single top production, can also result in this signature, if one of the jets is misidentified as a lepton. Both

processes are depicted in Figure 5-4 and involve a leptonically decaying W boson, and these backgrounds are predicted using the flavor symmetry of the W decay, described in the subsequent chapter. These backgrounds, referred to as Flavor Symmetric (FS) in this thesis, can be heavily suppressed by a cut at the end point of the M_{T2} distribution.
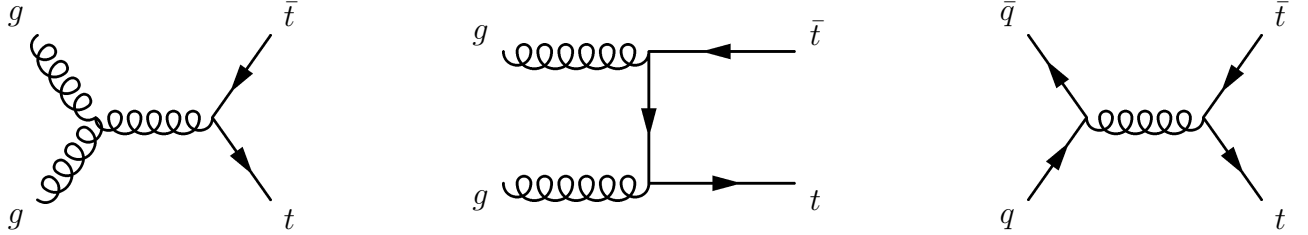
Figure 5-4. The leading order diagrams of t$\bar{\text{t}}$ production, through gluon fusion (left and middle) and quark-antiquark annihilation (right).
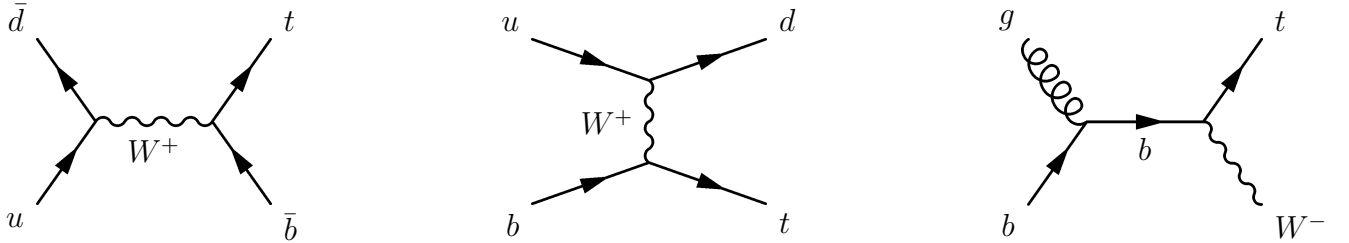
Figure 5-5. Single top quark production through s-channel (left), t-channel (middle) and in association with a W boson (right).

5.2.2   Drell-Yan

The large cross section Drell-Yan process (DY), where a virtual photon or a Z boson decays to two leptons, is a major background in the search for Charginos and Neutralinos as a lepton pair compatible with a Z boson is required in the final state, and is shown in Fig. 5-6. This process contain no production of neutrinos, with the result that the p_T^miss is solely due to jet resolution and detector effects. As will be thoroughly described in the chapter on the p_T^miss performance, the DY is very similar to the single photon production in the sense that it does not contain any real p_T^miss from neutrinos. Therefore, the p_T^miss contribution from DY can be estimated through the p_T^miss from single photon production. This data-driven technique is referred to as "p_T^miss template method" and profits from the high statistical power of the single γ process. In the direct slepton production search, a

Z boson veto is applied that heavily suppresses the DY, and the very minor contribution from DY is taken from simulation.
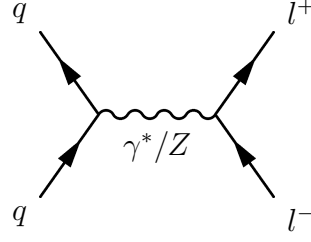
Figure 5-6. Leading order DY production.

5.2.3 Diboson production

Diagrams for diboson production are shown in Fig. 5-9, where V_1 and V_2 are either W and W, Z and Z, or W and Z. The WW process, where both W bosons decay leptonically, is flavor symmetric and is estimated using the flavor symmetric background prediction method described in the subsequent chapters. The WW process and ZZ process, when one Z boson decays to charged leptons and one Z boson decay to neutrinos, are dominant backgrounds in the direct slepton search as they fulfill the criteria of no hadronic activity and large p_T^{miss} from neutrinos. If both bosons decay leptonically, the WZ process can result in opposite sign same flavor pairs and enter both searches as a background if one of the leptons is out of η or p_T acceptance. The ZZ and WZ are referred to as "$Z + \nu$ backgrounds" in this thesis and are estimated through simulation with translation factors derived from dedicated control regions. As the strong, electroweak and slepton searches have different kinematic features, the translation factors for these processes are derived separatley for the three searches.
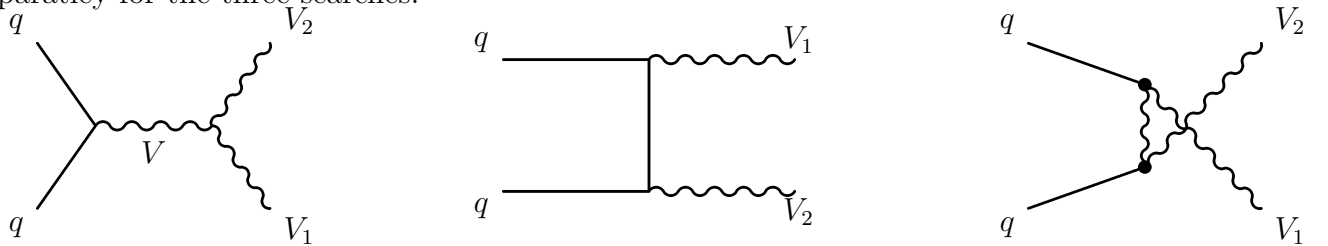
Figure 5-7. Leading order diboson production through s-channel (left), t-channel (middle) or u-channel (right).

5.2.4   Rare processes
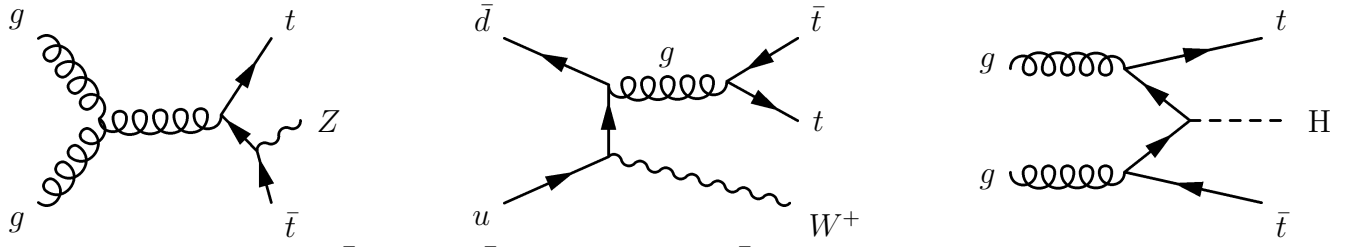
Figure 5-8. Leading order t$\bar{t}$Z (left), t$\bar{t}$W (middle) and t$\bar{t}$H (right) production.
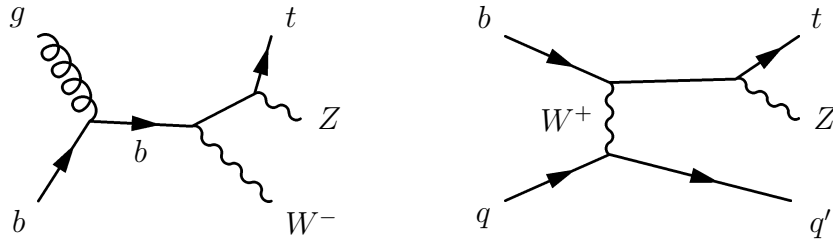
Figure 5-9. Leading order tWZ (left) and tZq (right) production.

CHAPTER 6

PHYSICS OBJECTS

Common to the strong, electroweak and slepton searches is the final state containing two leptons and $p_\mathrm{T}^\mathrm{miss}$ originating from the $\widetilde{\chi}_1^0$ or $\widetilde{G}$ LSP. The difference lies in the hadronic activity and whether the lepton pair is compatible with a Z boson or not. This chapter starts by presenting the Particle Flow algorithm (PF) that combines subdetector information to form the physics objects that are used in these searches. The rest of the chapter presents the reconstruction of the physics objects used in this search and the quality criteria imposed on them to be considered a signal event or suitable to predict various SM background processes.

6.1   The Particle Flow Algorithm

General purpose high energy particle physics detectors are built on the principle of layers of sub-detectors around the beam axis. The calorimeter layers and muon systems are designed for particles to get completely absorbed or produce hits in them. Intuitively, one would think that each layer would lend it self useful for reconstructing a certain type of particle: electrons and photons could be reconstructed from electromagnetic showers in the ECAL, hadrons reconstructed from hadronic showers in the HCAL, jets from combined calorimeter signals, muons from hits in the muon system. However, the idea behind the PF algorithm is to instead optimally combine the information from the

different sub-detector layers. The following is a simplified description of the PF algorithm

that closely follow that described in [35]. The design of the CMS detector as described

in Section 3.2 has proven to be well-suited for particle-flow (PF) reconstruction of the

physics objects. Reasons behind this statement include the excellent muon detectors

of CMS that provide efficient and pure identification of muons. The combination of a

fine-grained tracker and a strong magnetic field effectively measures charged particle

tracks that make up 65% of the jet energy. Additionally, the excellent resolution and

the segmentation of the ECAL provides separation of energy deposits from particles in

hadronic jets and accounts for 25% of the jet energy. The segmentation of the HCAL

allows to differentiate charged from neutral hadrons, and subsequently the measurement of

the remaining 10% of the jet energy. The algorithm can be summarized as follows:

- Tracks from the pixel and strip tracker are extrapolated from the last measured hit in the tracker to the calorimeters. The track is linked to a cluster in the calorimeters according to some boundary conditions, and if several clusters are linked to a track, the cluster with the smallest distance to the track is kept. Tracks are also extrapolated to the muon detector.

- In the second step of the algorithm muons are identified by a set of requirements on global/tracker muons, as described in Section 6.4. The tracks and clusters associated to these muons are not considered in the rest of the algorithm.

- After muons have been identified, the electron and isolated photon identification follow. Electron candidates are seeded from a GSF track if there are no additional tracks linked to that ECAL cluster. Photon candidates are seeded from ECAL superclusters with E_T greater than $10\,\mathrm{GeV}$ and no link to a GSF track. The tracks and clusters are then removed from the remainder of the algorithm.

- At this point, muons, electrons and isolated photons have been identified and their corresponding tracks and clusters removed. Charged hadrons ($\pi^{\pm}$, $K^{\pm}$, protons etc.), neutral hadrons (K^0_L, neutrons etc.) and non-isolated photons (from π^0 decays) are identified. Neutral hadrons and photons are identified by ECAL and HCAL clusters not linked to any track. More precisely, within the tracker volume, an ECAL cluster give rise to a photon whereas a HCAL cluster give rise to a neutral hadron. Outside of the tracker volume it is not possible to differentiate charged from neutral hadrons, and therefore ECAL and HCAL clusters that are linked are interpreted as charged or neutral hadrons, whereas only an ECAL cluster is identified as a photon.

- The remaining HCAL clusters can be linked to the remaining tracks (that are not already linked to another HCAL cluster), that subsequently can be linked to the remaining ECAL clusters. The total energy of the ECAL and HCAL clusters is then used to determine the *calibrated calorimetric energy* under a single charged hadron hypothesis. The sum of the track momenta is also computed and compared to the calibrated calorimetric energy. If the calibrated calorimetric energy is larger than the sum of the track momenta, this can be interpreted as a photon or a neutral hadron. If the calibrated calorimetric energy is compatible with the sum of the track momenta, this can be interpreted as a charged hadron. If the calibrated calorimetric energy is smaller than the sum of the track momenta, this can be interpreted as a muon.

To summarize; electrons, muons, photons, charged hadrons and neutral hadrons have been identified by the PF algorithm. These *PF candidates* are subsequently used to reconstruct jets, p_T^{miss} and τ lepton candidates.

6.2   Leptons

Two leptons are expected in all the studied signal processes, either from the decay of an on-shell Z boson or through the pair production of sleptons that decay to leptons and neutralinos. There are several SM background processes that can result in opposite sign same flavor leptons, and common to these are that they include a decay of a Z boson, virtual photon or a W boson. Although these bosons decay democratically to each lepton generation, as shown in Figure **??**, only the first and second lepton generations are considered in this thesis. In this work, dielectron (ee) and dimuon ($\mu\mu$) events are categorized as same flavor (SF) and events containing an electron and a muon ($e\mu$) are categorized as different flavor (DF). The reason to exclude the third lepton generation, the $\tau^{\pm}$ leptons and their corresponding ν_τ's, is due to the challenging final state. The τ lepton, being the most massive of the leptons flavors, decay hadronically 65% of the time. As an effect, including this decay mode in the various signal scenarios would result in the introduction of overwhelming backgrounds from multijet QCD processes, that in turn would decrease the sensitivity of the searches. However, the search for direct stau

production in semi-leptonic and hadronic final states has been performed by CMS[36, 37], where the challenging background due to τ leptons is taken into account.

6.3 Electrons

Electrons are reconstructed by associating ECAL clusters with a track reconstructed in the silicon detector. The following contains a description of the clustering of the ECAL deposits, the electron track and the combination of the two, and is roughly based on [38]. Finally, the identification and isolation requirements imposed on the electrons to be categorized as signal events or used for background estimation techniques are presented.

6.3.1 Electron reconstruction

Bremsstrahlung causes electrons to lose energy when interacting with detector material. Depending on the amount of tracker material before the ECAL, the electrons can lose more or less energy due to radiation of photons. For this reason, the energy of the measured electrons will be underestimated if one does not take into account the radiated photons.

6.3.1.1 Clustering

Clusters of ECAL crystals are constructed for this reason, with the aim to collect the energy deposits of the crystals surrounding a seed crystal. Two clustering algorithms are deployed that group crystals together into supercrystals (SC) following a set of threshold criteria on the E_T of the seeded crystal, and taking into account the differing crystal layouts of the ECAL barrel and endcap. Additionally, PF clusters are reconstructed by aggregating crystals surrounding a seed, following thresholds set on the electronic noise levels.

6.3.1.2 Seeding

The electron seed, the first step in the electron track reconstruction, can either be an ECAL-based seed or a tracker-based seed. As the naming suggests, the former seed is based on information from the SC that is extrapolated back to the vertex. The latter is a part of the PF reconstruction algorithm. The tracker-based procedure has advantages over the ECAL-based procedure for low p_{T} electrons that bend significantly in the magnetic field, the radiated photons are spread out and can not all be contained in a SC, making the extrapolation of the track starting from the SC in the ECAl-based approach sub-optimal, as not all radiated energy is contained in it. Additionally, for electrons in jets, the SC position and energy is affected by the other particle contributions, making the ECAL-based approach inefficient. To determine the tracker-based seed, the Kalman Filter (KF) track reconstruction can be used, but is sub-optimal when there is radiative losses of the electrons in the tracker material. In the cases where a lot of bremsthralung took place, the KF algorithm will not be able to follow the change in the curvature of the electron trajectory as a result of the energy lost. As a result, the hits can not be collected, or they are of bad quality. To recover the trajectory of these bremstrahlung electron, the KF tracks are refitted using the Gaussian sum filter (GSF) algorithm. The GSF algorithm takes into account that the bremsstrahlung electron energy loss is non-Gaussian, and model the bremstrahlung energy loss by weighted sums of Gaussians instead of just one Gaussian used in the KF algorithm. The χ^2 of the KF and GSF fitting, together with energy matching of the ECAL and tracker are used in an MVA to select the tracker seed as the electron seed.

6.3.1.3 Tracking

The ECAL-based seed and tracker-based seed are merged into a unique collection and are submitted to the full electron tracking with twelve GSF components. The track is

60

built using the combinatorial KF method which starts at the electron seed and proceeds iteratively in the next layer of the tracker, and the hits are collected.

6.3.1.4 Track and cluster association

Electron candidates are constructed by associating a GSF track to an ECAL cluster. ECAL-seeded electrons are now by construction associated to the ECAL cluster used to determine the ECAL-based seed. The tracker-seeded electrons are associated to the aforementioned PF clusters.

6.3.2 Electron identification and isolation

The rationale beind the electron identification is to efficiently differentiate prompt electrons from electrons from photon conversions, jets misidentified as electrons and electrons from b-quark decays. As there are many discriminating variables available, the electron identification can be performed by either a sequential set of selections on these variables, or combine them in an MVA analysis for improved discrimination. The electron identification used in these searches use the MVA identification technique, and below follow a description of the variables used in the analysis.

- $\sigma_\eta \sigma_\eta$: The shower shape variable is a measure of the spread of the electron shower along the η direction. [1] This variable exhibits a sharp peak for prompt electrons, whereas for misidentified electrons it is more spread, making it an ideal discrimination variable.

- $|\Delta\eta| = |\eta_{SC} - \eta_{in}^{extrap}|$: denotes the separation between the SC energy-weighted η position and the track η extrapolated from the innermost track position and direction to the point of closest approach (PCA) to the SC. The $|\Delta\eta|$ increases with the amount of bremstrahlung.

---

[1] defined as $(\sigma_\eta \sigma_\eta)^2 = \frac{[\Sigma(\eta_i - \bar{\eta})^2 w_i]}{\Sigma w_i}$ where the sums run over the 5x5 matrix of crystals around the highest E_T crystal of SC, and w_i denotes a weight that is logarithmically dependent on the contained energy.

- $|\Delta\phi| = |\phi_{SC} - \phi_{in}^{extrap}|$: denotes the azimuthal separation between the SC energy-weighted ϕ position and the track ϕ extrapolated from the innermost track position and direction to the PCA to the SC.

- H/E: The ratio H/E [2] is used to estimate the energy leakage into the HCAL. A well-identified electron would be expected to have a low H/E owing to the high X_0 of the CMS detector, thereby containing the EM showering before it reaches the HCAL. Hadron fakes instead exhibit a larger H/E.

- $|1/E - 1/p|$: This quantity expresses an energy-momentum matching requirement using the SC energy, E, and the track momentum, p, at the PCA to the track vertex. The requirement helps to reject backgrounds from hadronic activity where the spread of the E is not localized resulting in a low E/p, but also backgrounds where a $\pi^0 \rightarrow \gamma\gamma$ decay occurs in the close vicinity of a charged hadron, resulting in a very high E/p ratio.

- d_{xy} and d_z: The transverse and longitudinal distance between the electron track and the primary interaction vertex.

- Missing hits: Missing hits are used to suppress electrons from photon conversions. As photon conversions take place close to the beampipe or the pixel system, missing hits can occur due to the large change in curvature of the electron trajectory. At most, one missing hit is allowed for an accepted trajectory candidate, and, to avoid including hits from converted bremsstrahlung photons in the reconstruction of primary electron tracks, an increased χ^2 penalty is applied to trajectory candidates with one missing hit.

- Conversion veto: In order to reject secondary electrons produced in the conversion of photons in the tracker material, a vertexing algorithm is used. The hits in the tracker from the converted photon are fit to a common vertex using the well-defined topological constraint that tracks from conversions have virtually the same tangent at the conversion vertex in both the (r, ϕ) and (r, z) planes. The convereted photon candidates are rejected according to the χ^2 probability of the fit.

As isolation variable, the mini-isolation is used which features a shrinking cone-size with increasing p_{T} of the lepton. Thus, the cone size in which the PF particles are summed to calculate the relative isolation is no longer constant, but a function of the p_{T}

---

[2] where E is the energy of the SC, and H the sum of the HCAL tower energies within $\Delta R = \sqrt{((\Delta\eta)^2 + (\Delta\phi)^2)} = 0.15$

of the lepton

$$R = \frac{10.}{min\left[max\left(p_\mathrm{T}, 50\right), 200\right]} \quad.$$

For p_T values below $50\,\mathrm{GeV}$, this leads to a constant cone size of 0.2. For p_T values

between $50\,\mathrm{GeV}$ and $200\,\mathrm{GeV}$, the cone size shrinks from 0.2 to 0.05 at which it remains

for higher p_T leptons. Corrections to the isolation are applied by subtracting the average

energy density ρ from the effective geometrical area of the lepton's isolation cone. The

variable cone size is taken into account for this correction.

Table 6-1 summarizes the identification and isolation criteria imposed on the electrons

in order for them to be considered in the analysis. The MVA trained discriminator is

optimized on electrons from prompt W-boson decays in $t\bar{t}$ versus leptons stemming from

so-called "fakes"[3] in $t\bar{t}$. The working point used for the MVA identification discriminator

corresponds to a "tight" value, developed at the end of 2016. The actual MVA cut value

depends on the lepton p_T and $|\eta|$. In each $|\eta|$ bin the lower value is used for electrons with

$p_\mathrm{T} > 25$ GeV while the cut decreases linearly from the upper to the lower value for p_T

between 15 and 25 GeV. Additionally, conversion rejection cuts are applied.

---

[3] "Fake" leptons are mostly semi-leptonic b-quark decays where $b \to cW \to c\ell\nu$

Table 6-1. Electron selection criteria.

cut	value
Identification	
MVA	Tight 2016 Working point
Conversion rejection	
number of lost hits	0
conversion veto	pass
Isolation	
mini Isolation	< 0.1
Impact parameter	
d_{xy}	$< 5\,\mathrm{mm}$
d_z	$< 10\,\mathrm{mm}$
SIP3D	< 8

6.4 Muons

The calorimeters of CMS are efficiently stopping electrons, photons, charged and neutral hadrons, resulting in muons (and of course neutrinos) being the only particles reaching the muon systems. The muon track reconstruction and identification described below is based on [39]. Inner tracks and tracks in the muon systems are used as input for the muon-track reconstruction, that can be categorized in the following three ways:

- Standalone-muon track: DT and CSC hits are clustered and the track segments formed are used as seeds for the fitting. DT, CSC and RPC hits are used in the final fitting to reconstruct a standalone-muon track.

- Tracker-muon track: An "inside-out" approach is used for the tracker-muon tracks. If an inner track has $p_\mathrm{T} \mathrel{¿} 0.5\,\mathrm{GeV}$ and $p \mathrel{¿} 2.5\,\mathrm{GeV}$, then it is extrapolated to the muon system. If the extrapolated track and at least one track in a muon segment have an absolute difference in their position in x-coordinates less than $3\,\mathrm{cm}$ or the ratio of this distance to the uncertainty is less than 4, then tracks are considered to be matched.

- Global-muon track: In contrast to the tracker-muon track, an "outside-in" approach is used for global muon tracks. A standalone-muon track is matched to an inner track, if the position and momentum of the two tracks are compatible. The two tracks are combined and fitted using KF technique to form a global muon track.

The global muon track procedure improves the momentum resolution for muons of p_T ¿ 200 GeV compared to a tracker-muon track only fit. This is because large p_T muons have a higher probability to reach more than one muon segments, which is required for the global muon reconstruction. Lower p_T muons benefit from the tracker-muon reconstruction as only one muon segment is required in the matching. However, most muons are reconstructed as global muons, tracker muons or as both. The standalone-muon tracks have worse momentum resolution and can more often pick up cosmic muons.

6.4.1 Muon identification and isolation

From an experimental point of view, muons are much easier to measure than electrons. As the radiated power due to Bremsstrahlung is much lower for muons than for electrons[4] there are less amibiguities in the muon reconstruction than in the electron reconstruction. Additionally, the unlikely process of photon conversion to two muons make a conversion veto redundant. Instead, the different variables used for the muon identification is listed below:

- Reconstruction type: A muon can be identified using the various muon-track reconstruction categorization, i.e. require standalone-muon track, tracker-muon track, or global-muon track. PF muon identification requirement can also used imposed.

- Fraction of valid tracker hits: Fraction of hits from inner tracker layers that the muon traverses.

- Segment compatibility: The muon segment compatibility is computed by propagating the tracker track to the muon system, and evaluating both the number of matched segments in all stations and the closeness of the matching in position and direction

---

[4] The radiated power due to Bremsstrahlung goes as m^{-4}, meaning that it is suppressed for more massive particles such as muons and protons.

- Kink-finding: A kink-finding algorithm splits the tracker track into two separate tracks at several places along the trajectory For each split the algorithm makes a comparison between the two separate tracks, with a large χ^2 indicating that the two tracks are incompatible with being a single track.

- Position match: the tracker and standalone tracks are matched according to their position.

- Normalized χ^2: Requirement on the global fit required to have goodness-of-fit per degree of freedom

- d_{xy} and d_z: Requirements on the transversal and longitudinal distance between the tracker track and the location of the primary vertex. This cut is applied to suppress events with cosmic muons, tracks from pile-up or muons originating from in-flight decays.

There are several identification categorizations developed in CMS for various analysis needs, with working points aimed to more or less efficienciently suppress fakes from punch through hadrons while keeping a high muon identification efficiency. In this work, the medium muon ID is used. The medium muon ID is optimized for prompt muons and for muons from heavy flavor decay, and the definition is summarized in Table 6-2. The isolation variable used for electrons, the mini-isolation, is also used for muons.

Table 6-2. Muon selection criteria.

cut	value
Loose Muon ID	
PF Muon ID	True
Is Global OR Tracker Muon	True
Medium Muon ID	
Loose Muon ID	True
Global Muon	True
Fraction of valid tracker hits	> 0.8
Normalized global-track χ^2	< 3
Tracker-Standalone position match χ^2	< 12
Kick-finder	< 20
Tight Segment compatibility	> 0.451
Isolation	
mini Isolation	< 0.2
Impact parameter	
d_{xy}	$< 5\,\mathrm{mm}$
d_z	$< 10\,\mathrm{mm}$
SIP3D	< 8

6.5 Jets

The PF candidates found through the PF algorithm described in Section 6.1 are used as input in a clustering algorithm for hadronic jets. A variety of clustering algorithms are on the market, with different optimizations for high or low p_T jets or different shapes, with different computation speeds. In the following the ingredients needed for well-measured jets are described, starting with pileup mitigation, followed by the sequential jet clustering algorithms, and finally the identification and corrections needed.

6.5.1 Charged hadron subtraction

The clustering algorithms are efficient in forming jets while keeping the contamination of effects from pileup and underlying events at a minimum. During the clustering, there is a first line of defense against pileup called Charged Hadron Subtraction (CHS). CHS

is a type of particle-by-particle pileup subtraction performed on the PF candidates, that assigns charged hadrons to the primary vertex or pileup vertices, using tracking information. Charged hadrons clearly associated to a pileup vertex are removed.

6.5.2   Jet clustering algorithms

The clustering algorithm most commonly used today is the so-called anti$-k_t$ algorithm[40], which is a member of a class of sequential recombination algorithms used at hadron colliders. The idea behind these sequential clustering algorithms is to utilize the distance d_{ij} between particles i and j, the distance d_{iB} between particle i and the beam B and the transverse momenta of the particles $k_{\text{t}i,j}$ when forming a jet. The distance parameters d_{ij} and d_{iB} are defined as:

$$d_{ij} = \min(k_{\text{t}i}^{2a}, k_{\text{t}j}^{2a}) \times \frac{(\Delta_{ij}^2)}{R^2}, \tag{6–1}$$

$$d_{iB} = k_{\text{T}i}^a, \tag{6–2}$$

where $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, η and ϕ the pseudorapidity and azimuth distance between the particles, and R the radius parameter of the jet cone. The parameter a determines the type of clustering algorithm. The algorithm follows these steps: compute all distances d_{ij} and d_{iB}, and find the smallest one. If smallest is a d_{ij}, combine (sum four momenta) the two particles i and j, update distances, proceed to find the next smallestdistance. If smallest is a d_{iB}, remove particle i and call it a jet. Repeat the following steps until all particles are clustered into jets. The predecessors to the anti$-k_t$ algorithms are recovered for $a = 1$ (k_t algorithm[41]) or $a = 0$ (Cambridge/Aachen (CA) algorithm[42]). Setting the $a = -1$ is the basis of the anti$-k_t$ algorithm. The choice of a negative a is motivated by the following reasoning. Considering an event with a few hard particles and many soft particles. The distance parameter d_{i1} between the hard particle 1 and the various soft particles is dominated by the p_{T} of the hard particle, as a result of the negative exponent that makes the d_{ij} small for a large p_{T} hard particle 1, and larger

for equally separated soft particles. As a result, the soft particles will more likely cluster with hard ones before they cluster with another soft particle. A further consequence is that soft particles do not modify the shape of the jet whereas hard ones do, meaning that the jet boundary is resilient to soft radiation. This is a desired feature, also known as infrared and collinear (IRC) safety. Namely that neither soft emissions or collinear splittings should not change the boundary of jets. Comparing the shape of jets clustered with CA, k_t and anti$-k_t$ algorithms respectively, the anti$-k_t$ algorithm results in more conical jets. In this work, jets are clustered with the anti$-k_t$ algorithm with a radius parameter of 0.4 (AK4).

6.5.3   Jet identification

Identification criteria are imposed on the PF jets in order to suppress noise contributions from the calorimeters. The selections are based on relative energy fractions carried by the PF candidates with respect to their total jet energy, and the number of PF candidates in a jet. When tracking information is available ($| < 2.4|$) additional cuts are applied on the charged candidates. The Loose Jet ID working point is used in this work and its definition is summarized in Table 6-3.

Table 6-3. Definition of Loose Jet ID working point.

Variable	$|\eta| < 2.7$	$< 2.7 |\eta| < 3$	$|\eta| \geq 3$
Neutral hadron fraction	< 0.99	< 0.98	-
Neutral EM fraction	< 0.99	> 0.01	< 0.90
Number of constituents	> 1	-	-
Number of neutrals	-	> 2	> 10
Additional cuts for $|\eta| < 2.4$			
Charged hadron fraction	> 0	> 0	> 0
Charged multiplicity	> 0	> 0	> 0
Charged EM fraction	< 0.99	> 0.99	< 0.90

69

6.5.4 Jet calibration

As jets are complex objects consisting of highly energetic quarks and gluons that built

on input from several subdetector, the correct energy assignment to jets is a challenge.

Therefore, a set of jet energy corrections (JECs) are determined to account for various

effects and applied to jets in data and simulation. The different corrections are listed

below and follow the description outlined in [43].

- Pileup corrections: The so-called pileup offset corrections or L1 corrections are determined to reduce the effect of pileup on the jet energy. The corrections are determined in a simulated QCD dijet events, with and withoout pileup overlayed, before and after CHS applied. The corrections are parametrized as a function of a set of jet related quantities such as the jet area A_i, jet pseudorapidity η, jet p_T and the energy density ρ [5]

- MC-truth corrections: The MC-truth corrections, also known as L2L3 MC-truth corrections, is based on the comparison of a particle level jet and its corresponing reconstructed version. The corrections are derived from a QCD dijet simulated sample as a function of jet p_T and η.

- Residual corrections: The residual corrections, or L2L3 correction, are corrections to be applied on data only. The relative, or η dependent, correction is determined with QCD dijet events, where a tag jet in the barrel region is compared to a probe jet with no η restriction of similar p_T. The absolute correction is determined in γ+jets, Z($ee/\mu\mu$)+jets or multijet events, where a well-measured γ, Z boson or jet is compared to the recoiling jet

6.5.5 b-jet tagging

Heavy flavor jet identification algorithms are heavily used in CMS as a means to

select events containing top quark decays or SM bosons decaying through b quarks, and

the following description of the algorithms follow that introduced in [10]. As the lifetime

of a hadron containing b-quarks is in the order of 1.5 ps, this results in a displacement

---

[5] The energy density ρ is defined as the median of the transverse momentum of the jets over their area, $\rho = \mathrm{median}(p_{T,i}/A_i)$

of around a few millimeters up to a centimeter depending on the momentum available. As a result, this displacement leads to the possibility to reconstruct an additional vertex where the b hadron decay takes place. This secondary vertex (SV) is reconstructed from the displaced tracks that are characterized by their impact parameter (IP)[6] The SV reconstruction algorithm used in CMS in the LHC Run 2 is the inclusive vertex finder (IVF) algorithm. This algorithm uses as inout all reconstructed tracks in the event with $p_{\mathrm{T}} \geq 0.8\,\mathrm{GeV}$ and with longitudinal IP $\leq 0.3\,\mathrm{cm}$. The IVF algorithm is seeded by tracks with a three dimensional IP greater thatn $50\,\mu\mathrm{m}$ and the significance of the two dimensional IP (IP divided by its uncertainty) greater than 1.2. After this, the track clustering is performed by associating the seed track to any other track by imposing requirements on the distance at the point of closest approach and the angle between the tracks. The resulting track clusters are fitted using an adaptive vertex fitter. To resolve the track ambiguity when a track can be both associated to the PV and a SV, a track is discarded from the SV is it is more compatible with the PV. Following this track arbitration, the SV position is refitted. After the IVF algorithm is performed, the vertices found are used as input to the combined secondary vertex tagger version 2 (CSVv2), where in addition to the IVF vertex, two tracks per jet is required separated from the jet axis by $\Delta\mathrm{R}$ less than 0.3. The CSVv2 algorithm contains a step of training that is performed in the three independent vertex categories listed below.

- RecoVertex: a jet that contains more than one SV.

- PseudoVertex: no SV is found but at least two tracks with 2D IP significance larger than 2 and the combined invariant mass not compatible with a K_S^0.

- NoVertex: when none of the above categories are filled.

The variables combined in the algorithms are:

---

[6] The impact parameter is defined as the distance between the primary vertex and the tracks at their closest point of approach.

- SV 2D flight distance significance: the significance of the 2D flight distance of the SV with least uncertainty on its flight distance for jets in the RecoVertex category.

- Number of SVs: the number of SVs for jets in the RecoVertex category.

- Track η_{rel}: the track η relative to the jet axis for the track with the highest 2D IP significance for jets in the RecoVertex and PseudoVertex categories.

- Corrected SV mass: the corrected mass of the SV with the smallest uncertainty on its flight distance for jets in the RecoVertex category or the invariant mass obtained from the total summed four-momentum vector of the selected tracks for jets in the PseudoVertex category.

- Number of tracks from SV: the number of tracks associated with the SV for jets in the RecoVertex category or the number of selected tracks for jets in the PseudoVertex category.

- SV energy ratio: the energy of the SV with least uncertainty on its flight distance over the energy of the total summed four momentum vector of the selected tracks.

- ΔR(SV, jet): defined as the distance between the flight direction of the SV with least uncertainty on its flight distance and the jet axis for jets in the RecoVertex category, or the distance between the total summed four-momentum vector of the selected tracks for jets in the PseudoVertex category.

- 3D IP significance of first four tracks: the 3D IP significances of the four tracks with the highest 2D IP significance.

- Track $p_{T,rel}$: the track p_T relative to the jet axis, i.e. the track momentum perpendicular to the jet axis, for the track with the highest 2D IP significance.

- ΔR(track, jet): the distance between the track and the jet axis for the track with the highest 2D IP significance.

- Track $p_{T,rel}$ ratio: the track p_T relative to the jet axis divided by the magnitude of the track momentum vector for the track with the highest 2D IP significance.

- Track distance: the distance between the track and the jet axis at their point of closest approach for the track with the highest 2D IP significance.

- Track decay length: the distance between the PV and the track at the point of closest approach between the track and the jet axis for the track with the highest 2D IP significance.

- Summed tracks E_T ratio: the E_T of the total summed four-momentum vector of the selected tracks divided by the transverse energy of the jet.

- ΔR(summed tracks, jet): the ΔR between the total summed four momentum vector of the tracks and the jet axis.

- The number of selected tracks.

- The jet p_T and η.

The above discriminating variables are combined in a neural network and the final disciminator is shown in Figure 6-1
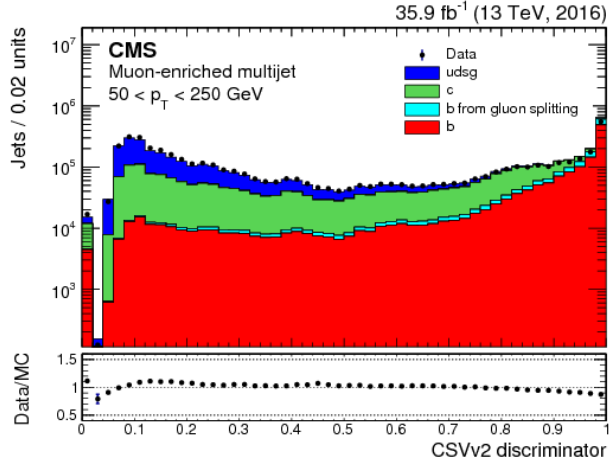
Figure 6-1. The CSVv2 discriminator variable[10]

6.5.6   Isotracks

Isolated tracks are used in the analysis as a means to improve the third lepton veto efficiency. These tracks are defined using charged PF candidates with different requirements depending on the flavor. PF electrons and PF muons, are required to pass $p_T > 5\,\text{GeV}$, $|dz| < 0.1\,\text{cm}$, as well as being well associated to the PV, and track isolation cuts of isolation/$p_T < 0.2$ and iso < 8 GeV. The track isolation sum is computed from all charged PF candidates within a cone of ΔR< 0.3, requiring them to pass $|dz| < 0.1\,\text{cm}$

with respect to the PV. Charged PF hadrons are required to pass $p_\mathrm{T} > 10\,\mathrm{GeV}$, $|dz| <$ $0.1\,\mathrm{cm}$, be well associated to the PV, and track isolation cuts of isolation$/p_\mathrm{T} < 0.1$ and isolation$< 8\,\mathrm{GeV}$. The track isolation is computed in the same way as for PF leptons above.

6.6   Datasets

As the searches presented in this thesis has the commonality that two same flavor opposite sign leptons are produced, naturally, the dielectron and dimuon streams of $13\,\mathrm{TeV}$ pp collision data are used. Further, additional datasets are used to predict a main SM background, collected with electron-muon, $p_\mathrm{T}^\mathrm{miss}$ and JetHT triggers. All datasamples are summarized in Table 6-4.

Table 6-4. Datasets used in the strong, electroweak and slepton searches and $p_\mathrm{T}^\mathrm{miss}$ study

Signal events
/DoubleEG/Run2016B-03Feb2017_ver2-v2/MINIAOD
/DoubleEG/Run2016(C-G)-03Feb2017-v1/MINIAOD
/DoubleEG/Run2016H-03Feb2017_ver2-v1/MINIAOD
/DoubleEG/Run2016H-03Feb2017_ver3-v1/MINIAOD
/DoubleMuon/Run2016B-03Feb2017_ver2-v2/MINIAOD
/DoubleMuon/Run2016(C-G)-03Feb2017-v1/MINIAOD
/DoubleMuon/Run2016H-03Feb2017_ver2-v1/MINIAOD
/DoubleMuon/Run2016H-03Feb2017_ver3-v1/MINIAOD
Datasets for background prediction
/MuonEG/Run2016B-03Feb2017_ver2-v2/MINIAOD
/MuonEG/Run2016(C-G)-03Feb2017-v1/MINIAOD
/MuonEG/Run2016H-03Feb2017_ver2-v1/MINIAOD
/MuonEG/Run2016H-03Feb2017_ver3-v1/MINIAOD
/JetHT/Run2016B-03Feb2017_ver2-v2/MINIAOD
/JetHT/Run2016(C-G)-03Feb2017-v1/MINIAOD
/JetHT/Run2016H-03Feb2017_ver2-v1/MINIAOD
/JetHT/Run2016H-03Feb2017_ver3-v1/MINIAOD
/MET/Run2016B-03Feb2017_ver2-v2/MINIAOD
/MET/Run2016(C-G)-03Feb2017-v1/MINIAOD
/MET/Run2016H-03Feb2017_ver2-v1/MINIAOD
/MET/Run2016H-03Feb2017_ver3-v1/MINIAOD

6.7   Triggers

The trigger selection for the searches presented in this thesis is driven by the requirement of at least two leptons at the HLT. Due to the changes in instantaneous luminosity, different dilepton triggers were active at different times and with varying prescales. This results in the need for a variety of triggers with slighlty different requirements. Non isolated double lepton paths are included to increase the efficiency in events with large dilepton system p_T. Triggers with an isolation requirement on the leptons enable for the recording of lower p_T leptons. The p_T requirements are asymmetric and depend on the flavor composition of the triggers. Supporting triggers,

with requirements on the jet H_T or online $p_\mathrm{T}^{\mathrm{miss}}$, are used for the study of the trigger efficiencies used in the FS background prediction method, taken from hadronic events. Additionally, triggers with requirements on the presence of an electron and a muon, are used to collect a sample dominated by $t\bar{t}$ events for the FS prediction background method. All signal and supporting triggers are documented in Table 6-5.

Several MC event generators are used to simulate the background and signal processes in this analysis, with the different parts of the generators introduced in Section 4.3. The simulation is normalized to luminosity using cross sections from `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns`. The `PYTHIA8` [44] package is used for parton showering, hadronization and undelying event simulation with the tune `CUETP8M1`, as described in Section 4.3. The various simulated samples used are presented in Appendix A.

Table 6-5. Triggers used in the strong, electroweak and slepton searches. The first section are the triggers used in most control and signal regions, while the supporting triggers are used for the calculation of the trigger efficiencies of the signal triggers.

Signal triggers
Di-muon triggers
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v*
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
HLT_Mu27_TkMu8_v*
HLT_Mu30_TkMu11_v*
Di-electron triggers
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_v*
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_MW_v*
Electron-muon triggers
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu30_Ele30_CaloIdL_GsfTrkIdVL_v*
HLT_Mu33_Ele33_CaloIdL_GsfTrkIdVL_v*
Supporting triggers
HLT_PFHT125_v*
HLT_PFHT200_v*
HLT_PFHT250_v*
HLT_PFHT300_v*
HLT_PFHT350_v*
HLT_PFHT400_v*
HLT_PFHT475_v*
HLT_PFHT600_v*
HLT_PFHT650_v*
HLT_PFHT800_v*
HLT_PFHT900_v*

CHAPTER 7

EVENT SELECTIONS

Common to the SUSY searches presented in this thesis is the final state containing two leptons and p_T^{miss} originating from the LSP. The difference lies in the hadronic activity and whether the lepton pair is compatible with a Z boson or not. This chapter contains an overview event selections for the signal regions for the strong, electroweak and slepton searches. Apart from the basic selections on leptons, jets and b-jets, more specific variables used to define the signal regions is also presented later in the chapter.

7.1 Common signal region variables

7.1.1 Lepton pair selection

The principle behind the main background prediction method, relies on the lepton flavor symmetry of the W decay, the identification and isolation requirements of the leptons are chosen so that they are as similar as possible between the flavors. This principle is reflected in the selections of the trigger requirement of the leptons of $p_T > 23$, 17, 12, and 8 GeV depending on the exact path. Full efficiency for any of these values is reached at a p_T of 25(20) GeV for the leading (trailing) lepton, which is the requirement on the leptons for the control region. One specific selection of this analysis is that not only the electrons are rejected if they appear in the transition region between the barrel and the endcap, but also the muons. The reason for this is that the flavor symmetric

background is taken from $e\mu$ events, thus the necessity of having symmetric cuts not only on the efficiency but also on the fiducial regions. For this reason any lepton within the $|\eta|$ region of 1.4 to 1.6 is rejected, to keep the reconstruciton of electrons and muons as similar as possible. Since there are some events with multiple lepton pairs, it is important to define an unambiguous way of selecting the "relevant" opposite-sign, same-flavor lepton pair. The implemented algorithm selects the two highest p_{T} leptons which are fully identified and that have a distance between them of 0.1 in ΔR. This is to say, there is no cross-cleaning or prioritization between electrons and muons applied, and non-identified leptons (including the ECAL transition region) do not enter in the consideration of the lepton pair selection. Further, the leptons can be used to define the signal regions for the different searches, by imposing requirements on the invariant mass of the leptons. In the searches where an on-shell production of a Z boson is expected, naturally the signal region is defined to have the lepton pair compatible with the Z boson mass. This is the case for the Electroweak search, and the gluino induced GMSB search. On the other hand, the strong SUSY search with a sbottom induced process and an intermediate slepton decay result in a lepton pair that is not compatible with the Z boson mass. This search instead utilizes bins in $m_{\ell\ell}$ outside of the Z mass window of 86-96 GeV. Finally, the search for direct slepton production inlcudes two leptons that are not compatible with the Z mass. To supress the DY SM background, a generous veto on the Z boson mass of 76-106 GeV is applied in the signal region. All the signal and control region criteria on the leptons is summarized in Table 7-2

7.1.2   Jet and b-jet selection

The jets used throughout this thesis are clustered from PF objects using anti-k_t algorithm with a distance parameter of 0.4, after excluding charged hadrons originating from pileup, and corrected with JECs, as described in Section 6.5. Identification of jets originating from b quark decays is done using the CSVv2 algortihm, introduced in Section

Table 7-1. Lepton kinematic criteria.

Strong GMSB and Electroweak SR lepton selection	
Flavor	$e^+e^-/\mu^+\mu^-$
Leading p_T	$> 25\,\mathrm{GeV}$
Subleading p_T	$> 20\,\mathrm{GeV}$
$|\eta|$	< 2.4 and $\ni [1.4,\ 1.6]$
$m_{\ell\ell}$	$\in [86,\ 96]$ GeV
Strong sbottom search lepton selection	
Flavor	$e^+e^-/\mu^+\mu^-$
Leading p_T	$> 25\,\mathrm{GeV}$
Subleading p_T	$> 20\,\mathrm{GeV}$
$|\eta|$	< 2.4 and $\ni [1.4,\ 1.6]$
$m_{\ell\ell}$	bins from 20 to 400+ GeV
Slepton SR lepton selection	
Flavor	$e^+e^-/\mu^+\mu^-$
Leading p_T	$> 50\,\mathrm{GeV}$
Subleading p_T	$> 20\,\mathrm{GeV}$
$|\eta|$	< 2.4 and $\ni [1.4,\ 1.6]$
$m_{\ell\ell}$	$\ni [76,\ 106]$ GeV
CR lepton selection	
Flavor	$e^+e^-/\mu^+\mu^-/e^\pm\mu^\mp$
Leading p_T	$> 25\,\mathrm{GeV}$
Subleading p_T	$> 20\,\mathrm{GeV}$
$|\eta|$	< 2.4 and $\ni [1.4,\ 1.6]$

6.5.5. Further, the scalar sum of all jet p_T is referred to as H_T and is used to design some of the signal regions. The p_T of the jets used for the control regions used for the flavor symmetric background prediction methods is 35 GeV and the jets are required to be within $|\eta| < 2.4$. The jets are required to be seperated from selected leptons by 0.4 in $\Delta\mathrm{R}$. In the slepton search no hadronic activity is expected, therefore a veto on jets of p_T greater than $25\,\mathrm{GeV}$ and $|\eta| < 2.4$ is imposed.

Table 7-2. Jet kinematic criteria.

Strong and electroweak search jet selection	
AK4 jets ≥ 2	
p_T	$> 35\,\mathrm{GeV}$
$|\eta|$	< 2.4
b-tagged jets ≥ 0	
CSVv2 discriminator	Medium WP
p_T	$> 25\,\mathrm{GeV}$
$|\eta|$	< 2.4
Slepton search jet selection	
AK4 jets $= 0$	
p_T	$> 25\,\mathrm{GeV}$
$|\eta|$	< 2.4
CR jets	
AK4 jets ≥ 2	
p_T	$> 35\,\mathrm{GeV}$
$|\eta|$	< 2.4

7.1.3 $p_\mathrm{T}^\mathrm{miss}$

Common to the various searches presented in this thesis is the existence of LSPs, as only R-parity conserving models are considered. The LSPs are massive and escape detection, and would manifest themselves through the imbalance in the transverse plane as $p_\mathrm{T}^\mathrm{miss}$. For this reason, all signal regions are exploit the $p_\mathrm{T}^\mathrm{miss}$ either explicitly through a binning in the variable, or implicitly in the M_{T2} variable that uses it as an input. The "Type-1 corrected" $p_\mathrm{T}^\mathrm{miss}$ is used, with all $p_\mathrm{T}^\mathrm{miss}$ filters applied to remove anomalous $p_\mathrm{T}^\mathrm{miss}$ events, see Chapter 8 for more details.

7.1.4 M_{T2}

The leptonic M_{T2} variable is used to define all signal regions. It is a generalization of the transverse mass for pair-produced particles which decay into visible and invisible objects, as described in [45, 46]. The variable is efficient in differentiating SM $t\bar{t}$ from the

signal final state scenarios, as it has a clear endpoint at the W boson mass for SM $t\bar{t}$. The idea behind this powerful variable is summarized below. The a leptonic decay of a top quark, the process can be summarized as follows

$$t \to W^+ b \to l^+ \nu b \tag{7–1}$$

From a theoretical point of view, it is easy to construct the transverse mass (m_T) of the W boson using the kinematic variables of the lepton and the neutrino and energy and momentum conservation in the transverse plane,

$$M_T = \sqrt{m_l^2 + m_\nu^2 + 2(E_T^l E_T^\nu - \vec{p}_T^l \vec{p}_T^\nu)} \tag{7–2}$$

where the $\vec{p}_T^l$ and $\vec{p}_T^\nu$ are the transverse momenta of the lepton and the neutrino, and E_T^l and E_T^ν the transverse energy of the same. In this case, the m_T gets a value very close to the W boson mass. In the case of di-leptonic $t\bar{t}$,

$$pp \to t\bar{t} \to W^+ b + W^- \bar{b} \to l^+ \nu b + l^- \bar{\nu} \bar{b} \tag{7–3}$$

the situation gets complicated. First of all, the pairing of the leptons and neutrinos that form the W bosons is not known. Further, the neutrinos escape detection and the momenta they carry is just collectively summed up in the p_T^{miss}. But let us disregard these experimental difficulties for a second. If one instead assume that the correct pairing is known, then one could create an upper bound dictated by the W boson mass according to

$$M_W^2 \geq \max\{M_T^2\left(\vec{p}_T^{l^+}, \vec{p}_T^\nu\right), M_T^2\left(\vec{p}_T^{l^-}, \vec{p}_T^{\bar{\nu}}\right)\}. \tag{7–4}$$

Meaning that in the case of di-leptonic $t\bar{t}$, if the lepton pairing is right the both transverse masses found would be close to the W boson mass. In order to cope with the problem of not knowing the p_T of the neutrinos, one can instead scan over all possible p_T^{miss} partitions:

$$M_W \geq \min_{\vec{p}_T^{miss} = \vec{p}_{T1}^{miss} + \vec{p}_{T2}^{miss}} \left(\max\{M_T\left(\vec{p}_T^{l_1}, \vec{p}_{T1}^{miss}\right), M_T\left(\vec{p}_T^{l_2}, \vec{p}_{T2}^{miss}\right)\}\right). \tag{7–5}$$

This is the formal definition of the M_{T2}

$$M_{T2} = \min_{\vec{p}_T^{miss} = \vec{p}_{T1}^{miss} + \vec{p}_{T2}^{miss}} \left(\max\{M_T\left(\vec{p}_T^{l_1}, \vec{p}_{T1}^{miss}\right), M_T\left(\vec{p}_T^{l_2}, \vec{p}_{T2}^{miss}\right)\} \right). \qquad (7\text{--}6)$$

For dileptonic $t\bar{t}$, the M_{T2} will be very close to the W boson mass, whereas for signal scenarios, where the leptons are paired with the momentum of the LSPs, which will result in values much larger than the W boson mass. The M_{T2} variable is used to define all signal regions in all searches, as it effectively reduce contributions from SM $t\bar{t}$ by imposing a cut around $80\,\text{GeV}$.

7.2   Strong search

7.2.1 $t\bar{t}$ likelihood

7.2.2   Strong SR definition

7.3   Electroweak search

7.3.1   Boosted bosons

7.3.2   Electroweak SR definition

7.4   Slepton search

7.4.1   Slepton SR definition

CHAPTER 8

PERFORMANCE OF MISSING TRANSVERSE MOMENTUM

At the forefront of SUSY physics program are searches where R-parity is conserved, resulting in one or more lightest SUSY particles that have no SM particles to decay into. The result is LSPs that are unable to interact with the detector material and thus escape detection. The existence of such particles can be inferred by the momentum imbalance in the transverse plane, $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, with its magnitude denoted $p_{\mathrm{T}}^{\mathrm{miss}}$. When the LSPs are massive, the $p_{\mathrm{T}}^{\mathrm{miss}}$ provides an excellent search tool for SUSY. But other sources can contribute to a large momentum imbalance. Any process with a leptonically decaying W-boson produces a neutrino that escape the detector similarly as the LSP. Additionally, as jets are complex objects to measure, and their energy are corrected through JECs, any over or under measurement in the jets will result in $p_{\mathrm{T}}^{\mathrm{miss}}$. In order to perform a SUSY search where R-parity is conserved, a deep understanding of the $p_{\mathrm{T}}^{\mathrm{miss}}$ object is needed to distinguish the $p_{\mathrm{T}}^{\mathrm{miss}}$ originating from LSPs from SM neutrinos and jet mismeasurements and detector inefficiencies. A challenge for reconstructing physics objects is differentiating tracks from the primary vertex with tracks from overlapping bunch crossings in multiple pp collisions (pileup). A detailed study of the performance of two commonly used $p_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction algorithms is presented in this chapter, along with a specific study analyzing the performance of the algorithms under extreme pileup conditions, as is expected in the High Luminosity phase of the LHC.

8.1 Missing transverse momentum algorithms in CMS

In collision events, the transverse momentum of the partons is small compared to the energy available in the center of mass, and does not depend on their longitudinal energy. Thus, an assumption that the initial transversal momentum of the system formed by the partons is zero can be made. As an effect of this assumption, if particles escape detection, a transverse energy inequilibrium is created. The final states containing one or more neutrinos therefore result in a significant missing energy corresponding to the vectorial sum of the neutrino momenta. When no neutrinos are created in the event, all missing transverse energy is due to detector inefficiencies and reconstruction issues. Therefore, final states without neutrinos are ideal for the study of the performance of the missing transverse momentum reconstruction algorithms originating from detector effects. The $p_\mathrm{T}^\mathrm{miss}$ is defined as the negative vectorial sum of the particles in the event

$$\vec{p}_\mathrm{T}^\mathrm{miss} = -\sum \vec{p}_\mathrm{T} \qquad (8\text{--}1)$$

and its magnitude is denoted $p_\mathrm{T}^\mathrm{miss}$. In CMS, two algorithms for the $p_\mathrm{T}^\mathrm{miss}$ reconstruction are used, PF $p_\mathrm{T}^\mathrm{miss}$ and PUPPI $p_\mathrm{T}^\mathrm{miss}$, and both will be introduced in the following sections.

8.1.1 Particle Flow $p_\mathrm{T}^\mathrm{miss}$ reconstruction

The first reconstruction algorithm is PF $p_\mathrm{T}^\mathrm{miss}$, which is the magnitude of the negative of the vectorial sum of all PF candidates in an event:

$$\vec{p}_\mathrm{T}^\mathrm{miss} = -\sum_{i \in PF} \vec{p}_\mathrm{T,i} \qquad (8\text{--}2)$$

As will be shown in the following, the PF $p_\mathrm{T}^\mathrm{miss}$ algorithm is highly performant and is therefore used in the majority of CMS analyses.

8.1.2 PUPPI $p_\mathrm{T}^\mathrm{miss}$ reconstruction

The PUPPI $p_\mathrm{T}^\mathrm{miss}$ algorithm uses the 'pileup per particle identification' method [47]. This method has been developed to reduce the dependence of pileup on physics objects.

8.1.2.1 PUPPI algorithm

In this section, the PUPPI algorithm is summarized, for further clarification please refer to [47]. The idea is to estimate how likely the PF candidates are to be originating from pileup, and reweight the particle four-momentum accordingly, with a weight, w_i, close to 1 if the candidate is from the hard scatter and close to 0 for particles from pileup. The procedure to calculate the w_i starts with defining a shape α_i for each particle,

$$\alpha_i = \log \sum_{\substack{j \in \text{event} \\ j \neq i}} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2 \times \Theta(\Delta R_{ij} - R_{min}) \times \Theta(R_0 - \Delta R_{ij}), \qquad (8\text{--}3)$$

where Θ is the Heaviside step function. The α of the i-th particle is thus depending on the p_T of the surrounding particles, and the distance between them in $\eta - \phi$ space, defined as the cone ΔR_{ij}. Only particles within some R_0 around particle i are considered. Surrounding particles j are discarded that are within some minimum radius R_{min} close to the particle i, to reduce the effect from collinear splittings. When the particle i is from hard scattering, the surrounding particles tend to be close in ΔR because of the collinear singularity of the parton shower, resulting in a relatively larger α_i. On the other hand, a wider separation in ΔR is expected for particles originating from pileup, as they should have no correlation with the direction of particle i, resulting in a smaller value for α_i. The p_T of the j-th particles is also used in the calculation of α_i, and the characteristic of this variable is that it is generally softer for particles originating from pileup, yielding the desired smaller value of α_i, and the opposite for when the particle is from the hard scattering. Now that the α_i is defined, the question of what particles should be summed over arise. For this, two regions are used, reflecting the design of the detector; the central

region ($|\eta| \leq 2.4$), in which the tracking can distinguish charged tracks from the primary vertex from charged tracks from pileup vertices, and the forward region ($|\eta| > 2.4$) where this discrimination is not possible. Where tracking is available, the PF algorithm can provide the following PF candidates; neutral particles, charged hadrons from the primary vertex and chanrged hadrons from pileup vertices. This results in two different computations of α_i, namely

$$\alpha_i^C = \log \sum_{\substack{j \in \text{Ch, LV} \\ j \neq i}} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2 \times \Theta(\Delta R_{ij} - R_{min}) \times \Theta(R_0 - \Delta R_{ij}), \tag{8–4}$$

$$\alpha_i^F = \log \sum_{\substack{j \in \text{event} \\ j \neq i}} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2 \times \Theta(\Delta R_{ij} - R_{min}) \times \Theta(R_0 - \Delta R_{ij}), \tag{8–5}$$

where α_i^C is the sum over all PF candidates, whereas the α_i^F is the sum over all particles in the event. The difference between these two computations is that in the central case, a particle j originating from pileup is discarded from the event, whereas this distinction can not be done in the forward region. However, both methods calculate the w_i from the α_i similarly for each particle, that is used to rescale its four momentum. The actual translation to a weight ranging from 0 to 1 is done by introducing the following quantity

$$\chi_i^2 = \Theta(\alpha_i - \bar{\alpha}_{\text{PU}}) \times \frac{(\alpha_i - \bar{\alpha}_{\text{PU}})^2}{\sigma_{\text{PU}}^2} \tag{8–6}$$

where the $\bar{\alpha}_{\text{PU}}$ and $\bar{\sigma}_{\text{PU}}$ are used to characterize the distributions on an event-by-event basis, and defined as

$$\bar{\alpha}_{\text{PU}}^{\eta} = \text{median}\{\alpha_{i \in \text{Ch,PU}}^{\eta}\} \tag{8–7}$$

and

$$\bar{\sigma}_{\text{PU}}^{\eta} = \text{RMS}\{\alpha_{i \in \text{Ch,PU}}^{\eta}\}. \tag{8–8}$$

The super script η is C for central or F for forward regions, indicating what region is used for the computation. As can be seen in Eq. 8–6, the χ_i^2 distribution quantifies how much the α_i value is fluctuating from the pileup median $\bar{\alpha}_{\text{PU}}$. Any value of α_i below the $\bar{\alpha}_{\text{PU}}$ is

considered pileup like, and due to the definition involving the Heaviside function, these values will result in a χ_i^2 of 0. Conversely, large values of α_i that are far from the $\bar{\alpha}_{\mathrm{PU}}$ will result in a large χ_i^2. Finally, the w_i is defined by

$$w_i = F_{\chi^2,\mathrm{NDF}=1}(\chi_i^2) \qquad (8\text{–}9)$$

with $F_{\chi^2,\mathrm{NDF}=1}$ being the cumulative distribution function of the χ^2 distribution. As a result, whenever the χ_i^2 is 0, the final w_i is 0, whenever the χ_i^2 is large, the final w_i is 1, and all values of χ_i^2 in between results in a fractional weight between 0 and 1.

8.1.2.2 PUPPI $p_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction

Now that the particle weights, that are the foundations of the PUPPI algorithms, have been defined the actual rescaling of the particles and how they enter the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation will be covered. For each event, the value of α_i^{η} are computed for all charged pileup, and the corresponding median and RMS distributions $\bar{\alpha}_{\mathrm{PU}}^{\eta}$ and $\bar{\sigma}_{\mathrm{PU}}^{\eta}$. When using PF algorithm, the particles available are neutral particles, charged hadrons from the primary vertex, and charged hadrons from pileup. Where tracking is available, the particles originating from pileup can be easily distinguished, and those receive a weight of 0, and are completely discarded in the remainder of the calculation, whereas the charged hadrons from the primary vertex receive a weight of 1. The weights w_i of the remaining particles are calculated, and the four-momentum of these particles is rescaled by this w_i. The charged hadrons from the primary vertex and the rescaled remaining particles used to reinterpret the event, in the context of jet clustering, or in this case, the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation, according to:

$$\text{PUPPI } \vec{p}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{i \in PF} w_i \times \vec{p}_{\mathrm{T,i}} \qquad (8\text{–}10)$$

8.2 Calibration of p_T^{miss}

As the p_T^{miss} reconstruction is depending on the accurate measurement of all the reconstructed physics objects, any inefficiency in the reconstruction or minimum energy or p_T thresholds will bias the energy scale of the p_T^{miss}. As described in Section 6.5.4, the energy of the jets are corrected with JECs. If these corrections are not taken into account in the computation of the p_T^{miss}, there will be a significant bias and imbalance in the event. Therefore, the energy scale of p_T^{miss} is improved by propagating the correction of the p_T of the jets, $\vec{p}_{T,\mathrm{jet}}^{\,\mathrm{corr}}$ to p_T^{miss} in the following way:

$$\text{Type 1}\ \vec{p}_T^{\,\mathrm{miss}} = \vec{p}_T^{\,\mathrm{miss}} - \sum_{\mathrm{jets}}(\vec{p}_{T,\mathrm{jet}}^{\,\mathrm{corr}} - \vec{p}_{T,\mathrm{jet}}) \tag{8–11}$$

In the rest of the thesis, the "Raw p_T^{miss} " is the uncorrected p_T^{miss}, and the corrected p_T^{miss}, commonly known as the "Type-1 p_T^{miss}", will be referred to as just p_T^{miss}. As jets are complex objects to measure, the choice of the jets whose corrections should be taken into account is a question of optimization, and subject to future improvement. The motivation behind the 15 GeV p_T threshold of the jets is to reduce the contribution of jets from pileup. As will be shown in the next section, this choice of the p_T threshold gives a response very close to unity. Further, ambiguity can arise if a jet is very close to a reconstructed muon, or ressembles an electron or a photon. If a muon reconstructed using the outer tracking system overlaps with a jet, its four momentum is subtracted from the four momentum of the jet, and the JES correction appropriate for the modified jet momentum is used in the p_T^{miss} calculation. Jets are reconstructed from energy deposited in both the HCAL and ECAL, with various fractions of the energy in each calorimeter. In order to not correct jets that are in fact an electron or a photon, i.e. with a large electromagnetic (EM) energy fraction, a choice is made to only correct jets with an EM fraction of less than 90%. The choice of 90% has been providing a well calibrated p_T^{miss} object but can be improved, as will be shown in the next section. The p_T^{miss} relies

on the accurate measurement of the reconstructed physics objects, namely muons, electrons, photons, hadronically decaying τ leptons, jets, and unclustered energy (UE). By factorizing the p_T^{miss} into these physics objects, and vary each object within its momentum scale and resolution uncertainties, provides a good estimate of the uncertainty that each of the object contribute to the p_T^{miss}. In the rest of this chapter, the uncertainty in the p_T^{miss} is evaluated by comparing the recalculated p_T^{miss} to the nominal, not varied, p_T^{miss}. As already hinted, the uncertainty related to the energy measurement of the jets is the dominant uncertainty in these measurements. These uncertainties are splitted into those relating to the JES uncertainties, which are up to 3% (12%) for jets inside (outside) the tracker acceptance, and the JER uncertainties that range between 5–20%. A subdominant uncertainty is that related to the measurement of the muon energy scale, which amuonts to 0.2%, and electron and photon energy scale, which amounts to 0.6% (1.5%) in the barrel (endcap). The UE uncertainty is evaluated based on the momentum resolution of each PF candidate, which depends on the type of the candidate. A detailed description of the PF candidate calibration can be found in Refs. [5, 35, 48]. The p_T measurement of PF charged hadrons is dominated by the tracker resolution. For PF neutral hadrons, the p_T resolution is dominated by the resolution of the HCAL. The ECAL resolution dominates the PF photon p_T measurement, whereas HF intrinsic resolution dominates that for the PF particles in the HF.

8.3 Event selection

Dilepton and single-photon samples are used to study the p_T^{miss} response and resolution. These samples are chosen as they contain events where no genuine p_T^{miss} from neutrinos is expected, and serves as a good tool to measure the performance of the p_T^{miss} originating from detector inefficiencies or jet mismeasurements.

8.3.1 Dilepton event samples

The dilepton samples are subdivided into two categories based on the flavor of the lepton, namely $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ samples. The dileptonic datasets are presented in Table 8-1

Table 8-1. Datasets used for the p_T^{miss} study

Dilepton datasets used for p_T^{miss} performance study
Dielectron samples
/DoubleEG/Run2016B-03Feb2017_ver2-v2/MINIAOD
/DoubleEG/Run2016(C-G)-03Feb2017-v1/MINIAOD
/DoubleEG/Run2016H-03Feb2017_ver2-v1/MINIAOD
/DoubleEG/Run2016H-03Feb2017_ver3-v1/MINIAOD
Dimuon samples
/DoubleMuon/Run2016B-03Feb2017_ver2-v2/MINIAOD
/DoubleMuon/Run2016(C-G)-03Feb2017-v1/MINIAOD
/DoubleMuon/Run2016H-03Feb2017_ver2-v1/MINIAOD
/DoubleMuon/Run2016H-03Feb2017_ver3-v1/MINIAOD

The events for the $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ samples are recorded using dimuon and dielectron triggers, as presented in Table 8-2 that select events where the p_T of the two leading leptons are above asymmetric thresholds. The simulated samples used are presented in Appendix B.

Candidate events are required to have both the leading (subleading) lepton p_T greater than 25 (20) GeV and an invariant mass in the range of 80 to 100 GeV, compatible with the mass of the Z boson. In order to have a pure sample of dilepton events originating from Drell–Yan production, a veto is applied on any event containing a third lepton of $p_T > 20$ GeV. The spectrum of the Z boson transverse momentum, q_T, is shown in Fig. 8-1 where only the statistical uncertainty in the simulated samples is considered as the dilepton energy resolution is very good.

Table 8-2. Triggers used for the p_T^miss performance study.

Dilepton triggers used for p_T^miss study
Dimuon and single muon triggers
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v*
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
HLT_Mu27_TkMu8_v*
HLT_Mu30_TkMu11_v*
HLT_Mu50_v*
HLT_IsoTkMu22_v*
HLT_IsoMu27_v*
HLT_IsoMu24_v*
HLT_IsoMu22_v*
Dielectron triggers
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_v*
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_MW_v*
HLT_Ele27_eta2p1_WPLoose_Gsf_v*
HLT_Ele27_WPTight_Gsf_v*
HLT_Ele35_WPLoose_Gsf_v*

8.3.2   Single-photon event sample

The events in the single-photon sample are presented in Table 8-3, and are selected using a set of prescaled isolated single-photon triggers with varying thresholds, presented in Table 8-4. The simulated samples used are presented in Appendix B.

Table 8-3. Datasets used for p_T^miss study

Datasets used for p_T^miss study
/SinglePhoton/Run2016B-03Feb2017_ver2-v2/MINIAOD
/SinglePhoton/Run2016(C-G)-03Feb2017-v1/MINIAOD
/SinglePhoton/Run2016H-03Feb2017_ver2-v1/MINIAOD
/SinglePhoton/Run2016H-03Feb2017_ver3-v1/MINIAOD
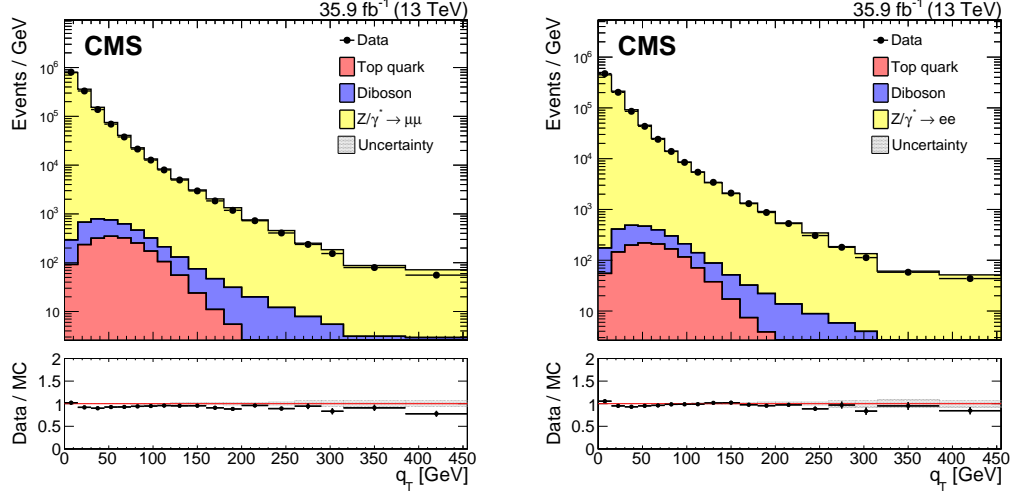
Figure 8-1. Upper panels: Z boson q_T in $Z \to \mu^+\mu^-$ (left) and $Z \to e^+e^-$ (right) samples. The Diboson contribution corresponds to processes with two electroweak bosons produced in the final state. The Top quark contribution corresponds to the top anti-top pair and single top production processes. Lower panel: Data to simulation ratio. The band corresponds to the statistical uncertainty in simulated samples.

Table 8-4. Triggers used for the p_T^{miss} performance study.

Single photon triggers used for p_T^{miss} study
HLT_Photon30_R9Id90_HE10_IsoM_v*
HLT_Photon50_R9Id90_HE10_IsoM_v*
HLT_Photon75_R9Id90_HE10_IsoM_v*
HLT_Photon90_R9Id90_HE10_IsoM_v*
HLT_Photon120_R9Id90_HE10_IsoM_v*
HLT_Photon165_R9Id90_HE10_IsoM_v*

The p_T thresholds of the triggers are 30, 50, 75, 90, 120, and 165 GeV, and the first five triggers had different L1 accept rate (prescale) during the data taking periods, following the luminosity. Candidate events are weighted based on the prescale values of the triggers.

One tight ID photon with $p_T > 50$ GeV is required and events are vetoed that contain leptons of $p_T > 20$ GeV. In order to isolate the events needed for this study, a requirement to have at least one jet with p_T greater than 40 GeV that recoils off of the photon. To match the trigger conditions, the leading photon is further required to have the ratio of

the energy deposited in a 3×3 crystal region of the ECAL, centered around the crystal containing an energy deposit greater than all of its immediate neighbors, to the energy of the entire deposit of the photon greater than 0.9. The photon q_T spectrum is shown in Fig. 8-2. Similarly to Fig. 8-1 only the statistical uncertainty in the simulated samples is considered as the photon energy resolution is very good.
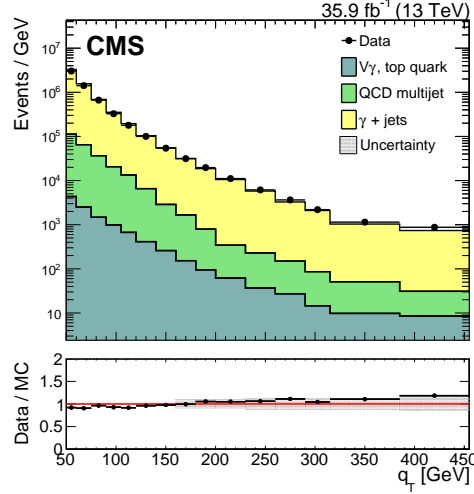
Figure 8-2. Upper panel: Distribution of the photon q_T in the single-photon sample. The Vγ+Top quark contribution corresponds to the Zγ, Wγ, top anti-top pair and single top production processes. Lower panel: Data to simulation ratio. The band corresponds to the statistical uncertainty in the simulated samples.

8.4 p_T^{miss} filters

Non-physical, anomalous high-p_T^{miss} events can arise because of a variety of reconstruction failures or malfunctioning detectors. In the p_T^{miss} study presented below, and in the two analyses covered in this thesis, these anomalous p_T^{miss} events are filtered away on an event basis. In the following, the various filters are presented and the underlying source of the anomalous p_T^{miss}.

8.4.1 HCAL filters

The geometrical patterns of HPD or RBX channels as well as the pulse shape and timing information are utilized by various HCAL barrel and endcap (HBHE) algorithms to identify and eliminate noise. These filter algorithms operate both in "noise filtering" and "event filtering" modes. In the noise filtering mode, the anomalous energy deposits are removed from the event reconstruction; in the event filtering mode, the bunch crossing is removed from the data set. In addition, there is an isolation-based noise filter that utilizes a topological algorithm, where energy deposits in HCAL and ECAL are combined and compared with measurements from the tracker to identify isolated anomalous activity in HBHE.

8.4.2 ECAL filters

One large source of anomalous $p_\mathrm{T}^\mathrm{miss}$ signals can be created if a hadron hits the front-end electronics of the ECAL super crystals, creating a large false signal. Additionally, anomalously high energetic deposits in the supercrystals, and the lack of information for channels that have nonfunctioning readout electronics, are removed through dedicated noise filters. During the datataking relevant for this thesis, five ECAL endcap supercrystals produced large, anomalous pulses, leading to spurious $p_\mathrm{T}^\mathrm{miss}$. These crystals were removed from the readout, and their energies were not considered. Furthermore, in about 0.7% of ECAL towers, the crystal-by-crystal information is not available. The trigger primitive (TP) information, however, is still available, and is used to estimate the energy. The TP information saturates above 127.5 GeV. Events with a TP close to saturation in one of these crystals are removed.

8.4.3 Beam halo filter

A final large source of anomalous large p_T^{miss} is due to so called machine-induced backgrounds or beam halo, meaning the production of muons when beam protons undergo collisions upstream of the detector. The characterisitcs of beam halo particles is that they travel parallel to the collision axis. If they leave energy deposits in the calorimeters, those will be along a line with constant ϕ. Similarly, any interaction of the beam halo particles in the CSC, will be in line with the calorimeter deposits. The filter exploits information from both the CSC and the calorimeters, and an example event display for a beam halo event is shown in Fig. 8-3, where collinear hits in the CSC are visible. To visualize the
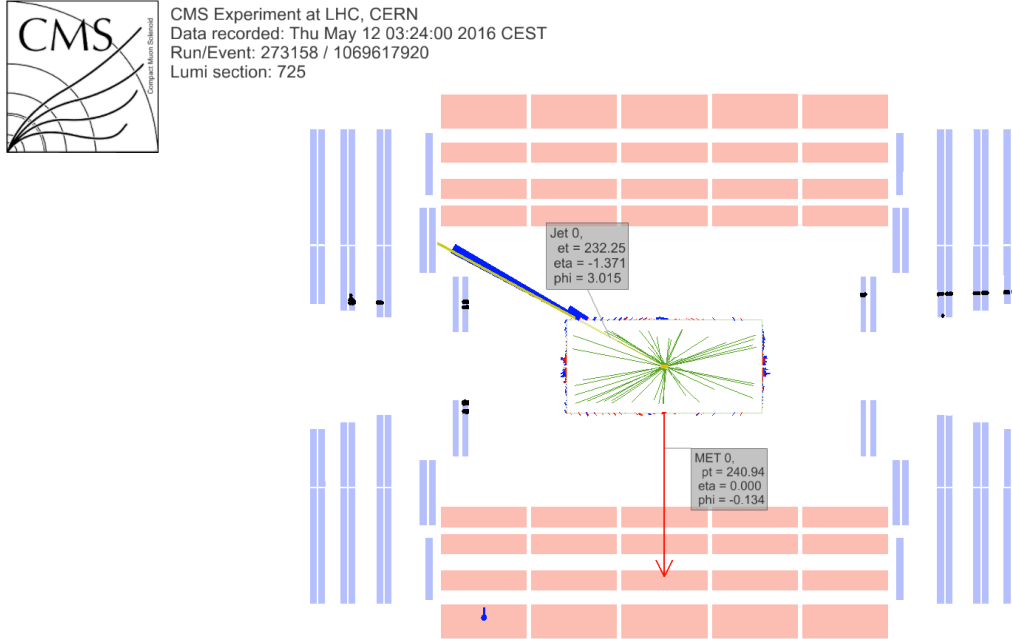
Figure 8-3. Event display for a beam halo event with collinear hits in the CSC (black), p_T^{miss} of $\sim 241\,\text{GeV}$, and a jet with $p_T =\sim 232\,\text{GeV}$.

need of these event filters, dijet and monojet events are used. Figure 8-4 (left) shows a comparison of the p_T^{miss} before and after the application of the event filters for the dijet sample, where the events with large p_T^{miss} are found to mostly to be due to electronic noise

in the calorimeters. Figure 8-4 (right) shows a comparison of the jet ϕ distribution before and after the application of the event filters for the monojet sample, where the excess of events with jet $\phi \approx 0$ or $\phi \approx \pi$ are typical for beam halo events.
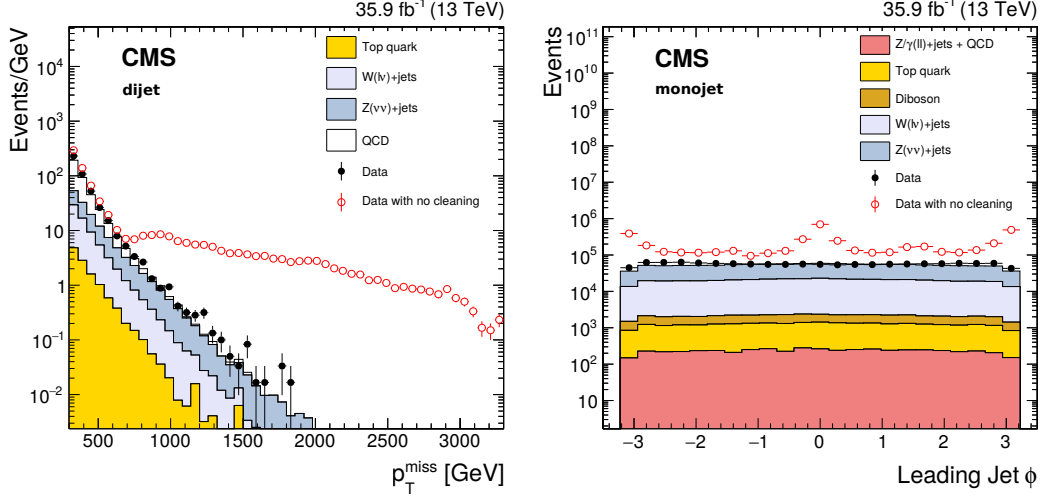
Figure 8-4. The p_T^{miss} (left) and jet ϕ (right) distributions for events passing the dijet (left) and monojet (right) selection with the event filtering algorithms applied, including that based on jet identification requirements (filled markers), without the event filtering algorithms applied (open markers), and from simulation (solid histograms).

8.5 p_T^{miss} performance

After having covered the p_T^{miss} reconstruction algorithms, their calibration and the filtering of anomalous p_T^{miss} events, the main part of the study will be presented. As has already been conveyed, the p_T^{miss} is a very sensitive observable, and is relying on accurate object reconstruction and efficient sub-detectors. Therefore it is a great tool for monitoring the detector and reonstruction performance during data-taking. The performance of the p_T^{miss} can be summarized in three parts. By monitoring the data and simulation of the p_T^{miss} in events with no genuine p_T^{miss} from neutrinos, any anomalies will show up as p_T^{miss} tails, or an overall disagreement between the data and the simulation. The p_T^{miss} response is crucial for validating the JECs or any issues related to the muon, electron or photon

energy scale. The $p_\mathrm{T}^\mathrm{miss}$ resolution is valuable for monitoring how the performance degrades as a function of some variable, such as the number of vertices. Initially, the studies presented in this thesis were performed during the 2016 data-taking, and was crucial in uncovering various reconstruction inefficiencies, and was later refined and summarized in a publication.

8.5.1 $p_\mathrm{T}^\mathrm{miss}$ performance using hadronic recoil

A well-measured Z/γ boson provides a unique event axis and a precise momentum scale. Such events should have little or no genuine $p_\mathrm{T}^\mathrm{miss}$, and the hadronic recoil is projected onto the axis of this well measured boson, as illustrated in Fig. 8-5. Formally, the hadronic recoil (u) is defined as the vector p_T sum of all PF candidates except for the vector boson (or its decay products in the case of the Z boson decay). The assumption of momentum conservation in the transverse plane imposes $\vec{q}_\mathrm{T} + \vec{u}_\mathrm{T} + \vec{p}_\mathrm{T}^\mathrm{miss} = 0$. The
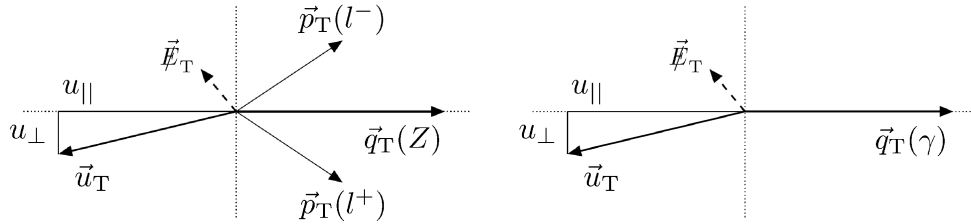
Figure 8-5. Illustration of the Z boson (left) and photon (right) event kinematics in the transverse plane. The vector $\vec{u}_\mathrm{T}$ denotes the vectorial sum of all particles reconstructed in the event except for the two leptons from the Z decay (left) or the photon (right).

hadronic recoil can be split into its parallel and perpendicular components with respect to the boson axis, and these quantities, $u_\parallel$ and $u_\perp$, are used to study the $p_\mathrm{T}^\mathrm{miss}$ response and resolution. The $p_\mathrm{T}^\mathrm{miss}$ response is defined as $-\langle u_\parallel \rangle / \langle q_\mathrm{T} \rangle$ where $\langle \rangle$ indicates the mean,

and reflects how well balanced the boson is to the hadronic recoil. The p_T^{miss} resolution is estimated through the RMS of the $u_\parallel + q_T$ and $u_\perp$ distributions, and are denoted by $\sigma(u_\parallel)$ and $\sigma(u_\perp)$, respectively.

8.5.2 Performance of PF p_T^{miss} algorithm

The p_T^{miss} in selections with no genuine p_T^{miss} from neutrinos are shown in Fig. 8-6, and show a good agreement between the data and the simulation. The lower pads of Fig. 8-6 show the magnitude of the uncertainties in these events, where the p_T^{miss} resolution is dominated by the resolution of the hadronic activity, ranging up to 10–15% for the jet momentum resolution [43]. Since the momentum resolution for leptons (photons) is σ_{p_T}/p_T of 1–4% (1–3%) [39, 48] is sub-dominant to the resolution of the hadronic actitvity, these uncertainties are not taken into account in the ratios, and the final uncertainty shown in the figures include uncertainties in the JES, the JER, and the UE, added in quadrature. The increase in the uncertainty band around $40\,\mathrm{GeV}$ is related to the JES and the JER sources in events with at least one jet and no genuine p_T^{miss}. For higher values of p_T^{miss}, where processes with genuine p_T^{miss}, e.g. top quark background, become more dominant, the uncertainty decreases. Distributions of $u_\parallel + q_T$ and $u_\perp$ in $Z \to \mu^+\mu^-$ $Z \to e^+e^-$ and γ+jets events are shown in Fig. 8-7. The kinematic definition of $u_\parallel$ dictates that for processes with no genuine p_T^{miss}, $u_\parallel$ is balanced with the boson q_T. Therefore, the vectorial sum of $u_\parallel$ and q_T results in a symmetric distribution, centered at zero; any deviations from this behavior imply imperfect calibration of p_T^{miss}. In events with genuine p_T^{miss}, due to the presence of the neutrinos, $u_\parallel$ and q_T are not balanced, leading to an asymmetric distribution as can be seen as the Top quark and Electroweak contribution in the distributions. Due to the assumed isotropic nature of the energy fluctuations of the detector noise and underlying event, the $u_\perp$ distribution is symmetric with a mean value of 0. Figure 8-8 shows the p_T^{miss} response as a function of q_T, in data and simulation, in $Z \to \mu^+\mu^-$ $Z \to e^+e^-$ and photon events. Two types of response are displayed. On
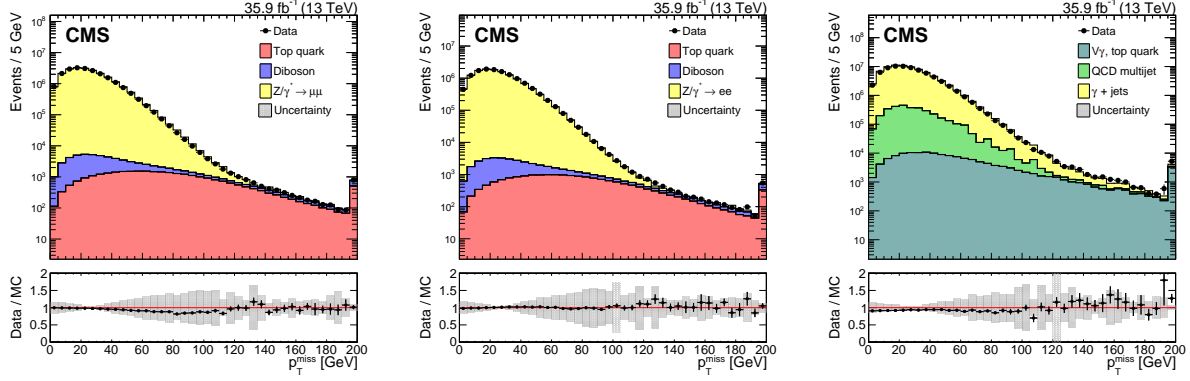
Figure 8-6. The p_T^{miss} for events passing the dimuon (left), dielectron (middle) and single photon (right) selections, in data (black markers) and simulation (solid histograms). The lower bands show the data to simulation ratio with the systematic uncertainties due to the JES, the JER, and variations in the UE are added in quadrature.
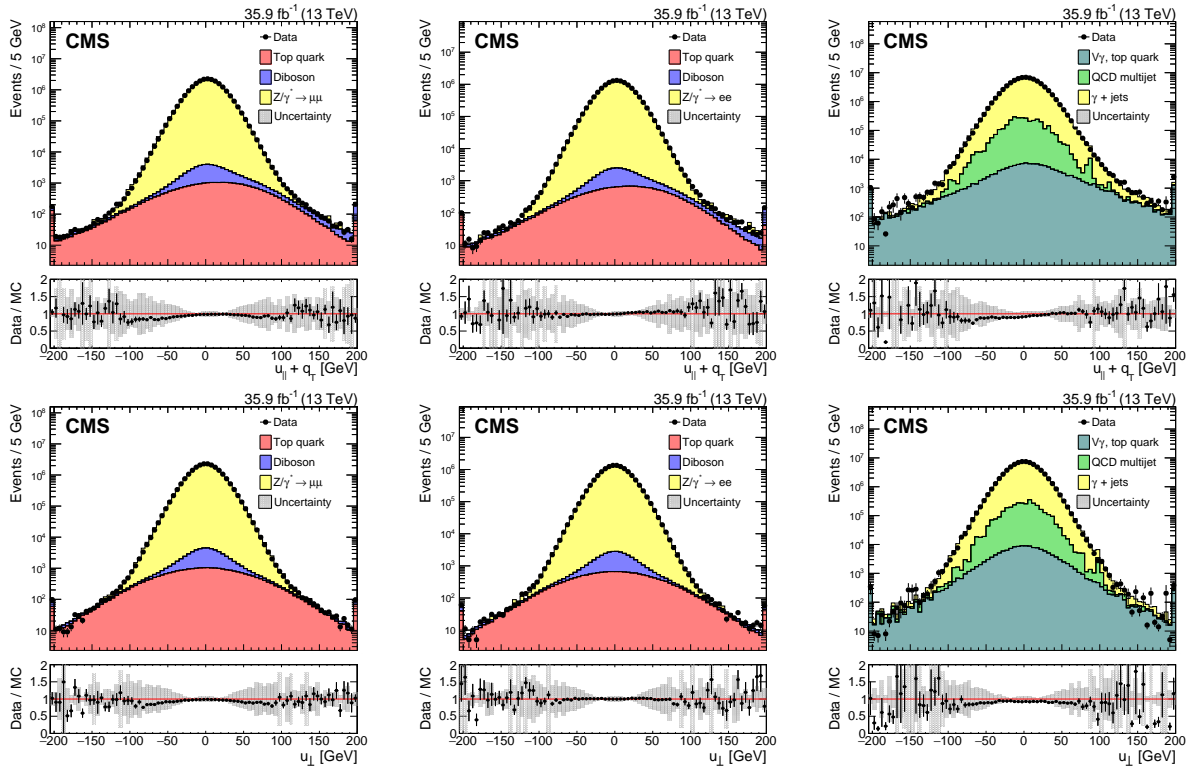
Figure 8-7. Distribution of $u_{\parallel}+q_T$ (upper) and $u_{\perp}$ (lower) components of the hadronic recoil, in data (filled markers) and simulation (solid histograms), in the $Z \to \mu^+\mu^-$ (left), $Z \to e^+e^-$ (middle), and γ+jets (right) samples.

the left is the response of the uncorrected p_T^{miss} (Raw p_T^{miss}), in which the JECs have not been propagated to the p_T^{miss}. The effect of correcting the p_T^{miss}, the so called "Type-1"

correction, is shown on the right, where the response is now much closer to unity, proving

that the corrrected jets results in a well balanced event. The underestimation of the

hadronic response observed at smaller $q_{\mathrm{T}} \leq 100\,\mathrm{GeV}$ is due to the significant contribution

of the uncalibrated component of $p_{\mathrm{T}}^{\mathrm{miss}}$, which mainly consists of jets with $p_{\mathrm{T}} < 15$ GeV

and unclustered particles. The response of $p_{\mathrm{T}}^{\mathrm{miss}}$ is found to agree between all three samples

within 2%. The residual response difference between the samples stems from the different

mechanism used to disambiguate muons, electrons, and photons from jets used in the

correction of the $p_{\mathrm{T}}^{\mathrm{miss}}$, as discussed in Section 8.2. Simulation studies has shown that

in the case of the electrons and photons, a small fraction ($\lesssim 10\%$) of jets survive the

disambiguation criteria yet overlap with prompt electrons and photons. As a result, these

jets wrongly contribute to the $p_{\mathrm{T}}^{\mathrm{miss}}$ calibration, leading to a 1–2% lower response in the

electron and photon channels. The uncorrected $p_{\mathrm{T}}^{\mathrm{miss}}$ response show a perfect agreement

between the electron and muon events, further proving that the disagreement in the

corrected response is due to the propagation of the JECS and the disambiguation criteria

of what jets to correct. The resolution of the $p_{\mathrm{T}}^{\mathrm{miss}}$ is evaluated as a function of the boson

p_{T} (q_{T}), the number of vertices and of the scalar p_{T} sum of all PF candidates ($\sum E_{\mathrm{T}}$).

The upper row of Fig. 8-9 shows the resolution as a function of q_{T}. In order to compare

the resolution of $p_{\mathrm{T}}^{\mathrm{miss}}$ consistently across the samples, the resolution in each sample is

corrected for the differences observed in the response, with a resulting negligible impact

on the results. The relative resolution both in $u_{\parallel}$ and $u_{\perp}$ is found to improve as a function

of q_{T} because of the improved energy resolution in the calorimeters. Furthermore, due

to the isotropic nature of energy fluctuations stemming from detector noise and the

underlying event, the dependence of the resolution of $u_{\perp}$ on q_{T} is smaller than for $u_{\parallel}$.For

$q_{\mathrm{T}} > 200\,\mathrm{GeV}$, the $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution is $\approx 13\%$ and $\approx 9\%$, for the $u_{\parallel}$ and $u_{\perp}$, respectively.

The middle row of Fig. 8-9 displays the resolution as a function of $\mathrm{N_{vtx}}$, showing a

significant dependence on $\mathrm{N_{vtx}}$, since only pileup mitigation techniques are employed to
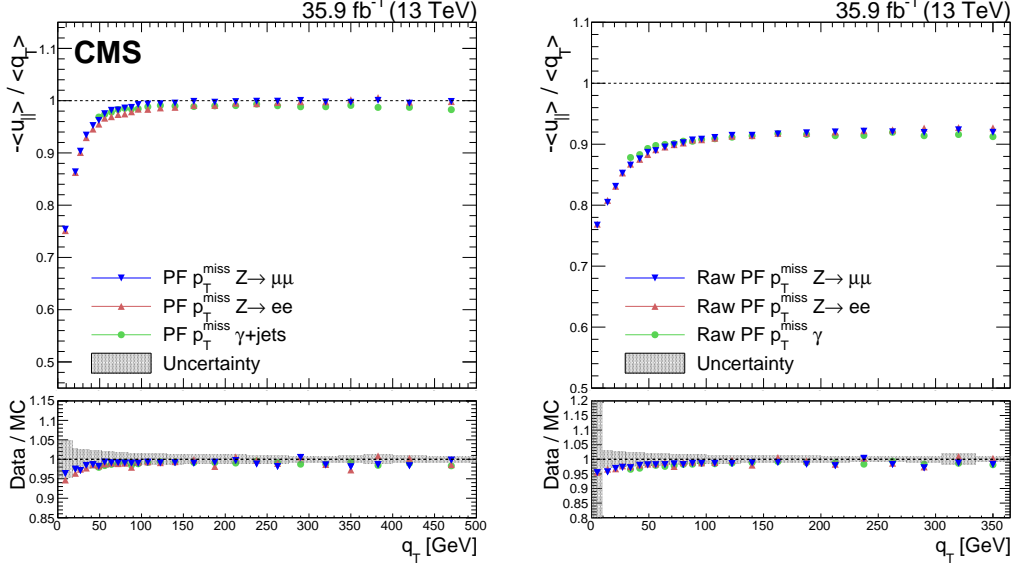
Figure 8-8. Upper panels: Response of the corrected (left) and uncorrected (right) p_T^{miss} in data in $Z \to \mu^+\mu^-$ $Z \to e^+e^-$ and γ+jets events. Lower panels: Ratio of the p_T^{miss} response in data and simulation. The band corresponds to the systematic uncertainties due to the JES, the JER, and variations in the UE added in quadrature, estimated from the $Z \to e^+e^-$ sample.

the PF jets and not the PF p_T^{miss} algorithm. The resolution is parametrized as:

$$f(N_{vtx}) = \sqrt{\sigma_c^2 + \frac{N_{vtx}}{0.70}\sigma_{PU}^2}, \qquad (8\text{--}12)$$

where σ_c is the resolution term induced by the hard scattering interaction and σ_{PU} is the resolution term induced on average by an additional pileup interaction. The factor 0.70 accounts for the vertex reconstruction efficiency [49]. Results of the parametrization for the $u_\parallel$ and $u_\perp$ components are given in Table 8-5. Each additional pileup vertex is found to degrade the resolution of each component by 3.8–4.0 GeV. Lastly, Fig. 8-9 (lower row) shows the resolution of $u_\parallel$ and $u_\perp$ as a function of the scalar p_T sum of all PF candidates ($\sum E_T$). The resolutions measured in different samples, and in data and simulation, are found to be in good agreement. The relative p_T^{miss} resolution improves with increasing $\sum E_T$, driven by the amount of the activity in the calorimeters. The resolution in different

samples is parametrized as:

$$\sigma_{u_\perp, u_\parallel} = \sigma_0 + \sigma_s \sqrt{\sum E_T}, \qquad\qquad (8\text{--}13)$$

where σ_0 is the resolution term induced by intrinsic detector noise and σ_s is the stochastic resolution term. Results of the parametrization for the $u_\parallel$ and $u_\perp$ components are given in Table 8-6.

Table 8-5. Parametrization results of the resolution curves for the $u_\parallel$ and $u_\perp$ components as a function of N_{vtx}. The parameter values for σ_c are obtained from data and simulation, and the values for σ_{PU} are obtained from data, along with a ratio R_{PU} of data and simulation. The uncertainties displayed for both components are obtained from the fit, and for simulation the JES, the JER, and UE uncertainties are added in quadrature.

Process	σ_c(data)[GeV]	σ_c(MC)[GeV]	σ_{PU}(data)[GeV]	$R_r = \sigma_{PU}$(data)$/\sigma_{PU}$(MC)
		$u_\parallel$ component		
$Z \to \mu^+\mu^-$	13.9 ± 0.07	11.9 ± 1.53	3.82 ± 0.01	0.95 ± 0.04
$Z \to e^+e^-$	14.6 ± 0.09	12.0 ± 1.09	3.80 ± 0.02	0.95 ± 0.03
γ+jets	12.2 ± 0.10	10.2 ± 1.98	3.97 ± 0.02	0.97 ± 0.05
		$u_\perp$ component		
$Z \to \mu^+\mu^-$	10.3 ± 0.08	8.58 ± 2.20	3.87 ± 0.01	0.97 ± 0.04
$Z \to e^+e^-$	10.7 ± 0.10	8.71 ± 1.76	3.89 ± 0.01	0.96 ± 0.03
γ+jets	9.04 ± 0.11	6.93 ± 2.70	3.94 ± 0.01	0.97 ± 0.04

8.5.3 Performance of PUPPI p_T^{miss} algorithm

The PUPPI p_T^{miss} distributions in the dilepton samples are shown in Fig. 8-10. The data distributions are modeled well by the simulation, in both the muon and the electron channels. Similar to the case of PF p_T^{miss}, the p_T^{miss} resolution in these events is dominated by the resolution of the hadronic activity, but the PUPPI-weighted PF candidates yield a much improved resolution for jets compared to the PF case. This is also reflected in the uncertainty shown in the figures, which includes the uncertainties due to jet energy scale and resolution, and the energy scale of the unclustered particles.
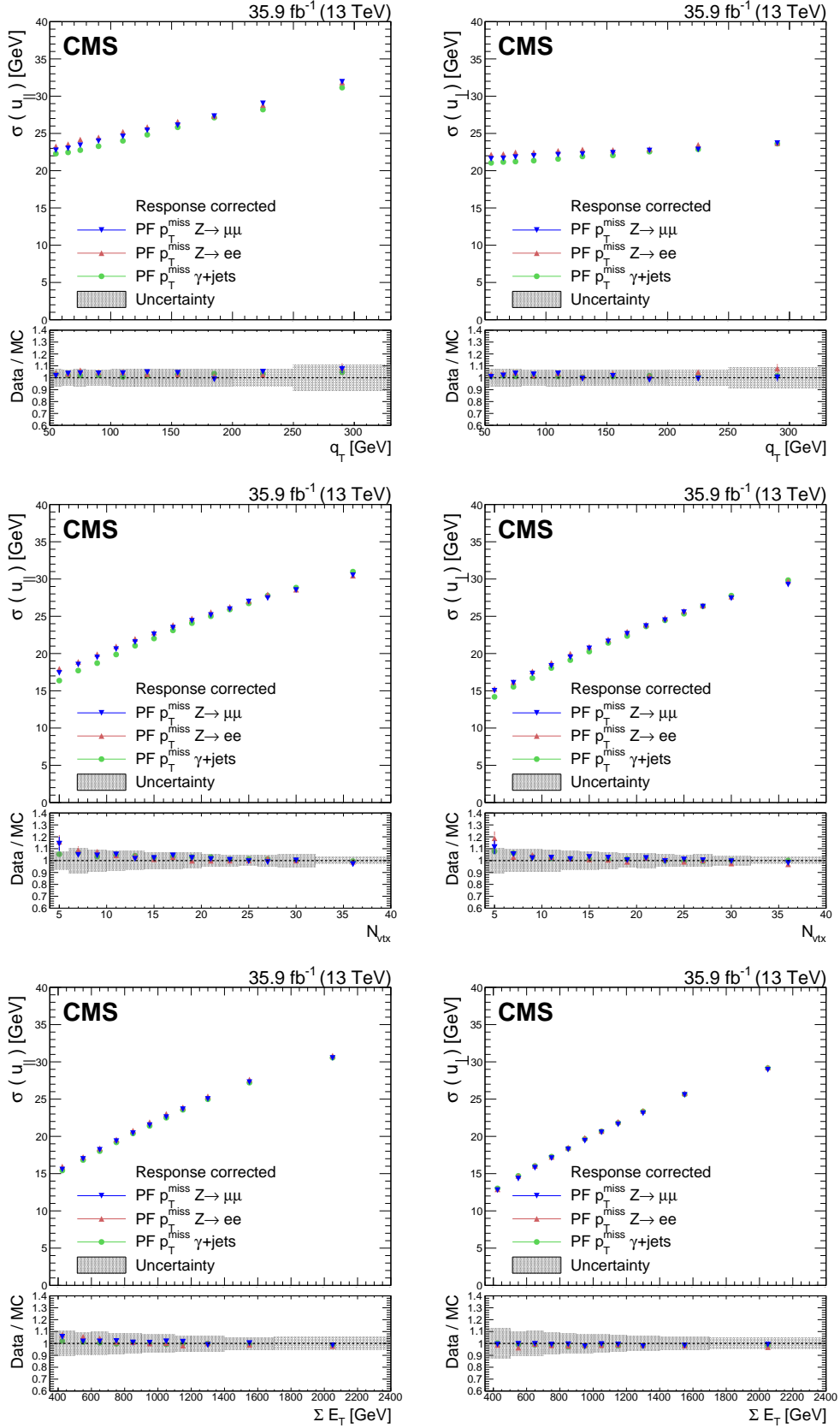
Figure 8-9. Resolution of the $u_\parallel$ and $u_\perp$ components of the hadronic recoil as a function of, q_T (upper row), the reconstructed vertices (middle row), and the scalar p_T sum of all PF candidates (lower row),104 $Z \to \mu^+\mu^-$ $Z \to e^+e^-$ and γ+jets events. In each plot, the upper panel shows the resolution in data, whereas the lower panel shows the ratio of data to simulation. The band corresponds to the
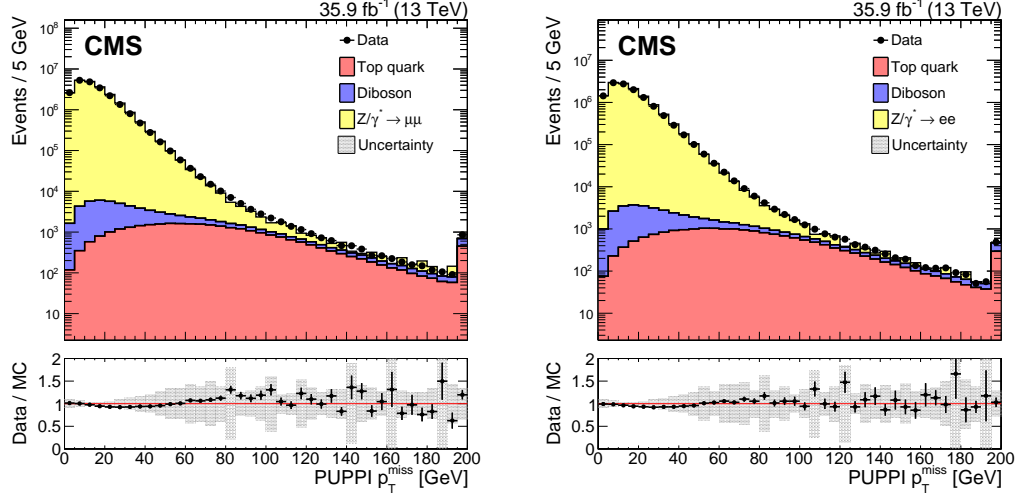
Figure 8-10. Upper panels: Distributions of PUPPI p_T^{miss} in $Z \to \mu^+\mu^-$ (left) and $Z \to e^+e^-$ (right) events. The last bin includes all events with $p_T^{miss} > 195$ GeV. Lower panels: Data-to-simulation ratio. The band corresponds to the systematic uncertainties due to the JES, the JER, and variations in the UE added in quadrature, estimated from the $Z \to e^+e^-$ sample.

The distributions in $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ events of the vectorial sum $u_\parallel + q_T$ and of $u_\perp$ using PUPPI p_T^{miss}, are shown in Fig. 8-11. Following the same arguments as in the PF p_T^{miss} case, in events with no genuine p_T^{miss}, the vectorial sum of $u_\parallel$ and q_T is

Table 8-6. Parametrization results of the resolution curves for $u_\parallel$ and $u_\perp$ components as a function of the scalar p_T sum of all PF candidates. The parameter values for σ_0 are obtained from data and simulation, whereas the σ_s are obtained from data along with the ratio R_s, the ratio of data and simulation. The uncertainties displayed for both components are obtained from the fit, and for simulation the JES, the JER, and UE uncertainties are added in quadrature.

Process	$\sigma_0(\text{data})[\text{GeV}]$	$\sigma_0(\text{MC})[\text{GeV}]$	$\sigma_s[\text{GeV}^{1/2}]$	$R_s = \sigma_s(\text{data})/\sigma_s(\text{MC})$
		$u_\parallel$ component		
$Z \to \mu^+\mu^-$	1.98 ± 0.07	0.85 ± 2.45	0.64 ± 0.01	0.95 ± 0.11
$Z \to e^+e^-$	2.18 ± 0.09	0.19 ± 2.90	0.64 ± 0.01	0.92 ± 0.11
γ+jets	1.85 ± 0.09	0.94 ± 2.52	0.64 ± 0.01	0.96 ± 0.11
		$u_\perp$ component		
$Z \to \mu^+\mu^-$	-1.63 ± 0.06	-1.72 ± 2.53	0.68 ± 0.01	0.99 ± 0.11
$Z \to e^+e^-$	-1.42 ± 0.08	-1.98 ± 2.95	0.69 ± 0.01	0.96 ± 0.12
γ+jets	-1.16 ± 0.08	-1.31 ± 2.53	0.68 ± 0.01	0.98 ± 0.11

symmetric around zero, whereas for processes with genuine p_T^{miss} an asymmetric behavior is observed. The distribution of $u_\perp$ is symmetric around zero. Simulation describes data well for all distributions.
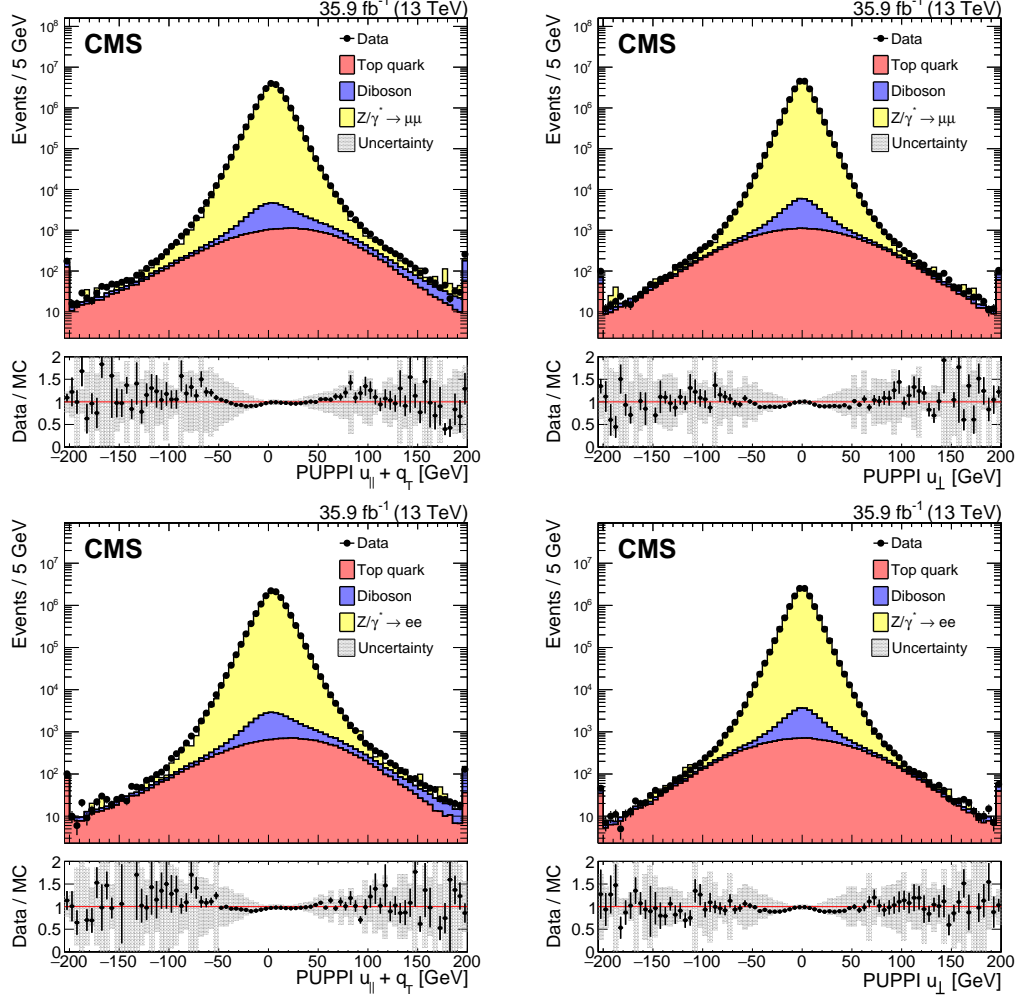
Figure 8-11. Upper panels: Distributions of the $u_\parallel + q_T$ and $u_\perp$ components of the hadronic recoil, in data (filled markers) and simulation (solid histograms), for the $Z \to \mu^+\mu^-$ (upper) and $Z \to e^+e^-$ (lower) events. The first and the last bins include all events below -195 and above +195, respectively. Lower panel: Data-to-simulation ratio. The band corresponds to the systematic uncertainties due to the JES, the JER, and variations in the UE added in quadrature, estimated from the $Z \to e^+e^-$ sample.

Figure 8-12 shows the PUPPI p_T^{miss} response as a function of q_T, extracted from data and simulation in $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ events. The response reaches unity for $Z \to \mu^+\mu^-$ events at a boson p_T of 150 GeV; while for PF p_T^{miss} the response is close to

unity at 100 GeV. The slower rise of the response to unity is due to the removal of PF candidates that are wrongly associated with pileup interactions by the PUPPI algorithm. Similarly to PF $p_\mathrm{T}^\mathrm{miss}$, there is no response correction for the UE for PUPPI $p_\mathrm{T}^\mathrm{miss}$, which results in an underestimated response for low q_T. The response of $p_\mathrm{T}^\mathrm{miss}$ is found to agree between the different samples within 2%.
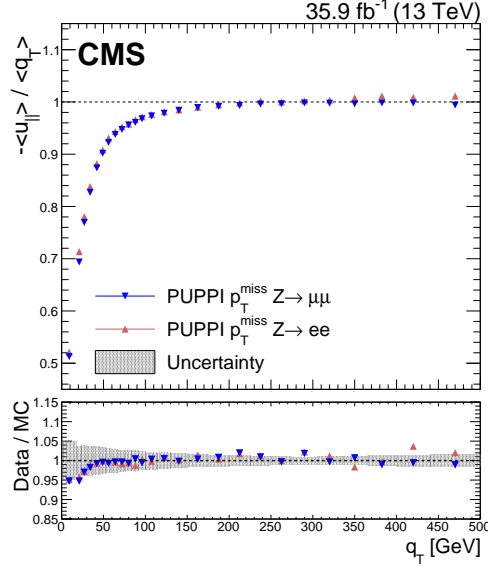
Figure 8-12. Upper panel: Response of PUPPI $p_\mathrm{T}^\mathrm{miss}$, defined as $-\langle u_\parallel\rangle/\langle q_\mathrm{T}\rangle$, in data in $Z \to \mu^+\mu^-$ and $Z \to \mathrm{e}^+\mathrm{e}^-$ events. Lower panel: ratio of the PUPPI $p_\mathrm{T}^\mathrm{miss}$ response in data and simulation. The band corresponds to the systematic uncertainties due to the JES, the JER, and variations in the UE added in quadrature, estimated from the $Z \to \mathrm{e}^+\mathrm{e}^-$ sample.

In Fig. 8-13, the results obtained for the case of PUPPI $p_\mathrm{T}^\mathrm{miss}$ are overlayed with the ones obtained using PF $p_\mathrm{T}^\mathrm{miss}$. Compared to the case of PF $p_\mathrm{T}^\mathrm{miss}$, the resolutions show a much reduced dependence on the number of pileup interactions. The resolution in different samples is parametrized using Eq. (8–12), and the result of the parameterization are given in Table 8-7. Each additional pileup interaction is found to degrade the resolution of each component by up to 2 GeV. This resolution degradation corresponds to half of what is observed in the case of PF $p_\mathrm{T}^\mathrm{miss}$.
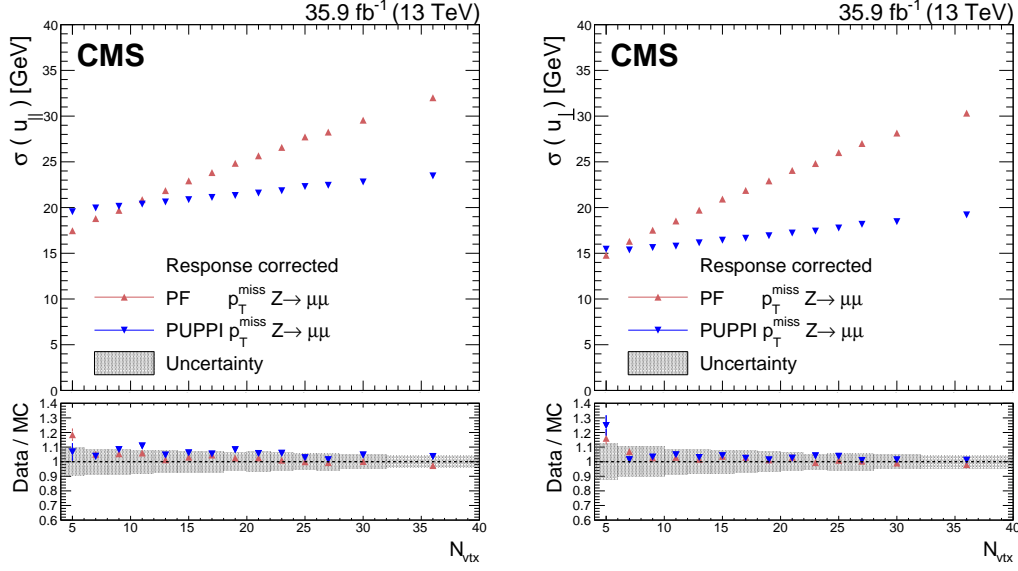
Figure 8-13. Upper panels: PUPPI and PF p_T^{miss} resolution of $u_\parallel$ (left) and $u_\perp$ (right) components of the hadronic recoil as a function of N_{vtx}, in $Z \to \mu^+\mu^-$ events. Lower panels: Data-to-simulation ratio. The systematic uncertainties due to the JES, the JER, and variations in the UE are added in quadrature and displayed with a band.

Table 8-7. Parameterization results of the resolution curves for PUPPI $u_\parallel$ and $u_\perp$ components as a function of N_{vtx}. The parameter values for σ_c are obtained from data and simulation, and the values for σ_{PU} are obtained from data, along with the ratio R_{PU} of data and simulation. The uncertainties displayed for both the components are obtained from the fit, and for simulation the JES, the JER,

Process	σ_c(data)[GeV]	σ_c(MC)[GeV]	σ_{PU}(data)[GeV]	$R_{PU} = \sigma_{PU}$(data)/σ_{PU}(MC)
$u_\parallel$ component				
$Z \to \mu^+\mu^-$	18.9 ± 0.05	17.5 ± 0.74	1.93 ± 0.02	0.97 ± 0.11
$Z \to e^+e^-$	18.9 ± 0.06	17.4 ± 0.80	1.94 ± 0.03	0.98 ± 0.12
$u_\perp$ component				
$Z \to \mu^+\mu^-$	14.2 ± 0.04	13.6 ± 0.59	1.78 ± 0.01	0.97 ± 0.09
$Z \to e^+e^-$	14.3 ± 0.05	13.6 ± 0.59	1.80 ± 0.02	0.96 ± 0.09

8.5.4 High pileup studies
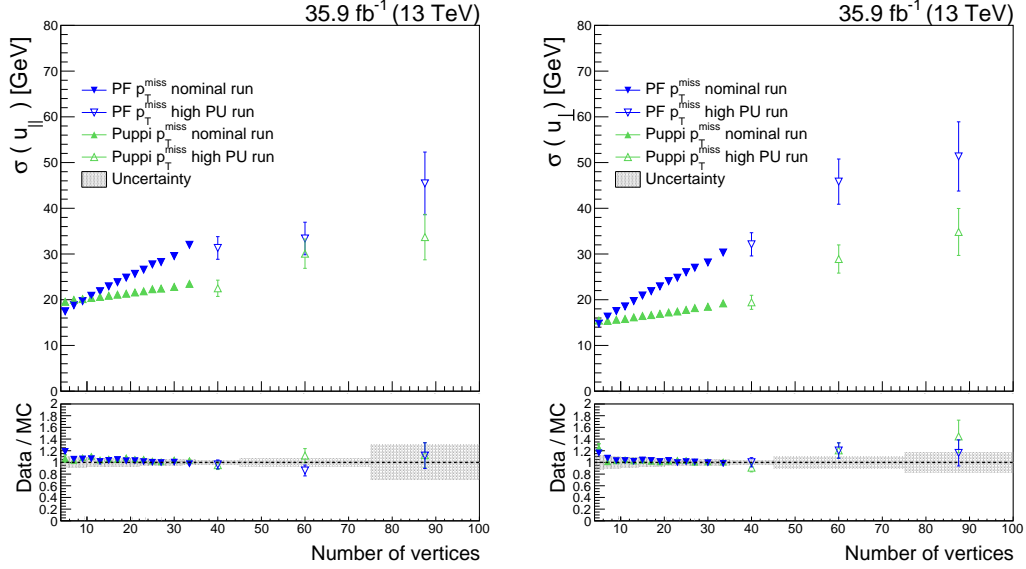
Figure 8-14. Upper panels: PUPPI and PF $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution of $u_{\parallel}$ (left) and $u_{\perp}$ (right) components of the hadronic recoil as a function of N_{vtx}, in $Z \to \mu^{+}\mu^{-}$ events. The blue (green) markers correspond to the PF (PUPPI) $p_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction algorithm, with filled (open) markers for the nominal run (high pileup run). Lower panels: Data-to-simulation ratio. The systematic uncertainties due to the JES, the JER, and variations in the UE are added in quadrature and displayed with a band.

By 2025, the LHC will be upgraded with a goal to increase the integrated luminosity by a factor of 10 beyond the original design. The new design, known as High Luminosity LHC (HL-LHC), will pose a major challenge due to the unprecedented increase of pileup expected (sim140) in the collision events per bunch crossing. In preparation, LHC delivered a special pp collisions data with conditions similar to the ones expected at the HL-LHC. The "high pileup" data set, as it will be referred in what follows, corresponds to an integrated luminosity of 5pb^{-1}. The bunch setup for the data was three isolated bunches of an average pileup of 70-100, and 2×48 bunch trains, corresponding to an average pileup of 35-50. Dedicated simulated samples were produced with similar conditions. Using the high pileup data set and the dedicated MC simulations, the performance of the PF $p_{\mathrm{T}}^{\mathrm{miss}}$ and Puppi $p_{\mathrm{T}}^{\mathrm{miss}}$ algorithms are studied in dimuon samples.

The $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution of the $u_\parallel$ and $u_\perp$ components of the hadronic recoil as a function of the number of reconstructed vertices is shown in Fig. 8-14. The results obtained from the high pileup data is overlayed with the ones obtained from the nominal data and found to be in agreement within the statistical uncertainties. Furthermore, the Puppi $p_{\mathrm{T}}^{\mathrm{miss}}$ is found to have more stable resolution across the full range.

APPENDIX A
Simulated samples used for SUSY searches

Several MC event generators are used to simulate the background and signal processes in this analysis, with the different parts of the generators introduced in Section 4.3. The simulation is normalized to luminosity using cross sections from `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns`. The `PYTHIA8` [44] package is used for parton showering, hadronization and undelying event simulation with the tune `CUETP8M1`, as described in Section 4.3. In Tables A-8, A-9, A-10 and A-11, the various MC simulations used for the different background prediction methods are listed.

A.0.5   Higher order corrections

For the $WZ \rightarrow l^+ l^- l\nu$ and $WZ \rightarrow l^+ l^- 2q$ processes, a NLO to NNLO correction factor of 1.109 is applied [?]. For the qq→ZZ process, a QCD NLO to NNLO correction factor is applied as a function of generator-level p_{T} of the diboson system, which is described more in depth in Section ??.

Table A-8. Simulated SM datasets used for the flavor symmetric (FS) background prediction. All samples are of the `MINIAOD` data format and of the version `RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/` and `Tune` is short for the pythia8 tune `CUETP8M1`.

Process	Dataset	σ (pb)
ttbar		
$t\bar{t} \to l^+\nu b + l^-\bar{\nu}\bar{b}$	/TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	$831.76{\times}0.1086^2{\times}9$
$t\bar{t} \to l^-\bar{\nu}{+}\mathrm{jets}$	/TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.2
$t\bar{t} \to l^+\nu{+}\mathrm{jets}$	/TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.2
Single Top		
$W^+ \to t\bar{b}$	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8	3.36
$\bar{b} \to \bar{t}W^+$	/ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg	11.7
$b \to tW^-$	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg	11.7
$q\bar{b} \to q'\bar{t}$	/ST_t-channel_antitop_4f_inclDecays_13TeV-powhegV2-madspin-pythia8	124.0
$qb \to q't$	/ST_t-channel_top_4f_inclDecays_13TeV-powhegV2-madspin-pythia8	208.0
Diboson (FS)		
$WW \to l^+\nu l^-\bar{\nu}$	/WWTo2L2Nu_13TeV-powheg-pythia8	$(118.7\text{-}3.974){\times}0.1086^2{\times}9$
$gg \to WW \to l^+\nu l^-\bar{\nu}$	/GluGluWWTo2L2Nu_MCFM_13TeV	$(3.974{\times}0.1086^2{\times}9{\times}1.4$
$gg \to H \to WW$	/GluGluHToWWTo2L2Nu_M125_13TeV_powheg_JHUgen_pythia8	1.002
WW	/WW_DoubleScattering_13TeV-pythia8	1.617
WW	/WpWpJJ_EWK-QCD_TuneCUETP8M1_13TeV-madgraph-pythia8	0.037
$q\bar{q} \to l\nu\gamma$	/WGToLNuG_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	405.3
Triboson (FS)		
WWW	/WWW_Tune_13TeV-amcatnlo-pythia8	0.209
WWγ	/WWG_Tune_13TeV-amcatnlo-pythia8	0.215
Rare (FS)		
$t\bar{t}W$	/TTWJetsToLNu_Tune_13TeV-amcatnlo-madspin-pythia8	0.204
$t\bar{t}W$	/TTWJetsToQQ_Tune_13TeV-amcatnloFXFX-madspin-pythia8	0.406
$t\bar{t}H$	/ttHToNonbb_M125_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	0.215
VH	/VHToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.952
$tttt$	/TTTT_TuneCUETP8M2T4_13TeV-amcatnlo-pythia8	0.009
W+jets	/WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	61527

Table A-9. Simulated SM datasets used for the ZZ to 4 lepton control regions. All samples are of the `MINIAOD` data format and of the version `RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/` and `Tune` is short for the pythia8 tune `CUETP8M1`. The k-factor referred to is specified in Subsection A.0.5

Process	Dataset	σ (pb)
$ZZ \rightarrow 4l$		
$ZZ \rightarrow 4l$	/ZZTo4L_13TeV_powheg_pythia8	$1.256\times$ k-factor
$gg \rightarrow H \rightarrow ZZ$	/GluGluHToZZTo4L_M12513TeV_powheg2_JHUgenV6_pythia8	0.013
$qq \rightarrow H \rightarrow ZZ$	/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUgenV6_pythia8	0.001
$gg \rightarrow ZZ \rightarrow 4e$	/GluGluToContinToZZTo4e_13TeV_MCFM701_pythia8	0.001586×2.3
$gg \rightarrow ZZ \rightarrow 4\mu$	/GluGluToContinToZZTo4mu_13TeV_MCFM701_pythia8	0.001586×2.3
$gg \rightarrow ZZ \rightarrow 4\tau$	/GluGluToContinToZZTo4tau_13TeV_MCFM701_pythia8	0.001586×2.3
$gg \rightarrow ZZ \rightarrow 2e2\tau$	/GluGluToContinToZZTo2e2tau_13TeV_MCFM701_pythia8	0.003194×2.3
$gg \rightarrow ZZ \rightarrow 2e2\mu$	/GluGluToContinToZZTo2e2mu_13TeV_MCFM701_pythia8	0.003194×2.3
$gg \rightarrow ZZ \rightarrow 2\mu2\tau$	/GluGluToContinToZZTo2mu2tau_13TeV_MCFM701_pythia8	0.003194×2.3
$ZZ \rightarrow 2l2\nu$		
$ZZ \rightarrow 2l2\nu$	/ZZTo2L2Nu_13TeV_powheg_pythia8	$0.564\times$ k-factor
$gg \rightarrow ZZ \rightarrow 2e2\nu$	/GluGluToContinToZZTo2e2nu_13TeV_MCFM701_pythia8	0.001720×2.3
$gg \rightarrow ZZ \rightarrow 2\mu2\nu$	/GluGluToContinToZZTo2mu2nu_13TeV_MCFM701_pythia8	0.001720×2.3
Others		
$ZZ \rightarrow 2l2q$	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.28
ZZZ	/ZZZ_Tune_13TeV-amcatnlo-pythia8	0.0139
VH	/VHToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.952
$q\bar{q} \rightarrow l^{+}l^{-}\gamma$	/ZGTo2LG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	123.9

Table A-10. Simulated SM datasets used for the WZ control regions. All samples are of the `MINIAOD` data format and of the version `RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/` and `Tune` is short for the pythia8 tune `CUETP8M1`.

Process	Dataset	σ (pb)
$WZ \rightarrow l^{+}l^{-}l\nu$	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.429×1.109
$WZ \rightarrow l^{+}l^{-}2q$	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595×1.109

Table A-11. Various non flavor symmetric processes. All samples are of the `MINIAOD` data format and of the version `RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6/` and `Tune` is short for the pythia8 tune `CUETP8M1`.

| Process	Dataset	σ (pb)
Drell-Yan		
$Z/\gamma^* \to l^+l^- (50)$	/DYJetsToLL_M-50_Tune_13TeV-madgraphMLM-pythia8	1921.8×3
$Z/\gamma^* \to l^+l^- (10-50)$	/DYJetsToLL_M-10to50_Tune_13TeV-madgraphMLM-pythia8	18610
Various non FS		
$t\bar{t}Z$	/TTZToLL_M-1to10_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.049
$t\bar{t}Z$	/TTZToLLNuNu_M-10_Tune_13TeV-amcatnlo-pythia8	0.253
$t\bar{t}Z$	/TTZToQQ_Tune_13TeV-amcatnlo-pythia8	0.530
tZq	/tZq_ll_4f_13TeV-amcatnlo-pythia8	0.076
tWZ	/ST_tWll_5f_LO_13TeV-MadGraph-pythia8	0.011
WWZ	/WWZ_Tune_13TeV-amcatnlo-pythia8	0.165
WZZ	/WZZ_Tune_13TeV-amcatnlo-pythia8	0.056
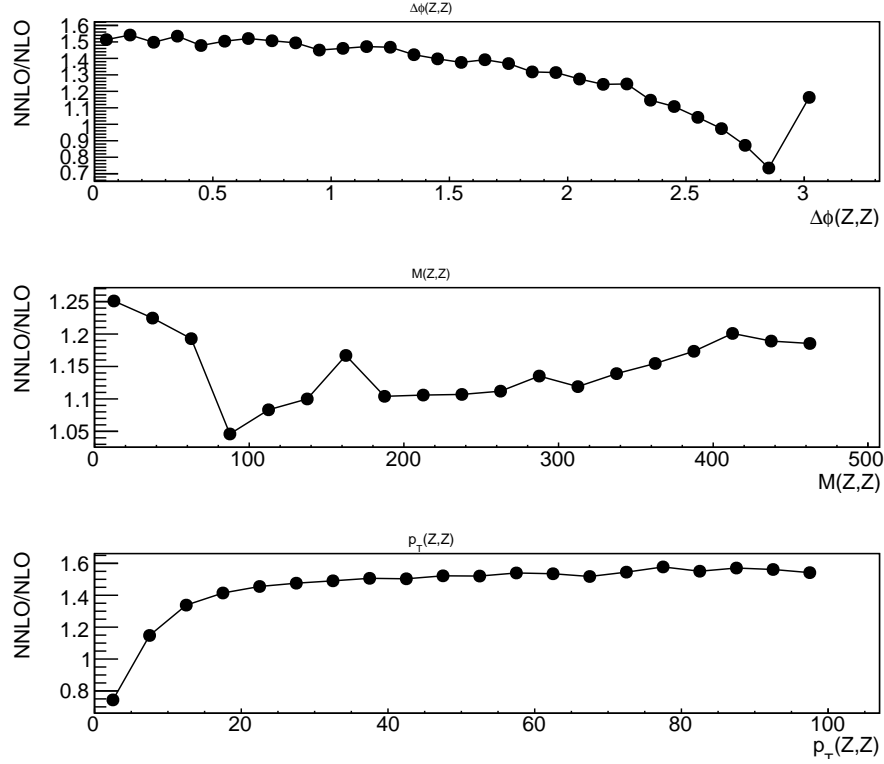
Figure A-15. QCD NNLO/NLO k factors for the qq→ZZ process in generator level variables of the diboson system.

APPENDIX B
Simulated samples for $p_\mathrm{T}^\mathrm{miss}$ performance study

Table A-12. Simulated SM datasets used for the $p_\mathrm{T}^\mathrm{miss}$ performance study with final states containing two leptons.

Process	Dataset	σ (pb)
DY		
Drell--Yan	/DYJetsToLL_M-50_Tune_13TeV-amcatnloFXFX	6025.2
Top		
$t\bar{t} \to l^+\nu b + l^-\bar{\nu}\bar{b}$	/TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	$831.76 \times 0.1086^2 \times 9$
$W^+ \to t\bar{b}$	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8	3.36
$\bar{b} \to \bar{t}W^+$	/ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powheg	80.95
$qb \to q't$	/ST_t-channel_top_4f_inclusiveDecays_13TeV-powheg	136.02
$b \to tW^-$	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg	11.7
EWK		
$ZZ \to 2l2q$	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.28
$ZZ \to 2l2\nu$	/ZZTo2L2Nu_13TeV_powheg_pythia8	0.564
$ZZ \to 4l$	/ZZTo4L_13TeV_powheg_pythia8	1.256
$WZ \to l^+l^-l\nu$	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.429×1.109
$WZ \to l^+l^-2q$	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595×1.109
$WW \to l^+\nu l^-\bar{\nu}$	/WWTo2L2Nu_13TeV-powheg-pythia8	$(118.7\text{-}3.974)\times0.1086^2\times9$
WWW	/WWW_Tune_13TeV-amcatnlo-pythia8	0.209
WWZ	/WWZ_Tune_13TeV-amcatnlo-pythia8	0.165
WZZ	/WZZ_Tune_13TeV-amcatnlo-pythia8	0.055
ZZZ	/ZZZ_Tune_13TeV-amcatnlo-pythia8	0.014

Table A-13. Simulated SM datasets used for the $p_{\mathrm{T}}^{\mathrm{miss}}$ performance study with final states containing two leptons.

Process	Dataset	σ (pb)
γ+jets		
γ+jets	/GJets_HT-40To100_Tune_13TeV_madgraphMLM	20730
γ+jets	/GJets_HT-100To200_Tune_13TeV_madgraphMLM	9226
γ+jets	/GJets_HT-200To400_Tune_13TeV_madgraphMLM	2300
γ+jets	/GJets_HT-400To600_Tune_13TeV_madgraphMLM	277
γ+jets	/GJets_HT-600ToInf_Tune_13TeV_madgraphMLM	93.4
QCD		
QCD	/QCD_HT-100to200_Tune_13TeV_madgraphMLM	2.799×10^{7}
QCD	/QCD_HT-200to300_Tune_13TeV_madgraphMLM	17.35×10^{5}
QCD	/QCD_HT-300to500_Tune_13TeV_madgraphMLM	3.668×10^{5}
QCD	/QCD_HT-500to700_Tune_13TeV_madgraphMLM	2.937×10^{4}
QCD	/QCD_HT-700to1000_Tune_13TeV_madgraphMLM	6524
QCD	/QCD_HT-1000to1500_Tune_13TeV_madgraphMLM	1064
QCD	/QCD_HT-1500to2000_Tune_13TeV_madgraphMLM	119.9
QCD	/QCD_HT-2000toInf_Tune_13TeV_madgraphMLM	25.2
Vγ+top quark		
W+jets	/WJetsToLNu_HT-100to200_Tune_13TeV_madgraphMLM	1345×1.21
W+jets	/WJetsToLNu_HT-200to400_Tune_13TeV_madgraphMLM	359.7×1.21
W+jets	/WJetsToLNu_HT-400to600_Tune_13TeV_madgraphMLM	48.9×1.21
W+jets	/WJetsToLNu_HT-600to800_Tune_13TeV_madgraphMLM	12.1×1.21
W+jets	/WJetsToLNu_HT-800to1200_Tune_13TeV_madgraphMLM	5.50×1.21
W+jets	/WJetsToLNu_HT-1200to2500_Tune_13TeV_madgraphMLM	1.33×1.21
W+jets	/WJetsToLNu_HT-2500toInf_Tune_13TeV_madgraphMLM	0.03×1.21
Wγ	/WGToLNuG_13TeV-amcatnloFXFX	489
Zγ	/ZGTo2LG_13TeV-amcatnloFXFX	117.9
Zγ	/ZGTo2NuG_13TeV-amcatnloFXFX	28.1
top+γ	/TTGJets_Tune_13TeV_amcatnloFXFX_madspin	4621
top+γ	/TGJets_Tune_13TeV_amcatnlo_madspin	2.97

REFERENCES

[1] O. S. Brning, *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs (CERN, Geneva, 2004). Available from: `https://cds.cern.ch/record/782076`.

[2] S. Chatrchyan, *et al.*, *JINST* **5**, T03001 (2010).

[3] M. Schott, M. Dunford, *Eur. Phys. J.* **C74**, 2916 (2014).

[4] CMS Luminosity Measurements for the 2016 Data Taking Period, *Tech. Rep. CMS-PAS-LUM-17-001*, CERN, Geneva (2017). Available from: `http://cds.cern.ch/record/2257069`.

[5] S. Chatrchyan, *et al.*, *JINST* **9**, P10009 (2014).

[6] S. Chatrchyan, *et al.*, *JINST* **8**, P09009 (2013). [JINST8,9009(2013)].

[7] S. Abdullin, *et al.*, *Eur. Phys. J.* **C55**, 159 (2008).

[8] S. Chatrchyan, *et al.*, *JINST* **7**, P10002 (2012).

[9] R. D. Ball, *et al.*, *Eur. Phys. J.* **C77**, 663 (2017).

[10] A. M. Sirunyan, *et al.*, *JINST* **13**, P05011 (2018).

[11] Technical proposal: L3, *Tech. Rep. CERN-LEPC-83-5. LEPC-P-4* (1983). Available from: `https://cds.cern.ch/record/297266`.

[12] S. Chatrchyan, *et al.*, *JINST* **3**, S08004 (2008).

[13] W. W. Armstrong, *et al.* (1994).

[14] (1995).

[15] *LHCb : Technical Proposal*, Tech. Proposal (CERN, Geneva, 1998). Available from: `https://cds.cern.ch/record/622031`.

[16] D. Boussard, T. Linnecar, *Adv. Cryog. Eng.* **45A**, 835 (2000).

[17] *The CMS magnet project: Technical Design Report*, Technical Design Report CMS (CERN, Geneva, 1997). Available from: `https://cds.cern.ch/record/331056`.

[18] M. D'Alfonso, *et al.* (2009).

[19] D. J. A. Cockerill, *Proceedings, 34th International Conference on High Energy Physics (ICHEP 2008): Philadelphia, Pennsylvania, July 30-August 5, 2008* (2008).

[20] *The CMS hadron calorimeter project: Technical Design Report*, Technical Design Report CMS (CERN, Geneva, 1997). The following files are from ¡a href=. Available from: `https://cds.cern.ch/record/357153`.

[21] V. Khachatryan, *et al.*, *JINST* **12**, P01020 (2017).

[22] R. D. Ball, *et al.*, *JHEP* **04**, 040 (2015).

[23] J. Alwall, *et al.*, *JHEP* **07**, 079 (2014).

[24] M. L. Mangano, M. Moretti, F. Piccinini, M. Treccani, *JHEP* **01**, 013 (2007).

[25] C. Oleari, *Nucl. Phys. Proc. Suppl.* **205-206**, 36 (2010).

[26] S. Alioli, P. Nason, C. Oleari, E. Re, *JHEP* **09**, 111 (2009). [Erratum: *JHEP* **02** (2010) 011].

[27] T. Sjstrand, *et al.*, *Comput. Phys. Commun.* **191**, 159 (2015).

[28] B. Andersson, G. Gustafson, G. Ingelman, T. Sjostrand, *Phys. Rept.* **97**, 31 (1983).

[29] A. M. Sirunyan, *et al.*, *JHEP* **03**, 076 (2018).

[30] A. M. Sirunyan, *et al.*, *Phys. Lett.* **B790**, 140 (2019).

[31] K. T. Matchev, S. D. Thomas, *Phys. Rev. D* **62**, 077702 (2000).

[32] P. Meade, M. Reece, D. Shih, *JHEP* **05**, 105 (2010).

[33] J. T. Ruderman, D. Shih, *JHEP* **08**, 159 (2012).

[34] P. Z. Skands, *et al.*, *JHEP* **07**, 036 (2004).

[35] A. M. Sirunyan, *et al.*, *JINST* **12**, P10003 (2017).

[36] Search for supersymmetry in events with tau leptons and missing transverse momentum in proton-proton collisions at sqrt(s)=13 TeV, *Tech. Rep. CMS-PAS-SUS-17-002*, CERN, Geneva (2017). Available from: `https://cds.cern.ch/record/2297162`.

[37] A. M. Sirunyan, *et al.*, *JHEP* **11**, 151 (2018).

[38] V. Khachatryan, *et al.*, *JINST* **10**, P06005 (2015).

[39] A. M. Sirunyan, *et al.*, *JINST* **13**, P06015 (2018).

[40] M. Cacciari, G. P. Salam, G. Soyez, *JHEP* **04**, 063 (2008).

[41] G. P. Salam, *Eur. Phys. J.* **C67**, 637 (2010).

[42] Y. L. Dokshitzer, G. D. Leder, S. Moretti, B. R. Webber, *JHEP* **08**, 001 (1997).

[43] V. Khachatryan, *et al.*, *JINST* **12**, P02014 (2017).

[44] T. Sjöstrand, S. Mrenna, P. Skands, *JHEP* **05**, 026 (2006).

[45] C. G. Lester, D. J. Summers, *Phys. Lett.* **B463**, 99 (1999).

[46] A. Barr, C. Lester, P. Stephens, *Journal of Physics G: Nuclear and Particle Physics* **29**, 2343 (2003). Available from: `https://doi.org/10.1088%2F0954-3899%2F29%2F10%2F304`.

[47] D. Bertolini, P. Harris, M. Low, N. Tran, *JHEP* **10**, 059 (2014).

[48] V. Khachatryan, *et al.*, *JINST* **10**, P08010 (2015).

[49] (2017). Available from: `https://cds.cern.ch/record/2290524`.