University of Cyprus
Department of Physics

Calculation of the electron charge sign mismeasurement probability at the CMS detector using $Z^0 \to e^- e^+$ decays

Anastasia Konstantinou
Supervisor: Assistant Professor Halil Saka

A thesis submitted in partial fulfilment of the requirements for the Bachelor of Science Degree in Physics

May 2025

ABSTRACT

Accurate measurements of particles produced during proton-proton collisions at the LHC experiment at CERN are crucial for understanding the processes that took place. However, mismeasurements can be reduced but not eliminated. Specifically, at the CMS detector, there is a possibility for charge sign misidentification for electrons, primarily due to energy loss from Bremsstrahlung radiation.

While in the general case, it cannot be certain whether a charge flip has occurred, this thesis aims to calculate the probability of such an event happening, using Z^0 boson decays into dielectron pairs. Since Z^0 bosons can only decay into oppositely charged electrons, when a decay into two same sign electrons is measured, it is a clear sign of charge misidentification. By developing methods to distinguish events that originated from Z^0 boson decays from events that came from other background processes, the ratio of the mismeasured to the correctly measured events can be extracted. This is achieved by fitting appropriate functions that describe both the Z^0 and the background events to the data. By further analysis, the statistical probability of such mismeasurements, as well as statistical and systematic uncertainties, can be calculated. The uncertainties are calculated by taking into consideration the errors of the fitting process and the choice of the function that describes the background events.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Assistant Professor Halil Saka, for his constant guidance and help throughout this project.

I would also like to thank my family for their support and encouragement during all these years. Their belief in me has been truly inspirational.

Contents

1 Introduction

1.1 The LHC

CERN houses the Large Hadron Collider (LHC), the largest particle accelerator globally. It is a 27-kilometer ring that accelerates beams of protons to speeds approaching the speed of light, which are then made to collide. This is possible due to superconducting electromagnets, which are responsible for bending the beams along the ring and focusing them, since the beams tend to spread due to the protons' mutual repulsion. In addition, superconducting electromagnets compress the beams even further just before the collision, to increase the chance of particle collisions. Along the accelerator ring are placed four particle detectors: ATLAS, CMS, ALICE, and LHCb. The particle beams are guided into a collision inside the detector that corresponds to the ongoing experiment. [1]

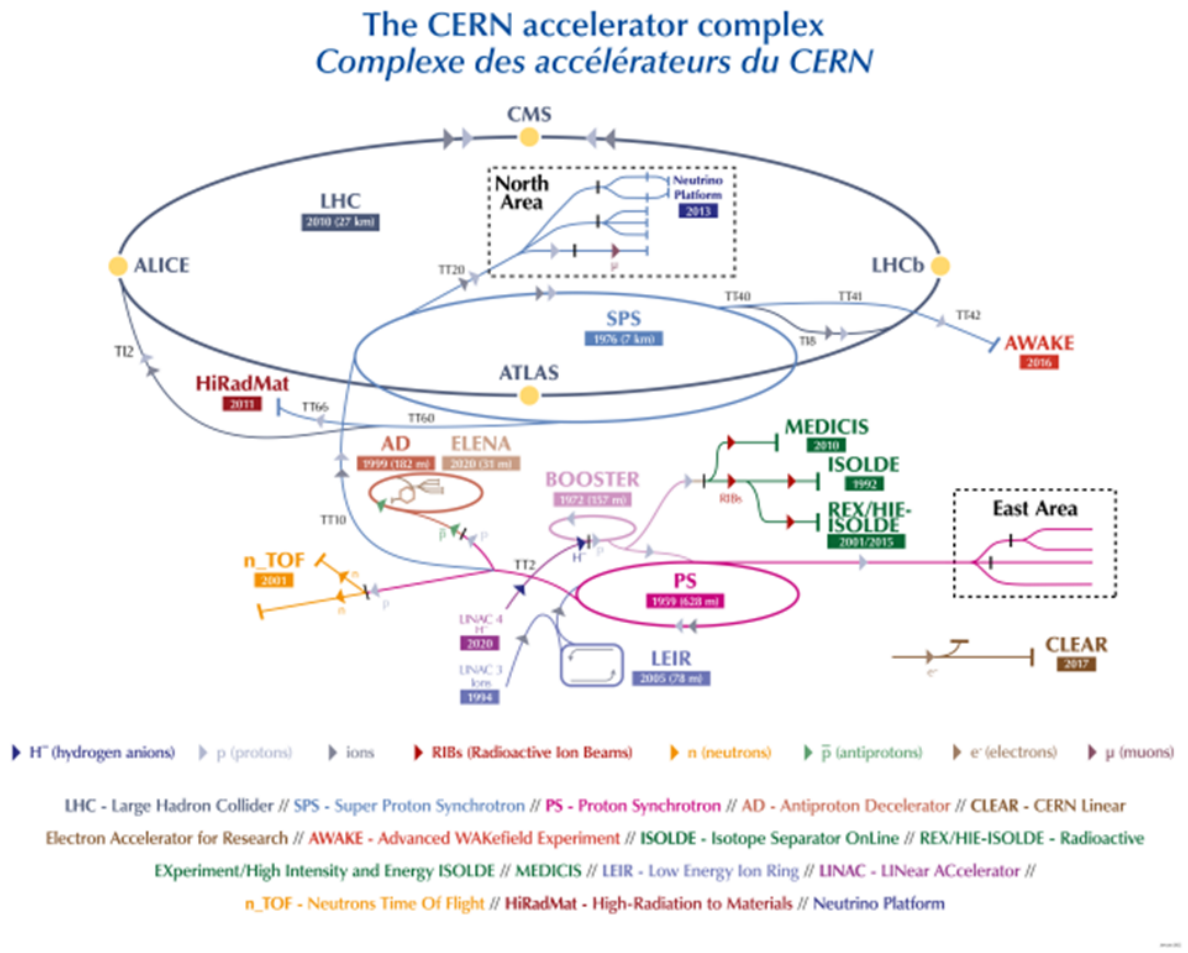

Figure 1: The Large Hadron Collider accelerator chain [2]

Each proton can be accelerated up to 7 TeV in the LHC. The centre-of-mass energy can reach up to 14 TeV. In the data analysed later in this thesis, the protons were accelerated to 6.5 TeV. Since they have opposite momenta, the centre-of-mass frame of reference of the colliding protons is equivalent to the laboratory frame of reference.[3] When those high-energy particles collide, they release energy, which is then converted into matter in the form of unstable heavy particles that instantly decay into lighter particles that can be detected. [4]

1.2 Proton composition

Protons consist primarily of two up quarks and one down quark, as well as gluons and sea quarks. Sea quarks are pairs of one quark and its respective anti-quark, created through gluon splitting. Gluons and sea quarks do not participate in the proton's core in the same way the 3 valence quarks do, meaning they do not determine the quantum numbers of the proton. When under no interactions with high-energy probes, gluons and sea quarks pop in and out of existence. On the contrary, when such an interaction occurs, they partake in hadronization processes, which means that they create hadrons as they separate. Even though only valence quarks determine the proton's structure functions, all partons (which include the valence quarks, gluons, and sea quarks) carry part of the proton's momentum. [5]

1.3 Proton-proton collisions

When a collision of two high-energy proton beams occurs, there is no guarantee that individual protons will interact. However, when an interaction does happen, it is described by an exchanged particle, according to the Standard Model of Physics. This exchanged particle carries a four-momentum vector Q, which represents the amount of momentum transferred in the interaction. The value of Q^2 (the squared four-momentum transfer) determines the nature of the interaction. For $Q^2 < 1GeV^2$, the protons will behave as a single particle rather than a collection of partons. For $1GeV^2 < Q^2 < 104GeV^2$, there is a clearer image of the composition of a proton, since the interactions will not involve protons as whole entities, but their valence quarks. Finally, for $Q^2 > 104GeV^2$, gluons and sea quarks will dominate the interactions. To sum up, if Q^2 is low, then the interaction is soft and the proton behaves as one entity, whereas if Q^2 is high, the interaction is a hard process and the proton's partonic structure becomes visible. [3]

During a hard collision at the LHC, partons from the colliding protons interact via a high-momentum transfer process, producing high-transverse momentum particles [6], such as Z and W bosons [7]. This occurs when a valence quark from one proton collides with a sea-anti-quark from the second proton, interacting via the weak force with one of the bosons as its carrier [8]. This happens according to these relations:

$$\bar{d}u \to W^+ \qquad d\bar{u} \to W^- \qquad \bar{u}u \to Z^0 \qquad \bar{d}d \to Z^0$$

These are the leading processes involved in the production of the corresponding bosons. Since they are heavy bosons, they immediately decay according to the following relations:

$$W^+ \to l^+\nu_l \qquad W^- \to l^-\bar{\nu}_l \qquad Z^0 \to l^+l^-$$

In this thesis, we will focus on the decays of Z^0 bosons. [42]

1.4 Frame of reference and conservation of momentum

Before this thesis dives deeper, it is important to establish the frame of reference used during proton-proton collisions. The two identical proton beams travel in opposite directions with opposite momenta. Thus, the centre-of-mass frame of reference coincides with the laboratory frame, since the detectors are at rest in this frame.

The coordinate system used is cylindrical, where the z-coordinate is the direction along the beam pipe, while the x-y plane is perpendicular to it. θ is the polar angle, and ϕ is the azimuthal angle. Pseudorapidity, η, is defined as $-log(tan(\theta/2))$, and it is constant on conical surfaces around the beam axis [3].

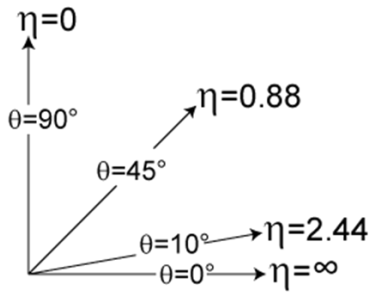

Figure 2: Pseudorapidity [31]

The data collected from the CMS are analysed by assuming that the visible transverse momentum before and after the collision is conserved. This is crucial because any violation of the conservation of momentum

in the transverse plane would mean that some particles went undetected, which is a very intriguing observation.

Concerning the Z^0 boson decays, the lepton and its corresponding anti-lepton form a dilepton system. According to the energy and transverse momentum of the dilepton system measured by the detectors (see section 1.6), the invariant mass is calculated. Consequently, it can be determined whether the dilepton system resulted from a Z^0 boson decay, simply by comparing that invariant mass with the rest mass of the Z^0 boson, which is approximately equal to 91.2 GeV/c^2.

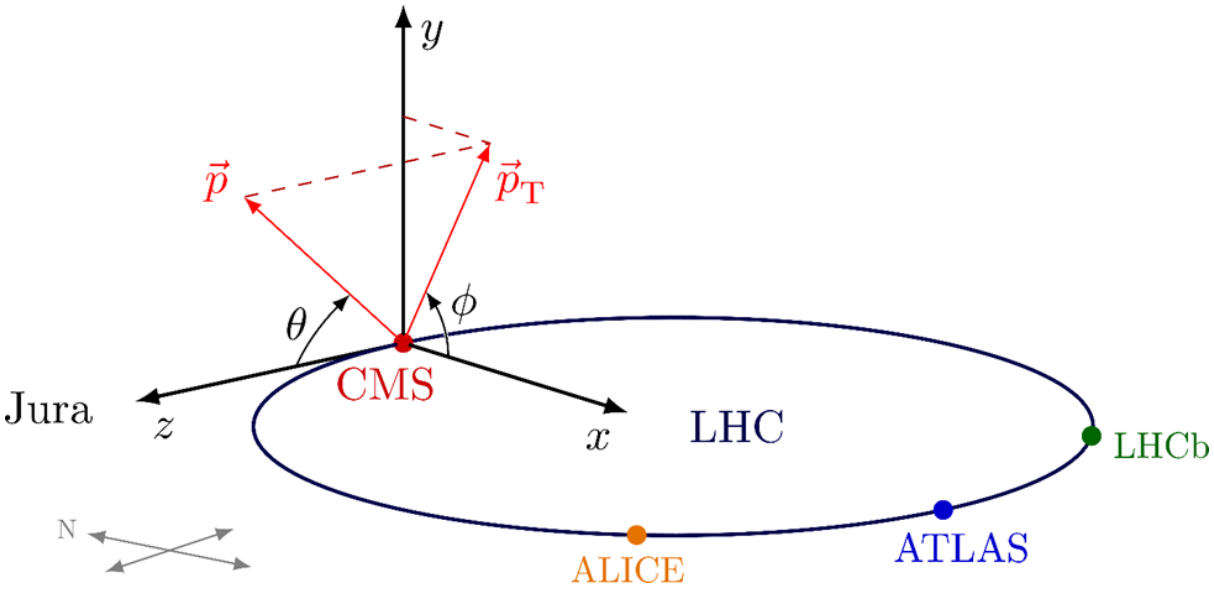

Figure 3: The CMS coordinate system [10]

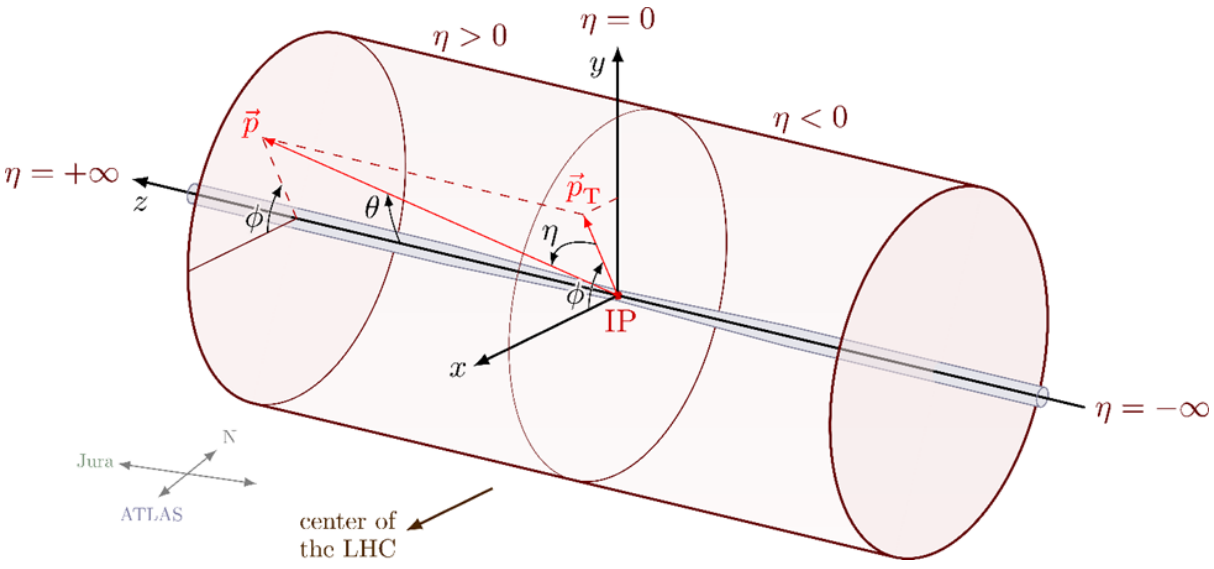

Figure 4: The coordinate system at the CMS with a cylindrical detector [10]

1.5 The CMS experiment

As mentioned in section 1.1, one of the detectors at the LHC is the CMS (Compact Muon Solenoid) detector. It is a general-purpose detector that aims to study the Standard Model, investigate particles that could make up dark matter, and search for extra dimensions. When a collision of two proton beams occurs at the centre of the CMS, numerous possible outcomes arise. For this thesis, we focus exclusively on the production of Z^0 bosons. These events are accompanied by at least one b-jet detected by the calorimeters (see section 1.6 for more information) [11].

1.6 The layers of the CMS – the subdetectors

Undoubtedly, the CMS must have a variety of detectors to detect and measure all the possible outcomes of a proton-proton collision. Specifically, the aim is to determine the particles' species, momentum, energy, and charge. To achieve that, the CMS includes tracking detectors that measure the charge, trajectory, and momentum of charged particles. There is also an electromagnetic calorimeter, which is responsible for measuring the energy and position of electromagnetic particles, such as electrons and photons. Next in line is a hadronic calorimeter, which is the equivalent of an electromagnetic calorimeter for hadrons, and lastly, there are muon chambers, which measure the momentum and trajectory of muons. The first three components fit inside a superconducting coil that provides the magnetic field necessary for the measurements, while the muon chambers surround it. [3], [40]

1.6.1 Tracking detectors

Tracking detectors are placed inside a calibrated magnetic field of B=3.8 Tesla directed along the longitudinal component of the detector [43], provided by a superconducting solenoidal magnet that envelopes the inner detectors of the experiment, which bends the trajectory of charged particles. There are generally two types of detectors, gas/wire drift chambers and silicon detectors. The latter is used at the CMS.

A tracking detector is made of small silicon detectors called pixels. Silicon detectors are semiconductors that have undergone doping to form a p-n junction. Then, a substantial reverse-bias voltage is applied to the junction, meaning that the depletion zone expands and the electric field inside the junction increases. As charged particles traverse the detector,

they ionise the silicon atoms inside the depletion zone, where there are only a few other free carriers to combine with and neutralise. The new charge carrier particles will be immediately affected by the depletion zone's strong electric field, creating electric signals, which are then detected and act as marks about where the particle has passed. Those marks are later used to reconstruct the particles' trajectory, a rather complicated process considering the large number of particles that pass through the detector for each collision. By determining the trajectory, it is possible to ascertain the charge sign of each particle, using the common knowledge that the magnetic field present in the detector will act on the particle with a force perpendicular to its velocity, known as the Lorentz force. This will cause the particle to move in a circular motion in the x-y transverse plane, and helical motion in the 3 dimensions, during its presence in the tracker with a radius r that can be obtained from its path. The more curved the path, the lower the momentum of the particle. After that, a quest begins to identify the particle, based on its interactions and behaviour in the detector, using algorithms [43]. Usually, for simplicity, the charge magnitude is taken to be $1|e|$, and the transverse momentum is calculated using the following relationship:

$$P_T = 0.3qBr,$$

where q is the charge of the particle in units of the absolute value of the electron charge e, B is the magnitude of the magnetic field, and r is the radius of curvature of the particle's path inside the tracker. During the particle identification, if the momentum calculated does not agree with the other measurements, the charge magnitude is re-evaluated, and the momentum is again calculated for the new charge magnitude value. It is also important to note that the trackers are built as thin as possible to avoid energy loss due to scattering. [3], [12], [13]

1.6.2 Electromagnetic calorimeter

An electromagnetic calorimeter fully absorbs the particles, aiming to measure the energy they carry. There exist two types of calorimeters, homogeneous and sampling. A homogeneous calorimeter consists only of one material, which is both the absorbing material and the detector, whereas a sampling calorimeter has alternating layers of passive absorbers and detectors. Detectors consist of scintillating, low-Z materials that radiate light when excited by ionising radiation.

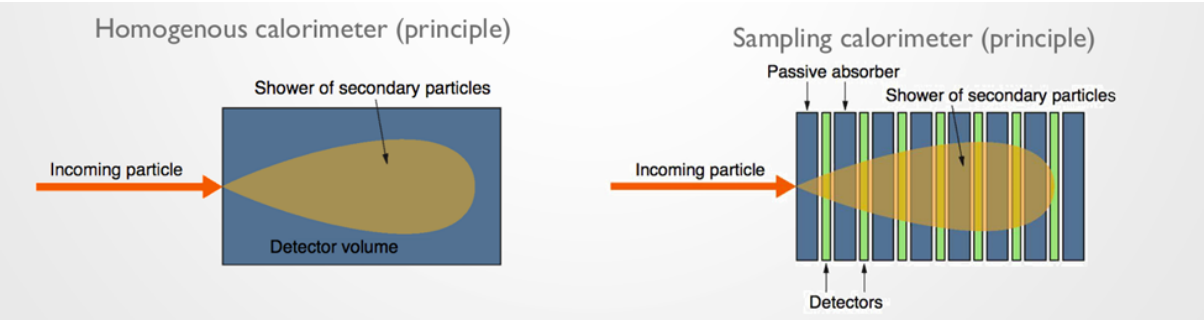

Figure 5: Homogeneous and sampling calorimeters [14]

Due to their lower price compared to the homogeneous calorimeters, sampling calorimeters are used at the CMS. They are composed of dense, high-Z passive absorber materials to initiate an electromagnetic shower when a particle travels through them.

Electromagnetic showers arise when particles that interact primarily through the electromagnetic force, known as electromagnetic particles, enter the calorimeter. Those particles include electrons, positrons, and photons. When those high-energy particles pass near the passive absorber's atoms, they interact via Bremsstrahlung radiation. As a result, they change their direction and lose energy via radiation- in other words, a photon is emitted. This interaction is demonstrated in the following figure.[15]

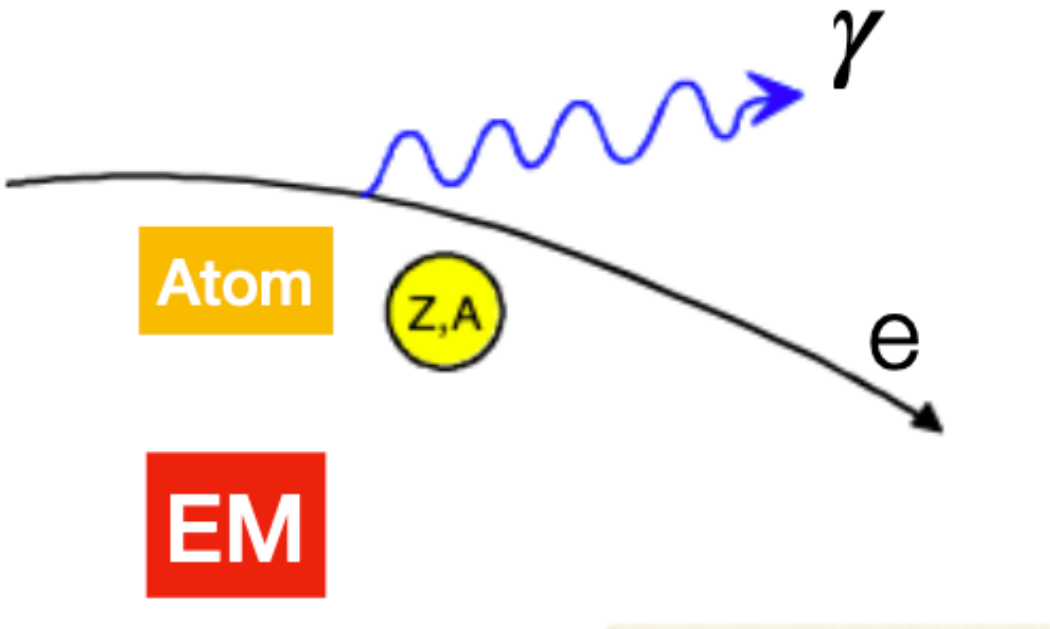

Figure 6: Bremsstrahlung radiation [17]

The photon undergoes pair production, converting into an electron-positron pair. The three electromagnetic particles go through the same process repeatedly; however, as this process continues, resulting in the secondary particles having lower energies, they begin to lose energy via ionisation, until the energy drops enough for the ionisation to be comparable to the Bremsstrahlung radiation. Specifically, when the rate of energy loss due to ionisation equals the rate of energy loss due to Bremsstrahlung radiation, the energy of the shower has dropped to a critical value, E_C, and the shower can no longer continue. During the ionisation process, the electrons lose energy by colliding with the detector's atoms and transferring the energy needed to eject one or more electrons, leaving behind an ionised atom.

A very important parameter regarding the electromagnetic calorimeters is the radiation length X_0, $X_0 = \frac{716.4 g cm^{-2} A}{Z(Z+1) ln(287/\sqrt{Z})}$, where A is the atomic mass and Z is the atomic number of the atoms composing the absorber. The radiation length is measured in centimetres, and it is defined as the average distance an electron will cover before reducing its energy by 1/e [16]. Calorimeters of this type typically have widths of about 15-30 X_0. As mentioned previously, high-Z materials are preferred when constructing electromagnetic calorimeters because they have smaller radiation lengths, reducing the size of the calorimeter itself. Photons also correlate with the radiation length, for their mean free path before they undergo a pair production is equal to $9X_0/7$ [3]. The distance a shower covers inside a calorimeter is given by

$$X_{max} = X_0 \times ln(E_0/E_C),$$

where E_0 is the energy of the initial particle.

Of course, a sampling calorimeter does not absorb all the energy of the particle, since the materials that initiate electromagnetic showers do not partake in the energy detection process. Consequently, part of the initial particle's energy is lost in the absorber layers because some of the secondary particles produced in the electromagnetic shower do not reach the closest detector layer. The energy resolution, which describes the relative uncertainty of the energy measurements, is given by the following relation:

$$\frac{\sigma(E)}{E} \approx \sqrt{(\frac{c_1}{\sqrt{E}})^2 + (\frac{c_2}{E})^2 + c_3^2}$$

The first term under the square root is the stochastic term, which describes the fluctuations in energy deposition due to the random nature

of particle interactions, which means that even if the initial energy is the same, the number of secondary particles produced is not always the same. The second term is the noise term, representing the noise in the readout electronics, as well as the pile-up from particles produced in other collision events that arrive almost simultaneously with the particles that need to be measured. Finally, the third term is called the constant term because it expresses variations in the calorimeter's dimensions, leading to inhomogeneous responses, calibration errors, and energy loss in dead channels.

Some of the factors contributing to the energy resolution are a result of the nature of the sampling calorimeters. Hence, homogeneous calorimeters have a better energy resolution ($\sim 1\%$) compared to the sampling calorimeters ($\sim 10\%$). In addition, it is obvious that the more energy the particles have, the higher the resolution of the calorimeter. [3], [14], [17]

1.6.3 Hadronic calorimeter

As stated before, a hadronic calorimeter measures the energy and position of heavy hadrons. The only type of hadronic calorimeter is a sampling one. Hadron calorimeters roughly work in the same way electromagnetic calorimeters do, however, they rely on hadronic showers that divide the energy of the initial hadron entering the calorimeter instead of electromagnetic showers. Furthermore, hadrons interact via strong interactions with matter as opposed to electromagnetic interactions that govern the electromagnetic calorimeters. Here, the important parameter equivalent to the radiation length is the interaction length λ_n , $\lambda_n \approx 35 g cm^{-2} A^{1/3}$, where A is the atomic mass. The interaction length is the average path of a particle between two inelastic collisions, and it depends on the material the calorimeter is made of [18]. For heavy or high-Z materials, the interaction length is greater than the radiation length, meaning that the electromagnetic showers precede the hadronic ones [3]. For charged hadrons and particles in general, calorimetry is complementary to the track measurements, whereas for neutral hadrons, hadronic calorimetry is the only way to measure their energy.

When a hadron enters the calorimeter, it interacts via the strong force with a nucleus in the calorimeter. There are two possible outcomes from this interaction: It will either be elastic and not produce any secondary particles, meaning the total kinetic energy of the system is conserved, or it will be inelastic and result in the production of secondary particles.

In the second scenario, a hadron interacts via strong interaction with a nucleon, which increases its energy and interacts with other nucleons inside the nucleus, forming an intra-nuclear cascade. This process results in the production of hadrons, mostly neutral and charged pions, that escape the nucleus along with some high-energy nucleons, causing a cascade of secondary particles. These cascades may also result in the production of neutrinos or low-energy neutrons that go undetected, or, in the case of neutrons, deposit their energy inefficiently. The neutral pions produced decay into two photons that undergo pair production. This generates an electromagnetic shower, one of the two components of the shower caused by a hadron. The rest of the hadrons continue to interact with other nuclei, repeating the same process described previously, creating the hadronic component of the shower. Eventually, the hadrons' interactions with nuclei will result in the production of neutral pions that will end in electromagnetic showers. It is also important to note that charged hadrons ionise both electrons and nuclei, with the latter being another photon-producing mechanism which initiates an electromagnetic shower. To follow a hadron from the moment it enters the calorimeter to the moment the cascade it produces ends, a hadronic calorimeter must be $10\lambda_n$ thick. [14], [19]

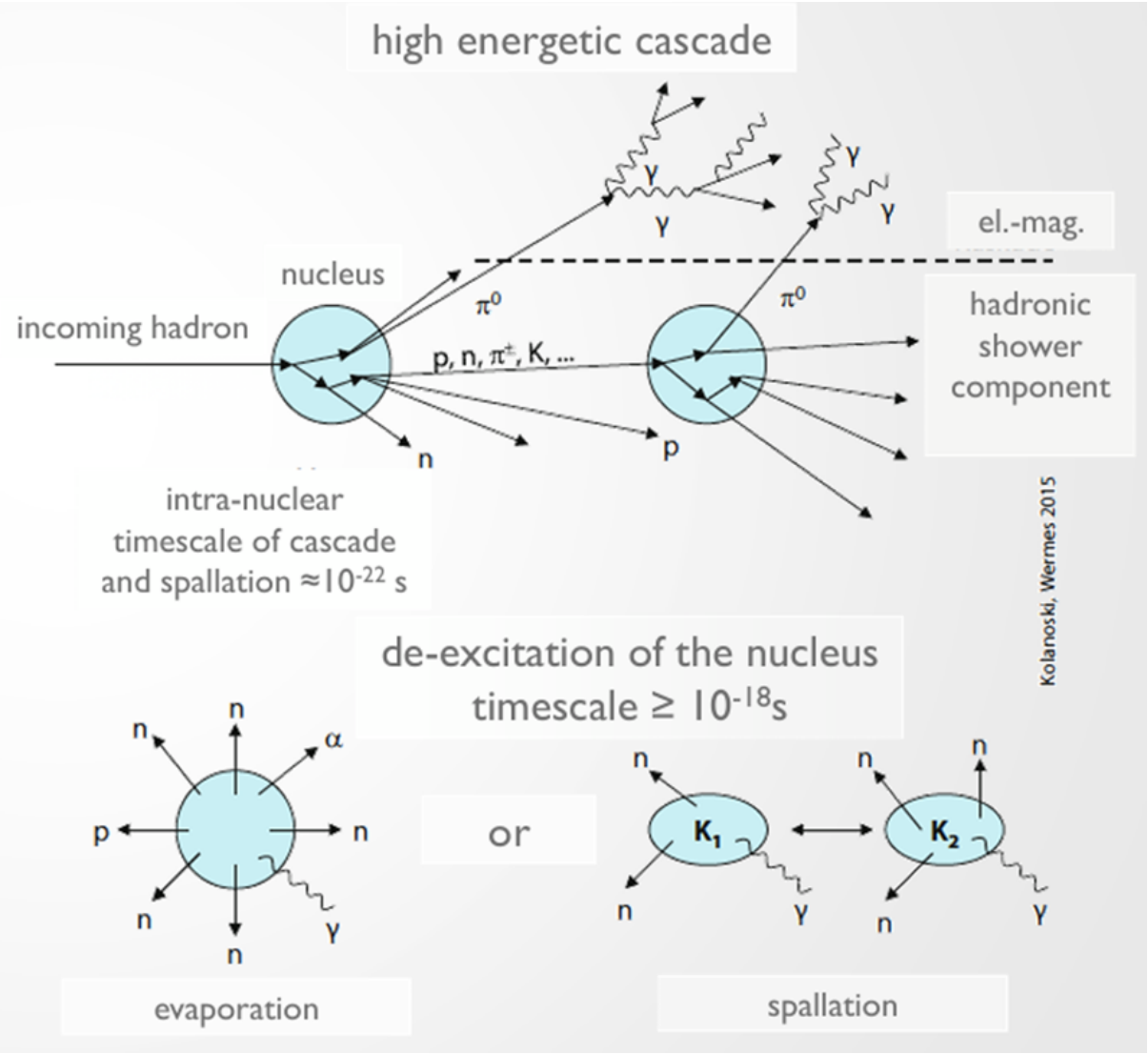

Figure 7: Hadronic interactions [14]

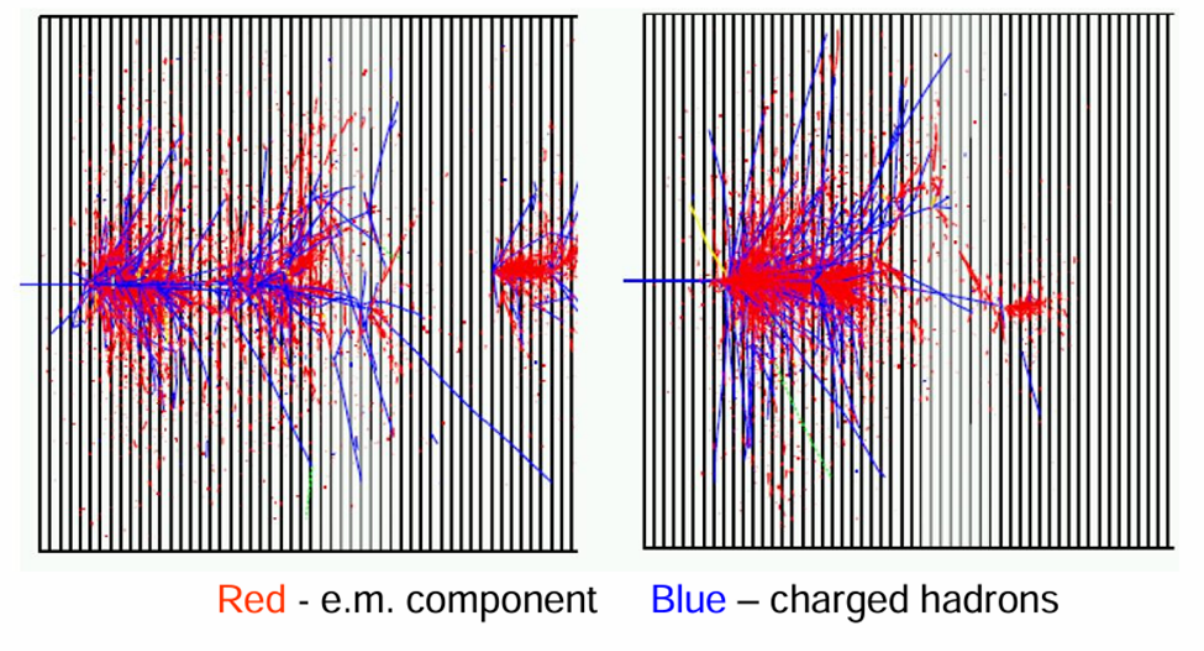

Figure 8: Simulation of hadronic showers [17]

The energy resolution of a hadronic calorimeter is again given by

$$\frac{\sigma(E)}{E} \approx \sqrt{(\frac{c_1}{\sqrt{E}})^2 + (\frac{c_2}{E})^2 + c_3^2},$$

where the terms have the exact same meaning as described in the section about electromagnetic calorimeters; however, more terms may appear, making the resolution worse compared to the electromagnetic calorimeters. [19], [20]

1.6.4 Muon chambers

Muon chambers are responsible for muon detection and recording. As suggested by the name CMS, this is one of the detector's main aims since it searches for the Higgs boson, which decays into four muons.

Muon chambers are placed in the outermost part of the detector because muons are highly penetrating and interact very weakly with matter compared to other charged particles. Consequently, calorimeters cannot stop muons, unlike most particles. The muon detector is constructed outside the magnetic coil, which results in the magnetic field present being around 2 Tesla. This is due to the steel yoke layers alternating with the muon detectors. The yoke is responsible for redirecting the magnetic field lines existing outside the solenoid back to it, to close the loops. These stations are responsible for tracking the trajectory of the muons. The chambers

consist of 250 drift tubes and 540 cathode strip chambers, responsible for monitoring the muon's trajectory, and 610 resistive plate chambers and 72 gas electron multiplier chambers that clear out unwanted data. [13], [21], [22], [23], [27]

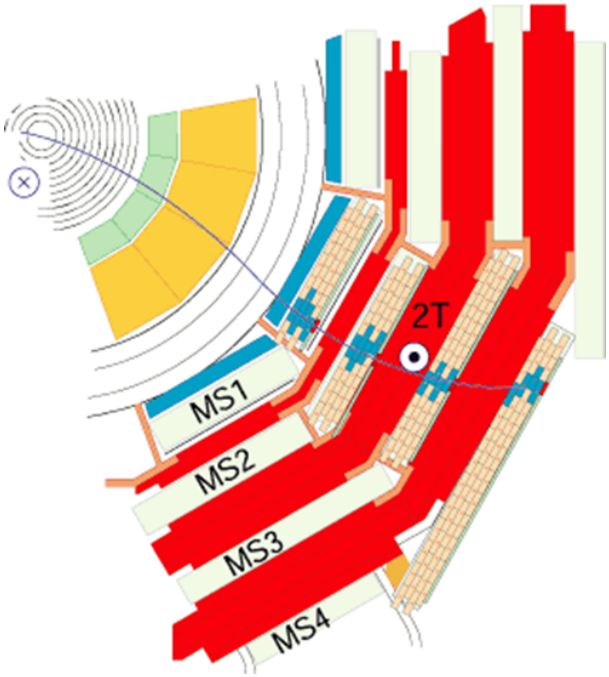

Figure 9: The four muon chambers [24]

Like every other part of the detector, the CMS has both a barrel and an endcap region. The barrel region is placed at the region where $\eta < 1.4$, while the endcap region is placed at $\eta > 1.4$. The barrel region contains Drift Tubes (DT) along with Resistive Parallel Plate Chambers (RPC), and the endcap region contains Cathode Strip Chambers (CSC) and RPSs.

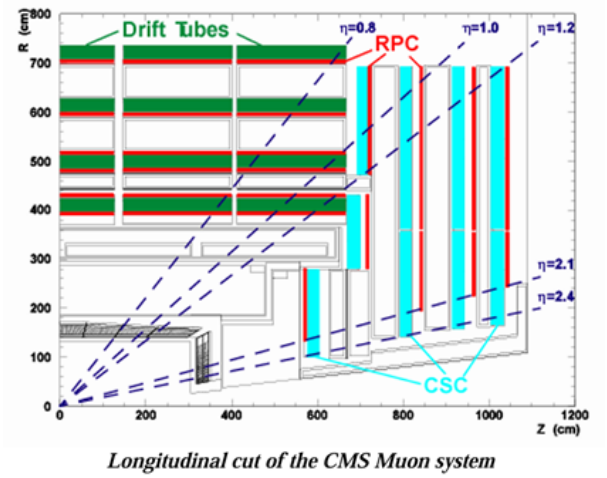

Figure 10: DTs, CSCs, and RPCs in the barrel and the endcap regions. The barrel and endcap regions are shown with respect to different pseudorapidity values. [30]

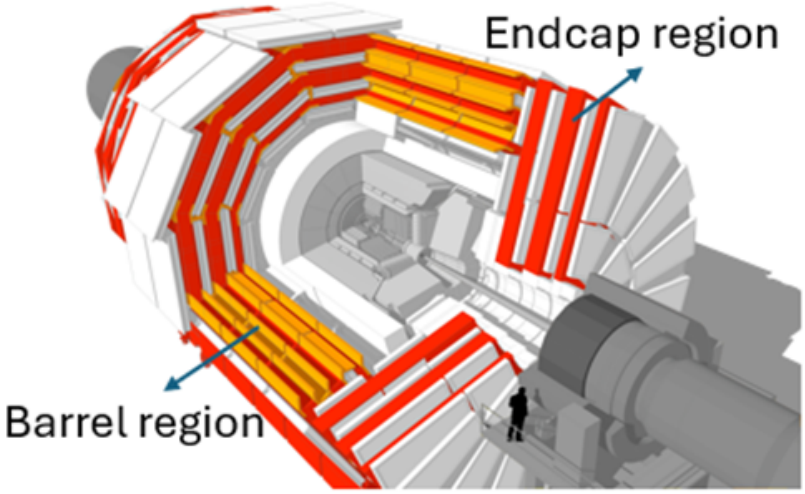

Figure 11: Barrel and endcap regions-image adapted from [26], with modifications by the author of this thesis.

As seen in Figure 10, barrel regions have alternating layers in the direction perpendicular to the beam pipe. The barrel is a yoke, which is responsible not only for guiding the magnetic field lines into loops but also for muon detection. From closest to furthest from the pipe, there is a muon chamber, followed by an iron (steel) slab, and this repeats until the fourth muon chamber. Specifically, the iron slabs are guiding the magnetic field lines, while the muon chambers are recording the muons' trajectories. The chambers are made of Drift Tubes, which are stretched wires within gas volumes. [23], [22], [13], [28], Drift Tubes are negatively charged tubes

(cathodes) filled with gas and have a positively charged wire (anode) at their centre. When a muon passes through the gas, it marks the beginning of an event, and it ionises the atoms in its path, creating free electrons and positive ions. The electrons drift radially through the gas towards the anode wire, under the influence of the electric field created between the anode wire and the cathode tubes. When the electron reaches the wire, it marks the end of the event. Those electrons later induce a current in the wire, creating an electric pulse. The drift time of the electron is identified as the time it takes for an electron to reach the anode from the moment of the ionisation. Based on the drift time, the space-to-drift time relation is used to calculate the position where the ionisation took place, which is a point of the muon's trajectory. [29]

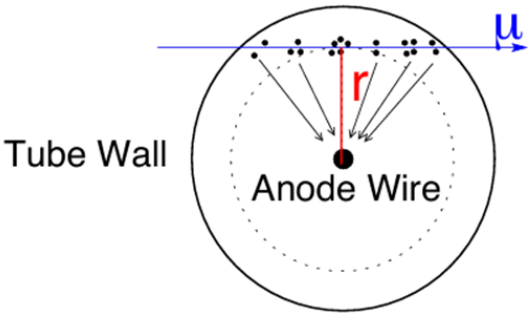

Figure 12: Principle of a spatial measurement inside a drift tube [29]

Of course, since the pulse travels along the wire, a drift tube can only give information about one spatial component of the muon's trajectory. A drift chamber consists of 12 aluminium layers segmented into long aluminium drift cells and stacked on top of each other, in the direction of the beam pipe. The layer stack is divided into three sections of four layers, called superlayers; the two outer superlayers have parallel drift tubes, while the middle superlayer has drift tubes perpendicular to them, as shown in the figure below:

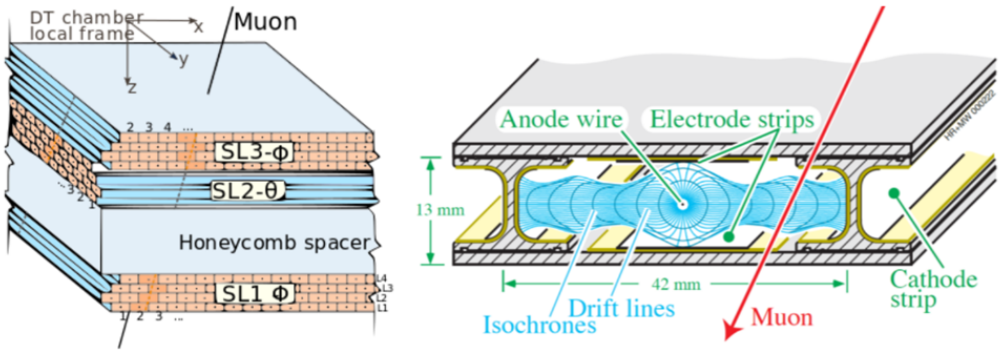

Figure 13: The muon chamber (left) and a drift tube (right) [33]

The z coordinate is measured according to the layer the muon has passed through, while the coordinates on the transverse plane are measured through gas ionisation, as described previously. In the chamber's local frame (Fig. 13), the x coordinate is calculated from the two outer superlayers while the y coordinate is calculated from the middle superlayer. In general, each coordinate is detected by the drift tubes that are oriented perpendicular to the corresponding axis. [26], [33]

Moving to the endcap region, the cathode strip chambers are constructed as 6-layer trapezoids that overlap to cover the entire space, ensuring that no muons are left undetected. In that region, the magnetic field is intense, with intensities reaching up to 3.5 Tesla, and the particle rates are large. Like the barrel region, the endcap region is also a magnetic yoke that is responsible not only for detecting muons but also for returning the magnetic field lines to complete the magnetic field circuit. Roughly, the structure of the yoke is the same as in the barrel region, with alternating layers of cathode strip chambers and steel slabs. The chambers are made of rigid panels that seal gas inside. Within that gas volume, there is also an alternating assembly of anode wire layers and copper cathode strip layers stacked on top of one another, with gas volumes between them. The cathode strips are arranged radially to the beam pipe, while the anode wires are positioned perpendicular to them.

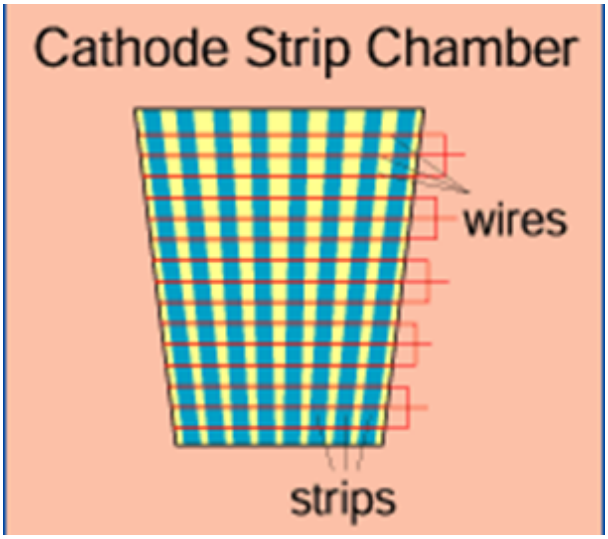

Figure 14: The cathode strip chambers wire-strip grids [32]

When a muon enters the chamber, it ionises the gas, giving free electrons and positive ions. As with the drift tubes, when this happens, a trigger is generated that marks the start of an event. The negative electrons move towards the anode wire, and the positive ions move towards the cathode strips. As the electrons approach the wire, the electric field is more intense, causing them to accelerate and collide with the gas atoms and ionise them. Through this mechanism, more electrons are produced, resulting in an electron avalanche, which acts as the amplification of the pulse generated by the initial ionisation. When they reach the wire, another trigger is created that signals the end of the event. The positive ions left from the ionisation caused by muons, as well as the electron avalanche, move more slowly due to their greater mass. Their motion is enough to induce a displacement current in the strip electronics. Again, based on the drift time of the electrons, the ionisation position can be calculated. When it comes to the positive ions, the position is directly measured by the pattern of induced charge across the strips. As a result, the detector can measure 2 coordinates, with the third being implied by the position of the chamber. [32], [28], [34], [35], [36], [26], [37], [13], [38]

Finally, the last category of muon chambers is the Resistive Plate Chambers (RPCs). The structure of an RPC consists of a gas-tight space constructed with two parallel plates, made of phenolic resin (bakelite), which is a highly resistant material. The outer surfaces of the two plates are covered with graphite paint, which is a conductive material. Touching the graphite, there is an insulating PET film that prevents the contact of the graphite with the aluminium strips, responsible for the read-out of

the signal. One of the two graphite-coated surfaces is connected to high voltage, while the other is connected to ground electrodes. This results in the existence of a strong electric field inside the gas-tight space. As mentioned before, RPCs are both in the barrel and endcap regions of the detector, placed next to some of the DTs and CSCs.

When a muon passes through an RPC, it ionises the gas, producing pairs of electrons and positive ions. Under the influence of the electric field, the electrons are accelerated towards the anode, creating an avalanche of electrons, while the positive ions move towards the cathode, like in the CSCs. This causes local changes in the electric field, which end up inducing a signal in the nearby aluminium strips, called a hit. Based on the structure of the hits, information can be extracted about the muon's trajectory, and therefore, transverse momentum. [28], [39]

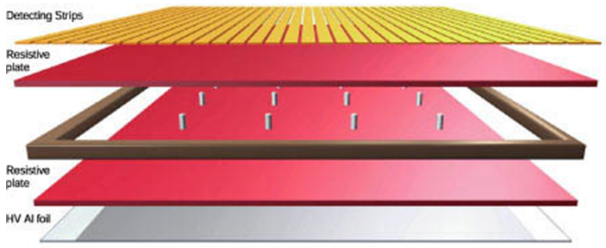

Figure 15: The structure of a Resistive Plate Chamber [39]

To sum up, the CMS detector consists of a tracker, responsible for detecting the trajectories of charged particles, an electromagnetic calorimeter for measuring the energy of electromagnetic particles, a hadronic calorimeter for measuring the energy of hadrons, and a muon detector for detecting muons. The following figure illustrates the complete picture of the CMS detector, including all the different parts described in this section.

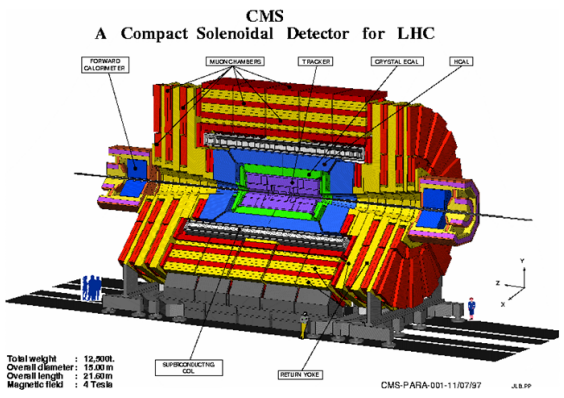

Figure 16: The components of the CMS detector [25]

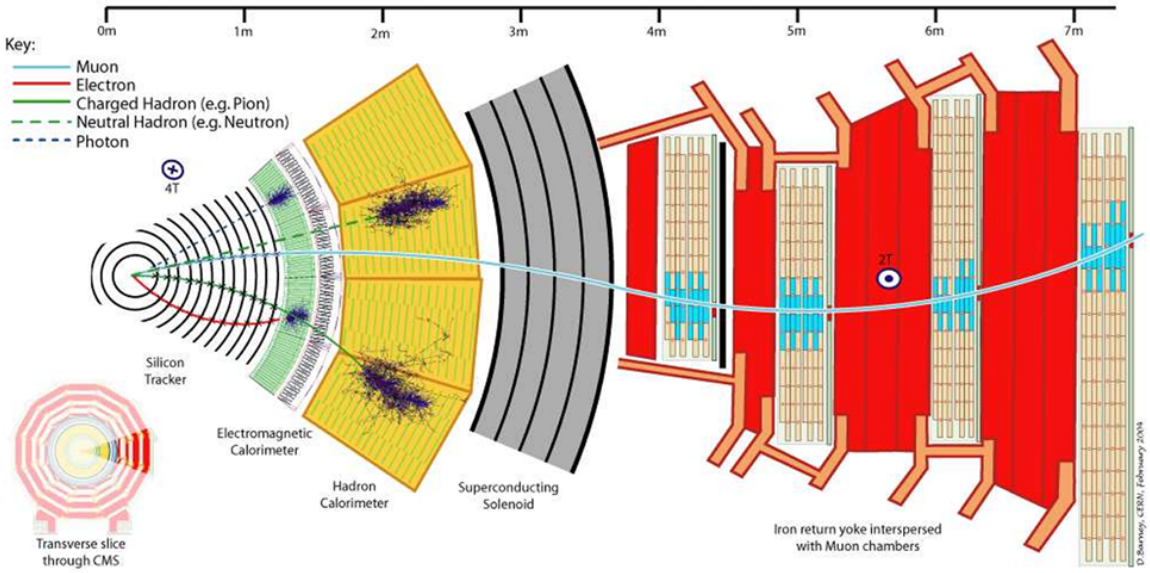

Figure 17: Particle detection in the CMS detector [3]

2 Electron measurements

2.1 Electron trajectory reconstruction

The path reconstruction for electrons begins in the tracker, the innermost part of the CMS detector. As described in section 1.6.1, each pixel is a detector that can produce a signal in case a charged particle passes through it. The signal consists of charge carrier pairs of electrons and holes, and it must have a minimum charge equivalent to 3200 electrons in the barrel region and 3000 electrons in the endcap region inside each pixel, to be distinguishable from background noise and not be filtered out as not important. The signal spreads out as it drifts towards the pixel electrodes responsible for reading it, which consequently generates a signal from neighbouring pixels. Furthermore, the magnetic field present bends the trajectories of the carriers, making them land on adjacent pixels. A hit is defined as a cluster of neighbouring pixels that recorded a signal above the threshold. For a cluster to be formed, the threshold charge is equivalent to that of 4000 electrons. The clusters' positions are found using two algorithms.

The first algorithm, called Combination Track Finder, is fast. It is based on the Kalman filter method, and it builds track candidates by extending seeds layer-by-layer through the tracker. Primarily, the track candidates are constructed for a few hits to get an initial estimate for the track parameters. The Kalman filter is used to predict the hits of each

specific track in each layer. If more than one hit is compatible with a prediction, different possibilities for different tracks arise. When a track candidate is fully reconstructed, the track parameters are again estimated for all the hits included. Then, different criteria like the chi-squared method, number of hits, or missing hits are used to assess the quality of the track and reject the ones that ranked low. This method is suitable for identifying the best quality, high transverse momentum tracks, which are more reliable and are used to establish the primary vertices.

The second algorithm is called Iterative Tracking, and it includes the iterative use of the first algorithm. It aims to reconstruct more challenging tracks with lower transverse momenta or missing hits in some layers, or even displaced tracks, without using hits that have already been included in other tracks. [37]

2.2 Measurements of electron charge sign

As discussed in Section 1.6.1, the sign of the electric charge of electrons and positrons produced in proton–proton collisions is determined within the tracker, where the 3.8 Tesla magnetic field generated by the superconducting coil bends the trajectory of those charged particles. The magnetic field acts on a charged particle via the Lorentz force, $\vec{F_L} = q(\vec{E} + \vec{U} \times \vec{B})$, where $\vec{E}$ (electric field) is taken to be zero. Since it is perpendicular to the trajectory and the magnetic field, $\vec{B}$, the Lorentz force makes a charged particle move in a circular motion in the transverse plane, with the bending depending on the charge sign. Electrons and positrons have opposite charges, resulting in their trajectories curving in opposite ways.

While traversing the electromagnetic calorimeters, where they deposit most of their energy, electromagnetic particles leave energy deposits that are analysed in conjunction with the trajectory information from the tracker. These measurements allow the reconstruction of the particles' trajectories inside the CMS detector. Based on the radius of curvature in the tracker and the intensity of the magnetic field, the transverse momentum is calculated under the assumption of a unit charge magnitude $|e|$. Suppose the inferred momentum from the reconstructed trajectory, and thus the energy, does not match the calorimetric energy or contradicts established particle identification. In that case, it may indicate track mis-reconstruction or particle misidentification. Subsequently, the particle's momentum is recalculated to ensure coherence with all experimental observations. However, while the charge magnitude may be correct, a residual possibility exists

that the charge sign might be mismeasured.

2.3 Mismeasurements of electron charge sign

It has already been established that high-momentum particles have a bigger radius of curvature under the influence of the magnetic field in the tracker than low-momentum ones. From that, it becomes clear that there are some challenges when it comes to measuring the charge sign of such a particle, since it does not have much space to unravel its trajectory. As a result, determining the way the trajectory bends is ambiguous, leading to false measurements.

Along with that comes the Bremsstrahlung radiation, because even if it is as limited as possible inside the tracker, it is not non-existent. As described in section 1.6.2, Bremsstrahlung radiation can result in the change of the electron's direction, as well as the production of new particles via pair production. The change in the electron's direction means that the electron no longer moves in its predicted circular motion, and this confuses the reconstruction algorithms. Even more, if the Bremsstrahlung photon takes up most of the electron's energy, it will probably end up conducting a pair production. In this case, there is a probability that the electron with the opposite charge compared to the electron that underwent Bremsstrahlung interaction will take most of the Bremsstrahlung photon's energy. If this happens, then this 'wrong' electron will create hits inside the tracker, while moving in the opposite direction compared to the initial electron, due to the magnetic field. Consequently, the algorithms will take into consideration those 'false' hits, leading to false tracking, and therefore, measurement of the wrong curvature, and thus, charge sign of the initial particle. This not only creates the problem of creating more hits than the initial particles would produce, but also the fact that the particles lose energy via the radiation, thus their momentum is reduced while they traverse through the detector. To sum up, in the case where the Bremsstrahlung photon takes a significant amount of energy from the initial electron, since the radius is directly proportional to the transverse momentum, the particle's radius will be reduced as well, and the further it traverses through the tracker, the more curved its path will be. However, the algorithms responsible for reconstructing the tracks fit all the hits in one curve, which means that if the hits before radiation dominate, the charge will probably be measured correctly. On the contrary, if the hits after the radiation dominate, which happens if a significant amount of

energy is lost via Bremsstrahlung radiation, there is a higher probability for the curve to be misfitted and for the charge sign to be mis-measured. This raises a question: how often does this mismeasurement happen, and how can one know it occurred?[41]

3 Analysis

3.1 Data

3.1.1 Electron-antielectron and Muon-antimuon data

To answer the question set in Section 2.3, something with a known charge that decays into those two particles (e^+e^-) has to be measured. Since the Z^0 boson has a zero electric charge, the decaying system also has a total charge of zero, and therefore, if it does not, it is a valid indication that something was measured wrong. This is why Z^0 boson decays were used to study these charge mismeasurements.

In this project, we received measurements of electrons and muons that originated from Z^0 boson decays. As noted previously, a Z^0 boson decays into two oppositely charged leptons: a lepton and its corresponding antilepton. The CMS can detect e^+e^- and $\mu^+\mu^-$ using trackers, electromagnetic calorimeters, and muon chambers. Each event studied in this thesis involved the detection of only two electrons or two muons, which are candidates for Z^0 boson decays. The next step involves calculating the invariant mass m of the dilepton system using the energy formula from special relativity:

$$E^2 = (mc^2)^2 + (pc)^2$$

which can be written as

$$E^2 = m^2 + p^2$$

for c=1, where the total energy of the system is $E = E_1 + E_2$, with E_1 and E_2 being the energies measured by the calorimeter for the lepton and antilepton, respectively, and $\vec{p} = \vec{p_1} + \vec{p_2}$ is the vector sum of the momenta of the lepton and antilepton.

When a particle decays, the system of products will have the same invariant mass as the initial particle.

The invariant mass of the Z^0 boson is 91.1876 GeV/c^2, so when the invariant mass of the dilepton system is approximately equal to this value, it is highly likely to have originated from a Z^0 boson decay. If a dielectron pair has the invariant mass of the Z^0 boson, we should expect the electrons

to have opposite charge. If a same-sign dielectron pair has the Z^0 boson's invariant mass, there has likely been a mismeasurement.

3.1.2 Graphical representations of the dilepton pairs

In this section, the mass spectra of the dilepton pairs will be displayed.

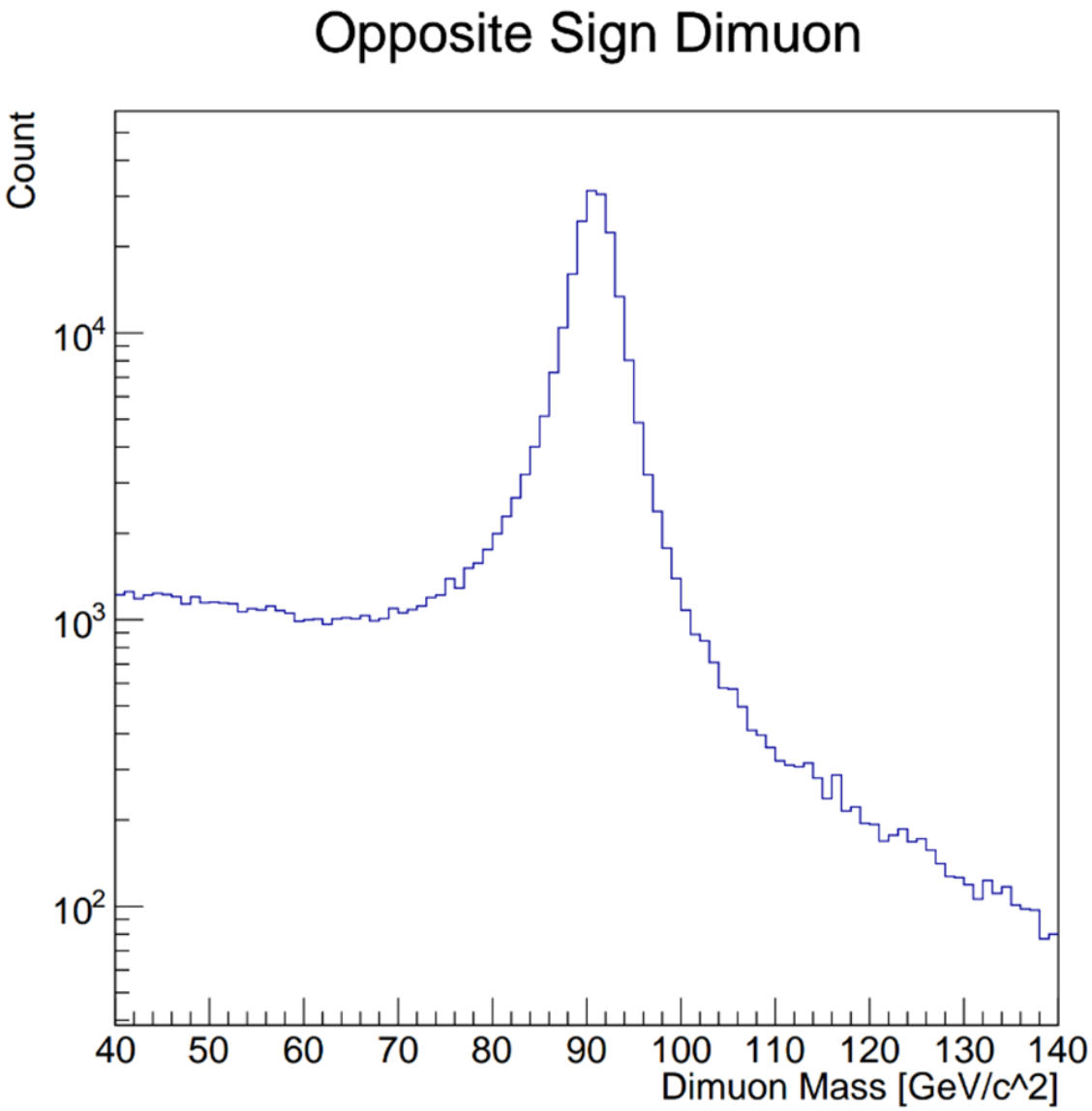

Figure 18: The graphical representation of the distribution of the opposite-sign dimuon events versus their invariant mass. Each event consists of a muon and an anti-muon. The graph is on a logarithmic scale for a better visual interpretation of the event distribution. There is a resonance around 90 GeV/c^2, meaning that those events most likely originated from Z^0 boson decays. The number of events in the resonance is far greater than the rest of the graph, which can be considered as background.

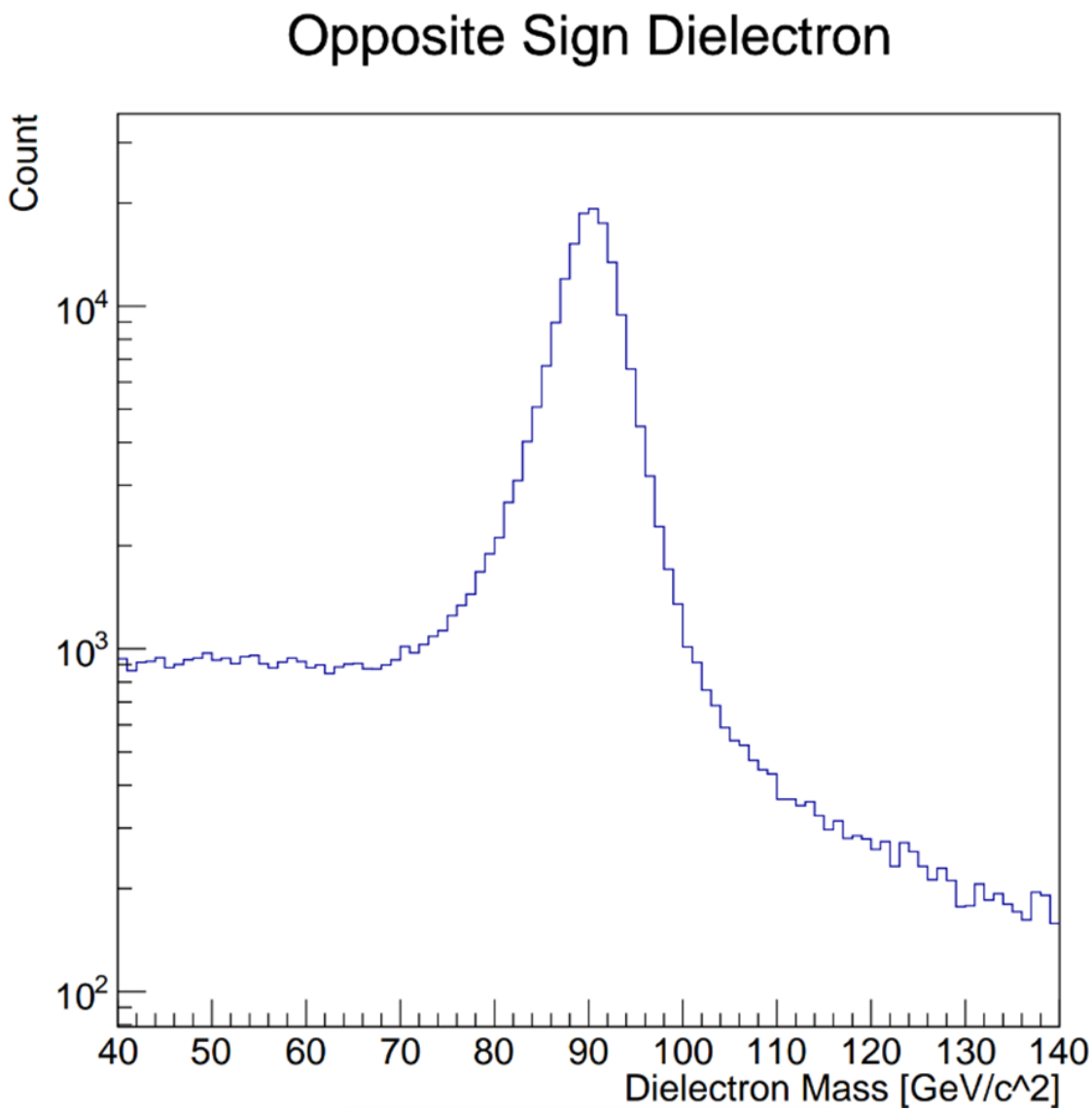

Figure 19: The graphical representation of the distribution of the opposite-sign dielectron events versus their invariant mass. Each event consists of an electron and a positron. The graph is on a logarithmic scale for a better visual interpretation of the event distribution. There is a resonance around 90 GeV/c^2, meaning that those events most likely originated from Z^0 boson decays. The number of events in the resonance is far greater than the rest of the graph, which can be considered as background.

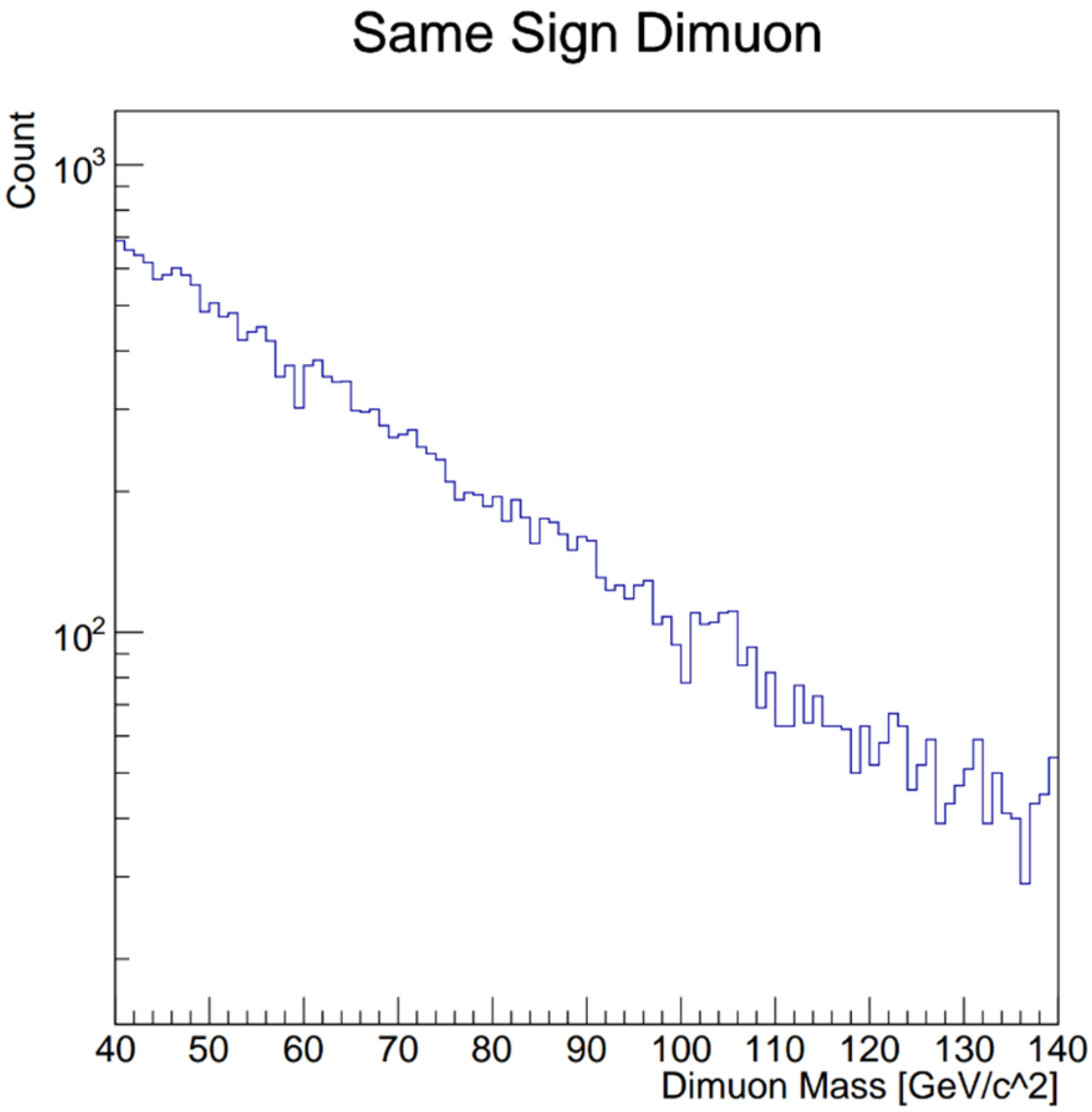

Figure 20: The graphical representation of the distribution of the same-sign dimuon events versus their invariant mass. Each event consists of two same-sign leptons with leptonic number $l_\mu = \pm 1$, meaning each pair is either a pair of muons or ant-imuons. The graph is on a logarithmic scale for a better visual interpretation of the events. There is no resonance observed in this graph, which leads to the result that in this dataset, there is no dominant resonance production mechanism for same-sign dimuon events.

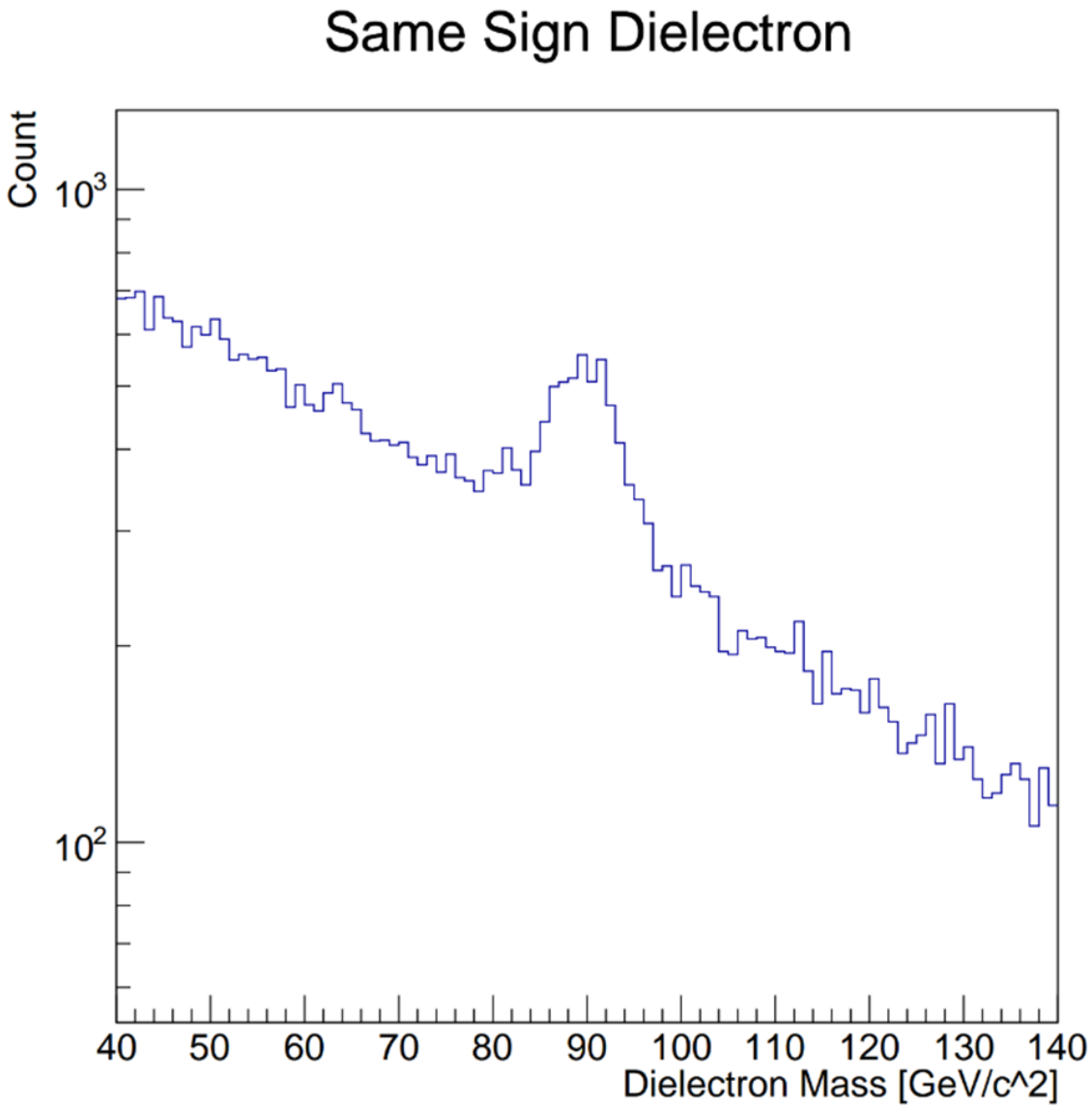

Figure 21: The graphical representation of the distribution of the same-sign dielectron events versus their invariant mass. Each event consists of two same-sign leptons, either electrons or positrons. The graph is on a logarithmic scale for a better visual interpretation of the events. There is a resonance around 90 GeV/c^2, meaning that there is a mechanism producing those pairs that originates from a particle that has an invariant mass of approximately 90 GeV/c^2. There are no other anomalies in the graph, which means that the rest of the events are considered to have originated from background processes.

From the previous discussions in Sections 2.3 and 3.11, it is safe to say that the resonances observed in Figures 18, 19, and 21 resulted from Z^0 boson decays. It is important to note that, since those events should all have an invariant mass equal to $\sim 91.2 GeV/c^2$, the peak would have been expected to be a delta function around that value. However, due to the errors of the calorimetric measurements, which follow the Gaussian distribution, the peak has a non-zero spread.

Of course, the resonance in Figure 21 is an anomaly, since an electrically neutral Z^0 boson cannot produce two electrons of the same sign. Therefore, it follows that this resonance results from the mismeasurement of the charge sign of an electron or a positron in an electron-positron pair. From this arises the need to measure the charge sign mismeasurement in data. This thesis aims to find the probability of this mismeasurement occurring.

3.1.3 The extraction of the mismeasurement probability formula

The Z boson decay of interest is as follows:

$$Z^0 \rightarrow e^- e^+$$

However, when charge-sign mismeasurements take place, there are three possible outcomes: either the electron is going to be measured as a positron, the positron will be measured as an electron, or both mismeasurements will happen simultaneously. In the latter case, the result will be an opposite-sign dielectron pair, which cancels the mismeasurement.

Each dielectron signal event corresponds to a Z^0 boson decay, so to find all the decays that resulted in the production of e^-e^+, there is a need to measure the events in the signal regions of Figures 19 and 21 and take their sum. This number will be represented by N. Now, assume that p is the probability that the charge of one lepton is mismeasured. For the same sign events, there are two possibilities: either the charge of the electron will be mismeasured, with a probability of $p(1-p)$, or the charge of the positron will be mismeasured, with a probability of $(1-p)p$. So, in total, the number of same-sign signal events will be $N_{SS} = N \times 2p(1-p)$. For the opposite sign events, there are also two possibilities: either the charge of both leptons will be correctly measured, with a probability of $(1-p)(1-p)$, or both leptons will be mismeasured with a probability of p^2. So, the total number of opposite sign signal events is $N_{OS} = N \times [p^2 + (1-p)^2]$. Assuming that the probability is very small, the results will be $N_{SS} \approx 2Np$ and $N_{OS} \approx N(1-2p)$, while $N = N_{SS} + N_{OS}$. From these, it can be extracted that the statistical probability is $p = \frac{N_{SS}}{2(N_{SS}+N_{OS})}$. $N = N_{SS} + N_{OS}$ is the total number of Z^0 bosons that decayed into e^+e^-. The error propagation for N goes as follows: $\sigma_N = \sqrt{{\sigma_{N_{SS}}}^2 + {\sigma_{N_{OS}}}^2}$, while the error propagation for p is:

$$\sigma_N = \sqrt{(\frac{2N_{OS}\sigma_{N_{SS}}}{4(N_{SS}+N_{OS})^2})^2 + (\frac{2N_{SS}\sigma_{N_{OS}}}{4(N_{SS}+N_{OS})^2})^2}$$

which gives the statistical uncertainty of the probability.

To measure this probability, there is a need for developing methods that will allow us to differentiate the signal data from the background data, to get the number of events that originated from Z^0 boson decays.

3.2 Data analysis

3.2.1 Probability calculation and statistical uncertainty

For the probability calculation, we assumed that the background data can be described by a quadratic function, and the signal data follow a Gaussian distribution. There is no particular reason for the choice a quadratic function for the background data. After all, any well-fitting function on the data should not notably affect our probability calculations. However, the Gaussian function was a conscious decision that resulted from the discussion in Section 3.1.2.

Firstly, a function was fitted to the same sign dimuon data to see whether it would fit the data well. The reasoning behind that was that the mechanisms that produce electrons and muons are quite symmetric, so their background processes are similar. As a result, if the function was a good choice for the dimuon data, it would also be for the dielectron data.

For the same sign di-muon data, the function used was the following:

$$f = [0] + [1]x + [2]x^2 + [3](1/[4])e^{-\frac{1}{2}(\frac{x-[5]}{[4]})^2}$$

where the numbers in square brackets represent free parameters. The quadratic part was supposed to describe the background processes, and the Gaussian represented the signal events. The numbers represent the free parameters that were extracted from the fitting process. Even though there was no visible peak in the same sign dimuon data, we still gave the opportunity for the fit function to show a peak. The reason is that muons are approximately 200 times heavier than electrons, and since Bremsstrahlung radiation is inversely proportional to the square of the mass of the particle, muons do not emit Bremsstrahlung photons as often as electrons.

Next, we fitted a function with free parameters that was a sum of a quadratic and a Gaussian to the same sign dielectron data, and a function with free parameters that was a sum of a quadratic and two Gaussian functions for the opposite sign dielectron data. The reason two Gaussians were used for the fitting of the opposite sign data was that the peak was very high, and it would not fit the data well otherwise.

Following that, the free parameters were extracted from the fitting process and they were used to define the functions describing the background and signal events separately. , the fitting parameters were used to split the function into two separate functions, one for the peak, which included the two Gaussians, and one for the background data, which included the quadratic part. Finally, the integral of the functions that fitted the signal data (f_S) for the opposite and same sign events, respectively, would give the number of events that originate from Z boson decays.

The results of the fitting process on the same sign dimuon data are presented below:

Degrees of freedom: 94
Chi^2: 127.927

Parameter	**Value**	**Uncertainty**
0	3332.6	187.318
1	-25.7801	1.40913
2	0.0480609	0.00364588
3	-137276	14088.6
4	79.3888	2.4941
5	27.0508	3.71605

Table 1: The fitting of a quadratic and a Gaussian function with free parameters on the same sign dimuon data. The parameters and their uncertainties as found through the fitting process.

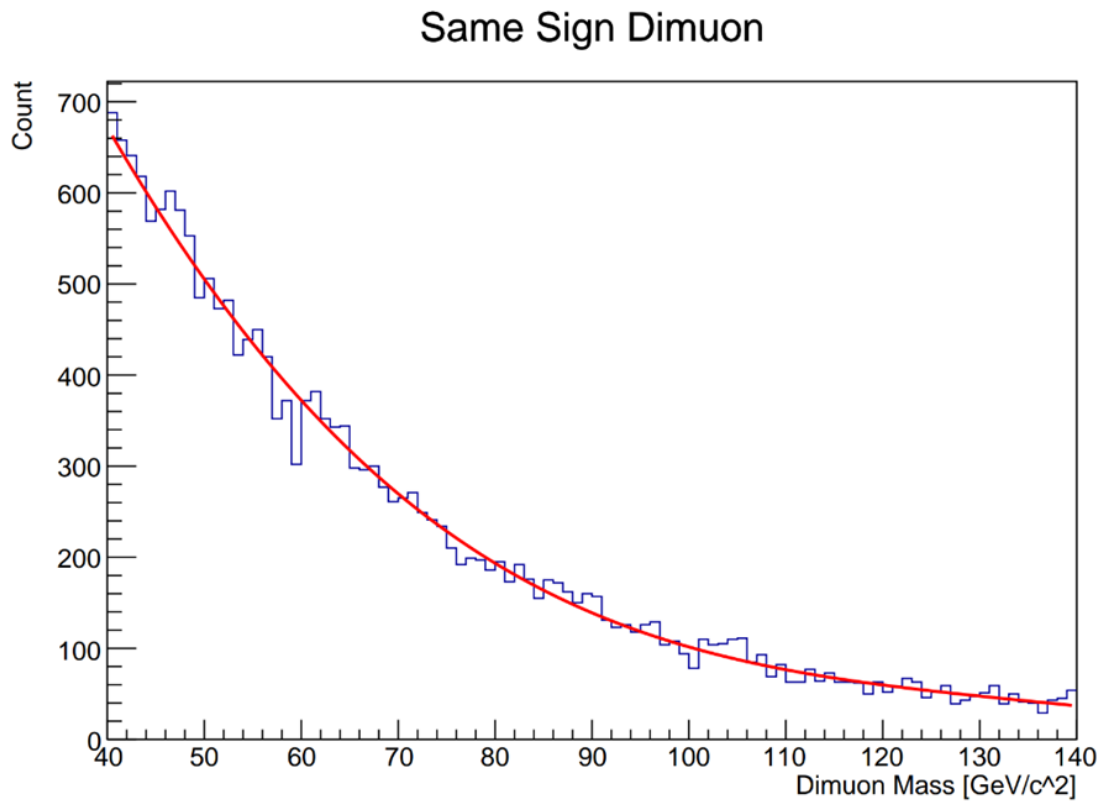

Figure 22: The function that fits the same sign dimuon data

The fit parameters extracted by the fitting process were of great importance because they were used as the initial guesses for the fit parameters of the background parts of the fit functions used for the same sign and opposite sign dielectron data, to help the fitting process. The values, of course, were different for each dataset.

For the opposite sign dielectron data, the function used for the fitting process was the following:

$$f_{OS} = [0] + [1]x + [2]x^2 + [3](1/[4])e^{-\frac{1}{2}(\frac{x-[5]}{[4]})^2} + [6](1/[7])e^{-\frac{1}{2}(\frac{x-[8]}{[7]})^2}$$

Once again, the numbers in square brackets represent free parameters. The fitting process gave the following:

Degrees of freedom: 91
Chi^2: 594.467

Parameter	Value	Uncertainty
0	1408.44	12.942
1	-9.26865	0.288467
2	0.000287017	0.00227793
3	37727.3	492.735
4	2.56241	0.0208731
5	90.273	0.0135804
6	23404.2	458.212
7	6.43295	0.0770218
8	88.5178	0.0586382

Table 2: The fitting of a quadratic and two Gaussian functions with free parameters on the opposite sign dielectron data. The parameters and their uncertainties as found through the fitting process are presented in the table.

Using the parameters from the fitting process, this fit function was split into the following two:

$$f_{b_{OS}} = [0] + [1]x + [2]x^2 \qquad \text{for the background data}$$

$$f_{s_{OS}} = [3](1/[4])e^{-\frac{1}{2}(\frac{x-[5]}{[4]})^2} + [6](1/[7])e^{-\frac{1}{2}(\frac{x-[8]}{[7]})^2} \qquad \text{for the signal data}$$

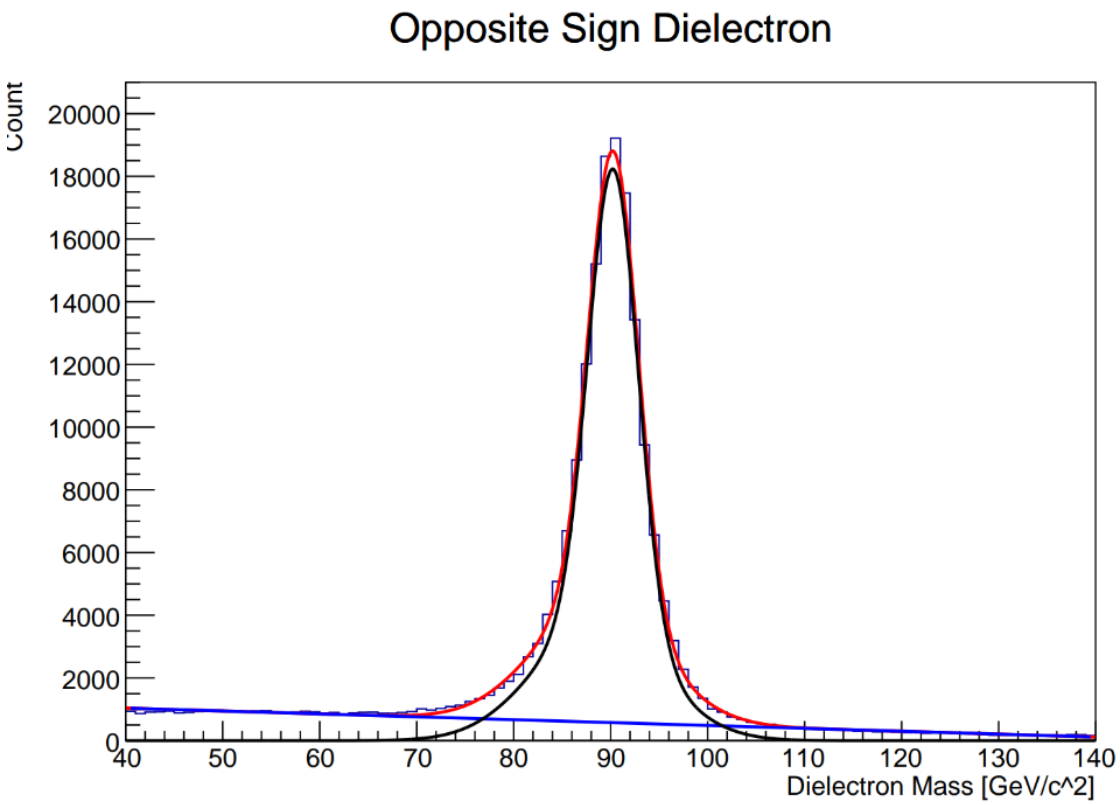

Figure 23: The fitting process for the opposite sign data. The red line represents the fit function f_{OS}, which was used for the fitting process. The blue line represents the $f_{b_{OS}}$ function, which was the quadratic part of f_{OS}. The black line represents the $f_{s_{OS}}$ function, which was the Gaussian part of f_{OS}.

The next step was to perform an integration for the $f_{S_{OS}}$ function to find the number of signal data in the resonance, N_{OS}. The range of integration was taken to be (68,110).

The same process was followed for the same sign dielectron data. The function used for the fitting process was as follows:

$$f_{SS} = [0] + [1]x + [2]x^2 + [3](1/[4])e^{-\frac{1}{2}(\frac{x-[5]}{[4]})^2}$$

The numbers in square brackets represent the free parameters.

Again, using the parameters extracted from the fitting process, this function was split into the following two:

$$f_{b_{SS}} = [0] + [1]x + [2]x^2 \qquad \text{for the background data}$$

$$f_{s_{SS}} = [3](1/[4])e^{-\frac{1}{2}(\frac{x-[5]}{[4]})^2} \qquad \text{for the signal data}$$

The results of the fitting process can be seen below:

Degrees of freedom: 94
Chi^2: 91.5779

Parameter	Value	Uncertainty	Relative uncertainty
0	1152.15	21.7686	0.02
1	-13.5288	0.494537	0.04
2	0.0438901	0.00259495	0.06
3	897.618	43.4223	0.05
4	3.56219	0.177844	0.05
5	89.7806	0.163593	0.002

Table 3: The fitting of the sum of a quadratic and a Gaussian function, f_{SS}, with free parameters on the same sign dielectron data. The free parameters and their uncertainties as found via the fitting process are presented in the table. The relative uncertainty of the parameters is also provided.

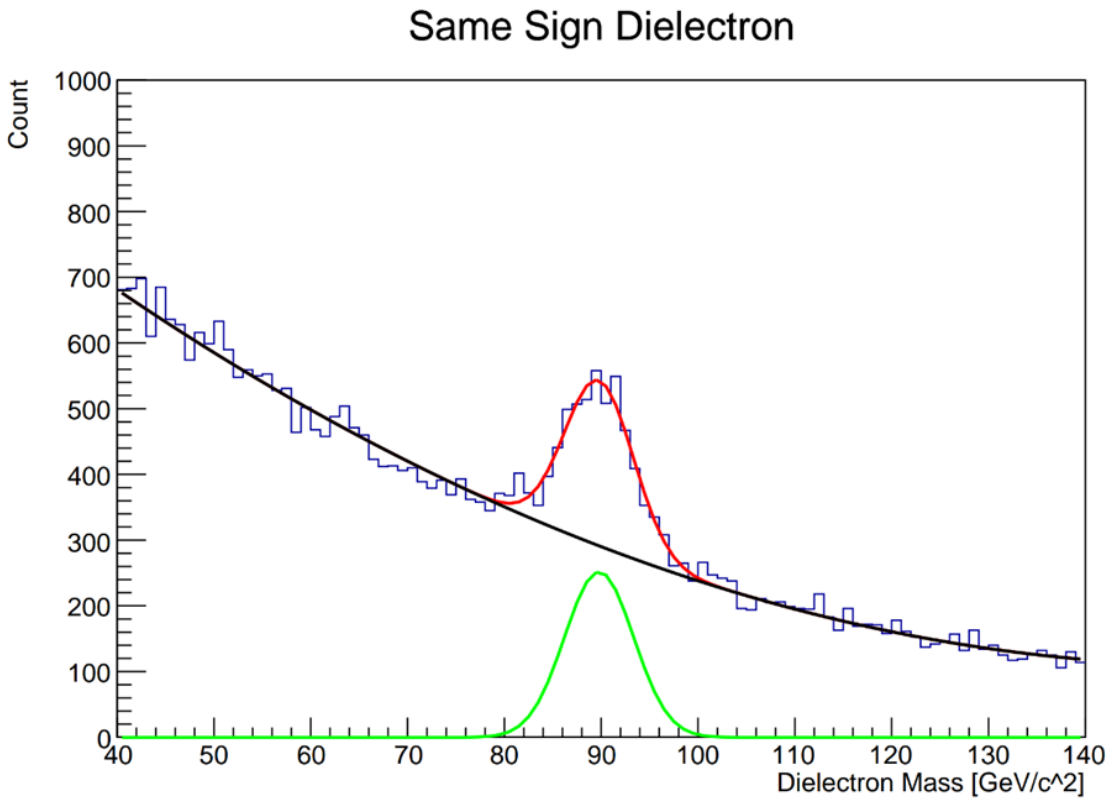

Figure 24: The fitting process for the opposite sign data. The red line represents the f_{SS} function, which was fitted to the data. The black line represents the $f_{b_{SS}}$ function, which was the quadratic part of f_{SS}. The green line represents the $f_{s_{SS}}$ function, which was the Gaussian part of f_{SS}.

An integration was performed on the signal function, f_S, in the range of (68,110) to find the number of signal events, N_{SS}, that are suspected to have originated from Z^0 boson decays.

The statistical uncertainty of the integrals was due to the uncertainties given by the fitting process. Specifically, it was found via error propagation on the integral using partial derivatives on the parameters of the fitting process.

3.2.2 1^{st} Systematic uncertainty

The next step was to approach the systematic uncertainty of this method. To do so, we found the relative uncertainties for each parameter of the signal function as calculated by the fitting process, as shown in Table 3.

The parameters with the most significant uncertainties were [3] and [4] for the same-sign events. Variations in parameter [3] change the amplitude of the curve, while variations in [4] affect the width and height. For those parameters, we took their respective uncertainties and made independent, correlated, and anti-correlated up or down variations of one standard deviation and checked how this affected the signal events.

In total, we made 8 different variations, as stated in the table below:

Variation	**Parameter [3]**	**Parameter [4]**
1	[3]+$\sigma_{[3]}$	[4]+$\sigma_{[4]}$
2	[3]+$\sigma_{[3]}$	[4]-$\sigma_{[4]}$
3	[3]-$\sigma_{[3]}$	[4]+$\sigma_{[4]}$
4	[3]-$\sigma_{[3]}$	[4]-$\sigma_{[4]}$
5	[3]	[4]-$\sigma_{[4]}$
6	[3]	[4]+$\sigma_{[4]}$
7	[3]-$\sigma_{[3]}$	[4]
8	[3]+$\sigma_{[3]}$	[4]

Table 4: The different variations of the signal function that arise from the different combinations of the parameters with the biggest significant uncertainties, when one standard deviation is added or subtracted from them (where one standard deviation is the uncertainty of each parameter as given by the fitting process).

The aim was to observe how the integral, and consequently the number of N_{SS} differentiated for every variation. According to the error propagation formula extracted in Section 3.1.3, the uncertainty of the N_{SS} dominates the uncertainty of the probability rather than the uncertainty of the N_{OS}, since the N_{OS} is much larger than the N_{SS}. Therefore, this process was not performed for the opposite sign data, since the error coming from

that would have an insignificant impact on the overall uncertainty of the probability.

The graphic results of each variation are shown compared to the original signal function.

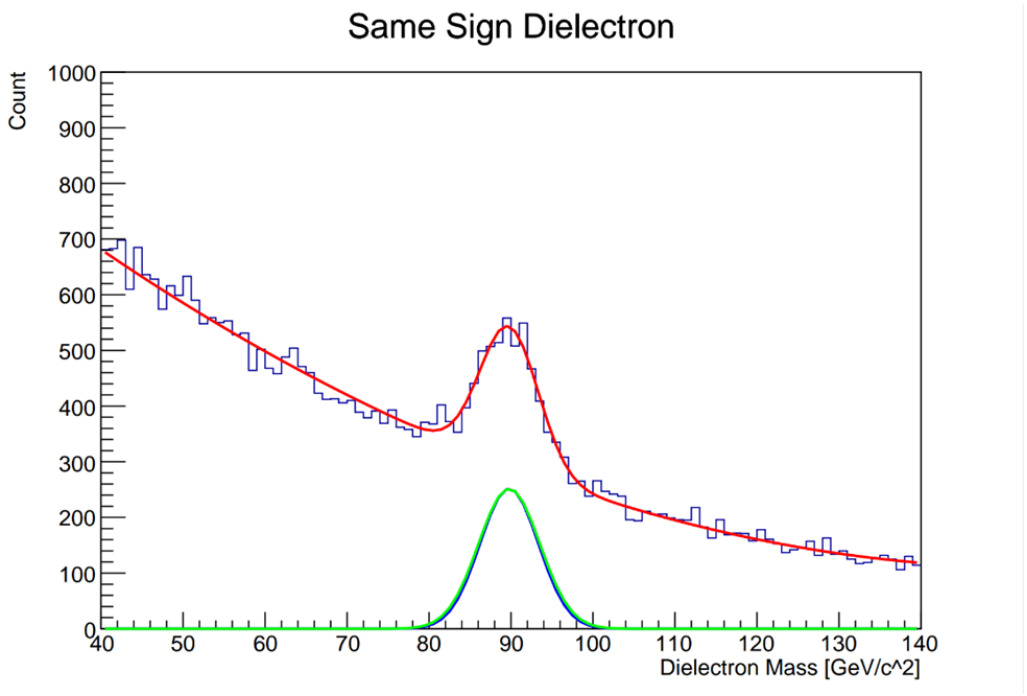

Figure 25: Variation 1. The blue function represents the original signal function, while the green represents the variation.

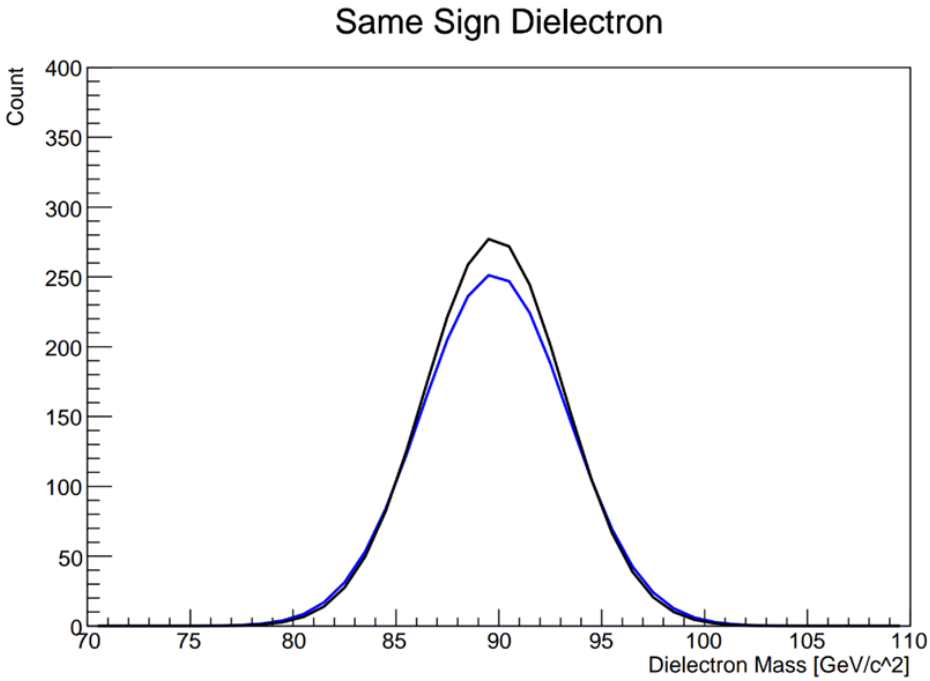

Figure 26: Variation 2. The blue function represents the original signal function, while the black function represents the variation.

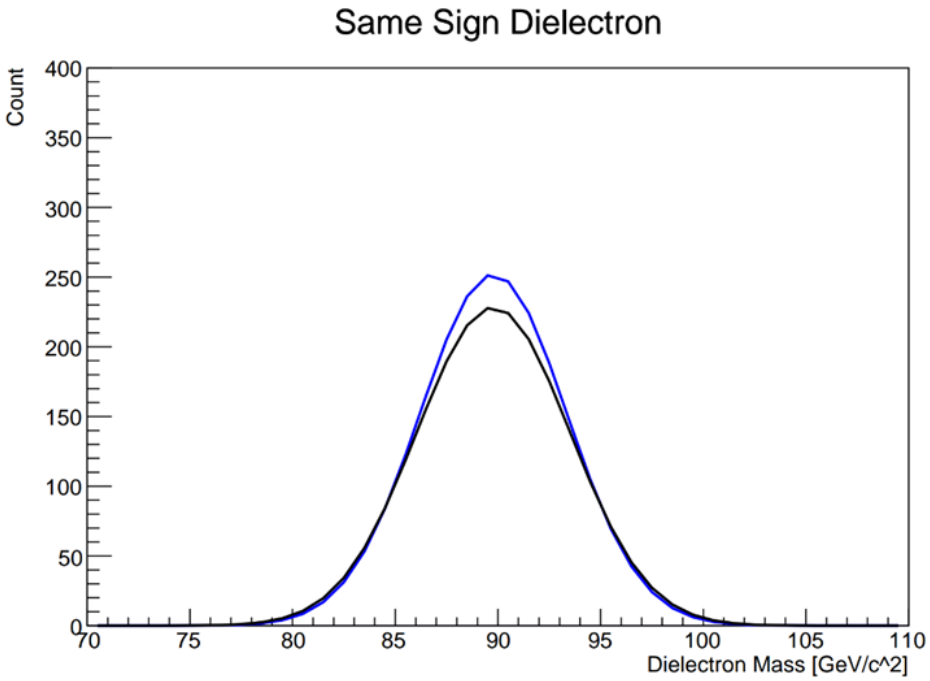

Figure 27: Variation 3. The blue line represents the initial signal function, and the black one represents the variation.

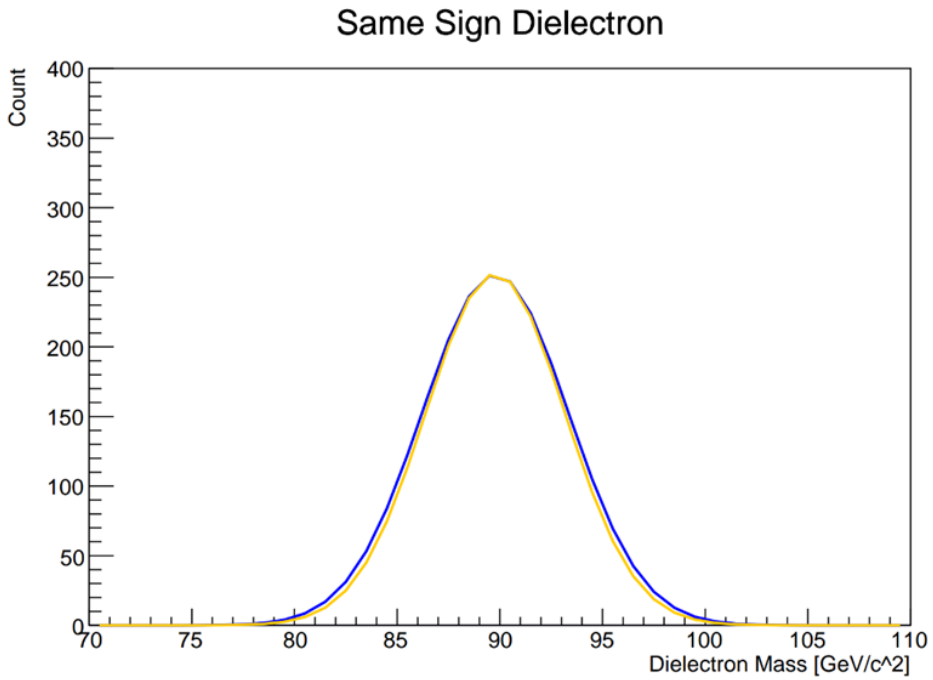

Figure 28: Variation 4. The blue line represents the initial function, and the yellow one represents the variation.

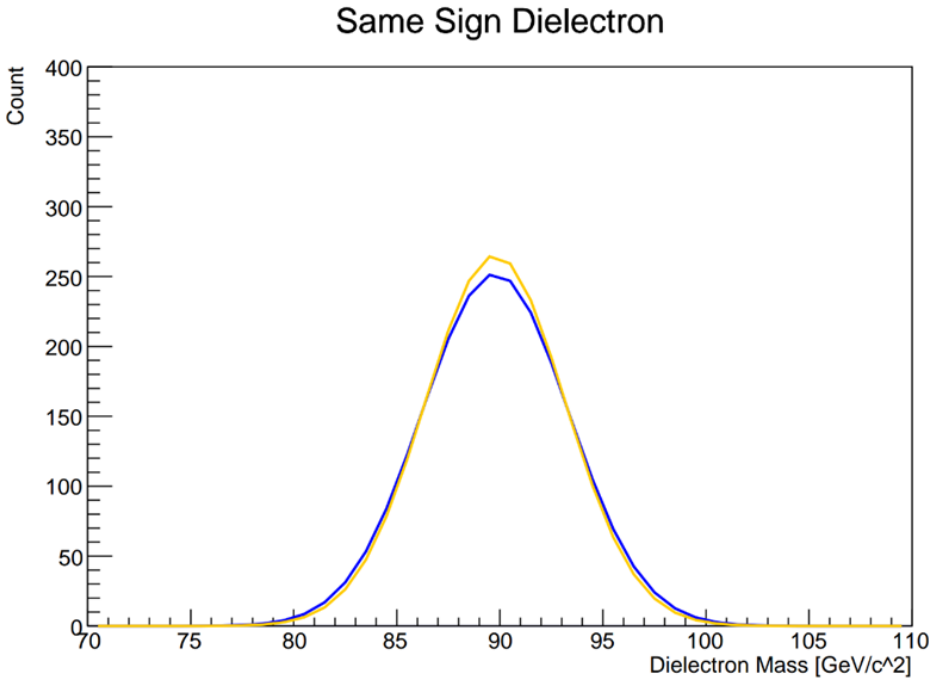

Figure 29: Variation 5. The blue line is the function that uses the original parameters, while the yellow one is the variation.

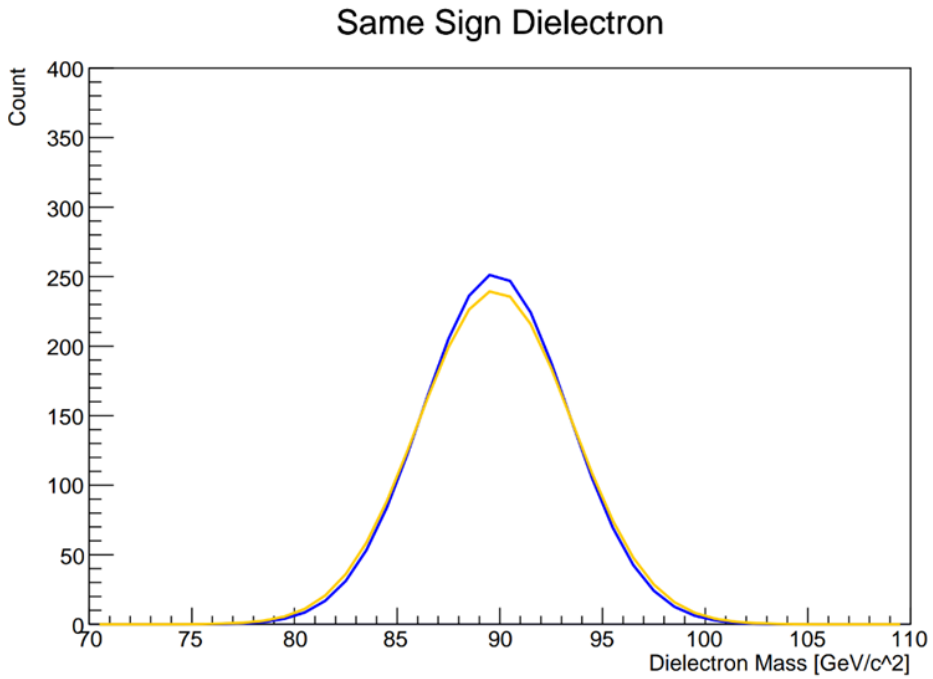

Figure 30: Variation 6. The blue line is the function that uses the original parameters, while the yellow one is the variation.

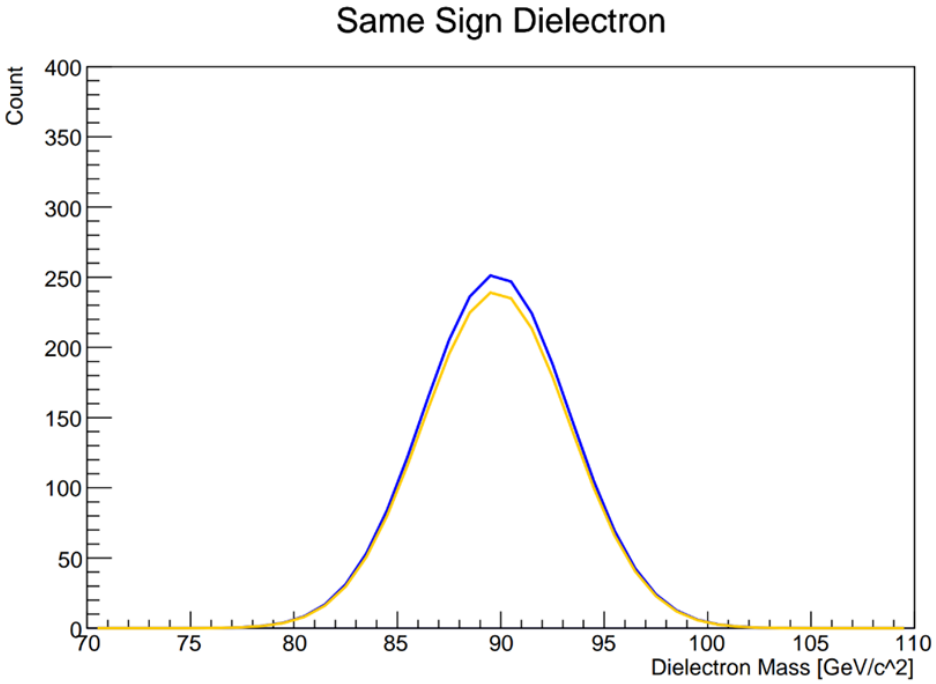

Figure 31: Variation 7. The blue line is the function that uses the original parameters, while the yellow one is the variation.

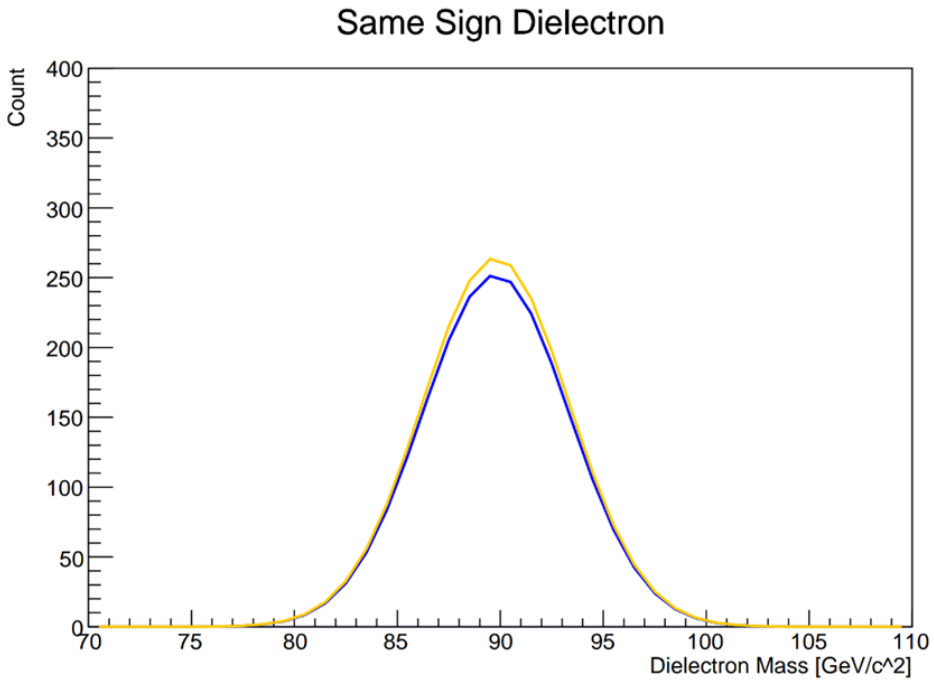

Figure 32: Variation 8. The blue line is the function that uses the original parameters, while the yellow one is the variation.

3.2.3 2^{nd} Systematic uncertainty

In the final step of this analysis, we aimed to calculate the systematic uncertainty due to the choice of function that represents the background processes. As mentioned in Section 3.2.1, the choice of the background function should not have an important impact on the extraction of the probability, since it is completely arbitrary. So, we aimed to study the impact the background function has on the calculation of the probability.

First, we performed a summation of every bin content in the resonance region, both for the opposite sign and the same sign data. This summation gave the total number of both the background and signal events in the peak region, for each data set. The next step was to find the number of background events and subtract it from the total number of events, to get N_{OS} and N_{SS}, respectively.

To do so, the data in the resonance region were masked, so there was only background data left in each data set. Subsequently, three new functions were defined for each set, which were randomly chosen. However, they were chosen to have a relevant form to that of the background data distribution.

For the same sign data, the following functions were defined:

$$f_{1_{SS}} = [0] + [1](1/x) + [2]x^2$$

$$f_{2_{SS}} = [0] + [1]x + [2]x^2 + [3](1/x)$$

$$f_{3_{SS}} = [0] + [1]x + [2]logx$$
$$f_{4_{SS}} = [0] + [1]x + [2]x^2$$

These were fitted to the background data. The results of the fitting process are presented below:

FUNCTION	**PARAMETER**	**VALUE**	**UNCERTAINTY**
$f_{1_{SS}}$	0	16.965	28.8174
	1	29215.5	1393.55
	2	-0.0062944	0.00112152
Degrees of freedom = 55			
Chi^2 **= 69.4297**			
$f_{2_{SS}}$	0	806.921	194.281
	1	-9.44596	2.29743
	2	0.0287821	0.0086046
	3	9073.67	5093.21
Degrees of freedom = 54			
Chi^2 **= 52.5277**			
$f_{3_{SS}}$	0	3066.03	175.162
	1	2.24858	0.607374
	2	-1524.63	119.68
Degrees of freedom = 55			
Chi^2 **= 54.7456**			
$f_{4_{SS}}$	0	1149.97	25.8285
	1	-13.3828	0.628602
	2	0.0428827	0.00337593
Degrees of freedom = 55			
Chi^2 **= 55.7011**			

Table 5: The outcomes of the fitting process for each background function of the same sign dielectron data set.

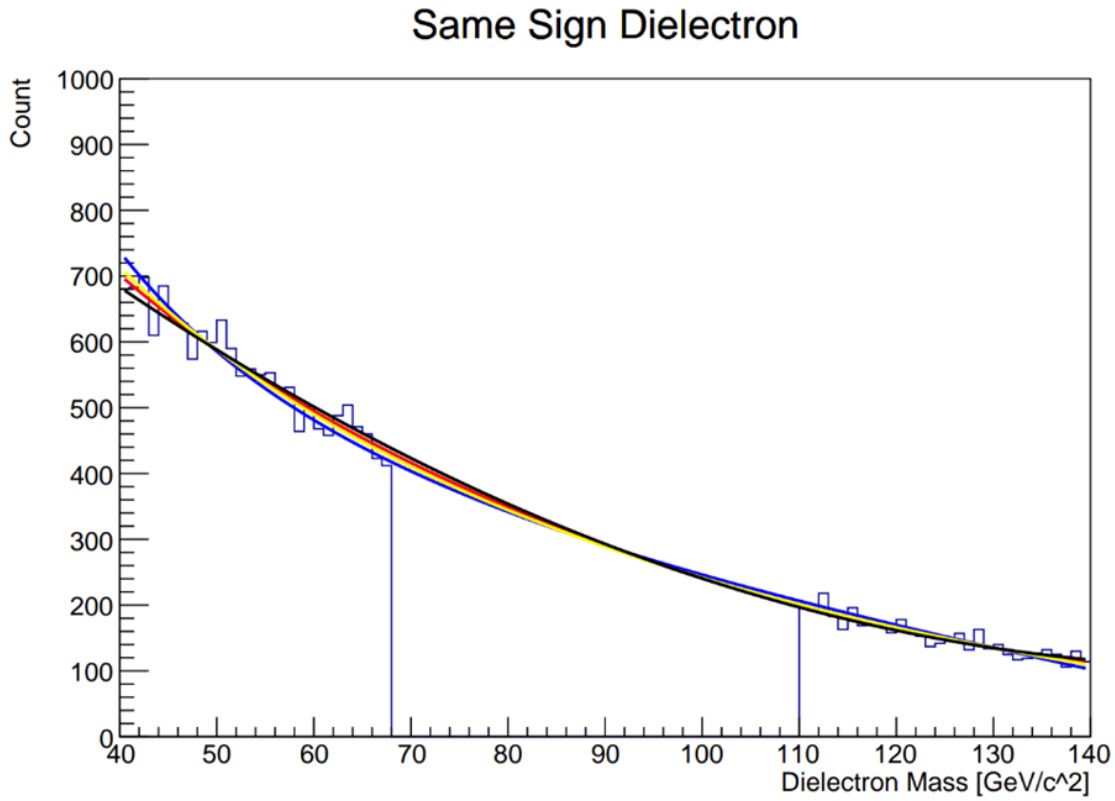

Figure 33: The background functions defined above are fitted to the masked data. The blue line corresponds to $f_{1_{SS}}$, the red line corresponds to $f_{2_{SS}}$, the yellow line corresponds to $f_{3_{SS}}$, and the black line corresponds to $f_{4_{SS}}$.

Even though it has been stated that the uncertainty due to the opposite sign data set can be negligible, we still did the same process for that set, only for visualisation purposes. This process was not used in later calculations. For the opposite sign data, the following background functions were defined:

$$f_{1_{OS}} = [0] + [1](1/x) + [2]x^2$$

$$f_{2_{OS}} = [0] + [1]x + [2]x^2 + [3](1/x)$$

$$f_{3_{OS}} = [0] + [1]x + [2]logx + [3](1/x)$$

$$f_{4_{OS}} = [0] + [1]x + [2]x^2$$

Those functions were also fitted to the masked background opposite-sign data. The results of the fitting process are presented below:

FUNCTION	PARAMETER	VALUE	UNCERTAINTY
$f_{1_{OS}}$	0	1031.44	314.886
	1	5344.27	16854.8
	2	-0.0570674	0.0126934
Degrees of freedom = 15			
Chi^2 **= 12.4591**			
$f_{2_{OS}}$	0	4788.55	246.557
	1	-48.9827	2.92034
	2	0.145661	0.0109312
	3	-88000.3	6418.75
Degrees of freedom = 54			
Chi^2 **= 56.3566**			
$f_{3_{OS}}$	0	1436.61	68.5707
	1	-17.7651	0.477766
	2	573.222	33.5966
	3	-28982.6	3014.22
Degrees of freedom = 14			
Chi^2 **= 10.8053**			
$f_{4_{OS}}$	0	1318.71	266.065
	1	-4.52908	6.41484
	2	-0.0359292	0.0356338
Degrees of freedom = 55			
Chi^2 **= 12.0612**			

Table 6: The outcomes of the fitting process for each background function of the opposite sign dielectron data set.

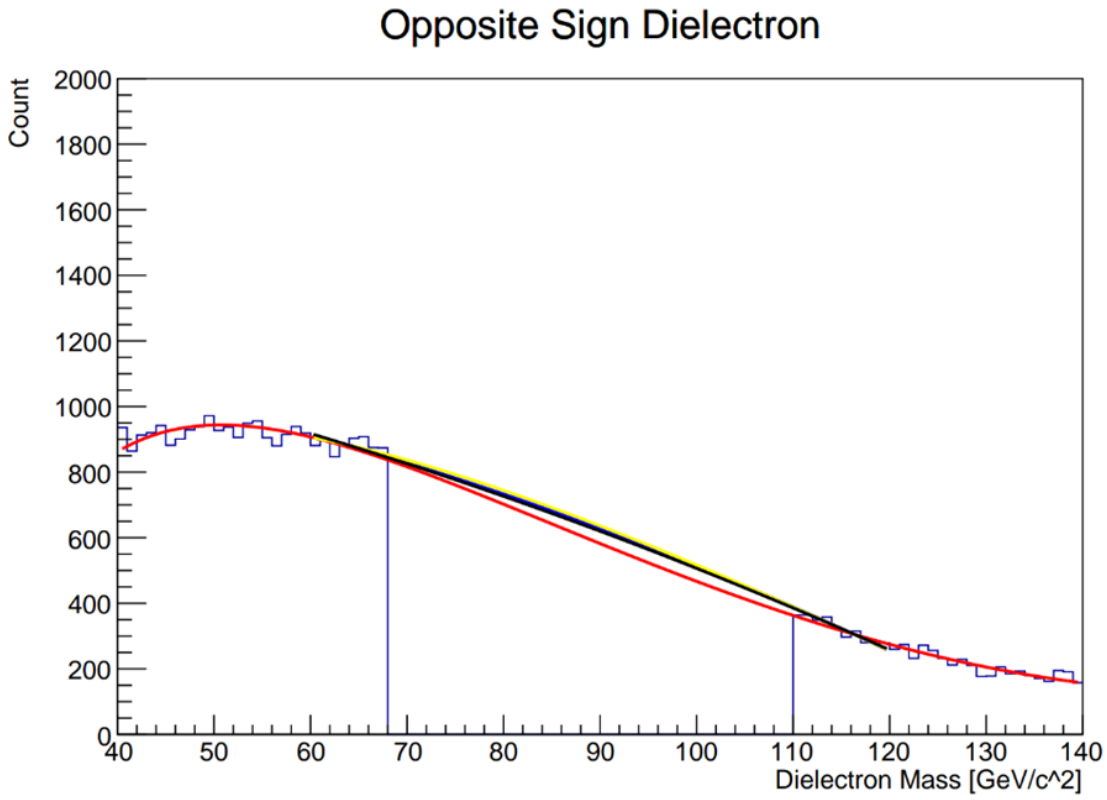

Figure 34: The opposite sign dielectron data set, where the data in the peak region are masked. The four background functions defined above are fitted as follows: $f_{1_{OS}}$ is represented by the blue line, $f_{2_{OS}}$ is represented by the red line, $f_{3_{OS}}$ is represented by the yellow line, and $f_{4_{OS}}$ is represented by the black line.

Thereafter, integration was performed for each function of each data set in the peak region. The integral of each function gave the number of background events predicted by each function inside the peak region. So, by subtracting the integral from the sum of the bin contents, the number of signal events each function predicts was extracted.

4 Results

4.1 Individual results

In this section, the results of each methodology followed previously will be presented.

4.1.1 Results of the calculation of the statistical probability and the statistical uncertainty

Data set	Integrated function	Range of integration	Integral	Uncertainty
Opposite sign	Signal	(68,110)	153168	449
Same sign	Signal	(68,110)	2250	109

Table 7: The results of the integration of the signal functions for the calculation of the probability, and their uncertainties calculated via error propagation.

The probability was calculated based on the formula extracted in Section 3.1.3 (statistical probability), and the uncertainty was found using the error propagation formula, also found in Section 3.1.3 (statistical uncertainty).

$$\mathbf{P=0.0072 \pm 0.0003}$$

4.1.2 Results of the calculation of the 1^{st} systematic uncertainty

Following the variation functions, the results of the integrations and the probability of mismeasurement calculated for each variation are presented in the table below:

VARIATION	RANGE OF INTEGRATION	INTEGRAL (N_{SS})	PROBABILITY
1	(68,110)	2359	0.0076
2		2359	0.0076
3		2141	0.0069
4		2141	0.0069
5		2250	0.0072
6		2250	0.0072
7		2141	0.0069
8		2359	0.0076

Table 8: The integral for each variation and the corresponding probability. For the probability calculation, the opposite sign data were taken to be the same as in Table 4.1, while the same sign data were taken to be the integral of each variation. The probability was calculated based on the formula extracted in Section 3.1.3.

The systematic uncertainty was calculated by finding the largest distance between the central value of the probability calculated in section 4.1.1, and the probabilities calculated for the different variations of the signal function:

$$\mathbf{P=0.0072 \pm 0.0004}$$

4.1.3 Results of the calculation of the 2^{nd} systematic uncertainty

The integration results of all the background functions and the bin content summation in the peak region (68,110) are presented in the table below:

Data set	Sum of bin contents	Background function	Integral	Signal events
Same sign	14894	$f_{1_{SS}}$	12631	2263
		$f_{2_{SS}}$	12699	2195
		$f_{3_{SS}}$	12613	2281
		$f_{4_{SS}}$	12805	2089
Opposite sign	178617	$f_{1_{OS}}$	26553	152064
		$f_{2_{OS}}$	25054	153563
		$f_{3_{OS}}$	26825	151792
		$f_{4_{OS}}$	26281	152336

Table 9: The integral for each background function in the peak region. The number of signal events as predicted by each function was calculated by subtracting the integral from the sum of the bin contents in the peak region.

Because the number of opposite-sign events is much greater than the number of same-sign events, the uncertainty of the probability calculation due to the opposite sign events will be negligible, as it is obvious from the corresponding equation extracted in Section 3.1.3. So, to find the statistical uncertainty caused by the background function selection, the

number of opposite-sign events was taken from Table 7, and the probability was calculated for every number of same sign signal events presented in Table 9.

Same sign background function	**Probability**
$f_{1_{SS}}$	0.0073
$f_{2_{SS}}$	0.0071
$f_{3_{SS}}$	0.0073
$f_{4_{SS}}$	0.0067

Table 10: The probability as calculated for the number of signal events that each same-sign signal function predicts. For the opposite sign events, the value was extracted from Table 7.

The systematic uncertainty was calculated by finding the largest distance between the central value of the probability as calculated in section 4.1.1, and the probabilities calculated for the different background functions of the same sign dielectron data set:

$$\mathbf{P=0.0072 \pm 0.0005}$$

4.2 Combination of the results

By taking into consideration the statistical and systematic uncertainties calculated in this analysis, the probability can be written as:

$$\mathbf{P=0.0072 \pm 0.0003 \pm 0.0004 \pm 0.0005}$$

As seen by comparing the uncertainties, the choice of the background function generated the largest uncertainty. This indicates that the fitting process itself does not affect the probability as much as the function selected for the fitting process. However, it has to be noted that the fitting of the background functions did not include the signal function and the data inside the peak region, which, of course, contributes to the differentiation of the probability. The fourth background function was taken to be a quadratic one, as it was when the fitting happened for the whole range of data and with the signal function. The probabilities calculated for each case differ by 0.0005, which indicates that indeed the exclusion of the peak region contributes to the uncertainty.

The probability can also be written as:

$$\mathbf{P=0.0072 \pm 0.0007}$$

if the uncertainties are added in quadrature.

5 Conclusions

As it can be understood from this thesis, electron charge misidentification is a real problem for analyses at the CMS experiment. There are processes, such as Z^0 boson decays, where this misidentification can be predicted, and it can act as a guideline for the measurement of the probability of such misidentifications occurring. However, during other events for which there are no predictions for their outcome, it is not possible to know whether the measurement is truthful or not. All there is to know is the mismeasurement probability.

Even though it was not presented in this thesis, the probability of charge sign mismeasurement is different between the barrel and the endcap region. This is a perfectly logical conclusion, considering the following: the barrel region is at $\eta < 1.4$, while the endcap region exceeds at $\eta > 1.4$. The value of the transverse momentum itself does not play a crucial role, since whether the particle traverses the endcap or the barrel region depends on its longitudinal momentum as well. What affects the probability is the amount of material the electron goes through in each region. Penetrating larger amounts of material means more opportunities for Bremsstrahlung radiation, which triggers false tracking.

Obviously, the more hits there are inside the tracker, the more accurate the trajectory reconstruction will be. This would require more material in the tracker to increase the number of hits. Nevertheless, even though tracker materials have minimal interactions with the particles, more material would lead to more Bremsstrahlung radiation, and eventually, to false tracking.

Finally, what we aimed to measure was the incorrectly measured electrons. Usually, measurements are also simulated, and the simulations are compared to the actual data. If those two do not coincide, then there is a factor that causes the data not to follow the prediction made by the simulation, and it needs to be studied. However, since the mismeasurements in this case are false measurements, they cannot be simulated very well. The Monte-Carlo method is used for simulations; however, when there is a difference between the simulation-generated signal and the actual data signal, it is not easy to distinguish whether the simulation was wrong or if this has a physical meaning.

Bibliography

[1]CERN. (2019). The Large Hadron Collider. Cern; CERN. https://home.cern/science/accelerators/large-hadron-collider

[2]CERN. (2025). Facts and Figures about the LHC — CERN. Home.cern; CERN. https://home.cern/resources/faqs/facts-and-figures-about-lhc

[3]Halkiadakis, E.(2009).Proceedings for TASI 2009 Summer School on "Physics of the Large and the Small" : Introduction to the LHC experiments. Ar5iv. https://ar5iv.labs.arxiv.org/html/1004.5564

[4]CERN. (2019, July 16). Accelerators. CERN. https://home.cern/science/accelerators

[5]Vogt, R. (2000). Physic of the nucleon sea quark distributions. Progress in Particle and Nuclear Physics. arXiv. https://doi.org/10.1016/s0146-6410(00)90012-7

[6]F. Rimondi. (2004). Soft and hard interactions in p -$\bar{p}$ collisions at $\sqrt{s}$ = 1800 and 630 GeV. Physics of Atomic Nuclei, 67(1), 130–139. https://doi.org/10.1134/1.1644016

[7]Maurice, J. (1984, July). Hard collisions and jets(CERN-TH-3953). CERN Document Server. https://cds.cern.ch/record/154775?ln=el

[8]The Z and the b — CMS Experiment. (2016). Cms.cern. https://cms.cern/news/z-and-b

[9]Di Lella, L., & Rubbia, C. (2015). The Discovery of the W and Z Particles. Advanced Series on Directions in High Energy Physics, 137–163. https://doi.org/10.1142/9789814644150_0006

[10]Neutelings, I. (n.d.). CMS coordinate system. https://tikz.net/axis3d_cms/

[11]The Z and the b — CMS Experiment. (2016). Cms.cern. https://cms.cern/news/z-and-b

[12]Tracking — CMS Experiment. (2025, March 17). Cern.ch. https://cmsexperiment.web.cern.ch/detector/identifying-tracks

[13]Klyukhin, V. (2021). Design and Description of the CMS Magnetic System Model. Symmetry, 13(6), 1052. https://doi.org/10.3390/sym13061052

[14]Calorimeters Lecture 4 Part 4 Content [Lecture slides]. (n.d.). CERN. Retrieved March 2025, from https://indico.cern.ch/event/1209475/

contributions/5231414/attachments/2579513/4448714/2020-Lecture-4-4-Calo
pdf

[15]17.4.3 Bremsstrahlung Radiation. (2023, January 10). XmPhysics.
https://xmphysics.com/2023/01/10/17-4-3-bremsstrahlung-radiation/

[16]Gupta, M. (2010, July 22). Calculation of radiation length in materials(PH-EP-Tech-Note-2010-013). CERN Document Server.
https://cds.cern.ch/record/1279627?ln=el

[17]Cockerill, D.(2016, May 4). Introduction to Calorimeters. Southampton Lecture. https://indico.cern.ch/event/518474/contributions/1198681/attachments/1267581/1877122/Calorimetry-lecture-to-Southampton-pdf

[18]Krivkova, P., & Leitner, R. (1999, March 15). Measurement of the interaction length of pions and protons in the TileCal Module 0(ATL-TILECAL-99-007). CERN Document Server. https://cds.cern.ch/record/683812

[19]Sefkow, F., & Simon, F. (2021). Calorimeters. Handbook of Particle Detection and Imaging, 1–33.
https://doi.org/10.1007/978-3-319-47999-6_53-1

[20]Lecture Detectors and analysis methods. (2018). 06 Hadronic Calorimeters. Desy.de. Retrieved March 2025, from
https://www.desy.de/~schleper/lehre/Det_Dat/SS_2018/

[21]Detecting Muons — CMS Experiment. (n.d.). Cms.cern.
https://cms.cern/detector/detecting-muons

[22]Student Sessions. (2005, August 10). tranpsarensies-6.Indico.
https://indico.cern.ch/event/424458/

[23]Barney, D.(27 November 2003).Presentation for public - Introduction to CMS for CERN guides - CERN Document Server. (2018). Cern.ch.
https://cds.cern.ch/record/2629323/files

[24]Detecting Muons. (2011, November 23). Cms.cern.
https://cms.cern/book/export/html/1203

[25]Introduction To End Cap Muon Group. (n.d). Ufl.edu. Retrieved March 2025, from
https://www.phys.ufl.edu/cms/introendcap.html

[26]Muon detectors. (2024, December 10). Github.io. Retrieved March 2025, from

https://cms-opendata-workshop.github.io/workshop2024-lesson-cms-detecto
instructor/06-muon

[27]Superconducting magnet. (2024, December 10). Github.io. Retrieved March 2025, from
https://cms-opendata-workshop.github.io/workshop2024-lesson-cms-detecto
instructor/05-magnet.html

[28]Layter, J. G. (1997). The CMS muon project : Technical Design Report.(CERN. CERN-LHCC-97-032 ; CMS-TDR-3). In CERN Document Server.
https://cds.cern.ch/record/343814

[29]Dubbert, J. (31 January 2001). EDIT 2011-DriftTubes. Indico CERN.
https://indico.cern.ch/event/124394/contributions/98538/subcontribution
9360/attachments/74603/107004/edit_2011_drift_tubes_WriteUp.pdf

[30]Chauhan, S. S., & Choudhary, B. C. (2009). Search for Quark Compositeness at $\sqrt{s} = 14$ TeV at the Large Hadron Collider.
https://www.researchgate.net/publication/48410310_Search_for_Quark_
Compositeness_at_sqrts_14_TeV_at_the_Large_Hadron_Collider

[31]Bakhet, N., Yu, K. M., & Hussein, T. (2015). Neural Networks Search for Charged Higgs Boson of Two Doublet Higgs Model at the Hadrons Colliders. ArXiv.org. https://arxiv.org/abs/1507.06547

[32]Cathode Strip Chambers — CMS Experiment. (n.d.). Cms.cern.
https://cms.cern/detector/detecting-muons/cathode-strip-chambers

[33]Muon Drift Tubes — CMS Experiment. (n.d.). Cms.cern.
https://cms.cern/detector/detecting-muons/muon-drift-tubes

[34]Hauser, J. (1996). Cathode strip chambers for the CMS endcap muon system. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 384(1), 207–210. https://doi.org/10.1016/S0168-9002(96)00905-9

[35]Argyropoulos, T., Assamagan, K. A., Benedict, B. H., Chernyatin, V., Cheu, E., Deng, J., Gordeev, A., Gough Eschrich, I., Grachev, V., Johns, K., Junnarkar, S., Kandasamy, A., Kaushik, V., Khodinov, A., Lankford, A., Lei, X., Murillo Garcia, R., Nikolopoulos, K., O'Connor, P., & Parnell-Lampen, C. L. (2008). Cathode strip chambers in ATLAS : Installation, commissioning and in situ performance. 2008 IEEE Nuclear Science Symposium Conference Record, 2819–2824.

https://doi.org/10.1109/nssmic.2008.4774958

[36]Elashri, M. (2017, June 14). Strip hit resolution of CMS Tracker analysis. https://doi.org/10.13140/RG.2.2.11136.84480

[37]Collaboration, C. (2010). Performance of the CMS cathode strip chambers with cosmic rays. Journal of Instrumentation, 5(03), T03018–T03018. https://doi.org/10.1088/1748-0221/5/03/t03018

[38]Masetti, G. (2010, May 21). Muon studies at CMS (CMS-CR-2010-063). CERN Document Server.https://cds.cern.ch/record/1285535

[39]Resistive Plate Chambers — CMS Experiment. (n.d.). Cms.cern. https://cms.cern/detector/detecting-muons/resistive-plate-chambers

[40]Detector Overview. (n.d.). Cms.cern. Retrieved in March 2025, from
https://cms.cern/book/export/html/1193

[41]Sirunyan, A. M., Tumasyan, A., Adam, W., Bergauer, T., Dragicevic, M., Escalante Del Valle, A., Frühwirth, R., Jeitler, M., Krammer, N., Lechner, L., Liko, D., Mikulec, I., Pitters, F. M., Rad, N., Schieck, J., Schöfbeck, R., Spanring, M., Templ, S., Waltenberger, W., & Wulz, C.-E. . (2021). Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. Journal of Instrumentation, 16(05), P05014. https://doi.org/10.1088/1748-0221/16/05/p05014

[42]Hilderth, M. (August 16-27 2010) *An Introduction to Charged Particle Tracking* [Summer School Presentation]. Hadron Collider Physics Summer School. Fermilab. https://indico.fnal.gov/event/3532/contributions/82029/attachments/52078/62335/Tracking_Lecture_3.pdf

[43]Tracking in a nutshell — Acts documentation. (2025). Readthedocs.io. Retrieved March 2025, from https://acts.readthedocs.io/en/latest/tracking.html#particle-propagation