MEASUREMENT OF THE INCLUSIVE $Z \to ee$ PRODUCTION CROSS SECTION IN PROTON-PROTON COLLISIONS AT $\sqrt{s} = 7$ TeV AND $Z \to ee$ DECAYS AS STANDARD CANDLES FOR LUMINOSITY AT THE LARGE HADRON COLLIDER

Jeremy Scott Werner

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF

PHYSICS

Adviser: Valerie Halyo

November 2011

To my father, Captain Paul Anthony Werner.

The work described in this dissertation has been published
in the following articles and presented at the following conferences:

arXiv:1107.4789 [hep-ex] (CMS-PAS-EWK-10-005, submitted to JHEP)

arXiv:1106.5393 [hep-ex]

10.1007/JHEP01(2011)080 (arXiv:1012.2466 [hep-ex])

CMS-PAS-EWK-10-002

CMS-PAS-EWK-10-004

Lepton Photon 2011: XXV International Symposium on Lepton Photon Interactions at High Energies, August 2011, Mumbai, Maharashtra, India

Rencontres de Moriond on Electroweak Interactions and Unified Theories, March 2011, La Thuile, Valle d'Aosta, Italy

LHC Lumi Days, January 2011, CERN, Geneva, Switzerland

35th International Conference on High Eenergy Physics, July 2010, Paris, France

American Physical Society Meeting, February 2010, Washington, D.C., USA

Abstract

This thesis comprises a precision measurement of the inclusive $Z \to ee$ production cross section in proton-proton collisions provided by the Large Hadron Collider (LHC) at a center-of-mass energy of $\sqrt{s} = 7$ TeV and the absolute luminosity based on $Z \to ee$ decays. The data was collected by the Compact Muon Solenoid (CMS) detector near Geneva, Switzerland during the year of 2010 and corresponds to an integrated luminosity of $\int \mathcal{L} dt = 35.9 \pm 1.4$ pb^{-1}.

Electronic decays of Z bosons allow one of the first electroweak measurements at the LHC, making the cross section measurement a benchmark of physics performance after the first year of CMS detector and LHC machine operations. It is the first systematic uncertainty limited $Z \to ee$ cross section measurement performed at $\sqrt{s} = 7$ TeV. The measured cross section pertaining to the invariant mass window $M_{ee} \in (60, 120)$ GeV is reported as $\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee) = 997 \pm 11(\text{stat}) \pm 19(\text{syst}) \pm 40(\text{lumi})$ pb, which agrees with the theoretical prediction calculated to NNLO in QCD.

Leveraging $Z \to ee$ decays as "standard candles" for measuring the absolute luminosity at the LHC is examined; they are produced copiously, are well understood, and have clean detector signatures. Thus the consistency of the inclusive $Z \to ee$ production cross section measurement with the theoretical prediction motivates inverting the measurement to instead use the $Z \to ee$ signal yield to measure the luminosity. The result, which agrees with the primary relative CMS luminosity measurement calibrated using Van der Meer separation scans, is one of the most precise absolute luminosity measurements performed to date at a hadron collider and the first based on a physics signal at the LHC.

Acknowledgements

I am forever indebted to my family for their enduring love and support: My wife Vanessa; father Paul; mother Dorothy; stepmother Connie; siblings Paul, John, Shannon, and Ryan; and all other family members, including my uncle Bob who needs recognition for his editing assistance. I owe much gratitude to my adviser Valerie Halyo for her steady guidance, deep physical insight, and infectious enthusiasm. I thank Dan Marlow for his support and direction, as well as all of the other outstanding professors at Princeton University with whom I've been fortunate enough to interact, including James Olsen for serving on my thesis committee. Likewise, I thank Paul Langacker of the Institute for Advanced Study for serving on my committee. I am also grateful to all of the excellent professors at UCLA who provided me with a high-caliber undergraduate physics foundation—especially Bob Cousins, whose continuing mentorship I've benefited greatly from. Thank you to Dmitry Bandurin, Jeff Berryhill, and Kalanand Mishra from whom I've learned a great deal and had long lasting and profitable collaborations. I'd like to thank Helmut Burkhardt and Simon White of the LHC team: Together we made Van der Meer scans at the LHC a reality. This research would not have been possible without CERN, Fermilab, the LHC team, the CMS collaboration, and all supporting agencies including the DOE: I am very much grateful for them. Finally, I must pay respect to my dear friend Aaron Cornwell, whose kind soul is sorely missed in this world.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Standard Model

The Standard Model (SM) of elementary particle physics is the framework describing matter and its interactions at the fundamental level. It was developed throughout the twentieth century, and is accepted as the current description of particle physics. This framework is a quantum field theory in which quantum mechanics is combined with special relativity. It includes three out of four of the known forces of nature: electromagnetic, weak nuclear, and strong nuclear.[1] In particular, the standard model combines quantum chromodynamics (QCD) [9] with the electroweak theory [10] to describe nature using particles called quarks and leptons, which are fermions constituting matter, interacting via force-mediating particles known as gauge bosons.

The SM has been experimentally verified to a high level of accuracy, but nevertheless, is not a complete theory. As mentioned, it lacks a description of Gravity. The model also contains many free parameters, which may suggest the existence of first principles from which these could be derived. In addition, the Higgs boson, which is a key part of the SM theorized to endow massive particles with their mass, still eludes experimental detection.

[1] Gravity is excluded, as to date it has not been possible to develop a renormalizable quantum field theory of gravity.

1.1.1 Particles and Forces

The SM is a gauge theory with group structure $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$. The particle family of the SM is shown in Figure 1.1. The mass, charge, and spin are listed for each particle. An interesting empirical observation is that the quarks and leptons may be grouped into three generations with increasing mass scales, as in the figure. That is, the particle properties of each of the three generations are entirely identical except that the masses increase with each generation. Since more massive particles decay to less massive ones—assuming there is an allowed decay mode—the fascinating fact arises that the everyday world around us is composed of only the first, lightest generation of matter (i.e., the ordinary world of matter is composed of u and d quarks and electrons). The additional generations are more exotic and usually need to be teased out at high energy scales, such as those created by particle accelerators.[2] Table 1.1 summarizes which forces act on which particles and lists the mediators for each interaction, as described below.

Table 1.1: Summary of forces acting on the quarks and leptons of the SM and the mediators for each interaction.

Particle		Force		
		Strong	Electromagnetic	Weak
		gluons	γ	W^+, W^-, and Z
Quarks (u, d, c, s, t, b)		✓	✓	✓
Leptons	Charged (e, μ, τ)		✓	✓
	Neutrinos $(\nu_e, \nu_\mu, \nu_\tau)$			✓

The Strong Force

The $\mathrm{SU(3)_C}$ factor corresponds to QCD or the strong force. There are six unique flavors of quarks (u, d, c, s, t, b), as seen in Figure 1.1, along with their accompanying anti-particles. Under the strong force, particles interact via "color" charge, and each quark carries one of

[2]The preference of matter over anti-matter in the universe is another fascinating empirical observation; that is, the ordinary world around us is composed of matter rather than anti-matter and this trend appears to be universal. Presumably this is because there was slightly more matter than anti-matter in the early universe, so that when the two annihilated the excess matter remained to form the cosmos. The mechanism of how this asymmetry came to be in the first place remains a mystery, however.

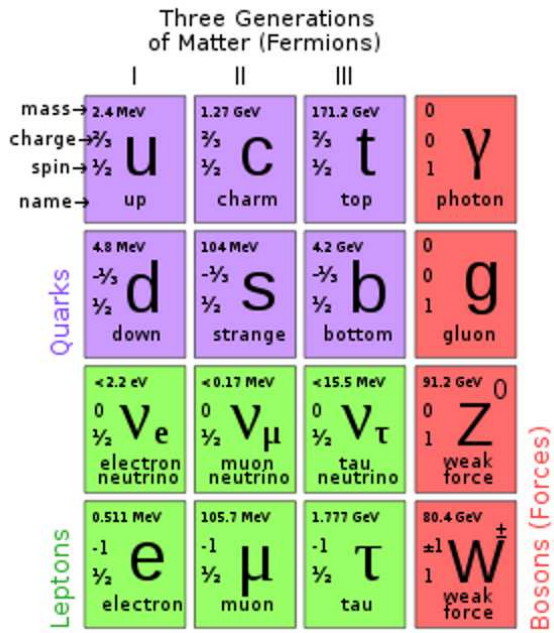


Figure 1.1: The Standard Model of elementary particle physics.

the three colors (r, g, b). The $\mathrm{SU}(3)_\mathrm{C}$ group gives rise to the eight massless mediators of the strong interaction, called gluons, which themselves carry a combination of color and anti-color, implying they interact with each other. Quarks and gluons are collectively referred to as partons.

Asymptotic freedom is a defining feature of QCD: The force between color carrying particles approaches zero (infinity) as the distance between them approaches zero (infinity). In particular, the strong coupling constant α_s is given at leading-order (LO) by:

$$\alpha_\mathrm{s}(Q^2) = \frac{12\pi}{(11c - 2n_f)\log\left(Q^2/\Lambda^2\right)} \tag{1.1}$$

where Q is the momentum transfer of the interaction, $c = 3$ is the number of quark colors, $n_f = 6$ is the number of quark flavors, and Λ is the QCD scale parameter. At high momentum transfer (i.e., small distance scales) $\alpha_\mathrm{s} \to 0$ as $Q^2 \to \infty$, and at low momentum transfer (i.e., large distance scales), $\alpha_\mathrm{s} \to \infty$ as $Q^2 \to \Lambda^2$.

Asymptotic freedom drives confinement: Solitary quarks are expressly forbidden, and hence quarks may only exist in color neutral groupings. They are found in colorless

quark-antiquark pairs called mesons (e.g., the $J/\psi = c\bar{c}$ particle) or in colorless groupings of three called baryons (e.g., the proton $p = uud$ or neutron $n = udd$).[3] Collectively, mesons and baryons are referred to as hadrons. The primary constituent quarks of a hadron are the valence quarks, which continuously interact via gluon exchange. These gluons can emit even more gluons and split into virtual quark anti-quark pairs, which are called sea quarks. The momentum distributions of the partons inside a given hadron are described by its parton distribution function (PDF).

The Electroweak Force

The electromagnetic and weak interactions have been unified into a single electroweak force by Glashow, Weinberg, and Salam [11, 12, 13]. This interaction corresponds to the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ factor of the SM. The postulated Higgs boson is thought to be responsible for the spontaneous symmetry breaking $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y} \rightarrow \mathrm{U(1)_{EM}}$ that "breaks" the gauge group into its unbroken electromagnetic and broken weak components observed at low energy scales. Through this mechanism, the scalar Higgs field is believed to endow particles with their masses. The $\mathrm{SU(2)_L}$ group gives rise to three massless vector gauge bosons W^1, W^2, and W^3. The $\mathrm{U(1)_Y}$ group yields another massless vector gauge boson B. The Higgs mechanism leads these bosons to mix via the following relations, resulting in the three experimentally observed massive states W^+, W^-, and Z, along with the massless state γ:

$$W^{\pm} = \frac{W^1 \pm iW^2}{\sqrt{2}} \tag{1.2}$$

$$Z = W^3 \cos(\theta_{\mathrm{W}}) - B \sin(\theta_{\mathrm{W}}) \tag{1.3}$$

$$\gamma = W^3 \sin(\theta_{\mathrm{W}}) + B \cos(\theta_{\mathrm{W}}) \tag{1.4}$$

where θ_{W} is the weak mixing angle.

There are 12 types of leptons, which all interact exclusively via the electroweak force: the electrically charged leptons (e, μ, τ) and the electrically neutral neutrinos $(\nu_e, \nu_\mu, \nu_\tau)$ of

[3] Exotic groupings of five or even more quarks may also be possible, though they have never been definitively observed.

the three mass generations, as seen in Figure 1.1, and their corresponding antiparticles.

The electromagnetic force—$\mathrm{U}(1)_{\mathrm{EM}}$ after symmetry breaking—acts on the charged leptons and is mediated via the photon γ. Quarks (and often their groupings) carry electric charge, implying the electromagnetic force also acts on them. A defining feature of the electromagnetic force is the well-known "$1/r^2$" law, by which particles having the same (opposite) electric charge are repelled from (attracted to) each other with a force proportional to the inverse square of the distance r between them.

The $W^{\pm}$ and Z bosons mediate the weak force, which is best known for its role in nuclear decays (e.g., neutron decay mediated by the W^- boson: $n \to pe\bar{\nu}_e$). The $W^{\pm}$ and Z bosons interact with all quarks and leptons, including the neutrinos, which are near-massless[4] and interact only weakly. The large masses of the mediating bosons drive the short distance scale of this interaction.

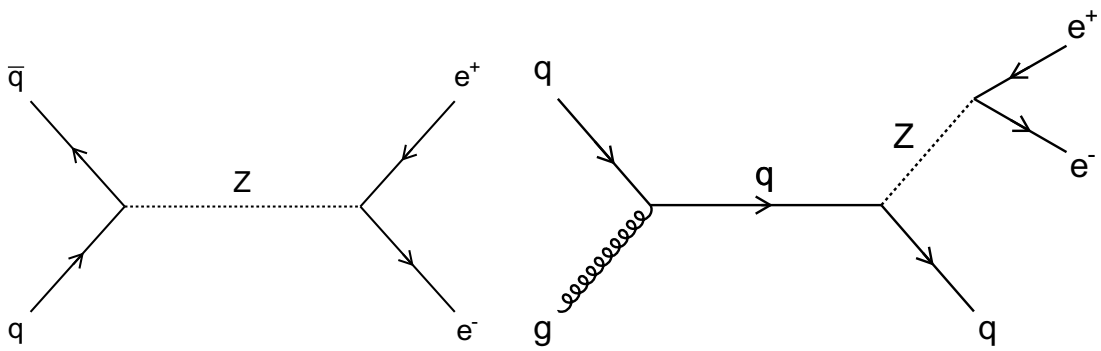
1.2 Z Boson Production at the LHC

Weak neutral currents were predicted by the electroweak theory of Glashow, Weinberg, and Salam [11, 12, 13]. They were later observed by the Gargamelle bubble chamber at CERN in 1974 [14]. The mediating particle of these interactions—the Z boson—was then produced and detected at CERN's Super Proton Synchrotron's UA1 and UA2 experiments along with the $W^{\pm}$ bosons in 1983 [15, 16]. The current world average values for the mass and width of the Z resonance are 91.1876 ± 0.0021 GeV and 2.4952 ± 0.0023 GeV, respectively [6].

Inclusive Z boson production at the LHC, that is $pp \to Z + X$, is dominated by the LO Drell-Yan process $q\bar{q} \to Z \to ee$, in which a quark and antiquark from the incoming protons fuse into a Z (Figure 1.2 (a)). The Drell-Yan process may also include initial-state gluon radiation. The other main Z production mode is NLO quark gluon scattering: $qg \to qZ \to qee$ (Figure 1.2 (b)). The Drell-Yan process constitutes about 65% of Z

[4] Neutrinos are treated as massless in the SM, but recent observations of neutrino flavor oscillation indicate they have non-zero mass. Minor modifications to the SM can accommodate this.

production at the LHC, while $qg \to qZ \to qee$ constitutes the vast majority of the remaining 35%.



(a) (b)

Figure 1.2: Feynman diagrams of the dominant Z production modes at the LHC: (a) the LO Drell-Yan Z production process, (b) an example of NLO $qg \to qZ \to qee$ production.

1.3 Inclusive $Z \to \ell\ell$ Production Cross Section

The production cross section represents the likelihood of a given interaction. That is, at the LHC the inclusive $Z \to \ell\ell$ production cross section characterizes the probability of creating this resonance in a proton-proton (pp) collision that then decays to the specified charged lepton pair. Theoretically, it may be calculated by convolving the parton-level cross section $\hat{\sigma}_{ab}(x_1, x_2)$ for partons a and b with their PDFs f_a and f_b, and then summing over all partons:

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to \ell\ell) = \tag{1.5}$$
$$\sum_{a,b} \int_0^1 \int_0^1 f_a(x_1) f_b(x_2) \hat{\sigma}_{ab}(x_1, x_2) dx_1 dx_2 \times \frac{1 - 4\sin^2\theta_{\mathrm{W}} + 8\sin^4\theta_{\mathrm{W}}}{21 - 40\sin^2\theta_{\mathrm{W}} + \frac{160}{3}\sin^4\theta_{\mathrm{W}}}$$

Here $\mathcal{B}(Z \to \ell\ell)$ is the branching ratio, which represents the probability of a Z boson decaying to the specified dilepton $\ell\ell$ and is given by the factor following the multiplication sign on the right hand side of the equation. The parton-level cross section of the dominant

Drell-Yan process $q\bar{q} \to Z$, for example, is given in terms of the square of the center-of-mass energy $\hat{s}$ of the parton-parton system at LO by:

$$\hat{\sigma}_{q\bar{q}}(\hat{s}) = \pi \frac{G_{\mathrm{F}}}{\sqrt{2}}(1 - 4|Q_q|\sin^2\theta_{\mathrm{W}} + 8Q_q^2\sin^2\theta_{\mathrm{W}})M_Z^2\delta(\hat{s} - M_Z^2) \tag{1.6}$$

where G_{F} is the Fermi constant and Q_q is the electric charge of the quark.

Experimentally, the cross section is determined via the relation:

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to \ell\ell) = \frac{N_{\mathrm{S}}}{A\epsilon_Z \int \mathcal{L}dt} \tag{1.7}$$

where N_{S} is the total number of $Z{\to}\ell\ell$ decays observed for a given integrated luminosity $\int \mathcal{L}dt$, A is the fiducial and kinematic acceptance for the signal, and ϵ_Z is the efficiency of collecting and selecting the $Z{\to}\ell\ell$ decays. The integrated luminosity characterizes the total number of inelastic pp collisions N_{pp} and is inversely proportional to the total inelastic pp collision cross section $\sigma(pp \to X)$:

$$\int \mathcal{L}dt = \frac{N_{pp}}{\sigma(pp \to X)} \tag{1.8}$$

The branching ratios for Z decays are provided in Table 1.2 [6]. Although the branching ratios to charged leptons are rather small at approximately 3.4% for each flavor, these channels provide an efficient means for identifying the final state with low background contamination at hadron colliders.[5] The $Z{\to}ee$ cross section is approximately 1000 pb for $\sqrt{s} = 7$ TeV at the LHC. Using the CMS selection criteria, the signal acceptance and efficiency are $A \approx 0.4$ and $\epsilon_Z \approx 0.6$. Thus about 200–300 $Z{\to}ee$ events are expected to be selected per pb^{-1} of luminosity.

Inclusive Z boson production cross section measurements have been an important early measurement at new colliders ever since their discovery:

- They have well-known cross sections and are produced copiously.
- Their leptonic decays (along with those of $W^{\pm}$ bosons) have clean detector signatures and provide the first electroweak precision measurements, making them benchmark physics processes.

[5] The branching ratios are identical (within experimental errors) for each lepton flavor. This property is known as lepton universality.

Table 1.2: The Z decay mode branching ratios [6].

Mode	$\mathcal{B}$
e^+e^-	0.03363 ± 0.00004
$\mu^+\mu^-$	0.03366 ± 0.00007
$\tau^+\tau^-$	0.03367 ± 0.00008
$\nu\bar{\nu}$	0.2000 ± 0.0006
$q\bar{q}$	0.6991 ± 0.0006

- They allow verification of high transverse momentum electron reconstruction and identification.

In addition, precision measurements of the Z at the LHC are important in testing the Standard Model more rigorously than ever before, constraining the proton PDF, and potentially uncovering signs of new physics that could appear through radiative corrections. Finally, the copious production of these well understood and clean signatures may suggest the use of Z bosons as "standard candles" for measuring the luminosity at the LHC. Thus measuring the inclusive $Z \to ee$ production cross section, which is the primary topic of this thesis, is of critical importance to the LHC physics program.

The inclusive $Z \to \ell\ell$ production cross sections reported in this thesis pertain to the invariant mass range $M_{\ell\ell} \in (60, 120)$ GeV, and are corrected for the fiducial and kinematic acceptance, but not for γ^* exchange. That is, what is referred to as Z production throughout this thesis also includes interference from the Drell-Yan virtual photon spectrum; technically it is Z/γ^* production.

Figure 1.3 provides an overview of inclusive $Z \to \ell\ell$ production cross section measurements at previous hadron collider experiments [1, 2, 3, 4] and the next-to-next-to-leading-order (NNLO) theoretical prediction (Chapter 10). The measured cross sections are represented by solid symbols with error bars as indicated in the legend, and the NNLO theoretical prediction is represented by a blue line.

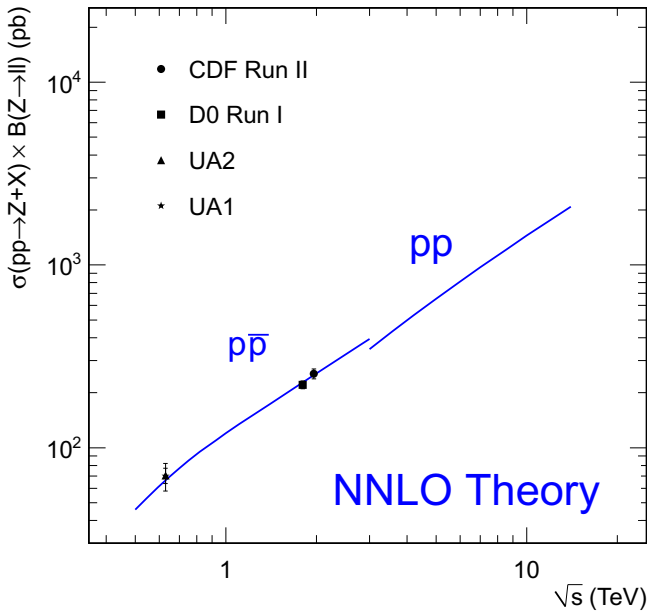


Figure 1.3: Measurements of the $Z \to \ell\ell$ inclusive production cross section from experiments at lower energy hadron colliders [1, 2, 3, 4]. The measured cross sections are represented by solid symbols with error bars as indicated in the legend, and the NNLO theoretical prediction is represented by a blue line.

Chapter 2

Experimental Apparatus: the LHC and the CMS Detector

2.1 The Large Hadron Collider

The LHC [17] is a two ring superconducting hadron accelerator and collider installed in the pre-existing 26.7 km tunnel that hosted the CERN LEP machine [18]. It is located near Geneva, Switzerland, and lies between 45–170 m below the ground on a plane inclined at 1.4% sloping towards Lake Geneva.

The LHC is designed to provide pp collisions at a center-of-mass energy of $\sqrt{s} = 14$ TeV and luminosity of $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1}.[1] The high beam intensities required to achieve this luminosity exclude counter-rotating proton and antiproton beams in a single ring, such as exists at the Tevatron [19]. Instead, the LHC has two rings with opposing magnetic dipole fields that bend the counter-rotating proton beams around the collider. The nominal strength of these magnetic dipole fields is 8.33 T. There are four interaction points (IPs) along the tunnel, each housing an experiment: two high luminosity IPs house the CMS [20] and ATLAS [21] detectors, while two lower luminosity IPs house the LHCb [22] and ALICE [23] detectors. A machine schematic is shown in Figure 2.1.

[1] The LHC can also run in a heavy ion mode in which lead ions are collided.

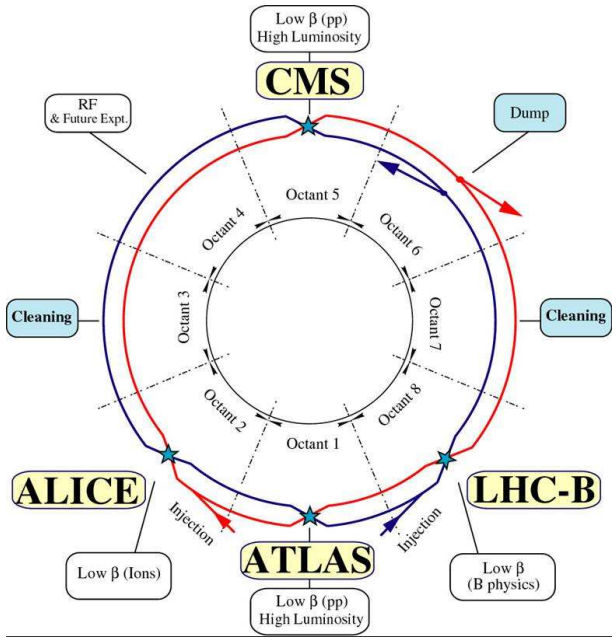


Figure 2.1: Schematic of the LHC machine.

2.1.1 Commissioning Timeline

The first pp collisions at the LHC were delivered in November, 2009 at $\sqrt{s} = 900$ GeV, and then later at $\sqrt{s} = 2.136$ TeV. This latter center-of-mass energy just exceeded that of the Tevatron (1.96 TeV), making the LHC the highest energy collider ever operated. The year 2010 saw the first extended period of data taking, beginning in March and ending in November. During this time the machine operated at $\sqrt{s} = 7$ TeV, far exceeding the Tevatron center-of-mass energy, and delivered approximately 45–50 pb^{-1} of integrated luminosity to both the ATLAS and CMS experiments.

The LHC suffered a quench in September, 2008 and was subsequently damaged. This incident was caused by a faulty weld connecting two sections of superconducting wire. The machine was repaired in the following year, and several improvements were made, including the installation of a new magnet protection system. Precaution suggested operation at $\sqrt{s} = 7$ TeV until 2012, when a shutdown will allow further upgrades to enable $\sqrt{s} = 14$ TeV collisions.

2.1.2 Magnets

The LHC is constructed of 1232 superconducting NbTi dipole bending magnets cooled by superfluid helium to a temperature of 1.9 K. Each magnet has a length of 16.5 m, a diameter of 0.57 m, and a mass of 2.75×10^4 kg. At the design beam energy of 7 TeV, the magnets have a current of 11.85 kA and a corresponding magnetic field strength of 8.33 T. The cross section of a dipole is illustrated in Figure 2.2.

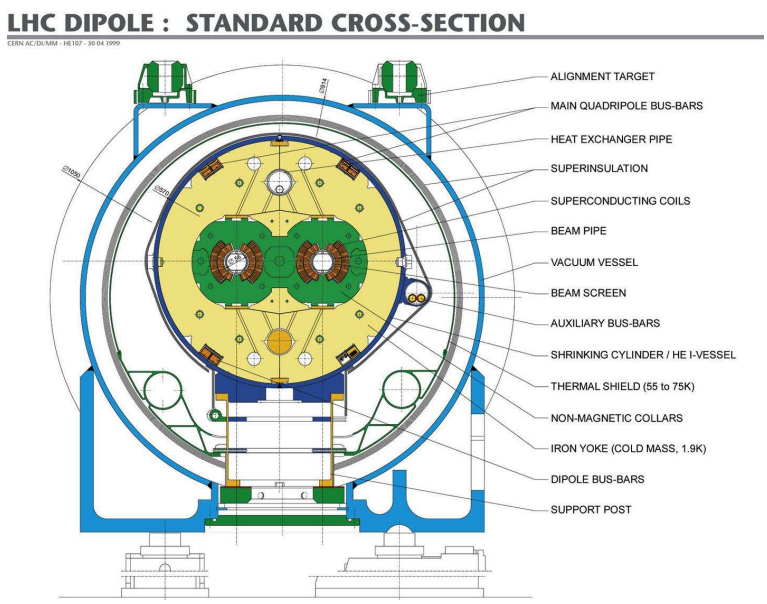


Figure 2.2: Cross section of a LHC dipole magnet.

Other magnets composing the LHC include 392 focusing quadrupole magnets, about 4800 corrector magnets, eight triplet quadrupole magnets that "squeeze" the beams into collision at the IPs, and dozens of kicker magnets used for beam injection and dumping.

2.1.3 Acceleration Systems

Beams injected into the LHC are captured, accelerated, and stored using a 401 MHz superconducting radio frequency (RF) cavity system. The single-cell cavities are constructed of niobium sputtered on copper. There is a separate system for each beam, and they are housed at LHC Point 4. Each system is composed of eight cavities, which are

independently driven by 300 kW klystrons. During acceleration, the RF system delivers an energy gain of 485 keV to the protons of each beam per LHC revolution. Synchrotron radiation energy losses are minimal due to the large mass of the proton: At a beam energy of 7 TeV, each proton loses only 7 keV per LHC revolution.

2.1.4 Injector Complex and Bunch Structure

The LHC is supplied with protons via the following injector chain (Figure 2.3):

Linac2 → Proton Synchrotron Booster (PSB) → Proton Synchrotron (PS)
→ Super Proton Synchrotron (SPS) → LHC

These accelerators, excluding the LHC, were preexisting at CERN and were upgraded to meet the stringent specifications of the new machine; a large number of high intensity proton bunches having small transverse and well-defined longitudinal emittances. The main challenges for LHC injection are:

- The unprecedented transverse beam brightness (ratio of the intensity to the emittance).
- The production of a bunch structure with the LHC spacing of 25 ns before extraction from the 25 GeV PS machine.

The transverse emittances of the LHC beams have to be maintained at their unusually small size throughout the injector chain. Small amounts of mis-steering and mismatch among the accelerators of the chain, virtually negligible for past operations, is relevant at the LHC. Their effect has to be measured by high-resolution beam profile monitors. Various position measurement systems have been modified to deal with the new harmonics in the circular machines and allow bunch-by-bunch observations.

The PS and SPS prepare the LHC proton bunch structure. Each beam is composed of 3564 bunches spaced by 25 ns. This spacing sets the 40 MHz sampling frequency of the experiments. Some bunches are empty (i.e., contain no protons), forming gaps to give the

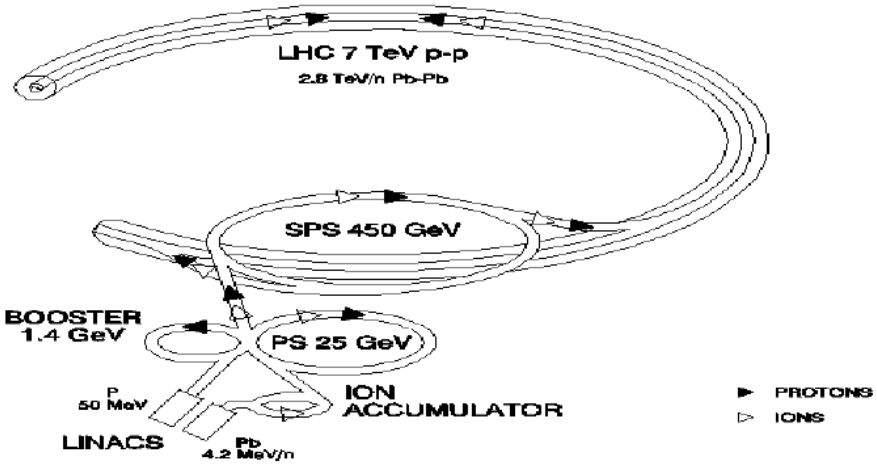


Figure 2.3: The LHC injector complex.

kicker magnets time to ramp up or down for beam injection and dumping. At design luminosity, 2808 of the 3564 bunches will be filled, corresponding to about 25 *pp* interactions per bunch crossing.

2.1.5 Luminosity Estimates

The instantaneous luminosity decreases throughout an LHC fill dominantly due to beam loss caused by the *pp* collisions. Accounting only for this collision induced decay, the initial lifetime of the beam intensity is:

$$\tau_I = \frac{N_0}{\sigma(pp \to X)\mathcal{L}_0 k} \tag{2.1}$$

where N_0 is the initial beam intensity in terms of number of protons, $\sigma(pp \to X)$ is the total *pp* collision cross section, $\mathcal{L}_0$ is the initial instantaneous luminosity, and k is the number of IPs. Assuming two high luminosity experiments with an initial instantaneous luminosity of $\mathcal{L}_0 = 10^{34}\text{cm}^{-2}\text{s}^{-1}$, yields a beam lifetime of $\tau_I = 45$ h. Equation 2.1 leads to the following decay of the beam intensity and luminosity as functions of time t:

$$N(t) = \frac{N_0}{1 + t/\tau_I} \tag{2.2}$$

$$\mathcal{L}(t) = \frac{\mathcal{L}_0}{(1 + t/\tau_I)^2} \tag{2.3}$$

The luminosity lifetime is then:

$$\tau_{\mathcal{L}} = (\sqrt{e} - 1)\tau_I \tag{2.4}$$

which yields a value $\tau_{\mathcal{L}} = 29$ h. Factoring in other losses[2] leads to a more accurate estimate of $\tau_{\mathcal{L}} = 15$ h, however.

The integrated luminosity over one fill is given by:

$$\int_{t=0}^{T_{\text{fill}}} \mathcal{L}(t)dt = \mathcal{L}_0\tau_{\mathcal{L}}(1 - e^{-T_{\text{fill}}/\tau_{\mathcal{L}}}) \tag{2.5}$$

where T_{fill} is the duration of the fill. Given a turnaround time of 7 h between fills and the luminosity lifetime of 15 h, the optimal fill duration is 12 h. If the LHC runs accordingly for 200 days per year, and $\mathcal{L}_0 = 10^{34}\text{cm}^{-2}\text{s}^{-1}$ for each fill, then the total integrated luminosity collected per year will be about 80 fb^{-1}. Assuming a much more optimistic turnaround time of 1.5 h still leads to a similar result: 120 fb^{-1}.

2.2 The Compact Muon Solenoid

2.2.1 Introduction: Detector Requirements, and Goals

CMS is a general purpose particle detector designed for the physics environment provided by the LHC. It is situated about 100 m underground near the village of Cessy, France, between Lake Geneva and the Jura mountains. The LHC will deliver approximately 10^9 inelastic *pp* collisions per second to the detector at design energy and luminosity. The sampling frequency of the experiment is fixed at 40 MHz since the detector is read out once every bunch crossing. This implies about 25 *pp* interactions are superimposed on each event of interest at full design energy and luminosity. (This phenomenon of multiple interactions occurring per bunch crossing is called pileup.) Very high specifications are required in particle detection and read-out systems to cope with the high interaction rate, high occupancy, and associated synchronization.

[2]Other sources of beam loss include slow emittance blow-up and Touschek scattering [24].

While CMS is a general purpose detector, its primary aim is to elucidate the nature of electroweak symmetry breaking and search for new phenomena at the TeV scale. The requirements for the detector to meet these aims may be summarized as:

- Good muon identification and momentum resolution over a broad range of momenta with wide geometric coverage ($|\eta| < 2.5$). Good dimuon mass resolution (approximately 1% at 100 GeV), and the ability to determine the charge of muons with $p < 1$ TeV unambiguously.

- Good charged particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ particles and b-jets, requiring pixel detectors close to the interaction region.

- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution (approximately 1% at 100 GeV), wide geometric coverage ($|\eta| < 2.5$), precise measurement of the location and direction of photons and/or correct localization of the primary interaction vertex, π^0 rejection, and efficient photon and lepton isolation at high luminosities.

- Good missing transverse energy and dijet mass resolution, requiring hadron calorimetry with a large hermetic geometric coverage ($|\eta| < 5$) and fine lateral segmentation ($\Delta\eta \times \Delta\phi < 0.1 \times 0.1$).

Figure 2.4 illustrates the detector layout. At the core of the detector sits a large-bore superconducting solenoid with a 3.8 T magnetic field surrounding an all silicon pixel and strip inner tracker, a lead-tungstate scintillating crystal electromagnetic calorimeter, and a brass-scintillator sampling hadron calorimeter. The iron yoke of the magnetic flux return is instrumented with four stations of muon detectors covering most of the 4π solid angle. Forward sampling calorimeters extend the pseudorapidity coverage from $|\eta| < 3$ to $|\eta| < 5$, assuring good hermeticity. The overall dimensions of the detector are a length of 21.6 m, a diameter of 14.6 m, and a total weight of about 1.25×10^7 kg.

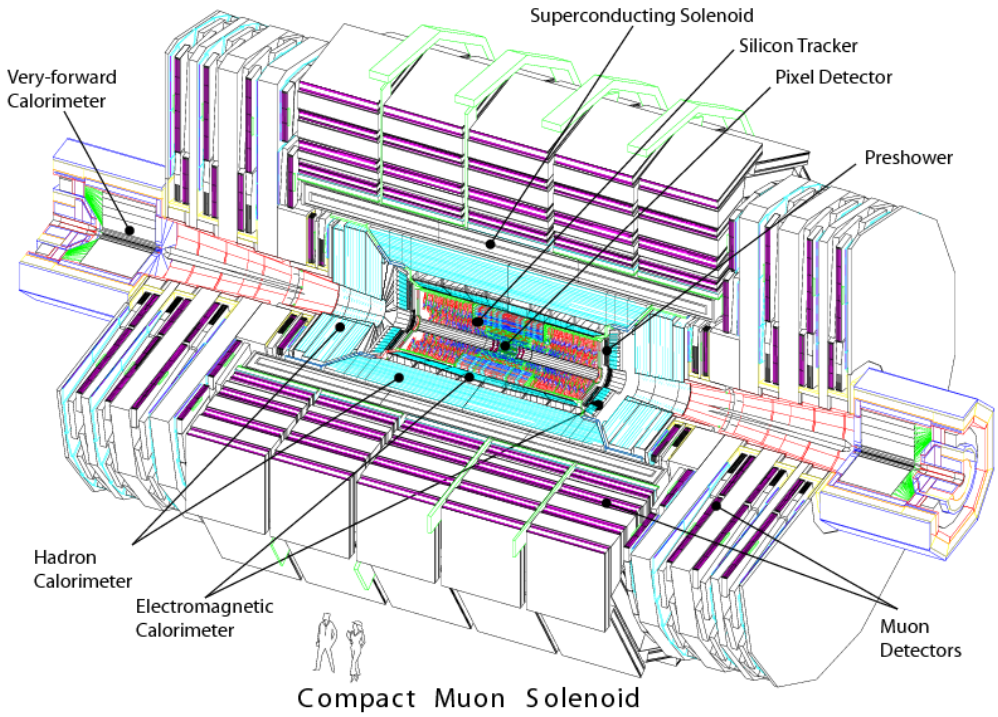


Figure 2.4: Illustration of the CMS detector.

2.2.2 Coordinate Conventions

The adopted coordinate system has the origin located at the nominal IP inside the detector, the x-axis pointing radially inward toward the center of the LHC, and the y-axis pointing vertically upward. Thus the z-axis points west towards the Jura mountains from the CMS cavern at "LHC Point 5", along the direction of the counterclockwise traveling beam (as seen from above). The azimuthal angle ϕ is measured from the x-axis in the x-y plane, and has the range $(-\pi, \pi)$. The polar angle θ is measured from the z-axis. Pseudorapidity is defined as $\eta \equiv -\ln(\tan(\theta/2))$. Thus, the momentum and energy measured transverse to the beam direction, denoted by P_T and E_T, respectively, are computed from the x and y components. The imbalance of energy measured in the transverse plane is denoted by $\not{E}_T$. Natural units with c $= \hbar = 1$ are used so that energy and momenta are expressed in units of eV.

2.2.3 Superconducting Magnet

The superconducting solenoidal magnet (Figure 2.4) has been designed to provide a uniform 4 T field in a free bore of 6 m diameter and 12.5 m length.[3] The stored energy corresponds to 2.6 GJ at full current. The radial thickness of the cold mass is 312 mm, corresponding to 3.9 radiation lengths. The flux is returned through a 10^7 kg yoke comprising five wheels and two endcaps. Each endcap is composed of three disks. The distinctive feature of the 2.2×10^5 kg cold mass is its four layer winding made from a stabilized, reinforced NbTi conductor. The ratio between the stored energy and cold mass is a hefty 11.6 kJ/kg, which is well beyond the value of any previous solenoidal detector magnet and causes a large mechanical deformation (0.15%) during energizing.

2.2.4 Inner Tracking System

The inner tracking system is designed to provide a precise and efficient measurement of the trajectories of charged particles emerging from the *pp* collisions, as well as a precise reconstruction of secondary vertices. It surrounds the IP and has a length (diameter) of 5.8 m (2.5 m). The solenoid provides a homogeneous field of 3.8 T over the full volume of the tracker. At design luminosity there will be an average of about 1000 particles traversing the tracker each bunch crossing. Therefore, a highly granular detector technology with fast response is required to identify the trajectories reliably and attribute them to the correct bunch crossing. This, however, drives the high power consumption of the on-detector electronics, which in turn require efficient cooling, increasing the material budget in direct conflict with the aim of minimizing the amount of multiple scattering, bremsstrahlung, photon conversion, and nuclear interactions. Thus these competing interests had to be mutually optimized. These intense particle fluxes also cause severe radiation damage to the tracking system. The main design challenge of the tracking system was to develop detector components able to operate in this harsh environment for an expected lifetime of 10 years. These requirements lead to a tracker design based

[3] To date, the magnet has only been operated at a field of 3.8 T.

entirely on silicon detectors.

The tracker is composed of a pixel detector with three barrel layers at radii between 4.4–10.2 cm and a silicon strip tracker with 10 barrel detection layers extending outwards to a radius of 1.1 m. Each system is completed by endcaps consisting of two disks in the pixel detector and three plus nine disks in the strip tracker on each side of the barrel, extending the acceptance of the tracker up to a pseudorapidity of $|\eta| < 2.5$ (Figure 2.5). With approximately 200 m^2 of active silicon area, the CMS tracker is the largest silicon tracker ever built [7, 25].

As mentioned, at design luminosity each bunch crossing brings on average about 1000 particles hitting the tracker, leading to a hit rate density of 1 MHz/mm^2 at a radius of 4 cm, 60 kHz/mm^2 at a radius of 22 cm, and 3 kHz/mm^2 at a radius of 115 cm. In order to keep the occupancy below 1%, pixelated detectors are used at radii less than 10 cm. For a pixel size of 100×150 μm^2 in either the r-ϕ or z planes, the occupancy is of the order 10^{-4} per pixel and LHC bunch crossing. This pixel size choice is driven by the desired impact parameter resolution. At intermediate radii of 20–55 cm, the reduced particle flux allows the use of silicon micro-strip detectors with a typical cell size of 10 cm $\times$ 80 μm, leading to an occupancy of up to 2–3% per strip and LHC bunch crossing. In the outer region corresponding to radii of 55–110 cm, the strip pitch is further increased. Given the large areas that are instrumented in this outer region, the strip length is increased in order to limit the number of read-out channels. However, the strip capacitance scales with its length and therefore the electronics noise is a linear function of the length. In order to maintain a signal-to-noise ratio well above 10, thicker silicon sensors are used for the outer tracker region (500 μm thickness as opposed to the 320 μm in the inner tracker). These thicker sensors would in principle have a higher depletion voltage, but since the radiation levels in the outer tracker are smaller, a higher initial resistivity can be chosen such that the initial depletion voltages of thick and thin sensors are in the same range of 100–300 V. In this way, cell sizes up to about 25 cm $\times$ 180 μm can be used in the outer region of the tracker, with an occupancy of about 1%. These occupancy-driven design choices for the

strip tracker also satisfy the requirements on position resolution. CMS is the first experiment to use silicon detectors in this outer region.

The radiation damage introduced by the high particle fluxes at the IP is a severe design constraint. Table 2.1 shows the expected fast hadron fluence and radiation dose in the barrel tracker for an integrated luminosity of 500 fb^{-1}, corresponding to about 10 years of LHC operation [7, 25]. Neutrons generated by hadronic interactions in the electromagnetic calorimeter (Ecal) crystals make up a substantial contribution to the fast hadron fluence, which actually dominates in the outer tracker close to the Ecal surface.

Table 2.1: The expected hadron fluence and radiation dose in different radial layers of the barrel tracker for an integrated luminosity of 500 fb^{-1}, corresponding to approximately 10 years of LHC operation. The fast hadron fluence is a good approximation to the 1 MeV neutron equivalent fluence [7].

Radius (cm)	Fast Hadron Fluence (10^{14} cm^{-2})	Dose (kGy)	Charged Particle Flux ($cm^{-2}s^{-1}$)
4	32	840	10^8
22	1.6	70	6×10^6
115	0.2	1.8	3×10^5

Tracker Layout

A schematic drawing of the tracker is shown in Figure 2.5. At radii of 4.4, 7.3, and 10.2 cm, three cylindrical layers of hybrid pixel detector modules surround the IP. They are complemented by two endcap disks of pixel modules on each side. The pixel detector delivers three high precision space points on each charged particle trajectory. It covers an area of about 1 m^2 and has 66×10^6 pixels.

The radial region between 20–116 cm is occupied by the silicon strip tracker, which is composed of three different subsystems. The Tracker Inner Barrel and Disks (TIB/TID) extend in radius 55 cm and are composed of four barrel layers, supplemented by three endcap disks on each side. TIB/TID delivers up to four r-ϕ measurements on a trajectory using 320 μm thick silicon micro-strip sensors with their strips parallel (radial) to the

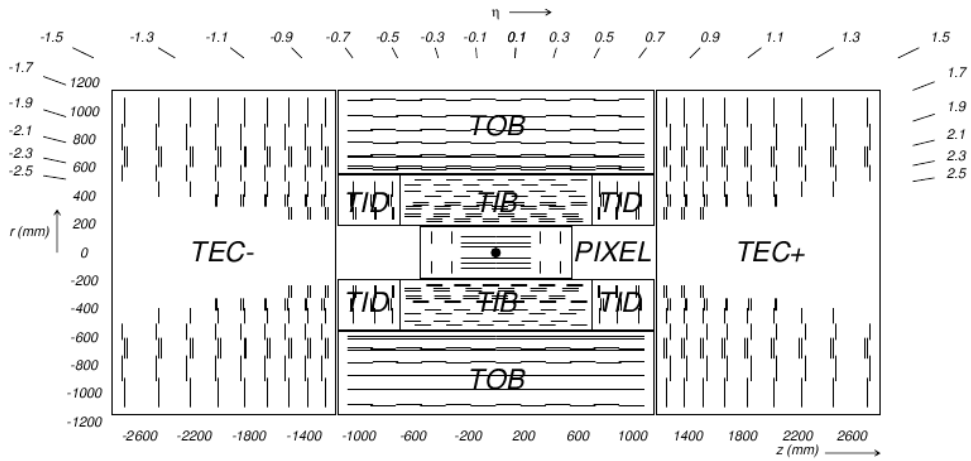


Figure 2.5: Schematic cross section through the tracker. Each line represents a detector module. Double lines indicate back-to-back modules that deliver stereo hits.

beam axis in the barrel (disks). The strip pitch is 80 μm (120 μm) on layers one and two (three and four) in the TIB, leading to a single point resolution of 23 μm (35 μm). In the TID, the mean pitch varies between 100–141 μm. The TIB/TID is surrounded by the Tracker Outer Barrel (TOB). It has an outer radius of 116 cm and consists of six barrel layers of 500 μm thick micro-strip sensors with strip pitches of 183 μm on the first four layers and 122 μm on layers five and six. It provides another six r-ϕ measurements with single point resolution of 53 μm and 35 μm, respectively. The TOB extends in z between ± 118 cm. Beyond this z range, the Tracker Endcaps (TEC+ and TEC$-$, where the sign indicates the location along the z-axis) cover the region $|z| \in (124, 282)$ cm and radius $|r| \in (22.5, 113.5)$ cm. Each TEC is composed of nine disks, carrying up to seven rings of silicon micro-strip detectors (320 μm thick on the inner four rings and 500 μm thick on rings five to seven) with radial strips of 97–184 μm average pitch. Thus, they provide up to nine ϕ measurements per trajectory.

In addition, the modules in the first two layers and rings, respectively, of TIB, TID, and TOB, as well as rings one, two, and five of the TECs carry a second micro-strip detector module which is mounted back-to-back with a stereo angle of 100 mrad in order to provide a measurement of the second coordinate (z in the barrel and r on the disks).

The single point resolution of this measurement is 230 (530) μm in the TIB (TOB), and varies with pitch in TID and TEC. This tracker layout ensures at least nine hits in the silicon strip tracker in the full range $|\eta| < 2.4$, with at least four of them being two-dimensional (Figure 2.6). The acceptance of the tracker ends at $|\eta| \approx 2.5$. The silicon strip tracker has a total of 9.3×10^6 strips and 198 m^2 of active silicon area.

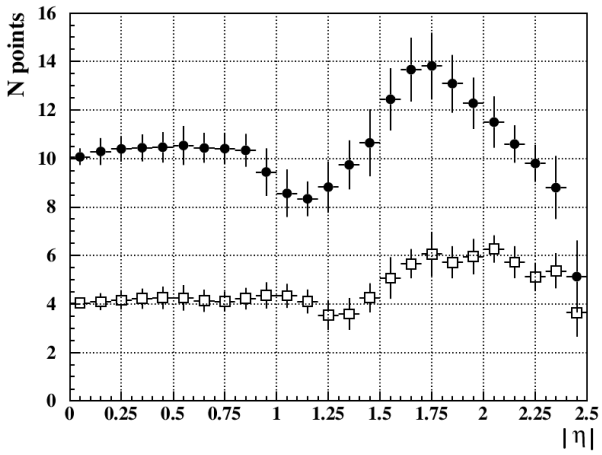
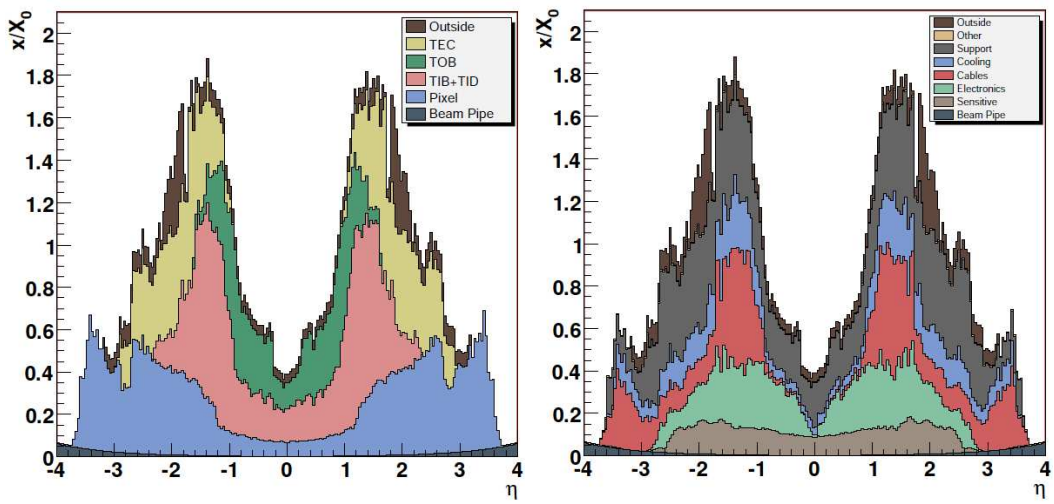


Figure 2.6: Number of measurement points in the strip tracker as a function of $|\eta|$. Filled circles show the total number (back-to-back modules count as one), while open squares show the number of stereo layers.

Figure 2.7 shows the material budget of the CMS tracker in units of radiation length. It increases from approximately 0.4 X_0 at $|\eta| \approx 0$ to about 1.8 X_0 at $|\eta| \approx 1.4$, beyond which it falls to approximately 1 X_0 at $|\eta| \approx 2.5$.

Expected Performance

Figure 2.8 shows the expected resolution of (a) transverse momentum, (b) transverse impact parameter, and (c) longitudinal impact parameter, as a function of $|\eta|$ [7] for single muons of transverse momenta 1, 10, and 100 GeV. For high momentum tracks (100 GeV), the transverse momentum resolution is about 1–2% up to $|\eta| \approx 1.6$, beyond which it degrades due to the reduced lever arm. At a transverse momentum of 100 GeV, multiple scattering in the tracker material accounts for 20–30% of the transverse momentum



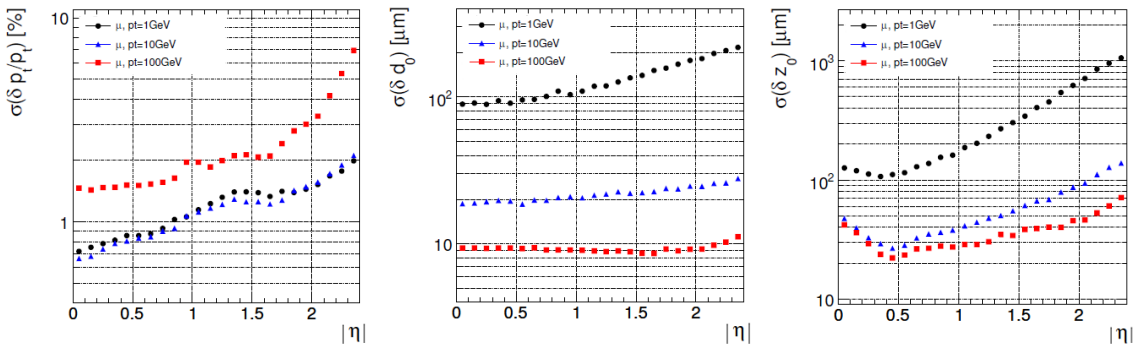
(a) (b)

Figure 2.7: Material budget in units of radiation length as a function of pseudorapidity: (a) for the different sub-detectors, (b) broken down into functional contributions.

resolution, while at lower momentum it is dominated by multiple scattering. The transverse impact parameter resolution reaches 10 μm for high P_T tracks, and is dominated by the resolution of the first pixel hit, while at lower momentum it is degraded by multiple scattering. Similar conclusions hold for the longitudinal impact parameter. Figure 2.9 shows the expected track reconstruction efficiency for single (a) muons and (b) pions as a function of $|\eta|$. The efficiency is about 99% over most of the acceptance for muons. The efficiency decreases slightly near $\eta \approx 0$ due to gaps between the ladders of the pixel detector at $z \approx 0$. The efficiency drop at high $|\eta|$ is mainly driven by the reduced coverage of the pixel forward disks. The efficiency is generally lower for pions and hadrons because of interactions with the tracker material.

2.2.5 Electromagnetic Calorimeter

The Ecal is a hermetic, homogeneous calorimeter made of 61200 lead tungstate ($PbWO_4$) crystals mounted in the central barrel part, closed by 7324 crystals in each of the two



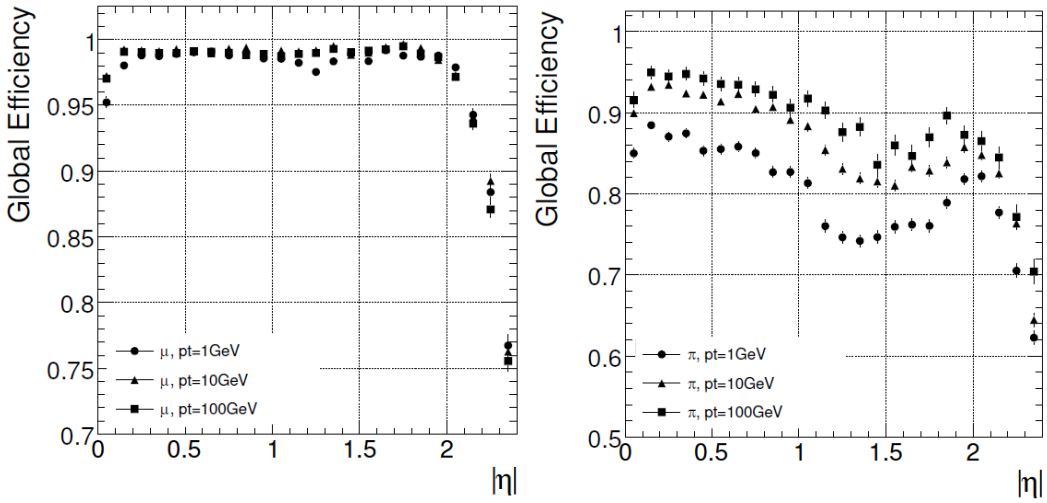
(a) (b) (c)

Figure 2.8: Expected resolution of track parameters as a function of $|\eta|$ for single muons with transverse momenta of 1 (black), 10 (blue), and 100 (red) GeV: (a) transverse momentum, (b) transverse impact parameter, (c) longitudinal impact parameter.

endcaps. A preshower detector is placed in front of the endcap crystals. Avalanche photodiodes (APDs) are used as photodetectors in the barrel and vacuum phototriodes (VPTs) are used in the endcaps. The use of high density crystals allows a calorimeter that is fast and radiation resistant with fine granularity, which are all important characteristics in the LHC environment. The capability to detect the decay of the postulated Higgs boson to two photons was a driving criterion of the design. This capability is enhanced by the good energy resolution provided by a homogeneous crystal calorimeter.

PBWO$_4$ Crystals

The characteristics [26, 27, 28] of the PbWO$_4$ crystals make them an appropriate choice for the Ecal. Their high density (8.28 g/cm^3), short radiation length (0.89 cm), and small Moliere radius (2.2 cm) result in a compact calorimeter with fine granularity. The scintillation decay time of these crystals is of the same order of magnitude as the LHC bunch crossing time: approximately 80% of the light is emitted within 25 ns. The light output is relatively low and varies with temperature: At 18°C, about 4.5 photoelectrons per MeV are collected in both APDs and VPTs. The crystals emit blue-green scintillation



(a) (b)

Figure 2.9: Expected global track reconstruction efficiency as a function of $|\eta|$ for particles with transverse momenta of 1 (circle), 10 (triangle), and 100 (square) GeV: (a) muons, (b) pions.

light with a broad maximum at 420–430 nm [28, 29].

Ecal Layout

Figure 2.10 shows the layout of the calorimeter. The Ecal Barrel (EB) covers the pseudorapidity range $|\eta| < 1.479$. The barrel granularity is 360-fold in ϕ and (2×85)-fold in η, resulting in a total of 61200 crystals. The crystals have a tapered shape, slightly varying with position in η. They are mounted in a quasi-projective geometry to avoid cracks aligned with particle trajectories, so that their axes make a small angle of about 3° with respect to the vector from the nominal IP in both the η and ϕ projections. The crystal cross section corresponds to approximately 0.0174×0.0174 in η–ϕ or 22×22 (26×26) mm^2 at the front (rear) face of crystal. The crystal length is 230 mm, corresponding to 25.8 X_0. The barrel crystal volume is 8.14 m^3 and its weight is 6.74×10^4 kg.

The centers of the front faces of the crystals are at a radius of 1.29 m. The crystals are

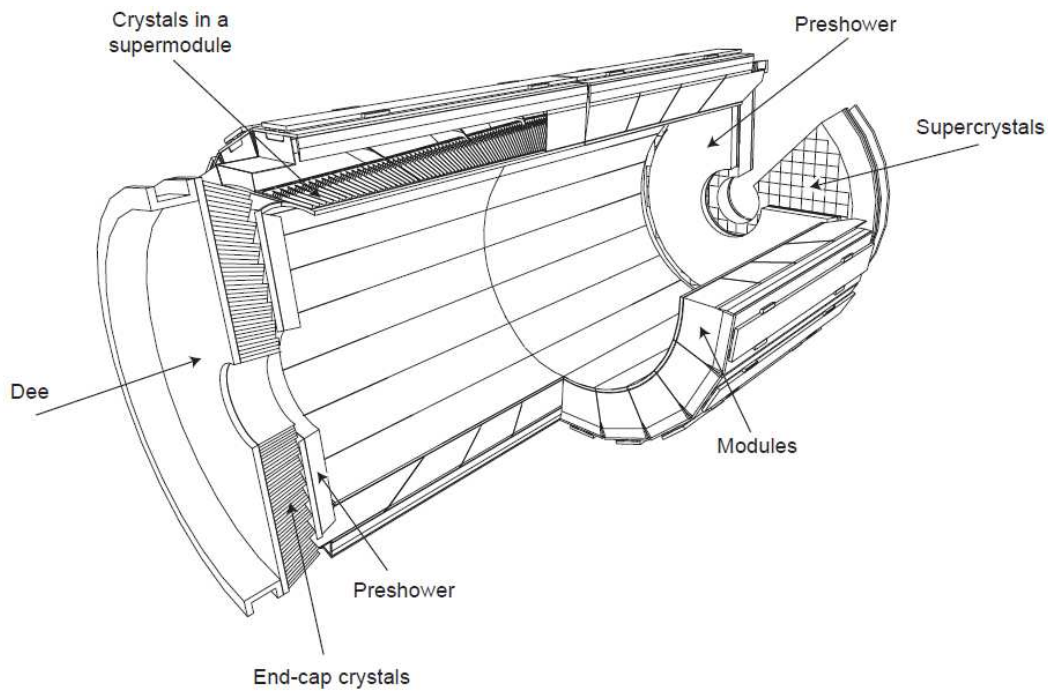


Figure 2.10: Layout of the electromagnetic calorimeter showing the arrangement of crystal modules, supermodules, and endcaps, with the preshower in front.

contained in a thin walled alveolar structure, called a submodule. The submodules are assembled into modules according to their *eta* position. Each module contains 400 or 500 crystals. Four modules, separated by aluminum conical webs 4 mm thick, are assembled into a supermodule containing 1700 crystals.

The Ecal Endcaps (EE) cover the pseudorapidity range $|\eta| \in (1.479, 3.0)$. The longitudinal distance between the IP and the endcap envelope is 315.4 cm, taking into account the estimated 1.6 cm shift toward the IP when the magnetic field is switched on. The two endcaps consist of identically shaped crystals grouped into mechanical units of 5×5 crystals (supercrystals or SCs) consisting of a carbon-fiber alveolar structure. Each endcap is divided into two halves, called Dees. Each Dee holds 3662 crystals. These are contained in 138 standard supercrystals and 18 special partial-supercrystals on the inner and outer circumference. The crystals and supercrystals are arranged in a rectangular x-y grid, with the crystals pointing at a focus 1300 mm beyond the IP, giving off-pointing

angles ranging from 2–8°. The crystals have a front (rear) face cross section of 28.62×28.62 mm^2 (30×30 mm^2) and a length of 220 mm, corresponding to 24.7 X_0. The endcap crystal volume is 2.90 m^3 and its weight is 2.40×10^4 kg.

The principal aim of the preshower detector is to identify neutral pions in the endcaps within the fiducial region of $|\eta| \in (1.653, 2.6)$. It also helps in discriminating electrons from minimum ionizing particles, and improves the position determination of electrons and photons with high granularity. The preshower is a sampling calorimeter with two layers: Lead radiators initiate electromagnetic showers from incoming photons/electrons, while silicon strip sensors placed after each radiator measure the deposited energy and the transverse shower profiles. The total thickness of the preshower is 20 cm. The material thickness of the preshower traversed at $|\eta| = 1.653$ before reaching the first sensor plane is 2 X_0, followed by a further 1 X_0 before reaching the second plane. Thus about 95% of single incident photons start showering before the second sensor plane. The orientation of the strips in the two planes is orthogonal. A major design consideration is for all lead to be covered by silicon sensors, including the effects of shower spread, primary vertex spread, etc. For optimum Level-1 trigger performance, the profile of the outer edge of the lead follows the shape of the Ecal crystals behind it. The exact profiling of the lead is far less critical for the inner radius, and thus a circular shape has been chosen. The lead planes are arranged in two Dees, one on each side of the beam pipe, with the same orientation as the crystal Dees.

Calibration

Calibration is a severe technical challenge for the operation of the Ecal. Many small effects that are negligible at low precision need to be treated with care as the level of precision of a few per mil is approached. Calibration is naturally seen as composed of a global component, giving the absolute energy scale, and a channel-to-channel relative component referred to as inter-channel calibration. The calibrated energies must be uniform and stable so that showers recorded in different locations and/or at different

times are accurately related to each other. The estimated particle energy released in the electromagnetic calorimeter is expressed as:

$$E_{e/\gamma} = F_{e/\gamma}(\eta, E_T) \times \sum_i G(GeV/ADC) \times c_i \times A_i(ADC) \tag{2.6}$$

where the index i runs over the crystals used in the energy sum for the reconstructed particle object, A_i are the reconstructed amplitudes in analog-to-digital converter (ADC) counts, and c_i are the inter-channel calibration factors. The constant G is the absolute energy scale for either the Barrel or the Endcaps. $F_{e/\gamma}$ is a function depending on the particle type, and includes all energy corrections related to shower containment, Bremsstrahlung energy loss, presence of dead channels and geometrical cracks, etc. The Ecal has previously been calibrated in test beams, but now the energy scale and inter-channel calibrations are achieved in situ with physics events. For instance, the Ecal can be calibrated by reconstructing the mass of known resonances such as $\pi^0 \to \gamma\gamma$, $\eta \to \gamma\gamma$, $J/\Psi \to ee$, and $Z \to ee$ [30, 31, 32].[4]

The varying conditions of LHC running drive a cyclical behavior in the crystal transparency between LHC collision runs and machine fills. That is, although the $PbWO_4$ crystals are radiation resistant, they are subject to temporary, limited, rapid losses of optical transmission due to irradiation. Irradiation causes temporary color centers in the crystals that absorb a fraction of the transmitted light. The magnitude of these fluctuations is dose rate dependent, and ranges from 1–2% in the barrel at low luminosity, to tens of percent in the high $|\eta|$ regions of the endcap at design luminosity. The evolution of the crystal transparency is measured using laser pulses injected into the crystals via optical fibers.

2.2.6 Energy Resolution

For energies below about 500 GeV, the resolution can be parameterized as:[5]

[4] The reference in the middle was co-written by the author.

[5] Beyond 500 GeV, shower leakage from the rear of the calorimeter becomes non-negligible.

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{2.7}$$

where S is the stochastic term, N is the noise term, and C is the constant term. There are three basic contributions to the stochastic term:

- Event-to-event fluctuations in the lateral shower containment.
- Fluctuations in the number of photoelectrons released by the photodetectors (contribution of 2.1%).
- Fluctuations in the energy deposited in the preshower absorber with respect to what is measured in the preshower silicon detector.

The three contributions to the noise term are:

- Electronics noise.
- Digitization noise.
- Pileup noise.

The most important contributions to the constant term are:

- Non-uniformity in the longitudinal light collection.
- Inter-channel calibration errors.
- Leakage of energy from the back of the crystal.

A typical empirical parameterization of the resolution found in a test beam for the EB is:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%)^2 \tag{2.8}$$

where E is given in units of GeV [33].

2.2.7 Hadron Calorimeter

The hadron calorimeter (Hcal) is particularly important for the measurement of hadron jets and neutrinos or exotic particles resulting in apparent missing transverse energy [34]. Figure 2.11 shows the longitudinal view of the CMS detector, including the locations of the hadron barrel (HB), endcap (HE), outer (HO), and forward (HF) calorimeters. The dashed lines are at fixed η values. HB and HE sit behind the tracker and the Ecal as seen from the IP. HB is radially restricted between the outer extent of the Ecal ($r = 1.77$ m) and the inner extent of the magnet coil ($r = 2.95$ m). This constrains the total amount of material that can be put in to absorb the hadronic showers. Therefore, an outer hadron calorimeter or tail catcher is placed outside the solenoid to complement the barrel calorimeter. Beyond $|\eta| = 3$, the HF calorimeters extend the pseudorapidity coverage up to $|\eta| = 5.2$ using a Cherenkov based, radiation-hard technology. They are placed at a longitudinal distance of 11.2 m from the nominal IP.

Figure 2.11: Longitudinal view of the CMS detector showing the locations of the HB, HE, HO, and HF calorimeters.

Barrel Design

The HB is a sampling calorimeter covering the pseudorapidity range $|\eta| < 1.3$. The calorimeter consists of 36 identical azimuthal wedges that form the two half-barrels (HB+ and HB−). The wedges are constructed out of flat brass absorber plates aligned parallel to the beam axis. Each wedge is segmented into four azimuthal angle sectors. The plates are bolted together in a staggered geometry, resulting in a configuration that contains no projective dead material for the full radial extent of a wedge (Figure 2.12). The innermost and outermost plates are made of stainless steel for structural strength. The plastic scintillator is divided into 16 η sectors, resulting in a segmentation of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The wedges themselves are bolted together in such a fashion as to minimize the cracks between the wedges to less than 2 mm.

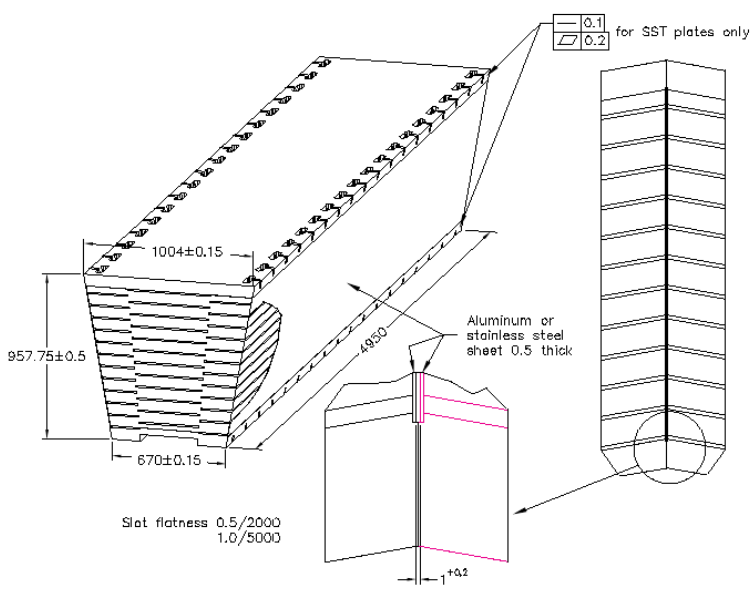


Figure 2.12: Isometric view of the HB wedges, which shows the hermetic design of the scintillator sampling.

The absorber (Table 2.2) consists of a 40 mm thick front steel plate, followed by eight 50.5 mm thick brass plates, six 56.5 mm thick brass plates, and one 75 mm thick steel back plate. The total absorber thickness at 90° is 5.82 interaction lengths (λ_I). The HB

effective thickness increases with polar angle as $1/\sin\theta$, resulting in 10.6 λ_I at $|\eta| = 1.3$. The Ecal [35] sits in front of HB and adds about 1.1 λ_I of material.

Table 2.2: Absorber thickness in the HB wedges.

Layer(s)	Material	Thickness (mm)
front plate	steel	40
1-8	brass	50.5
9-14	brass	56.5
back plate	steel	75

The active scintillating medium uses the well-known tile and wavelength shifting fiber concept to collect the light. The CMS hadron calorimeter consists of about 70000 scintillator tiles. The tiles of a given ϕ layer are grouped into a single mechanical scintillator tray unit in order to limit the number of individual elements to be handled. The HB baseline active material is 3.7 mm thick Kuraray SCSN81 plastic scintillator, chosen for its long-term stability and moderate radiation hardness. The zeroth layer of scintillator is located in front of the steel support plate, and is made of 9 mm thick Bicron BC408. Its purpose is to sample hadronic showers developing in the inert material between the EB and HB. The larger thickness of layer 16 serves to correct for late developing showers leaking out the back of HB. The scintillators are summarized in Table 2.3.

Table 2.3: HB Wedge scintillator layers.

Layer(s)	Material	Thickness (mm)
0	Bicron BC408	9
1-15	Kuraray SCSN81	3.7
16	Kuraray SCSN81	9

After exiting the scintillator, the wavelength shifting fibers are spliced to clear fibers. The clear fibers are terminated by optical connectors at the ends of the trays. Optical cables then direct the light to optical decoding units. The decoding units arrange the fibers into read-out towers and bring the light to hybrid photodiodes [36], which convert the light into currents.

Endcap Design

The hadron calorimeter endcaps [37] cover a substantial portion of the rapidity range, $|\eta| \in (1.3, 3)$ (13.2% of the solid angle). This region contains about 34% of the particles produced in the final state. The high design luminosity of the LHC requires HE to handle MHz-level counting rates and have high radiation tolerance (e.g., The expected dose after 10 years of operation at design luminosity is 10 Mrad at $|\eta| \approx 3$). Since the calorimeter is inserted into the ends of a 4 T solenoidal magnet, the absorber must be made from a non-magnetic material. It must also have a maximum number of interaction lengths to contain hadronic showers, robust mechanical properties, and reasonable cost. These lead to the choice of C26000 cartridge brass. The endcaps are attached to the muon endcap yoke as shown in Figure 2.13.

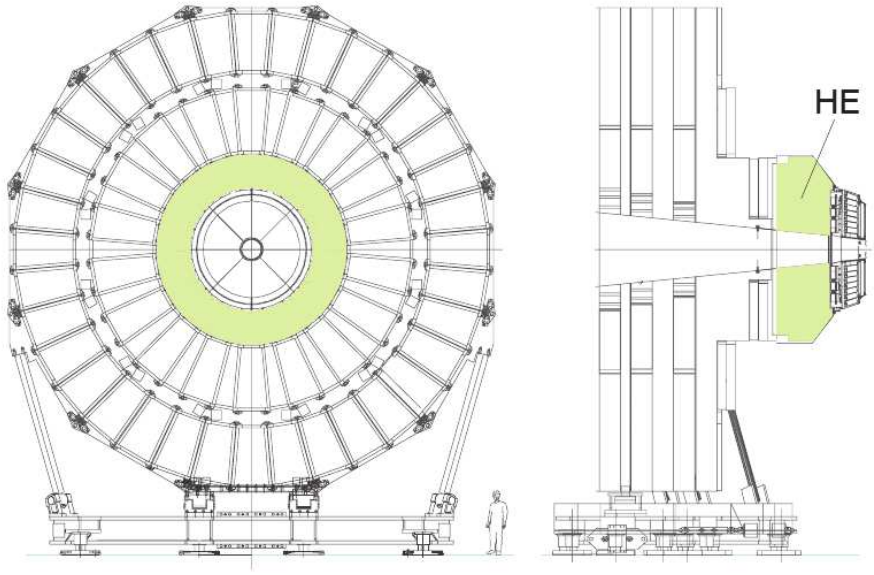


Figure 2.13: The hadronic endcap calorimeter mounted on the endcap iron yoke.

The design of the absorber is driven by the need to minimize the cracks between HB and HE, rather than single particle energy resolution, since the resolution of jets in the HE will be limited by pileup, magnetic field effects, and parton fragmentation [38, 39]. The plates are bolted together in a staggered geometry resulting in a configuration that

contains no projective "dead" material. The design provides a self-supporting hermetic construction. The brass plates are 79 mm thick with 9 mm gaps to house the scintillators. The total length of the calorimeter, including the electromagnetic crystals, is about $\lambda_I = 10$.

The scintillation light is collected by wavelength shifting fibers [40, 41] inserted into trapezoidal shaped scintillators 3.7 mm (9 mm) thick SCSN81 (Bicron BC408) for layers 1-17 (0). The ends of the fibers are spliced to a clear fiber, which is terminated in an optical connector. The light is converted to a current by hybrid photodiodes. The granularity of the HE is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ for $|\eta| \in (1.3, 1.6)$ and $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ for $|\eta| \in (1.6, 3)$.

Outer Calorimeter Design

The combined stopping power of EB plus HB does not provide sufficient containment for hadron showers in the central pseudorapidity region. To ensure adequate sampling depth for $|\eta| < 1.3$, the hadron calorimeter is extended outside the solenoid with a tail catcher called the HO or outer calorimeter. The HO utilizes the solenoid coil as an additional absorber equal to $1.4/\sin\theta$ interaction lengths. It is used to identify late starting showers and measure the shower energy deposited after HB. The detector is situated as the first sensitive layer of each of the five rings (along the z-axis) composing the iron yoke that returns the magnetic field of the solenoid. Figure 2.14 shows the position of HO layers in the rings of the magnetic field return.

HB has the minimal absorber depth at $\eta = 0$. Therefore, the central ring (ring zero) has two layers of HO scintillators on either side of a 19.5 cm thick piece of iron (the tail catcher iron) at radial distances of 3.82 m and 4.07 m, respectively. The other four rings (rings ± 1 and ± 2) have a single HO layer at a radial distance of 4.07 m. The total depth of the calorimeter system is thus extended to a minimum of 11.8 λ_I except at the barrel-endcap boundary region. The sizes and positions of the scintillator tiles in HO roughly map to the layers of HB, making towers of granularity $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$.

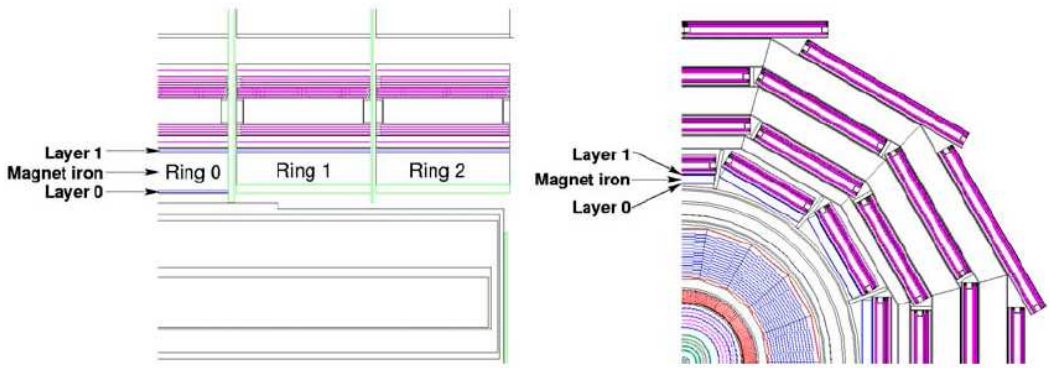


Figure 2.14: Longitudinal and transverse views of the CMS detector showing the position of HO layers.

Scintillation light from the tiles is collected using wavelength shifting fibers and transported to the photodetectors located on the structure of the return yoke by splicing a clear fiber with the wave length shifting fibers.

Forward Calorimeter Design

HF will experience unprecedented particle fluxes. On average, 760 GeV per pp interaction are deposited into the two forward calorimeters, compared to only 100 GeV for the rest of the detector. Moreover, this energy is not uniformly distributed but rather has a pronounced maximum at the highest pseudorapidities. The HF will have sustained a radiation dose of 10 MGy at $|\eta| = 5$ after 10 years of LHC operation. The charged hadron flux will also be extremely high. In the same duration, the flux will exceed 10^{11} per cm^2 inside the HF absorber at 125 cm from the beam line [37]. This hostile environment presents a considerable challenge to calorimetry, and the design of the HF calorimeter was first and foremost guided by the necessity to survive in these harsh conditions for at least a decade. Successful operation critically depends on the radiation hardness of the active material. This was the principal reason why quartz fibers (fused-silica core and polymer hard-cladding) were chosen as the active medium.

The signal is generated when charged shower particles propagate through the quartz

fibers above the Cherenkov threshold ($E > 190$ keV for electrons) and generate Cherenkov light. This renders the calorimeter mostly sensitive to the electromagnetic component of showers [42]. The two HF calorimeters (HF+ and HF−) are composed of these fibers encased in cylindrical steel matrices with outer radii of 130.0 cm. The light is converted to currents by photomultiplier tubes. The front face of the HF± calorimeter is located at ±11.2 m from the interaction point along the z-axis, as seen in Figure 2.11. The hole for the beam pipe is cylindrical, with radius 12.5 cm from the center of the beam line. This structure is azimuthally subdivided into 20° modular wedges: 36 such wedges (18 on either side of the interaction point) make up the HF calorimeters. The quartz fibers run parallel to the beam line, and are bundled to form towers of granularity $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$. The detector is housed in a hermetic radiation shielding consisting of a 40 cm thick layer of steel, a 40 cm thick layer of concrete, and a 5 cm thick layer of polyethylene. A large plug structure in the back of the detector provides additional shielding.

2.2.8 Muon System

Muon detection is a powerful tool for recognizing signatures of interesting processes over the very high background rate present at the LHC. For example, the predicted decay of the Standard Model Higgs boson into ZZ or ZZ*, which in turn decay into four leptons, has been called "gold plated" for the case in which all the leptons are muons. Besides the relative ease in detecting muons, the best four-particle mass resolution can be achieved if all the leptons are muons because they are less affected than electrons by radiative losses in the tracker material. This example, and others from supersymmetric [43] models, emphasize the discovery potential of muon final states and the necessity for muon detection with wide angular coverage.

Therefore, as is implied by the experiment's middle name, the detection of muons is of central importance to CMS. The muon system has three functions: muon identification, momentum measurement, and triggering. Good muon momentum resolution and trigger capability are enabled by the high field solenoidal magnet and its flux-return yoke. The

latter also serves as a hadron absorber for the identification of muons. The material thickness crossed by muons is shown in Figure 2.15 as a function of $|\eta|$.

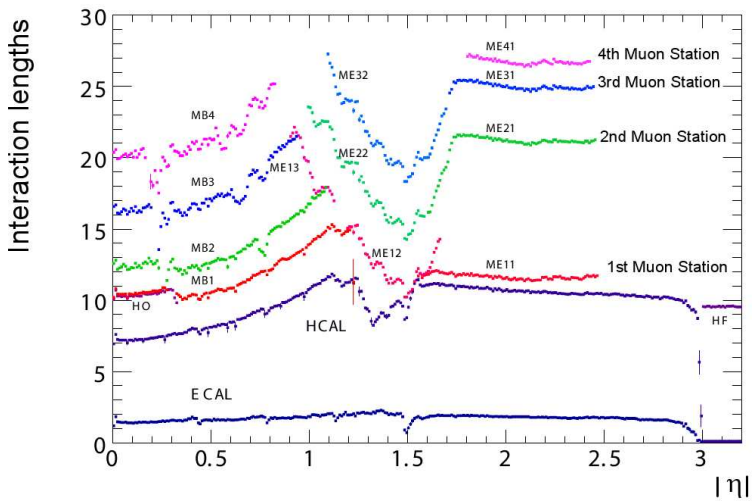


Figure 2.15: Material thickness in interaction lengths at various depths as a function of absolute pseudorapidity.

The muon system is designed to reconstruct the momentum and charge of muons over the entire kinematic range of the LHC. Three types of gaseous particle detectors are used for muon identification [44]. The muon system has a cylindrical barrel section closed by two planar endcaps. Because the system consists of about 2.5×10^4 m^2 of detection planes, the muon chambers had to be inexpensive, reliable, and robust.

Drift Tubes

Drift chambers with standard rectangular drift cells are used in the barrel region, where the neutron-induced background is small, the muon rate is low, and the 4 T magnetic field is uniform and mostly contained in the steel yoke. The barrel drift tube (DT) chambers cover the pseudorapidity region $|\eta| < 1.2$ and are organized into four stations interspersed among the layers of the flux return plates, forming concentric cylinders around the beam line: The three inner cylinders have 60 drift chambers each and the outer cylinder has 70. There are about 172000 sensitive wires. It is possible to use drift chambers as the tracking

detectors for the barrel muon system because of the low expected rate and the relatively low strength of the local magnetic field.

The wire length is approximately 2.4 m in the chambers measured in an r-ϕ projection, and is constrained by the longitudinal segmentation of the iron barrel yoke. The transverse dimension of the drift cell, i.e., the maximum path and time of drift, was chosen to be 21 mm (corresponding to a drift time of 380 ns in a gas mixture of 85% Ar + 15% CO_2). This value is small enough to produce a negligible occupancy and to avoid the need for multi-hit electronics, yet the cell is large enough to limit the number of active channels to an affordable value. A tube was chosen as the basic drift unit to obtain protection against damage from a broken wire and to partially decouple contiguous cells from the electromagnetic debris accompanying the muon itself.

The amount of iron in the return yoke was dictated by the decision to have a large and intense solenoidal magnetic field at the core of CMS. Two detector layers, the first inside the yoke and the other outside, would be insufficient for reliable identification and measurement of a muon. Therefore, two additional layers are embedded within the yoke iron (Figure 2.16). There are four muon chambers per wheel, labeled MB1, MB2, MB3, and MB4, in each of the 12 sectors of the yoke. The yoke iron supports that are between the chambers of a station generate 12 unavoidable dead zones in the ϕ coverage, although the supports are placed so as not to overlap in ϕ.

A DT chamber consists of two or three superlayers (SLs, Figure 2.17), itself being composed of four layers of rectangular drift cells staggered by half a cell. The SL is the smallest independent unit of the design. The wires in the two outer SLs are parallel to the beam line and provide a track measurement in the magnetic-bending r-ϕ plane. In the inner SL, the wires are orthogonal to the beam line and measure the z position along the beam. This third z-measuring SL is not present in the fourth station, which therefore measures only the ϕ-coordinate. A muon coming from the interaction point first encounters a ϕ-measuring SL, passes through the honeycomb plate, then crosses the z-measuring SL and the second ϕ-measuring SL. In this scenario, there still exist limited

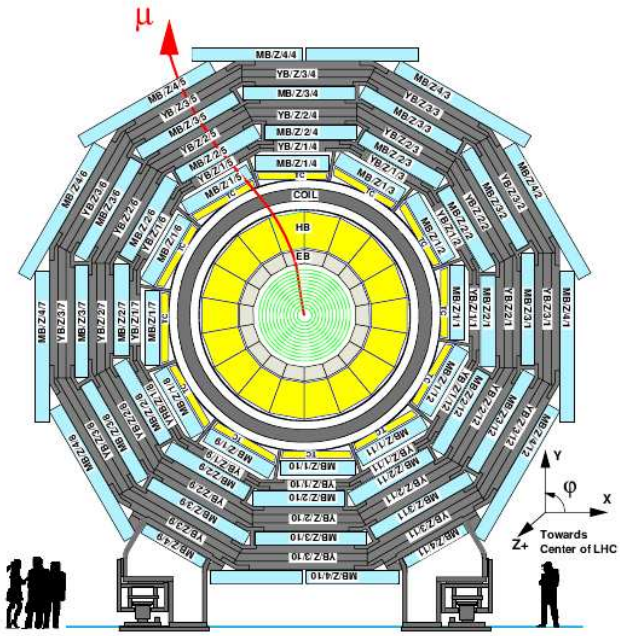


Figure 2.16: Layout of the CMS barrel muon DT chambers in one of the five wheels. The chambers in each wheel are identical with the exception of wheels -1 and $+1$, where the presence of cryogenic chimneys for the magnet shortens the chambers in two sectors. The MB4 chambers are cut in half in sectors 4 (top) and 10 (bottom) to simplify the mechanical assembly and the global chamber layout.

regions of η in which the combined effect of the ϕ and z discontinuities limits the number of stations crossed by a muon to only two out of four.

At high momenta of at least 40 GeV, the probability of electromagnetic cascades accompanying the parent muon becomes relevant. A reliable way to cope with this effect in the regions where only two stations are available is to have a good tracking efficiency in each station, even in the presence of electromagnetic debris. Redundancy is also needed to cope with the uncorrelated background hits generated by neutrons and photons, whose rates are much larger than those from prompt muons. Redundancy is obtained by having several layers of separated drift cells per station. The separation, i.e., the thickness of the tube walls, should be large enough to decouple the basic units from low energy electrons. The relatively thick 1.5 mm walls of the DTs give an effective decoupling among the

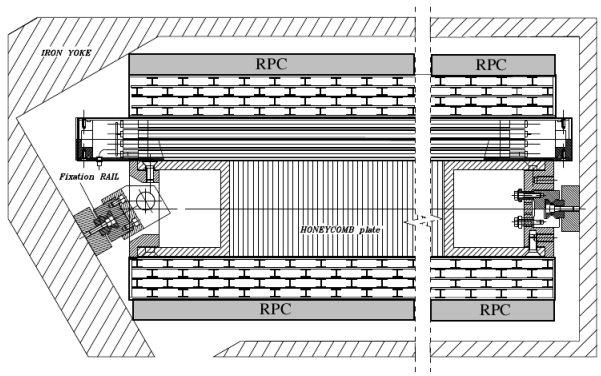


Figure 2.17: r-ϕ view of A DT chamber in position inside the iron yoke. One can see the two SLs with wires along the beam direction and the other perpendicular to it. In between is a honeycomb plate with supports attached to the iron yoke. Not shown are the resistive plate chambers, which are attached to the DT chambers via support plates glued to the bottom and/or top faces, depending on chamber type.

several layers of tubes inside the same station. The efficiency to reconstruct a high P_T muon track with a momentum measurement delivered by the barrel muon system alone is better than 95% in the pseudorapidity range covered by four stations, i.e., $|\eta| < 0.8$. The constraints of mechanical stability, limited space, and redundancy led to the choice of a tube cross section of 13×42 mm^2.

Cathode Strip Chambers

The Muon Endcap system consists of 468 cathode strip chambers (CSCs) arranged in groups as follows: 72 ME1/1, 72 ME1/2, 72 ME1/3, 36 ME2/1, 72 ME2/2, 36 ME3/1, 72 ME3/2, and 36 ME4/1 (Figure 2.18). The chambers are trapezoidal and cover either 10° or 20° in ϕ. All chambers, except for the ME1/3 ring, overlap and provide contiguous ϕ coverage. A muon in the pseudorapidity range $|\eta| \in (1.2, 2.4)$ crosses three or four CSCs. In the endcap-barrel overlap range of $|\eta| \in (0.9, 1.2)$, muons are detected by both the barrel DTs and endcap CSCs. Muons with $|\eta| < 2.1$ are also detected by resistive plate chambers in the baseline design; however, in the initial detector setup this coverage is reduced to $|\eta| < 1.6$.

The CSCs are multi-wire proportional chambers comprised of six anode wire planes

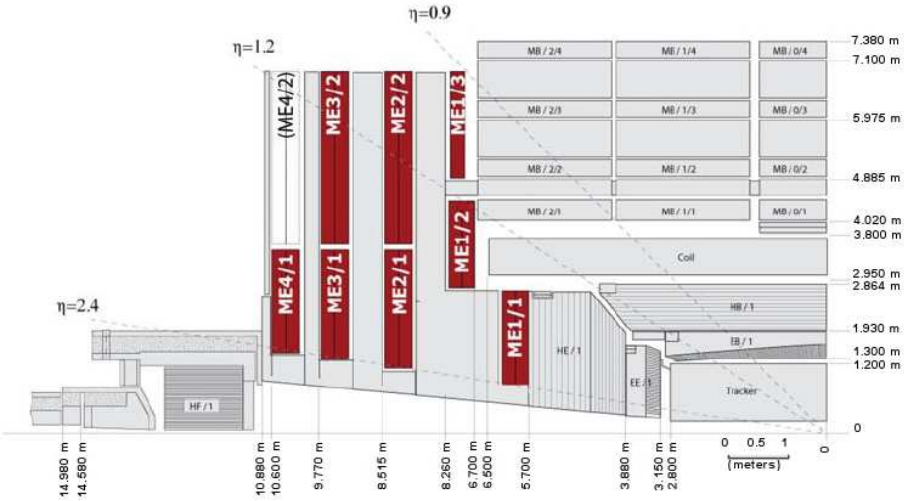


Figure 2.18: Quarter-view of the CMS detector. Cathode strip chambers of the Endcap Muon system are highlighted.

interleaved among seven cathode panels. Wires run azimuthally and define the radial coordinate of a track. Strips are milled on cathode panels and run lengthwise at constant $\Delta\phi$ width. Following the original CSC concept [45], the muon coordinate along the wires is obtained by interpolating charges induced on strips. The largest chambers, ME2/2 and ME3/2, are about 3.4×1.5 m^2 in size. The overall area covered by the sensitive planes of all chambers is about 5×10^3 m^2. The gas volume is greater than 50 m^3 and the total number of wires is about 2×10^6. There are about 9×10^3 high-voltage channels in the system, about 2.2×10^5 cathode strip read-out channels with 12-bit signal digitization, and about 1.8×10^5 anode wire read-out channels.

The CSCs provide the functions of precision muon measurement and muon trigger in one device. They can operate at high rates and in large, non-uniform magnetic fields. They do not require precise gas, temperature, or pressure control. Moreover, a radial fan-shaped strip pattern, natural for measurements in the endcap region, can easily be arranged on the cathode planes.

The performance requirements for the CSC system include:

- Reliable and low maintenance operation for at least 10 years at LHC design luminosity; i.e., at estimated random hit rates up to 1 kHz/cm^2.
- At least 99% efficiency per chamber for finding track stubs by the Level-1 trigger.
- At least 92% probability per chamber of identifying correct bunch crossings by the Level-1 trigger. With such an efficiency per chamber and three to four CSCs on a muon track path, a simple majority rule ensures that the reconstructed muons will be assigned to the correct bunch crossing number with a probability in excess of 99%.
- About 2 mm resolution in r-ϕ at the Level-1 trigger.
- About 75 μm offline spatial resolution in r-ϕ for ME1/1 and ME1/2 chambers, and about 150 μm for all others.

Resistive Plate Chambers

Resistive Plate Chambers (RPCs) are gaseous parallel-plate detectors that combine adequate spatial resolution with a time resolution comparable to that of scintillators [46, 47]. RPCs are capable of tagging the time of an ionizing event on a much shorter scale than the 25 ns between two consecutive LHC bunch crossings. Therefore, a fast dedicated muon trigger device based on RPCs can unambiguously identify the bunch crossing to which a muon track belongs even in the presence of the high rate and background present at the LHC. Signals from these devices also provide the position of a muon hit with the required accuracy for triggering.

The RPCs provide fast, independent, and highly segmented triggers with a sharp P_T threshold. They cover the large pseudorapidity range $|\eta| < 1.6$ and are double-gap chambers (Figure 2.19), operated in avalanche mode. Their excellent time resolution is countered by their coarse position resolution, which is poorer than the DTs or CSCs. Six layers of RPCs are embedded in the barrel muon system: two in each of the first two stations, and one in each of the last two stations. The redundancy in the first two stations allows triggering of low P_T tracks that stop before reaching the outer two stations. In the

endcap region, each of the first three stations contains a plane of RPCs. This allows the trigger to use the coincidences between stations to reduce background, improves the time resolution for bunch crossing identification, and improves the P_T resolution.

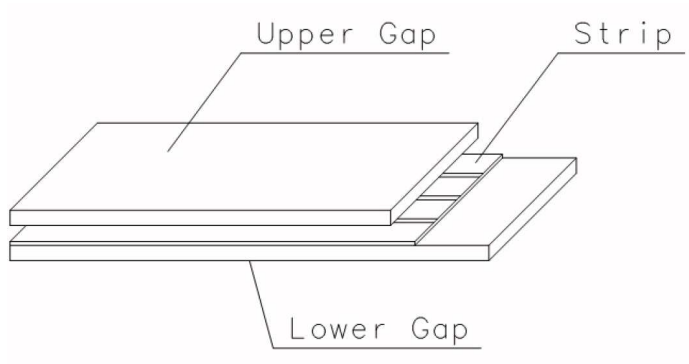


Figure 2.19: Layout of a double-gap RPC.

2.2.9 Trigger

The LHC provides *pp* collisions at high interaction rates. The beam crossing interval is 25 ns, corresponding to a crossing frequency of 40 MHz. Approximately 25 simultaneous *pp* collisions occur each crossing at the nominal design luminosity of 10^{34} cm^{-2} s^{-1}. Since it is impossible to store and process the large amount of data associated with the resulting high number of events, a drastic rate reduction has to be achieved. This task is performed by the trigger system, which is the start of the physics event selection process. The rate is reduced in two steps corresponding to the the Level-1 Trigger (L1) [48] and High-Level Trigger (HLT) [5, 49]. The L1 consists of custom designed, largely programmable electronics. The HLT is a software system implemented in a filter farm of about 1000 commercial processors with a latency of approximately 30 ms. The rate reduction capability is designed to be at least a factor of 10^6 for the combined L1 and HLT. The design output rate limit of the L1 is 100 kHz, which translates in practice to a calculated maximal output rate of 30 kHz, assuming an approximate safety factor of three. The L1 trigger uses coarsely segmented data from the calorimeters and the muon system, while

holding the high-resolution data in pipelined memories in the front-end electronics. The final output rate limit of the HLT is 100 Hz. The HLT has access to the complete read-out data and can therefore perform complex calculations similar to those made in offline analyses. The L1 and HLT electron algorithms are discussed in Chapter 4 and [49][6] . For reasons of flexibility, the L1 trigger hardware is implemented in FPGA technology where possible, but ASICs and programmable memory look-up tables are also widely used where speed, density, and radiation resistance requirements are important. A software system called XDAQ controls the configuration and operation of the trigger components.

The L1 has local, regional, and global components. At the bottom end, the Local Triggers, also called Trigger Primitive Generators (TPGs), are based on energy deposits in the calorimeter trigger towers and track segments or hit patterns in the muon chambers. Regional Triggers combine their information and use pattern logic to determine ranked and sorted trigger objects such as electron or muon candidates in limited spatial regions. The rank is determined as a function of energy or momentum and quality, which reflects the level of confidence attributed to the L1 parameter measurements, based on detailed knowledge of the detectors and trigger electronics and on the amount of information available. The Global Calorimeter and Global Muon Triggers determine the highest-rank calorimeter and muon objects across the entire experiment and transfer them to the Global Trigger, which is the top entity of the L1 trigger hierarchy. The Global Trigger makes the decision to reject an event or to accept it for further evaluation by the HLT. The decision is based on algorithm calculations and on the readiness of the sub-detectors and the data acquisition (DAQ), which is determined by the Trigger Control System (TCS). The Level-1 Accept (L1A) decision is communicated to the sub-detectors through the Timing, Trigger, and Control (TTC) system. The architecture of the L1 is depicted in Figure 2.20. The L1 trigger has to analyze every bunch crossing. The allowed latency, between a given bunch crossing and the distribution of the trigger decision to the detector front-end electronics, is 3.2 μs. The processing must therefore be pipelined in order to

[6] Co-written by the author.

enable quasi-deadtime-free operation. The L1 electronics are housed partly on the detectors and partly in the underground control room located at a distance of approximately 90 m from the experimental cavern.

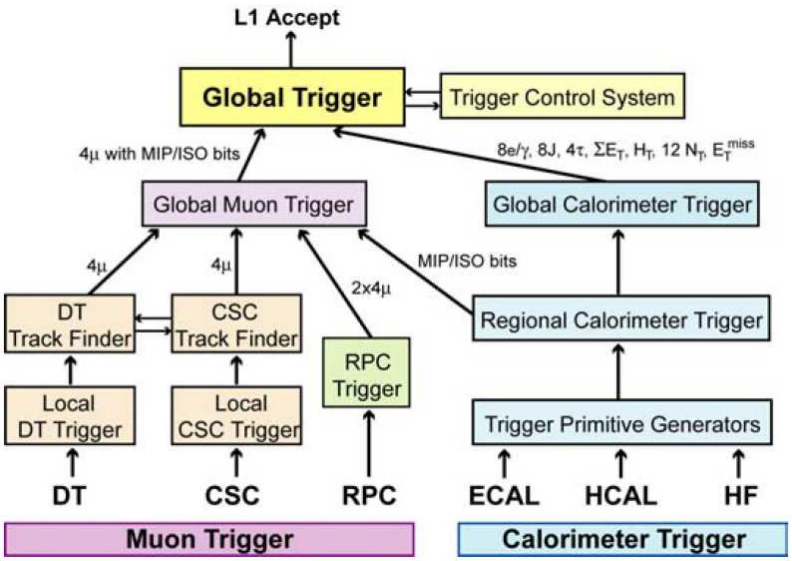


Figure 2.20: Architecture of the Level-1 Trigger.

Chapter 3

The Luminosity System

3.1 Goals and Requirements

The luminosity measurement is used to monitor the performance of the LHC on a bunch-by-bunch basis in real-time and provide an overall normalization for physics analyses. The design goal of the real-time measurement is to determine the average luminosity with a 1% statistical accuracy with an update rate of 1 Hz. For offline analyses, the goal is to achieve a systematic accuracy of 5%, although every effort is being made to produce a more accurate result. Both of these requirements must be met over a very large range of luminosities, extending from roughly 10^{28}–10^{34} cm^{-2}s^{-1} and possibly beyond.

3.2 Real-Time Luminosity Monitoring Using the Hadronic Forward Calorimeters

The Hadronic Forward calorimeters (HF, Subsection 2.2.7) are used as the primary real-time luminosity monitors [50][1] . The number of interactions per bunch crossing is Poisson distributed with a mean μ that is proportional to the luminosity $\mathcal{L}$:

$$\mu = \frac{\sigma(pp \to X)\mathcal{L}}{f_{\mathrm{BX}}} \tag{3.1}$$

[1] Co-written by the author.

where $\sigma(pp \to X)$ is the total inelastic proton-proton collision cross section and f_{BX} is the bunch crossing rate. There are two methods for extracting a real-time luminosity signal from the HF by inferring μ.

Zero Counting Method

In this method, the occupancy of the individual HF towers is examined each bunch crossing in order to infer μ. Each of the 864 HF towers has some probability p of being empty in a single interaction event and the number of empty towers is binomially distributed with probability p. In a given bunch crossing with k interactions, the expected fraction of empty towers is p^k. Averaged over a large number of bunch crossings M, with the number of interactions per bunch crossing distributed according to a Poisson distribution with mean μ, the expected fraction of empty towers $\langle f_0 \rangle$ is:

$$\langle f_0 \rangle = \frac{1}{M} \sum_{k=0}^{\infty} M \frac{e^{-\mu}\mu^k}{k!} p^k = \sum_{k=0}^{\infty} \frac{e^{-\mu}\mu^k}{k!} p^k = e^{\mu(p-1)} \tag{3.2}$$

In the limit that p is close to unity (which is the case for most HF rings [50]), this becomes simply p^μ. $\langle f_0 \rangle$ is then the measurable quantity and may be expressed as:

$$\langle f_0 \rangle = e^{\mu(p-1)} \approx p^\mu \tag{3.3}$$

The logarithm of Equation 3.3 is proportional to the mean number of interactions per bunch crossing and hence to the luminosity.

Thus in principle one needs only measure the average fraction of empty towers per bunch crossing in order to determine the relative luminosity. A slightly more detailed analysis is required to actually apply this technique though: The effects of noise—spurious signals in the HF not caused by incoming particles—must be accounted for. Even towers that have not received energy from a particle generated in a pp collision may record some energy in the presence of noise. A low energy threshold is introduced, below which towers are considered empty. Such a threshold removes some of the signal distribution, changing the form of Equation 3.3.

To fully determine the effects of noise and the energy threshold, one must start from the probability density f as a function of the energy x of a given tower. For a single interaction in the presence of zero noise, this is given by:

$$f(x) = p\delta(x) + (1-p)s(x) \tag{3.4}$$

where p is again the probability that a tower is empty and $s(x)$ is the energy probability density of a hit tower. Let $n(x)$ be the probability density as a function of energy for noise hits. The total energy probability density for a single interaction F is then the convolution of these two functions:

$$\begin{aligned} F(x) &= (f \otimes n)(x) \\ &= \int f(x\,')n(x-x\,')dx\,' \\ &= \int (p\delta(x\,')n(x-x\,') + (1-p)s(x\,')n(x-x\,'))dx\,' \\ &= pn(x) + (1-p)(s \otimes n)(x) \end{aligned} \tag{3.5}$$

Generally, a given bunch crossing contains k interactions, in which case the total probability density is:

$$F^{(k)}(x) = (\overbrace{f \otimes f \otimes \ldots \otimes f}^{k \text{ times}} \otimes n)(x) \tag{3.6}$$

Expanding and simplifying, the following relation is obtained for the general case:

$$F^{(k)}(x) = p^k n(x) + \sum_{i=1}^{k} \binom{k}{i} p^{k-i}(1-p)^i (\overbrace{s \otimes s \otimes \ldots \otimes s}^{i \text{ times}} \otimes n)(x) \tag{3.7}$$

The probability P that a given tower is below the chosen threshold T may be evaluated using this formula. For a single interaction it is:

$$P(x < T) = \int_0^T (pn(x) + (1-p)(s \otimes n)(x))dx \tag{3.8}$$

Defining $r_n \equiv \int_0^T n(x)dx$ and $r_{s\otimes n} \equiv \int_0^T (s \otimes n)(x)dx$, this reduces to:

$$P(x < T) = pr_n + (1-p)r_{s\otimes n} \tag{3.9}$$

In an event with k interactions this becomes:

$$P^{(k)}(x < T) = \int_0^T (p^k n(x) + kp^{k-1}(1-p)(s \otimes n)(x) + O((1-p)^2))dx \tag{3.10}$$

Since $(1-p)$ is small (e.g., from simulation p is between 0.82–0.99 depending on the η ring), and if it is assumed that the threshold is small compared to the mean of $s(x)$ (i.e., $r_{s\otimes n}$ is small), this may be approximated as:

$$P^{(k)}(x < T) \approx p^k r_n + kp^{k-1}(1-p)r_{s\otimes n} \tag{3.11}$$

Averaging over the Poisson distribution in k as before yields:

$$\langle f_0 \rangle = \sum_{k=0}^{\infty} \frac{e^{-\mu}\mu^k}{k!}(p^k r_n + kp^{k-1}(1-p)r_{s\otimes n}) = e^{\mu(p-1)}\left(r_n + (1-p)\mu r_{s\otimes n}\right) \tag{3.12}$$

So that:

$$-\ln\langle f_0 \rangle = (1-p)\mu - \ln\left(r_n + (1-p)\mu r_{s\otimes n}\right) \tag{3.13}$$

Assuming $(1-p)\mu r_{s\otimes n}$ is small:

$$-\ln\left(r_n + (1-p)r_{s\otimes n}\mu\right) \sim -\ln r_n - \frac{(1-p)r_{s\otimes n}\mu}{r_n} + O(((1-p)r_{s\otimes n}\mu)^2) \tag{3.14}$$

which gives the final result:

$$-\ln\langle f_0 \rangle = (1-p)(1 - \frac{r_{s\otimes n}}{r_n})\mu - \ln r_n \tag{3.15}$$

So long as the appropriate quantities are small (as suggested by the simulation [50]) , then linearity in the average number of interactions per bunch crossing and hence luminosity is maintained, with some offset due to noise given by the second term in the right hand side of Equation 3.15. Because of the correction of order $r_{s\otimes n}/r_n$, the slope is sensitive to changes in the pedestal width and gain, however.

The fraction of empty towers diminishes for $\mu \gg 1$. In practice, this effects the measurement only for luminosities at least 5 times greater than the design luminosity. That is, the tower occupancy measurement may be untenable for luminosities above 5×10^{34} cm^{-2}s^{-1}.

$\sum E_T$ Method

A second method for measuring the real-time luminosity exploits the linear relationship between the average transverse energy E_T deposited in an HF tower and the number of interactions per bunch crossing [50]. This method has two important advantages:

- It does not saturate at high luminosities.
- An energy threshold cut is not required as in the zero counting method.

The second advantage is important since threshold cuts can make the measurement sensitive to pedestal levels and widths, photomultiplier tube gains, etc. From Equation 3.7, the expected E_T deposited per tower in a single interaction can be determined as:

$$\begin{aligned}
\langle E_T \rangle^{(1)} &= \frac{1}{\cosh\eta}\int xF(x)dx \\
&= \frac{1}{\cosh\eta}\int (pxn(x) + (1-p)x(s\otimes n)(x))dx \\
&= \frac{1}{\cosh\eta}(p\nu_n{}' + (1-p)(\nu_s{}' + \nu_n{}')) \\
&= \nu_n + \nu_s(1-p) \qquad (3.16)
\end{aligned}$$

where:

$$\nu_n \equiv \nu_n'(\cosh\eta)^{-1} = (\cosh\eta)^{-1}\int xn(x)dx \qquad (3.17)$$

$$\nu_s \equiv \nu_s'(\cosh\eta)^{-1} = (\cosh\eta)^{-1}\int xs(x)dx \qquad (3.18)$$

For a bunch crossing with k interactions it is easy to show that:

$$\langle E_T \rangle^{(k)} = \nu_n + k\nu_s(1-p) \qquad (3.19)$$

Once again the number of interactions per bunch crossing is Poisson distributed with mean μ, leading to:

$$\langle E_T \rangle = \frac{1}{M}\sum_{k=0}^{\infty} M\frac{e^{-\mu}\mu^k}{k!}(\nu_n + k\nu_s(1-p)) = \nu_n + \nu_s(1-p)\mu \qquad (3.20)$$

Thus the average transverse energy deposited per tower is indeed proportional to the number of interactions per bunch crossing.

3.3 Hardware Implementation

The luminosity system requires only a small amount of extra hardware since the HF and its read-out are already part of the detector design. A mezzanine board, called the HF Luminosity Transmitter (HLX), is mounted in each of the 36 Hadronic Trigger Read-Out (HTR) boards. The HLXs collect the luminosity data and transmit it to servers via Ethernet. The HTR receives long and short fiber signals from each of the physical HF towers: The long fibers, which are most sensitive to hadronic activity, from the inner four HF azimuthal rings are used by the HLX to create tower occupancy and $\sum E_T$ histograms. The inner four HF rings correspond to $|\eta| \in (3.5, 4.2)$ and were chosen to avoid suffering from non-linearities inherent in averaging over a large pseudorapidity range in which the average tower occupancy varies strongly. The histograms are time-accumulating and take at least 2^{14} LHC orbits (about 1.45 s) to overflow. Leaving a safety margin, the data is transmitted by the HLXs to the luminosity servers and the histograms reset at a rate of about 2.7 Hz.

3.4 Luminosity Read-Out Path

The design of the luminosity read-out is shown in Figure 3.1. The HLX data is transmitted by an Ethernet switch to servers that collect and distribute the data. The first server receives the data from the switch and is responsible for archiving it every luminosity section (approximately 23 s) in the online database. An additional layer on this server provides communication with the central CMS run control in order to start and stop the luminosity data acquisition software. The luminosity system may still operate independently of the central run control, so that an "Always ON" functionality is achieved. The second server receives the luminosity data from the first and is responsible for publishing it to various clients such as the CMS and LHC control rooms. Fail-over servers stand ready in case of problems with the primary machines.

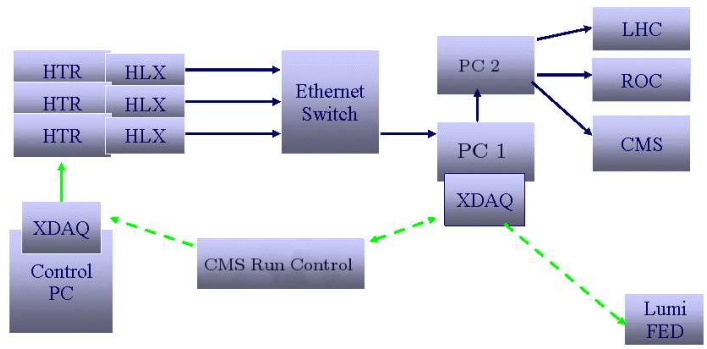


Figure 3.1: Design of the luminosity read-out path.

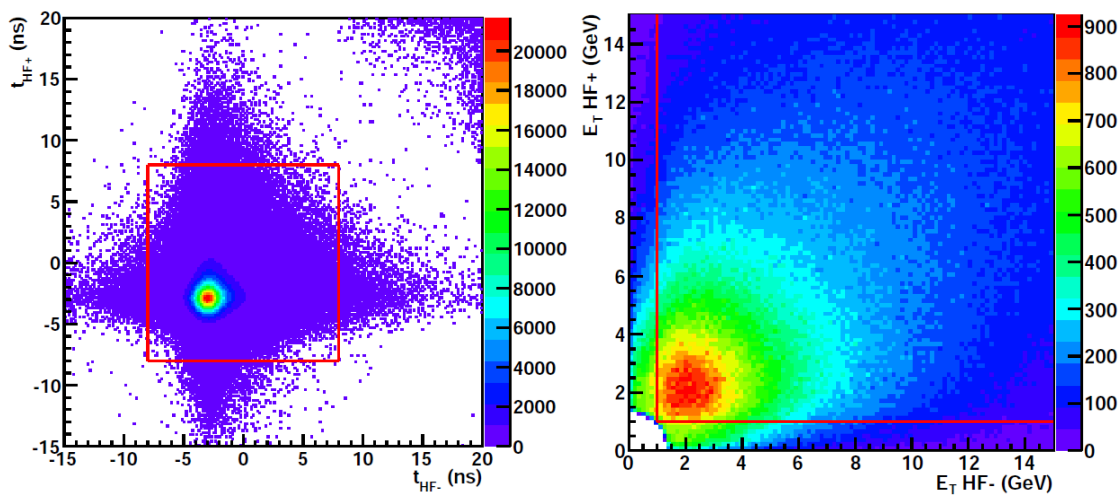
3.5 Offline Method

As a cross check of the HF-based online luminosity monitor, there is an offline method for monitoring applied to two different sub-detectors. The measurements use a zero counting method similar to that used online, except in this case only one measurement is performed per sub-dector. (Online, separate measurements are performed for each HF tower and are then combined.) That is, μ is inferred by counting the number of empty bunch crossings and inverting the Poisson zero probability. The offline measurements have the drawbacks of a relatively low number of samplings and long latencies (typically about 24 h elapse before the offline information from a given run is available). Nevertheless, the offline technique employs a largely independent data-handling path, and in the case of one of its applications, involves a completely separate set of systematic uncertainties.

Offline HF Measurement

This measurement is based on the coincidence of the $\sum E_T$ depositions in the HF+ and HF− calorimeters above a threshold of 1 GeV. Timing cuts, where $|t_{\mathrm{HF}}| < 8$ ns for both HF+ and HF−, are imposed to eliminate non-collision backgrounds. Bunch crossings in which such a coincidence satisfying these criteria does *not* occur then contribute to the zero counting. Figure 3.2 shows timing and energy information from a run at $\sqrt{s} = 7$ TeV (LHC fill 1089). Figure 3.2 (a) is a scatter plot of the time in HF+ vs. the time in HF−

for all events in the sample (the offset of the time scale is arbitrary). The sharp cluster at $t_{\mathrm{HF}+} \approx t_{\mathrm{HF}-} \approx -2$ ns is the signal. Figure 3.2 (b) shows the transverse energy in HF+ vs. the transverse energy in HF−. Entries in the region $E_T^{\mathrm{HF}+} \approx E_T^{\mathrm{HF}-} \approx 0$ have been eliminated to suppress a large noise (pedestal) peak near the origin.



(a) (b)

Figure 3.2: The distributions of the variables cut on in the offline HF event selection: (a) $t_{\mathrm{HF}+}$ vs. $t_{\mathrm{HF}-}$, (b) $E_T^{\mathrm{HF}+}$ vs. $E_T^{\mathrm{HF}-}$. The cut values are represented by red lines.

Vertex Measurement

This measurement requires at least one vertex having at least two tracks to be found in the event. The z-position of the vertex is required to be less than 150 mm from the nominal IP. Otherwise, the event contributes to the zero counting. Additional details regarding tracking and vertexing in CMS are available here [51].

3.6 Absolute Calibration using Van der Meer Separation Scans

The separation scan method for absolute luminosity determination was pioneered by S. van der Meer at CERN's Intersecting Storage Rings in the 1960s [52]. The size and shape of the interaction region is measured by recording the relative interaction rate as a function of the transverse beam separations. If the beam profile (in terms of proton density) in x and y is given by the function $F(x, y) = f_x(x)f_y(y)$, then the instantaneous luminosity is:

$$\mathcal{L}_0 = \frac{N_1 N_2 \nu_{\text{orbit}} N_{\text{b}} F(0,0)}{\int f_x(\Delta x) d\Delta x \int f_y(\Delta y) d\Delta y} \tag{3.21}$$

where $\mathcal{L}_0$ is the peak instantaneous luminosity, N_i is the bunch intensity in beam i, ν_{orbit} is the LHC orbit frequency, N_{b} is the number of colliding bunches per beam, and Δx (Δy) is the beam separation in the x (y) plane.

The beam intensities are determined using Fast Beam Current Transformers (FBCTs), which measure the current in each LHC bunch [53]. The FBCT measurements provide accurate bunch-to-bunch values and are normalized to a low-bandwidth measurement of the total circulating current made by DC current transformers.

In order to fit the tails of the distributions observed, it is necessary to use double-Gaussian distributions for the functions f_x and f_y:

$$f_x(x) = \frac{h_x}{\sqrt{2\pi}\sigma_{1x}} e^{\frac{-x^2}{2\sigma_{1x}^2}} + \frac{(1-h_x)}{\sqrt{2\pi}\sigma_{2x}} e^{\frac{-x^2}{2\sigma_{2x}^2}} \tag{3.22}$$

where h_x is the fractional area of the double-Gaussian belonging to the first term, and similarly for f_y. Inserting these functional forms into Equation 3.21 yields:

$$\mathcal{L}_0 \equiv \frac{N_1 N_2 \nu_{\text{orbit}} N_{\text{b}}}{2\pi\sigma_{\text{eff}}(x)\sigma_{\text{eff}}(y)} \equiv \frac{N}{\sigma} \tag{3.23}$$

where the effective beam size[2] $\sigma_{\text{eff}}(j)$ for each scan plane j is given by:

$$\sigma_{\text{eff}}(j) \equiv \frac{\sigma_{1j}\sigma_{2j}}{h_j\sigma_{2j} + (1-h_j)\sigma_{1j}} \tag{3.24}$$

[2]The beam size here is the convolution of both beams.

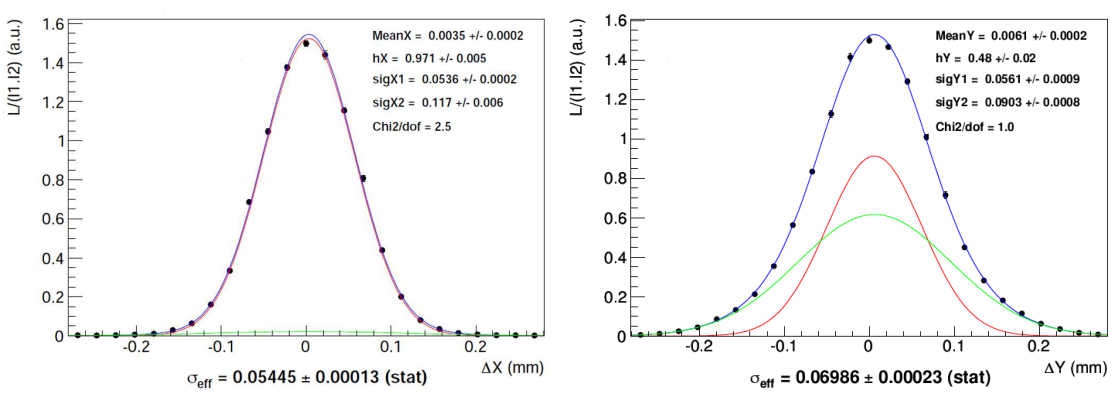
In general the luminosity as a function of the beam separation d is:

$$\mathcal{L}(d) = \mathcal{L}_0 \left(\frac{h_j}{\sqrt{2\pi}\sigma_{1j}} e^{\frac{-d^2}{2\sigma_{1j}^2}} + \frac{(1-h_j)}{\sqrt{2\pi}\sigma_{2j}} e^{\frac{-d^2}{2\sigma_{2j}^2}} \right) \tag{3.25}$$

In addition to calibration, the scans also permit luminosity optimization by determining the $d = 0$ beam separation point in the horizontal and vertical axes.[3]

3.7 Separation Scan Measurements

Multiple separation scans were performed during 2010 to calibrate and optimize the luminosity at LHC Point 5. Figure 3.3 shows the result of a calibration scan taken during LHC fill 1058 for the (a) horizontal and (b) vertical axes. The response of the HF luminometer in terms of uncalibrated interaction rates is plotted as a function of the beam separation. The plot is fit to a double-Gaussian to allow determination of the effective beam size as well as the absolute luminosity via Equation 3.23.



(a) (b)

Figure 3.3: Van der Meer separation scan results from LHC fill 1058. The blue curve is the total double-Gaussian, the red curve is the core Gaussian (σ_1), and the green curve is the Gaussian for the tails (σ_2): (a) horizontal axis, (b) vertical axis.

[3]This proved invaluable in the early days of LHC start-up: The beams were quite separated and repeated scanning performed by the author along with his colleagues in the LHC control room allowed proper positioning of the beams, increasing the luminosity at LHC Point 5 by a factor of approximately 3.

These calibrations, and the 2010 luminosity measurement in general, are documented in [54, 8, 55]. The systematic uncertainty breakdown of the calibration is in Table 3.1. The total systematic uncertainty on the luminosity measurement is 4.0%.

Table 3.1: Summary of the luminosity calibration systematic uncertainties [8].

Source	$\Delta\mathcal{L}/\mathcal{L}$
Beam Background	< 0.001
Method and Fill Variation	0.025
Beam Shape	0.003
Vertical Scale Non-Linearity	0.004
Length-Scale Calibration	0.003
Emittance Slope	0.002
Non-Linear Correlations	0.009
Beam Intensity	0.029
HF Drift	0.005
Afterglow	0.005
Total	0.040

3.8 Summary

The design and performance of the CMS luminosity system have been detailed. Methods for estimating the relative luminosity both on and offline have been described. In 2010, these methods were observed to provide good consistency [54]. The absolute luminosity has been calibrated using Van der Meer separation scans to a systematic accuracy of 4.0%, already below the design goal of 5%. Thus a high accuracy normalization has been provided for physics analyses performed at CMS. (Still, this 4.0% error is the dominant uncertainty on the $Z \to ee$ cross section measurement presented in this thesis.) Scans have also been performed to maximize the delivered luminosity. The delivered and recorded integrated luminosity profiles for 2010 are shown in Figure 3.4.[4] The total integrated

[4]Recall that the delivered luminosity is all that provided to the detector by the LHC, while the recorded luminosity is the part of the delivered luminosity actually collected by CMS. Thus the ratio of recorded to delivered is a measure of data taking efficiency (about 92% in 2010).

luminosity for the full 2010 dataset analyzed in this thesis is:

$$\int \mathcal{L}dt = 35.9 \pm 1.4(\text{syst})\ \text{pb}^{-1} \tag{3.26}$$

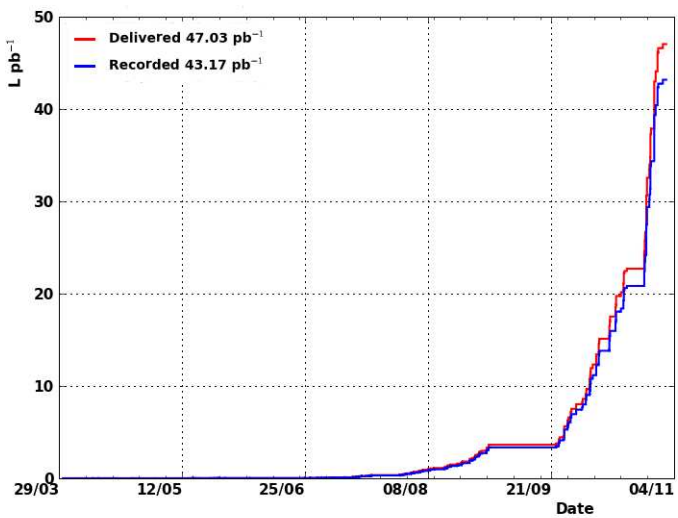


Figure 3.4: The 2010 delivered (red) and recorded (blue) integrated luminosity profiles.

Chapter 4

Offline Electron Reconstruction and Single Electron Triggers

4.1 Offline Electron Reconstruction

An offline electron object [56] is defined as clusters of Ecal energy deposits matched to a track from the silicon tracker. To reconstruct an electron object, first its energy deposition is recovered in the Ecal, then hits in the pixel tracker are searched for about these energy depositions, and finally an associated track is built in the silicon tracker from the pixel hits.

4.1.1 Ecal Clustering and Superclustering

Electron and photon showers deposit their energy in several crystals in the Ecal. Approximately 94% (97%) of the incident energy of a single electron or photon is contained in a 3×3 (5×5) crystal array [56]. Summing the energy measured in such fixed arrays provides the best performance for photons or electrons in a field-free environment that excludes material in front of the crystals. However, in the realistic detector environment material in front of the calorimeter leads electrons to undergo bremsstrahlung and photons to convert, and the energy reaching the calorimeter is spread

in ϕ due to the magnetic field. This spread energy is recovered for electron and photon objects by building a cluster of clusters of such energy depositions that is extended in ϕ, referred to as a supercluster.

Superclustering Algorithms

The reconstruction of electrons starts with building clusters seeded by hot cells in the Ecal. These clusters are then used to form superclusters that further collect the energy radiated by bremsstrahlung in the tracker volume. The Hybrid algorithm is used in the Ecal Barrel ($|\eta| < 1.5$). It leverages the η-ϕ geometry of the crystals to exploit the knowledge of the lateral shower shape in the η direction (taking a fixed bar of three or five crystals in η), while searching dynamically for separated energy deposits in the ϕ direction. Starting from the highest energy seed crystal, ϕ varies to a maximal extension of 17° in both directions. The Multi-5×5 algorithm is used in the Ecal Endcaps ($|\eta| > 1.5$). It first collects the energy deposited within 5×5 crystal arrays and then forms superclusters by grouping such arrays whose positions lie within a ϕ road of extension 17° in both directions.

4.1.2 Pixel Seed Association

The first step of electron track reconstruction is finding seeds in the pixel detector. For electrons, which suffer radiative losses, this requires dedicated strategies to preserve efficiency while minimizing fake track rates. The finding of the first two track hits, or seeds, in the pixel detector is driven by the electromagnetic superclusters. The supercluster to pixel hit matching takes advantage of the fact that the energy-weighted average impact point of the electron and associated bremsstrahlung photons, as calculated using information from the supercluster in the Ecal, coincides with the impact point that would have been measured for a non-radiating electron of the same initial momentum. Hits in the pixel layers are predicted by propagation of the energy-weighted mean position of the supercluster backward through the magnetic field towards the pixel detector under

both charge hypotheses. A first compatible hit is then searched for in the innermost (barrel) pixel layer within loose $\Delta\phi$ and Δz windows. When the first compatible hit is found, a new estimate for the z-coordinate of the primary track vertex is calculated by combining information from this hit with that from the calorimeter in the r-z plane. The predicted trajectory is then propagated to search for a second pixel hit in the next pixel layer(s), within some narrower $\Delta\phi$ and Δz window. The average efficiency of the pixel finding for single electrons typically reaches about 90% for electrons with $P_T = 10$ GeV, depending on the exact definitions of the pixel window criteria. The two pixel hits thus found serve as the seeds for the building and fitting of the silicon tracker track to be associated with the electron object.

4.1.3 Track Association

The default track reconstruction method in CMS relies on a Kalman Filter (KF) algorithm. The algorithm has been shown to work for high P_T electrons provided that a tight χ^2 cut is used in the trajectory building [49]. In this way, the emphasis is put on the early stages of the electron track evolution that contain the most significant information on the electron initial momentum and direction from the primary vertex. A different approach is needed, however, for detailed electron analyses.

The KF algorithm is equivalent to a global least-square minimization based on a linear model for the track evolution and Gaussian fluctuations. The assumption of considering all random fluctuations of track parameters due to material effects as Gaussian is roughly valid when the dispersion is mainly caused by multiple scattering processes, but clearly fails for the energy radiated in the thin layers of the tracker material. A non-linear filter approach such as the Gaussian Sum Filter (GSF) [57] thus better describes the propagation of electrons through the tracker. The GSF is a non-linear generalization of the KF in which the distributions of all state vectors and errors are Gaussian mixtures. The weights of the components of the mixture depend on the measurements.

The performance for transverse momentum reconstruction is illustrated for low

$P_T = 10$ GeV electrons in Figure 4.1. The best measurement of the electron P_T at the primary vertex is obtained from the most probable values of the mixture of Gaussians characterizing the track state at each tracker layer. More detailed discussions on the electron GSF tracks and measurement performances at the primary vertex can be found in [58]. When compared to the simple KF approach mentioned above, the GSF tracks are found to provide a very similar momentum resolution and slightly improved η and ϕ measurements from $P_T = 5$ GeV up to at least 30 GeV.

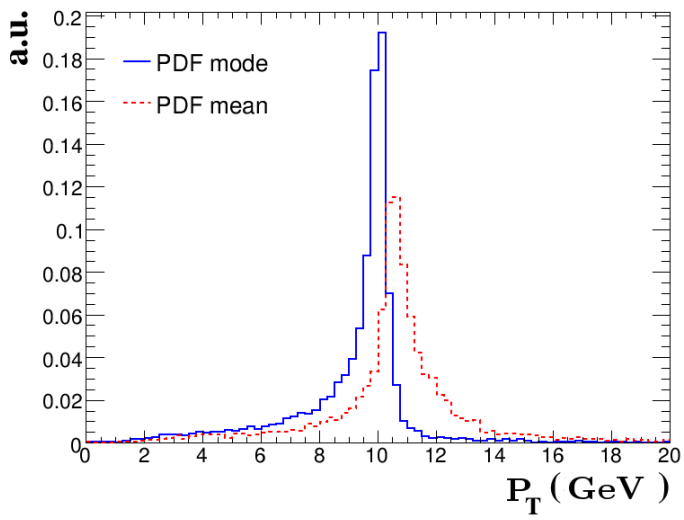
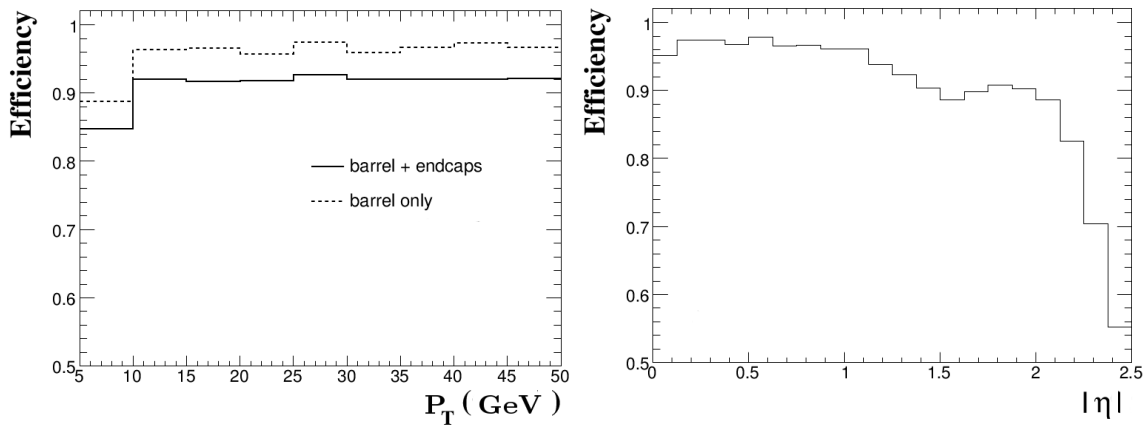


Figure 4.1: Track reconstruction performance for electrons of $P_T = 10$ GeV: reconstructed P_T for Gaussian Sum Filter tracks as obtained from the most probable value (blue), and the weighted mean (red) of the Gaussian mixture.

The efficiency of the electron track reconstruction is shown as functions of P_T and η in Figure 4.2. It is above 90% with a few exceptions: at very low $P_T < 10$ GeV, in the transition region between the Ecal Barrel and Endcap ($|\eta| \approx 1.5$), and towards the edge of the η acceptance.

A benefit of the GSF tracking algorithm comes from the combined facts that hits are collected efficiently along the full trajectory through the tracker volume and meaningful track parameter errors are available at both track ends. Thus a good estimation of the electron track parameters at the Ecal boundary is made available. But most importantly,



(a) (b)

Figure 4.2: Electron track reconstruction efficiency: (a) as a function of P_T either averaged over the full η range of the Ecal Barrel and Endcap (solid) or just the Ecal Barrel (dotted), (b) as a function of $|\eta|$ for a uniform distribution in P_T in the range $(5, 50)$ GeV.

the fractional amount of momentum carried away by bremsstrahlung photons can be evaluated from the outermost and innermost track parameters. This helps in improving the electron energy measurements, distinguishing various electron patterns, and electron identification. For example, classes of low-radiating electrons can be defined in a novel way and used for Ecal calibration purposes using $J/\psi \to ee$ decays [31][1] . The bremsstrahlung based classification of electrons also provides a means of correcting the supercluster energy. Generally, the energy assigned to an electron object is calculated from a weighted average of the corrected supercluster energy and track momentum measurements.[2]

4.2 Single Electron Triggers

This section provides an overview of the high E_T single electron triggers used in the $Z \to ee$ analysis. All events passing the L1 trigger are read-out and processed in the HLT farm.

[1] Co-written by the author.

[2] This weighted average is dominated by the supercluster energy in the case of the high $E_T > 25$ GeV electrons selected in this analysis.

4.2.1 Level-1 Electromagnetic Trigger

The primary design goal of the L1 trigger is to achieve a rejection factor of order 10^3 with a latency of 3.2 μs, while still maintaining reasonable efficiencies. The L1 electromagnetic (EM) trigger [48] is based on Ecal trigger towers; no attempt is made to distinguish between electrons and photons at this stage. Energy deposits in trigger towers (corresponding to 5×5 crystal arrays in the Ecal Barrel and varying collections in the Endcaps) are classified as isolated or non-isolated. An overview of the L1 EM trigger isolation algorithm is shown in Figure 4.3. The algorithm involves only the eight nearest neighbors around the central hit trigger tower and is applied over the entire (η, ϕ) plane. The electron/photon candidate E_T is determined by summing the E_T in the hit tower with the maximum E_T tower of its four broad side neighbors (white in Figure 4.3). Using this summed transverse energy provides a sharper efficiency turn-on with the true E_T of the particles than using only the E_T of the hit tower.

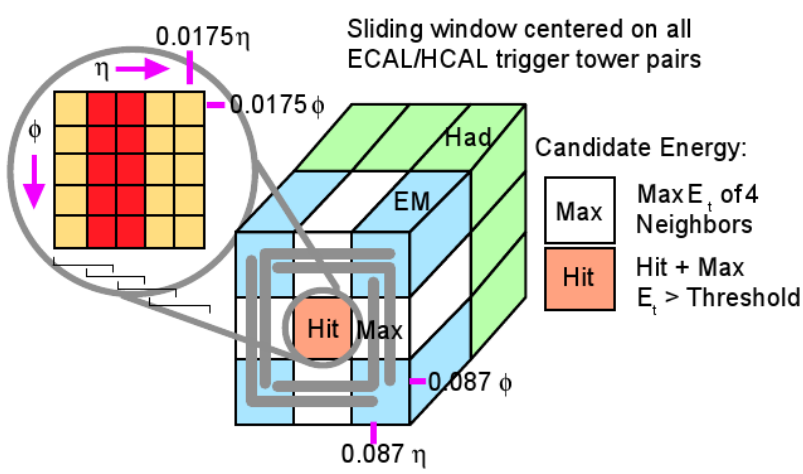


Figure 4.3: The electron/photon L1 Trigger isolation algorithm.

A non-isolated candidate is required to pass two shower profile vetoes. The first veto is based on the Fine-Grain Ecal crystal energy profile. The second veto is based on a comparison of the Hcal to Ecal trigger tower energies, e.g., H/E less than 5%, and is called the Hadronic Calorimeter (HAC) veto. The Fine-Grain veto is evaluated as follows:

The energy released in each pair of adjacent 2×5 crystal strips (in $\eta \times \phi$, as shown in red in Figure 4.3) is calculated for the hit trigger tower. The maximum energy found in such a strip is compared to the total energy deposition in the trigger tower. If the ratio of these two quantities is below the programmed threshold, the electron/photon candidate is vetoed; otherwise it is kept. The HAC veto is evaluated by comparing the ratio of the E_T in the Hcal trigger tower immediately behind the hit Ecal trigger tower to that of the hit Ecal trigger tower. If this H/E value is above the programmed threshold, then the electron/photon candidate is vetoed; otherwise it is kept. A minimum threshold may be programmed for either the Fine-Grain or HAC veto, so that the veto is not applied when the E_T of the hit trigger tower is below threshold.

In addition to having the hit trigger tower pass the Fine-Grain and HAC vetoes, an isolated candidate must have all eight neighboring towers pass these vetoes as well. One of the four adjacent trigger towers (blue in Figure 4.3) must also be "quiet", having all crystals below a programmable threshold (e.g., 1.5 GeV).

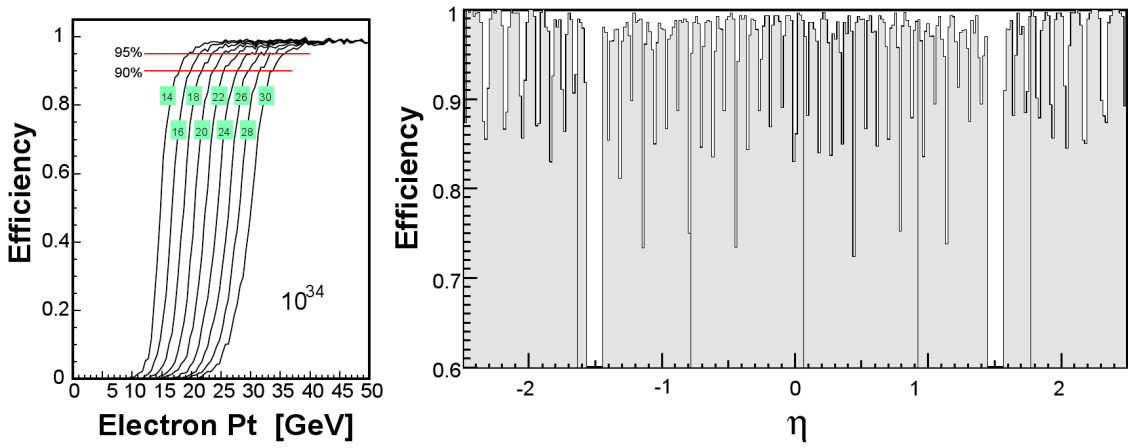
Each candidate is characterized by the (η, ϕ) indices of the calorimeter region where the hit trigger tower is located. In each calorimeter region (4×4 trigger towers) the highest E_T non-isolated and isolated electron/photon candidates are found separately. The 16 candidates of both types found in a regional trigger crate (corresponding to 16 calorimeter regions and covering $\Delta\eta \times \Delta\phi = 3.0 \times 0.7$) are further sorted by transverse energy. The four highest E_T candidates of both types from each crate are transferred to the global calorimeter trigger, where the highest four E_T candidates are retained for processing by the global trigger.

Level-1 EM Trigger Summary

The nominal electron/photon algorithm allows both non-isolated and isolated streams. The non-isolated stream uses only the hit tower information, except for adding in any leakage energy from the maximum neighbor tower. The isolation cuts are programmable and can be adjusted to the running conditions. For example, at high luminosity the

isolation cuts could be relaxed to take into account higher pileup energies. The specification of the electron/photon triggers also includes the definition of the η-ϕ region where it is applicable. In particular, it is possible to define different trigger conditions (energy thresholds, isolation cuts, etc.) in different pseudorapidity regions.

Figure 4.4 (a) plots The efficiency of the L1 electron/photon algorithm as a function of the electron transverse momentum for different E_T thresholds (green boxes). Figure 4.4 (b) plots the L1 trigger efficiency as a function of pseudorapidity for electrons with P_T = 35 GeV. Only the non-isolated L1 EM triggers were used during 2010.



(a) (b)

Figure 4.4: The efficiency of the L1 EM algorithms [5]: (a) as a function of the electron P_T with different E_T thresholds, (b) as a function of η for electrons with $P_T = 35$ GeV.

4.2.2 Electron High Level Trigger

After an event passes the L1, it is sent to the HLT [49][3] . The primary design goal of the HLT is to achieve a rejection factor of order 10^3 with a latency of approximately 30 ms, while still maintaining reasonable efficiencies. The electron HLT comprises a sequence of increasingly complex and sophisticated reconstruction and filter algorithms that are referred to as modules and executed on the HLT CPU farm. As the complexity of the

[3]Co-written by the author.

modules increases, so does the processing time. Hence events are rejected and processing stopped as early on as possible to maintain low latency. For CPU performance, the reconstruction is performed regionally in small volumes of the detector corresponding to the locations of the L1 trigger objects in the event. Each module takes as input the candidates passing the previous filter step, following the chain:

1. Regional Ecal supercluster reconstruction, as described in Section 4.1.
2. Supercluster E_T threshold filter.
3. Ecal based electron isolation and identification filters.
4. Regional Hcal tower reconstruction.
5. Hcal based electron isolation and identification filters.
6. Global pixel hit reconstruction, as described in Section 4.1, and association of the electron object to two pixel hits.
7. Reconstruction of the electron track and association with the supercluster, as described in Section 4.1, but using a Kalman Filter for speed rather than a Gaussian Sum Filter.
8. Electron track based electron identification filters (e.g., filter according to whether the supercluster was successfully associated to a track).
9. Regional tracker reconstruction: Tracks are seeded from the pairs of pixel hits found in the pixel hit reconstruction within a region near to that of the electron object.
10. Track based electron identification and isolation filters.

The online electron isolation and identification variables all have close offline analogues. The offline variables are described in Chapter 6. Keeping the discrimination variables used online as coherent as possible with those used offline is a primary concern for electron analyses. In addition to ensuring the best overall electron object performance,

this minimizes the difficulty in understanding the associated systematic uncertainties. More details concerning this and how the thresholds of the isolation/identification variables can be optimized are available here [59].[4]

While this chapter is concerned with the single high E_T electron triggers, it should be noted that other types of electron triggers are present at CMS. For instance, both L1 and HLT can handle multiple object triggers. One example running at CMS is a low E_T dielectron filter that triggers on $J/\psi \to ee$ and $\Upsilon \to ee$ decays [31].[5]

4.3 Summary

This chapter has described the electron reconstruction and single electron triggers. The first step in reconstructing an electron object is building a supercluster, which contains the associated energy deposition in the Ecal. Next, near-by hits in the tracker are found so that they can seed the construction of a full track, which is then associated with the electron object. At L1, no distinction between electrons and photons is made. Both isolated and non-isolated EM objects are selected according to their E_T and the isolation algorithm described in Subsection 4.2.1. The selected L1 EM objects are processed by the HLT, which runs a fast, regionalized version of the electron reconstruction. Events are filtered as quickly as possible in the HLT by software modules that cut on the E_T of the object in addition to various electron identification and isolation variables, which are similar to those used in offline analyses. The specific triggers used in the $Z \to ee$ cross section measurement are detailed in Chapter 6.

[4] Co-written by the author.
[5] Co-written by the author.

Chapter 5

Data and Monte-Carlo Samples

5.1 Data Sample

The data analyzed in this thesis corresponds to the full 2010 CMS dataset collected during LHC $\sqrt{s} = 7$ TeV pp collision running. It was collected between March 30 and October 31, 2010, and corresponds to an integrated luminosity of 35.9 ± 1.4 pb^{-1}. The luminosity profile rapidly increased throughout data taking: Hence the electron trigger used changed multiple times. The changes in the trigger are documented in Chapter 6; the average electron trigger efficiency throughout data taking was about 98%. All data analyzed[1] has been put through a quality validation chain requiring all detector subsystems to be flagged as good.

5.2 Monte-Carlo Samples

While the techniques used in this analysis rely primarily on the data—that is, they are "data-driven"—Monte-Carlo (MC) is used for comparisons to the data and evaluation of several systematic uncertainties. The signal and background MC samples are listed in Table 5.1. The electroweak samples with W and Z production have been produced using

[1] This analysis heavily relies on the CMS software framework (CMSSW) [56], the ROOT data analysis package [60], and the RooFit data modeling toolkit [61].

the Powheg [62] NLO generator interfaced with Pythia [63] for the parton showering.[2] The other samples have been produced using the Pythia LO generator. The QCD dijet samples have been generated using an electromagnetic-enriching filter that requires an isolated electron in the final state at generator-level [64][3] . The filter increases the probability of a reconstructed electron candidate being present in the event and hence decreases the production time. The electromagnetic-enriched QCD dijet samples do not contain decays of c or b quarks to electrons. These heavy flavor events and the $t\bar{t}$ events have been produced separately. The Powheg samples use the CT10 [65] PDF set, while the Pythia samples use the CTEQ6L1 [66] set. The quoted cross sections for the electroweak samples with W and Z production have been calculated to NNLO using FEWZ [67, 68]. The quoted cross section for the $t\bar{t}$ sample has been calculated to NLO using MCFM [69]. The remaining samples have their cross sections taken at LO directly from Pythia. All events are processed through the full Geant4 [70, 71] detector simulation, trigger emulator [72], and event reconstruction chain. The samples use the Z2 tune of the underlying event structure performed on the first data [73].

Table 5.1: Signal and background MC samples.

Category	Process	Generator	PDF Set	σ (pb)	$\int L$dt (pb^{-1})
Signal	$Z \to ee$ ($M_{ee} > 20$ GeV)	Powheg	CT10	1.67×10^3	1.27×10^3
QCD	Dijets ($\hat{P}_T \in (20, 30)$ GeV)	Pythia	CTEQ6L1	2.45×10^6	15.0
	Dijets ($\hat{P}_T \in (30, 80)$ GeV)	Pythia	CTEQ6L1	3.87×10^6	18.6
	Dijets ($\hat{P}_T \in (80, 170)$ GeV)	Pythia	CTEQ6L1	1.40×10^5	43.6
QCD Heavy Flavor	$b/c \to e$ ($\hat{P}_T \in (20, 30)$ GeV)	Pythia	CTEQ6L1	1.32×10^5	17.0
	$b/c \to e$ ($\hat{P}_T \in (30, 80)$ GeV)	Pythia	CTEQ6L1	1.37×10^5	14.6
	$b/c \to e$ ($\hat{P}_T \in (80, 170)$ GeV)	Pythia	CTEQ6L1	9.36×10^3	111
$t\bar{t}$	$t\bar{t}$	Pythia	CTEQ6L1	158	6.98×10^3
Electroweak	$Z \to \tau\tau$ ($M_{\tau\tau} > 20$ GeV)	Powheg	CT10	1.67×10^3	1.23×10^3
	W+jets	Powheg	CT10	1.04×10^4	1.45×10^3
	WW	Powheg	CT10	43.0	4.64×10^4
	WZ	Powheg	CT10	18.2	4.77×10^4
	ZZ	Powheg	CT10	5.9	3.22×10^5

[2]The Powheg Z samples include interference from γ^*, but following the convention described in Section 1.3 they are referred to simply as Z rather than Z/γ^* samples.

[3]Co-written by the author.

Chapter 6

$Z \rightarrow ee$ Event Selection and Performance

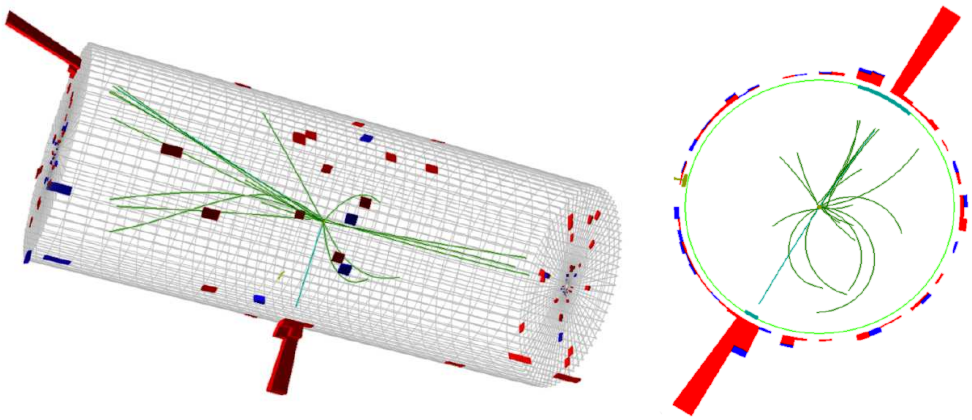
The $Z \rightarrow ee$ event selection and its performance are detailed in this chapter. The acceptance criteria for the selected $Z \rightarrow ee$ candidates are defined, and the electron triggers used are summarized. The electron identification variables are introduced and their optimization is described. The cross section measured in this analysis pertains to the invariant mass window $M_{ee} \in (60, 120)$ GeV. Thus the $Z \rightarrow ee$ candidates are selected within this window.[1] A total of 8441 $Z \rightarrow ee$ candidates were collected during 2010 data taking. The plots of these candidates highlight the performance of the detector and reveal that the data is in good agreement with the MC.

6.1 $Z \rightarrow ee$ Event Selection

The $Z \rightarrow ee$ signal is characterized by the presence of two isolated high P_T electrons having M_{ee} near the mass of the Z resonance (about 91.2 GeV). This is highlighted in Figure 6.1, which shows the CMS detector event display for a typical signal event. The energy

[1] The Z boson four-vector is defined as the sum of the four-vectors of its two daughter electron objects. (e.g., The electron P_T and direction of its associated track are used instead of the Ecal supercluster transverse energy and direction).

deposits in the Ecal (Hcal) are represented by red (blue) bars, tracks are represented by green curves, and reconstructed electrons are represented by cyan curves. The Ecal deposits corresponding to the two electrons from the Z decay are well separated from other calorimeter deposits, as are the reconstructed electrons from other tracks.



(a) (b)

Figure 6.1: The CMS detector event display for a typical $Z \to ee$ event ((run, luminosity section, event) = $(133877, 387, 28405693)$). The energy deposits in the Ecal (Hcal) are represented by red (blue) bars, tracks are represented by green curves, and reconstructed electrons are represented by cyan curves: (a) three-dimensional view, (b) r-ϕ plane view.

6.1.1 Fiducial and Kinematic Acceptance

Both superclusters associated with the daughter electrons of the $Z \to ee$ decays are required to be within the Ecal fiducial region of $|\eta| \in (0, 1.4442)$ or $|\eta| \in (1.566, 2.5)$, where it is assumed the superclusters point to the nominal IP. Both electrons are selected above a supercluster E_T threshold of 25 GeV, where the superclusters point to the primary vertex of the event.

6.1.2 Online Electron Selection

The electron triggers have been described in Section 4.2. For a $Z \rightarrow ee$ candidate to pass the selection, one of the associated electrons must be matched within a cone of aperture $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.3$ to an object passing the electron trigger being used when the event was collected. The luminosity profile increased dramatically during 2010, forcing multiple changes in the trigger as the run progressed to control the output rate. The trigger used in each period is documented in Table 6.1, and may be summarized as:

1. L1: Ecal trigger tower with $E_T > 5$ GeV.
   HLT: supercluster with $E_T > 10$ GeV matched to the L1 trigger tower.

2. L1: Ecal trigger tower with $E_T > 5$ GeV.
   HLT: supercluster with $E_T > 15$ GeV matched to the L1 trigger tower.

3. L1: Ecal trigger tower with $E_T > 5$ GeV.
   HLT: electron object with $E_T > 15$ GeV matched to the L1 trigger tower and passing loose identification criteria.

4. L1: Ecal trigger tower with $E_T > 8$ GeV.
   HLT: electron object with $E_T > 17$ GeV matched to the L1 trigger tower and passing loose identification criteria.

5. L1: Ecal trigger tower with $E_T > 8$ GeV.
   HLT: electron object with $E_T > 17$ GeV matched to the L1 trigger tower and passing identification criteria.

6. L1: Ecal trigger tower with $E_T > 8$ GeV.
   HLT: electron object with $E_T > 22$ GeV matched to the L1 trigger tower and passing identification and isolation criteria.

7. L1: Ecal trigger tower with $E_T > 8$ GeV.
   HLT: electron object with $E_T > 22$ GeV matched to the L1 trigger tower and passing identification and isolation criteria.

The table includes the run range, total integrated luminosity, HLT trigger name, L1 trigger name, and trigger rate from a high luminosity run characteristic of the run range for each trigger used. The average efficiency of the triggers relative to the offline selection is approximately 98%, as will be demonstrated in Chapter 9.

Table 6.1: The single electron triggers during 2010 data taking.

Period	Run Range	$\int L\mathrm{dt}$ (pb^{-1})	HLT Name	L1 Name	HLT Rate (Hz)
1	132440-137028	0.005	Photon10_L1R	SingleEG5	8.29
2	138564-140401	0.26	Photon15_Cleaned_L1R	SingleEG5	5.32
3	141956-144114	2.9	Ele15_SW_CaloEleId_L1R	SingleEG5	6.28
4	144115-147145	5.1	Ele17_SW_CaloEleId_L1R	SingleEG8	20.81
5	147146-148058	9.4	Ele17_SW_TightEleId_L1R	SingleEG8	31.09
6	148103-149065	10.2	Ele22_SW_TighterCaloIdIsol_L1R_v1	SingleEG8	13.06
7	149180-149442	8.0	Ele22_SW_TighterCaloIdIsol_L1R_v2	SingleEG8	11.73

6.1.3 Offline Electron Selection

The two isolated high P_T electrons characterizing the signal are selected via cuts on a set of selection variables that discriminate electrons originating in Z decays from background events. At the LHC, the dominant background is QCD multijet events, which have a very different detector signature than the signal. For instance, QCD multijet events generally are not well isolated and have broader shower profiles than $Z \rightarrow ee$ decays. The electron selection variables used at CMS are described below.

Isolation Variables

There are three isolation variables: $TrackIso$, $EcalIso$, and $HcalIso$, which describe how isolated in the detector the transverse momenta or energy deposits associated with the electron candidate are. $TrackIso$ is the sum of the transverse momenta in the tracker falling within a cone of aperture $\Delta R < 0.3$ about the direction of the electron candidate. The track associated with the electron is excluded from the sum. $EcalIso$ ($HcalIso$) is the sum of the transverse energy in the Ecal (Hcal) falling within a cone of aperture

$\Delta R < 0.3$ about the electron candidate momentum direction. The energy deposits associated with the electron candidate are excluded from the $EcalIso$ sum. The energy deposit in the Hcal immediately behind ($\Delta R < 0.15$) the Ecal energy deposits of the electron candidate is excluded from the $HcalIso$ sum. These variables are then divided by the P_T of the electron candidate to form relative isolation variables. Detailed definitions of the relative isolation variables are provided in Table 6.2.

Table 6.2: Definitions of the relative electron isolation variables.

$TrackIso/P_T$	Isolation variable defined as the $\sum P_T$ of all tracks within an annular cone of $\Delta R \in (0.015, 0.3)$ about the electron track satisfying $P_T > 0.7$ GeV and distance from the vertex < 0.2 cm, and then divided by the electron P_T. A slice in η centered at the electron track and having a width of 0.015 is excluded from the $\sum P_T$.
$EcalIso/P_T$	Isolation variable defined as the $\sum E_T$ of all Ecal crystal energy deposits within an annular cone of $\Delta R \in (X, 0.3)$ about the electron supercluster satisfying $E > 0.1$ ($E_T > 0.08$) GeV in the Ecal Barrel (Endcap), and then divided by the electron P_T. Here X corresponds to the width of three Ecal crystals. A slice in η centered at the electron seed crystal and having a width of 1.5 crystals is excluded from the $\sum E_T$.
$HcalIso/P_T$	Isolation variable defined as the $\sum E_T$ of all Hcal tower energy deposits within an annular cone of $\Delta R \in (0.15, 0.3)$ about the electron supercluster and then divided by the electron P_T.

Identification Variables

There are four identification variables, which deal with the shower profiles of the electron candidates. The first is $\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$, which is the energy log-weighted shower width of the electron supercluster in the η direction. The next is $\Delta\phi_{\mathrm{in}}$ ($\Delta\eta_{\mathrm{in}}$), which is the absolute difference in the ϕ (η) direction between the supercluster and the associated track as extrapolated to its vertex. The $\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$, $\Delta\phi_{\mathrm{in}}$, and $\Delta\eta_{\mathrm{in}}$ variables are related to the transverse shower profiles of the electron candidates, while the next variable describes the longitudinal shower profiles. It is called H/E and is the ratio of hadronic energy deposits in the Hcal to electromagnetic energy deposits in the Ecal in a cone of aperture $\Delta R < 0.15$ about the

electron supercluster. The cone used for H/E has been defined such that it corresponds to the hollow region excluded in the transverse energy sum of the $HcalIso$ variable. Detailed definitions of the identification variables are provided in Table 6.3.

Table 6.3: Definitions of the electron identification variables.

$\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$	The energy log-weighted Ecal shower width of the electron calculated from the 5×5 crystal array about the supercluster seed crystal.
$\Delta\phi_{\mathrm{in}}$	$\Delta\phi_{\mathrm{in}} = \lvert\phi_{\mathrm{SC}} - \phi_{\mathrm{track@vertex}}\rvert$, where ϕ_{SC} is the energy weighted ϕ direction of the supercluster and $\phi_{\mathrm{track@vertex}}$ is the ϕ direction of the electron track as extrapolated to its vertex.
$\Delta\eta_{\mathrm{in}}$	$\Delta\eta_{\mathrm{in}} = \lvert\eta_{\mathrm{SC}} - \eta_{\mathrm{track@vertex}}\rvert$, where η_{SC} is the energy weighted η direction of the supercluster and $\eta_{\mathrm{track@vertex}}$ is the η direction of the electron track as extrapolated to its vertex.
H/E	The ratio of hadronic energy deposits in the Hcal to electromagnetic energy deposits in the Ecal in a cone of $\Delta R < 0.15$ about the electron supercluster.

Photon Conversion Rejection Variables

An electron candidate is assumed to come from a converted photon if both the distance—$Dist$—in the r-ϕ plane between it and its nearest partner track is less than 0.02 and the $\Delta\cot\theta$ between these two objects is also less than 0.02. That is, if an electron satisfies both of these criteria, it is assumed to come from a converted photon and is vetoed. Partner tracks are searched for within a cone of aperture $\Delta R < 0.5$ about the electron track and must have charges opposite that of the electron candidate. In the absence of any partner track, the $Dist$ and $\Delta\cot\theta$ variables are initialized to 1, so that the photon conversion veto is not applied. If the electron is missing too many tracker hits in front of its first valid hit in the inner tracker layers, then it is also vetoed as a converted photon. This variable—the number of missing tracker hits in front of the electron's first valid hit—is referred to simply as $MissingHits$. Detailed definitions of the conversion rejection variables are provided in Table 6.4.

Table 6.4: Definitions of the photon conversion rejection variables.

Variable	Description
$Dist$	Distance in the r-ϕ plane between the electron and its nearest partner track.
$\Delta \cot \theta$	Absolute difference in $\cot \theta$ between the electron and its nearest partner track.
$MissingHits$	Number of missing tracker hits in front of the electron's first valid hit in the inner tracker.

Selection Criteria Optimization

The selection criteria are optimized by tuning the cut thresholds of a discrete set of selection variables in such a way that maximal signal efficiency is achieved for a given, fixed background rejection. The tune is carried out iteratively to provide different performances, with each step defining a working point. These points lie on a curve such that as the cuts become tighter the background rejection increases and the signal efficiency falls. In this way a physics analysis involving electrons in the final state need only adopt the appropriate working point on the curve rather than develop its own analysis-specific electron selection. The curve corresponding to the series of electron selection points developed at CMS is drawn in black in Figure 6.2. The baseline selection points were extracted using MC before data taking commenced. They were later observed to still be near-optimal with data. Indeed, although this is a MC-tuned cut-based selection, other selection schemes, including cuts in categories [74] and multi-variate methods such as boosted decision trees and likelihoods, have been observed on data to yield only marginal improvements over this baseline. Hence the baseline was maintained unchanged. The curves corresponding to these alternative schemes are also drawn in Figure 6.2, with the colors indicated in the legend.

Working Points for the $Z \to ee$ Cross Section Measurement

The working point for the $Z \to ee$ cross section measurement is the one in which reconstructed electrons with $E_T > 20$ GeV are about 80% efficient in the MC.

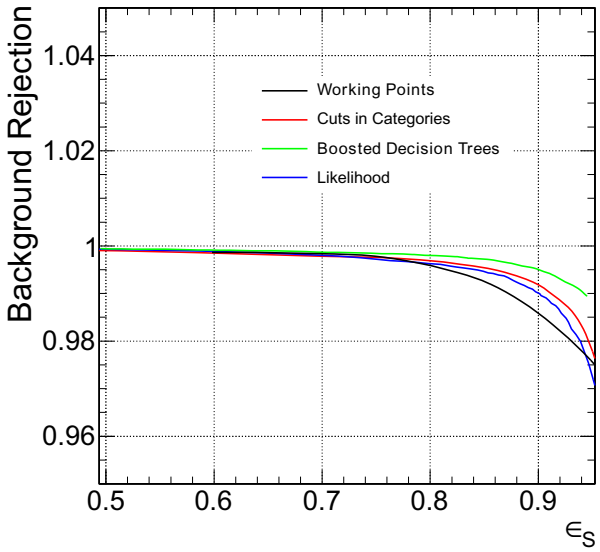


Figure 6.2: The electron selection criteria working point curves. The electron efficiency is plotted on the horizontal axis, while the background rejection (defined as the difference between unity and the background efficiency) is plotted on the vertical axis for each curve. The curve corresponding to the default CMS working point optimization is drawn in black. The curves corresponding to a few alternative optimization schemes are also drawn: cuts in categories (green), boosted decision trees (red), and likelihood (blue).

Accordingly, this working point is referred to as WP80. It was selected for this analysis since it is the loosest point (maximizing the signal yield) that still admits a negligible QCD background, implying that the corresponding systematic uncertainty is minimal. The cut thresholds of WP80 are provided in Table 6.5.

The looser WP90 and tighter WP70 selection points will be used during the measurement for various systematic studies. The looser WP95 and WP85 selection points will be used as well in an example of the background subtraction methodology. The cut thresholds for WP95, WP90, WP85, and WP70 are provided in Tables 6.6, 6.7, 6.8, and 6.9, respectively.

Table 6.5: The working point 80% selection criteria.

Variables	Selection in Barrel	Selection in Endcaps
$MissingHits$	< 1	< 1
$Dist$	> 0.02	> 0.02
or		
$\Delta\cot\theta$	> 0.02	> 0.02
$TrackIso/P_T$	< 0.09	< 0.04
$EcalIso/P_T$	< 0.07	< 0.05
$HcalIso/P_T$	< 0.10	< 0.025
$\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$	< 0.01	< 0.03
$\Delta\phi_{\mathrm{in}}$	< 0.06	< 0.03
$\Delta\eta_{\mathrm{in}}$	< 0.004	< 0.007
H/E	< 0.04	< 0.025

6.1.4 $Z\to ee$ Candidate Arbitration

As the selection criteria are tight, it is rare ($\ll 1\%$) to find multiple Z candidates in an event. However, when this happens the following arbitration is made: If there are two possible Z candidates, then one is chosen at random. If there are more than two possible Z candidates, then the event is rejected to avoid biases associated with the arbitration.

6.1.5 Selected $Z\to ee$ Signal Yield

After applying the selection criteria as described above, including the cut on the invariant mass window, application of the WP80 criteria on both electrons, and the fiducial and kinematic acceptance cuts, the final selected yield is $N = 8441$ $Z\to ee$ candidates for the 35.9 pb^{-1} dataset.[2] [3]

6.2 $Z\to ee$ Performance Plots

The criteria used to select the $Z\to ee$ decays is evaluated in Figures 6.3-6.19. The MC samples are stacked and normalized to the integrated luminosity of the data. The data is

[2] Note there is no opposite-sign charge requirement on the electron pairs.

[3] A small ad-hoc, post-reconstruction correction to the energy scale that was determined using $Z\to ee$ decays is applied in order to select the final $Z\to ee$ events within the defined mass window and acceptance. Without this correction, 8454 $Z\to ee$ events would be selected.

Table 6.6: The working point 95% selection criteria.

Variables	Selection in Barrel	Selection in Endcaps
$MissingHits$	< 2	< 2
$Dist$	> 0.02	> 0.02
or		
$\Delta\cot\theta$	> 0.02	> 0.02
$TrackIso/P_T$	< 0.15	< 0.08
$EcalIso/P_T$	< 2	< 0.06
$HcalIso/P_T$	< 0.12	< 0.05
$\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$	< 0.01	< 0.03
$\Delta\phi_{\mathrm{in}}$	< 0.8	< 0.7
$\Delta\eta_{\mathrm{in}}$	< 0.007	< 0.01
H/E	< 0.15	< 0.07

plotted in black, signal MC in yellow, the various electroweak MC samples in orange, and the $t\bar{t}$ MC sample in red. The QCD MC samples are not included, as their contribution is negligible.

6.2.1 Z Boson Kinematic Distributions

Figures 6.3-6.9 show the kinematic distributions of the selected Z bosons, including the invariant mass M_{ee}, P_T, η, ϕ, rapidity y, and $\cos\theta$. The invariant mass plot on the log scale includes an extended mass window outside of that selected. The selected window of $M_{ee} \in (60, 120)$ GeV is marked by vertical green lines in this plot. The data generally agrees well with the MC, except that the data has a broader mass and hence energy resolution than the MC.

Table 6.7: The working point 90% selection criteria.

Variables	Selection in Barrel	Selection in Endcaps
$MissingHits$	< 2	< 2
$Dist$	> 0.02	> 0.02
or		
$\Delta\cot\theta$	> 0.02	> 0.02
$TrackIso/P_T$	< 0.12	< 0.05
$EcalIso/P_T$	< 0.09	< 0.06
$HcalIso/P_T$	< 0.1	< 0.03
$\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$	< 0.01	< 0.03
$\Delta\phi_{\mathrm{in}}$	< 0.8	< 0.7
$\Delta\eta_{\mathrm{in}}$	< 0.007	< 0.009
H/E	< 0.12	< 0.05

Table 6.8: The working point 85% selection criteria.

Variables	Selection in Barrel	Selection in Endcaps
$MissingHits$	< 2	< 2
$Dist$	> 0.02	> 0.02
or		
$\Delta\cot\theta$	> 0.02	> 0.02
$TrackIso/P_T$	< 0.09	< 0.05
$EcalIso/P_T$	< 0.08	< 0.05
$HcalIso/P_T$	< 0.1	< 0.025
$\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$	< 0.01	< 0.03
$\Delta\phi_{\mathrm{in}}$	< 0.06	< 0.04
$\Delta\eta_{\mathrm{in}}$	< 0.006	< 0.007
H/E	< 0.04	< 0.025

Table 6.9: The working point 70% selection criteria.

Variables	Selection in Barrel	Selection in Endcaps
$MissingHits$	< 1	< 1
$Dist$	> 0.02	> 0.02
or		
$\Delta\cot\theta$	> 0.02	> 0.02
$TrackIso/P_T$	< 0.09	< 0.04
$EcalIso/P_T$	< 0.07	< 0.05
$HcalIso/P_T$	< 0.10	< 0.025
$\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$	< 0.01	< 0.03
$\Delta\phi_{\mathrm{in}}$	< 0.06	< 0.03
$\Delta\eta_{\mathrm{in}}$	< 0.004	< 0.007
H/E	< 0.04	< 0.025

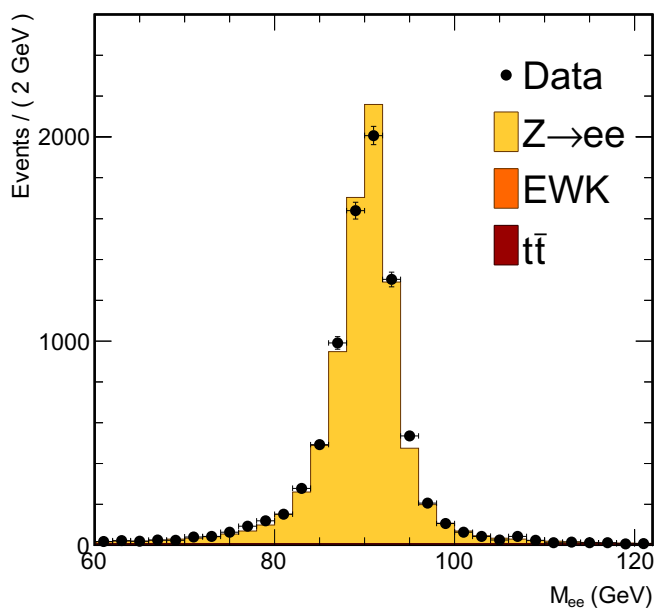


Figure 6.3: The invariant mass distribution of the $Z \to ee$ candidates on a linear scale.

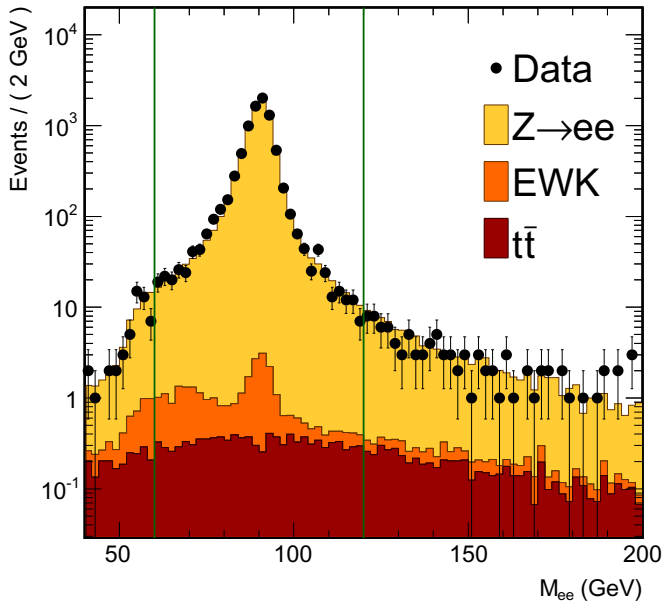


Figure 6.4: The invariant mass distribution of the $Z \rightarrow ee$ candidates on a log scale with an extended window.

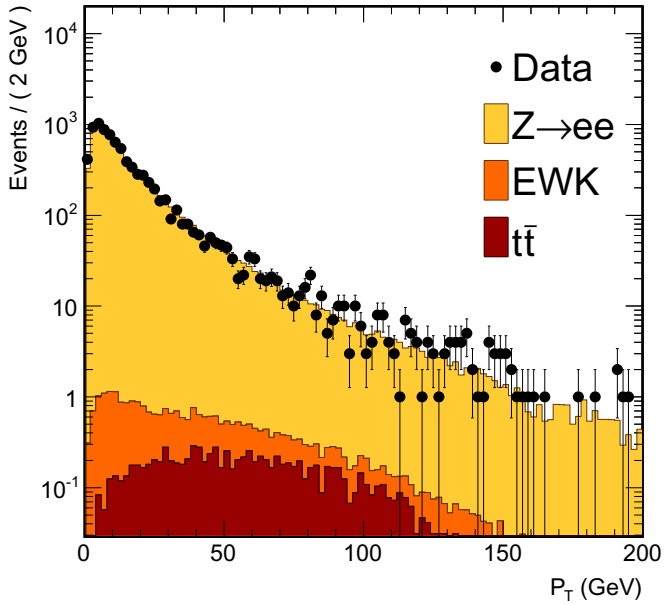


Figure 6.5: The P_T distribution of the $Z \rightarrow ee$ candidates.

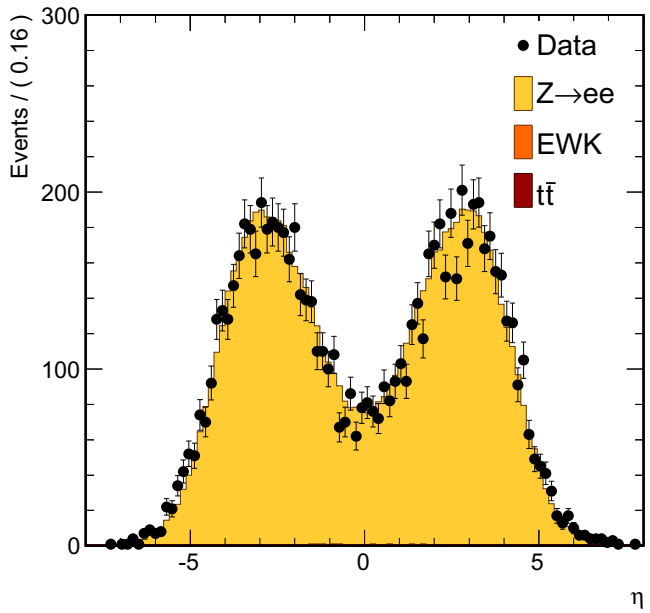


Figure 6.6: The η distribution of the $Z \to ee$ candidates.

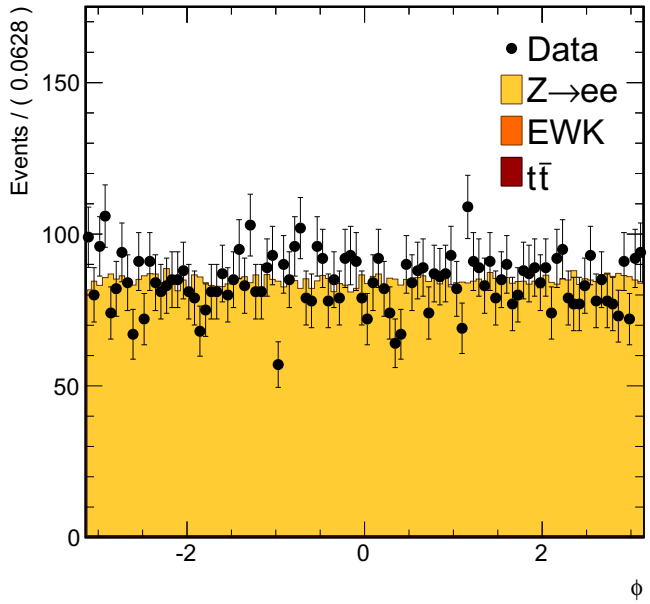


Figure 6.7: The ϕ distribution of the $Z \to ee$ candidates.

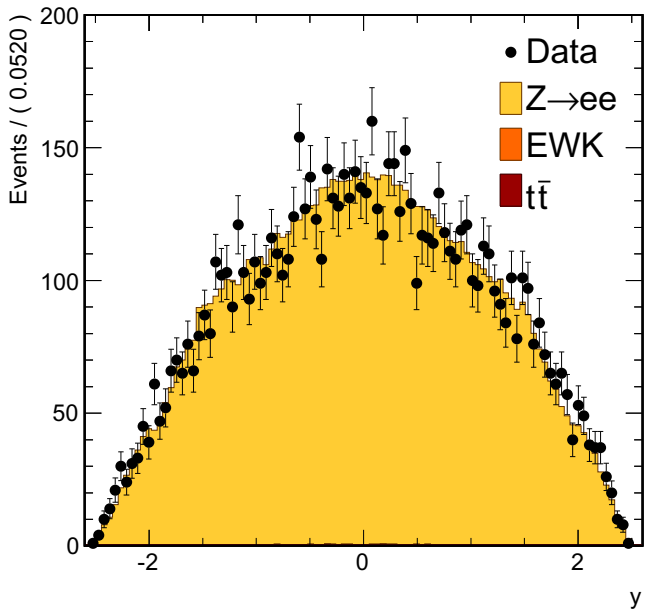


Figure 6.8: The y distribution of the $Z \to ee$ candidates.

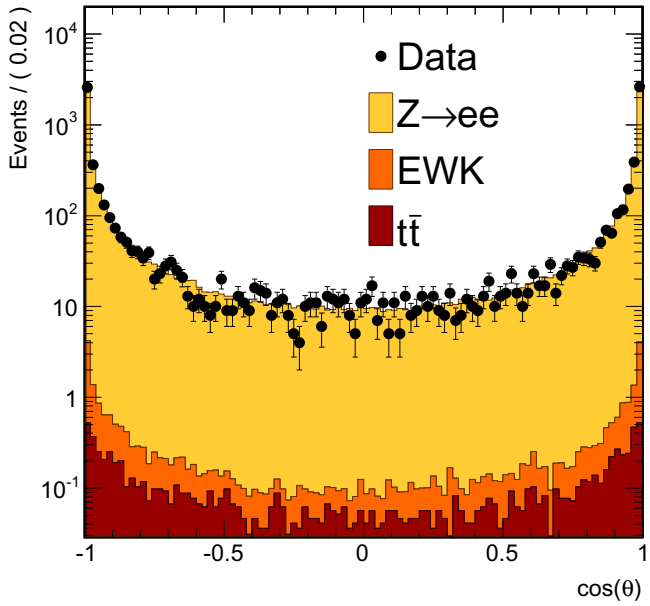


Figure 6.9: The $\cos\theta$ distribution of the $Z \to ee$ candidates.

6.2.2 Associated Electrons Kinematic Distributions

Figures 6.10-6.12 show the kinematic distributions of the electrons associated with the Z bosons. These are "N-1" style plots, in which the event must pass all selection criteria excluding the cut under study for the plotted electron candidate The plotted distributions include P_T, η, and ϕ. The data agrees well with the MC.

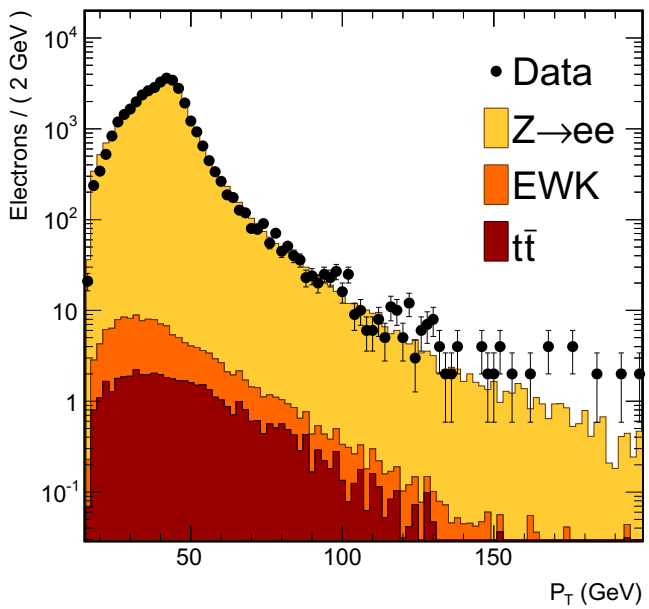


Figure 6.10: The P_T distribution of the associated electrons.

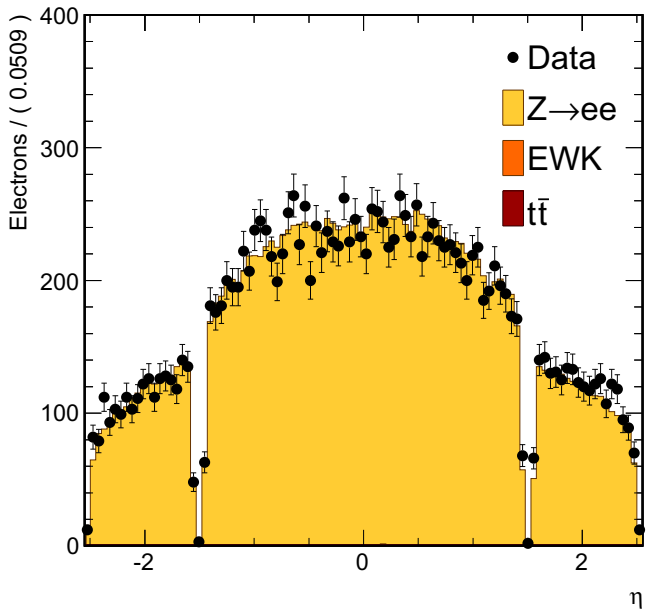


Figure 6.11: The η distribution of the associated electrons.

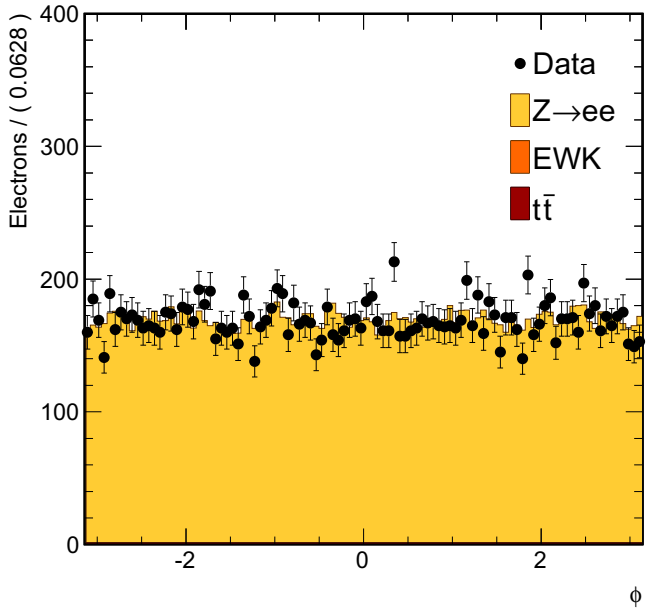
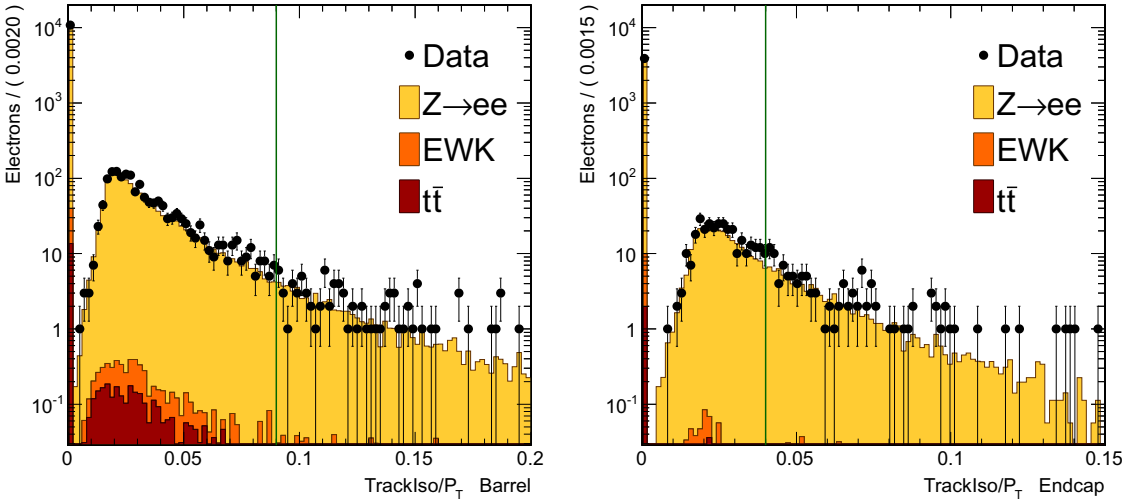


Figure 6.12: The ϕ distribution of the associated electrons.

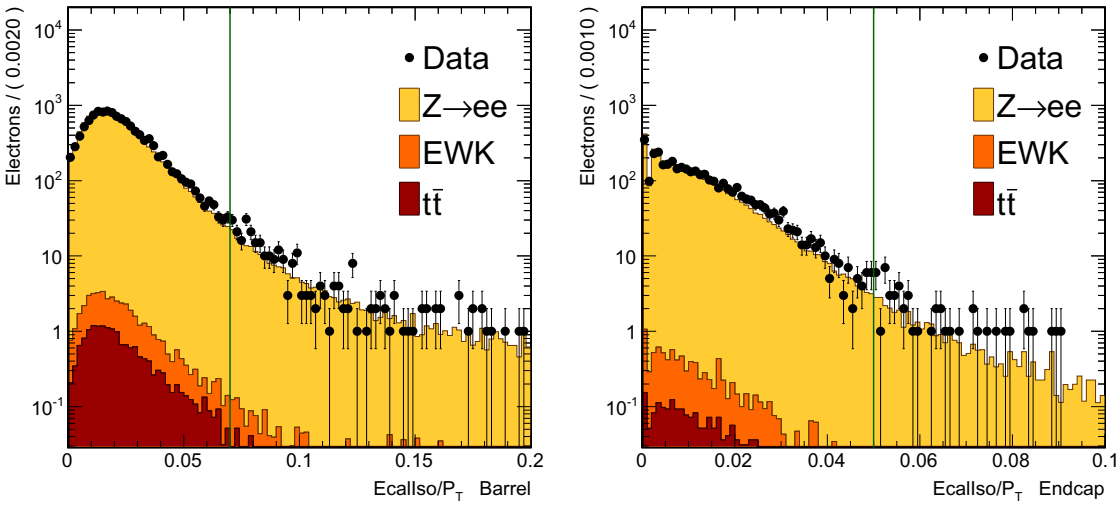
6.2.3 Associated Electrons Selection Variables Distributions

In Figures 6.13-6.19 one finds the distributions of the isolation and identification variables for the electrons associated with the Z bosons. These are also "N-1" style plots, as described in Subsection 6.2.2 immediately above. The plotted variables include $TrackIso/p_T$, $HcalIso/p_T$, $EcalIso/p_T$, $\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$, $\Delta\phi_{\mathrm{in}}$, $\Delta\eta_{\mathrm{in}}$, and H/E. All cut thresholds are marked by vertical green lines. The data generally agrees well with the MC, except for in the tails of a few of the distributions.



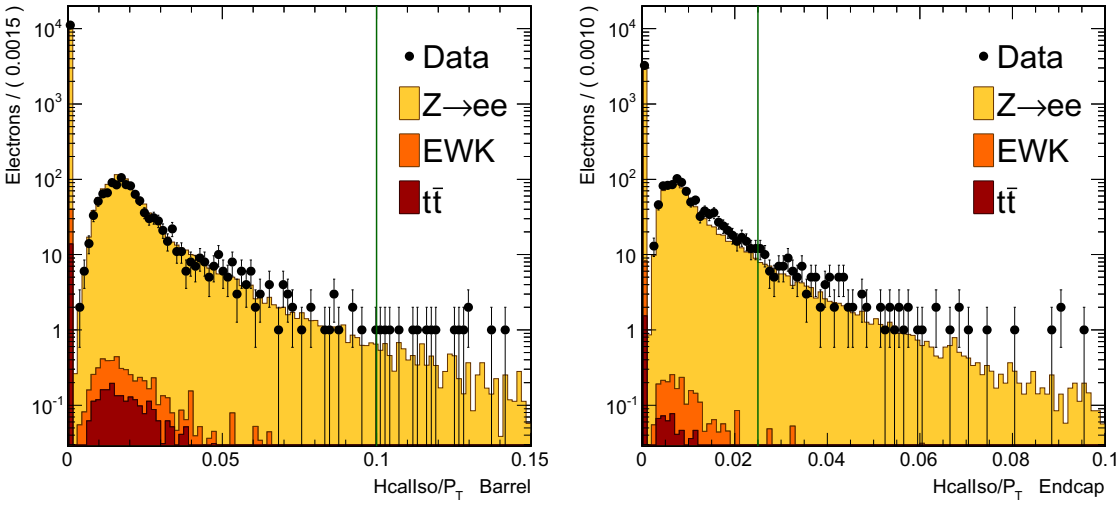
(a) (b)

Figure 6.13: The $TrackIso/P_T$ distribution of the associated electrons: (a) Ecal Barrel, (b) Ecal Endcap.



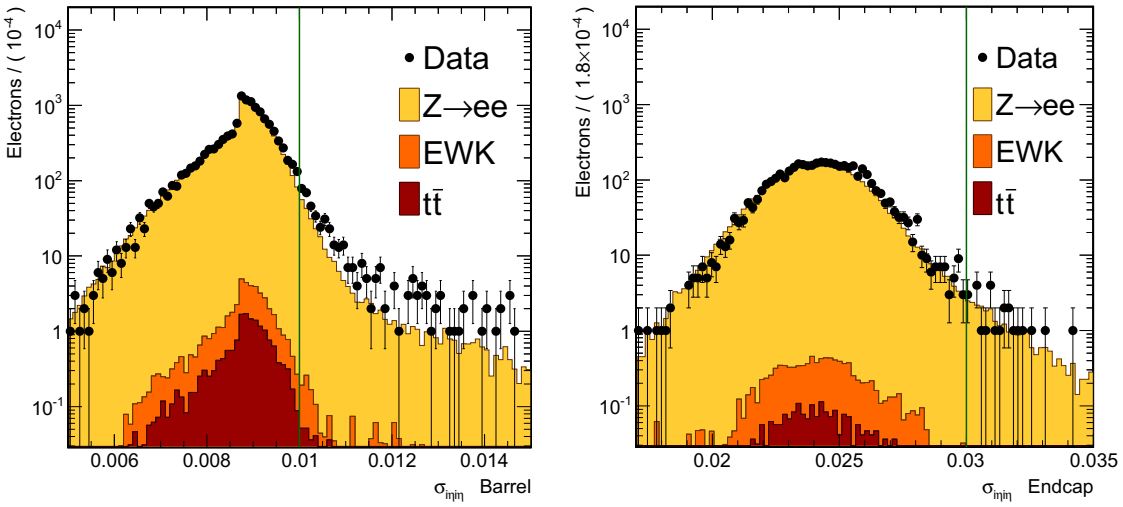
(a) (b)

Figure 6.14: The $EcalIso/P_T$ distribution of the associated electrons: (a) Ecal Barrel, (b) Ecal Endcap.



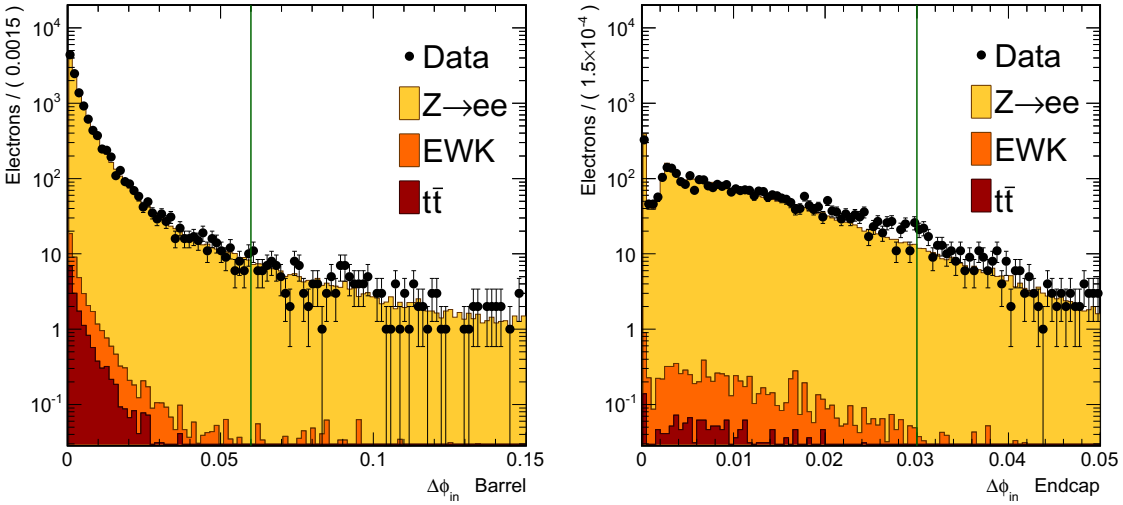
(a) (b)

Figure 6.15: The $HcalIso/P_T$ distribution of the associated electrons: (a) Ecal Barrel, (b) Ecal Endcap.



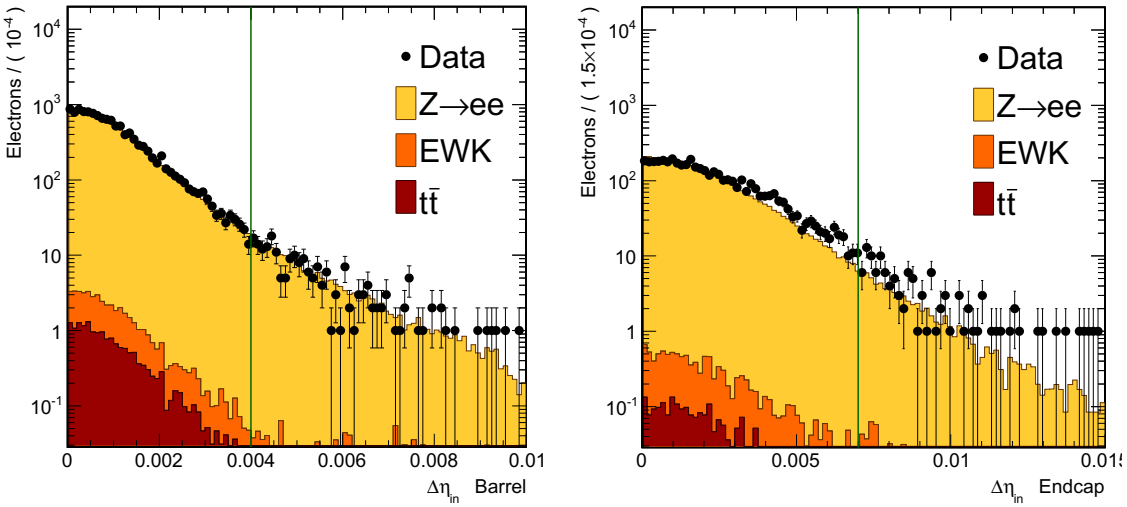
(a) (b)

Figure 6.16: The $\sigma_{i\eta i\eta}$ distribution of the associated electrons: (a) Ecal Barrel, (b) Ecal Endcap.



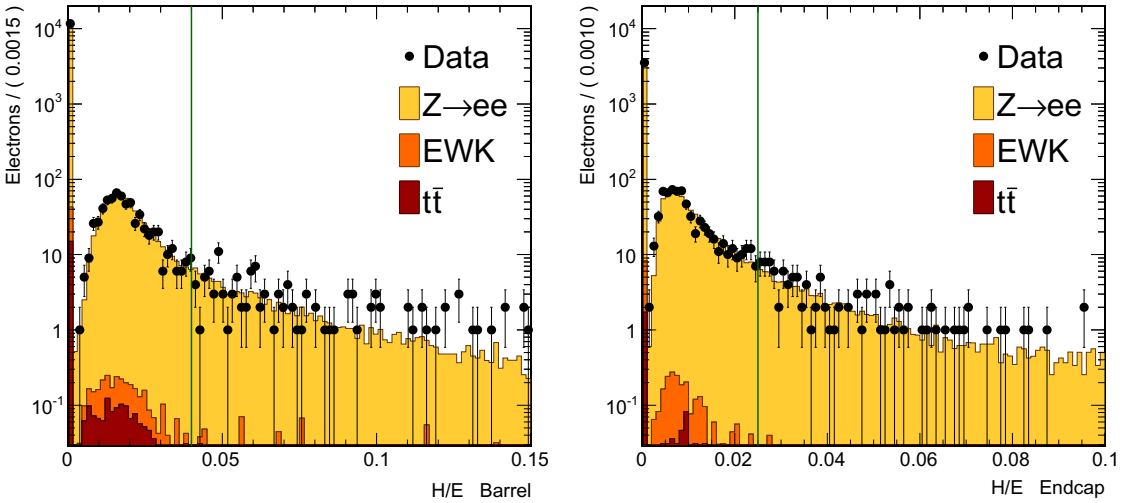
(a) (b)

Figure 6.17: The $\Delta\phi_{\rm in}$ distribution of the associated electrons: (a) Ecal Barrel, (b) Ecal Endcap.



(a) (b)

Figure 6.18: The $\Delta\eta_{\text{in}}$ distribution of the associated electrons: (a) Ecal Barrel, (b) Ecal Endcap.



(a) (b)

Figure 6.19: The H/E distribution of the associated electrons: (a) Ecal Barrel, (b) Ecal Endcap.

6.3 Summary

The $Z \rightarrow ee$ event selection has been described and a suite of performance plots has been presented. The online trigger requirements have been discussed in addition to the offline selection variables and criteria. The selection used is cut-based and has been tuned on the MC. Nevertheless, the selection criteria are near optimal on data and perform similarly to other possible selection schemes. A total of 8441 $Z \rightarrow ee$ candidates have been selected on the full 2010 CMS dataset under study. The data generally agrees well with the MC, providing a strong confirmation of the detector simulation.

Chapter 7

Data Driven Background Estimations

7.1 Introduction

The number of background events remaining beneath the signal needs to be estimated after application of the event selection. The dominant source of background is QCD multijet events where the jets fake electrons coming from a Z decay, electroweak processes such as $Z \to \tau\tau$, and $t\bar{t}$. The jet backgrounds (i.e., QCD multijet and W+jets events) are estimated using data-driven methods to avoid dependencies on the MC that initially need to be checked in the early data; i.e., the MC must first be verified to provide the correct detector response for these events. Generally, the remaining non-jet background contributions may be taken directly from the MC.

Two independent methods are considered to estimate the backgrounds in data. The primary method is sensitive only to the jet backgrounds and uses a variable having very different shapes in signal and background so that it can be used to discriminate between the two. The second method is used as a cross-check and is based on the electric charge signs of the two electron candidates associated with the selected Z boson. It is primarily sensitive to charge product symmetric backgrounds in which half of the events have

same-sign electron pairs and half have opposite-sign, such as QCD multijet events, but lacks information to disentangle other backgrounds (such as $t\bar{t}$, which are dominantly opposite-sign) from these. Instead, a systematic uncertainty is assigned to give coverage for all backgrounds.

Table 7.1 provides the yields of the signal and backgrounds after application of the standard WP80 event selection in data and MC. The raw number of entries is also provided for each sample. The MC predicts no QCD background contribution and a negligible contribution from W+jets events. The electroweak and $t\bar{t}$ background is expected to amount to 0.4% of the signal. To demonstrate the methods in a MC study, the selection is loosened so that the jet backgrounds dominate. In particular, the WP95 selection criteria (Table 6.6) with a lower supercluster E_T threshold of 20 GeV is employed. This selection is denoted as the "reference" one, so as not to confuse it with the standard selection that requires the superclusters associated with both electrons to have an E_T >25 GeV and pass the WP80 selection point.

Table 7.1: The contributions of the various samples after the standard WP80 selection for 35.9 pb^{-1}.

Channel	Yield	Raw Entries
Data $Z \to ee$	8441 ± 92	8441
MC $Z \to ee$	8284 ± 91	294732
QCD	0	0
$t\bar{t}$	10 ± 3	1910
$W + jet$	0 ± 1	15
$Z \to \tau\tau$	7 ± 3	232
WW	3 ± 2	3286
WZ	6 ± 3	8907
ZZ	4 ± 2	36249
All Backgrounds	30 ± 5	50599

7.2 Template Fitting Method

The template method is used to estimate the background under the Z peak by exploiting a discriminating variable that has very different shapes in signal and background. Background and signal templates of this variable are extracted from control samples. These templates are then used as probability distribution functions (PDFs) to fit the data and extract the signal fraction. There are a few discriminating variables that could be used with this method, such as relative track isolation ($TrackIso/p_T$), relative calorimeter isolation, or $\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$. In the following, $TrackIso/p_T$ is used to estimate the background contribution beneath the signal.

Data Template

This template is made from the $TrackIso/p_T$ distribution composed of both electrons in the events passing the selection denoted SemiTight+SemiTight, which is identical to the reference selection (Section 7.1) except that there is no cut on $TrackIso/p_T$.

Signal Template

The signal template is extracted from a control sample selected using the SemiTight+SemiTight selection, but with an added opposite-sign (OS) requirement for the associated tracks, and also tighter cuts on the weakly correlated variables H/E and $\Delta\eta_{\mathrm{in}}$. The H/E threshold is lowered to 0.075 (0.035) in the Ecal Barrel (Endcap), and the $\Delta\eta_{\mathrm{in}}$ threshold is lowered to 0.0035 (0.005) in the Ecal Barrel (Endcap). These criteria define the OS-Tight+Tight selection.

The efficiency of the OS-Tight+Tight selection relative to the SemiTight+SemiTight selection is approximately 83.8% and 14.9% for the signal and jet background events, respectively. The $TrackIso/p_T$ templates for the SemiTight+SemiTight and OS-Tight+Tight selections obtained from the $Z\rightarrow ee$ MC in the signal invariant mass region of $M_{ee} \in (60, 120)$ GeV are shown in Figure 7.1. There is no tangible bias due to the tightening of the selection criteria.

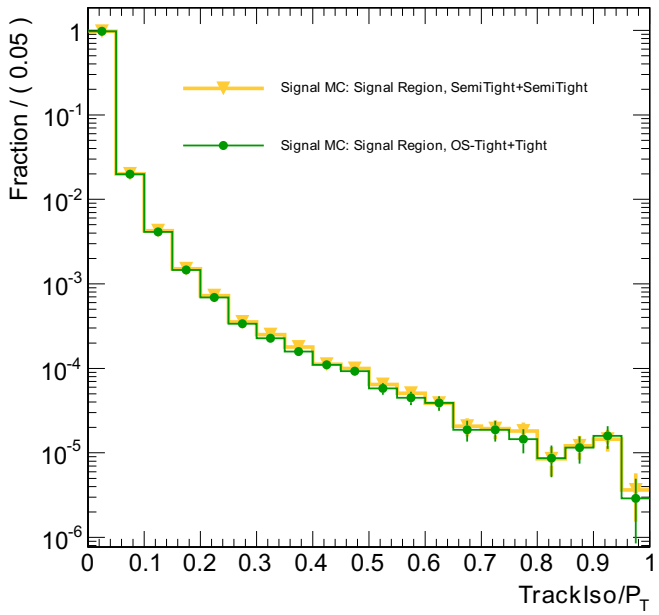


Figure 7.1: Relative track isolation of the $Z \rightarrow ee$ MC for the SemiTight+SemiTight (yellow) and OS-SemiTight+SemiTight (green) selections in the signal invariant mass region $M_{ee} \in$ (60, 120) GeV.

Background Template

The jet background template is derived from a control sample in the sideband invariant mass region $M_{ee} \in (40, 60)$ GeV. An additional same-sign (SS) requirement is employed to reduce the fraction of $Z/\gamma^{\star}$ events. The Drell-Yan contamination is reduced substantially after application of the SS requirement, as shown in Figure 7.2. Specifically, about 86 background and 5 $Z/\gamma^{\star}$ events are expected from MC for 35.9 pb^{-1} in the sideband mass region of $M_{ee} \in (40, 60)$ GeV.

To examine whether the background $TrackIso/p_T$ template is independent of the invariant mass region, the shape extracted from the sideband region is compared to that from the signal region $M_{ee} \in (60, 120)$ GeV using the jet background MC samples. Figure 7.3 shows that the template from the sideband region, where the Drell-Yan contamination is small, agrees with that from the signal region. That is, the $TrackIso/p_T$ templates are the same in both invariant mass regions and there is not a substantial bias.

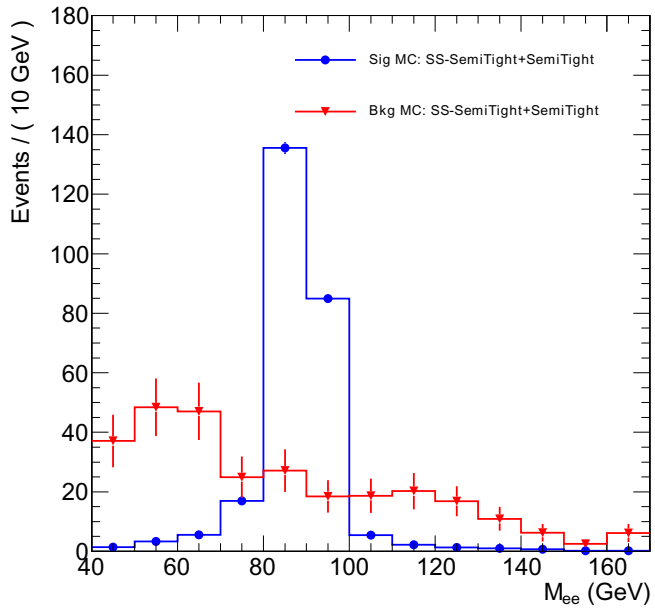


Figure 7.2: Background (red) and Drell-Yan (blue) MC invariant mass distributions scaled to 35.9 pb^{-1} for the SS-SemiTight+SemiTight selection criteria.

In order to have a larger number of background events (and also further reduce the relative Drell-Yan contamination), the cuts on one of the daughter electrons may be loosened. Figure 7.3 also compares the $TrackIso/p_T$ shapes between the SemiTight+SemiTight selection and this SS-SemiTight+Loose selection. Here "Loose" corresponds to the set of cuts where the thresholds of the SemiTight selection are loosened by a factor of 10. The result is that the $TrackIso/p_T$ shape is stable with respect to loosening one of the legs. In loosening one of the legs, however, sensitivity to the non-jet backgrounds is lost. Thus these other backgrounds must be added separately from the MC estimates. The expected number of background events passing the SS-SemiTight+Loose selection for 35.9 pb^{-1} in the region $M_{ee} \in (40, 60)$ GeV is about 1093 with an admixture of about 13 Drell-Yan events. The impact of the Drell-Yan admixture on the background estimation is studied in the next subsection.

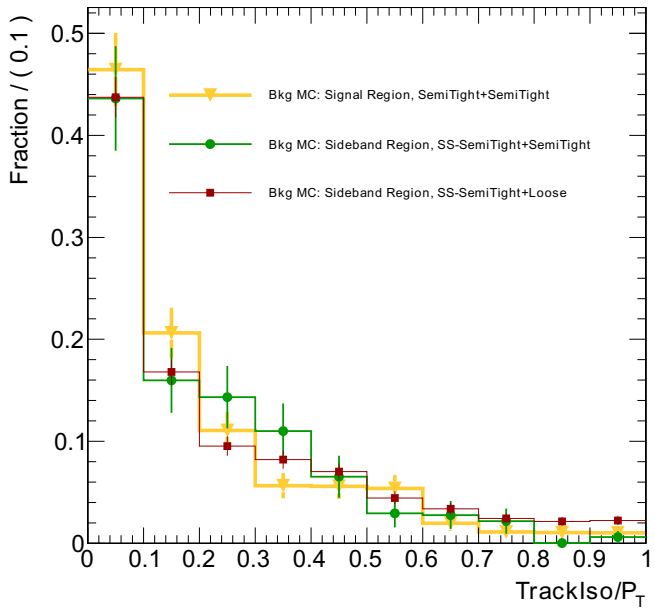


Figure 7.3: Relative track isolation of the background MC for the SemiTight+SemiTight selection in the signal region $M_{ee} \in (60, 120)$ GeV (yellow), the SS-SemiTight+SemiTight selection in the sideband region $M_{ee} \in 40, 60)$ GeV (green) and the SS-SemiTight+Loose selection in the sideband region $M_{ee} \in (40, 60)$ GeV (red).

7.2.1 Closure Test of the Template Method

In this subsection MC signal and jet background samples are mixed to evaluate the accuracy of the template method. Quasi-data MC samples are prepared by mixing the $Z \rightarrow ee$ signal and background events with fractional contributions f_{S} and $1 - f_{\mathrm{S}}$, respectively. The $TrackIso/p_T$ signal and background templates are then fit to the quasi-data using the fractional fitter TFractionFitter [75]. In this way, estimates of the initial signal and background fractions are obtained.

For this exercise, a uniform cut on the $TrackIso/p_T$ variable with a threshold of $T = 0.1$ is used in both the Ecal Barrel and Endcap for the final, selected sample. The fit is still performed over the full, populated $TrackIso/p_T$ range (0,1), however. The result of the fit over the entire range is then extrapolated back to obtain the signal fraction

beneath the threshold T, which is given by:

$$f_S(x|T) = \frac{f_S{}' \times \int_0^T S(x)\,dx}{f_S{}' \times \int_0^T S(x)\,dx + (1 - f_S{}') \times \int_0^T B(x)\,dx} \tag{7.1}$$

Here the signal and background templates are denoted S and B, respectively, and it is assumed that their integrals over the entire region $TrackIso/p_T \in (0, 1)$ have been normalized to unity. $f_S{}'$ is defined as the signal fraction over the entire range.

The results are shown in Table 7.2, where the signal fraction f_S has been varied from 0.95–0.999. The line "Bkg + N% DY " shows results of the fit with errors when N% Drell-Yan contamination is added to the background template (i.e., N% of the template is the Drell-Yan shape and the rest is the background shape). The Drell-Yan contamination is expected to be about 1% from the MC. However, in Table 7.2 the Drell-Yan fraction is varied up to 10%. It is apparent from the table that the predicted signal fractions are in agreement with the true values in all cases, even when much larger than expected Drell-Yan contamination is admitted. The maximum bias observed due to the Drell-Yan contamination is 0.01%, and this value is assigned as a systematic uncertainty. The expected combined statistical and systematic uncertainty returned by the fit for 35.9 pb^{-1} is about 0.1–0.2% in the high signal purity limit $f_S > 0.99$ expected for this analysis. This uncertainty is dominated by the relatively low number of background events present in the SS-SemiTight+Loose sideband control sample.

Table 7.2: Closure test on the signal fractions for 35.9 pb^{-1}.

Background Sample	Signal Fraction f_S			
	0.95	0.99	0.995	0.999
Bkg + 1% DY	0.9552 ± 0.0078	0.9920 ± 0.0025	0.9962 ± 0.0018	0.9991 ± 0.0010
Bkg + 2% DY	0.9552 ± 0.0079	0.9920 ± 0.0025	0.9962 ± 0.0018	0.9992 ± 0.0010
Bkg + 4% DY	0.9552 ± 0.0079	0.9920 ± 0.0025	0.9962 ± 0.0018	0.9991 ± 0.0011
Bkg + 10% DY	0.9551 ± 0.0078	0.9919 ± 0.0026	0.9962 ± 0.0018	0.9992 ± 0.0010

Although there does not appear to be any bias due to the extrapolation from the SemiTight+SemiTight to the OS-Tight+Tight control sample to obtain the signal template, this statement requires quantification and any bias needs to have a systematic

uncertainty assigned. To evaluate this, the closure test is repeated with 1% Drell-Yan contamination and the MC signal template extracted from the OS-Tight+Tight sample is varied to that extracted from the SemiTight+SemiTight sample. The results with both signal templates are compared in Table 7.3. The table indicates that the results agree within the quoted errors, confirming the original claim that there is no bias due to the extrapolation to the signal control sample. Hence no systematic uncertainty is assigned for the signal template extrapolation.

Table 7.3: Closure test using different signal templates.

Signal Template	Signal Fraction f_S			
	0.95	0.99	0.995	0.999
SemiTight+SemiTight	0.9552 ± 0.0078	0.9920 ± 0.0025	0.9962 ± 0.0018	0.9991 ± 0.0010
OS-Tight+Tight	0.9555 ± 0.0078	0.9921 ± 0.0025	0.9963 ± 0.0018	0.9993 ± 0.0009

Similarly, in Table 7.4 the background MC template is varied from the default SS-SemiTight+Loose sideband control sample to the alternative SemiTight+SemiTight sample in the signal invariant mass region. From the table it is seen that the results from the two templates agree within the quoted errors, proving the claim that there is no bias due to the extrapolation of the background template to the sideband region with the SS-SemiTight+Loose selection. Accordingly, no systematic uncertainty is assigned for the background template extrapolation.

Table 7.4: Closure test using different background templates.

Background Template	Signal Fraction f_S			
	0.95	0.99	0.995	0.999
Sideband Region, SS-SemiTight+Loose	0.9552 ± 0.0078	0.9920 ± 0.0025	0.9962 ± 0.0018	0.9991 ± 0.0010
Signal Region, SemiTight+SemiTight	0.9497 ± 0.0093	0.9898 ± 0.0029	0.9948 ± 0.0020	0.9988 ± 0.0012

To summarize, all relevant systematic biases of the template method have been examined and evaluated. The dominant uncertainty is the combined statistical and

systematic error returned by the fit (0.1–0.2% for $f_S > 0.99$ at 35.9 pb^{-1}). This error is driven by the small number of events in the background sample. There is a very small systematic bias (0.01%) arising from possible Drell-Yan contamination in the background control sample. Systematic biases due to the extrapolations to the signal and background control samples were evaluated, but no such biases were found. The expected total uncertainty on the jet background estimate for an analysis with a signal purity greater than 99% is about 0.1–0.2%. This proves the applicability of the template method to estimate the signal beneath the Z peak using a high purity selection to within an accuracy of 0.1–0.2% with 35.9 pb^{-1} of data.

7.3 Same/Opposite-Sign Method

The second method used to estimate the background beneath the $Z \to ee$ peak relies on the electron track charges. Specifically, the number of signal events under the Z peak is calculated from the number of events with opposite- and same-sign track charges, N_{OS} and N_{SS}, respectively:[1]

$$N_S = \frac{N_{OS} - N_{SS}}{(1 - 2q_1^{misid})(1 - 2q_2^{misid})} \tag{7.2}$$

where $q_{1(2)}^{misid}$ is the charge misidentification for the first (second) electron from the Z decay. By integrating over detector η and electron P_T, Equation 7.2 can be rewritten as:

$$N_S = \frac{N_{OS} - N_{SS}}{(1 - 2q^{misid})^2} \tag{7.3}$$

where q^{misid} is the average charge misidentification over both electrons.

In order to evaluate q^{misid} in Equation 7.3, a sample of dielectron events selected using much tighter criteria than the reference selection (Section 7.1) is constructed. This is necessary since the sample passing the reference selection may still be contaminated with QCD multijet background events. The tight sample is constructed using the WP85 (Table 6.8) selection. Using this sample, the charge misidentification for a single electron

[1] This result—Equation 7.2—is derived in Appendix A.

track can be extracted from the average product of the two electron charges:[2]

$$q^{\text{misid}} = \frac{1 - \sqrt{|\langle q_1 \times q_2 \rangle|}}{2} \tag{7.4}$$

Distributions for $q_1 \times q_2$ in the $Z \to ee$ signal MC events passing the reference and tight selections are shown in Figure 7.4. Approximately 11590 $Z \to ee$ events are expected to pass the reference selection (for the full 35.9 pb^{-1} dataset), with about 247 of them being same sign. This yields $\langle q_1 \times q_2 \rangle = -0.957 \pm 0.001$, and correspondingly $q^{\text{misid}} = 0.0108 \pm 0.0004$. For the tight selections, $\langle q_1 \times q_2 \rangle = -0.971 \pm 0.001$ and $q^{\text{misid}} = 0.0072 \pm 0.0002$ is obtained. This equates to a bias on the signal yield of about 1.45%, which is propagated as a systematic uncertainty.

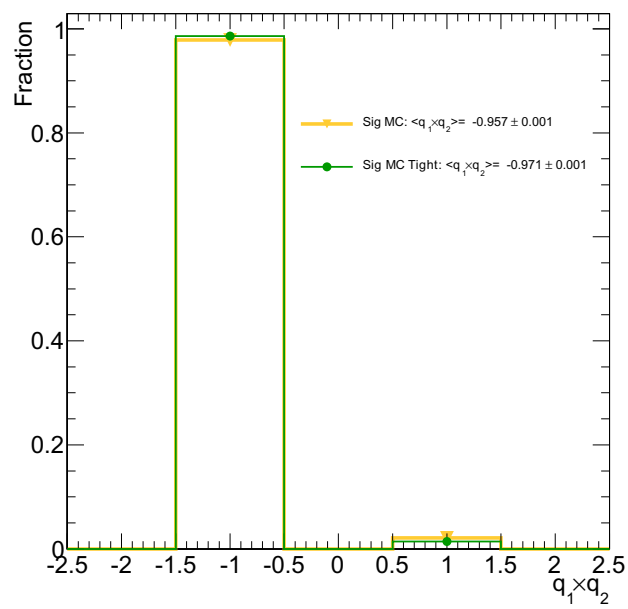


Figure 7.4: $q_1 \times q_2$ distributions for the MC $Z \to ee$ events with the reference (yellow) and tight (green) selections.

Figure 7.5 plots the $q_1 \times q_2$ distribution for all signal and background MC events after application of the reference selection. In this case, $\langle q_1 \times q_2 \rangle = -0.952 \pm 0.001$ and $q^{\text{misid}} = 0.0121 \pm 0.0004$. The signal and background MC contributions for the reference

[2]This result—Equation 7.4—is derived in Appendix A.

selection are provided in Table 7.5. These estimates lead to an expectation of 128 ± 11 background events in the mass range $M_{ee} \in (60, 120)$ GeV for 35.9 pb^{-1}.

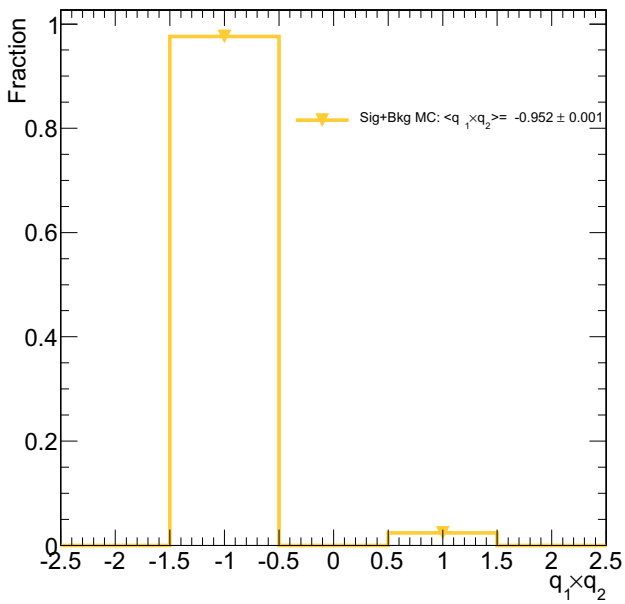


Figure 7.5: $q_1 \times q_2$ distribution for the $Z \to ee$ signal plus the expected background events in the MC after application of the reference selection criteria.

The background events may lead to a bias in the N_{S} value in Equation 7.3 in two cases:

- Violation of the charge symmetry assumption used in the derivation of Equation 7.3. (i.e., The numerator of this equation.)
- Violation of the assumption that there is a pure signal event yield in the tight sample used for the determination of q^{misid}.

To estimate the maximum uncertainty in the bias due to the charge symmetry assumption, all 58 non-QCD background events are assumed to be either OS or SS. N_{S} is then recalculated for each of these two assumptions and compared with the default value. In the case that all non-QCD background events are assumed to be SS (OS), the variation in N_{S} is about 1.06% (0.02%). The larger value is assigned as a systematic uncertainty.

Table 7.5: The contributions of the signal and background MC Samples after the reference WP95 selection for 35.9 pb^{-1}.

Channel	Yield	Raw Entries
MC $Z \to ee$	11590 ± 108	407824
QCD	70 ± 8	42
$t\bar{t}$	17 ± 4	3174
W+jets	6 ± 2	244
$Z \to \tau\tau$	16 ± 4	529
WW	4 ± 2	4820
WZ	9 ± 3	30374
ZZ	6 ± 2	38255
All Backgrounds	128 ± 11	77438

After the tight selection, the expected background contribution is 40 ± 6 events for 35.9 pb^{-1}, as shown in Table 7.6. This is a 3.2-fold reduction compared to the reference selection. To estimate the size of a maximum bias on q^{misid} due to the background still remaining in this tight sample, the following conservative scenario is employed: It is assumed that after application of the tight criteria there will be three times the background events expected from the MC. (i.e., It is assumed there will be $3 \times 40 = 120$ background events). After admixture of the threefold background events with the $Z \to ee$ events, the values $\langle q_1 \times q_2 \rangle = -0.971 \pm 0.001$ and $q^{\text{misid}} = 0.0074 \pm 0.0003$ are obtained. Figure 7.6 shows the $q_1 \times q_2$ distribution for this admixture. These numbers are in good agreement with the initial estimates for just signal events with the tight selection ($\langle q_1 \times q_2 \rangle = -0.971 \pm 0.001$ and $q^{\text{misid}} = 0.0072 \pm 0.0003$). The difference in these values implies a change of 0.06% in the signal yield, and is assigned as a systematic uncertainty. All the variations studied are listed in Table 7.7.

To summarize, the relevant systematic biases of the method have been examined and evaluated: These are listed in Table 7.8 in terms of the fractional uncertainty on the signal yield $\Delta N_{\text{S}}/N_{\text{S}}$. The total systematic uncertainty is 1.8%. This proves the applicability of the same/opposite sign charge method to estimate the signal yield beneath the Z peak using the reference selection criteria with an expected overall uncertainty of 1.8%.[3]

[3] At low integrated luminosity one must also account for the Poisson counting error in the q^{misid} value

Table 7.6: The contributions of the MC samples after the tight WP85 selection for 35.9 pb^{-1}.

Channel	Yield	Raw Entries
MC $Z/\gamma^* \to ee$	9699 ± 99	336562
QCD	0	0
$W + jet$	1 ± 1	42
$Z/\gamma^* \to \tau\tau$	11 ± 3	353
$t\bar{t}$	12 ± 4	2345
WW	3 ± 2	3827
WZ	8 ± 3	25004
ZZ	5 ± 2	31231
All Backgrounds	40 ± 6	62802

Table 7.7: $\langle q_1 \times q_2 \rangle$ and q^{misid} values for various samples.

	Signal		Signal+Background (Tight Selection)	
Variable	Reference	Tight	S+B	S+3×B
$\langle q_1 \times q_2 \rangle$	-0.9574 ± 0.0014	-0.9713 ± 0.0012	-0.9711 ± 0.0012	-0.9707 ± 0.0012
q^{misid}	0.01077 ± 0.00036	0.00724 ± 0.00032	0.00728 ± 0.00032	0.00737 ± 0.00032

7.4 Results on Data

Both the template and SS/OS methods have been applied on the full 35.9 pb^{-1} dataset using the standard WP80 $Z \to ee$ selection. The results of the background estimations using both methods are described in the following subsections.

due to the small number of observed same-sign events. For 1 pb^{-1}, this leads to an additional uncertainty of approximately 1.7%, for instance.

Table 7.8: The systematic uncertainties of the same/opposite sign method in terms of $\Delta N_S/N_S$ using the reference selection.

Source	$\Delta N_S/N_S$
Extrapolation of q^{misid} from Reference to Tight Sample	0.0145
Charge Symmetry Violation	0.0106
Background Remaining in the Tight Sample	0.0006
Total	0.0180

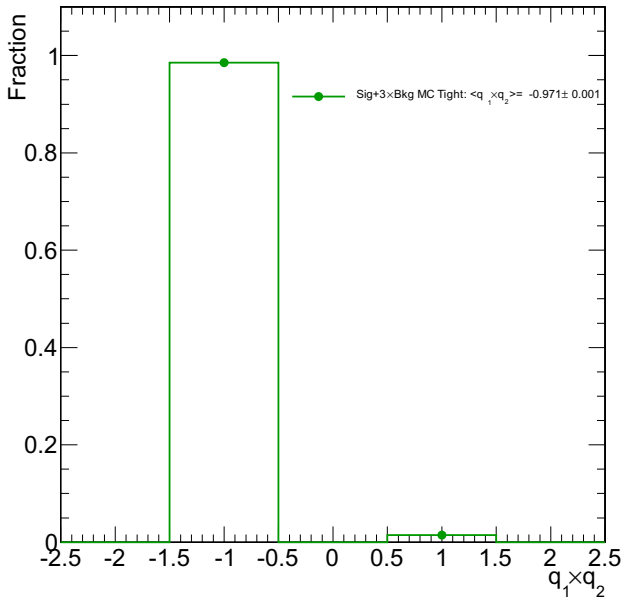


Figure 7.6: $q_1 \times q_2$ distribution for the $Z \rightarrow ee$ signal plus three times the expected background events in the MC, as they are found in the last column of Table 7.7, after application of the tight selection criteria.

7.4.1 Template Method Results

To apply this method, both signal and background templates for the $TrackIso/p_T$ variable were extracted directly from data control samples per the methodology described in Section 7.2. These can be seen in Figure 7.7, and the plot shows that the signal and background templates are well separated. The SemiTight+SemiTight signal region distribution (i.e., the selected events, but without a cut on the $TrackIso/p_T$) is also plotted, and its signal purity is estimated in this subsection. Figure 7.8 (7.9) compares the signal (background) template extracted from the data to that of the MC. The templates extracted from data agree well with their MC counterparts. The number of entries for the signal OS-Tight+Tight and background SS-SemiTight+Loose sideband templates is 15432 and 526 electrons, respectively. These templates are fit directly to the SemiTight+SemiTight sample in the signal region over the relative track isolation range that is populated in all histograms, which corresponds to $TrackIso/p_T \in (0, 0.35)$. This

yields a high quality fit with a $\chi^2/DOF < 1$, which is displayed in Figure 7.10. The fit yields the signal fraction in the SemiTight+SemiTight sample beneath a $TrackIso/p_T$ threshold of $T = 0.1$ as $f_{\mathrm{S}} = 0.9995 \pm 0.0008(\text{fit error}) \pm 0.0001(\text{syst})$. Although this threshold is higher than that applied in the standard selection (see Table 6.5), it already indicates the presence of zero jet background. The result of the fit may be interpreted in this way so that the final estimate for the jet backgrounds in the selected standard sample is quoted as an upper bound of 0.13%. This is consistent with the jet contribution of zero predicted by the MC. As described earlier, the non-jet backgrounds from the MC (30 ± 5 events for 35.9 pb^{-1}) must still be added in. This leads to a final template method based estimate for the background beneath the signal of $N_{\mathrm{B}} = 30 \pm 12(\text{syst})$, where the Poisson counting error from the non-jet background estimate has been propagated as a systematic uncertainty.

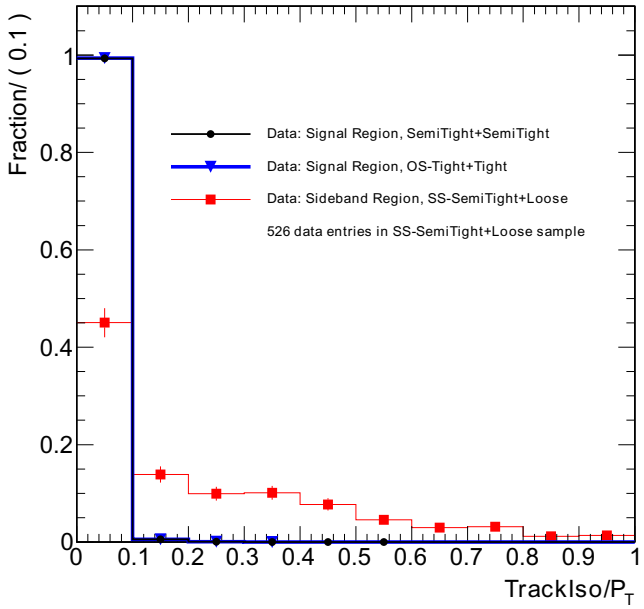


Figure 7.7: The signal (blue) and background (red) templates extracted from the data, along with the SemiTight+SemiTight sample in the signal region (black).

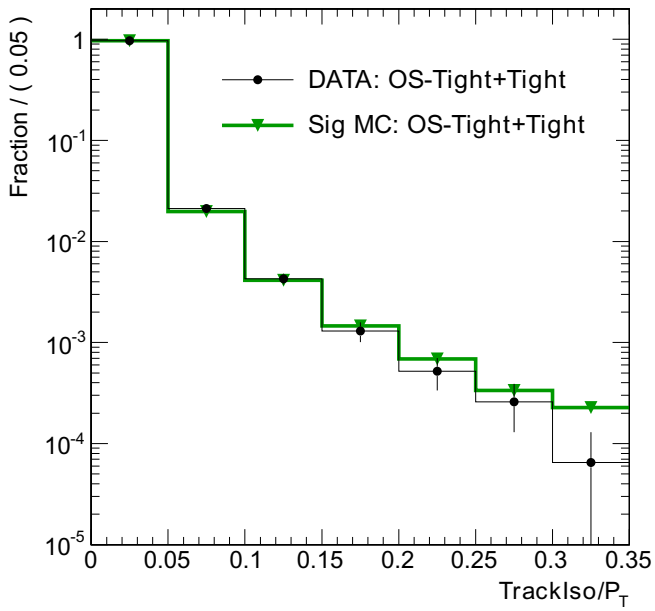


Figure 7.8: The signal template extracted from the data (black) and the signal MC (green).

7.4.2 Same/Opposite Sign Method Results

Figures 7.11–7.14 show distributions of the charge product $q_1 \times q_2$ in the selected $Z \to ee$ events in the data and signal MC, after application of the tight selection criteria in the signal invariant mass region $M_{ee} \in (60, 120)$ GeV. Here the tight selection uses the WP70 criteria (Table 6.9). A few topologies are considered:

- Both electrons anywhere in the Ecal (Figure 7.11).
- Both electrons in the Ecal Barrel (Figure 7.12).
- One electron in the Ecal Barrel and the other in the Ecal Endcap (Figure 7.13).
- Both electrons in the Ecal Endcap (Figure 7.14).

The MC distributions agree well with the data for each case.

In the selected data there are $N_{\mathrm{OS}} = 8325$ opposite-sign and $N_{\mathrm{SS}} = 116$ same-sign events. There are 6503 events remaining after the tight selection, for which the $q_1 \times q_2$

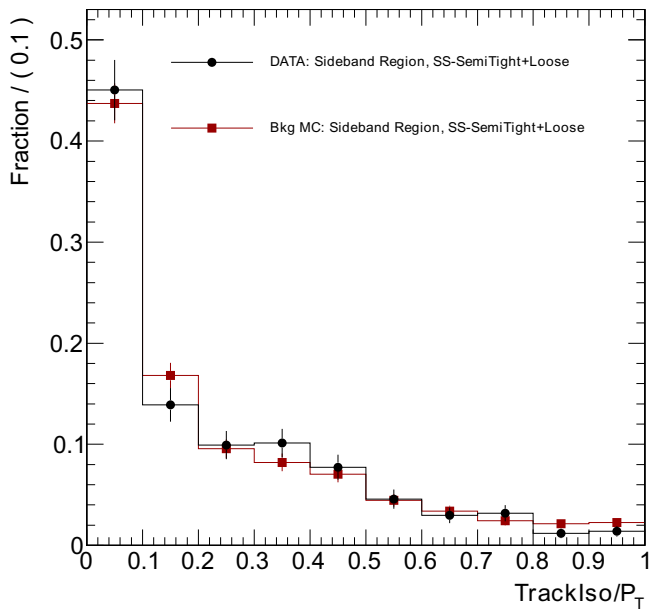


Figure 7.9: The background template extracted from the data (black) and the background MC (red).

distribution is shown in Figure 7.11. From the tight selection, the values $\langle q_1 \times q_2 \rangle = -0.979 \pm 0.002$ and $q^{misid} = 0.0052 \pm 0.0006$ are obtained. This is in agreement with the signal MC estimates of $\langle q_1 \times q_2 \rangle = -0.983 \pm 0.001$ and $q^{misid} = 0.0044 \pm 0.0003$. Following Equation 7.4 yields the total number of $Z \to ee$ events as $N_{\mathrm{S}} = 8382$.

The systematic biases have been recalculated with the MC for these selections according to the methods described in Section 7.3: These are listed in Table 7.9 in terms of the fractional uncertainty on the signal yield $\Delta N_{\mathrm{S}}/N_{\mathrm{S}}$. The Poisson counting error for the q^{misid} determination has also been propagated as a systematic uncertainty. Thus the total systematic uncertainty is 1.08% and the final result for the background beneath the $Z \to ee$ signal is $N_{\mathrm{B}} = 59 \pm 91$. This is consistent with the low background (30 ± 5 events) predicted by the MC.

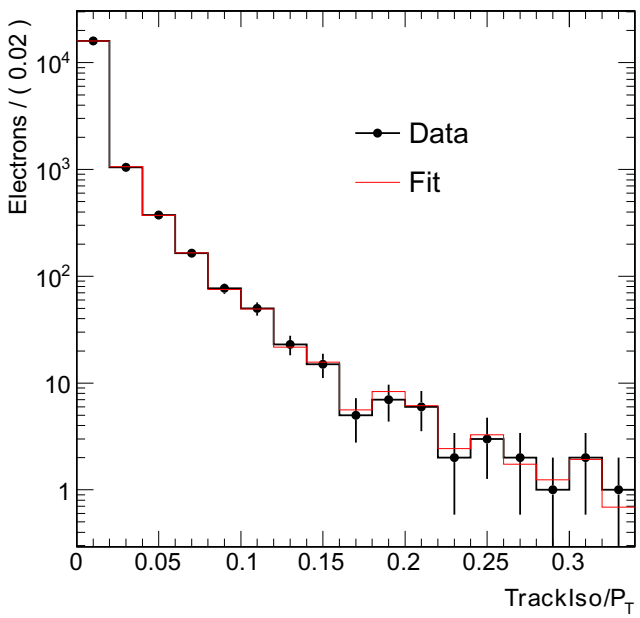


Figure 7.10: The fit (red) of the data (black) to the signal and background templates.

Table 7.9: The systematic uncertainties of the same/opposite sign method in terms of $\Delta N_S/N_S$ using the standard selection.

Source	$\Delta N_S/N_S$
Extrapolation of q^{misid} from Standard to Tight Sample	0.0076
Charge Symmetry Violation	0.0072
Background Remaining in the Tight Sample	0.0003
Poisson Counting Error in q^{misid} Determination	0.0025
Total	0.0108

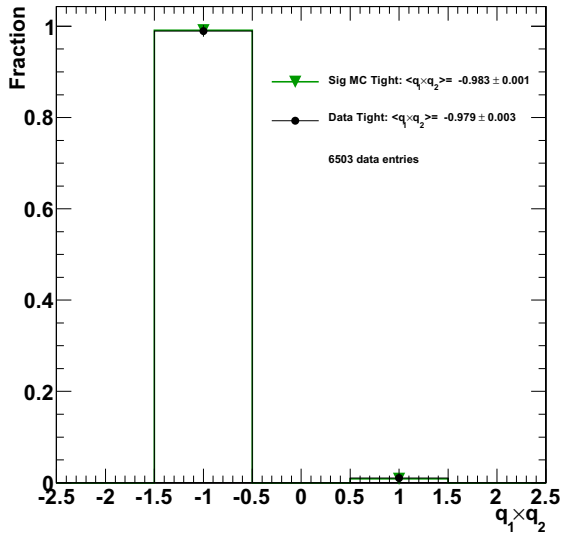


Figure 7.11: The $q_1 \times q_2$ distributions in data and MC for $Z \rightarrow ee$ candidates of all topologies passing the tight selection. The total number of events in data is shown on the plot.

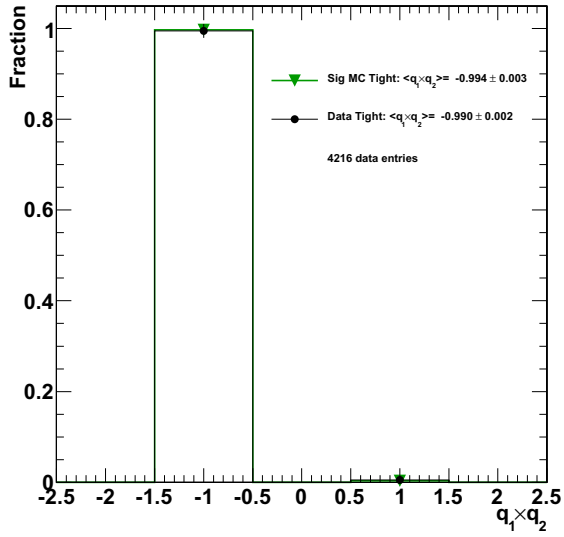


Figure 7.12: The $q_1 \times q_2$ distributions in data and MC for $Z \rightarrow ee$ candidates passing the tight selection and having both associated electrons in the Ecal Barrel. The total number of events in data is shown on the plot.

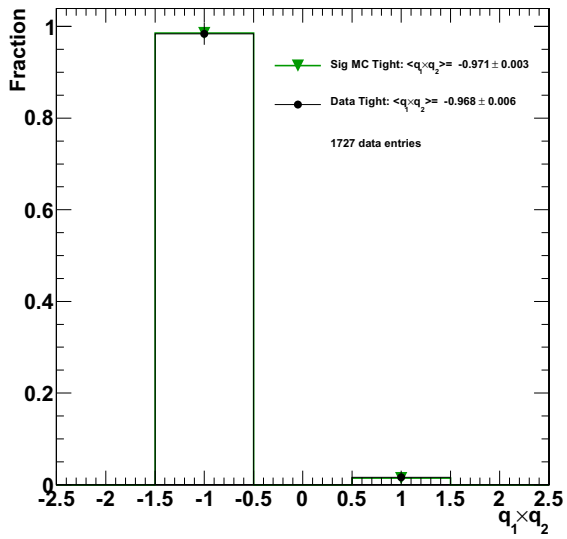


Figure 7.13: The $q_1 \times q_2$ distributions in data and MC for $Z \rightarrow ee$ candidates passing the tight selection and having one associated electron in the Ecal Barrel and the other in the Ecal Endcap. The total number of events in data is shown on the plot.

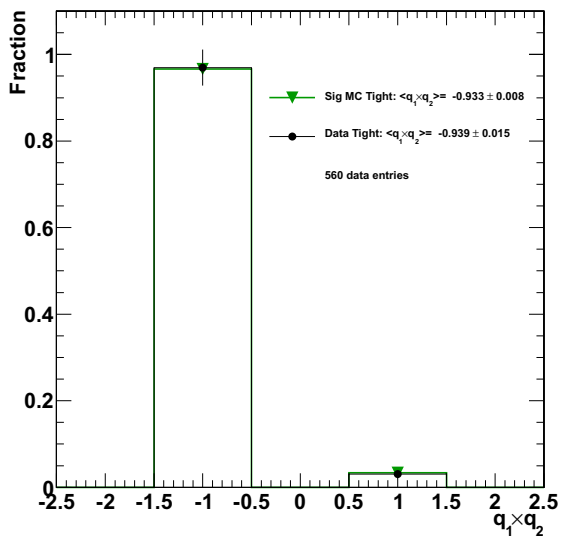


Figure 7.14: The $q_1 \times q_2$ distributions in data and MC for $Z \rightarrow ee$ candidates passing the tight selection and having both associated electrons in the Ecal Endcap. The total number of events in data is shown on the plot.

7.5 Summary

Two methods have been developed to estimate the background beneath the $Z \to ee$ peak. The primary method relies on signal and background templates of a chosen variable that are extracted from control samples. These templates are then fit to the selected data distribution to determine the jet background contribution. The remaining non-jet background contribution is then added from the MC. The signal and background templates were observed to be well separated in data, and also agreed well with their MC counterparts. The result of this method is a background estimate of $N_{\mathrm{B}} = 30 \pm 12(\mathrm{syst})$.

The second method incorporates an estimation of the q^{misid} so that the charge products of the dielectrons in the selected sample can be used to estimate the background. It lacks information to disentangle charge product symmetric from non-charge product symmetric backgrounds: A systematic uncertainty is propagated instead to provide coverage for all backgrounds. The q^{misid} values were observed to agree with the MC expectations. The result of this method is a background estimate of $N_{\mathrm{B}} = 59 \pm 91(\mathrm{syst})$.

The results of both data-driven background estimation methods are in good agreement with each other, and also with the MC prediction of $N_{\mathrm{B}} = 30 \pm 5(\mathrm{stat})$ from Table 7.1. This comprises a strong confirmation of the methods, the MC, and the CMS detector simulation. The quoted result that will be used in the cross section determination is that of the template method:

$$N_{\mathrm{B}} = 30 \pm 12(\mathrm{syst}) \tag{7.5}$$

or

$$N_{\mathrm{S}} = 8411 \pm 92(\mathrm{stat}) \pm 12(\mathrm{syst}) \tag{7.6}$$

This result is chosen since it has much less uncertainty than that of the same/opposite-sign method.

Chapter 8

$Z \to ee$ Invariant Mass Line Shape Model and Systematic Bias Study

A parameterization of the signal invariant mass line shape will be used to fit the $Z \to ee$ invariant mass distributions observed in data. This parameterization is needed both for the electron efficiency measurements and for a simultaneous fit that will be used to determine the cross section. The shape is parameterized by convolving the underlying physics shape with a detector resolution model. The physics shape is extracted as a histogram from the output of the Powheg [62] NLO MC generator. The chosen resolution model consists of a Crystal Ball function modified to include an extra half-Gaussian in the high-mass tail.[1] This choice is purely empirical: Before data taking commenced it was found to describe the resolution resulting from the detector simulation well. The background is described by a decaying exponential. The chosen signal plus background $(S + B)$ parameterization describes the data well and its statistical robustness is verified in this chapter.

[1] The (unnormalized) modified crystal ball resolution function is defined as:

$$f(M|\alpha, n, \bar{M}, \sigma, \sigma_2, \beta) = \begin{cases} \beta e^{\frac{-(M-\bar{M})^2}{2\sigma^2}} + (1-\beta) e^{\frac{-(M-\bar{M})^2}{2\sigma_2^2}}, \text{ for } \frac{M-\bar{M}}{\sigma} \geqslant -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{|\alpha|^2}{2}} \left(\frac{n}{|\alpha|} - |\alpha| - \frac{M-\bar{M}}{\sigma}\right)^{-n}, \text{ for } \frac{M-\bar{M}}{\sigma} < -\alpha \end{cases} \tag{8.1}$$

where $\beta \in (0, 1)$.

8.1 Fit to the Invariant Mass Distribution of the Selected $Z \to ee$ events

Figure 8.1 shows the invariant mass distribution of the 8441 selected $Z \to ee$ events fit to the chosen $S + B$ parameterization using an unbinned extended maximum log-likelihood fit. The background is fixed to the level estimated by the template method ($N_{\mathrm{B}} = 30 \pm 12$), and is assumed flat. The $\bar{M}$ and σ parameters of the modified Crystal-Ball are floating. The fit has $\chi^2/DOF = 0.990$, corresponding to a probability value of $p = 0.549$. The errors of the floating parameters have been evaluated with MINOS [76], and the absolute values of the positive and negative errors of the parameters are near to each other. This confirms that the fit has good quality. In the following chapters it will be necessary to fit $Z \to ee$ invariant mass distributions that have been selected with much looser criteria: It will be shown in turn that these fits also have good quality.

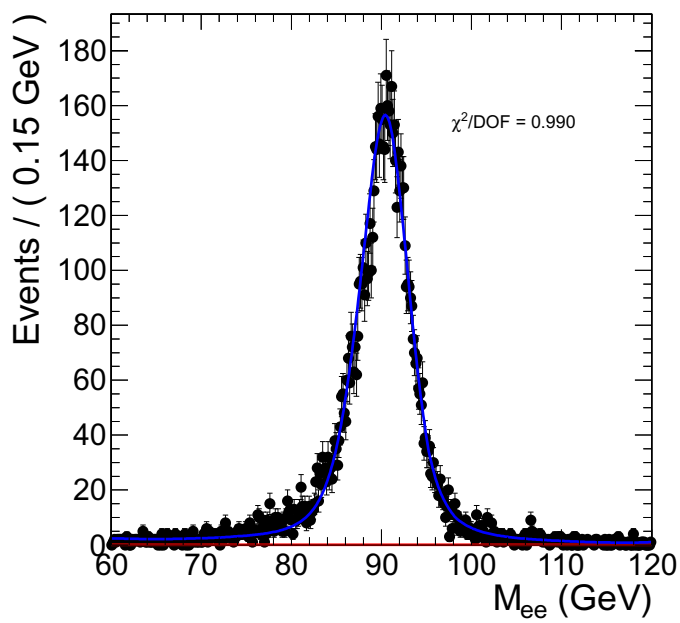


Figure 8.1: The invariant mass distribution of the selected $Z \to ee$ events fit to the $S + B$ parameterization. The data is in black, the background fit result in red, and the $S + B$ fit result in blue. The $\chi^2/DOF = 0.990$, corresponding to a probability value of $p = 0.549$.

8.2 Systematic Bias Study

In addition to confirming that the fits have high probability values and well-evaluated parameter errors, it must also be verified that they accurately reproduce the signal yields and their errors. This is done by performing thousands of pseudo-experiments. Toy MC ensembles are generated using the parameter values found in data for a loose sample with low signal over background (S/B).[2] Fits to these ensembles are then performed as they would be in the data. Figure 8.2 displays a histogram of the ratio of the signal yields reported by the fits to the true signal yields of the toy MC. This histogram is fit to a Gaussian, which has a mean close to one, demonstrating that the fits accurately reproduce the true signal yields. Figure 8.3 plots the pull distribution in the yield variable of the pseudo-experiments.[3] This histogram is fit to a Gaussian, which has a mean close to zero and width close to one, demonstrating that the fits accurately reproduce the true errors on the signal yields.

8.3 Summary

The $S + B$ invariant mass line shape parameterization has been developed. The signal model is the generator-level physics shape convolved with a resolution function determined empirically from the detector resolution in the simulation: a Crystal Ball modified to include an extra half-Gaussian in the high-mass tail. The background is modeled as a falling exponential. This choice has been seen to describe the data well, yielding a good quality fit and properly evaluated parameter errors. Moreover, the statistical robustness of the parameterization has been verified in a systematic bias study using pseudo-experiments to demonstrate that the signal yields and errors are accurately reproduced by the fits.

[2] This low S/B case corresponds to the SC→Reco Tag+Fail sample in the Ecal Barrel for the efficiency calculation, as described in Chapter 9.

[3] The pull value for a given variable is defined as the difference between its observed and true value divided by its reported error; i.e., the pull value is $(N_{\text{observed}} - N_{\text{truth}})/\Delta N_{\text{observed}}$.

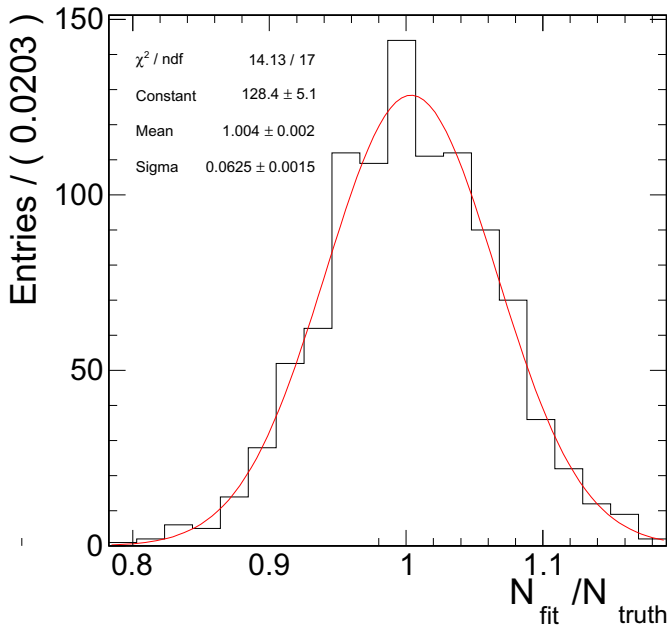


Figure 8.2: Histogram of the ratio of the signal yields reported by the fits to the true signal yields of the pseudo-experiments (black). The histogram is fit to a Gaussian (red).

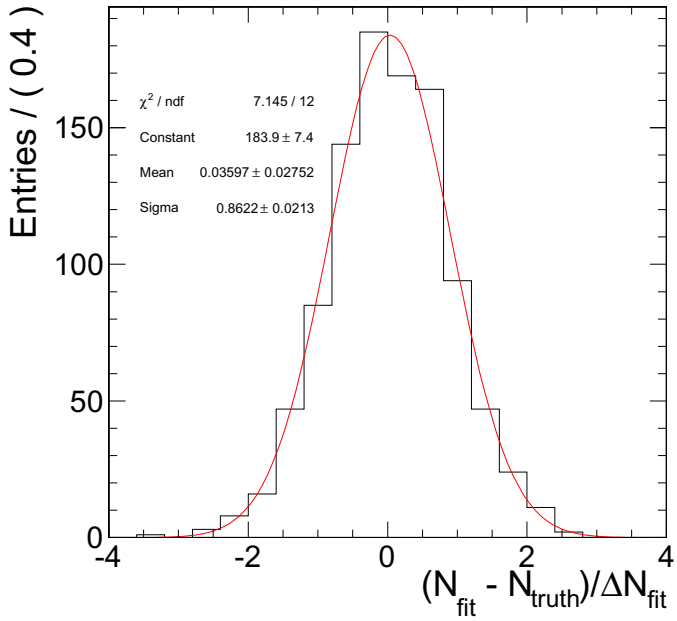


Figure 8.3: The pull distribution in the yield variable of the pseudo-experiments (black). The histogram is fit to a Gaussian (red).

Chapter 9

Tag and Probe Electron Efficiency Measurements

9.1 Motivation

Electron efficiency measurements are a central component of any physics analysis involving electrons in the final state, including the $Z \rightarrow ee$ cross section measurement, where the electron efficiency enters into the cross section calculation of Equation 1.7. The Tag and Probe (TP) method used to determine the efficiencies is described in this chapter. All relevant systematic biases are evaluated, and the results are detailed in multiple tables and plots. Several cross-checks are performed to ensure the robustness of the results, and comparisons to the MC are made.

9.2 Introduction to the Tag and Probe Method

The TP method [77] uses $Z \rightarrow ee$ decays as a high-purity source of unbiased electrons from which to extract efficiencies. The tag electron is the control electron to which stringent selection criteria is applied. The probe electron is the test electron used to measure the efficiency of the particular selection criteria under study. Naturally, imposing an invariant mass cut on the Tag+Probe pair about the $Z \rightarrow ee$ peak ensures a high-purity sample of

Tag+Probe pairs.

Construction of the tag and probe criteria leads to two samples: pairs with one tag electron and one probe electron passing the particular selection requirements under study (Tag+Pass), and pairs with one tag electron and one probe electron failing the selection requirements under study (Tag+Fail). The efficiency calculation then reduces to an estimation of the signal yield in the passing and failing samples. This estimation is carried out via unbinned maximum log-likelihood fits to the dielectron invariant mass distributions in the two samples.[1]

The electron efficiency calculation is broken down into three consecutive steps:

1. Acc→Reco: Supercluster to reconstructed (GSF track-matched) electron.
2. Reco→ID: Reconstructed electron to the WP80 selection criteria (ID).[2]
3. ID→Trig: WP80 to the online trigger requirement.

These three efficiencies are denoted $\epsilon_{\text{Acc}\to\text{Reco}}$, $\epsilon_{\text{Reco}\to\text{ID}}$, and $\epsilon_{\text{ID}\to\text{Trig}}$, respectively. For a $Z\to ee$ event to pass the selection, both electron daughters must be reconstructed and pass the ID, while one must pass the trigger requirement. Thus the signal efficiency is given by:

$$\epsilon_Z = \epsilon^2_{\text{Acc}\to\text{Reco}}\epsilon^2_{\text{Reco}\to\text{ID}}(1-(1-\epsilon_{\text{ID}\to\text{Trig}})^2) \tag{9.1}$$

The electron efficiency measurements are performed for a two-binned scheme and a scheme finely binned in both P_T (6 bins) and η (10 bins). The two-binned scheme measures the efficiencies separately in the Ecal Barrel and Endcap, since the electron selection and background levels are both different in these two regions. The efficiency may have some residual charge dependence due to possible correlations between electron reconstruction/selection efficiency and the level of misalignment between the tracker and the Ecal: Thus results are provided for both positive and negative charges in the

[1] Other methods of estimating the signal yield are possible; e.g., either the template or same/opposite-sign method of Chapter 7 could be used.

[2] In Subsection 6.1.3, the "identification" variables referred to only a subset of the WP80 criteria, as is the convention of the CMS collaboration. More generic terminology is preferred from here on so that "ID" refers to all of the WP80 offline selection criteria (Table 6.5).

two-binned scheme. It is not actually necessary to adopt the fine-binned scheme, since the efficiency variations in the kinematic phase space are folded into an average effect. Nevertheless, performing this fine-binned evaluation provides a cross-check of the primary two-binned results and allows scanning for irregularities that may otherwise go unnoticed.

9.3 Tag and Probe Pair Selections

The Tag+Probe pairs are selected within the invariant mass window $M_{ee} \in (60, 120)$ GeV and the fiducial and kinematic acceptance (Subsection 6.1.1). The definitions of the Tag, Probe, and Passing Probe for each efficiency step are as follows.

9.3.1 Acc→Reco

- Tag: Reconstructed electron passing the ID and matched to an HLT object within a cone of aperture $\Delta R < 0.3$.
- Probe: Supercluster with $H/E < 0.15$.
- Passing Probe: Supercluster matched to a reconstructed electron in the event within a cone of aperture $\Delta R < 0.3$.

The Tag+Fail sample has a large background from QCD multijet events in the presence of a small signal ($S/B \ll 1$), making it difficult to reliably estimate the number of signal events. Therefore, the loose selection criterion $H/E < 0.15$ is applied to the probe to reduce the background. This cleaning cut is approximately 99.8% efficient for signal events in the MC. The bias it introduces is evaluated in Subsection 9.6.5 and is assigned a systematic uncertainty of 0.42%.

9.3.2 Reco→ID

- Tag: Reconstructed electron passing the ID and matched to an HLT object within a cone of aperture $\Delta R < 0.3$.

- Probe: Reconstructed electron.
- Passing Probe: Reconstructed electron passing the ID.

9.3.3 ID→Trig

- Tag: Reconstructed electron passing the ID and matched to an HLT object within a cone of aperture $\Delta R < 0.3$.
- Probe: Reconstructed electron passing the ID.
- Passing Probe: Reconstructed electron passing the ID and matched to an HLT object within a cone of aperture $\Delta R < 0.3$.

9.3.4 Arbitration in Case of Multiple Candidates

It is rare ($\ll 1\%$) to find more than one Z candidate in an event if the probe is also well identified (e.g., the ID→Trig step). Multiple Z candidates occur approximately 2–3% of the time when the probe is loosely selected (e.g., the Acc→Reco step), however, due to combinatorics with electrons from jets. In cases where there is more than one Z candidate to select the pair from, the following arbitration is made:

- If there are two probe candidates in the event and only one of them passes the tag criteria then the one which passes the tag criteria is chosen.
- If there are two probe candidates in the event and both pass the tag criteria then one is chosen at random.
- If there are two probe candidates in the event and both fail the tag criteria then one is chosen at random.
- If there are more than two probe candidates in the event then it is rejected from consideration (i.e., the event is dropped from both the numerator and denominator) to avoid biases that would be introduced by attempting to choose the "correct" probe.

9.4 Extraction of Efficiency and Signal Yields

The signal components of the Tag+Pass and Tag+Fail samples are extracted via unbinned maximum log-likelihood fits performed in the dielectron invariant mass variable using the signal and background models described in Chapter 8. The floating parameters of the fits are listed below:

- $f^{\text{pass}}_{\text{signal}}$: fraction of Tag+Pass pairs that are signal.
- $f^{\text{fail}}_{\text{signal}}$: fraction of Tag+Fail pairs that are signal.
- $P^{\text{pass}}_{\text{bkg}}(M_{ee}) \propto e^{-\chi^{\text{pass}} M_{ee}}$: background PDF in the Tag+Pass sample with the exponential shape parameter χ^{pass} floating.
- $P^{\text{fail}}_{\text{bkg}}(M_{ee}) \propto e^{-\chi^{\text{fail}} M_{ee}}$: background PDF in the Tag+Fail sample with the exponential shape parameter χ^{fail} floating.
- $P^{\text{pass}}_{\text{signal}}(M_{ee})$: signal PDF in the passing case that is floating as described below.
- $P^{\text{fail}}_{\text{signal}}(M_{ee})$: signal PDF in the failing case that is floating as described below.

The signal and background contributions are extracted in the Tag+Pass and Tag+Fail samples through the following relations:

$$N^{\text{pass}}_{\text{signal}}(M_{ee}) = f^{\text{pass}}_{\text{signal}} \times N^{\text{pass}} \times P^{\text{pass}}_{\text{signal}}(M_{ee}) \tag{9.2}$$

$$N^{\text{pass}}_{\text{bkg}}(M_{ee}) = (1 - f^{\text{pass}}_{\text{signal}}) \times N^{\text{pass}} \times P^{\text{pass}}_{\text{bkg}}(M_{ee}) \tag{9.3}$$

$$N^{\text{fail}}_{\text{signal}}(M_{ee}) = f^{\text{fail}}_{\text{signal}} \times N^{\text{fail}} \times P^{\text{fail}}_{\text{signal}}(M_{ee}) \tag{9.4}$$

$$N^{\text{fail}}_{\text{bkg}}(M_{ee}) = (1 - f^{\text{fail}}_{\text{signal}}) \times N^{\text{fail}} \times P^{\text{fail}}_{\text{bkg}}(M_{ee}) \tag{9.5}$$

where N^{pass} (N^{fail}) is the total number of tag and probe pairs in the Tag+Pass (Tag+Fail) sample. The efficiency is then given by:

$$\epsilon = \frac{\int N^{\text{pass}}_{\text{signal}}(M_{ee}) dM_{ee}}{\int N^{\text{pass}}_{\text{signal}}(M_{ee}) dM_{ee} + \int N^{\text{fail}}_{\text{signal}}(M_{ee}) dM_{ee}} = \frac{f^{\text{pass}}_{\text{signal}} \times N^{\text{pass}}}{f^{\text{pass}}_{\text{signal}} \times N^{\text{pass}} + f^{\text{fail}}_{\text{signal}} \times N^{\text{fail}}} \tag{9.6}$$

In the majority of the fits, the $\bar{M}$ and σ parameters of the modified Crystal Ball function are allowed to float in both the passing and failing samples. In a few instances, however, the fit quality or stability necessitates changing the floating parameters from this default. In all cases the MINOS [76] uncertainties are evaluated to confirm the stability of the fit. In addition to the MINOS evaluation, detailed systematic bias studies similar to those of Section 8.2 have been carried out. The contribution of background in the Tag+Pass sample of the Reco→ID step is very low (less than 1%, as observed in Chapter 7). In order to avoid biasing the fit for this step, the background is constrained in the Tag+Pass sample to the level determined in Chapter 7 and its distribution is assumed uniform. The background is also very low in both the Tag+Pass and Tag+Fail samples of the ID→Trig step. Hence the efficiency for this step is determined by just counting the background subtracted signal yields in the passing and failing samples rather than by performing a fit.[3]

9.5 Results

The two-binned results are provided in Tables 9.1–9.3 for the three efficiency steps Acc→Reco, Reco→ID, and ID→Trig, respectively. The corresponding fit plots are provided in Figures 9.1–9.18. The fine-binned results are provided in Tables 9.4–9.6. Results from the signal MC are provided in the tables for comparison, and are determined via counting. The data to MC efficiency ratios are also provided for comparison. All averages are calculated using weights determined from the generator-level MC occupancies. The fits return combined statistical and systematic errors.[4] It sometimes occurs that this combined error is less than the binomial statistical error, in which case the binomial error is quoted instead. The binomial error is also quoted for the ID→Trig and MC results, since these are obtained from counting.

[3]Fits are still carried out for this step, however, and are displayed in the next section to demonstrate the high quality of the signal parameterization in this tightest case.

[4]MINOS attempts to evaluate the total error on each floating parameter, so that the uncertainties reported by the fits generally have statistical and systematic components, with correlations being properly accounted for.

The results from the two-binned and fine-binned schemes are observed to be in good agreement. No significant charge dependence is observed in the results for the Reco→ID and ID→Trig steps. A 2.4σ deviation (1.5% significance) is observed in the Ecal Barrel between positively and negatively charged electrons for the Acc→Reco step, however, suggesting a minor misalignment between the Ecal Barrel and tracker that is impacting electron reconstruction. The normalizations of the efficiencies disagree between data and MC, but the trends in both P_T and η generally agree well, as is shown in Subsection 9.5.1. There is no indication of time-dependence in $\epsilon_{\text{Acc}\to\text{Reco}}$ or $\epsilon_{\text{Reco}\to\text{ID}}$, as shown in Subsection 9.5.2, while there is for $\epsilon_{\text{ID}\to\text{Trig}}$.

Table 9.1: The two-binned Acc→Reco efficiencies.

Topology	ϵ_{MC}	ϵ_{Data}	$\epsilon_{\text{Data}}/\epsilon_{\text{MC}}$
EB	0.9778	0.9703 ± 0.0029	0.9923 ± 0.0030
EE	0.9461	0.9430 ± 0.0035	0.9968 ± 0.0037
⟨EB⊕EE⟩	0.9690	0.9627 ± 0.0023	0.9935 ± 0.0024
EB e^-	0.9790	0.9632 ± 0.0042	0.9839 ± 0.0043
EE e^-	0.9448	0.9419 ± 0.0044	0.9970 ± 0.0046
EB e^+	0.9767	0.9766 ± 0.0033	0.9999 ± 0.0034
EE e^+	0.9475	0.9447 ± 0.0052	0.9971 ± 0.0055

Table 9.2: The two-binned Reco→ID efficiencies.

Topology	ϵ_{MC}	ϵ_{Data}	$\epsilon_{\text{Data}}/\epsilon_{\text{MC}}$
EB	0.8747	0.8397 ± 0.0032	0.9600 ± 0.0037
EE	0.7561	0.7307 ± 0.0065	0.9664 ± 0.0086
⟨EB⊕EE⟩	0.8415	0.8092 ± 0.0029	0.9616 ± 0.0035
EB e^-	0.8763	0.8354 ± 0.0046	0.9534 ± 0.0053
EE e^-	0.7560	0.7358 ± 0.0090	0.9732 ± 0.0120
EB e^+	0.8732	0.8409 ± 0.0045	0.9630 ± 0.0052
EE e^+	0.7563	0.7242 ± 0.0093	0.9576 ± 0.0123

Table 9.3: The two-binned ID→Trig efficiencies.

Topology	ϵ_{MC}	ϵ_{Data}	$\epsilon_{\mathrm{Data}}/\epsilon_{\mathrm{MC}}$
EB	0.9710	0.9799 ± 0.0013	1.0092 ± 0.0013
EE	0.9716	0.9730 ± 0.0025	1.0014 ± 0.0026
$\langle$EB$\oplus$EE$\rangle$	0.9712	0.9780 ± 0.0012	1.0070 ± 0.0012
EB e^-	0.9716	0.9782 ± 0.0019	1.0068 ± 0.0019
EE e^-	0.9727	0.9705 ± 0.0037	0.9978 ± 0.0038
EB e^+	0.9704	0.9816 ± 0.0017	1.0115 ± 0.0018
EE e^+	0.9705	0.9755 ± 0.0034	1.0052 ± 0.0035

Table 9.4: The fine-binned Acc→Reco efficiencies.

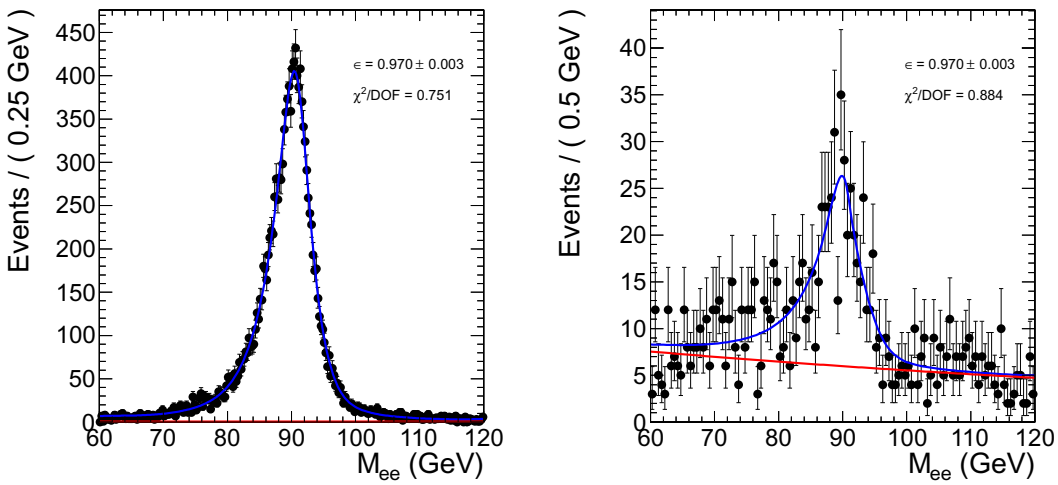
P_T Range (GeV)	η Range	ϵ_{MC}	ϵ_{Data}	$\epsilon_{\mathrm{Data}}/\epsilon_{\mathrm{MC}}$
(25, 30)	(-2.5, -2)	0.9135	0.9234 ± 0.0247	1.0108 ± 0.0271
	(-2, -1.5)	0.9478	0.9703 ± 0.0216	1.0237 ± 0.0228
	(-1.5, -1)	0.9522	0.9279 ± 0.0718	0.9745 ± 0.0754
	(-1, -0.5)	0.9610	0.9651 ± 0.0239	1.0043 ± 0.0249
	(-0.5, 0)	0.9611	0.9487 ± 0.0270	0.9871 ± 0.0281
	(0, 0.5)	0.9601	0.9659 ± 0.0459	1.0061 ± 0.0478
	(0.5, 1)	0.9607	0.9049 ± 0.0313	0.9420 ± 0.0325
	(1, 1.5)	0.9706	0.9951 ± 0.0152	1.0253 ± 0.0157
	(1.5, 2)	0.9361	0.9401 ± 0.0256	1.0042 ± 0.0273
	(2, 2.5)	0.8907	0.8860 ± 0.0231	0.9947 ± 0.0260
(30, 35)	(-2.5, -2)	0.9265	0.9469 ± 0.0180	1.0220 ± 0.0195
	(-2, -1.5)	0.9588	0.9462 ± 0.0141	0.9868 ± 0.0147
	(-1.5, -1)	0.9726	0.9782 ± 0.0113	1.0057 ± 0.0117
	(-1, -0.5)	0.9718	0.9863 ± 0.0083	1.0149 ± 0.0086
	(-0.5, 0)	0.9749	0.9931 ± 0.0099	1.0186 ± 0.0101
	(0, 0.5)	0.9729	0.9281 ± 0.0160	0.9540 ± 0.0164
	(0.5, 1)	0.9780	0.9641 ± 0.0156	0.9858 ± 0.0160
	(1, 1.5)	0.9712	0.9504 ± 0.0116	0.9786 ± 0.0119
	(1.5, 2)	0.9603	0.9727 ± 0.0146	1.0130 ± 0.0152
	(2, 2.5)	0.9267	0.9663 ± 0.0132	1.0428 ± 0.0143
(35, 40)	(-2.5, -2)	0.9326	0.9447 ± 0.0134	1.0130 ± 0.0143
	(-2, -1.5)	0.9539	0.9152 ± 0.0147	0.9594 ± 0.0154
	(-1.5, -1)	0.9810	0.9592 ± 0.0133	0.9778 ± 0.0135
	(-1, -0.5)	0.9794	0.9627 ± 0.0078	0.9830 ± 0.0080
	(-0.5, 0)	0.9775	0.9859 ± 0.0078	1.0086 ± 0.0079
	(0, 0.5)	0.9784	0.9719 ± 0.0078	0.9934 ± 0.0080
	(0.5, 1)	0.9741	0.9673 ± 0.0099	0.9930 ± 0.0101
	(1, 1.5)	0.9744	0.9613 ± 0.0253	0.9865 ± 0.0260
	(1.5, 2)	0.9679	0.9624 ± 0.0097	0.9943 ± 0.0100
	(2, 2.5)	0.9289	0.9105 ± 0.0836	0.9801 ± 0.0900
(40, 45)	(-2.5, -2)	0.9299	0.9177 ± 0.0745	0.9869 ± 0.0801
	(-2, -1.5)	0.9527	0.9834 ± 0.0181	1.0322 ± 0.0190
	(-1.5, -1)	0.9773	0.9597 ± 0.0094	0.9820 ± 0.0096
	(-1, -0.5)	0.9833	0.9684 ± 0.0069	0.9849 ± 0.0070
	(-0.5, 0)	0.9815	0.9496 ± 0.0309	0.9675 ± 0.0315
	(0, 0.5)	0.9789	0.9797 ± 0.0074	1.0008 ± 0.0076
	(0.5, 1)	0.9854	0.9667 ± 0.0073	0.9810 ± 0.0074
	(1, 1.5)	0.9787	0.9729 ± 0.0079	0.9941 ± 0.0080
	(1.5, 2)	0.9724	0.9667 ± 0.0304	0.9941 ± 0.0312
	(2, 2.5)	0.9386	0.9364 ± 0.0449	0.9977 ± 0.0478
(45, 50)	(-2.5, -2)	0.9393	0.9472 ± 0.0337	1.0085 ± 0.0359
	(-2, -1.5)	0.9692	0.9631 ± 0.0256	0.9938 ± 0.0264
	(-1.5, -1)	0.9790	0.9937 ± 0.0093	1.0150 ± 0.0095
	(-1, -0.5)	0.9886	0.9879 ± 0.0069	0.9993 ± 0.0070
	(-0.5, 0)	0.9865	0.9811 ± 0.0064	0.9945 ± 0.0065
	(0, 0.5)	0.9796	0.9745 ± 0.0082	0.9947 ± 0.0084
	(0.5, 1)	0.9813	0.9761 ± 0.0080	0.9947 ± 0.0081
	(1, 1.5)	0.9682	0.9778 ± 0.0155	1.0100 ± 0.0161
	(1.5, 2)	0.9746	0.9680 ± 0.0308	0.9932 ± 0.0316
	(2, 2.5)	0.9314	0.8929 ± 0.0241	0.9587 ± 0.0259
(50, ∞)	(-2.5, -2)	0.9287	0.9557 ± 0.0191	1.0291 ± 0.0206
	(-2, -1.5)	0.9718	0.9643 ± 0.0168	0.9922 ± 0.0173
	(-1.5, -1)	0.9869	0.9261 ± 0.0191	0.9384 ± 0.0193
	(-1, -0.5)	0.9912	0.9725 ± 0.0099	0.9812 ± 0.0100
	(-0.5, 0)	0.9794	0.9764 ± 0.0088	0.9969 ± 0.0090
	(0, 0.5)	0.9905	0.9870 ± 0.0077	0.9965 ± 0.0077
	(0.5, 1)	0.9879	0.9552 ± 0.0138	0.9669 ± 0.0140
	(1, 1.5)	0.9747	0.9757 ± 0.0156	1.0011 ± 0.0160
	(1.5, 2)	0.9776	0.9458 ± 0.0220	0.9675 ± 0.0226
	(2, 2.5)	0.9385	0.9350 ± 0.0643	0.9963 ± 0.0685
average		0.9701	0.9616 ± 0.0029	0.9912 ± 0.0030

Table 9.5: The fine-binned Reco→ID efficiencies.

P_T Range (GeV)	η Range	ϵ_{MC}	ϵ_{Data}	$\epsilon_{\mathrm{Data}}/\epsilon_{\mathrm{MC}}$
(25, 30)	(-2.5, -2)	0.6852	0.6065 ± 0.0423	0.8851 ± 0.0617
	(-2, -1.5)	0.6003	0.5864 ± 0.0569	0.9769 ± 0.0948
	(-1.5, -1)	0.7115	0.7605 ± 0.0404	1.0688 ± 0.0567
	(-1, -0.5)	0.7960	0.7104 ± 0.0547	0.8925 ± 0.0687
	(-0.5, 0)	0.7644	0.7401 ± 0.0347	0.9682 ± 0.0454
	(0, 0.5)	0.8058	0.8265 ± 0.0278	1.0257 ± 0.0345
	(0.5, 1)	0.7772	0.8486 ± 0.0420	1.0919 ± 0.0540
	(1, 1.5)	0.7171	0.6632 ± 0.0395	0.9248 ± 0.0551
	(1.5, 2)	0.6231	0.5925 ± 0.0481	0.9508 ± 0.0773
	(2, 2.5)	0.6972	0.7331 ± 0.0394	1.0514 ± 0.0565
(30, 35)	(-2.5, -2)	0.7618	0.6827 ± 0.0343	0.8963 ± 0.0450
	(-2, -1.5)	0.6615	0.6394 ± 0.0365	0.9666 ± 0.0551
	(-1.5, -1)	0.8015	0.8007 ± 0.0251	0.9990 ± 0.0313
	(-1, -0.5)	0.8483	0.7909 ± 0.0217	0.9323 ± 0.0256
	(-0.5, 0)	0.8340	0.7875 ± 0.0226	0.9442 ± 0.0271
	(0, 0.5)	0.8453	0.8098 ± 0.0203	0.9579 ± 0.0241
	(0.5, 1)	0.8448	0.7786 ± 0.0226	0.9216 ± 0.0267
	(1, 1.5)	0.7613	0.7929 ± 0.0254	1.0415 ± 0.0333
	(1.5, 2)	0.6656	0.5846 ± 0.0340	0.8783 ± 0.0511
	(2, 2.5)	0.7294	0.7108 ± 0.0316	0.9745 ± 0.0433
(35, 40)	(-2.5, -2)	0.7542	0.7058 ± 0.0308	0.9357 ± 0.0409
	(-2, -1.5)	0.7162	0.6682 ± 0.0296	0.9330 ± 0.0413
	(-1.5, -1)	0.8407	0.8067 ± 0.0204	0.9596 ± 0.0242
	(-1, -0.5)	0.8854	0.8684 ± 0.0153	0.9808 ± 0.0172
	(-0.5, 0)	0.8697	0.7810 ± 0.0189	0.8980 ± 0.0217
	(0, 0.5)	0.8686	0.8243 ± 0.0175	0.9490 ± 0.0202
	(0.5, 1)	0.8897	0.8576 ± 0.0157	0.9639 ± 0.0177
	(1, 1.5)	0.8427	0.8309 ± 0.0197	0.9860 ± 0.0234
	(1.5, 2)	0.7071	0.6789 ± 0.0300	0.9602 ± 0.0424
	(2, 2.5)	0.7769	0.7567 ± 0.0276	0.9740 ± 0.0355
(40, 45)	(-2.5, -2)	0.8165	0.7707 ± 0.0263	0.9438 ± 0.0322
	(-2, -1.5)	0.7917	0.7734 ± 0.0329	0.9768 ± 0.0415
	(-1.5, -1)	0.8886	0.9064 ± 0.0136	1.0200 ± 0.0153
	(-1, -0.5)	0.9078	0.8907 ± 0.0127	0.9812 ± 0.0140
	(-0.5, 0)	0.9101	0.8521 ± 0.0148	0.9362 ± 0.0163
	(0, 0.5)	0.9089	0.8718 ± 0.0133	0.9591 ± 0.0147
	(0.5, 1)	0.9049	0.8859 ± 0.0136	0.9790 ± 0.0151
	(1, 1.5)	0.8859	0.8616 ± 0.1168	0.9726 ± 0.1318
	(1.5, 2)	0.7851	0.7290 ± 0.0253	0.9285 ± 0.0322
	(2, 2.5)	0.8109	0.7720 ± 0.0354	0.9520 ± 0.0436
(45, 50)	(-2.5, -2)	0.8399	0.7788 ± 0.0328	0.9272 ± 0.0391
	(-2, -1.5)	0.7588	0.7911 ± 0.0293	1.0426 ± 0.0386
	(-1.5, -1)	0.9077	0.8748 ± 0.0202	0.9638 ± 0.0223
	(-1, -0.5)	0.9162	0.8550 ± 0.0185	0.9331 ± 0.0202
	(-0.5, 0)	0.9053	0.8542 ± 0.0179	0.9435 ± 0.0198
	(0, 0.5)	0.9124	0.8403 ± 0.0182	0.9210 ± 0.0199
	(0.5, 1)	0.9190	0.9063 ± 0.0215	0.9862 ± 0.0234
	(1, 1.5)	0.9083	0.8951 ± 0.0565	0.9854 ± 0.0622
	(1.5, 2)	0.7952	0.8241 ± 0.0271	1.0363 ± 0.0341
	(2, 2.5)	0.8430	0.8113 ± 0.0319	0.9624 ± 0.0378
(50, ∞)	(-2.5, -2)	0.8440	0.7786 ± 0.0393	0.9225 ± 0.0466
	(-2, -1.5)	0.8171	0.8305 ± 0.0310	1.0165 ± 0.0380
	(-1.5, -1)	0.9307	0.8942 ± 0.0259	0.9608 ± 0.0278
	(-1, -0.5)	0.9269	0.8817 ± 0.0174	0.9512 ± 0.0188
	(-0.5, 0)	0.9081	0.8556 ± 0.0193	0.9421 ± 0.0213
	(0, 0.5)	0.9185	0.8318 ± 0.0215	0.9056 ± 0.0235
	(0.5, 1)	0.9194	0.8510 ± 0.0204	0.9255 ± 0.0222
	(1, 1.5)	0.9216	0.9106 ± 0.0198	0.9881 ± 0.0215
	(1.5, 2)	0.8212	0.8076 ± 0.0310	0.9835 ± 0.0377
	(2, 2.5)	0.8689	0.8369 ± 0.0316	0.9632 ± 0.0364
average		0.8424	0.8085 ± 0.0044	0.9542 ± 0.0053

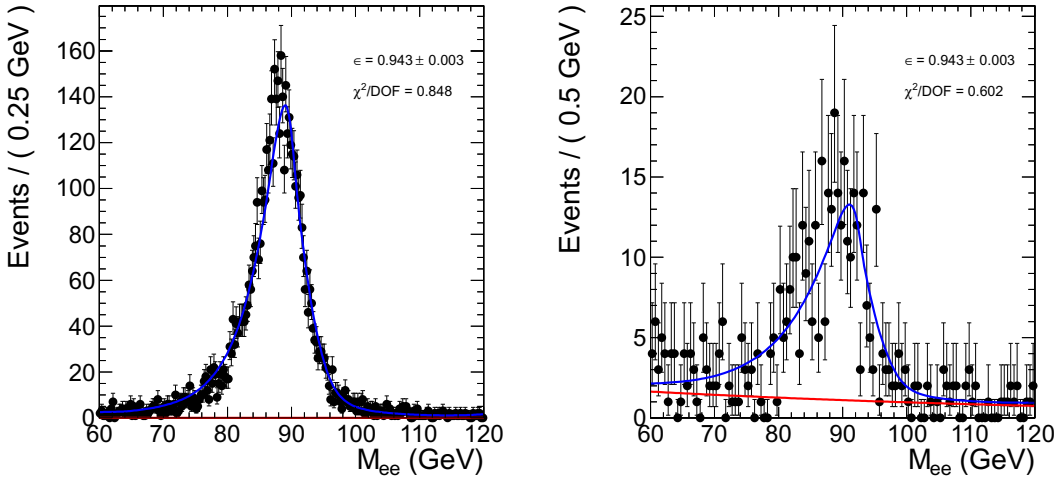
Table 9.6: The fine-binned ID→Trig efficiencies.

P_T Range (GeV)	η Range	ϵ_{MC}	ϵ_{Data}	$\epsilon_{\mathrm{Data}}/\epsilon_{\mathrm{MC}}$
(25, 30)	(-2.5, -2)	0.9692	0.9320 ± 0.0256	0.9616 ± 0.0264
	(-2, -1.5)	0.9692	0.9897 ± 0.0103	1.0212 ± 0.0106
	(-1.5, -1)	0.9633	0.9624 ± 0.0168	0.9991 ± 0.0174
	(-1, -0.5)	0.9658	0.9817 ± 0.0106	1.0164 ± 0.0109
	(-0.5, 0)	0.9860	0.9770 ± 0.0103	0.9908 ± 0.0104
	(0, 0.5)	0.9832	0.9667 ± 0.0126	0.9832 ± 0.0128
	(0.5, 1)	0.9660	0.9880 ± 0.0085	1.0227 ± 0.0088
	(1, 1.5)	0.9704	0.9632 ± 0.0150	0.9926 ± 0.0155
	(1.5, 2)	0.9854	0.9464 ± 0.0218	0.9605 ± 0.0222
	(2, 2.5)	0.9619	0.9435 ± 0.0213	0.9810 ± 0.0221
(30, 35)	(-2.5, -2)	0.9630	0.9634 ± 0.0149	1.0005 ± 0.0155
	(-2, -1.5)	0.9898	0.9732 ± 0.0134	0.9832 ± 0.0136
	(-1.5, -1)	0.9674	0.9873 ± 0.0073	1.0206 ± 0.0076
	(-1, -0.5)	0.9736	0.9541 ± 0.0118	0.9800 ± 0.0122
	(-0.5, 0)	0.9845	0.9967 ± 0.0033	1.0124 ± 0.0033
	(0, 0.5)	0.9852	0.9632 ± 0.0102	0.9776 ± 0.0104
	(0.5, 1)	0.9782	0.9717 ± 0.0094	0.9933 ± 0.0096
	(1, 1.5)	0.9594	0.9694 ± 0.0115	1.0105 ± 0.0120
	(1.5, 2)	0.9800	0.9422 ± 0.0182	0.9615 ± 0.0186
	(2, 2.5)	0.9586	0.9780 ± 0.0110	1.0202 ± 0.0115
(35, 40)	(-2.5, -2)	0.9621	0.9594 ± 0.0143	0.9972 ± 0.0149
	(-2, -1.5)	0.9746	0.9820 ± 0.0090	1.0076 ± 0.0092
	(-1.5, -1)	0.9590	0.9751 ± 0.0083	1.0168 ± 0.0087
	(-1, -0.5)	0.9678	0.9810 ± 0.0063	1.0136 ± 0.0066
	(-0.5, 0)	0.9868	0.9780 ± 0.0069	0.9911 ± 0.0070
	(0, 0.5)	0.9853	0.9778 ± 0.0070	0.9924 ± 0.0071
	(0.5, 1)	0.9726	0.9916 ± 0.0042	1.0196 ± 0.0043
	(1, 1.5)	0.9578	0.9770 ± 0.0081	1.0201 ± 0.0085
	(1.5, 2)	0.9698	0.9721 ± 0.0114	1.0024 ± 0.0117
	(2, 2.5)	0.9642	0.9733 ± 0.0109	1.0094 ± 0.0113
(40, 45)	(-2.5, -2)	0.9789	0.9959 ± 0.0041	1.0173 ± 0.0042
	(-2, -1.5)	0.9730	0.9786 ± 0.0087	1.0058 ± 0.0090
	(-1.5, -1)	0.9497	0.9824 ± 0.0062	1.0345 ± 0.0065
	(-1, -0.5)	0.9626	0.9813 ± 0.0056	1.0195 ± 0.0058
	(-0.5, 0)	0.988	0.9911 ± 0.0040	1.0031 ± 0.0040
	(0, 0.5)	0.9901	0.9838 ± 0.0051	0.9936 ± 0.0052
	(0.5, 1)	0.9643	0.9831 ± 0.0056	1.0195 ± 0.0058
	(1, 1.5)	0.9547	0.9647 ± 0.0091	1.0105 ± 0.0095
	(1.5, 2)	0.9743	0.9719 ± 0.0099	0.9975 ± 0.0102
	(2, 2.5)	0.9734	0.9804 ± 0.0088	1.0072 ± 0.0090
(45, 50)	(-2.5, -2)	0.9673	0.9934 ± 0.0066	1.0270 ± 0.0068
	(-2, -1.5)	0.9725	0.9620 ± 0.0144	0.9891 ± 0.0148
	(-1.5, -1)	0.9474	0.9768 ± 0.0095	1.0311 ± 0.0100
	(-1, -0.5)	0.9524	0.9772 ± 0.0081	1.0261 ± 0.0085
	(-0.5, 0)	0.9860	0.9841 ± 0.0065	0.9980 ± 0.0066
	(0, 0.5)	0.9822	0.9899 ± 0.0051	1.0078 ± 0.0051
	(0.5, 1)	0.9704	0.9917 ± 0.0048	1.0219 ± 0.0050
	(1, 1.5)	0.9508	0.9767 ± 0.0088	1.0272 ± 0.0093
	(1.5, 2)	0.9770	0.9784 ± 0.0108	1.0014 ± 0.0111
	(2, 2.5)	0.9718	0.9929 ± 0.0071	1.0217 ± 0.0073
(50, ∞)	(-2.5, -2)	0.9772	0.9714 ± 0.0165	0.9941 ± 0.0169
	(-2, -1.5)	0.9646	0.9786 ± 0.0124	1.0144 ± 0.0128
	(-1.5, -1)	0.9497	0.9869 ± 0.0076	1.0392 ± 0.0080
	(-1, -0.5)	0.9645	0.9848 ± 0.0068	1.0212 ± 0.0070
	(-0.5, 0)	0.9859	0.9752 ± 0.0088	0.9892 ± 0.0089
	(0, 0.5)	0.9851	0.9829 ± 0.0076	0.9978 ± 0.0077
	(0.5, 1)	0.9630	0.9898 ± 0.0059	1.0279 ± 0.0061
	(1, 1.5)	0.9548	0.9772 ± 0.0102	1.0234 ± 0.0107
	(1.5, 2)	0.9765	0.9744 ± 0.0128	0.9978 ± 0.0131
	(2, 2.5)	0.9595	0.9841 ± 0.0112	1.0257 ± 0.0117
average		0.9710	0.9777 ± 0.0012	0.9542 ± 0.0012



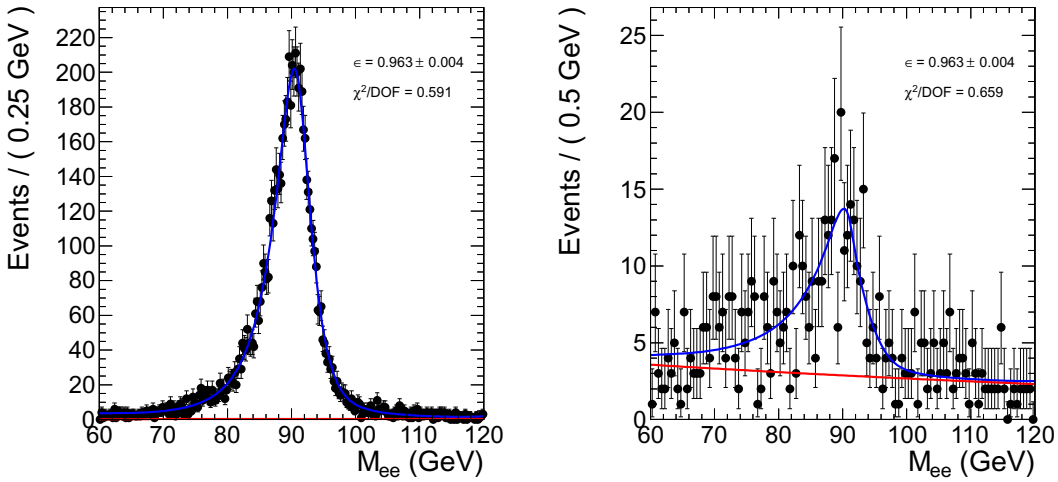
(a) (b)

Figure 9.1: The fits for the Acc→Reco step in the Ecal Barrel. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



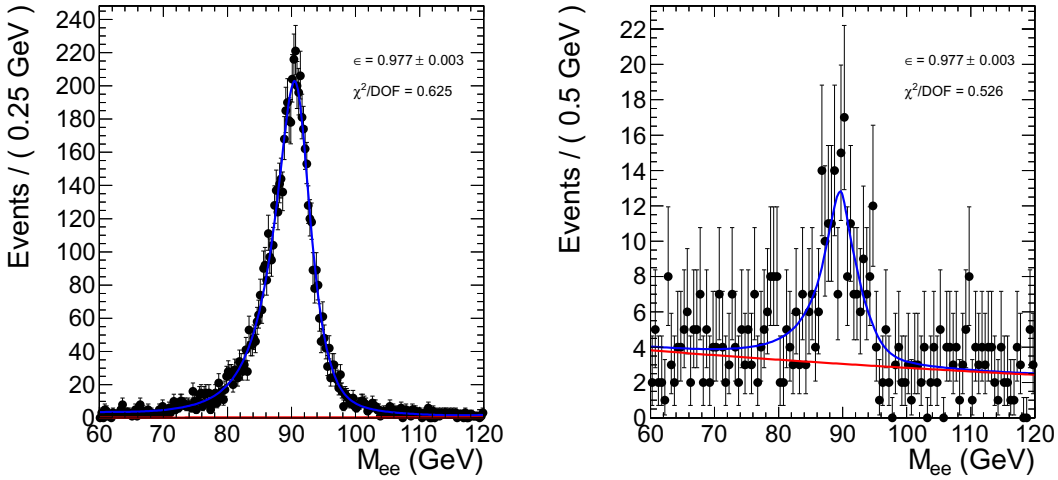
(a) (b)

Figure 9.2: The fits for the Acc→Reco step in the Ecal Endcap. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



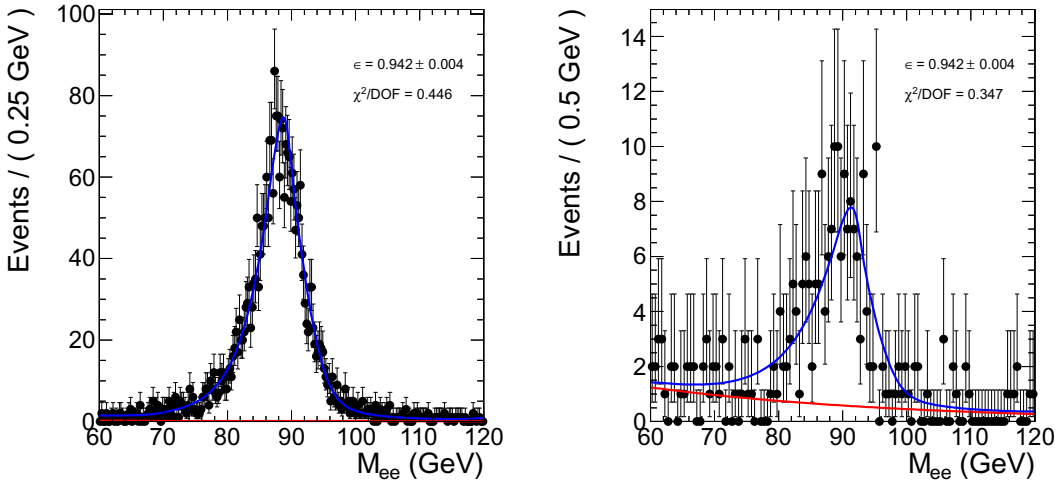
(a) (b)

Figure 9.3: The fits for the Acc→Reco step in the Ecal Barrel for electrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



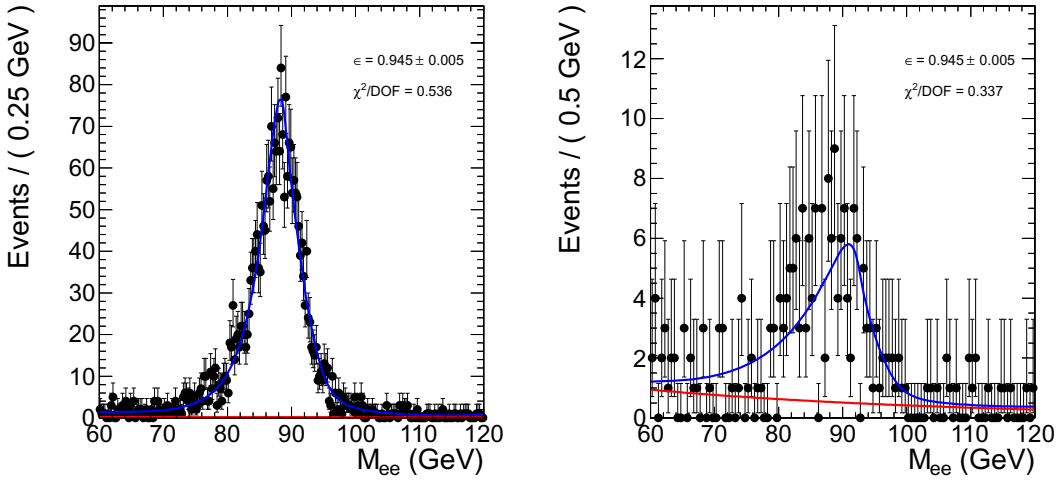
(a) (b)

Figure 9.4: The fits for the Acc→Reco step in the Ecal Barrel for positrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



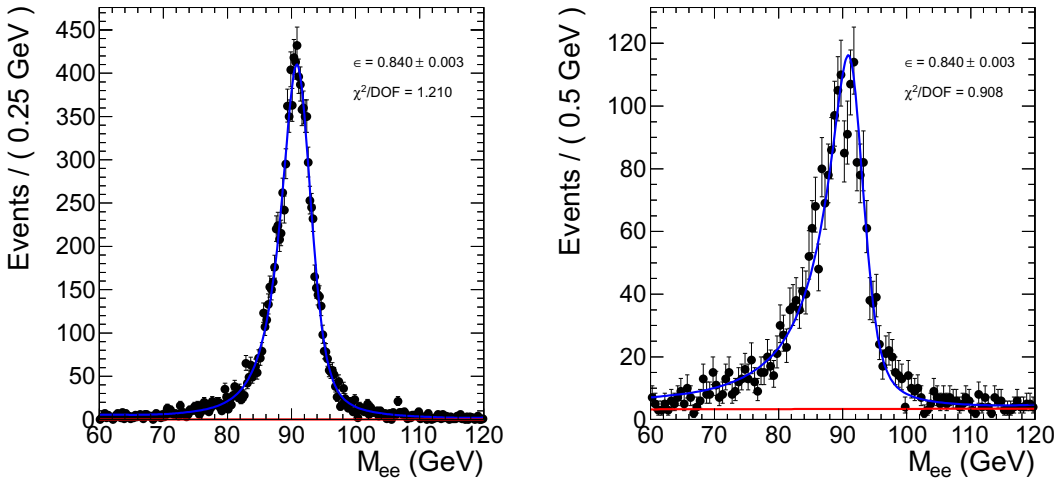
(a) (b)

Figure 9.5: The fits for the Acc→Reco step in the Ecal Endcap for electrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



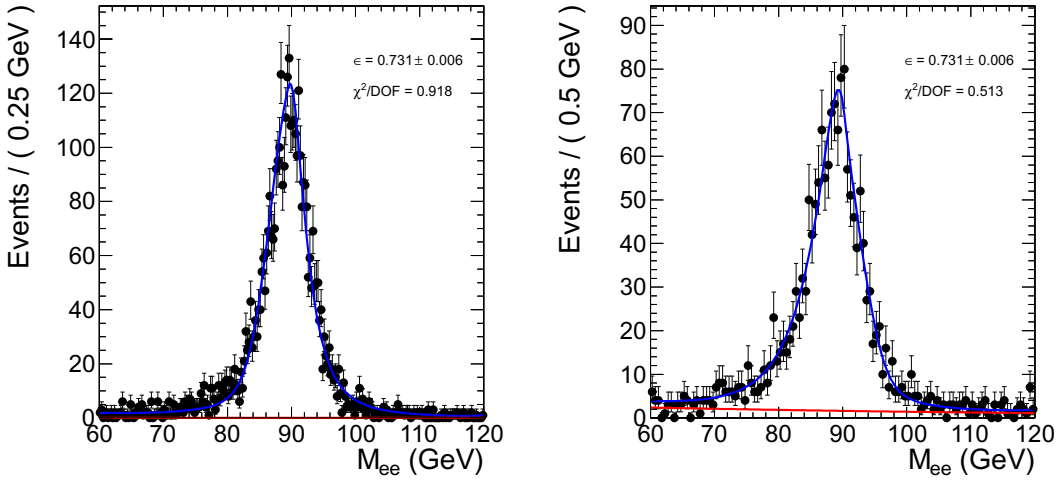
(a) (b)

Figure 9.6: The fits for the Acc→Reco step in the Ecal Endcap for positrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



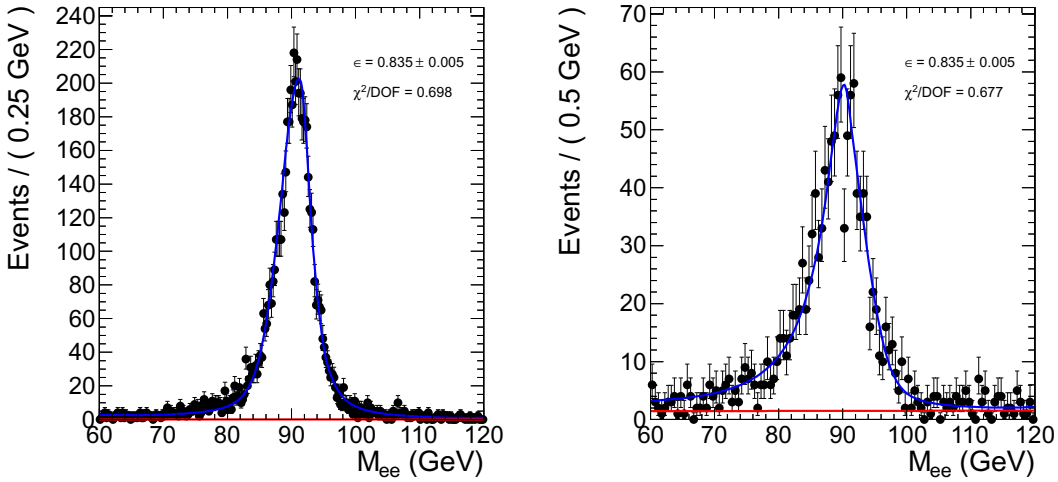
(a) (b)

Figure 9.7: The fits for the Reco→ID step in the Ecal Barrel. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



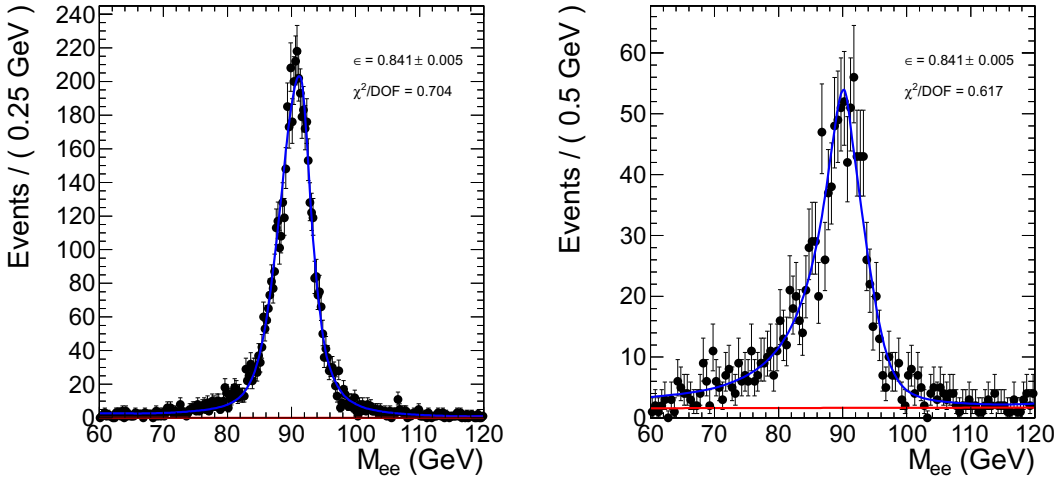
(a) (b)

Figure 9.8: The fits for the Reco→ID step in the Ecal Endcap. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



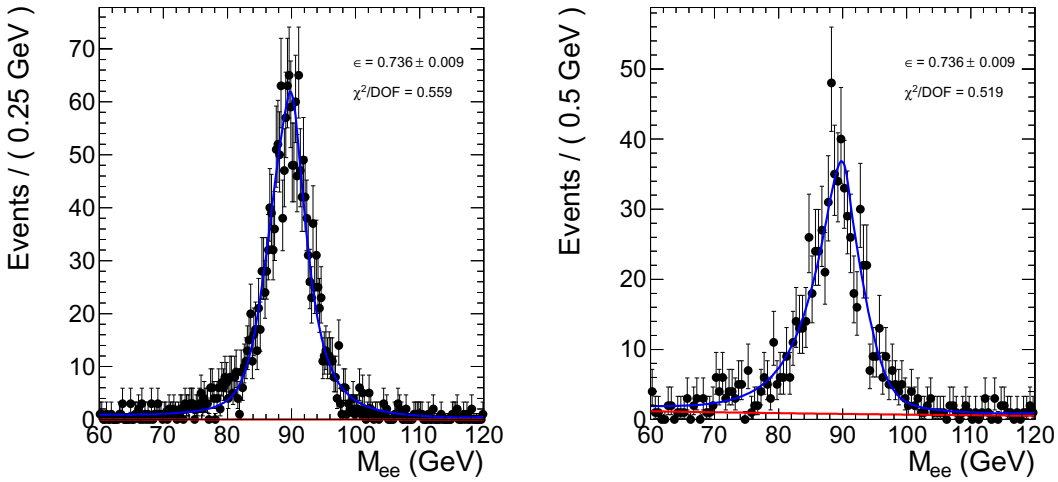
(a) (b)

Figure 9.9: The fits for the Reco→ID step in the Ecal Barrel for electrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



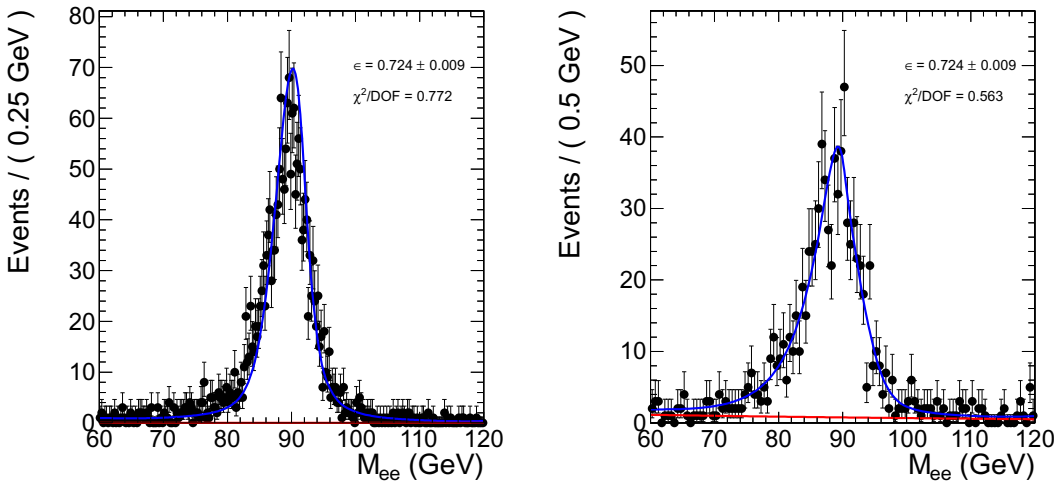
(a) (b)

Figure 9.10: The fits for the Reco→ID step in the Ecal Barrel for positrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



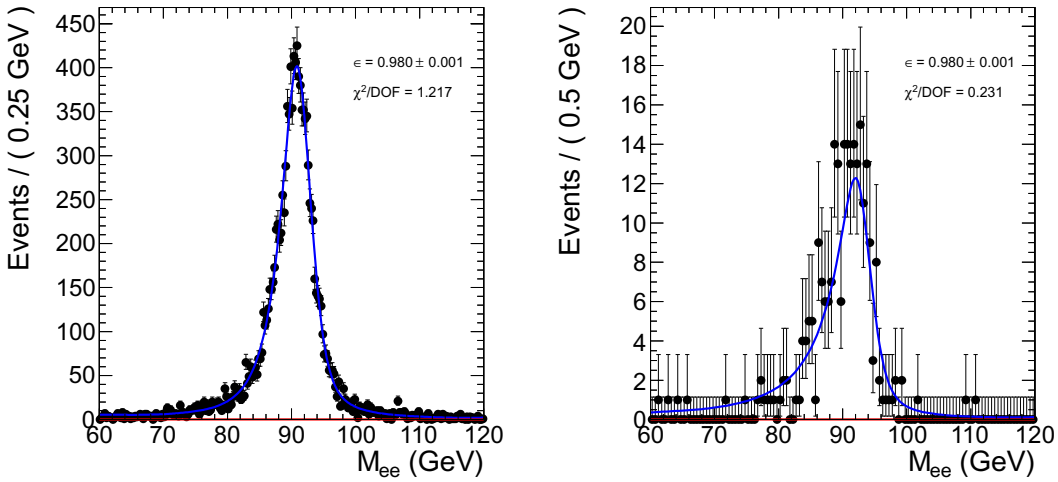
(a) (b)

Figure 9.11: The fits for the Reco→ID step in the Ecal Endcap for electrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



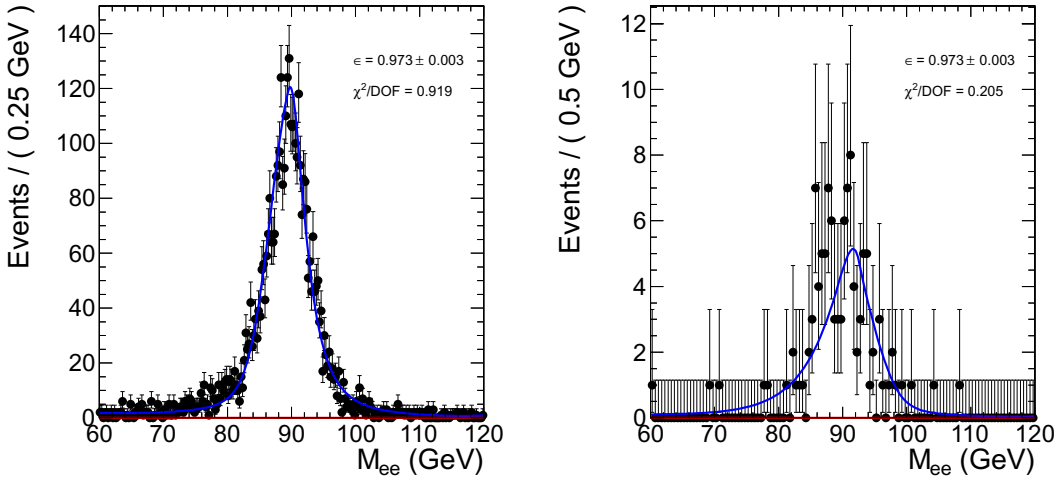
(a) (b)

Figure 9.12: The fits for the Reco→ID step in the Ecal Endcap for positrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



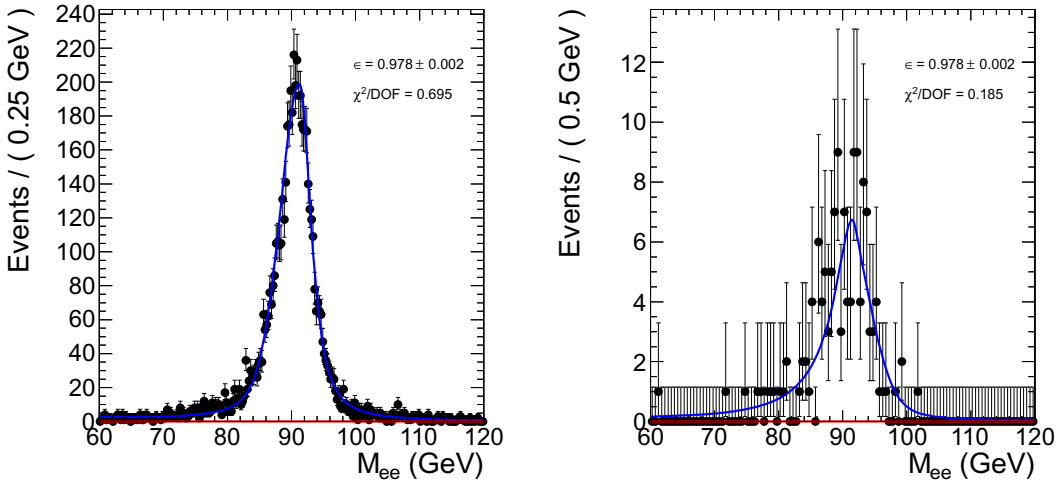
(a) (b)

Figure 9.13: The fits for the ID→Trig step in the Ecal Barrel. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



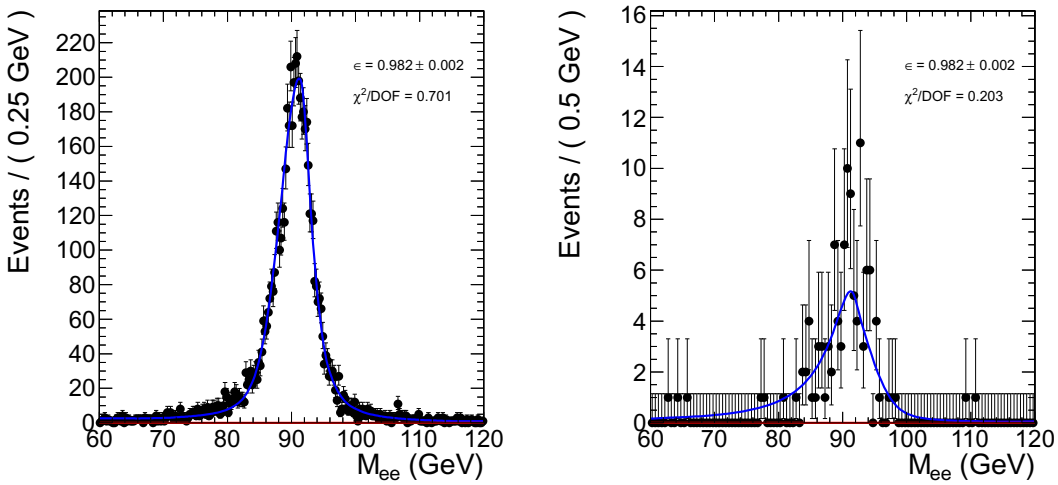
(a) (b)

Figure 9.14: The fits for the ID→Trig step in the Ecal Endcap. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



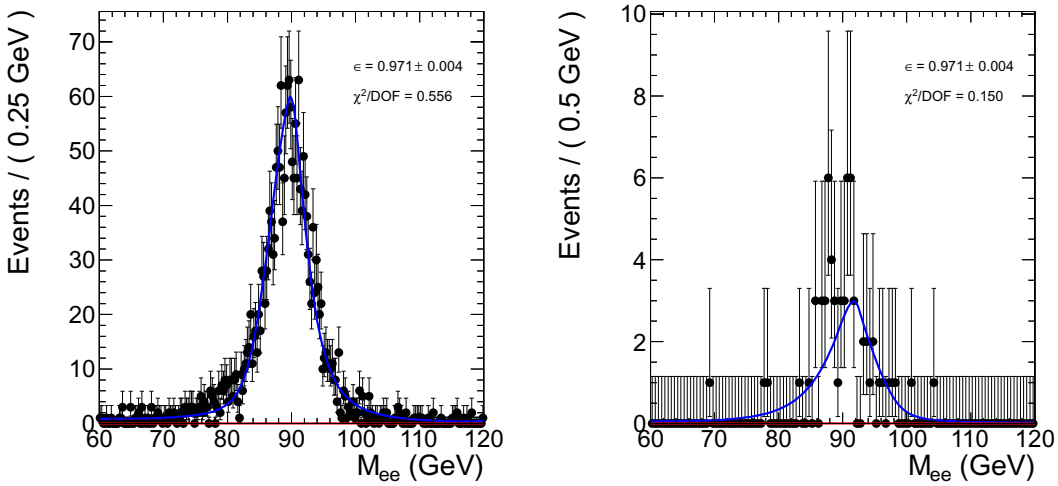
(a) (b)

Figure 9.15: The fits for the ID→Trig step in the Ecal Barrel for electrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



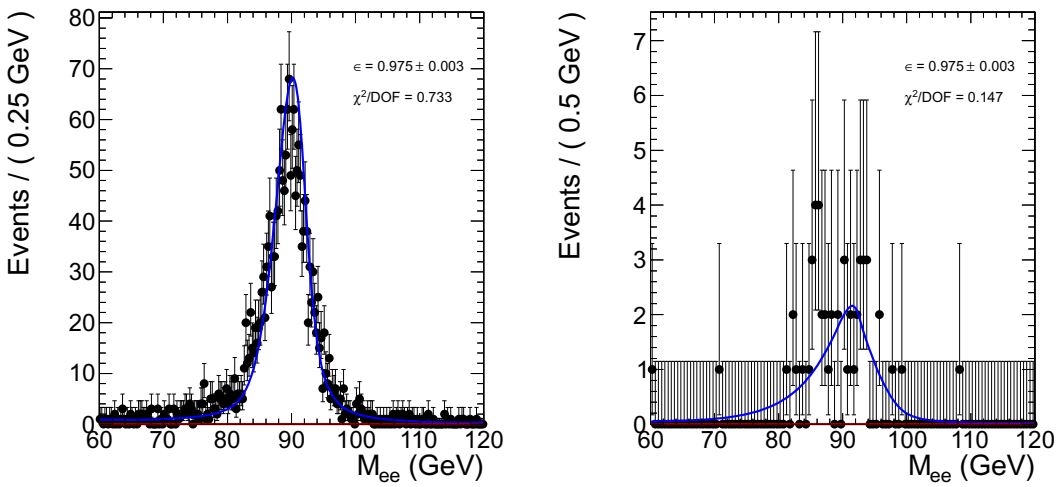
(a) (b)

Figure 9.16: The fits for the ID→Trig step in the Ecal Barrel for positrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.



(a) (b)

Figure 9.17: The fits for the ID→Trig step in the Ecal Endcap for electrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.

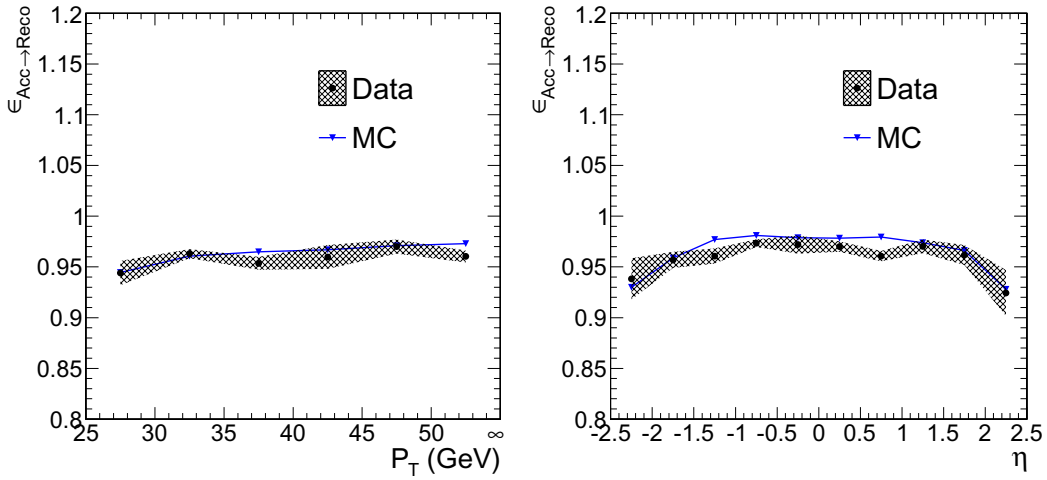


(a) (b)

Figure 9.18: The fits for the ID→Trig step in the Ecal Endcap for positrons. The data is in black, background fit in red, and signal+background fit in blue: (a) passing sample, (b) failing sample.

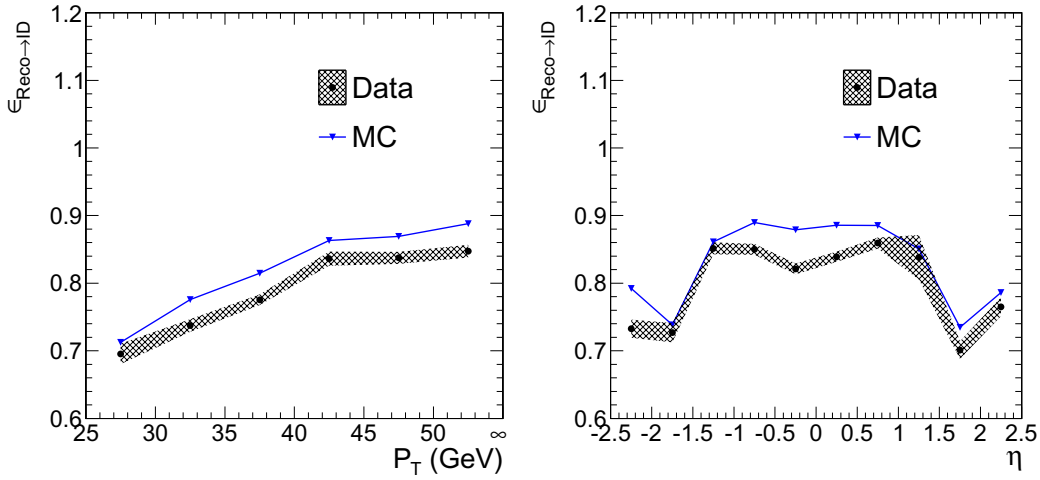
9.5.1 Fine-Binned Efficiency Results Profiled in Terms of P_T and η

The fine-binned efficiency results are profiled in terms of (a) P_T and (b) η in Figures 9.19–9.21 for the Acc→Reco, Reco→ID, and ID→Trig steps, respectively. The data profiles are shown in black. The corresponding fit uncertainties are represented by hatched bands. These are to be compared with the MC profiles, which are represented in blue. The trends in P_T and η observed in data are similar to those of the MC for the Acc→Reco and Reco→ID steps. Both the P_T and η trends observed in data differ from those of the MC for the ID→Trig step, however. The observed differences are not unusual since the trigger performance is very difficult to model reliably and the trigger emulator is known to have limited accuracy [72]. Indeed, the data indicates better performance in both profiles than expected.



(a) (b)

Figure 9.19: The Acc→Reco efficiency profiled in terms of: (a) P_T, (b) η. The data profiles are shown in black. The corresponding errors are represented by hatched bands. The MC profiles are represented in blue.

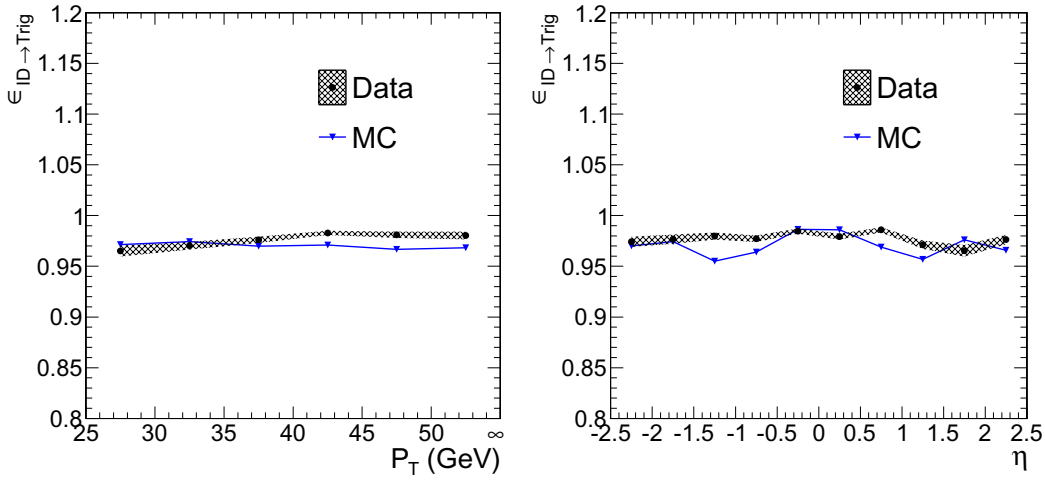


(a) (b)

Figure 9.20: The Reco→ID efficiency profiled in terms of: (a) P_T, (b) η. The data profiles are shown in black. The corresponding errors are represented by hatched bands. The MC profiles are represented in blue.

9.5.2 Two-Binned Efficiency Results Profiled in Terms of Integrated Luminosity

The two-binned efficiencies are measured in blocks of fixed integrated luminosity and profiled to assess whether there is any indication that the performance of the detector fluctuated during data taking. These profiles are provided in Figures 9.22–9.24 for the Acc→Reco, Reco→ID, and ID→Trig steps, respectively. The luminosity blocks chosen correspond to 2.4 pb^{-1}, except the last one which corresponds to 2.3 pb^{-1}. The data is shown in black. The corresponding fit uncertainties are represented by hatched bands. The averages of these measurements are represented by orange lines, and their fits to constants are represented by red lines. The χ^2/DOF values and corresponding probabilities of the fits are printed on the plots. Figure 9.22 (9.23) is consistent with a stable Acc→Reco (Reco→ID) efficiency, having no large deviations present, a $\chi^2/DOF = 1.14$ (1.25), and corresponding probability value $p = 0.308$ (0.225). Thus there is no indication of fluctuations in detector performance impacting electron reconstruction



(a) (b)

Figure 9.21: The ID→Trig efficiency profiled in terms of: (a) P_T, (b) η. The data profiles are shown in black. The corresponding errors are represented by hatched bands. The MC profiles are represented in blue.

or selection. On the other hand, a large positive fluctuation is seen in Figure 9.22 towards the left hand side of the plot. The fit to the data also has a high $\chi^2/DOF = 5.79$ and low corresponding probability value $p = 1.8 \times 10^{-11}$, indicating that the ID→Trig efficiency changed during data taking. This is the expected result since the electron trigger was successively tightened during data taking, as discussed in Subsection 6.1.2.

9.6 Systematic Uncertainties

The general expectation is that the fits return the proper combined systematic and statistical uncertainties, so that no other large fit-related systematic biases need be accounted. Nevertheless, several studies are carried out to confirm this and assess any small, residual bias. In particular, it is assessed whether there is any unaccounted for fit-related systematic bias coming from the signal shape or background shape parameterizations. Another possible source of bias evaluated is the choice of tag, which is

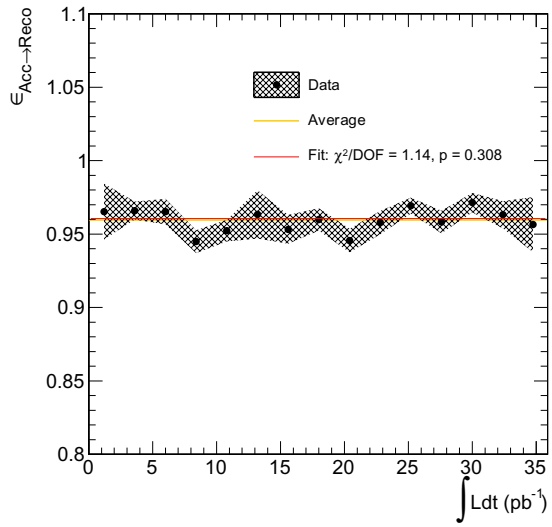


Figure 9.22: The Acc→Reco efficiency profiled in terms of integrated luminosity. The data profile is shown in black. The corresponding errors are represented by a hatched band. The average of the profile is represented by an orange line. The fit of the profile to a constant is represented by a red line. The χ^2/DOF and corresponding probability of the fit are printed on the plot.

also related to the signal shape parameterization. Possible biases arising from the energy scale correction and the H/E cleaning cut for the Acc→Reco step are also evaluated. The conclusion is as expected, and the unaccounted for residual biases related to the fit are found to be quite small. The energy scale bias is also found to be quite small. The H/E cleaning cut is found to be the dominant source of systematic bias.

9.6.1 Signal Shape Systematic Uncertainty

The fits used in the efficiency calculations are verified to be stable and have high goodness of fit. Nevertheless, to assess whether there is any systematic bias arising from the chosen signal parameterization, large numbers of toy MC pseudo-experiments are performed. To generate the toy MC ensembles, f_{signal}, P_{signal}, P_{bkg}, and N are set to the values reported by the fits to the data for both the Tag+Pass and Tag+Fail samples. These are then fit back to the $S + B$ shape whence they were generated, allowing the signal shape to float as in the actual efficiency measurement, while keeping the background shape fixed. The

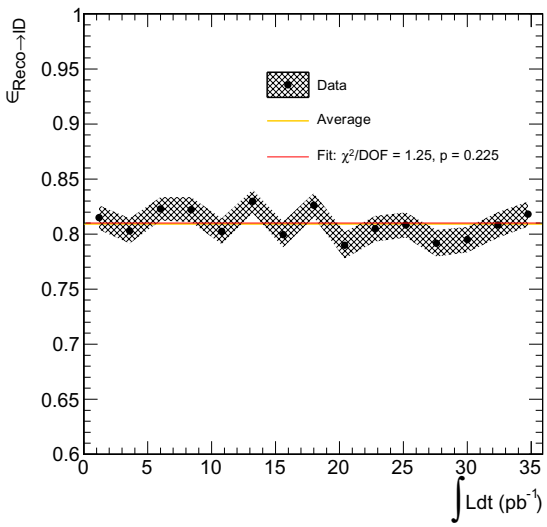


Figure 9.23: The Reco→ID efficiency profiled in terms of integrated luminosity. The data profile is shown in black. The corresponding errors are represented by a hatched band. The average of the profile is represented by an orange line. The fit of the profile to a constant is represented by a red line. The χ^2/DOF and corresponding probability of the fit are printed on the plot.

systematic uncertainty is taken as the absolute difference between the average mean of the efficiency reported back by the fits and the MC truth efficiency: $\Delta\epsilon = |\langle\epsilon_{\rm fit}\rangle - \epsilon_{\rm truth}|$. The resulting systematic biases are found to be negligible. They are listed in Table 9.7. (Recall that the ID→Trig efficiency is determined by counting, implying the uncertainty for this step is identically zero.) In addition, the pull distributions of the signal-yield variables have means and widths close to zero and one, respectively, indicating that there is no large systematic bias. Figure 9.25 shows a histogram of the fractional difference between the observed and true yields in the (a) Tag+Pass and (b) Tag+Fail samples for the Acc→Reco step in the Ecal Barrel, for example. Figure 9.26 plots the pull distributions of the yield variables for the same two cases. The plotted distributions are fit to Gaussians.

9.6.2 Choice of Tag Systematic Uncertainty

The systematic bias arising from the choice of tag is assessed by recalculating the efficiencies for tag criteria of varying tightness. This test is complementary to the study of

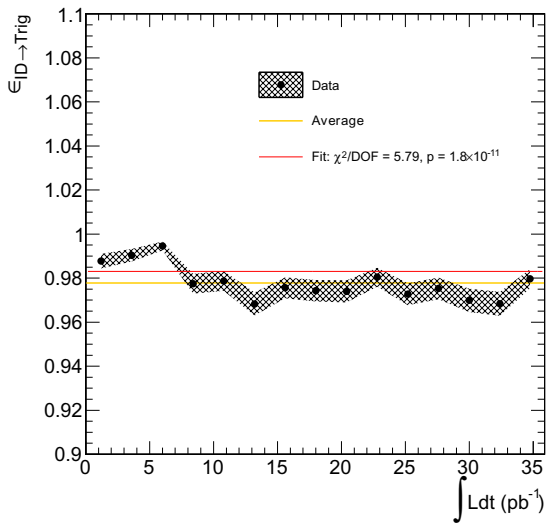
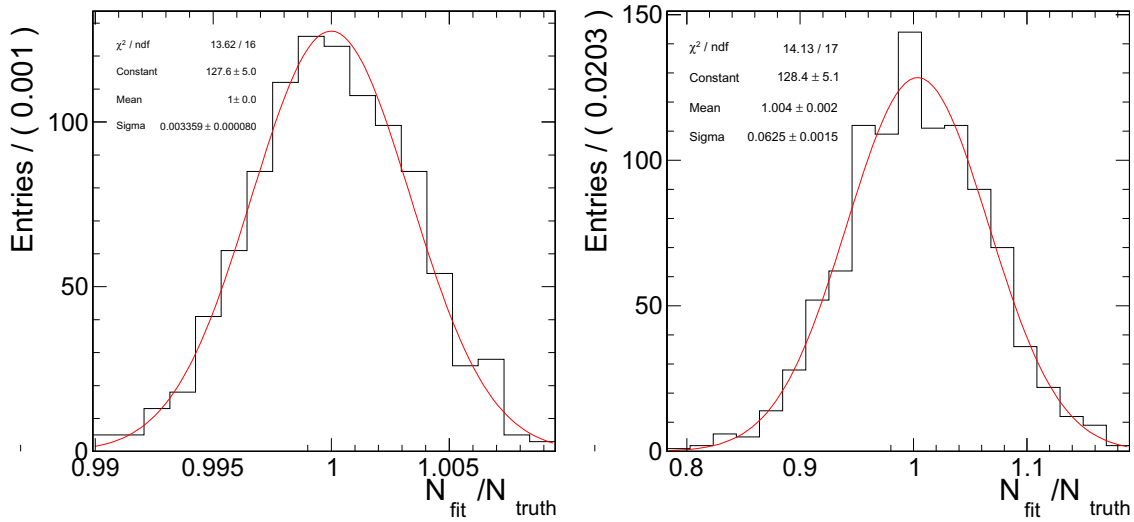


Figure 9.24: The ID→Trig efficiency profiled in terms of integrated luminosity. The data profile is shown in black. The corresponding errors are represented by a hatched band. The average of the profile is represented by an orange line. The fit of the profile to a constant is represented by a red line. The χ^2/DOF and corresponding probability of the fit are printed on the plot.

Table 9.7: The signal parameterization systematic uncertainty for each efficiency step in terms of $\Delta\epsilon/\epsilon$.

Step	$\Delta\epsilon/\epsilon$
Acc→Reco	0.0001
Reco→ID	0.0004
ID→Trig	0

Subsection 9.6.1, as it demonstrates on data that the efficiencies—and hence the signal shapes—don't change as the tag criteria is successively tightened to admit less background. Three tag criteria are considered in order of loosest to tightest: A reconstructed electron passing WP90, WP80, or WP70. The usual HLT matching criterion is also applied to each tag choice. Any difference in the efficiency results falling outside of the range permitted by the quoted systematic errors from the fit is assigned as a systematic uncertainty. Table 9.8 displays the efficiencies using the three different tag criteria along with the assigned systematic uncertainties. Results for the WP90 tag are



(a) (b)

Figure 9.25: The fractional difference between the observed and true yields for the Acc→Reco step in the Ecal Barrel (black) fit to Gaussians (red): (a) passing sample, (b) failing sample.

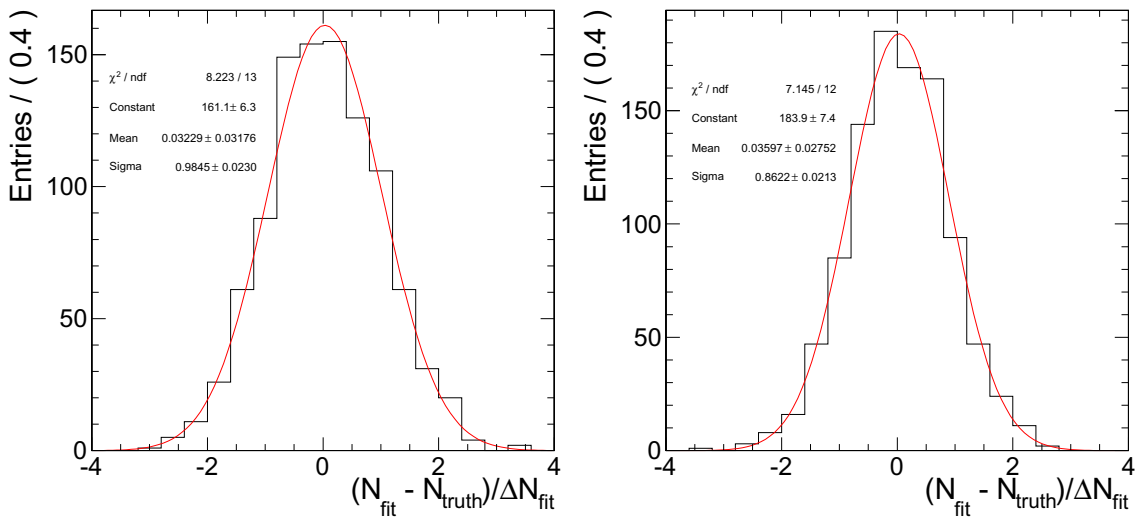
not given for the latter two steps since WP90 is not a reasonable tag choice for them.[5]

The only significant systematic bias is present in the Acc→Reco step at the level of 0.17%.

Table 9.8: The two-binned ⟨EB⊕EE⟩ electron efficiencies for tags of varying tightnesses and the corresponding choice of tag systematic uncertainty.

Step	$\epsilon \pm \Delta\epsilon$(fit syst)			$\Delta\epsilon/\epsilon$
	WP90 Tag	WP80 Tag	WP70 Tag	
Acc→Reco	0.9580 ± 0.0024	0.9627 ± 0.0019	0.9621 ± 0.0037	0.0017
Reco→ID	-	0.8092 ± 0.0000	0.8090 ± 0.0000	0.0002
ID→Trig	-	0.9780 ± 0.0000	0.9779 ± 0.0000	0.0001

[5]For the TP formalism to apply, the chosen tag criteria must be at least as tight as the passing probe criteria under study [78].



(a) (b)

Figure 9.26: The yield variable pull distributions of the pseudo-experiments for the Acc→Reco step in the Ecal Barrel (black) fit to Gaussians (red): (a) passing sample, (b) failing sample.

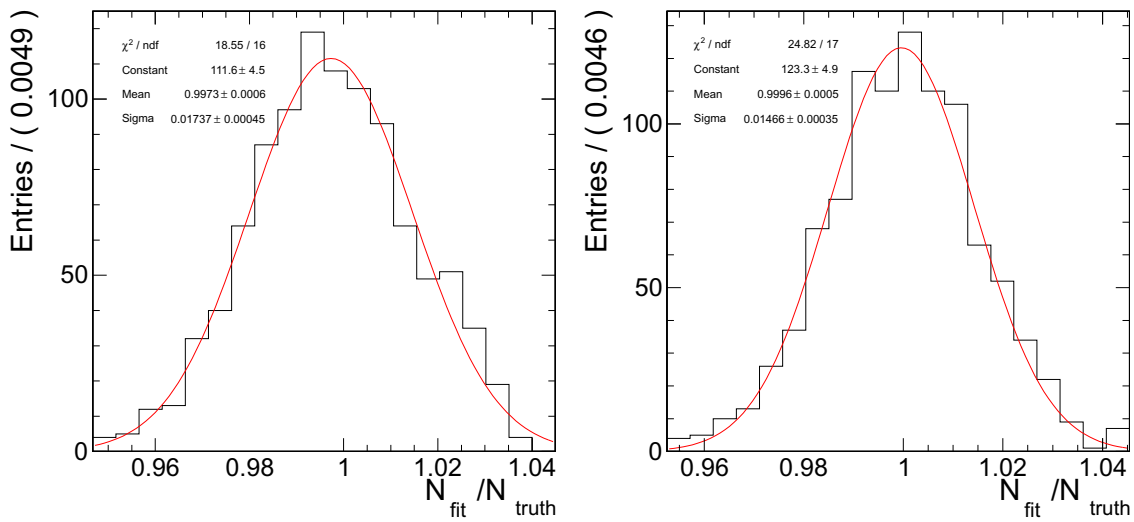
9.6.3 Background Shape Systematic Uncertainty

To estimate the systematic bias due to the background parameterization, large numbers of toy MC ensembles are generated with background PDFs of the form $1/M^{\alpha}$, where $\alpha = 5$.[6] The ensembles are generated by setting f_{signal}, P_{signal}, and N to the values reported by the fits to the data for both the Tag+Pass and Tag+Fail samples. These are then fit back to the sum of the signal shape whence they were generated and two different floating background shape hypotheses: an exponential decay and a power law (except for the Tag+Pass sample in the Reco→ID step since the background is fixed in this case). The systematic uncertainty is quoted as the absolute difference in the average mean reported by the fits between the two hypotheses: $\Delta\epsilon = |\langle\epsilon_{\text{exponential}}\rangle - \langle\epsilon_{\text{power law}}\rangle|$. The systematic biases are found to be small, and are listed in Table 9.9. (Recall that the ID→Trig efficiency is determined by counting, implying the systematic uncertainty is identically zero in this case.) In addition, the pull distributions of the signal yield variables have means and widths close to zero and one, respectively, indicating that there is no large systematic bias. Figure 9.27 histograms the fractional difference between the observed and true yields in the Tag+Fail sample for the Reco→ID step in the Ecal Endcap assuming an (a) exponential or (b) power law background shape, for example. Figure 9.28 plots the yield variable pull distributions of the pseudo-experiments for the same two cases. The plotted distributions are fit to Gaussians. The only significant systematic bias is present in the Acc→Reco step at the level of 0.13%.

Table 9.9: The background parameterization systematic uncertainty for each efficiency step in terms of $\Delta\epsilon/\epsilon$.

Step	$\Delta\epsilon/\epsilon$
Acc→Reco	0.0013
Reco→ID	0.0005
ID→Trig	0

[6] This is the expected form of the dominant, underlying multijet background.



(a) (b)

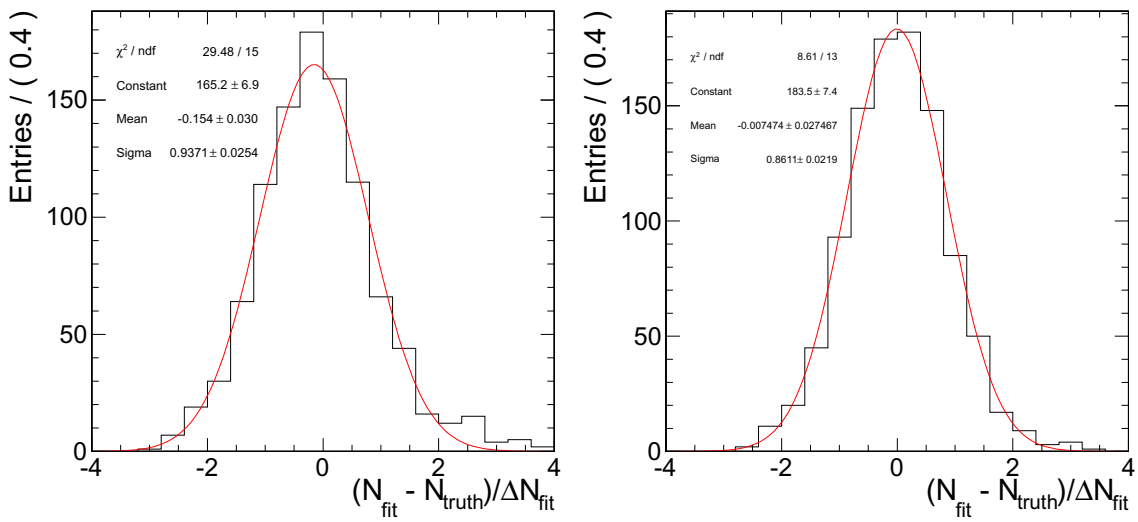
Figure 9.27: The fractional difference between the observed and true yields in the failing sample for the Reco→ID step in the Ecal Endcap fit to Gaussians, assuming the background shape is an: (a) exponential, (b) power law.

9.6.4 Energy Scale Systematic Uncertainty

To estimate the systematic uncertainty due to the energy scale, the electron energies in the MC are shifted by the uncertainty on the scale in both the positive and negative directions. The maximum difference among the three extracted efficiency values—nominal, positive shift, and negative shift—is assigned as the systematic uncertainty. The energy scale systematic uncertainties are found to be negligible, and are listed in Table 9.10.

Table 9.10: The energy scale systematic uncertainty for each efficiency step in terms of $\Delta\epsilon/\epsilon$.

Step	$\Delta\epsilon/\epsilon$
Acc→Reco	0
Reco→ID	0.0002
ID→Trig	0



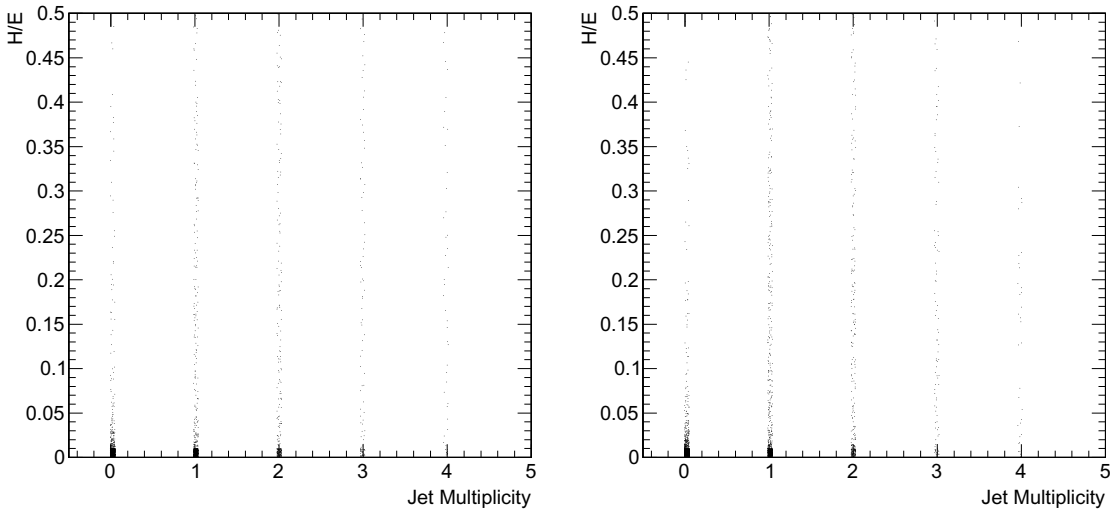
(a) (b)

Figure 9.28: The yield variable pull distributions of the pseudo-experiments in the failing sample for the Reco→ID step in the Ecal Endcap fit to Gaussians, assuming the background shape is an: (a) exponential, (b) power law.

9.6.5 H/E Cleaning Cut Systematic Uncertainty

As mentioned in Section 9.3, the Tag+Fail sample in the Acc→Reco step has a large contamination from QCD multijets. To enable an accurate signal estimation in this step, a cleaning cut of $H/E < 0.15$ is applied to the probe to reduce the background. To estimate the systematic bias arising from this cut, an orthogonal cleaning criterion is applied instead of the H/E cut. The efficiency is evaluated with only this orthogonal criterion, and then again after also applying the H/E cut. The chosen orthogonal criterion is that there be zero $JetMultiplicity$ in the event. The jets reconstructed in the Hcal must have $P_T > 20$ GeV and electromagnetic energy fractions less than 0.1 to contribute to the $JetMultiplicity$. Fig 9.29 shows scatter plots of H/E for the supercluster probe versus the event $JetMultiplicity$ for probes in the Ecal (a) Barrel and (b) Endcap, demonstrating that these variables are not strongly correlated. The full difference in the efficiency before and after the application of the H/E cut is assigned as the systematic

uncertainty (0.42%). Table 9.11 shows the result of this study.



(a) (b)

Figure 9.29: Scatter plots of supercluster H/E vs. event $JetMultiplicity$ in the Ecal: (a) Barrel, (b) Endcap (b).

Table 9.11: Results of the H/E cleaning cut systematic uncertainty study.

Probe Cuts	ϵ	$\Delta\epsilon/\epsilon$
$JetMultiplicity < 1$	0.9781	0.0042
$JetMultiplicity < 1$ and $H/E < 0.15$	0.9822	

9.6.6 Systematic Uncertainties Summary

A thorough series of tests has been conducted to assess whether there is any unaccounted for bias present in the efficiency calculation due to the signal shape parameterization, background parameterization, choice of tag, energy scale, or H/E cleaning cut. The conclusion, in line with expectations, is that the combined statistical and systematic uncertainties returned by the fits largely cover the fit-related systematic uncertainties: The remaining residual fit-related systematic uncertainties are all quite small. The choice

of tag and energy scale systematic uncertainties are also quite small. The dominant systematic bias is due to the H/E cleaning cut in the Acc→Reco step. The total non-fit systematic uncertainties are listed in Table 9.12.

Table 9.12: The summed non-fit systematic uncertainty for each efficiency step in terms of $\Delta\epsilon/\epsilon$.

Step	$\Delta\epsilon/\epsilon$
Acc→Reco	0.0047
Reco→ID	0.0007
ID→Trig	0.0001

9.7 Summary

The first systematic uncertainty limited electron efficiency measurements performed at CMS have been presented. A TP method was used that relied on unbinned maximum log-likelihood fits to the data. Several cross-checks were performed to ensure the robustness of the results, and all of the associated systematic uncertainty studies have been detailed. Comparisons between the data and MC indicated agreement in both the P_T and η trends, offering a strong confirmation of the method, the MC, and the CMS detector simulation.

Finally, taking the two-binned electron efficiency values from the fourth rows of Tables 9.1–9.3, and adding in quadrature the errors from these tables with the corresponding entries of Table 9.12, determines ϵ_Z (Equation 9.1) as:

$$\epsilon_Z = 0.6065 \pm 0.0078(\text{syst}) \text{ (two-binned efficiencies)} \tag{9.7}$$

where all errors on the efficiencies have been propagated as systematic uncertainties. Similarly, the fine-binned efficiency values from the last rows of Tables 9.4–9.6 give:

$$\epsilon_Z = 0.6041 \pm 0.0094(\text{syst}) \text{ (fine-binned efficiencies)} \tag{9.8}$$

These values will be used to determine the $Z \to ee$ cross section in Chapter 11.

Chapter 10

Theoretical Predictions

10.1 Introduction

The efficiency and background corrected signal yields measured in the previous chapters are used to perform two separate measurements: The inclusive $Z \to ee$ production cross section measurement, which is concluded in Chapter 11, and a measurement of the absolute luminosity presented in Chapter 12. The theoretical predictions and accompanying uncertainties impacting both measurements are estimated in this chapter. In particular, the cross section measurement depends on the acceptance A (Subsection 6.1.1), while the luminosity measurement depends on the effective theoretical cross section $A\sigma\mathcal{B}$ (Equation 12.1).

10.2 Acceptance Calculation

The prediction for the fiducial and kinematic acceptance is needed for the cross section measurement and the $Z \to ee$ decay based luminosity measurement, and represents an important source of systematic uncertainty on these measurements. It is calculated using the baseline signal MC sample (Section 5.2) generated with the NLO MC generator Powheg [62] interfaced with Pythia [63] for the parton showering and the PDF set

CT10 [65]. It has the value:

$$A = 0.3876 \pm 0.0062 \tag{10.1}$$

The error on the acceptance is calculated in the following subsection.

10.2.1 Uncertainties on the Acceptance

Uncertainties are quoted with respect to the Powheg baseline wherever applicable. It has been verified that this baseline is in agreement with the more complex integrator tool ResBos [79, 80, 81, 82, 83, 84]. The sources of uncertainty on the acceptance are:

- PDF Uncertainties.
- Effects from initial state radiation modeling and NNLO QCD.
- Effects from missing electroweak corrections.
- Effects from final state radiation modeling.
- NNLO scale uncertainties.

The studies [85] estimating these uncertainties are described in the following subsubsections.

PDF Uncertainty

PDFs are extracted from global fits to data from deep inelastic scattering, Drell-Yan, and jet processes. Error PDFs are packaged with each PDF set to describe the uncertainty of the proton structure. Three PDF sets are used: MSTW08 [86], CT10 [65], and NNPDF2.0 [87]. The uncertainty calculation follows the prescription of the PDF4LHC working group [88]. Positive and negative errors are determined for each set by calculating the spread in acceptance among the error PDFs. The error on α_s is considered simultaneously in this calculation to obtain the PDF+α_s uncertainties. The PDF+α_s uncertainties on the acceptance A are summarized in Table 10.1. Rows two to four represent the uncertainty (68% confidence level) within the specified PDF set. The value

quoted in row five corresponds to half of the maximum difference between the central values of any pair of the three listed sets. The final quoted systematic uncertainty in the bottom row considers half of the maximum difference between extreme values (central values plus positive or negative uncertainties), again for any pair of the three sets, plus remaining α_s uncertainties.

Table 10.1: Systematic uncertainties from PDF assumptions on the calculated acceptance.

PDF Set	$\Delta A/A$
CT10	0.007
MSTW	0.004
NNPDF	0.007
ΔSets	0.002
Systematic Uncertainty	0.009

Effects from Initial State Radiation Modeling and NNLO QCD (ISR+NNLO)

The baseline signal sample uses the Powheg Monte-Carlo generator interfaced with Pythia for the parton showering. Thus the parton showering is only accurate to leading-log order for soft, non-perturbative QCD and to NLO for perturbative QCD. ResBos implements the resummation procedure so that it is accurate to next-to-next-leading-log-order for non-perturbative QCD. This generator is also accurate to NNLO for perturbative QCD. ResBos does not provide the kinematics of the scattering or radiating partons, and so the comparison of it to the baseline is done at generator-level to quantify the systematic uncertainty.

Effects from Missing Electroweak Corrections (EWK)

At the energy scale of weak boson production, $\mathcal{O}(\alpha)$ corrections are of the same scale as $\mathcal{O}(\alpha_s^2)$ effects. Powheg interfaced with Pythia generates QED ISR and FSR using a parton shower approximation, but neglects one loop virtual corrections. Thus NLO electroweak corrections are only partially accounted for in the baseline sample. The HORACE generator implements Z production and subsequent decay into leptons accurate to NLO in

electroweak theory. It also uses a parton shower method to account for FSR beyond single photon emission. Thus HORACE is used with its full suite of corrections enabled and is compared to HORACE with only FSR enabled to estimate the effects of electroweak corrections missing in the baseline sample.

Effects from Final State Radiation Modeling (FSR)

FSR is described by Pythia in the baseline sample, and also by HORACE. Thus this systematic uncertainty is taken as the difference in acceptance between the baseline sample and the HORACE sample with only FSR enabled. The effects of QED ISR are negligible, as determined by comparing the baseline signal sample with QED ISR either enabled or disabled. Hence no systematic uncertainty is assigned for the QED ISR modeling.

NNLO Scale Uncertainties (>NNLO)

Cross sections calculated to a fixed order depend on the renormalization and factorization scales, μ_{R} and μ_{F}, respectively. The uncertainty on the cross section due to higher perturbative orders is estimated by varying these scales. This is done using the integrator FEWZ, which calculates cross sections accurate to NNLO in QCD, and letting $\mu \equiv \mu_R = \mu_F$ vary in the range $\mu \in (M_Z/2, 2M_Z)$. For a given value of μ, the acceptance is then the ratio of the cross section calculated for the restricted fiducial and kinematic phase space defined in Subsection 6.1.1 to the cross section calculated for the full phase space. The systematic uncertainty is assigned as half of the maximum variation in the acceptance from the default value with $\mu = M_Z$.

Results

The uncertainties on the acceptance are summarized in Table 10.2 [85]. The top three contributions are approximately equal; they are PDF, ISR+NNLO, and EWK. Aside from the 4.0% uncertainty on the luminosity, the 1.6% theoretical error on the acceptance is the largest uncertainty on the cross section measurement.

Table 10.2: Systematic uncertainties on the calculated acceptance.

Type	$\Delta A/A$
PDF	0.009
ISR+NNLO	0.008
EWK	0.008
FSR	0.005
>NNLO	0.004
Total	0.016

10.3 Theoretical $Z \to ee$ Cross Section Calculation

The prediction for the theoretical cross section is needed for the $Z \to ee$ decay based luminosity measurement presented in Chapter 12 and for comparison of the cross section measurement to the theory. Its associated uncertainties comprise the dominant error on the $Z \to ee$ decay based luminosity measurement. The prediction is calculated at NNLO using the program FEWZ [67, 68] with the MSTW08 [86] PDF set [89, 90]. The cross section pertaining to the invariant mass window $M_{ee} \in (60, 120)$ GeV is:

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee)_{\text{NNLO Theory}} = 972 \pm 26 \text{ pb} \tag{10.2}$$

The uncertainties, at 68% confidence level, include contributions from the PDF and strong coupling α_{s} [91, 90], the choice of heavy quark masses (charm and bottom quarks) [92], as well as neglected higher-order corrections beyond NNLO, calculated by allowing the renormalization and factorization scales to vary in a similar way to that described in Subsubsection 10.2.1.

The 3.1% error on the effective cross section $A\sigma\mathcal{B}$ —obtained by adding the fractional errors on each factor in quadrature—dominates the uncertainty of the $Z \to ee$ decay based luminosity measurement.

10.4 Summary

The theoretical predictions for the acceptance and cross section, including central values and derived systematic uncertainties, have been calculated in this chapter using the most modern tools currently available. Sections 10.2 and 10.3 described the calculations for the acceptance and cross section, respectively, while Equations 10.1 and 10.2 provided the predicted values. The 1.6% systematic uncertainty on the acceptance is the second leading error on the cross section measurement, following the 4.0% luminosity uncertainty. The 3.1% error on the effective cross section was obtained by adding the uncertainties on the acceptance and cross section in quadrature, and constitutes the dominant uncertainty of the $Z \rightarrow ee$ decay based luminosity measurement, as will be discussed in Chapter 12.

Chapter 11

Inclusive $Z \to ee$ Production Cross Section Determination

11.1 Introduction

The inclusive $Z \to ee$ production cross section is determined in this chapter in a few different ways. The reported values pertain to the invariant mass range $M_{ee} \in (60, 120)$ GeV, and are corrected for the fiducial and kinematic acceptance, but not for γ^* exchange. The cross section is given by the well-known expression (Equation 1.7):

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee) = \frac{N_{\mathrm{S}}}{A\epsilon_Z \int \mathcal{L}dt} \tag{11.1}$$

where N_{S} is the signal yield, A is the signal acceptance, ϵ_Z is the signal efficiency, and $\int \mathcal{L}dt$ is the integrated luminosity.

11.2 Cross Section Determination Using Counted Signal Event Yield and Tag and Probe Electron Efficiencies

All quantities on the right hand side of Equation 11.1 have been measured or calculated in the previous chapters. Their values are listed in Table 11.1. The estimate for the signal yield is taken from Equation 7.6. The acceptance value is taken from Equation 10.1. The

signal efficiency is taken from Equation 9.7, and was determined from the two-binned electron efficiencies. The integrated luminosity is taken from Equation 3.26. Its contribution to the systematic uncertainty on the cross section is denoted "lumi" and is quoted separately. Putting these values into the formula for the cross section yields:

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee) = 997 \pm 11(\text{stat}) \pm 21(\text{syst}) \pm 40(\text{lumi})\ \text{pb} \tag{11.2}$$

(two-binned efficiencies/counted signal yield)

The measured cross section agrees with the NNLO theoretical prediction of 972 ± 26 pb (Equation 10.2).

Table 11.1: The values of the variables used in the cross section determination.

Variable	Value	Description
N_{S}	8411 ± 92 (stat) ± 12 (syst)	Counted signal event yield
A	0.3876 ± 0.0062 (syst)	Kinematic and fiducial acceptance
ϵ_Z	0.6065 ± 0.0078 (syst)	$Z \to ee$ signal efficiency (two-binned)
$\int \mathcal{L} dt$	35.9 ± 1.4 pb (lumi)	Integrated luminosity

Alternatively, the fine-binned signal efficiency from Equation 9.8 yields:

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee) = 1001 \pm 11(\text{stat}) \pm 22(\text{syst}) \pm 40(\text{lumi})\ \text{pb} \tag{11.3}$$

(fine-binned efficiencies/counted signal yield)

This value agrees with the production cross section determined using the two-binned electron efficiencies, and also with the NNLO theoretical prediction.

11.3 Simultaneous Fit of the Selected Events to Extract Electron Efficiencies and the $Z \to ee$ Production Cross Section

A different approach to determine the inclusive $Z \to ee$ production cross section is to extract the signal yield and electron efficiencies directly from a single extended

simultaneous unbinned maximum log-likelihood fit. The advantage of this technique is that it accounts for the correlations between the signal yield and efficiencies within the same dielectron sample, minimizing the uncertainty on the cross section.

In this scheme, two mutually exclusive samples of events selected within the invariant mass window $M_{ee} \in (60, 120)$ GeV and the fiducial and kinematic acceptance (Subsection 6.1.1) are considered:

- Tag+Pass: At least one candidate satisfies the tag criteria of Section 9.3, while the other passes the ID. (These criteria are equivalent to the $Z \to ee$ selection described in Chapter 6).
- Tag+Fail: One candidate satisfies the tag criteria of Section 9.3, while the other fails them either at reconstruction or ID level.[1]

The mass spectrum is fit to the signal plus background parameterization developed in Chapter 8 to determine the signal contribution in each sample. In the Tag+Fail sample, the background is modeled by a decaying exponential. In the Tag+Pass case, the nearly vanishing background is fixed to the value reported by the template method in Section 7.5 ($N_{\mathrm{B}} = 30 \pm 12$), and its spectrum is assumed to be flat. The $\bar{M}$ and σ parameters of the modified Crystal Ball function are allowed to float in both the Tag+Pass and Tag+Fail samples. The cross section and electron efficiencies enter as explicit floating parameters. The floating parameters of the fit are listed below:

- $N^{\mathrm{pass}}_{\mathrm{signal}}$: total number of signal Tag+Pass events.
- $N^{\mathrm{fail}}_{\mathrm{signal}}$: total number of signal Tag+Fail events.
- $N^{\mathrm{fail}}_{\mathrm{bkg}}$: total number of background Tag+Fail events.
- $P^{\mathrm{pass}}_{\mathrm{signal}}(M_{ee})$: signal PDF in the Tag+Pass sample that is floating as described above.
- $P^{\mathrm{fail}}_{\mathrm{bignal}}(M_{ee})$: signal PDF in the Tag+Fail sample that is floating as described above.

[1] Superclusters have the $H/E < 0.15$ cleaning cut applied, as in the Tag and Probe efficiency measurements (Subsection 9.3.1).

- $P^{\text{fail}}_{\text{bkg}}(M_{ee}) \propto e^{-\chi^{\text{fail}} M_{ee}}$: background PDF in the Tag+Fail sample with the exponential shape parameter χ^{fail} floating.
- $\epsilon_{\text{Acc}\to\text{ID}}$: the Acc→ID electron efficiency.
- $\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee)$: the inclusive Z→ee production cross section.

The Acc→ID electron efficiency and production cross section are extracted by the fit using the following relations:

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee) = \frac{N^{\text{pass}}_{\text{signal}}}{\int \mathcal{L}dt \times A\epsilon^2_{\text{Acc}\to\text{ID}}(1 - (1 - \epsilon_{\text{ID}\to\text{Trig}})^2)} \tag{11.4}$$

$$\sigma(pp \to \ Z + X) \times \mathcal{B}(Z \to ee) = \frac{N^{\text{fail}}_{\text{signal}}}{\int \mathcal{L}dt \times A2\epsilon_{\text{Acc}\to\text{ID}}(1 - \epsilon_{\text{Acc}\to\text{ID}})\epsilon_{\text{ID}\to\text{Trig}}} \tag{11.5}$$

$$N^{\text{pass}}(M_{ee}) = N^{\text{pass}}_{\text{signal}} P^{\text{pass}}_{\text{signal}}(M_{ee}) + N^{\text{pass}}_{\text{bkg}} P^{\text{pass}}_{\text{bkg}}(M_{ee}) \tag{11.6}$$

$$N^{\text{fail}}(M_{ee}) = N^{\text{fail}}_{\text{signal}} P^{\text{fail}}_{\text{signal}}(M_{ee}) + N^{\text{fail}}_{\text{bkg}} P^{\text{fail}}_{\text{bkg}}(M_{ee}) \tag{11.7}$$

$N^{\text{pass(fail)}}(M_{ee})$ is the number of events in the Tag+Pass (Tag+Fail) sample as a function of M_{ee}, and floats implicitly. $N^{\text{pass}}_{\text{bkg}}$ is the total, fixed number of background events in the Tag+Pass sample. $P^{\text{pass}}_{\text{bkg}}(M_{ee})$ is the fixed, uniform background PDF in the Tag+Pass sample. The trigger efficiency $\epsilon_{\text{ID}\to\text{Trig}}$ is fixed: Its value is folded in directly from the fourth row of Table 9.3.

Figure 11.1 shows the fit to the (a) Tag+Pass and (b) Tag+Fail samples. The χ^2/DOF of the fit to the Tag+Pass (Tag+Fail) sample is 0.990 (1.126), corresponding to a probability value of $p = 0.549$ (0.086). The combined χ^2/DOF for the full fit to both samples is 1.041, corresponding to a probability value of $p = 0.227$. The positive and negative errors of the floating parameters have been evaluated using MINOS, and their absolute values are near to each other. The resultant signal yield, efficiency, and cross section are:

$$N_{\text{S}} = 8410 \pm 92 \text{ (simul fit)} \tag{11.8}$$

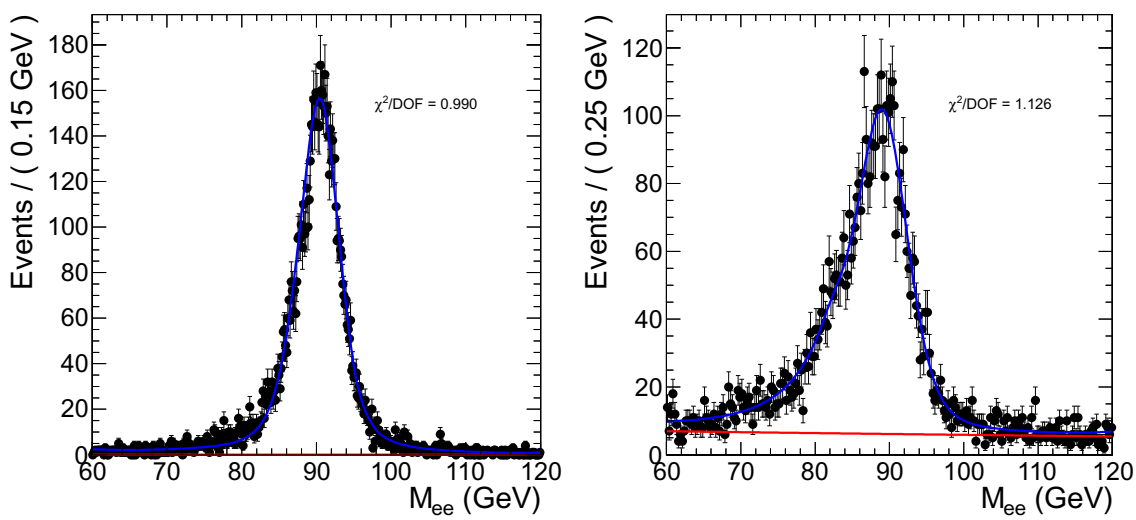
$$\epsilon_Z = \epsilon^2_{\text{Acc}\to\text{ID}}(1 - (1 - \epsilon_{\text{ID}\to\text{Trig}})^2) = 0.6061 \pm 0.0063 \text{ (simul fit)} \tag{11.9}$$

$$\sigma(pp \to Z + X) \times BR(Z \to ee) = 997 \pm 11 \text{ pb (simul fit)} \tag{11.10}$$

This cross section result is in good agreement with the prior ones obtained using either the two-binned or fine-binned electron efficiency measurements along with the counted signal yield:

$$\sigma(pp \to Z+X) \times \mathcal{B}(Z \to ee) = 997 \pm 11(\text{stat}) \pm 21(\text{syst}) \pm 40(\text{lumi}) \text{ pb (two-binned/counted)}$$

$$\sigma(pp \to Z+X) \times \mathcal{B}(Z \to ee) = 1001 \pm 11(\text{stat}) \pm 22(\text{syst}) \pm 40(\text{lumi}) \text{ pb (fine-binned/counted)}$$



(a) (b)

Figure 11.1: The simultaneous fit. The data is in black, background fit in red, and signal+background fit in blue: (a) Tag+Pass sample, (b) Tag+Fail sample.

The efficiency determination of the simultaneous fit is very similar to that of the Tag and Probe method, implying that the non-fit systematic uncertainty on the efficiency from the simultaneous fit can be taken directly from Table 9.12. This leads to a total non-fit fractional uncertainty on the efficiency of $\Delta\epsilon_Z/\epsilon_Z = 0.0095$. Propagating this along with the other uncertainties yields the final simultaneous fit result as:

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee) = 997 \pm 11(\text{stat}) \pm 19(\text{syst}) \pm 40(\text{lumi}) \text{ pb (simul fit)} \tag{11.11}$$

where the uncertainty returned by the fit has been taken as the statistical one since it is dominated by the Poisson counting error on N_{S}. This measurement has less systematic uncertainty than those using the counted signal yield, as expected. Hence the simultaneous fit result is taken as the final, quoted production cross section measurement.

A summary of the measurements using the different methods (simultaneous fit and counted signal yield using the two-binned or fine-binned tag and probe electron efficiencies) is given in Figure 11.2. The figure illustrates the consistency of the three results, as well as the confirmation of the NNLO theoretical prediction. For each reported measurement, the statistical error is represented in black and the total experimental uncertainty, obtained by adding in quadrature the statistical and systematic uncertainties, in dark blue. The luminosity uncertainty is added in quadrature to the experimental uncertainty, and is represented in green. The orange vertical line represents the theoretical prediction, and the light-yellow vertical band is the theoretical uncertainty, interpreted as a 68% confidence interval.

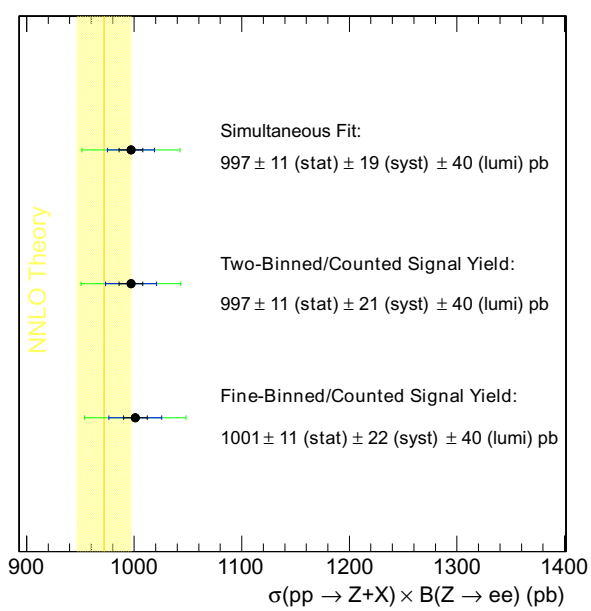


Figure 11.2: Summary of the $Z \to ee$ production cross section measurements.

Simultaneous Fit with Different ID Choices

The robustness of the simultaneous fit method is verified in this subsubsection by confirming the stability of the cross section measurement when using different ID criteria. The definitions of the Tag+Pass and Tag+Fail samples are the same as described above, but have the standard WP80 ID swapped for either the WP70 (Table 6.9) or WP90 (Table 6.7) ID. The measured production cross sections are listed in Table 11.2.[2] The uncertainties in the table are the combined errors returned by the fits. A summary of the measurements with the three different choices of ID criteria is provided in Figure 11.3. For each measurement, the combined fit uncertainty is represented in black. The orange vertical line represents the theoretical prediction, and the light-yellow vertical band is the theoretical uncertainty, interpreted as a 68% confidence interval. The results from the three ID choices are in good agreement with each other. The measurements also further confirm the NNLO theoretical prediction. The table indicates that the default WP80 ID yields an error at least as small as the other choices; suggesting that an optimal default working point was selected, as claimed in Subsection 6.1.3.

Table 11.2: The $Z \to ee$ production cross section using different ID criteria.

ID	$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee)$ (pb)
WP80	997 ± 11
WP70	998 ± 11
WP90	1002 ± 11

Pseudo-Experiment Systematic Bias Study

One further cross-check to perform is a simultaneous fit systematic bias study using a large number of pseudo-experiments, as was done in Section 8.2. Toy MC ensembles are generated using the parameters of the simultaneous fit found in data. The simultaneous fit routine is then run on these ensembles as it would be in the data to confirm that the cross

[2] The jet background is non-vanishing in the Tag+Pass sample for the WP90 case: It was subtracted using the template method developed in Chapter 7. The non-jet backgrounds in the Tag+Pass samples have been estimated from the MC for each ID choice.

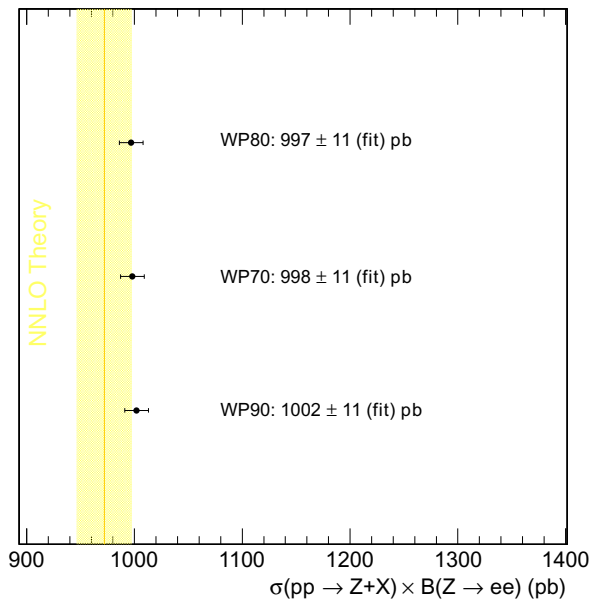
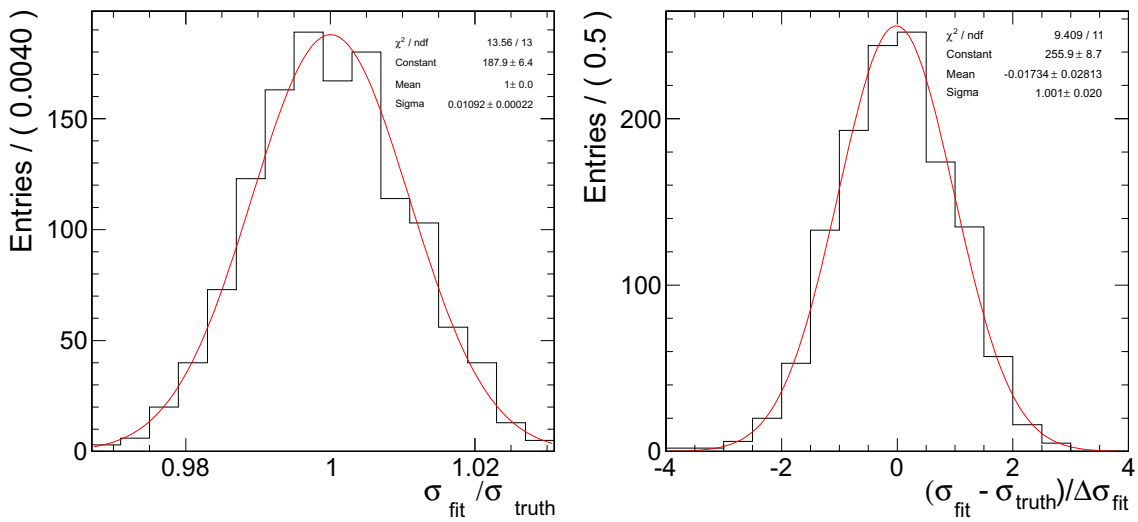


Figure 11.3: Summary of the $Z \rightarrow ee$ production cross section measurements using the simultaneous fit technique with different ID choices.

section and its error are accurately reproduced. Figure 11.4 (a) displays a histogram of the ratio of the production cross sections reported by the fits to the true production cross sections of the pseudo-experiments. This histogram is fit to a Gaussian, which has a mean close to one, indicating that the fit accurately reproduces the true cross sections.
Figure 11.4 (b) plots the pull distribution of the production cross sections reported by the pseudo-experiments. This histogram is fit to a Gaussian, which has a mean close to zero and a width close to one, indicating that the simultaneous fit accurately reproduces the true errors on the reported cross sections. These results, combined with the evaluation of the MINOS errors, confirm the statistical robustness of the simultaneous fit procedure.

Efficiency Results Profiled in Terms of Integrated Luminosity

The $Z \rightarrow ee$ efficiency is measured in blocks of fixed integrated luminosity using the simultaneous fit technique. The results are profiled in Figure 11.5 to assess whether fluctuations in detector performance impacted the efficiency. The luminosity blocks chosen



(a) (b)

Figure 11.4: (a) Histogram (black) of the ratio of the production cross sections reported by the fits to the true cross section of the pseudo-experiments. The histogram is fit to a Gaussian (red). (b) The pull distribution (black) of the production cross sections reported by the pseudo-experiments. The histogram is fit to a Gaussian (red).

correspond to 2.4 pb^{-1}, except the last one which corresponds to 2.3 pb^{-1}. The data is shown in black. The corresponding fit uncertainties are represented by a hatched band. The average of these measurements is represented by an orange line, and the fit of the profile to a constant is represented by a red line. The figure is consistent with a stable efficiency, having no large deviations present, a $\chi^2/DOF = 1.163$, and a corresponding probability value $p = 0.296$.

Simultaneous Fit Summary

The technique to extract the signal yield and electron efficiency directly from a single simultaneous fit was introduced in this section. The robustness of the technique was verified by performing measurements using selection criteria both looser and tighter than the standard ones. The results using these different selections agreed well with that

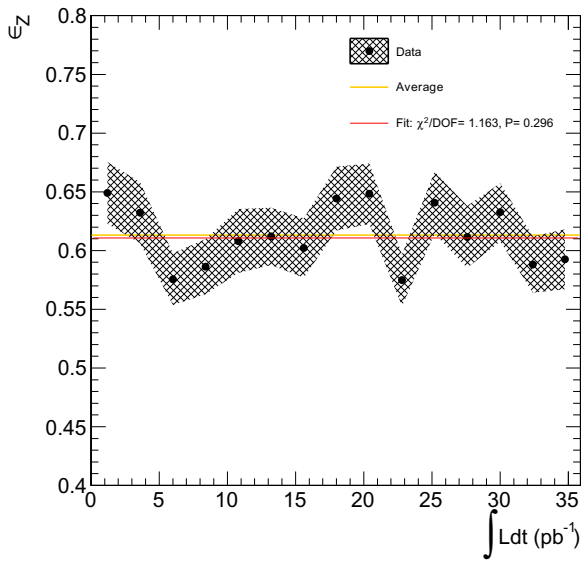


Figure 11.5: ϵ_Z profiled in terms of integrated luminosity. The data profile is shown in black. The corresponding errors are represented by a hatched band. The average of the profile is represented by an orange line. The fit of the profile to a constant is represented by a red line. The χ^2/DOF and corresponding probability of the fit are printed on the plot.

obtained using the standard selection. This comparison also suggested that the standard selection was optimal for the $Z \rightarrow ee$ production cross section measurement. The statistical stability of the fit was validated in a systematic bias study using large numbers of pseudo-experiments. The $Z \rightarrow ee$ efficiency was profiled in terms of integrated luminosity to assess whether it was impacted by fluctuations in detector performance. There was no indication of this. The measurements presented in this section agreed with the prior efficiency and production cross section measurements performed using the Tag and Probe method and counted signal yields. Moreover, they provided further confirmation of the NNLO theoretical prediction.

11.4 Summary

The $Z \rightarrow ee$ production cross section has been measured in a few different ways:

- Counted signal yield with two-binned Tag and Probe electron efficiencies.

- Counted signal yield with fine-binned Tag and Probe electron efficiencies.
- Simultaneous Fit: extended maximum log-likelihood fit that determined the cross section, signal yield, and electron efficiencies simultaneously.

The results of the three measurements are summarized in Figure 11.2. The reported $Z \to ee$ production cross sections were internally consistent and provided confirmation of the NNLO theoretical prediction. The simultaneous fit constituted the measurement with the lowest systematic uncertainty. Hence it has been taken as the final, quoted measurement:

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee) = 997 \pm 11(\text{stat}) \pm 19(\text{syst}) \pm 40(\text{lumi}) \text{ pb} \tag{11.12}$$

(quoted cross section result)

Figure 11.6 shows the quoted CMS inclusive $Z \to ee$ production cross section measurement together with $Z \to \ell\ell$ measurements at lower energy hadron colliders [1, 2, 3, 4] and the NNLO theoretical prediction. The measured cross sections are represented by solid symbols with error bars as indicated in the legend, and the NNLO theoretical prediction is represented by a blue line. The predicted increase of the cross section with energy is confirmed by the CMS measurement.

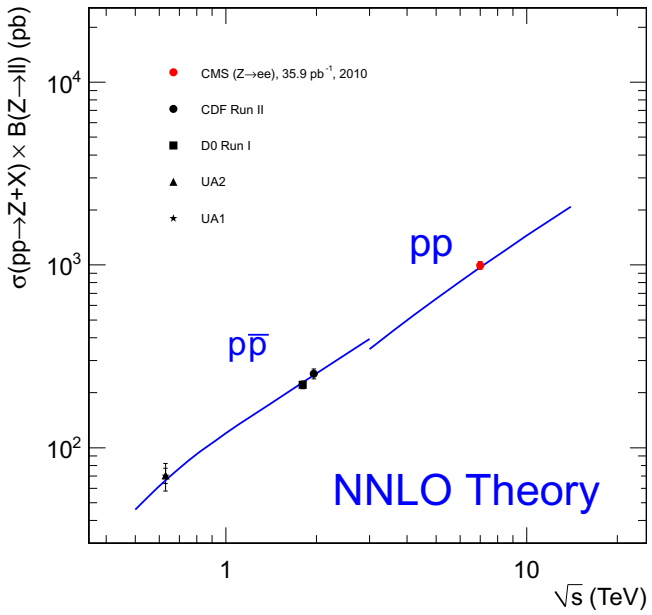


Figure 11.6: Measurements of the $Z \to ee$ inclusive production cross section from CMS and $Z \to \ell\ell$ measurements from experiments at lower energy hadron colliders [1, 2, 3, 4]. The measured cross sections are represented by solid symbols with error bars as indicated in the legend, and the NNLO theoretical prediction is represented by a blue line.

Chapter 12

$Z \to ee$ Decays as a Luminometer

12.1 Introduction

Electroweak gauge bosons are copiously produced at the LHC: Approximately 3000 Z decays per lepton channel are expected for every 10 pb^{-1} of integrated luminosity, after accounting for detector acceptance and efficiency. This translates to nearly 100 Z decays per lepton channel every 30 s at the LHC design luminosity of 10^{34} cm^{-2}s^{-1}.[1] Electroweak gauge boson production cross sections are known to a high level of accuracy [85], and their leptonic decays have clean detector signatures. These facts combine to suggest that leptonic gauge boson decays may be leveraged as "standard candles" at the LHC for measuring the luminosity [89, 93, 94]. A $Z \to ee$ based calibration may be done, for instance, by inverting the $Z \to ee$ cross section measurement: The observed signal yield is used along with the theoretical production cross section and acceptance values to instead solve for the luminosity in Equation 1.7:

$$\int \mathcal{L} dt = \frac{N_{\mathrm{S}}}{\epsilon_Z (A\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee))_{\text{ Theory}}} \tag{12.1}$$

and similarly for the other channels.

In the limit of high luminosity, the fractional systematic uncertainty on the $Z \to \ell\ell$ cross section measurement is $\delta\sigma_Z \sim 2\delta\epsilon_\ell$, while that of the $W \to \ell\nu$ cross section is

[1] The rate of $W \to \ell\nu$ decays is approximately 10 times higher.

$\delta\sigma_W \sim \sqrt{(\delta\epsilon_\ell)^2 + (\delta\not{E}_T)^2 + (\delta G)^2}$, where $\delta\epsilon_\ell$ is the fractional uncertainty on the single lepton reconstruction and identification efficiency, $\delta\not{E}_T$ is the fractional $\not{E}_T$ uncertainty, and δG is the fractional calorimeter energy scale and resolution systematic uncertainty. For high luminosity, $(\delta\epsilon_\ell)^2 \ll (\delta\not{E}_T)^2 + (\delta G)^2$, implying a W based measurement carries larger uncertainty than a Z based measurement in this limit [95]. In addition, soon it will no longer be possible to collect $W \to \ell\nu$ decays unprescaled: The single lepton triggers will have to be prescaled to control the trigger rate at instantaneous luminosities above about 10^{33} cm^{-2}s^{-1}. This is not the case for $Z{\to}\ell\ell$ decays, since they can be collected using double lepton triggers that have tolerable rates up to design luminosity. Such observations thus make $Z{\to}\ell\ell$ the most attractive channels for luminosity calibration or monitoring at high instantaneous luminosity.

12.2 Calibration of the 2010 Recorded Luminosity using $Z{\to}ee$ Decays

The $Z{\to}ee$ inclusive production cross section measurement was thoroughly validated and cross-checked (Chapters 5–11). It provided confirmation of the NNLO theoretical prediction. The result is also consistent with the other electroweak measurements performed at CMS on the same dataset [96, 97], which include inclusive production cross section measurements for $Z{\to}\mu\mu$, $W \to e\nu$, $W^+ \to e^+\nu$, $W^- \to e^-\nu$, $W \to \mu\nu$, $W^+ \to \mu^+\nu$, and $W^- \to \mu^-\nu$, as well as combined lepton channel results for Z, W, and $W^\pm$. The ratios of the W to Z and W^+ to W^- cross sections have also been measured. The results are all internally consistent, and further confirm the NNLO theoretical predictions. The consistency of the results suggests already inverting the $Z{\to}ee$ cross section measurement to instead measure the absolute recorded luminosity. $Z{\to}ee$ decays could then be used alongside the Van der Meer separation scans (Section 3.6) to calibrate the relative luminosity measured with the HF (Chapter 3).

Substituting the theoretical acceptance and cross section values (Equations 10.2 and

10.1, respectively), along with the yield and efficiency determined by the simultaneous fit (Equations 11.8 and 11.9, respectively), into Equation 12.1, yields the following result for the recorded luminosity:

$$\left(\int \mathcal{L}dt\right)_{Z\to ee} = 36.8 \pm 0.4(\text{stat}) \pm 1.2(\text{syst})\ \text{pb}^{-1} \tag{12.2}$$

The fractional error on the ratio of the signal yield to the signal efficiency is $\Delta(N_{\text{S}}/\epsilon_Z)/(N_{\text{S}}/\epsilon_Z) = 0.0109$, as determined by the fit. This error is propagated as the statistical one since it is dominated by the Poisson counting error on N_{S}. The contributions to the systematic error are dominated by the effective cross section: $\Delta(A\sigma\mathcal{B})/(A\sigma\mathcal{B}) = 0.031$. Other contributions include the non-fit errors on the signal efficiency $\Delta\epsilon_Z/\epsilon_Z = 0.0095$ and the uncertainty due to the background subtraction $\Delta N_{\text{S}}/N_{\text{S}} = 0.0014$.

The $Z\to ee$ based calibration above is in agreement with the Van der Meer separation scan calibration using the HF luminosity data (Equation 3.26):

$$\left(\int \mathcal{L}dt\right)_{\text{scan}} = 35.9 \pm 1.4(\text{syst})\ \text{pb}^{-1} \tag{12.3}$$

Thus with only about 35–40 pb^{-1} of data, the luminosity measurement using $Z\to ee$ decays already has an uncertainty of just 3.4%, which is competitive with the 4.0% uncertainty of the separation scan calibration. In fact, a $Z\to\ell\ell$ decay based luminosity calibration with an uncertainty of less than 4% should be possible on a daily basis if the LHC provides CMS with approximately 30–40 pb^{-1} of collisions per day. Such high luminosity data taking could occur before the end of 2011. In the future, when the LHC design luminosity is achieved and later exceeded by the Super LHC upgrade [98], it may even be possible to measure the instantaneous luminosity using Z boson decays in real-time.

12.3 Luminosity Measurement Stability

The stability of a luminosity measurement using electroweak boson decays must be continuously checked. Irregularities can be uncovered in a couple of ways, as discussed below.

The first method, known as a Kolmogrov-Smirnov (KS) omnibus test [99], is used to detect anomalies in the uncorrected signal yield. Its main advantage is that it can assess the stability of the yield on an event-by-event basis. For instance, calculating efficiency-corrected yields with such a fine granularity is not possible for the purpose of live event counting for real-time luminosity monitoring. Alternatively, the efficiency can be assumed constant and by evaluating the time-ordered frequency distribution of the uncorrected yield, the KS test can detect anomalous behavior in real-time, signaling fluctuations in efficiency.

The second method is a χ^2 test on the W or Z based luminosity (which is proportional to the efficiency corrected signal yield) profiled in terms of blocks of fixed integrated luminosity measured in an orthogonal way (e.g., the relative HF luminosity measurement calibrated by the separation scans). The test is useful at the level of granularity in which it is feasible to calculate the signal efficiency.[2] The distribution is assumed to be flat and fit to a constant. The resultant χ^2/DOF of the fit determines the probability that the distribution is consistent with the uniform hypothesis. If desired, the test can also be run on the uncorrected yields to alert a change in efficiency.

12.3.1 Kolmogrov-Smirnov Omnibus Tests

Performing KS omnibus tests [99] allows detection of irregularities in the uncorrected signal yield. These tests yield the significance or probability value of an observed or claimed deviation in a given frequency distribution from the expected one. KS tests are popular in high energy physics, being frequently used because they are relatively simple, robust, and offer clear visual interpretations. In KS tests, the empirical distribution function (EDF) of a given signal yield is plotted versus an orthogonal variable. The EDF is just the fractional yield of the signal ordered in some way. The orthogonal variable could be the separation scan calibrated luminosity or the EDF of another signal. This frequency distribution is then compared to the expected distribution (e.g., the yield

[2] This corresponds to about 2 pb^{-1}. Below this amount of luminosity there are too few events in the Tag+Fail sample for the signal contribution to be reliably estimated.

increasing linearly with luminosity). The signal yield N and maximal vertical difference D_{stat} between the observed and expected distributions determine the probability that the observed distribution is consistent with the hypothesis model according to the relation:

$$P_{\text{KS}} = 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2(jD_{\text{stat}}N)^2} \tag{12.4}$$

Such a test is shown in Figure 12.1, where the fractional $Z\rightarrow ee$ yield is plotted versus the separation scan calibrated luminosity. The data is shown in black, while the expected distribution is in blue. The observed distribution agrees well with the expectation and the probability value $P_{\text{KS}} = 0.170$ indicates a stable signal yield, consistent with the hypothesis that it increases linearly with luminosity.

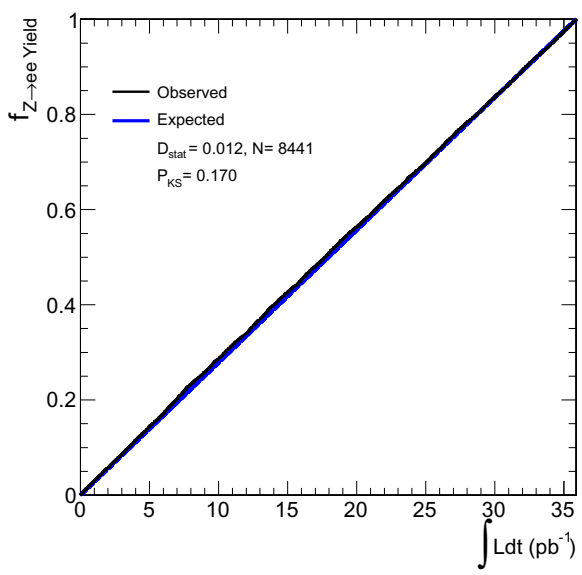


Figure 12.1: EDF plot: fractional $Z\rightarrow ee$ yield vs. integrated luminosity. The observed frequency distribution is plotted in black, while the expected distribution is plotted in blue. Results of the corresponding KS test are printed on the plot.

KS tests are also performed on the $Z\rightarrow\mu\mu$ (Figure 12.2) and $W \rightarrow e\nu$ (Figure 12.3) signal events reported in [96, 97]. The probability $P_{\text{KS}} = 0.176$ for the $Z\rightarrow\mu\mu$ yield indicates stability in this yield. The low probability $P_{\text{KS}} = 10^{-6}$ for $W \rightarrow e\nu$ indicates the yield of this signal is not stable. This suggests the collection of W decays is more sensitive

to fluctuations in detector performance than that of Z decays. This is expected since collecting a $Z \to \ell\ell$ decay requires only one of the two daughter leptons to be triggered, while the single lepton must be triggered to collect a $W \to \ell\nu$ decay.[3] [4]

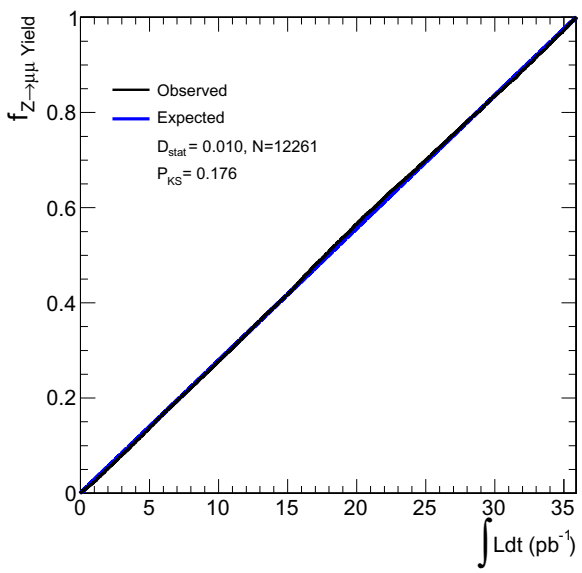


Figure 12.2: EDF plot: fractional $Z \to \mu\mu$ yield vs. integrated luminosity. The observed frequency distribution is plotted in black, while the expected distribution is plotted in blue. Results of the corresponding KS test are printed on the plot.

The signal yield should be compared to an additional orthogonal variable to confirm its stability independently of the separation scan calibrated luminosity. A natural choice is to compare it to the yield of some other signal, selected such that its triggered detector signature is not correlated to that of the original signal. Comparing the time-ordered yields of $Z \to ee$ and $Z \to \mu\mu$ is a good choice, for instance, since these signals are triggered by two entirely separate sub-detectors (the Ecal for $Z \to ee$ and the Muon chambers for $Z \to \mu\mu$). This plot is shown in Figure 12.4. The resultant probability $P_{\mathrm{KS}} = 0.070$ suggests the yields of both signals are stable independent of the separation scan calibrated luminosity.

[3] Recall, however, from Section 12.1 that Z decays will eventually be collected using dilepton triggers, in which case a first order dependence will also apply to them.

[4] Indeed, the irregularity in the $W \to e\nu$ EDF is further evidence for the deviation in the single electron trigger efficiency observed in Subsection 9.5.2.

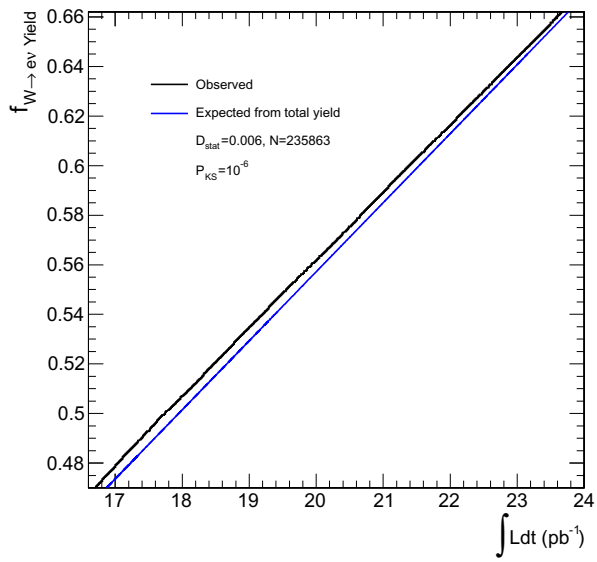


Figure 12.3: EDF plot: fractional $W \to e\nu$ yield vs. integrated luminosity. The observed frequency distribution is plotted in black, while the expected distribution is plotted in blue. Results of the corresponding KS test are printed on the plot. The plot is zoomed-in to display the maximum deviation, which is visible in the lower left hand corner.

12.3.2 W/Z Luminosity Profiled in Terms of Separation Scan Calibrated Luminosity

The W or Z based luminosity profiled in terms of blocks containing a fixed amount of separation scan calibrated luminosity can be examined to determine whether there is any time-dependence. The profile can be fit to the expected, flat distribution and a χ^2 test can be performed to evaluate whether the observed distribution is consistent with the expected one.

Figure 12.5 plots the measured $Z\to ee$ luminosity versus separation scan luminosity block. The separation scan calibrated luminosity blocks chosen correspond to 2.4 pb^{-1}.[5] The data profile is shown in black, and the corresponding fit uncertainties are represented by a hatched band. The average of these measurements is represented by an orange line, and the fit of the data to a constant is represented by a red line. The low

[5] As in Subsection 9.5.2, the last block is an exception: It contains 2.3 pb^{-1}, but has its results reweighted to correspond to 2.4 pb^{-1}.

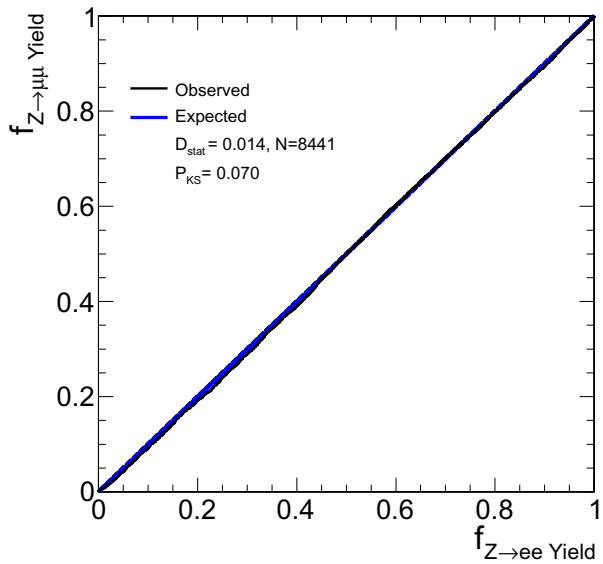


Figure 12.4: EDF Plot: time-ordered fractional $Z\to\mu\mu$ yield vs. time-ordered fractional $Z\to ee$ yield. The observed frequency distribution is plotted in black, while the expected distribution is plotted in blue. Results of the corresponding KS test are printed on the plot.

$\chi^2/DOF = 0.677$ and high corresponding probability $p = 0.798$ indicate the luminosity values extracted from the corrected $Z\to ee$ yields are stable over time with respect to the separation scan calibrated luminosity. The normalization is also in good agreement with the separation scan luminosity values. These observations combine to suggest that $Z\to ee$ decays could be used alongside the separation scans to provide an absolute calibration of the relative HF luminosity measurement.

Similarly, Figure 12.6 plots the $W \to e\nu$ luminosity versus separation scan calibrated luminosity block. In this case, the electron efficiencies are determined using the simultaneous fit method as above, while the uncorrected $W \to e\nu$ yield is taken from the raw event count used in [96, 97]. The hatched error band represents the errors on the efficiencies returned by the fits added in quadrature to the Poisson counting error on the raw event yields. The theoretical cross section and acceptance values are also taken from [96, 97]. The $\chi^2/DOF = 1.649$ and corresponding probability $p = 0.058$ suggest the luminosity values extracted from the corrected $W \to e\nu$ yields are stable with respect to

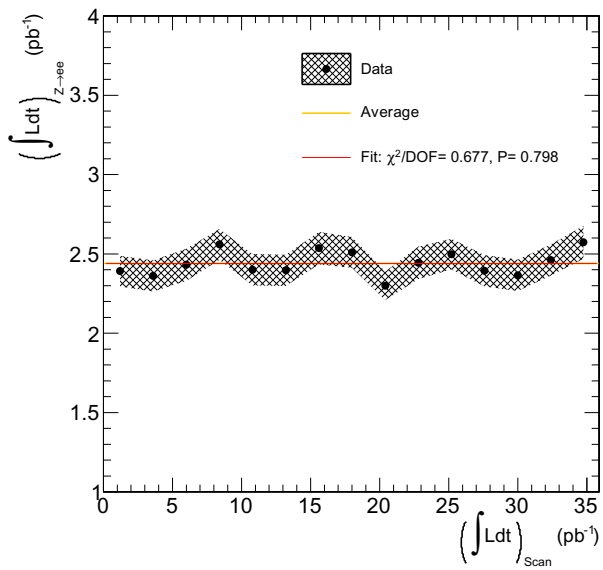


Figure 12.5: The $Z \to ee$ luminosity profiled in terms of separation scan calibrated luminosity. The data profile is shown in black. The corresponding errors are represented by a hatched band. The average of the profile is represented by an orange line. The fit of the profile to a constant is represented by a red line. The χ^2/DOF and corresponding probability of the fit are printed on the plot.

the separation scan calibrated luminosity. The normalization is also in good agreement with the separation scan luminosity values.

Finally, in order to check the stability of the $Z \to ee$ based luminosity measurement in a way independent of the separation scan calibrated luminosity, the ratio of the $W \to e\nu$ to $Z \to ee$ luminosities is profiled in Figure 12.7 versus separation scan calibrated luminosity. This choice of signals provides the comparison that is least sensitive to variations in the lepton efficiencies, since the ratio only includes one electron reconstruction and identification efficiency factor and the correlation between it and the $Z \to ee$ signal yield is accounted for. The hatched error band represents the errors on the efficiencies returned by the fits and the Poisson counting errors on the uncorrected event yields. The low $\chi^2/DOF = 0.611$ and high corresponding probability $p = 0.857$ indicate the ratio is stable. The normalization is also close to unity, indicating agreement between the $W \to e\nu$ and $Z \to ee$ luminosity measurements.

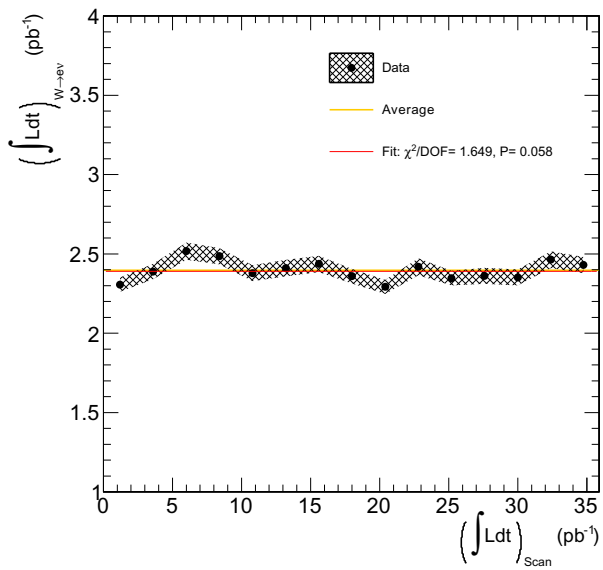


Figure 12.6: The $W \to e\nu$ luminosity profiled in terms of separation scan calibrated luminosity. The data profile is shown in black. The corresponding errors are represented by a hatched band. The average of the profile is represented by an orange line. The fit of the profile to a constant is represented by a red line. The χ^2/DOF and corresponding probability of the fit are printed on the plot.

12.4 Pileup Effects

Understanding pileup effects will become increasingly important to a physics signal based luminosity measurement as higher luminosities are achieved.[6] Higher pileup energy deposits in the calorimeters will lead to a lower signal efficiency for a given selection criteria. Preliminary studies of the 2011 data indicate a drop in ϵ_Z (Equation 9.1) of approximately 14% per 1 μb^{-1}s^{-1} increase in single bunch crossing luminosity.[7] No significant change in the uncertainty has been observed. Still, this preliminary study has only covered up to a maximum single bunch cross luminosity of about 1.2 μb^{-1}s^{-1}, which is much less than the design single bunch crossing luminosity: More detailed studies at higher instantaneous luminosity will be needed to quantify the effects of pileup. The tests

[6] If in the future, the existing HF luminosity monitor exhibits non-linearities resulting from pileup, they may prove quite challenging to adequately understand as compared to pileup effects in a Z based luminosity measurement. Such a physics signal based luminosity measurement may be most desirable if this proves to be the case.

[7] At LHC design luminosity and energy, the average single bunch crossing luminosity is about 4 μb^{-1}s^{-1}.

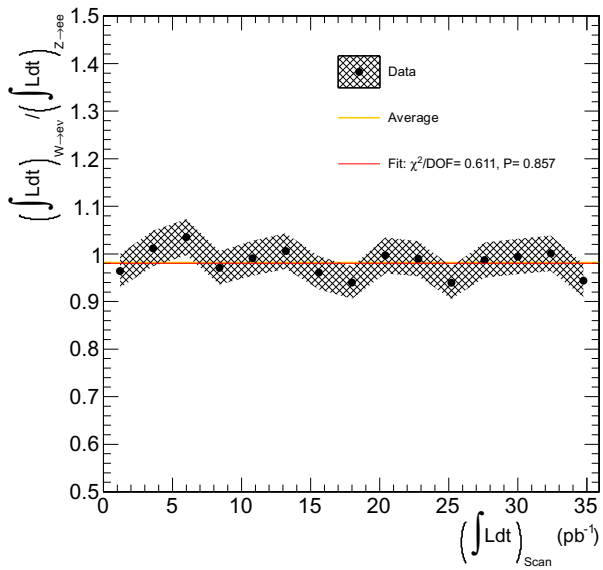


Figure 12.7: The ratio of the $W \to e\nu$ to $Z \to ee$ luminosities profiled in terms of separation scan calibrated luminosity. The data profile is shown in black. The corresponding errors are represented by a hatched band. The average of the profile is represented by an orange line. The fit of the profile to a constant is represented by a red line. The χ^2/DOF and corresponding probability of the fit are printed on the plot.

discussed in the previous subsection will be useful in disentangling the effects of pileup. For instance, KS tests can be performed on the uncorrected W or Z signal yields ordered in terms of the average bunch crossing instantaneous luminosity, rather than time, as shown for $Z \to ee$ in Figure 12.8. The W or Z based luminosities or lepton efficiencies can also be profiled versus bins of separation scan calibrated instantaneous luminosity.

12.5 Summary

Leptonic W and Z decays have ideal properties for measuring the absolute luminosity at the LHC and hence for calibrating the relative luminometers of the experiments, in particular the HF at CMS:

- Copious production.
- Well-known cross sections.

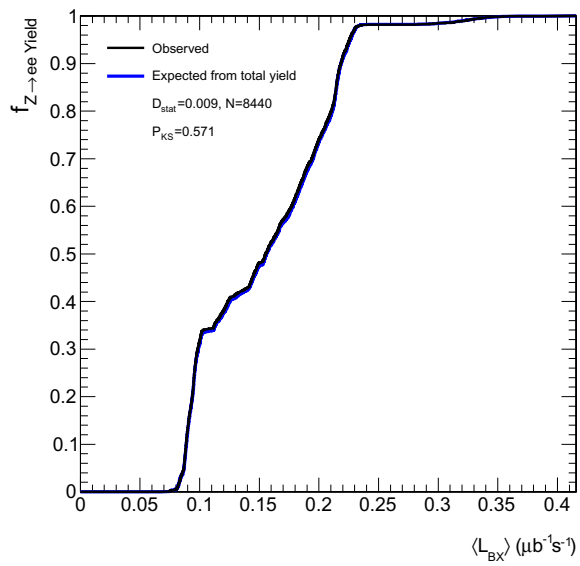


Figure 12.8: EDF plot: The fractional $Z \to ee$ yield vs. average bunch crossing instantaneous luminosity. The observed frequency distribution is plotted in black, while the expected distribution is plotted in blue. Results of the corresponding KS test are printed on the plot.

- Clean detector signatures.

Z decays in particular are excellent candidates for such a calibration because they can be triggered unprescaled at very high luminosities, and have smaller associated systematic uncertainties than do W decays in this limit.

The absolute, recorded luminosity is determined by the $Z \to ee$ decays according to Equation 12.1. A $Z \to ee$ decay based calibration of the full 2010 recorded luminosity was provided in Equation 12.2:

$$\left(\int \mathcal{L} dt\right)_{Z \to ee} = 36.8 \pm 0.4(\text{stat}) \pm 1.2(\text{syst}) \text{ pb}^{-1} \tag{12.5}$$

which was in good agreement with the quoted value of $\int \mathcal{L} dt = 35.9 \pm 1.4(\text{syst}) \text{ pb}^{-1}$ calibrated via the Van der Meer separation scans using the HF luminosity data. With an uncertainty of just 3.4%, this result constitutes one of the most precise absolute luminosity measurements performed to date at a hadron collider.[8] It is also the first physics signal

[8] A recent Van der Meer separation scan calibration at ATLAS also claims a 3.4% uncertainty [100]. The only absolute luminosity measurement with less uncertainty was performed at the Intersecting Storage Rings

based luminosity measurement performed at the LHC. High luminosity running, in which 30–40 pb^{-1} of integrated luminosity may be recorded each day, is possible before the end of 2011. If such running indeed occurs, $Z \to ee$ decays will be able to provide a calibration with an uncertainty of less than 4% on a daily basis, which is comparable to the 4% uncertainty of the separation scan calibration. As LHC design luminosity is approached, it will also become possible to count $Z \to ee$ decays for real-time luminosity monitoring.

The stability of these physics signal based luminosity measurements must be continuously confirmed. Hence two methods of checking the stability were introduced. The first relied on Kolmogrov-Smirnov omnibus tests performed on the uncorrected signal yields and the second relied on χ^2 tests performed on the gauge boson calibrated luminosity profiles. Both methods had the advantage that they could test the stability in ways either dependent or independent of the separation scan calibrated luminosity. Both methods confirmed the stability of the $Z \to ee$ based luminosity measurement on the 2010 dataset, ensuring the robustness of the calibration provided above. Understanding the effects of pileup will become increasingly important at higher instantaneous luminosities, and these tests will assist in their evaluation.

in the 1980s: Less stringent machine specifications at this collider allowed a measurement with an accuracy of about 1% [101]. In the future, it may also be possible to achieve a 1% error on the absolute luminosity using the very forward detectors (TOTEM at CMS and ALPHA at ATLAS) at the LHC [102].

Chapter 13

Conclusions

This thesis has comprised the first data-driven, systematic uncertainty limited, precision measurement of the inclusive $Z \to ee$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV and absolute luminosity based on $Z \to ee$ decays. The data was collected by the CMS detector at the LHC during 2010 and corresponded to an integrated luminosity of $\int \mathcal{L} dt = 35.9$ pb^{-1}.

The cross section pertaining to the invariant mass window $M_{ee} \in (60, 120)$ GeV has been reported as:

$$\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee) = 997 \pm 11(\text{stat}) \pm 19(\text{syst}) \pm 40(\text{lumi}) \text{ pb} \tag{13.1}$$

which is in agreement with the NNLO theoretical prediction of:

$$(\sigma(pp \to Z + X) \times \mathcal{B}(Z \to ee))_{\text{NNLO Theory}} = 972 \pm 26 \text{ pb} \tag{13.2}$$

The cross section measurement has been a benchmark physics analysis performed after the first year of CMS detector and LHC machine operations. It provided validation of high P_T electron reconstruction and identification at CMS, in addition to confirming the Standard Model at the new energy frontier. All systematic uncertainties have been evaluated, and the measurement has been heavily cross-checked using multiple techniques to ensure the robustness of the result.

The confirmation of the theoretical prediction for the $Z \to ee$ cross section suggested leveraging the copiously produced and well-understood $Z \to ee$ decays as standard candles to measure the absolute luminosity at the experiments. A measurement of the 2010 absolute luminosity using $Z \to ee$ decays at CMS was performed, which was in agreement with the Van der Meer separation scan calibration using the HF relative luminosity data. It was the first physics signal based absolute luminosity measurement performed at the LHC, and it carried an uncertainty of just 3.4%, making it among the most precise luminosity measurements performed to date at a hadron collider. The low uncertainty also made it competitive with the separation scan calibration, which had a 4.0% uncertainty. During the high luminosity LHC running planned for 2011, a $Z \to ee$ based calibration with an uncertainty below 4% could be performed every 24 h. As LHC design luminosity is approached, it will also become possible to perform live $Z \to \ell\ell$ event counting for real-time luminosity monitoring.

Appendix A

Derivation of the Same/Opposite Sign Background Estimation Method

The electric charge signs of the electron daughters of the Z candidates are used in this method to estimate the background beneath the signal peak. The dominant background in a loose selection is QCD multijet events, where jets fake electrons from a Z decay. This background is charge product symmetric; i.e., it contains the same amount of opposite-sign (OS) and same-sign (SS) charge pairs. The number of OS and SS events *observed* in data can be written as:

$$\begin{aligned} N^{\mathrm{OS}} &= N_{\mathrm{S}}^{\mathrm{OS}}(1-q_1^{\mathrm{misid}})(1-q_2^{\mathrm{misid}}) + N_{\mathrm{S}}^{\mathrm{OS}} q_1^{\mathrm{misid}} q_2^{\mathrm{misid}} \\ &+ N_{\mathrm{B}}^{\mathrm{OS}}(1-q_1^{\mathrm{misid}})(1-q_2^{\mathrm{misid}}) + N_{\mathrm{B}}^{\mathrm{OS}} q_1^{\mathrm{misid}} q_2^{\mathrm{misid}} \\ &+ N_{\mathrm{B}}^{\mathrm{SS}}[q_2^{\mathrm{misid}}(1-q_1^{\mathrm{misid}}) + q_1^{\mathrm{misid}}(1-q_2^{\mathrm{misid}})] \end{aligned} \tag{A.1}$$

and

$$
\begin{aligned}
N^{\mathrm{SS}} &= N_{\mathrm{B}}^{\mathrm{SS}}(1-q_1^{\mathrm{misid}})(1-q_2^{\mathrm{misid}}) + N_{\mathrm{B}}^{\mathrm{SS}} q_1^{\mathrm{misid}} q_2^{\mathrm{misid}} \\
&+ N_{\mathrm{B}}^{\mathrm{OS}}[q_2^{\mathrm{misid}}(1-q_1^{\mathrm{misid}}) + q_1^{\mathrm{misid}}(1-q_2^{\mathrm{misid}})] \\
&+ N_{\mathrm{S}}^{\mathrm{OS}}[q_2^{\mathrm{misid}}(1-q_1^{\mathrm{misid}}) + q_1^{\mathrm{misid}}(1-q_2^{\mathrm{misid}})]
\end{aligned} \tag{A.2}
$$

where N_{S} (N_{B}) is the true number of signal (background) events in the sample, and q_1^{misid} (q_2^{misid}) is the fraction of events in which the charge of the first (second) electron has been misidentified. Assuming that $N_{\mathrm{B}}^{\mathrm{OS}} = N_{\mathrm{B}}^{\mathrm{SS}}$, and subtracting Equation A.2 from Equation A.1, yields:

$$
N^{\mathrm{OS}} - N^{\mathrm{SS}} = N_{\mathrm{S}}^{\mathrm{OS}}(1-2q_1^{\mathrm{misid}})(1-2q_2^{\mathrm{misid}}) \tag{A.3}
$$

Solving for the number of true signal events gives:

$$
N_{\mathrm{S}}^{\mathrm{OS}} = \frac{N^{\mathrm{OS}} - N^{\mathrm{SS}}}{(1-2q_1^{\mathrm{misid}})(1-2q_2^{\mathrm{misid}})} \tag{A.4}
$$

which is just Equation 7.2. QED.

To derive an expression for the charge misidentification, let q_1^{misid} (q_2^{misid}) denote the probability that the charge of the first (second) electron is misidentified. Then the average charge product of the two electrons in the signal events can be written as:

$$
\begin{aligned}
\langle q_1 \times q_2 \rangle &= -1 \times q_1^{\mathrm{misid}} q_2^{\mathrm{misid}} + 1 \times (1-q_1^{\mathrm{misid}}) q_2^{\mathrm{misid}} \\
&+1 \times (1-q_2^{\mathrm{misid}}) q_1^{\mathrm{misid}} - 1 \times (1-q_1^{\mathrm{misid}})(1-q_2^{\mathrm{misid}})
\end{aligned} \tag{A.5}
$$

where the leading factor of each term is just the charge product corresponding to that term. Equation A.5 can be rewritten as:

$$
\langle q_1 \times q_2 \rangle = -(1-2q_1^{\mathrm{misid}})(1-2q_2^{\mathrm{misid}}) \tag{A.6}
$$

If both electrons have an equal probability for charge misidentification so that $q_1^{\mathrm{misid}} = q_2^{\mathrm{misid}} \equiv q^{\mathrm{misid}}$, then:

$$
\langle q_1 \times q_2 \rangle = -(1-2q^{\mathrm{misid}})^2 \tag{A.7}
$$

or

$$4(q^{\text{misid}})^2 - 4q^{\text{misid}} + (1 + \langle q_1 \times q_2 \rangle) = 0 \tag{A.8}$$

Solving this equation for q^{misid} finally yields:

$$q^{\text{misid}} = \frac{1 \pm \sqrt{|\langle q_1 \times q_2 \rangle|}}{2} \tag{A.9}$$

where the "$-$" solution is just Equation 7.4 and corresponds to signals having opposite-sign charge pairs (e.g., $Z \rightarrow ee$), while the "$+$" solution corresponds to signals with same-sign charge pairs. QED.

References

[1] The CDF Collaboration, "First Measurements of Inclusive W and Z Cross Sections from Run II of the Fermilab Tevatron Collider", *Phys. Rev. Lett.*, vol. 94, no. 091803, 2005.

[2] The D0 Collaboration, "Extraction of the Width of the W Boson from Measurements of $\sigma(p\vec{p} \to W + X) \times B(W \to e\nu)$ and $\sigma(p\vec{p} \to Z + X) \times B(Z \to ee)$ and their Ratio", *Phys. Rev.*, vol. D61, no. 072001, 2000.

[3] The UA1 Collaboration, "Studies of Intermediate Vector Boson Production and Decay in UA1 at the CERN Proton-Antiproton Collider", *Z. Phys.*, vol. G44, no. 15, 1989.

[4] The UA2 Collaboration, "Measurements of the W and Z Production Cross Sections at the CERN $p\bar{p}$ Collider", *Z. Phys.*, vol. C47, no. 11, 1990.

[5] The CMS Collaboration, "The TriDAS Project, Technical Design Report. Volume 2: Data Acquisition and High-Level Trigger", *CERN-LHCC-2002-026, http://cdsweb.cern.ch/record/578006*, 2000.

[6] Particle Data Group; K. Nakamura et al., "The Review of Particle Physics", *J. Phys.*, vol. G37, no. 075021, 2010.

[7] The CMS Collaboration, "The CMS Tracker: Addendum to the Technical Design Report", *CERN-LHCC-2000-016, http://cdsweb.cern.ch/record/490194*, 2000.

[8] N. Adam et al., "Absolute Calibration of the Luminosity Measurement at CMS", *CMS Detector Performance Note: LumiCal_DPN*, 2011.

[9] Particle Data Group; K. Nakamura et al., "Quantum Chromodynamics", *J. Phys.*, vol. G37, no. 075021, Chapter 9, 2010.

[10] Particle Data Group; K. Nakamura et al., "Electroweak Model and Constraints on New Physics", *J. Phys.*, vol. G37, no. 075021, Chapter 10, 2010.

[11] S. I. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.*, vol. 22, no. 579, 1961.

[12] S. Weinberg, "A Model of Leptons", *Phys. Lett.*, vol. 19, no. 1264, 1967.

[13] A. Salam, "Elementary Particle Physics", *Nobel Symp.*, vol. 8, no. 367, 1968.

[14] W. F. Fry and D. Haidt, "Calculation of the Neutron-Induced Background in the Gargamelle Neutral Current Search ", *CERN Yellow Report, no. 1*, 1975.

[15] The UA1 Collaboration, "Experimental Observation of Lepton Pairs of Invariant Mass around 95 GeV/c^2 at the CERN SPS Collider", *Phys. Lett.*, vol. B126, no. 398, 1983.

[16] The UA2 Collaboration, "Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ collider", *Phys. Lett.*, vol. B129, no. 130, 1983.

[17] L. Evans and P. Bryant (editors), "LHC Machine", *JINST*, vol. 3, no. S08001, 2008.

[18] LEP Design Group, "Design Study of a 22 to 130 GeV electron-positron Colliding Beam Machine (LEP)", *CERN/ISR-LEP/79-33*, 1979.

[19] R. R. Wilson, "The Tevatron", *FERMILAB-TM-0763*, 1978.

[20] The CMS Collaboration, "The CMS Experiment at the CERN LHC", *JINST*, vol. 3, no. S08004, 2008.

[21] The ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST*, vol. 3, no. S08003, 2008.

[22] The LHCb Collaboration, "The LHCb Detector at the LHC", *JINST*, vol. 3, no. S08005, 2008.

[23] The ALICE Collaboration, "The ALICE Experiment at the CERN LHC", *JINST*, vol. 3, no. S08002, 2008.

[24] F. Zimmermann and M. P. Zorzano-Mier, "Touschek Scattering in HERA and LHC", *CERN LHC-PROJECT-NOTE-244*, 2000.

[25] The CMS Collaboration, "The CMS Tracker System Project: Technical Design Report", *CERN-LHCC-98-006, http://cdsweb.cern.ch/record/368412*, 1998.

[26] Particle Data Group; S. Eidelman et al., "Review of Particle Physics", *Phys. Lett.*, vol. B592, no. 1, 2004.

[27] P. Lecoq et al., "Lead Tungstate ($PBWO_4$) Scintillators for LHC EM Calorimetry", *Nucl. Instrum. Meth.*, vol. A365, no. 291, 1995.

[28] A. A. Annenkov et al., "Lead Tungstate ($PBW0_4$) Scintillation Material", *Nucl. Instrum. Meth.*, vol. A490, no. 30, 2002.

[29] X. D. Qu et al., "Radiation Induced Color Centers and Light Monitoring for Lead Tungstate Crystals", *IEEE Trans. Nucl. Sci.*, vol. 47, no. 1741, 2000.

[30] Riccardo Paramatti for the CMS Collaboration, "Calibration of the CMS Electromagnetic Calorimeter at LHC startup", *CMS CR 2010/139*, 2011.

[31] R. Arcidiacon et al., "Calibration of the ECAL energy scale using the $J/\psi \to ee$ resonance", *CMS AN 2010/443*, 2010.

[32] S. Nourbakhsh et al., "Energy scale of CMS Electromagnetic Calorimeter with 2010 Data", *CMS AN 2011/038*, 2011.

[33] P. Adzic et al., "Energy Resolution of the Barrel of the CMS Electromagnetic Calorimeter", *JINST*, vol. 2, no. P04004, 2007.

[34] The CMS Collaboration, "The Compact Muon Solenoid Technical Proposal", *CERN-LHCC-94-38, http://cdsweb.cern.ch/record/290969*, 1994.

[35] The CMS Collaboration, "The Electromagnetic Calorimeter Project: Technical Design Report", *CERN-LHCC-97-033, http://cdsweb.cern.ch/record/349375*, 1997.

[36] P. Cushman et al., "Custom HPD Readout for the CMS HCAL", *Nucl. Instrum. Meth.*, vol. A442, no. 289, 2000.

[37] The CMS Collaboration, "The Hadron Calorimeter Project: Technical Design Report", *CERN-LHCC-97-031, http://cdsweb.cern.ch/record/357153*, 1997.

[38] A. Heister et al., "Measurement of Jets with the CMS detector at the LHC", *CMS-NOTE-2006-036, http://cdsweb.cern.ch/record/933705*, 2006.

[39] P. Cushman et al., "Measurement of Missing Transverse Energy with the CMS Detector at the LHC", *Eur. Phys. J.*, vol. C46, no. 45, CMS NOTE 2006/035, http://cdsweb.cern.ch/record/933706, 2006.

[40] V. I. Kryshkin and A. I. Ronzhin, "An Optical Fiber Readout for Scintillator Calorimeters", *Nucl. Instrum. Meth.*, vol. A247, no. 583, 1986.

[41] M. G. Albrow et al., "A Uranium Scintillator Calorimeter with Plastic-Fibre Readout", *Nucl. Instrum. Meth.*, vol. A256, no. 23, 1987.

[42] N. Akchurin and R. Wigmans, "Quartz Fibers as Active Elements in Detectors for Particle Physics", *Rev. Sci. Instrum.*, vol. 74, no. 2955, 2002.

[43] Peter Freund, "Introduction to Supersymmetry", *Cambridge Monographs on Mathematical Physics*, 1986.

[44] The CMS Collaboration, "The CMS Muon Project, Technical Design Report", *CERN-LHCC-97-032, http://cdsweb.cern.ch/record/343814*, 1997.

[45] G. Charpak and F. Sauli, "High-Accuracy, Two-Dimensional Read-Out in Multiwire Proportional Chambers", *Nucl. Instrum. Meth.*, vol. 113, no. 381, 1973.

[46] R. Santonico and R. Cardarelli, "Development of Resistive Plate Counters", *Nucl. Instrum. Meth.*, vol. 187, no. 377, 1981.

[47] R. Cardarelli et al., "Performance of a Resistive Plate Chamber Operating with pure CF_3Br", *Nucl. Instrum. Meth.*, vol. A333, no. 399, 1993.

[48] The CMS Collaboration, "The TriDAS Project, Technical Design Report. Volume 1: The Level-1 Trigger", *CERN-LHCC-2000-038, http://cdsweb.cern.ch/record/706847*, 2000.

[49] D. Acosta et al., "CMS High level trigger", *CMS AN 2007/009*, 2007.

[50] N. Adam et al., "The CMS Luminosity System", *CMS IN 2007/030*, 2006.

[51] The CMS Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions", *CMS Physics Analysis Summary, TRK-10-005*, 2010.

[52] S. Van der Meer, "Calibration of the Effective Beam Height in the ISR", *ISR-PO/68-31, KEK68-64*, 1968.

[53] D. Belohrad et al., "The LHC Fast BCT System: A comparison of Design Parameters with Initial Performance", *Proceedings of BIW10, Santa Fe, New Mexico, US*, 2010.

[54] The CMS Collaboration, "Measurement of CMS Luminosity", *CMS PAS EWK 2010/004*, 2011.

[55] N. Adam et al., "Measurement of CMS Luminosity in the 2010 Run", *CMS AN 2010/175*, 2010.

[56] The CMS Collaboration, "CMS Physics Technical Design Report Volume I: Detector Performance and Software", *CERN-LHCC-2006-001, http://cmsdoc.cern.ch/cms/cpt/tdr/ptdr1_final_colour.pdf*, 2006.

[57] R. Fruhwirth and T. Speer, "A Gaussian-Sum Filter for Vertex Reconstruction", *Nucl. Instrum. and Methods*, vol. A534, no. 217-221. doi:10.1016/j.nima.2004.07.090, 2004.

[58] S. Baffioni et al., "Electron Reconstruction in CMS", *CMS NOTE 2006/040*, 2006.

[59] D. Bandurin et al., "A Baseline Candidate Single Electron HLT Path for LHC Startup", *CMS AN 2009/021*, 2009.

[60] I. Antcheva et al., "ROOT – A C++ Framework for Petabyte Data Storage, Statistical Analysis and Visualization", *Comp. Phys. Comm.*, vol. 180, no. 12, pp. 2499–2512, 2009.

[61] W. Verkerke and D. Kirkby, "The RooFit Toolkit for Data Modeling", *Proceedings of PHYSTAT05, Oxford, UK, http://www.physics.ox.ac.uk/phystat05/proceedings/files/roofit.ps*, 2005.

[62] P. Nason, "A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms", *JHEP*, vol. 0411, no. 040, arXiv:hep-ph/0409146, 2004.

[63] T. Sjostrand et al., "PYTHIA 6.4 Physics and Manual", *JHEP*, vol. 05, no. 026, 2006.

[64] D. Bandurin et al., "Generator-Level Preselection of QCD Electromagnetic Jets for HEEP Analyses", *CMS AN 2008/023*, 2008.

[65] H. L. Lai et al., "New Parton Distributions for Collider Physics", *Phys. Rev.*, vol. D82, no. 074024, arXiv:1007.2241v3, 2010.

[66] J. Pumplin et al., "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis ", *JHEP*, vol. 07, no. 012, arXiv:hep-ph/0201195v3, 2002.

[67] K. Melnikov and F. Petriello, "Electroweak Gauge Boson Production at Hadron Colliders through $\mathcal{O}(\alpha_s^2)$", *Phys. Rev.*, vol. D74, no. 114017, arXiv:hep-ph/0609070, 2006.

[68] K. Melnikov and F. Petriello, "The W Boson Production Cross Section at the LHC through $\mathcal{O}(\alpha_S^2)$", *Phys. Rev. Lett.*, vol. 96, no. 231803, arXiv:hep-ph/0603182, 2006.

[69] John Campbell and Keith Ellis, "MCFM v5.7 - A Monte Carlo for FeMtobarn Processes at Hadron Colliders - Users Guide", *http://mcfm.fnal.gov/mcfm.pdf*, 2010.

[70] S. Agostinelli et al., "G4 A Simulation Toolkit", *Nucl. Instrum. Meth.*, vol. A506, no. 3, 2003.

[71] J. Allison et al., "Geant4 Developments and Applications", *IEEE Nucl. Sci.*, vol. 53, no. 1, 2006.

[72] K. Zachariadou et al., "The Global Trigger Emulator System for the CMS Experiment", *IEEE Nucl. Sci Symposium Conference Record*, vol. 2, 2004.

[73] R. Field, "Early LHC Underlying Event Data - Findings and Surprises", *arxiv:1010.3558 (The difference between the Z2 tune used in this analysis and the Z1 tune described in the article is just that Z2 uses the CTEQ6L PDF set while Z1 uses the older CTEQ5L PDF set.)*, 2010.

[74] J. Branson et al., "A Cut Based Method for Electron Identificaion at CMS", *CMS AN 2008/082*, 2008.

[75] R. Barlow and C. Beeston, "Fitting Using Finite Monte Carlo Samples", *Comp. Phys. Comm.*, vol. 77, pp. 219–228, 1993.

[76] D. Drijard, "Statistical Methods in Experimental Physics", *North-Holland*, 1971.

[77] N. Adam et al., "Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data", *CMS AN 2009/111*, 2009.

[78] J. Berryhill et al., "Measuring Electron Efficiencies at CMS with Early Data", *CMS AN 2007/019*, 2007.

[79] G. A. Ladinsky and C. P. Yuan, "The Nonperturbative Regime in QCD Resummation for Gauge Boson Production at Hadron Colliders", *Phys. Rev.*, vol. D50, no. 4239, arXiv:hep-ph/9311341, 1994.

[80] C. Balazs et al., "Effects of QCD Resummation on Distributions of Leptons from the Decay of Electroweak Vector Bosons", *Phys. Lett.*, vol. B355, no. 548-554, arXiv:hep-ph/9505203, 1995.

[81] C. Balazs and C. P. Yuan, "Testing Multiple Gluon Dynamics at the Tevatron", *Phys. Rev. Lett.*, vol. 79, no. 23982401, arXiv:hep-ph/9703405, 1997.

[82] C. Balazs and C. P. Yuan, "Soft Gluon Effects on Lepton Pairs at Hadron Colliders", *Phys. Rev.*, vol. D56, pp. 5558–5583, arXiv:hep–ph/9704258, 1997.

[83] R. Brock et al., "Tevatron Run-1 Z Boson Data and Collins-Soper-Sterman Resummation Formalism", *Phys. Rev.*, vol. D67, no. 073016, arXiv:hep-ph/0212159, 2003.

[84] A. V. Konychev and P. M. Nadolsky, "Universality of the Collins-Soper-Sterman Nonperturbative Function in Gauge Boson Production", *Phys. Lett.*, vol. B633, pp. 710–714, arXiv:hep–ph/0506225, 2006.

[85] G. Bauer et al., "Acceptance Uncertainties for the Inclusive W and Z cross section Measurements", *CMS AN 2011/055*, 2011.

[86] A. D. Martin et al., "Parton Distributions for the LHC", *arXiv:0901.0002*, 2009.

[87] R. D. Ball et al, "A First Unbiased Global NLO Determination of Parton Distributions and their Uncertainties", *arXiv:1002.4407v2*, 2010.

[88] PDF4LHC Working Group, "PDF4LHC Recommendations", *http://www.hep.ucl.ac.uk/pdf4lhc/PDF4LHCrecom.pdf*, 2010.

[89] N. Adam et al., "Theoretical Uncertainties in Electroweak Boson Production Cross Sections at 7, 10, and 14 TeV at the LHC", *JHEP*, vol. 11, no. 074, arXiv:1006.3766, 2010.

[90] G. Watt, "Parton Distribution Function Dependence on Benchmark Standard Model Total Cross Sections at the 7 TeV LHC", *CERN-PH-TH*, , no. 149, arXiv:1106.5788v1, 2011.

[91] A. D. Martin et al., "Uncertainties on α_s in Global PDF Analyses and Implications for Predicted Hadronic Cross Sections", *Eur. Phys. J.*, vol. C64, no. 653, 2009.

[92] A. D. Martin et al., "Heavy-Quark Mass Dependence in Global PDF Analyses and 3- and 4-Flavour Parton Distributions", *Eur. Phys. J.*, vol. C70, no. 51, 2010.

[93] N. Adam et al., "Evaluation of the Theoretical Uncertainties in the Z to ll Cross Sections at the LHC", *JHEP*, vol. 05, no. 062, arXiv:0802.3251v6, 2008.

[94] N. Adam et al., "Evaluation of the Theoretical Uncertainties in the W to Lepton and Neutrino Cross Sections at the LHC", *JHEP*, vol. 09, no. 133, arXiv:0808.0758v2, 2008.

[95] K. Mishra, *Private communication*, 2011.

[96] The CMS Collaboration, "Measurement of the Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV", *arXiv:1107.4789 (submitted to JHEP)*, 2010.

[97] S. Baffioni et al., "Measurements of the Inclusive W and Z Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV: Update with full 2010 statistics", *CMS AN 2010/395*, 2010.

[98] M. L. Mangano, "The Super-LHC", *Contemporary Physics*, vol. 51, no. 3, pp. 211–231, arXiv:0910.0030, 2010.

[99] W. T. Eadie et al., "Statistical Methods in Experimental Physics", *Amsterdam: North-Holland*, pp. 269271, ISBN 0444101, 1971.

[100] N. Kanaya and V. Hedberg, "Precision Calibration of the Luminosity Measurement in ATLAS", *Europhysics Conference on High-Energy Physics*, vol. B254, pp. 697–736, 2011.

[101] G. Carboni et al., "Precise Measurements of the Proton-Antiproton and Proton-Proton Total Cross Sections at the CERN Intersecting Storage Rings", *Nucl. Phys.*, vol. B254, pp. 697–736, 1985.

[102] M. Deile et al., "TOTEM: Prospects for Total Cross-Section AND Luminosity Measurements", *CERN-Proceedings*, vol. 1, pp. 7–10, arXiv:1108.0880v1, 2011.

[103] W. Bertl et al., "Feasibility of Intercalibration of CMS ECAL Supermodules with Cosmic Rays", *Eur. Phys. J.*, vol. C41, no. S2.11, 2005.

[104] I. Magrans de Arbril et al., "Conceptual Design of the CMS Trigger Supervisor", *IEEE Trans. Nucl. Sci.*, vol. 53, no. 474, 2006.

[105] The Vector Boson Task Force of the CMS Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV", *CMS AN 2010/116*, 2010.

[106] S. Baffioni et al., "Updated Measurements of Inclusive W and Z Cross Sections at 7 TeV", *CMS AN 2010/264*, 2010.

[107] D. Bandurin et al, "Data Driven Techniques to Estimate the Background in the $Z \to ee$ Events", *CMS AN 2010/277*, 2010.