

Search for the Higgs Boson in the  $ZZ \rightarrow ll\nu\nu$   
Channel with CMS

by

Kevin Kai Hong Sung

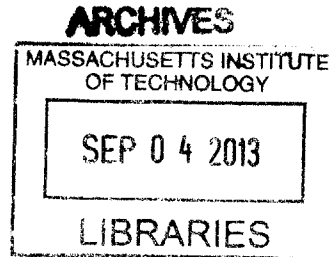
Submitted to the Department of Physics  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013




© Massachusetts Institute of Technology 2013. All rights reserved.

Author .....  .....

Department of Physics

March 11, 2013


Certified by .....  .....

Steve Nahn

Associate Professor

Thesis Supervisor



Accepted by ..  .....

John Belcher

Associate Department Head for Education



# Search for the Higgs Boson in the $ZZ \rightarrow \ell\nu\nu$ Channel with CMS

by

Kevin Kai Hong Sung

Submitted to the Department of Physics  
on March 11, 2013, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy in Physics

## Abstract

The Standard Model (SM) of particle physics is a quantum field theory of the strong, weak, and electromagnetic interactions. It successfully describes a large majority of observed phenomena at microscopic scales. The key issue that remains to be addressed is the mechanism for electroweak symmetry breaking (EWSB). Without EWSB in the theory, the massive weak force mediators – the  $W$  and  $Z$  bosons, are required to be massless to satisfy gauge symmetry. To incorporate EWSB into the SM, the Higgs mechanism is invoked and introduces a scalar field, the Higgs field, that couples to the  $W$  and  $Z$  bosons and gives them mass. Additionally, the SM postulates that the fundamental fermions acquire mass through interactions with the Higgs field. Observation of the field quantum, the Higgs boson, would greatly validate our ideas on electroweak symmetry breaking and complete the table of fundamental particles predicted by the Standard Model.

If it exists, the Higgs boson has eluded detection for decades. A primary objective of the Large Hadron Collider (LHC) project is to make a definitive statement about the existence of the Higgs boson. The LHC provided proton-proton collisions at  $\sqrt{s} = 7$  TeV for the 2011 run. The search for the SM Higgs with the 2011 dataset represented the first major foray into the search effort by the LHC project.

This thesis presents a search for the SM Higgs boson in the  $ZZ \rightarrow 2\ell 2\nu$  channel from the analysis of  $4.9 \text{ fb}^{-1}$  of data collected with the Compact Muon Solenoid (CMS) detector at the LHC. Overviews of electroweak theory and of Higgs production at the LHC are given, followed by descriptions of the CMS detector and the algorithms for event reconstruction. A measurement of the inclusive cross section for  $Z$  boson production is presented as a validation on the measured efficiencies of the electron and muon reconstruction and selection requirements. Lastly, the Higgs search analysis and results are presented.

Thesis Supervisor: Steve Nahn  
Title: Associate Professor





# Acknowledgments

I wish to thank my advisor, Professor Steve Nahn, for his guidance, encouragement, and patience. His enthusiasm and energy is infectious. Being part of the MIT CMS group was a real pleasure, not only because of the wealth of experience and expertise, but also because of the great camaraderie. I was fortunate enough to share an office and to explore physics, Geneva, and the Alps with Philip Harris and Pieter Everaerts.

To my Grade 5 teacher, Mrs. Scrimmes, who acquainted me with the fascinating world of particles.

To Daniel Kraemer, Anirban Mazumdar, Kimi Shirasaki, Yasuhiro Shirasaki, and Victor Wang, for all the laughs and support. I am especially grateful to Kimi and Yasuhiro for their generous hospitality during the final months of completing my degree.

To my sister, Connie.

To my parents, Paul and Ellen, for affording me every opportunity to pursue my interests. I dedicate this thesis to them.



# Contents

<b>1</b>	<b>Introduction</b>	<b>11</b>
1.1	Theory Synopsis of the Standard Model . . . . .	11
1.1.1	Electroweak Interactions . . . . .	12
1.1.2	Higgs Mechanism . . . . .	13
1.1.3	The Standard Model Higgs Boson . . . . .	14
1.2	Large Hadron Collider . . . . .	15
1.2.1	LHC Operation in 2011 . . . . .	16
1.2.2	Higgs Production at the LHC . . . . .	17
<b>2</b>	<b>The Compact Muon Solenoid Detector</b>	<b>21</b>
2.1	Superconducting Magnet . . . . .	23
2.2	Tracker . . . . .	24
2.2.1	Pixel Tracker . . . . .	25
2.2.2	Strip Tracker . . . . .	26
2.3	Electromagnetic Calorimeter . . . . .	28
2.3.1	ECAL Barrel . . . . .	30
2.3.2	ECAL Endcap . . . . .	31
2.3.3	Preshower . . . . .	31
2.3.4	Crystal Transparency . . . . .	32
2.3.5	Energy Resolution . . . . .	32
2.4	Hadronic Calorimeter . . . . .	33
2.4.1	HCAL Barrel . . . . .	33
2.4.2	HCAL Endcap . . . . .	34

2.4.3	HCAL Outer . . . . .	35
2.4.4	HCAL Forward . . . . .	35
2.5	Muon System . . . . .	35
2.5.1	Drift Tube . . . . .	36
2.5.2	Cathode Strip Chambers . . . . .	36
2.5.3	Resistive Plate Chambers . . . . .	38
2.6	Forward Calorimeters . . . . .	38
2.7	Trigger . . . . .	39
<b>3</b>	<b>Event Reconstruction</b>	<b>41</b>
3.1	Track Reconstruction . . . . .	41
3.2	Primary Vertex Reconstruction . . . . .	43
3.3	Supercluster Reconstruction . . . . .	44
3.4	Photons . . . . .	45
3.4.1	Photon Conversions . . . . .	45
3.5	Electrons . . . . .	46
3.6	Muons . . . . .	47
3.7	Particle Flow event description . . . . .	48
3.7.1	Jets . . . . .	49
3.7.2	Missing Transverse Energy . . . . .	50
3.7.3	Simulation . . . . .	50
<b>4</b>	<b>Inclusive <math>Z</math> cross section via decays to electron or muon pairs</b>	<b>53</b>
4.1	Introduction . . . . .	53
4.2	Event Selection . . . . .	54
4.2.1	Trigger Requirements . . . . .	55
4.2.2	Isolation . . . . .	56
4.2.3	Impact Parameter . . . . .	58
4.2.4	Electron Identification . . . . .	59
4.2.5	Muon Identification . . . . .	61
4.2.6	Yields . . . . .	62

4.2.7	Distributions . . . . .	64
4.3	Acceptance . . . . .	71
4.3.1	QCD resummation and NNLO corrections . . . . .	71
4.3.2	Higher order QCD corrections . . . . .	72
4.3.3	Parton distribution function . . . . .	73
4.3.4	Electroweak corrections . . . . .	74
4.3.5	Summary of Systematic Uncertainties on the Acceptance . . . . .	75
4.4	Lepton Efficiency . . . . .	75
4.4.1	Tag-and-Probe Method . . . . .	76
4.4.2	Measured Efficiencies and Scale Factors . . . . .	78
4.4.3	Efficiency Corrections to the Acceptance . . . . .	84
4.4.4	Systematic Uncertainties on the Efficiency . . . . .	86
4.5	Luminosity . . . . .	87
4.6	Results . . . . .	88
<b>5</b>	<b>Search for <math>H \rightarrow ZZ \rightarrow 2\ell 2\nu</math></b>	<b>91</b>
5.1	Introduction . . . . .	91
5.2	Event Selection . . . . .	92
5.3	Background Estimation . . . . .	94
5.3.1	Top and $WW$ backgrounds . . . . .	94
5.3.2	Single $Z$ +jets background . . . . .	95
5.4	Systematic Uncertainties . . . . .	103
5.4.1	Theoretical Uncertainties on Yields from Simulation . . . . .	103
5.4.2	Summary of Systematic Uncertainties on Normalisation . . . . .	104
5.4.3	$M_T$ Shape Uncertainties . . . . .	104
5.5	Results . . . . .	107
5.5.1	Yields . . . . .	107
5.5.2	Statistical Analysis . . . . .	111
5.5.3	Limits . . . . .	113
<b>6</b>	<b>Conclusion</b>	<b>115</b>

<b>A</b>	<b>Electron Energy Scale and Resolution</b>	<b>117</b>
<b>B</b>	<b>Lepton Efficiency Fits and Tables</b>	<b>121</b>
B.1	Electron Efficiencies . . . . .	121
B.1.1	GSF Reconstruction Efficiency . . . . .	121
B.1.2	Electron Selection Efficiency . . . . .	133
B.2	Muon Efficiencies . . . . .	144
B.2.1	Standalone Muon Efficiency . . . . .	144
B.2.2	Muon Track Efficiency . . . . .	155
B.2.3	Muon Selection Efficiency . . . . .	160

# Chapter 1

## Introduction

### 1.1 Theory Synopsis of the Standard Model

The Standard Model has been built over the last century through the interplay between theoretical postulation and experimental observation, and is the subject of many textbooks (such as [1] and [2]). Rather than repeating a thorough exposition on the theory, only the basic building blocks of electroweak theory [3–5] and crucial elements to the Higgs mechanism [6–8] are described.

The SM description of particle physics begins with the fundamental quarks and leptons,

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$$
$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}.$$

The theory is constructed under the  $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$  symmetry group. More specifically, the colour dynamics governing strong interactions of quarks and gluons exhibit  $SU(3)$  symmetry (hence the subscript “c”), while the electroweak interactions are described by  $SU(2)$  symmetry of weak isospin coupling “left-handed” chiral states only and  $U(1)$  symmetry of weak hypercharge (hence the “L” and “Y” subscripts).

### 1.1.1 Electroweak Interactions

The framework of electroweak theory can be outlined using the  $(\nu_e, e^-)$ -generation as an example. Consider a left-handed weak isospin doublet and a right-handed weak isospin singlet,

$$\mathbf{L} = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad \text{and} \quad \mathbf{R} = e_R, \quad (1.1)$$

with weak hypercharge  $Y_L = -1$  and  $Y_R = -2$ , respectively. Gauge fields must be introduced to satisfy local gauge invariance: a weak isovector  $\mathbf{W}_\mu$  with coupling constant  $g$  corresponding to  $SU(2)_L$ , and a weak isoscalar  $B_\mu$  with coupling constant  $g'$  corresponding to  $U(1)_Y$ . These lead to the tensor for weak isospin,

$$F_{\mu\nu}^\ell = \partial_\nu W_\mu^\ell - \partial_\mu W_\nu^\ell + g\varepsilon_{jkl}W_\mu^jW_\nu^k, \quad (1.2)$$

where  $\varepsilon_{jkl}$  is the Levi-Civita symbol, and the tensor for weak hypercharge,

$$f_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu. \quad (1.3)$$

The Lagrangian describing the electroweak interaction within the electron generation is,

$$\begin{aligned} \mathcal{L} = & \bar{\mathbf{R}} i\gamma^\mu \left( \partial_\mu + i\frac{g'}{2}B_\mu Y \right) \mathbf{R} \\ & + \bar{\mathbf{L}} i\gamma^\mu \left( \partial_\mu + i\frac{g'}{2}B_\mu Y + i\frac{g}{2}\boldsymbol{\tau} \cdot \mathbf{W}_\mu \right) \mathbf{L} \\ & - \frac{1}{4}F_{\mu\nu}^\ell F^{\ell\mu\nu} - \frac{1}{4}f_{\mu\nu}f^{\mu\nu}, \end{aligned} \quad (1.4)$$

where  $\boldsymbol{\tau}$  denote the Pauli spin matrices and  $Y$  is the hypercharge operator. While the above formulation unifies the weak and electromagnetic forces, the requirement of local gauge symmetry implies the quarks, leptons, and gauge bosons have to be massless (i.e. no mass terms involving  $\bar{e}e$ ,  $B^\mu B_\mu$ ,  $W_\mu^\ell W^{\ell\mu}$ ). This symmetry is spontaneously broken via the Higgs mechanism, allowing the bosons to acquire mass.



### 1.1.2 Higgs Mechanism

In the simplest scheme that implements the Higgs mechanism, a complex doublet of scalar fields is introduced into the theory,

$$\phi \equiv \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \quad (1.5)$$

with hypercharge  $Y_\phi = +1$ . Define the gauge-covariant derivative,

$$\mathcal{D}_\mu = \partial_\mu + i\frac{g'}{2}B_\mu Y + i\frac{g}{2}\boldsymbol{\tau} \cdot \mathbf{W}_\mu, \quad (1.6)$$

and consider Lagrangian terms

$$(\mathcal{D}^\mu \phi)^\dagger (\mathcal{D}_\mu \phi) - V(\phi^\dagger \phi), \quad (1.7)$$

for some interaction potential  $V$ . There is no prediction for what this potential should be, but to demonstrate the procedure of EWSB, the typical form for  $V$  is chosen to be,

$$V(\phi^\dagger \phi) = -\mu^2 (\phi^\dagger \phi) + \lambda^2 (\phi^\dagger \phi)^2. \quad (1.8)$$

The potential is at a minimum whenever  $\phi^\dagger \phi = \frac{\mu^2}{2\lambda^2}$ . A convenient choice for the vacuum state is

$$\langle \phi \rangle_0 = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}, \quad (1.9)$$

where  $v = \mu/\lambda$  is the Higgs vacuum expectation value. Re-expressing the scalar field in terms of deviations from the vacuum ground state,  $H \equiv \phi_1 - v/\sqrt{2}$ , results in explicit mass terms in the Lagrangian for the electroweak gauge fields. The quantity  $v$  sets the electroweak scale and is related to the Fermi constant by

$$v = \left(G_F \sqrt{2}\right)^{-\frac{1}{2}} \approx 246 \text{ GeV}. \quad (1.10)$$

The vacuum state  $\langle\phi\rangle_0$  is no longer invariant under symmetry operations generated by  $\tau$  and  $Y$ , but remains invariant under the symmetry generated by  $\frac{1}{2}(\tau_3 + Y) = Q$ , namely the electric charge operator. Consequently, the weak gauge bosons acquire masses while the photon remains massless. The  $W$  bosons are represented by

$$W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp iW^2), \quad (1.11)$$

with  $W$  mass given by  $M_W = gv/2$ . The  $Z$  boson and photon arise from the mixing of isospin vector component  $W^3$  and hypercharge scalar  $B$ , the degree of mixing parametrised by the weak mixing angle  $\theta_W$ , so that

$$Z = W^3 \cos \theta_W - B \sin \theta_W \quad \text{and} \quad A = W^3 \sin \theta_W + B \cos \theta_W, \quad (1.12)$$

with  $Z$  mass given by  $M_Z = M_W / \cos \theta$ .

### 1.1.3 The Standard Model Higgs Boson

From the choice of the Higgs potential in the preceding discussion, the mass of the Higgs boson is given by  $m_H = \sqrt{2}\mu$ . However, this is not really a precise prediction because the true form of the Higgs potential is unknown. While the Higgs mechanism provides the weak gauge bosons with mass, the Standard Model further postulates that the fundamental fermions (except, possibly, neutrinos) acquire mass,  $m$ , via Yukawa coupling with the Higgs field,

$$-\frac{m\sqrt{2}}{v} [\bar{\mathbf{R}} (\phi^\dagger \mathbf{L}) + (\bar{\mathbf{L}} \phi) \mathbf{R}]. \quad (1.13)$$

With all its couplings defined, the properties of the SM Higgs only depend on its mass value. Of considerable interest to experimentalists are the mass dependent branching ratios for Higgs decays, as summarised in Figure 1-1 [9]. Note that there is no  $Hgg$ ,  $H\gamma\gamma$ , or  $HZ\gamma$  vertex, so the decay into these final states proceed indirectly through a quark or weak boson loop.

Even though the Standard Model has proven to be a successful model of the

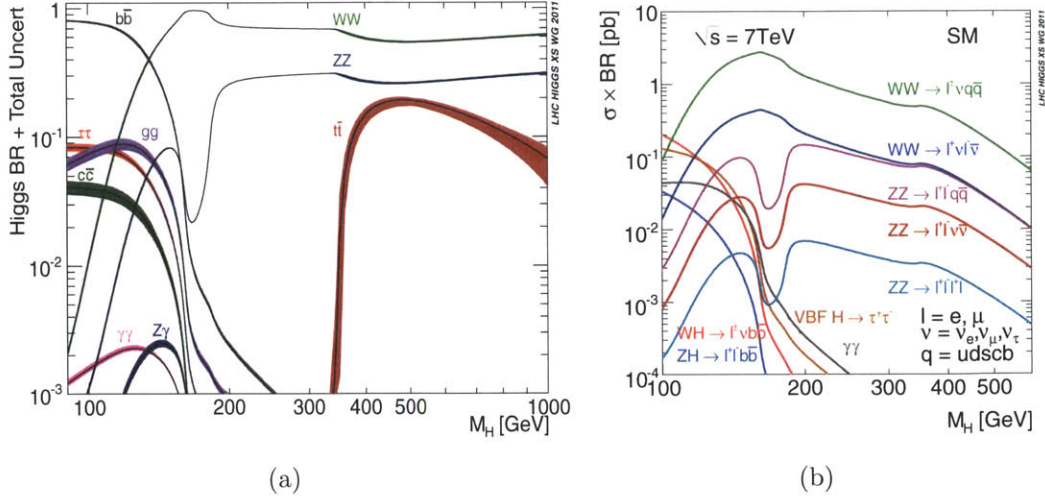


Figure 1-1: (a) Branching ratios of different SM Higgs decays (b) Cross section times branching ratios of Higgs decays separated by different experimental final states.

particle world, it is not without critical shortcomings, leading to various theoretical extensions with consequences for the Higgs sector. One predicament in the SM is that the Higgs boson mass receives quadratically divergent contributions from radiative corrections. A solution to this problem is supersymmetry [10, 11], and the simplest SM extension is the Minimal Supersymmetric Standard Model (MSSM). In MSSM, there are two Higgs doublets which results in five scalar bosons [12, 13]. Another SM extension, the Inert Higgs Doublet model [14–16], asserts a second Higgs doublet that does not couple to quarks or leptons, and therefore the model presents a candidate for dark matter, whereas the SM provides no such candidates.

## 1.2 Large Hadron Collider

Located at the facilities of the European Organization for Nuclear Research (CERN), the Large Hadron Collider (LHC) [17] is a synchrotron machine designed to explore particle physics at the highest collision energies yet, with hopes of providing clues to electroweak symmetry breaking and physics beyond the Standard Model. The circular accelerator is 27 km long and runs about 100 m underground through the countryside near the Franco-Swiss border between Lake Geneva and the Jura mountains. The

LHC was constructed with capabilities to circulate beams of protons or beams of lead nuclei for collisions. For proton-proton collisions, the design peak luminosity is  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  with a mere 25 ns between proton bunches. Guiding the protons through the LHC are 1232 superconducting dipole magnets; the magnets generate fields up to 8.33 T and are cooled to below 2 K with superfluid helium. Two counter-rotating beams circulate in the accelerator and are deflected into collision at four interaction regions, delivering data to four experiments: ATLAS, CMS, ALICE, and LHCb (Figure 1-2).

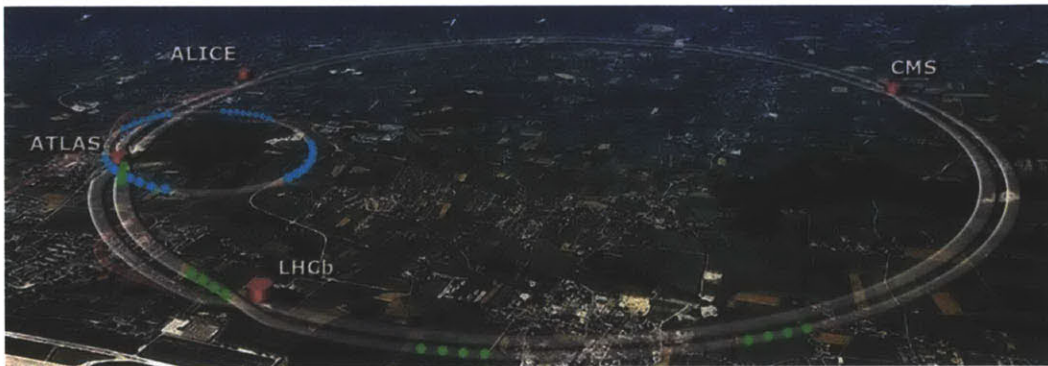


Figure 1-2: The LHC accelerator and detectors (ATLAS, CMS, ALICE, LHCb) superimposed over an aerial view of the Swiss and French regions.

### 1.2.1 LHC Operation in 2011

The LHC delivered over  $6 \text{ fb}^{-1}$  of proton-proton collision data from March until the end of October in 2011. Machine operations progressed through the year with instantaneous luminosities on the order of  $10^{32} \text{ cm}^{-2} \text{ s}^{-1}$  at the beginning of the year to over  $3 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  by the end. The integrated luminosity provided by the accelerator over the year is charted in Figure 1-3.

With about  $10^{11}$  protons per bunch, it is not uncommon for multiple proton-proton interactions to occur in each bunch crossing. This introduces complications for data analysis, as signal events may be obscured by several extraneous collisions that deposit energy in the detector that is uncorrelated with the hard scatter interaction. These additional inelastic collisions that accompany the hard interaction are called

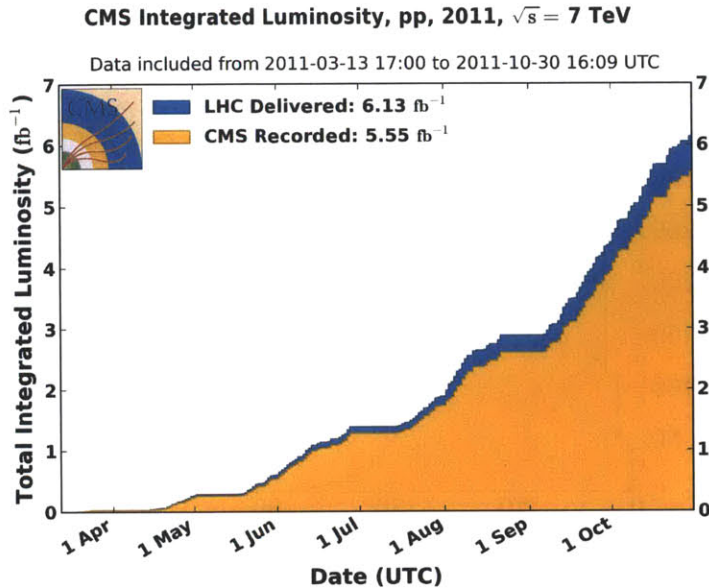


Figure 1-3: Cumulative integrated luminosity delivered over 2011.

“pile-up” (PU) interactions. Figure 1-4 shows the distribution of pile-up interactions averaged over the 2011 data. The distribution is computed from the product of the instantaneous luminosity and the expected cross section for inelastic proton-proton collision, integrated over the data taking time period.

### 1.2.2 Higgs Production at the LHC

The different mechanisms for Higgs production accessible at the LHC are depicted in Figure 1-5. The primary Higgs production mode at the LHC is gluon fusion (GF) via a quark loop, with the largest contribution from the top quark, due to its large coupling to the Higgs. Another important production channel is vector boson fusion (VBF), where the Higgs is produced in association with two hard jets in the forward regions of the detector. The “Higgs-strahlung” mechanism, where the Higgs is produced in association with a  $W$  or  $Z$  boson is relevant for low mass Higgs searches, but the expected cross section drops relatively quickly with increasing Higgs mass. A smaller contribution still comes from Higgs production in association with top-antitop production, whose role quickly diminishes for higher Higgs masses. The



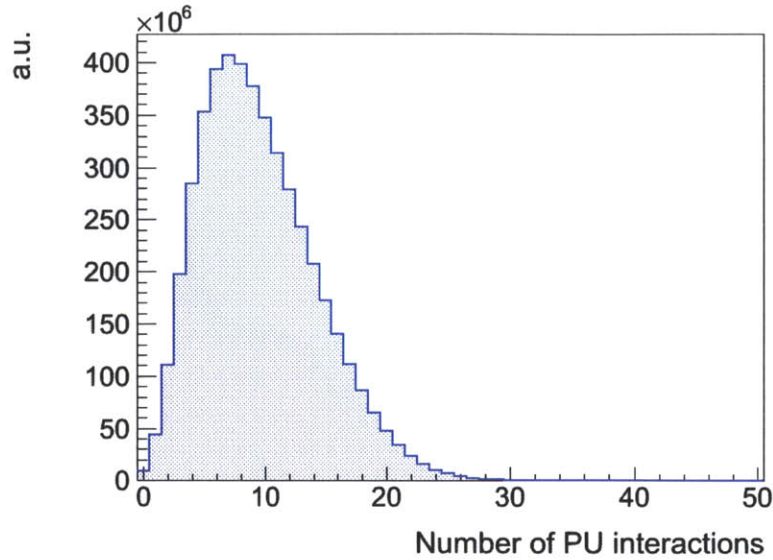


Figure 1-4: Distribution of the average expected number of pile-up interactions in 2011.

predicted Higgs production cross sections as a function of mass in each mode are shown in Figure 1-6 [9].

With the tremendous upgrade in collision energy and rate compared to previous accelerators, the LHC is expected to address the existence of the Higgs boson up to  $\mathcal{O}(1 \text{ TeV}/c^2)$  in mass. Its predecessor, the Large Electron-Positron Collider (LEP), had excluded a Higgs boson with mass below  $114.4 \text{ GeV}/c^2$  at the 95% confidence level [18]. Prior to the conclusion of the 2011 run of the LHC, the Tevatron experiment reported an exclusion for a mass between  $156 \text{ GeV}/c^2$  and  $177 \text{ GeV}/c^2$  [19].

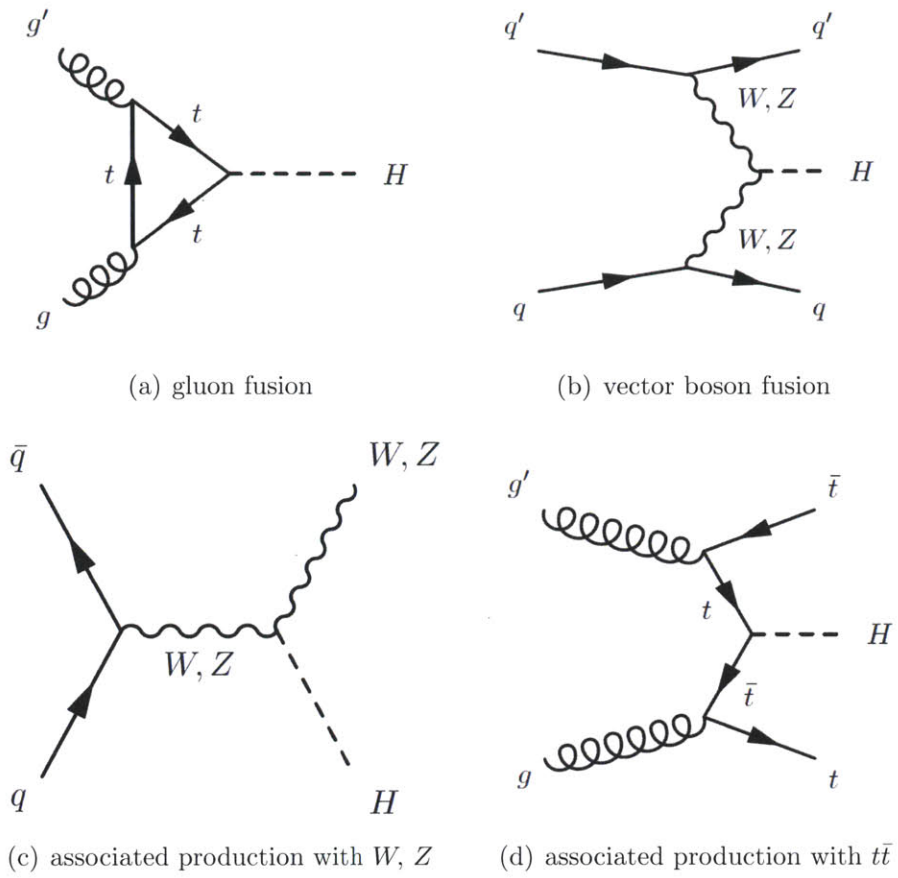


Figure 1-5: Higgs production modes.

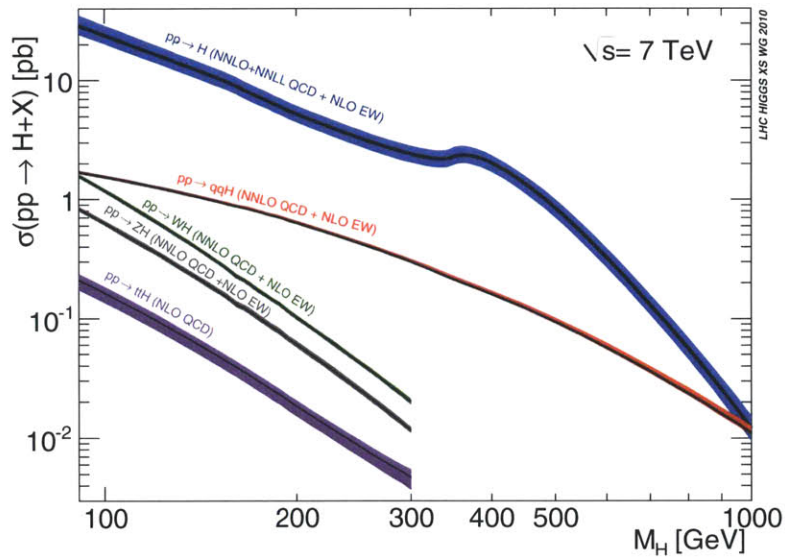


Figure 1-6: Higgs production cross section as a function of Higgs boson mass.





# Chapter 2

## The Compact Muon Solenoid Detector

The Compact Muon Solenoid detector is a general purpose detector on the LHC located 100 m underground near the French village of Cessy. The CMS detector was designed for studying fundamental particle physics in proton collisions as well as heavy ion physics in lead nuclei collisions. It is essentially a cylindrical apparatus that is 21.6 m in length with a diameter of 14.6 m and weighs 12 500 tonnes. A cutaway view of the CMS detector is provided in Figure 2-1. The CMS detector features a superconducting solenoid magnet that produces a very large 3.8 T field, a fully silicon-based inner tracking system, a homogeneous electromagnetic calorimeter (ECAL), a sampling hadronic calorimeter (HCAL), and an extensive muon tracking system employing multiple gas detector technologies. Figure 2-2 illustrates the distinct trails left by various particle species traversing through the series of CMS subdetectors, providing the basis for particle identification, which will be discussed in more detail in the chapter on Event Reconstruction. This chapter will be an overview of the detector components. A detailed description can be found in [20].

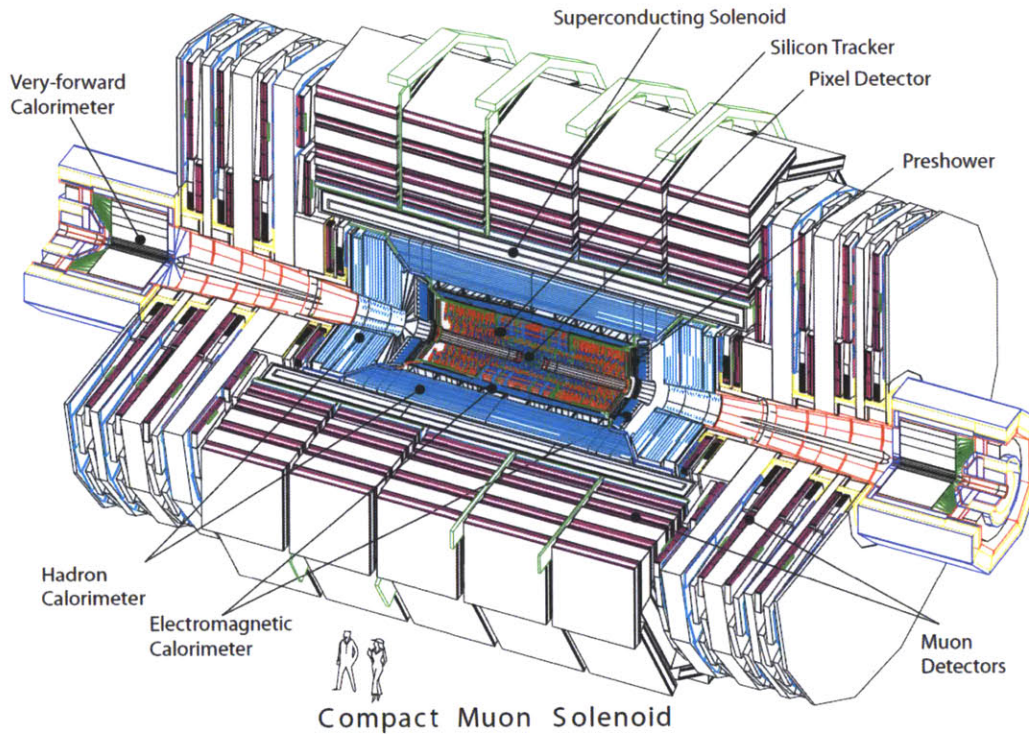


Figure 2-1: A cutaway diagram of the CMS detector.

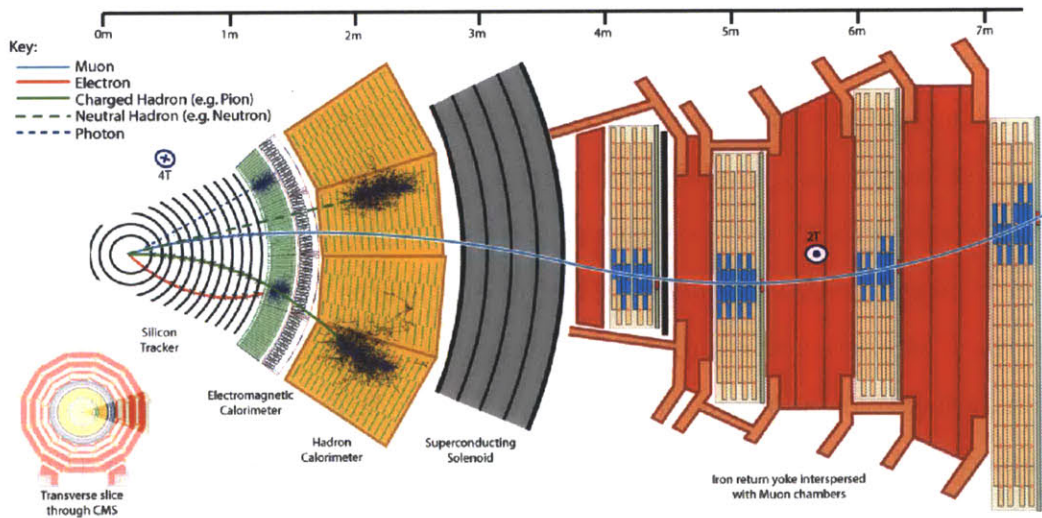


Figure 2-2: The expected signatures of different particles in the CMS detector. Only charged particles will produce a signal in the silicon tracker. Electrons and photons deposit their energies and are stopped in the ECAL. Hadrons deposit their energies and are stopped in the HCAL. Muons traverse beyond the solenoid and into the muon system; the magnetic field outside the solenoid is weaker than and opposite in polarity to the field inside the solenoid, so the muon trajectory changes accordingly.

## 2.1 Superconducting Magnet

The CMS magnet system consists of a solenoid that encases the inner tracker, ECAL, and HCAL, and an iron yoke that embeds the muon system and controls the return flux outside the solenoid. The solenoid has conducting wires winding in a 12.5 m long cylindrical cold mass with a bore diameter of 6.3 m. To produce the high magnetic field desired, there are four layers of winding conductors – instead of one layer of winding typically found in other experiments. During standard operations, a 3.8 T axial field is produced and the cold mass is cooled with liquid helium. The magnetic field is the strongest yet produced in a particle detector. The magnetic energy stored in the 220 tonne cold mass is 2.6 GJ, giving an energy-to-mass ( $E/M$ ) ratio of 11.6 kJ/kg; this puts the CMS magnet system beyond all current and previous detector magnets in terms of energy stored and  $E/M$ . A comparison of the CMS solenoid magnet with other detector magnets is given in Figure 2-3 [20]. A higher  $E/M$  value for the same magnetic energy means less passive material in the detector to degrade measurements for the same field strength. The  $E/M$  ratio is also proportional to the stress on the coils, so that a strong but lightweight support structure is desired; CMS uses an aluminum alloy to reinforce the conductor. The iron yoke weighs 10 000 tonnes, making up most of the total detector mass. The return field is half of the field strength inside the solenoid (and of opposite polarity).

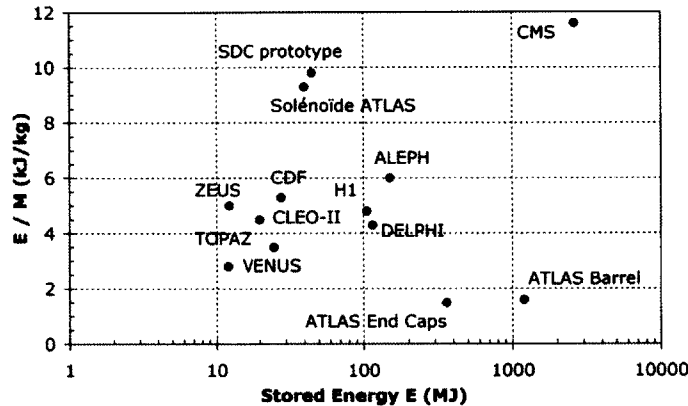


Figure 2-3: Comparison of energy stored and  $E/M$  for different detector magnets.

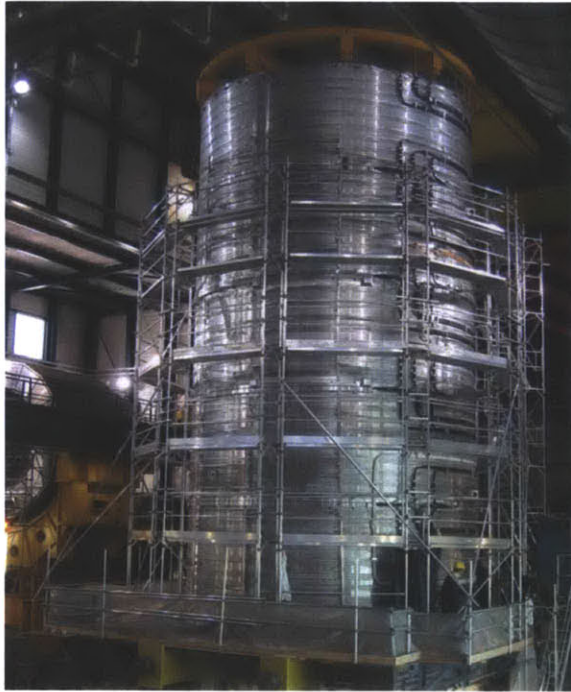


Figure 2-4: A photo of the cold mass (mounted vertically) prior to integration into the CMS detector.

## 2.2 Tracker

The inner tracking system is the subdetector closest to the interaction point. The tracker is 5.8 m in length with a diameter of 2.5 m. Situated well within the solenoid, the entire tracker volume is immersed in a uniform magnetic field. The role of the tracker is to provide measurements along the trajectory of charged particles to estimate their momenta. The tracker covers the pseudorapidity range  $|\eta| < 2.5$ , where  $\eta$  is defined based on the polar angle with respect to the beam line,  $\theta$ , by  $\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right]$ . The tracker consists of concentric cylindrical layers of sensors around the beam line in the barrel region, or circular disks perpendicular to the beam line in the endcap regions. A diagram illustrating the layout of the CMS tracker is shown in Figure 2-5.

The CMS tracker employs only silicon sensors; with over  $200 \text{ m}^2$  of active silicon, it is the largest silicon tracker ever built. A charged particle crossing a sensor deposits energy that produces electron-hole pairs in the silicon. Under an applied electric field,

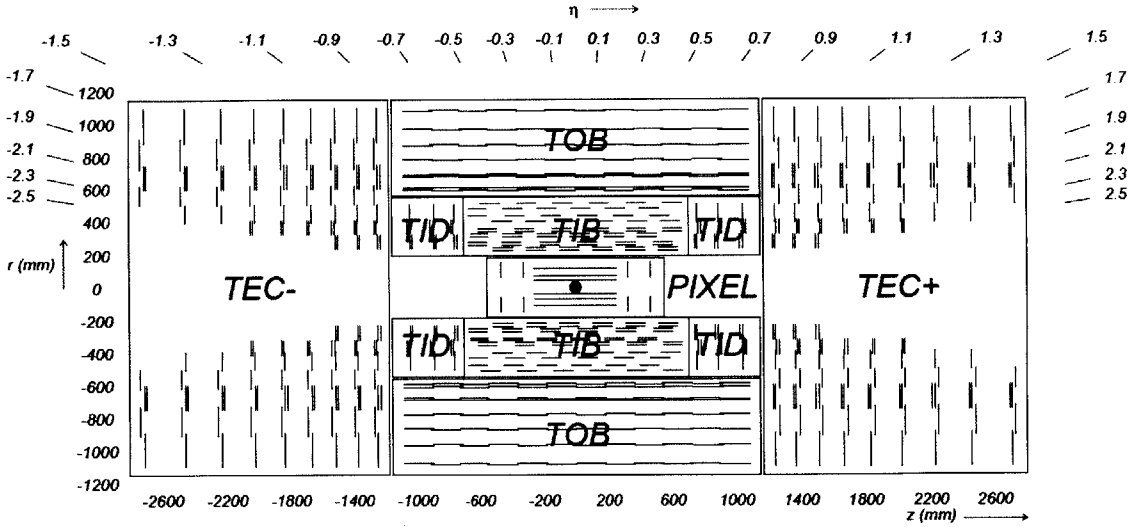


Figure 2-5: The layout of CMS tracker in the  $r - z$  plane. The direction from the interaction point for various values of  $\eta$  are indicated by the numbers at the top of the diagram.

the charge carriers move towards collection electrodes and induce a signal current. Silicon sensor technology was adopted because of its fast response, high granularity, and radiation hardness. There are two major components to the CMS tracker: the innermost layers of the tracker employ finer grained pixel sensors, while the rest of the tracker employ coarser grained strip sensors. To be effective, the tracker must operate at low temperatures (around  $4^{\circ}\text{C}$  to  $7^{\circ}\text{C}$ ) and this is achieved with a combination of liquid coolant ( $\text{C}_6\text{F}_{14}$ ) and pre-chilled gas ( $\text{N}_2$ ).

The expected performance for tracking is shown in Figure 2-6 [21]. The relevant lepton momentum scale for a Standard Model Higgs search is  $\mathcal{O}(10 \text{ GeV}/c) - \mathcal{O}(100 \text{ GeV}/c)$ ; the momentum resolution is a few percents and the impact parameter resolution is  $\mathcal{O}(10 \mu\text{m})$ .

### 2.2.1 Pixel Tracker

Within 15 cm of the beam line, the high particle flux necessitates fine grained sensors in order to distinguish between particles. The silicon sensors are divided into  $100 \times 150 \mu\text{m}^2$  units (or “pixels”), designed such that the expected fraction of pixels



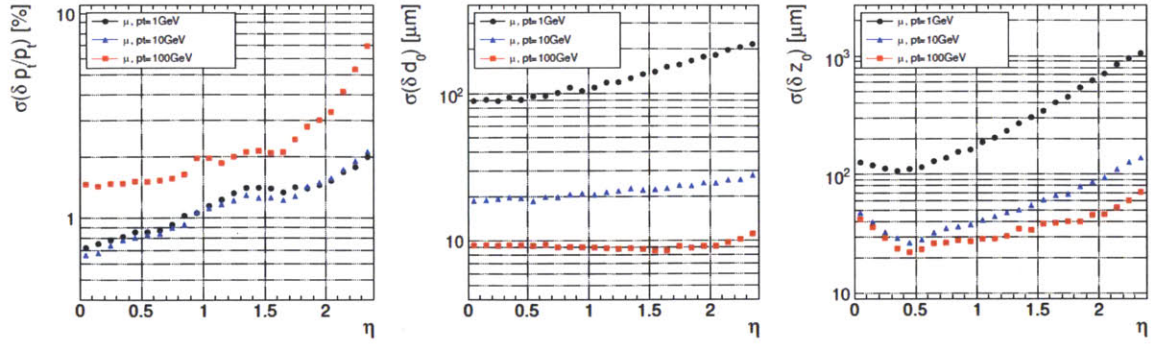


Figure 2-6: Expected resolution of track parameters as a function of  $\eta$  for muons with transverse momenta of 1 GeV/c (black circles), 10 GeV/c (blue triangles), and 100 GeV/c (red squares) *Left*: Transverse momentum resolution. *Middle*: Transverse impact parameter resolution. *Right*: Longitudinal impact parameters resolution.

that register a signal (typically termed *occupancy*) is about one percent. Figure 2-7 illustrates the layout of a pixel sensor on a readout chip. The pixel tracker consists of three barrel layers and two endcap disks at each end, altogether containing about 66 million pixels. The barrel layers are 53 cm long at radial distances of 4.4 cm, 7.3 cm, and 10.2 cm away from the beam line. The endcap disks are placed along the beam line at  $\pm 4.5$  cm,  $\pm 46.5$  cm from the expected interaction point, and have radial coverage between 6 cm to 15 cm. The pixel tracker has pseudorapidity coverage of  $|\eta| < 2.5$ . The layout of the pixel layers and disks are shown in the upper plot of Figure 2-8 [20]. The lower plot of Figure 2-8 shows the expected efficiency to record two or more hits as a function of  $\eta$ ; pairs and triplets of hits in the pixel tracker are vital to initiate the algorithms to reconstruct tracks. The resolution on the position of each hit from a charged particle crossing the layers of the pixel tracker is between 15  $\mu\text{m}$  to 20  $\mu\text{m}$ .

## 2.2.2 Strip Tracker

Beyond 15 cm from the beam line, the particle flux is sufficiently low so that the silicon sensors can be partitioned into larger units called, “strips”, and still maintain occupancy of no more than a few percent. A photo of the barrel component taken during construction (with some artistic lighting) is shown in Figure 2-9. The strip

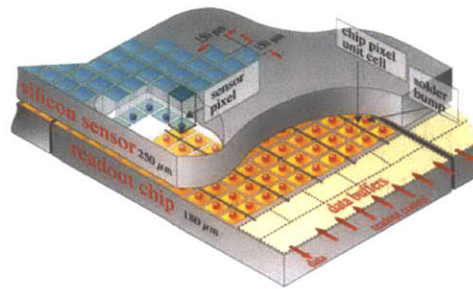


Figure 2-7: Illustration of a sensor with pixels on a readout chip.

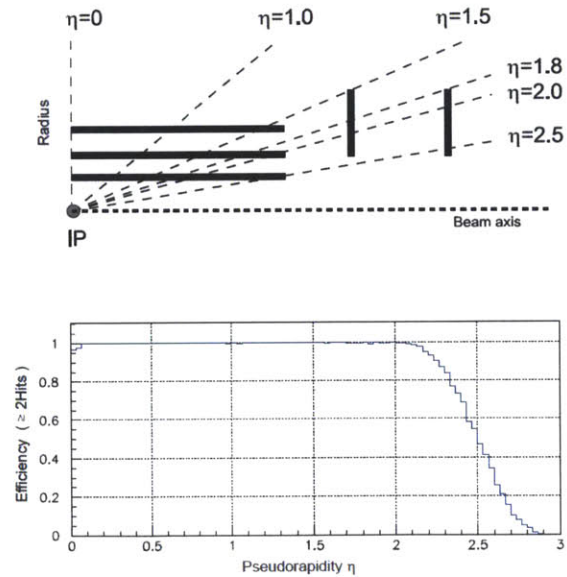


Figure 2-8: The layout of the barrel layers and endcap disks of the CMS pixel tracker and the expected coverage in  $\eta$ .

tracker consists of 10 barrel layers and 12 endcap disks, for a total of 9.3 million strips. The barrel of the strip tracker extends out to a radial distance of 110 cm away from beam line, while the endcaps extend out to  $z = \pm 282$  cm from the interaction point. The pitch of the strips range from  $80\ \mu\text{m}$  to  $184\ \mu\text{m}$ , where the pitch increases with distance away from the interaction point. The position resolution varies with strip pitch and is in the range of  $23\ \mu\text{m}$  to  $53\ \mu\text{m}$ . The layout of a strip tracker sensor is shown in Figure 2-10.

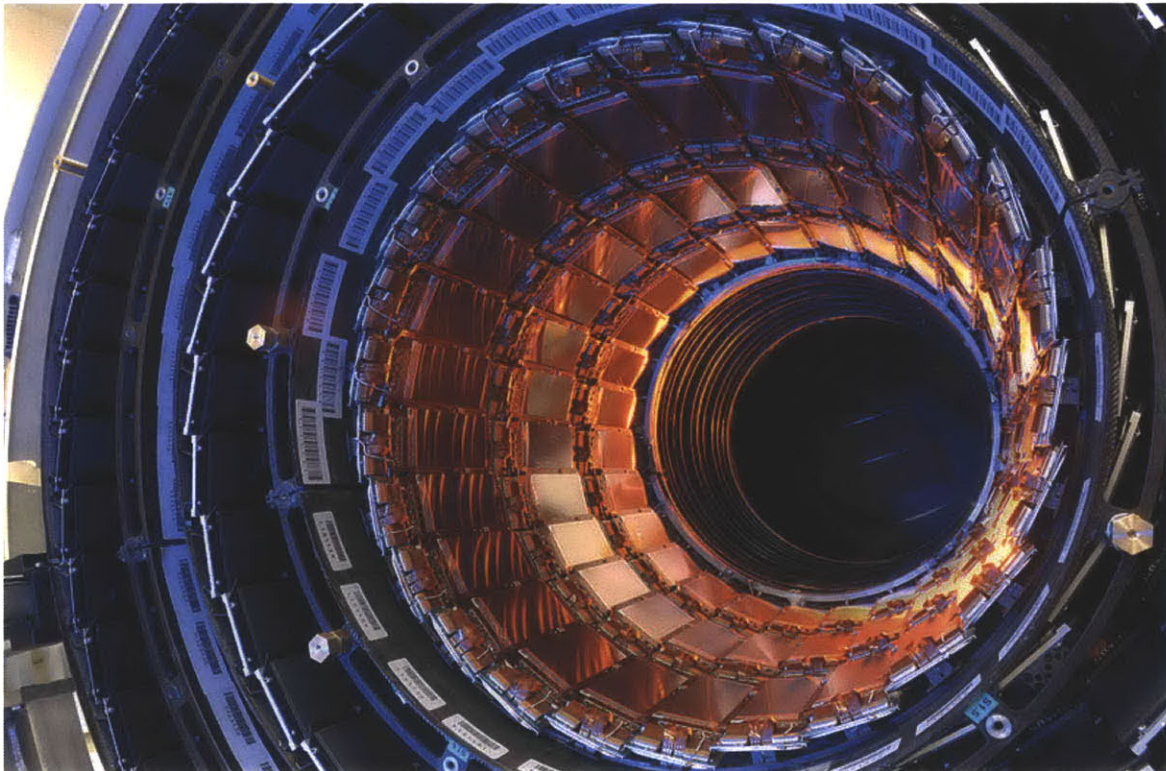


Figure 2-9: Photo of some barrel layers of the CMS strip tracker during construction.

## 2.3 Electromagnetic Calorimeter

The CMS ECAL is a homogeneous calorimeter constructed from scintillating lead tungstate ( $\text{PbWO}_4$ ) crystals. Lead tungstate is a relatively dense material ( $\rho = 8.28\ \text{g/cm}^3$ ) with a short radiation length  $X_0 = 0.89\ \text{cm}$  and Moliere radius  $X_M = 2.2\ \text{cm}$ . The crystals have fast decay times: approximately 80% of the scintillation



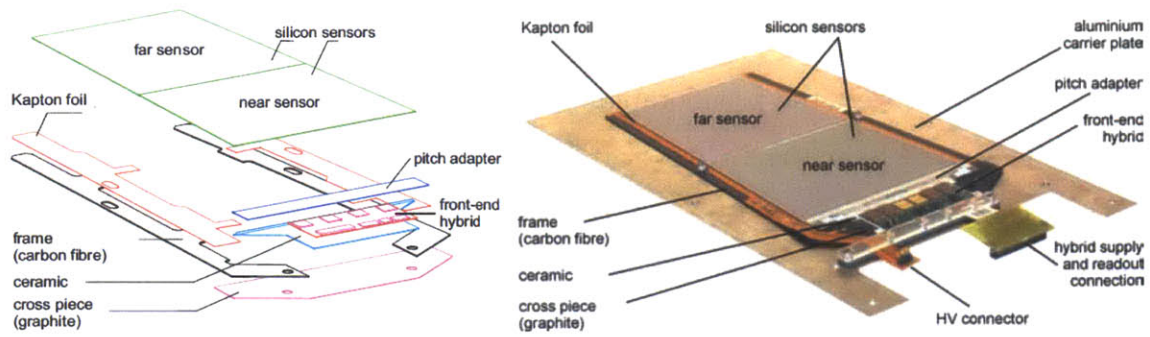


Figure 2-10: Layout of a strip tracker sensor.

light is emitted in 25 ns. The crystals emit a blue-green scintillation light that is collected by photodetectors attached to the crystal and the emission spectrum has a broad peak around 420 – 430 nm. At the operating temperature of 18 °C, an average of 4.5 photoelectrons are generated per MeV. The CMS ECAL consists of a barrel section, an endcap section on each side, and a preshower detector placed just in front of each endcap.

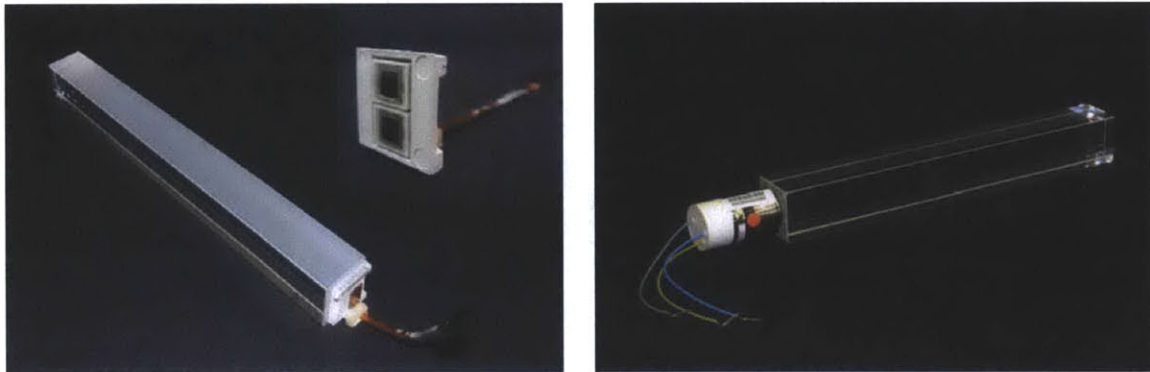


Figure 2-11: Photos of ECAL scintillating crystals. *Left:* A crystal for the ECAL Barrel attached to an APD (zoomed in view of APD in inset). *Right:* A crystal for the ECAL Endcap attached to a VPT.

### 2.3.1 ECAL Barrel

The ECAL Barrel covers the pseudorapidity region  $|\eta| < 1.479$  and contains 61 200 crystals. The crystals in the ECAL Barrel are organized into 36 “supermodules” of 1700 crystals (see Figure 2-12), each covering about  $20^\circ$  in  $\phi$ . The crystals are truncated pyramidal in shape, with front face dimensions of  $22 \times 22 \text{ mm}^2$  and back face dimensions of  $26 \times 26 \text{ mm}^2$  – essentially on the order of one Moliere radius, and cover approximately  $0.0174 \times 0.0174$  in  $\eta - \phi$  space. The crystals are 23 cm long, equivalent to  $25.8X_0$  in radiation lengths. The ECAL Barrel begins at a radial distance of 1.29 m from the beam line. The readout is performed using avalanche photodiodes (APD), which are silicon photodetectors configured to amplify a signal initiated by a photon through the process of an avalanche breakdown [22,23]. A pair of APDs is attached to the back face of each barrel crystal. A lead tungstate crystal and APD unit is shown in Figure 2-11.

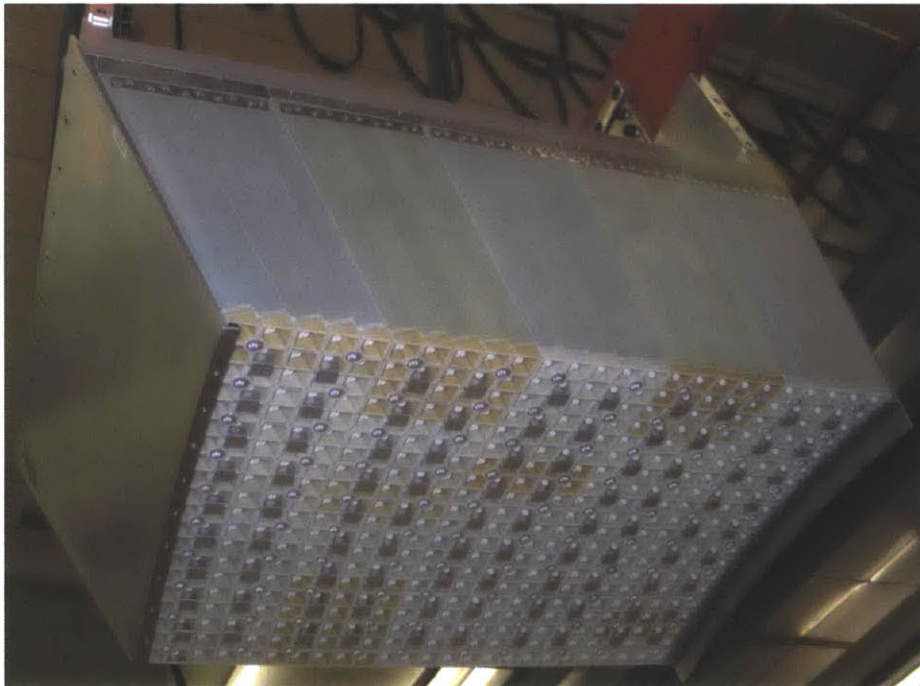


Figure 2-12: A supermodule of the ECAL barrel consisting of 1700 crystals.

### 2.3.2 ECAL Endcap

The ECAL Endcap covers the pseudorapidity region  $1.479 < |\eta| < 3$  and contains 7324 crystals per endcap, organised into a pair of “Dees” (see Figure 2-13). The crystals are  $28.62 \times 28.62 \text{ mm}^2$  on the front face and  $30 \times 30 \text{ mm}^2$  on the back face. The crystals are 22 cm long, equivalent to  $24.7X_0$  in radiation lengths. The ECAL Endcaps are located at  $z = \pm 315.4 \text{ cm}$  from the interaction point. The readout is performed using vacuum phototriodes (VPT), which are single-stage photomultipliers [22]. A single VPT is attached to the back face of each endcap crystal. A crystal and VPT unit is shown in Figure 2-11.

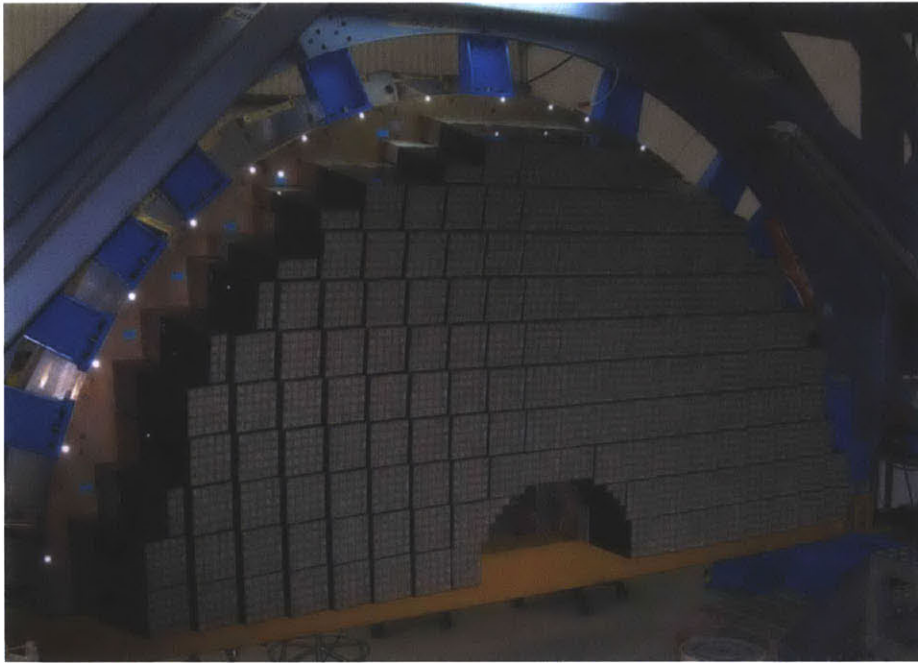


Figure 2-13: An ECAL endcap Dee during installation.

### 2.3.3 Preshower

The preshower detector consists of a lead radiator layer sandwiched between silicon strip sensors. Placed in front of the ECAL Endcap, it covers  $1.653 < |\eta| < 2.6$ . The purpose of the preshower is to help identify  $\pi^0$ s and charged hadrons that can contribute to producing a false signature mimicking an electron in the detector. The

energy deposits in the silicon sensors can be compared to expectations for a minimum ionizing particle (MIP) to identify particles such as  $\pi^\pm$ . The lead radiator induces pair production from photons that originate in a  $\pi^0$  decay. For the 2011 data taking, the energy measurements from the preshower are not used in reconstructing the event, but the energy loss due to particles traversing through the preshower material is accounted for.

### 2.3.4 Crystal Transparency

The major challenge of employing lead tungstate crystals is that exposure to ionizing radiation results in temporary degradation of transparency and a decrease in light yield [24]. The transparency loss is dependent on dose rate of irradiation, which is directly proportional to the instantaneous luminosity provided by the LHC. As the rate of collisions in the LHC increased during 2011 data taking, the crystal corrections needed to account for the transparency loss must also be adjusted accordingly. The state of a crystal is monitored by measuring the light yields of a set of three lasers. Two lasers operate at 440 nm, near the scintillation peak wavelength, to gauge the transparency due to irradiation; two such lasers are employed for redundancy and robustness against failure of a single laser. A third laser operates at 796 nm, far from the scintillation peak wavelength, in order to track the stability of other factors.

### 2.3.5 Energy Resolution

The general expression for the energy resolution of an electromagnetic calorimeter is,

$$\frac{\sigma}{E} = \frac{A}{E} \oplus \frac{B}{\sqrt{E}} \oplus C, \quad (2.1)$$

where  $A$  is the noise term,  $B$  is the stochastic term, and  $C$  is the constant term. The factors that contribute to these terms are as follows,

- *noise term*: noise from electronics, digitization, and pile-up
- *stochastic term*: event-by-event fluctuations in lateral shower containment, pho-

to statistics, fluctuations in energy deposit to preshower absorber with respect to energy measured in preshower silicon

- *constant term*: non-uniformity of longitudinal light collection, intercalibration errors, leakage of energy from the back of the crystal

From the 2006 test beam data, the resolution function was measured to be,

$$\frac{\sigma}{E} = \frac{3.37\%}{E} \oplus \frac{0.108}{\sqrt{E}} \oplus 0.25\%, \quad (2.2)$$

The energy scale and resolution of the ECAL are also monitored and calibrated using data from proton collisions with electrons and photons from several physics processes. These include calibration to the momentum measurements from the tracker of electrons from  $W$  decays, calibration to the masses of  $\pi^0$  and  $\eta$  when these particles decay to diphotons, and calibration to the mass of the  $Z$  in decays to dielectrons.

## 2.4 Hadronic Calorimeter

The CMS hadronic calorimeter is a sampling calorimeter and is divided into four major partitions: HCAL Barrel (HB), HCAL Endcap (HE), HCAL Outer (HO), and HCAL Forward (HF). The HB lies within the solenoid while the HE is inserted at the ends of the solenoid with the HF situated further beyond. The HO lies just outside the solenoid. All the partitions together give the HCAL acceptance up to  $|\eta| < 5.2$ . Figure 2-14 illustrates the pseudo-rapidity coverage of the HCAL and its various partitions. The relative pion energy resolution of the HCAL is on the order of  $\sigma/E \approx 100\%/\sqrt{E}$ .

### 2.4.1 HCAL Barrel

The HB lies inside the solenoid spanning radially 1.77 m to 2.95 m from the beam line and covers the pseudorapidity range  $|\eta| < 1.3$ . It uses brass absorber plates interleaved with plastic scintillators. The brass absorbers provide thickness in terms



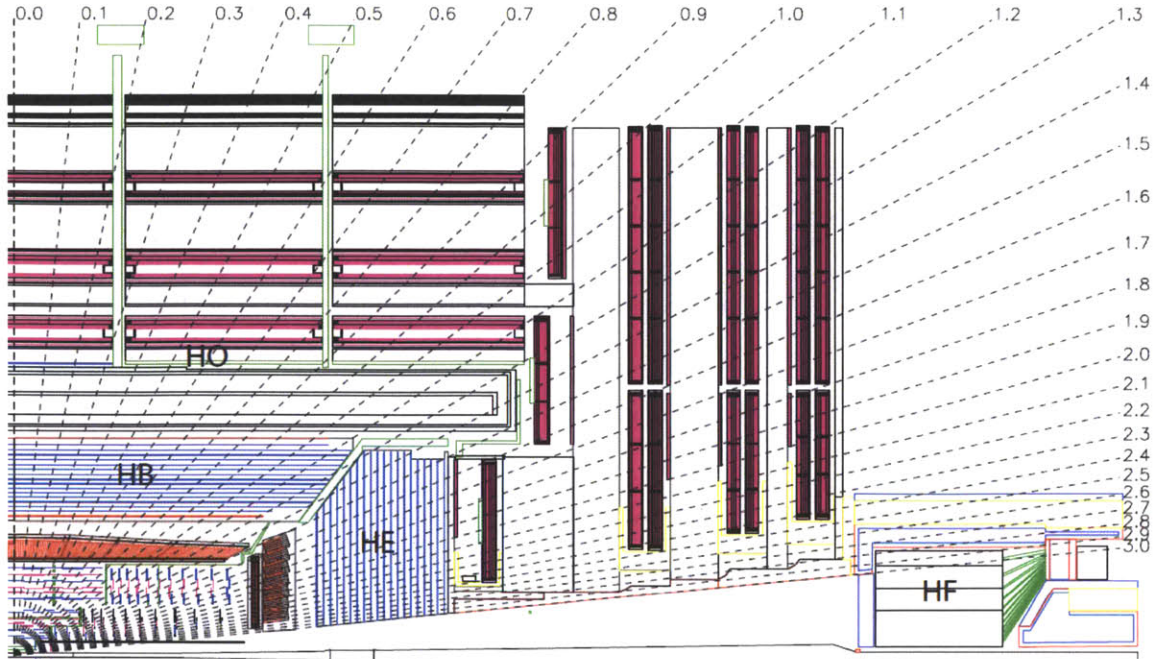


Figure 2-14: Layout of the CMS HCAL indicating  $\eta$  coverage for each partition.

of nuclear interaction lengths ( $\lambda_I$ ) from  $5.82\lambda_I$  at  $|\eta| \approx 0$  to  $10.6\lambda_I$  at  $|\eta| \approx 1.3$ . The plastic scintillators are square tiles containing wavelength shift (WLS) fibers to collect and guide the light to hybrid photodiodes (HPD) for readout. An HPD is a combination of photomultiplier and silicon photodetection principles, and is capable of effective operation at high magnetic fields [25]. For event reconstruction, the HB is divided into sectors or “towers” that cover an area of  $0.087 \times 0.087$  in  $\eta - \phi$  space.

## 2.4.2 HCAL Endcap

The HE are inserted at the ends of the solenoid and must be made of non-magnetic material due to gradients from the non-uniform magnetic field. It covers the range  $1.3 < |\eta| < 3$  and uses HPD for readout like the HB. The size of the towers in  $\eta - \phi$  are  $0.087 \times 0.087$  for  $|\eta| < 1.6$  and  $0.17 \times 0.17$  for  $|\eta| > 1.6$ . Including the ECAL endcap, the total thickness of the endcap calorimeter is about  $10\lambda_I$ .

### 2.4.3 HCAL Outer

The HO, also known as the “tail catcher”, serves to provide additional stopping power against hadronic debris entering the muon system. With the inclusion of the HO and also the solenoid as additional absorber material, the CMS calorimeter system guarantees at least a depth of  $11.8\lambda_I$  in interaction lengths over all of  $|\eta| < 3$ . Figure 2-15 shows a photograph of a plastic scintillator tile with embedded WLS fibres for the HO.

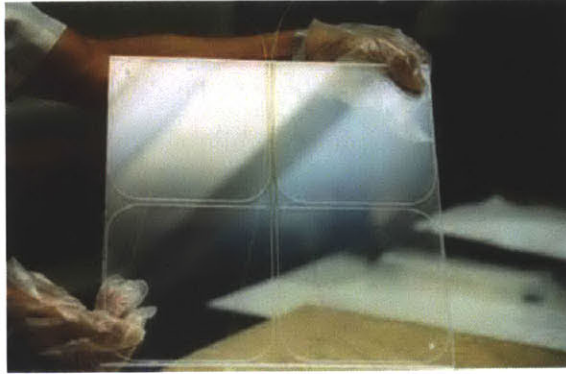


Figure 2-15: HCAL Outer tile and WLS

### 2.4.4 HCAL Forward

The HF is placed at  $z = \pm 11.2$  m from the interaction point and provides coverage out to  $|\eta| = 5.2$ . The design of the HF uses Cherenkov light emitting quartz fiber for the active medium and steel for the absorber material. The choice of quartz is because the higher hadronic flux in the the forward region requires more radiation hardness. The Cherenkov light is readout using photomultiplier tubes (PMT). The HF can be used as a luminosity monitor.

## 2.5 Muon System

Beyond the calorimeters lie the muon system. Using various gas detector technologies, it contains roughly  $25\,000\text{ m}^2$  of detection planes. The muon system is partitioned

into the barrel ( $|\eta| < 1.2$ ) and endcaps ( $0.9 < |\eta| < 2.4$ ), where drift tubes (DT) are employed in the barrel and cathode strip chambers (CSC) in the endcaps. Resistive plate chambers (RPC) are used throughout the barrel and part of the endcaps ( $|\eta| < 1.6$ ) and primarily serve to provide a timing measurement of the bunch crossing for relevant events.

### 2.5.1 Drift Tube

The DTs in the barrel are encased in four layers of concentric cylinders called “stations”. All stations contain 8 “chambers” of DTs parallel to the beam line to provide position measurements in the  $r - \phi$  plane, while the first three stations contain 4 chambers of DTs perpendicular to the beam line to provide measurements in the  $z$  direction. The three inner stations have 60 DTs each while the last station has 70 DTs.

A drift tube consists of a 2.4 m sensitive wire centered in a rectangular casing with transverse dimensions of  $13 \times 42 \text{ mm}^2$ . Figure 2-16 illustrates the layout of a drift tube. The drift tube is filled with gas made of 85% Ar and 15% CO<sub>2</sub> kept at atmospheric pressure. The drift time for a typical distance of 21 mm is 380 ns; both the drift velocity and signal efficiency are sensitive to contamination from N<sub>2</sub> and O<sub>2</sub> respectively, hence the drift cell must be air-tight to maintain stable operating conditions. The timing resolution of a superlayer (4 layers of DTs) is a few ns, allowing for tagging of bunch crossings. The  $r - \phi$  resolution of a chamber is approximately 100  $\mu\text{m}$ .

### 2.5.2 Cathode Strip Chambers

The endcaps contain 468 CSCs divided into four stations per endcap. A CSC is a multi-wire proportion chamber consisting of 6 anode wire planes interleaved among 7 cathode panels. The wires run azimuthally, thereby defining the radial coordinate. The position resolution ranges from 75  $\mu\text{m}$  to 150  $\mu\text{m}$  and the timing resolution is a few ns.



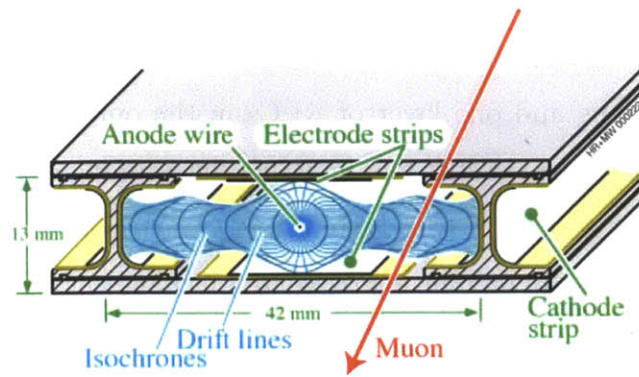


Figure 2-16: Layout of a drift tube. The red arrow indicates a typical traversal of a muon through the detector.

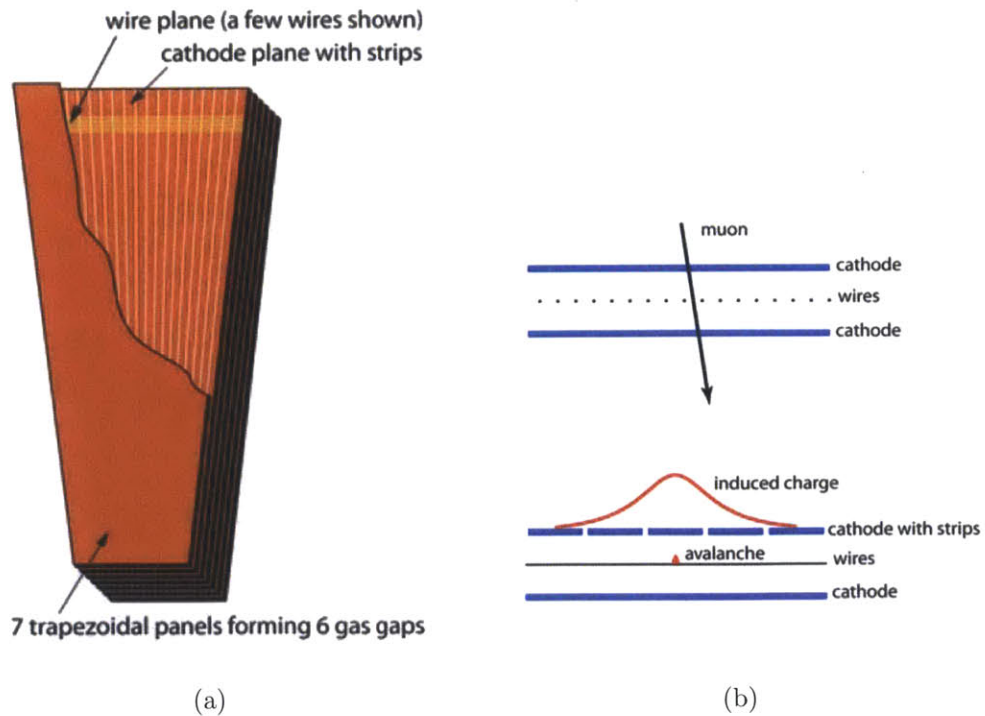


Figure 2-17: (a) Layout of a CSC. (b) Schematic of a gas gap between cathode panels, and a typical traversal of a muon through the detector.

### 2.5.3 Resistive Plate Chambers

The muon system contains 480 RPCs. In the barrel, there are two layers of RPCs in the inner two stations and one layer of RPCs in the outer two stations. In each endcap, there are three layers of RPCs. The RPCs have timing resolutions of a couple of ns, but the position resolution is roughly 1.5 cm, much poorer relative to the DTs and the CSCs. A depiction of an ionising particle traversing through an RPC is shown in Figure 2-18.

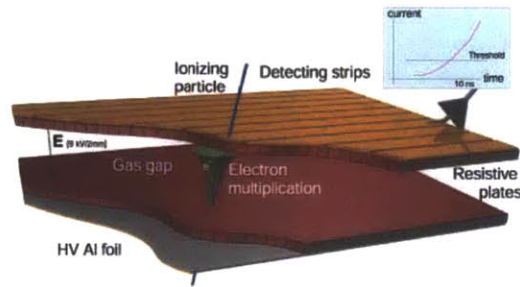


Figure 2-18: RPC

## 2.6 Forward Calorimeters

The CMS detector employs two calorimeters in the very forward regions to study diffractive and low- $x$  physics. These are CASTOR (Centauro and Strange Object Research) and ZDC (Zero degree calorimeter). CASTOR is a Cherenkov-based, quartz-tungsten sampling calorimeter with photomultiplier tubes for readout located at  $z = \pm 14.38$  m from the interaction point and covering  $5.2 < |\eta| < 6.6$ . ZDC is also a quartz-tungsten sampling calorimeter located at approximately 140 m away from the interaction point and covers  $|\eta| > 8.3$ . Neither of these very forward calorimeters are utilized in the Higgs search analysis.

## 2.7 Trigger

For the design bunch crossing interval of 25 ns at the LHC, the rate of proton-proton collisions is 40 MHz. The data taking and storage systems cannot handle recording every single collision event and filtering is required to reduce the processing load to a manageable rate. The trigger system identifies properties associated with potentially interesting physics events to determine whether more thorough processing and storage of events should proceed. The CMS trigger can be considered a sequence of two major steps,

- Level-1 (L1) trigger: implemented with custom-designed, programmable electronics using coarsely segmented data from the calorimeters and the muon system, designed to reduce the 40 MHz rate down to 100 kHz,
- High Level Trigger (HLT): implemented with a software filter farm using full detector information, designed to reduce the rate down to  $\mathcal{O}(100 \text{ Hz})$ .

The processing of detector readout by the trigger system is commonly called the “online” reconstruction, since this is performed during data taking and the output goes to determine if an event is accepted or rejected.

At L1 stage, only information in the calorimeters and the muon chambers are used essentially, to perform quick calculations and render a decision on whether or not to process the event further. ECAL crystals and HCAL towers are grouped together to form calorimeter towers, and various conditions to trigger an accept of the event can be defined with respect to the number of towers with energy deposits above given thresholds. These energetic calorimeter towers become the candidates for building objects representing electrons, photons, or jets at the HLT. In the muon system, crude tracking is performed to obtain muon candidates and a trigger decision can be made based on the number of candidates and their momenta.

At the HLT stage, the event information is utilised more thoroughly to build data objects that are associated with particles such as muons, electrons, photons, or jets. In particular, inner tracker information is introduced at this stage. In addition to

improved energy and momentum measurements compared to L1, attributes such as the amount of energy deposits around a particle (quantifying the level of isolation) or the angular separation between an inner tracker track and the ECAL energy cluster of an electron (quantifying the quality of the electron reconstruction) are computed and used by the HLT to make a more informed trigger decision. However, as a compromise between precision and speed, these algorithms are mainly streamlined versions of the techniques employed by “offline” reconstruction (Chapter 3).

Since the instantaneous luminosity increased by roughly an order of magnitude over the 2011 data taking campaign, the menu of trigger conditions had to evolve in order to cope. Typically, energy thresholds are increased or stricter quality requirements are imposed to compensate for the higher rate of collisions. For triggers that serve the purpose of picking out events from interesting physics processes, this is a feasible approach since such processes are typically rare. However, for triggers that are used to obtain background enriched control samples, tightening the requirements usually means closing up the phase space available to the control region. In such cases, a trigger “prescale” is implemented meaning only a fraction of the trigger accepts are actually recorded.

# Chapter 3

## Event Reconstruction

After an event has been accepted by the online trigger system, the information from the detector is stored and processed offline by sophisticated algorithms to obtain the most accurate and precise picture of the debris from the proton-proton collision. Hits in the inner tracker and muon chambers are connected to form tracks corresponding to trajectories of charged particles. Energy deposits in the ECAL are grouped together to form “superclusters”. These tracks and superclusters then initiate the incorporation of information from other subdetectors to construct data objects for analysis corresponding to photons, electrons, and muons; these data objects include the four-momentum and additional information about the quality and properties of the measurements used. Dedicated procedures for each of these particles have been developed to achieve the best reconstruction. In contrast, the “Particle Flow” (PF) algorithm is an alternative procedure that aims to identify and describe all particles in the event, regardless of particle species. This global event description facilitates the reconstruction of jets and computation of the missing transverse energy.

### 3.1 Track Reconstruction

As charged particles traverse through the inner tracker, they leave slight energy deposits in the silicon pixels and strips. The objective of track reconstruction is to identify the collection of hits corresponding to an individual particle and extract a

momentum measurement based on the curvature of the trajectory. To be more precise, the particle trajectory is locally a helix parameterised by five variables called *track parameters*,

- $q/|p|$ : electric charge divided by the momentum,
- $\lambda: \pi/2 - \theta$ , where  $\theta$  is the polar angle from the beam line ( $z$  axis),
- $\phi$ : azimuthal angle in the  $x - y$  plane,
- $d_0$ : distance in the  $x - y$  plane of the trajectory point closest to the origin of the coordinate system,
- $z_0$ : distance in the  $z$  coordinate of the trajectory point closest to the origin of the coordinate system.

The main steps of the tracking algorithm are seed finding, pattern recognition, and track fitting, and are summarised below.

Hits in the tracker are reconstructed from clusters of pixels and strips that have signal. From the set of reconstructed hits, track seeds are found from collinear triplets of hits or pairs of hits consistent with the interaction point. These track seeds provide a crude  $p_T$  estimation and a direction to search for other hits along the particle trajectory. The tracking algorithm follows an iterative procedure that begins with seeds consistent with a particle from the primary interaction point to build tracks, then proceeds with track building with seeds found progressively further away from the beam line using remaining hits not yet associated with a track.

From a track seed, the pattern recognition step uses the Kalman filter (KF) technique to build the collection of hits to form a track. The KF method is a linear least-squares estimator and consists of repetitions of “propagation” steps and “update” steps. The propagation step takes the current state of the partially built track and estimates the expected position and associated uncertainty of the hit on the next tracker layer, taking into account the equation of motion, inhomogeneity in the magnetic field, multiple Coulomb scattering, and energy loss in detector material from

ionization and bremsstrahlung. If hits are found on the next tracker layer that is compatible with expectation, then the measurements of the hits are incorporated into track and the track parameters and uncertainties are updated accordingly. Essentially, the KF method handles both the pattern recognition and track fitting steps, improving on the estimation of the track parameters as hits are added to the track. At the end of this procedure, the outermost hit on the track will have the best estimate of the track parameters, since the track state on that layer uses information from all the hits on the track. A track “smoothing” is done by applying the KF method “inwards” starting with the outermost hit. The expected hit positions are not recomputed in the propagation steps, but the updates allow for rejection of outliers along the track. At the end of the smoothing, the estimate of the track parameters at the innermost hit now also uses information from the full set of hits.

## 3.2 Primary Vertex Reconstruction

The identification of where the hard scattering process took place is vital to an analysis given the large number of pile-up interactions at the LHC. For strong discrimination against pile-up, the desired property for the primary vertex reconstruction is a small minimum separation,  $\Delta z$ , required to distinguish two interactions. The primary vertex reconstruction consists of the following steps,

- (1) select tracks compatible with the beam line and satisfy a requirement on the number of hits,
- (2) cluster tracks that are consistent with originating from the same interaction point; this is done with the deterministic annealing method,
- (3) for each cluster of tracks, perform a fit for the vertex position,
- (4) keep vertices which are within 1 cm from the beam line.

The minimum separation needed to distinguish two vertices is  $\Delta z \approx 1$  mm. Note that the minimum separation value is dependent on the clustering method and the typical



resolution of a track, and this is much larger than the position resolution of a vertex, which comes from the fit of an assigned set of tracks.

### 3.3 Supercluster Reconstruction

A proper clustering of the energy deposits in the ECAL is essential for accurate reconstruction of electrons and photons. Typically, about 94% of the energy from  $e/\gamma$  are contained in a  $3 \times 3$  array of crystals, and about 97% in a  $5 \times 5$  array. To make sure the energy from an electron or photon is fully accounted for, algorithms are developed to form clusters of crystals and then clusters of crystal clusters to form superclusters, which become the seeds for reconstructing an electron or photon.

Due to the strong magnetic field, particles in the electromagnetic showers tend to spread in the  $\phi$  direction. The crystals in the ECAL are oriented differently between the barrel (radially) and the endcap (along  $z$ ), so different algorithms are used to build superclusters in each ECAL partition since electromagnetic showers will develop differently relative to the detector geometry. In the barrel, the clustering method is straightforward. For a crystal above the seed energy threshold, a  $5 \times 1$  “domino” of crystals is considered, with the seed crystal at the center. Searching in the  $\phi$  direction, adjacent dominoes of sufficient energy are incorporated into the supercluster. Hence, superclusters in the ECAL barrel are rectangles in  $\eta - \phi$ , with a five crystal width in  $\eta$  and variable extent in  $\phi$ . In the endcap, a seed crystal determines a  $5 \times 5$  array centered around the seed. All crystals at the border of this array above an energy threshold will also spawn its own  $5 \times 5$  array of crystals, which overlaps with the initial array. Continuing in this manner with crystals at the border, these overlapping arrays form a supercluster.

After the superclusters are formed, several corrections to the energy are applied to account for geometry and detector material effects. In the barrel, the crystals do not face the interaction point, but are off-pointing by a few degrees. Hence, there can be leakage of energy across crystals and out of the supercluster. This is accounted for by an  $\eta$ -dependent correction. In both the barrel and endcap, there is a correction for

material in front of the ECAL, and a correction is applied based on the ratio of the supercluster size in  $\phi$  to the supercluster size in  $\eta$ . Also for the barrel and endcap, a correction dependent on  $\eta$  and  $E_T$  is made to account for the amount of tracker material, which varies in  $\eta$ , and the effects of bremsstrahlung and photon conversions, which depend on energy.

## 3.4 Photons

Every ECAL supercluster is a seed for photon reconstruction. For a supercluster to be promoted to a photon, it must satisfy  $E_T > 10$  GeV and  $H/E < 0.5$ , where  $H/E$  is the ratio between the energy deposits of the HCAL and the ECAL. To derive the four momentum of the photon, the direction is taken to be the direction from the interaction point of the supercluster (average of the crystal positions weighted by energy contribution) and the energy component to be assigned depends on the value of the ratio between the energy in the  $3 \times 3$  array of crystals centered around the seed to the total supercluster energy, denoted  $R_9$ . This variable provides a measure of the lateral spread of the electromagnetic shower. A low value of  $R_9$  is indicative of a photon that has undergone conversion in the tracker material resulting in a larger spread. If  $R_9 < 0.94$  for a barrel supercluster or  $R_9 < 0.95$  for an endcap supercluster, the total energy of the supercluster is assigned to the photon object. The threshold is tighter for the endcap because the crystals are larger. Otherwise, the photon energy is assigned to be the energy of the  $5 \times 5$  array of crystals centered around the seed crystal, to reduce noise.

### 3.4.1 Photon Conversions

Photons can undergo conversion when traversing through the beam pipe and tracker material, resulting in electron-positron pairs that produce tracks. When conversions occur before the first pixel layer, the electron (or positron) can be mistaken for a prompt particle originating from the collision. To discriminate against such a process, photon conversion candidates are identified by performing a vertex fit on pairs of

oppositely charged tracks. A constraint that the pair of tracks are parallel in three dimensions is imposed, and the position of the conversion vertex must be at least 2 cm away from the primary vertex to avoid the primary interaction as a conversion vertex.

### 3.5 Electrons

As charged particles, electrons will leave a track in the tracker before depositing most of their energy in the ECAL. However, the electrons typically produced in collisions at the LHC are in the regime where energy loss through bremsstrahlung radiation is substantial. Though this is well described by the Bethe-Heitler model, it results in a non-Gaussian distribution of energy loss for electrons traversing through detector material and sub-optimal performance of the Kalman filter method for tracking. The remedy is an extension of the KF technique using sums of Gaussian distribution to better approximate the Bethe-Heitler description, and this tracking algorithm dedicated to electrons is called the Gaussian Sum Filter (GSF).

The reconstruction of an electron begins with a supercluster and looks for hits in the pixel layers and strip endcap disks compatible within  $\phi$  and  $z$  windows around the supercluster. These trajectory seeds are used to initiate GSF track reconstruction. The increased complexity of the algorithm means GSF tracks take longer to reconstruct than standard KF tracks, so GSF reconstruction is only applied to electron candidates to keep processing times down. The difference between the momenta at the innermost hit and at the outermost hit is indicative of the amount of energy lost due to radiation.

It is possible for the curvature of the electron trajectory to change signs upon radiating a photon, hence charge determination is not trivial for an electron. There are three ways to determine the charge of an electron and the ultimate assignment is taken from the majority. These three methods are,

- charge of the GSF track,

- charge of the KF track that best matches the GSF track,
- compute the signed  $\Delta\phi$  between the vector joining the beam spot and the supercluster and the vector joining the beam spot and the first hit of the electron track.

Finally, the four momentum of the electron needs to be determined. The direction is derived from the electron track at the point of closest approach to the beam spot. The magnitude of the momentum is a weighted mean of the ECAL energy and track momentum measurements when the two quantities are in fair agreement. When the ECAL energy and track momentum disagree, if the ECAL energy is below 15GeV and the GSF track shows only small energy loss due to radiation, the track momentum is taken since it has better expected resolution; otherwise, the ECAL energy is taken.

## 3.6 Muons

The reconstruction of a muon object relies on measurements in the inner tracker and the muon chambers. There are two complementary algorithms to reconstruct a muon object: an “outside-in” approach starting with a track in the muon system then matching that with a track from the inner tracker, and an “inside-out” approach starting with an inner tracker track then matching that to hits in the muon system.

The “outside-in” approach starts with finding patterns of hits in the muon system consistent with individual particles to form “segments” in each DT or CSC station. From these segments, the Kalman filter method is used to accumulate compatible hits to build a trajectory and determine the track parameters. These tracks, reconstructed with measurements from the muon chambers only, are called “standalone muons”. Each standalone muon then serves as a seed to reconstruct a “global muon”, which combines muon system and inner tracker information, to achieve the best measurement of the muon trajectory. The standalone muon is matched to tracker tracks in roughly the same  $\eta-\phi$  direction and the parameters of the combined tracks are determined again from the Kalman filter technique. During the track fitting procedure, it

is possible for hits in the standalone muon to be removed from the global fit if the hit is deemed incompatible with the trajectory. Hit removal improves robustness against spurious hits in the muon system due to combinatorics or radiation from muons. The standalone muon and tracker track combination that gives the best  $\chi^2$  fit forms a global muon.

The “inside-out” approach is seeded by a tracker track and looks for at least one compatible segment in the muon chambers. The muon object reconstructed by this algorithm is called a “tracker muon”. The tracker muon momentum is taken from the tracker track as no combined fit is performed for the tracker track and the muon hits. The tracker muon method was primarily developed to improve efficiency for reconstructing low  $p_T$  muons (on the order of a few GeV/ $c$ ), which may not travel far enough into the muon system and produce enough hits to form a segment.

Almost all well measured muons in the relevant  $p_T$  range will be reconstructed by both algorithms. If a global muon and a tracker muon share the same tracker track, it is considered to be the same muon object. To improve robustness and purity, a selection requirement in the analysis will be that a muon object is both a global muon and a tracker muon, rather than relying on just one of the algorithms. For  $p_T \lesssim 200$  GeV/ $c$ , the best momentum resolution is obtained from the tracker track alone, hence for this analysis, the muon system functions primarily to identify the presence of muons in the event.

### 3.7 Particle Flow event description

The Particle Flow algorithm provides a complete description of the event in terms of particles. The PF event description is a list of particles of the following types: photon, charged hadron, neutral hadron, muon, and electron. This collection of particles can then be fed into higher level reconstruction algorithms to compute jets,  $\tau$  leptons, and missing transverse energy. For this analysis, the objects reconstructed using the dedicated algorithms for photons, electrons, and muons are used as they perform slightly better than the analogous objects from the PF algorithm.

The PF reconstruction uses the collection of tracker and muon tracks found by the methods described previously. However, the PF method has its own strategy for clustering energy in the calorimeters. A sophisticated algorithm then proceeds to link tracks and/or calorimeter clusters and assign particle types. Corrections are applied to calorimeter clusters associated with hadrons. Clusters consistent with a charged hadron are calibrated to the associated track momentum. Clusters corresponding to neutral hadrons have a correction applied derived from simulation studies.

Given the list of particles in the event, it becomes straightforward to construct jets and the missing transverse energy in the event.

### 3.7.1 Jets

A jet is a collimated group of particles resulting from the hadronisation of a quark or gluon. Hence, a jet object can be reconstructed by clustering particles given by the Particle Flow event description. The clustering algorithm used is the anti- $k_T$  method [26] with a cone radius in of 0.5 in  $\eta - \phi$ . The anti- $k_T$  algorithm is infrared safe and collinear safe, meaning jet finding is not significantly altered by soft gluon radiation or the number of hard particles carrying most of the parton energy. These properties are essential for making correspondences between experiment and theory. Furthermore, compared to other infrared safe and collinear safe techniques, the anti- $k_T$  algorithm is expected to be less sensitive to contributions from the underlying event and pile-up interactions.

To obtain a more accurate description of the parton from which the jet originated, the jet energy scale is corrected for pile-up contributions and detector response. The pile-up contribution is estimated by the product of a jet’s “area” and the “noise density” of the event. The jet area is the effective extent in  $\eta - \phi$  space in which particles will cluster to the same jet [27]. The noise density is determined by constructing the distribution of energy per area for jets in the event and taking the median of this distribution [28]. The median is less biased than the mean to contributions from the hard scatter process. The pile-up corrections are applied event-by-event and jet-by-jet. The detector response is compensated with jet energy scale corrections. These

values are derived by determining the corrections needed to achieve transverse energy balance in dijet,  $Z$ +jet, and  $\gamma$ +jet events [29].

### 3.7.2 Missing Transverse Energy

The signature of neutrinos in an event is missing transverse energy,  $\cancel{E}_T$ , in the detector. The inability to detect neutrinos results in an observed non-conservation of transverse momentum in the event. The missing transverse energy is computed as the negative vector sum of transverse momentum from all PF particles. It is worth noting that neutrinos are not the only source of transverse momentum imbalance, but there are experimental sources as well. Resolution effects result in mismeasurements that can induce significant missing energy. Inefficiencies to reconstruct all particles in the event will produce missing energy. The computation of  $\cancel{E}_T$  using PF particles reduces these experimental effects compared to calorimeter based or track based methods.

### 3.7.3 Simulation

Simulated event samples are used to help design the selection algorithms for signal versus background discrimination, to parametrise efficiencies and resolutions for various observables, and to study higher order theoretical uncertainties which affect the measurements. The production of simulated events begins with the generation of particle kinematics of the underlying physics using Monte Carlo event generators. The physics of the hard scatter process are simulated with POWHEG [30–32], MADGRAPH [33], or PYTHIA [34]. POWHEG accounts for matrix elements up to next-to-leading order (NLO) in QCD. MADGRAPH [33] computes matrix elements at leading order (LO) of the hard scatter process plus any number of emitted partons in the final state (up to three or four partons depending on the physics process). PYTHIA handles only diagrams at LO in QCD.

To complete the picture for the generated event, further processing with PYTHIA is done for parton showering, hadronisation, and simulation of the underlying event. The underlying event consists of the “spectator interactions” of everything that is



not involved in the hard scatter. PYTHIA provides many parameters to tune its model of the underlying event, and the suite of parameter settings used by the CMS collaboration is known as “Tune Z2” [35]. To avoid double production of jets, emitted partons generated with POWHEG or MADGRAPH are matched to clusters of particles produced from the showering and hadronisation of PYTHIA.

Pile-up interactions are incorporated into an event by including additional inelastic proton-proton collisions. The number of pile-up interactions in a simulated sample of events is a uniform distribution from 0 to 50. To make meaningful comparisons with data, the events of a simulated sample must be re-weighted to match the expected pile-up multiplicity distribution in data.

The event is then processed by GEANT4 [36, 37] to simulate the passage of particles through the CMS detector. This generates the inputs needed to run the reconstruction algorithms.



# Chapter 4

## Inclusive $Z$ cross section via decays to electron or muon pairs

### 4.1 Introduction

The production of  $Z$  bosons in collider experiments [38] is a well understood process. The experimental signature of a  $Z$  boson decaying into dielectrons or dimuons is very clean and simple to identify; selection of a pair of high- $p_T$  leptons satisfying a dilepton mass window constraint results in a high purity sample of  $Z$  events. Further requirements on the leptons to be isolated and to pass quality criteria reduce the QCD multi-jet background to negligible levels. A sample of  $Z$  events is used to measure the efficiencies of the trigger and lepton selection requirements. The measurement of the inclusive cross section at 7TeV with  $36\text{pb}^{-1}$  of proton-proton collisions in 2010 [39] is consistent with calculations at next-to-next-to-leading order (NNLO) in perturbative QCD, currently the most precise theoretical prediction [40–44]. This chapter describes the measurement of the inclusive  $Z$  cross section with the 2011 dataset of  $4.9\text{fb}^{-1}$ , using the same lepton selection as in the search for the Higgs, to cross check the measured electron and muon efficiencies.

A cross section measurement can be neatly summarised by the following equation,

$$\sigma = \frac{N_S}{A \cdot \varepsilon \cdot \mathcal{L}}, \quad (4.1)$$

where

- $\sigma$  is the cross section,
- $N_S$  is the number of observed signal events,
- $A$  is the acceptance – the expected fraction of events produced that can be observed,
- $\varepsilon$  is the efficiency of observing events that fall within acceptance,
- $\mathcal{L}$  is the integrated luminosity of data.

For the  $Z$  cross section measurement,  $N_S$  is extracted by counting events passing selection criteria and subtracting estimated backgrounds. The event selection is explained in Section 4.2. The acceptance is defined to be the fraction of  $Z$  events that pass offline event selection as computed from Monte Carlo simulation. Since trigger strategies evolved frequently during data taking, but these changes are not closely mimicked in simulation, the trigger selection is not included in the definition. A discussion of theoretical models used to calculate the acceptance and its systematic uncertainties is given in Section 4.3. The efficiency for reconstructing and selecting leptons from  $Z \rightarrow \ell\ell$  decays is measured in data. The procedure and systematic uncertainties on the efficiency measurement are detailed in Section 4.4. Note that in the acceptance definition, the lepton efficiency in simulation is implicitly included. The efficiency from data is incorporated by considering the ratio of efficiencies of data to simulation and then re-weighting the simulated events in computing the acceptance. Finally, the calibration of the integrated luminosity, performed by a dedicated working group in the CMS collaboration, is summarised in Section 4.5.

## 4.2 Event Selection

To select  $Z \rightarrow \ell\ell$  candidates, events with dielectrons or dimuons are considered where the leptons are oppositely charged and the dilepton mass satisfies  $60 \text{ GeV}/c^2 < m_{\ell\ell} < 120 \text{ GeV}/c^2$ . The electrons are required to have a supercluster with  $E_T > 20 \text{ GeV}$  and

$|\eta| < 2.5$  (within the ECAL coverage), but excluding  $1.4442 < |\eta| < 1.566$ , which is a region around the barrel-endcap interface where the measurement quality is poor due to the presence of additional material from service and support structures. The muons are required to have an inner tracker track with  $p_T > 20 \text{ GeV}/c$  and  $|\eta| < 2.4$  (within the muon system coverage). Events to be accepted for offline reconstruction are triggered by the presence of a muon, a muon pair, or an electron pair. Suppression of background from multi-jet processes is improved by requiring the reconstructed leptons to be isolated and satisfy lepton identification requirements.

### 4.2.1 Trigger Requirements

During data taking, beam conditions evolved several times towards increasing density of proton bunches and frequency of bunch crossings. Consequently, due to limited data bandwidth, the trigger requirements were modified accordingly to adapt to higher collision rates; lepton energy thresholds were raised or lepton identification and isolation criteria were tightened to decrease the rate of events accepted for processing.

#### Electron Triggers

Dielectron triggers are employed to select  $Z \rightarrow ee$  events. Due to the relatively high rate by which jets can fake an electron-like signal, isolation and electron identification requirements are imposed to keep the rates manageable. A useful single electron trigger was not available, as the energy thresholds quickly became too high and the isolation and identification requirements too strict to efficiently select events for a  $Z$  cross section measurement or the Higgs search. The dielectron triggers are summarised in Table 4.1.

#### Muon Triggers

In contrast to electrons, the low rate which QCD produces fake muon signals allows the use of both dimuon and single muon triggers. The single muon triggers provide a roughly 5% increase in signal yield over simply using the dimuon triggers alone. The

Runs	Lumi [fb <sup>-1</sup> ]	Level-1 requirements	HLT requirements
160329 - 170053	1.2	single $e/\gamma$ : $E_T > 12$	dielectrons: $E_T > 17$ , 8 with calo cuts
170071 - 170759	1.0	single $e/\gamma$ : $E_T > 12$	dielectrons: $E_T > 17$ , 8 with calo and track cuts
170826 - 170901 171484 - 180291	2.7	pair of $e/\gamma$ : $E_T > 12$ , 5	dielectrons: $E_T > 17$ , 8 with calo and track cuts

Table 4.1: Dielectron triggers used to select  $Z \rightarrow ee$  candidates. For each trigger, the amount of data collected (in fb<sup>-1</sup>) is listed and a brief summary of the Level-1 and HLT requirements are provided.  $E_T$  quantities are in units of GeV.

dimuon and single muon triggers are summarised in Tables 4.2 and 4.3 respectively. In contrast to electrons, few requirements beyond the  $p_T$  threshold are needed to maintain manageable event rates. Isolation is imposed for the single muon trigger.

Runs	Lumi [fb <sup>-1</sup> ]	Level-1 requirements	HLT requirements
160329 - 164273	0.2	dimuon: both $p_T > 3$	dimuon: both $p_T > 7$
165085 - 178380	3.8	dimuon: both $p_T > 3$	dimuon: $p_T > 13$ , 8
178420 - 180291	0.9	dimuon: leading $p_T > 10$	dimuon: $p_T > 17$ , 8

Table 4.2: Dimuon triggers used to select  $Z \rightarrow \mu\mu$  candidates. For each trigger, the amount of data collected (in fb<sup>-1</sup>) is listed and a brief summary of the Level-1 and HLT requirements are provided.  $p_T$  quantities are in units of GeV/ $c$ .

Runs	Lumi [fb <sup>-1</sup> ]	Level-1 requirements	HLT requirements
160329 - 167043	0.9	$p_T > 10$	$p_T > 17$ , isolated
167078 - 173198	1.1	$p_T > 12$	$p_T > 24$ , isolated
173243 - 180291	2.9	$p_T > 24$ , $ \eta  < 2.1$	$p_T > 30$ , $ \eta  < 2.1$ , isolated

Table 4.3: Single muon triggers used to select  $Z \rightarrow \mu\mu$  candidates. For each trigger, the amount of data collected (in fb<sup>-1</sup>) is listed and a brief summary of the Level-1 and HLT requirements are provided.  $p_T$  quantities are in units of GeV/ $c$ .

## 4.2.2 Isolation

Leptons originating from  $Z$  decay are typically isolated from other particles in the event, whereas lepton signatures in hadronic processes are not. To reduce the selection

of leptons induced by jets, the sum of transverse energy from tracks and calorimeter deposits around the lepton candidate within a  $\Delta R = 0.3$  cone is computed; an upper limit on this sum imposes isolation. Pile-up interactions contribute uncorrelated energy deposits to the isolation cone. To account for this effect, a subtraction from the isolation sum is made based on the event “energy density” computed with the FastJet algorithm [45, 46]. This corrected isolation sum is divided by the lepton  $p_T$  to obtain “relative isolation”: muons are selected with relative isolation less than 0.15, and electrons are selected with relative isolation less than 0.1. The tighter isolation cut for electrons is used to counteract the higher probability for jets to fake an electron.

To illustrate the behavior of the isolation variable in different physics processes, single muons or electrons passing impact parameter cuts and lepton identification requirements (to be discussed in subsequent sections) are selected and the isolation distribution is plotted in Figure 4-1. The simulated events are normalized to the yield in data. The primary source of real muons and electrons is  $W \rightarrow \ell\nu$ , with secondary contributions from  $Z$ , electroweak diboson, and top productions. The dashed line and arrow indicate the region to be selected with the relative isolation cut. The cut threshold is chosen to remove as much QCD background as possible while maintaining high efficiency for real leptons.

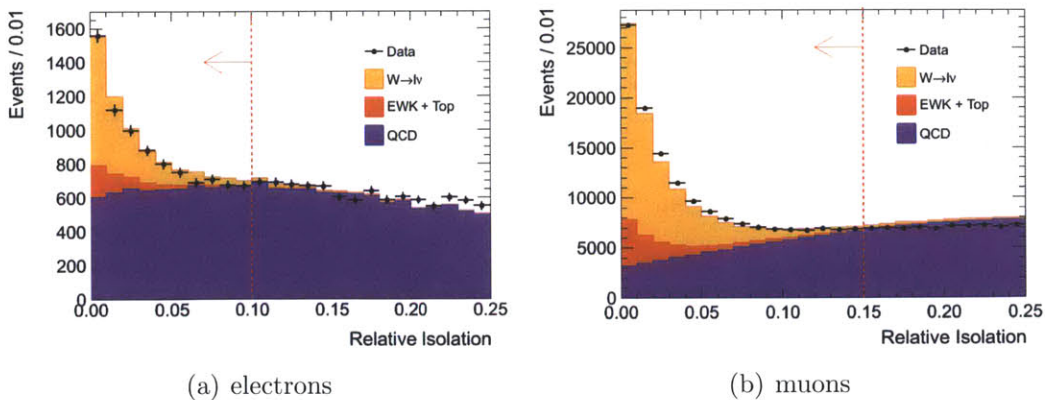


Figure 4-1: Distributions of relative isolation for electrons and muons. The dashed line and arrow indicates the region to be selected by the isolation cut.



### 4.2.3 Impact Parameter

To further reduce background events from non-prompt leptons (such as from semi-leptonic heavy flavour decays) and pile-up interactions, a small impact parameter with respect to the event primary vertex is required for both leptons. The lepton tracks must be within 0.02 cm in the transverse plane ( $d_0$ ), and within 0.1 cm along the  $z$ -direction ( $d_z$ ). The impact parameter distributions for electrons and muons are plotted in Figures 4-2 and 4-3 from a selection of events with isolated leptons.

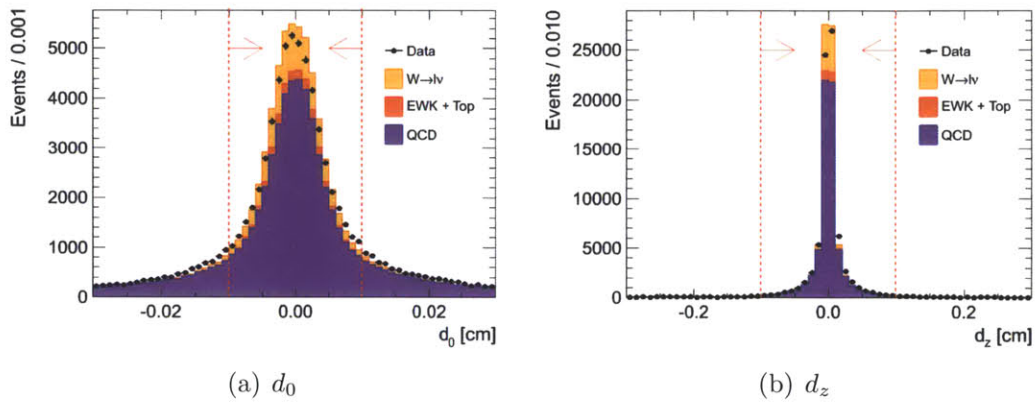


Figure 4-2: Distributions of the transverse ( $d_0$ ) and longitudinal ( $d_z$ ) impact parameter for electrons. The dashed line and arrow indicates the region to be selected by the impact parameter cut.

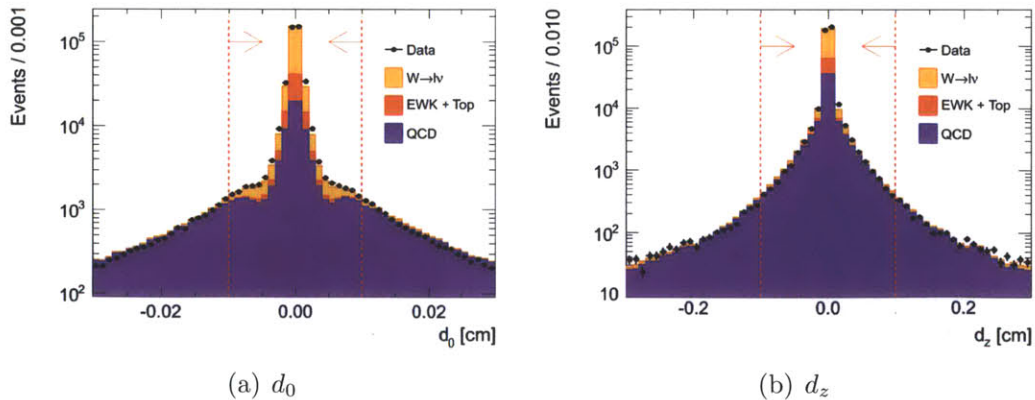


Figure 4-3: Distributions of the transverse ( $d_0$ ) and longitudinal ( $d_z$ ) impact parameter for muons. The dashed line and arrow indicates the region to be selected by the impact parameter cut.

## 4.2.4 Electron Identification

Electron identification starts with electron candidates reconstructed by both ECAL-seeded and track-seeded methods. Electrons from photon conversions are rejected by requiring no missing expected hits closer to the interaction point than the innermost hit on the electron track, as well as a low fit probability from the track-based conversion finder. The distribution of the number of missing expected hits is shown in Figure 4-4, from a selection of isolated electrons.

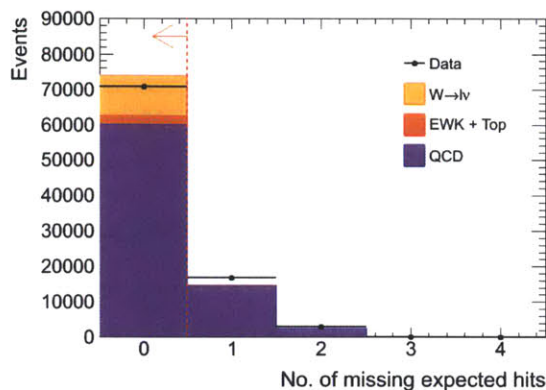


Figure 4-4: Distributions of the number of missing expected hits before the innermost tracker hit. The dashed line and arrow indicates the region to be selected by the cut.

To reduce contamination from hadronic processes, requirements are imposed on the shower development in the calorimeters. The longitudinal spread should be well contained within the ECAL, hence a low ratio ( $H/E$ ) of the HCAL deposit to the ECAL deposit in the calorimeter tower is required. The  $H/E$  distributions for isolated electrons are shown in Figure 4-5.

The transverse shower spread in the ECAL supercluster is expected to be smaller for electrons and photons compared to hadrons, and this is quantified by a low  $\sigma_{i\eta i\eta}$ , which is a quantity defined as,

$$\sigma_{i\eta i\eta}^2 = \frac{\sum_i^{5 \times 5} w_i (i\eta_i - i\eta_{\text{seed}})^2}{\sum_i^{5 \times 5} w_i}, \quad w_i = \max\left(0, 4.7 + \ln \frac{E_i}{E_{5 \times 5}}\right) \quad (4.2)$$

Essentially,  $\sigma_{i\eta i\eta}$  is the RMS shower width in the  $\eta$  direction, computed from the

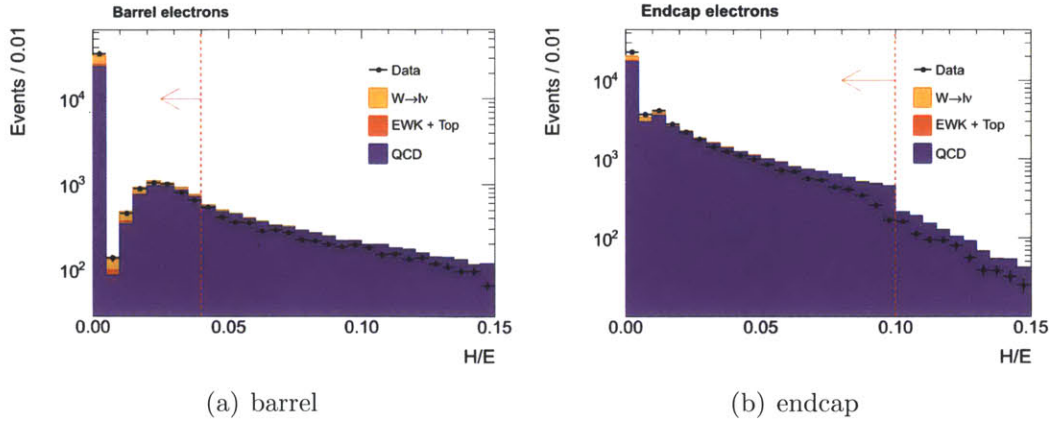


Figure 4-5: Distributions of  $H/E$  separated for electrons in the barrel and endcap. The dashed line and arrow indicates the region to be selected by the cut.

energy weighted displacements of the crystals in a  $5 \times 5$  array around the seed crystal. The  $\sigma_{i\eta i\eta}$  distributions for isolated electrons are shown in Figure 4-6.

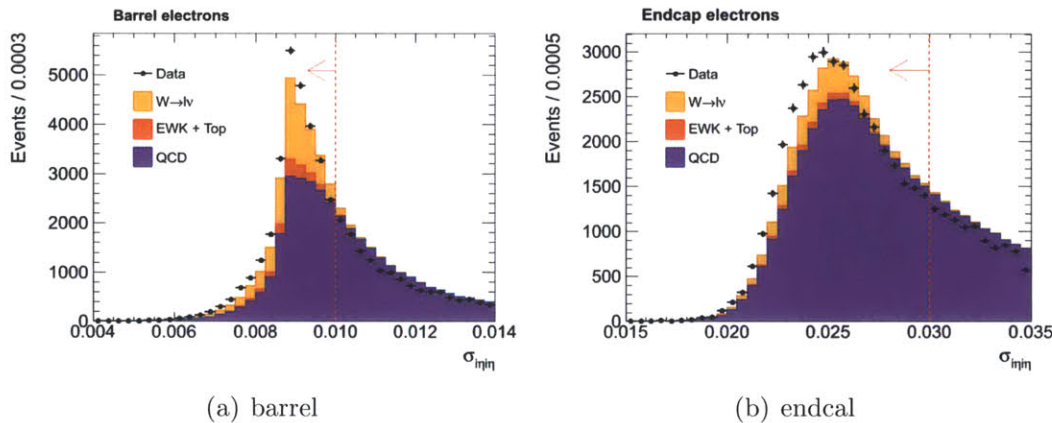


Figure 4-6: Distributions of  $\sigma_{i\eta i\eta}$  separated for electrons in the barrel and endcap. The dashed line and arrow indicates the region to be selected by the cut.

Finally, the track and supercluster should be well matched for an electron. Hence, the differences in  $\phi$  and in  $\eta$  between the supercluster position and the direction of the track at the primary vertex are required to be small. The reasoning for considering the track parameters at the primary vertex rather than at the tracker-ECAL interface is that the ECAL typically collects bremsstrahlung photons into the same supercluster, and the energy weighted position of the supercluster should be consistent with the electron track at its origin. The distributions are shown in Figures 4-7 and 4-8. The

suite of electron identification cuts are summarised in Table 4.4.

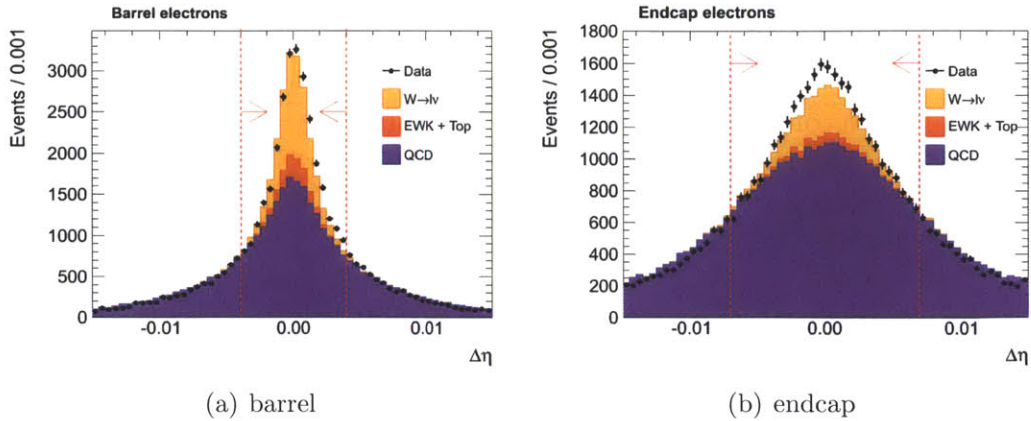


Figure 4-7: Distributions of  $\Delta\eta$  separated for electrons in the barrel and endcap. The dashed line and arrow indicates the region to be selected by the cut.

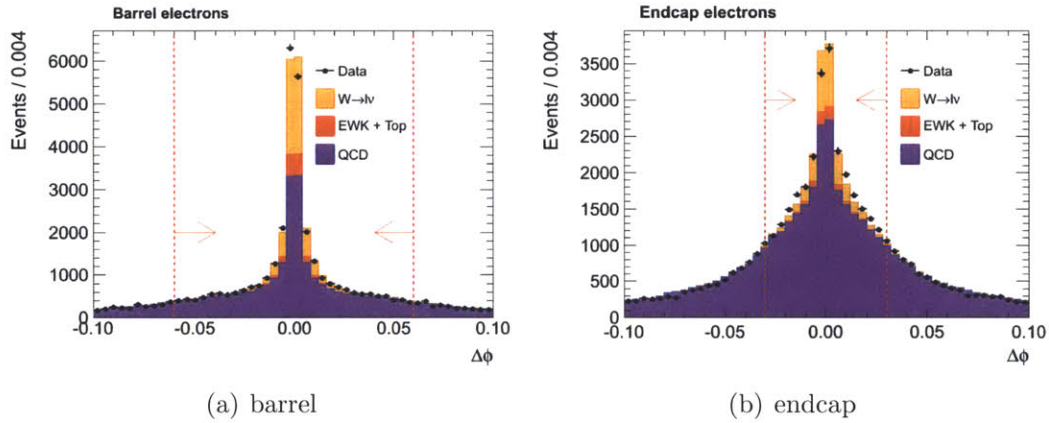


Figure 4-8: Distributions of  $\Delta\phi$  separated for electrons in the barrel and endcap. The dashed line and arrow indicates the region to be selected by the cut.

### 4.2.5 Muon Identification

The muons are required to be reconstructed by both the global muon (“outside-in”) and tracker muon (“inside-out”) algorithms. Compared to electron reconstruction, muons are much less likely to be faked by jets. The identification cuts for muons serve primarily to ensure the momentum is well measured by requiring sufficiently high number of hits in the inner tracker and muon systems, and that the track fit is

Quantity	Barrel	Endcap
Iso/ $p_T$	< 0.1	< 0.1
$ d_z $	0.1 cm	0.1 cm
$ d_0 $	0.02 cm	0.02 cm
No. of missing expected hits	= 0	= 0
Conversion fit probability	< $10^{-6}$	< $10^{-6}$
$\sigma_{i\eta i\eta}$	< 0.01	< 0.03
$ \Delta\phi_{\text{in}} $	< 0.06	< 0.03
$ \Delta\eta_{\text{in}} $	< 0.004	< 0.007
$H/E$	< 0.04	< 0.1

Table 4.4: Electron selection criteria

robust. The cut values are summarized in Table 4.5. The distributions are shown in Figures 4-9 to 4-11 for isolated muons.

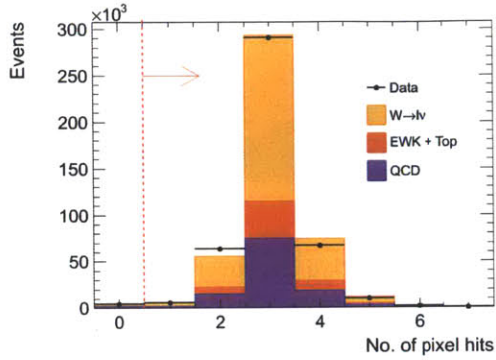
Quantity	Requirement
Iso/ $p_T$	< 0.15
$ d_z $	< 0.1 cm
$ d_0 $	< 0.02 cm
No. of tracker hits	> 10
No. of pixel hits	> 0
$\sigma(p_T)/p_T$	< 0.1
$\chi^2/\text{ndf}$	< 10
No. of chambers matched to segments	> 1
No. of valid muon hits in global fit	> 0

Table 4.5: Muon selection criteria

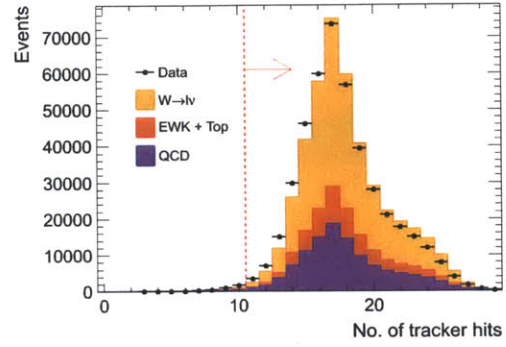
## 4.2.6 Yields

The expected and observed yields are detailed in Table 4.6. The contamination from background processes is estimated from simulation to be 0.45% of the total yield, in both the electron channel and the muon channel. The  $W$  and  $t\bar{t}$  samples are generated with POWHEG, the single top and  $Z \rightarrow \tau\tau$  samples are generated with POWHEG, and the  $ZZ$  and  $WZ$  samples are generated with PYTHIA.



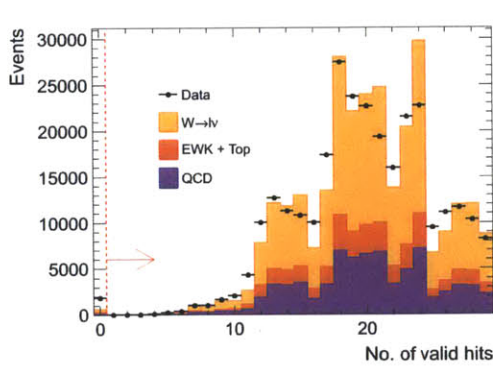


(a) pixel hits

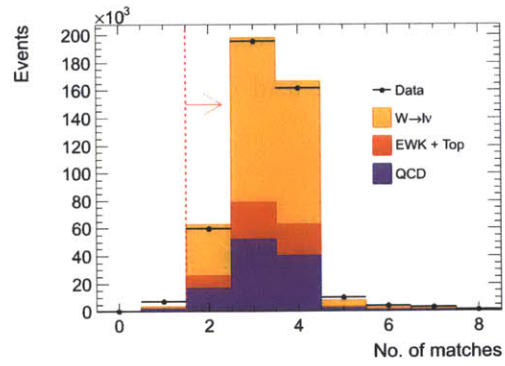


(b) inner tracker hits

Figure 4-9: Distributions of the number of pixel hits and inner tracker hits. The dashed line and arrow indicates the region to be selected by the cut.



(a) hits in global fit



(b) muon chambers

Figure 4-10: Distributions of the number of valid hits used in the global muon fit, and the number of chambers matched to muon segments. The dashed line and arrow indicates the region to be selected by the cut.

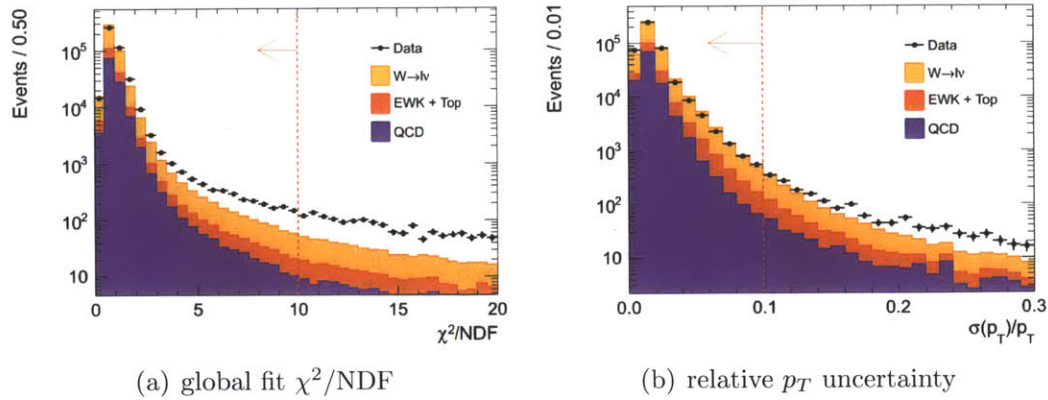


Figure 4-11: Distributions of the normalised  $\chi^2$  of the global muon fit and the relative uncertainty on the  $p_T$  measurement. The dashed line and arrow indicates the region to be selected by the cut.

Process	$ee$	$\mu\mu$
Top	$1655 \pm 5$	$2731 \pm 7$
$WW$	$348 \pm 3$	$611 \pm 4$
$WZ$	$868 \pm 4$	$1442 \pm 6$
$ZZ$	$679 \pm 3$	$1153 \pm 3$
$W \rightarrow \ell\nu$ ( $\ell = e, \mu, \tau$ )	$116 \pm 15$	$36 \pm 8$
$Z \rightarrow \tau\tau$	$1171 \pm 23$	$2066 \pm 31$
$Z \rightarrow \ell\ell$ ( $\ell = e/\mu$ )	$1066744 \pm 581$	$1878315 \pm 774$
Expected	$1071581 \pm 582$	$1886355 \pm 775$
Observed	1072710	1898664

Table 4.6: Expected and observed yields. Uncertainties on the expected yields are statistical only.

## 4.2.7 Distributions

The distributions for the dilepton mass and other kinematic quantities for events passing selection are presented in this subsection.

### Muon channel

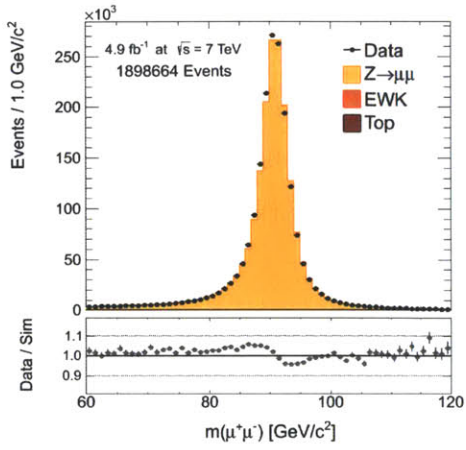
The mass, transverse momentum, and rapidity distributions for selected dimuon events are shown in Figure 4-12. The mass and rapidity distributions are well described by simulation. The boson  $p_T$  distribution has discrepancies with simulation at



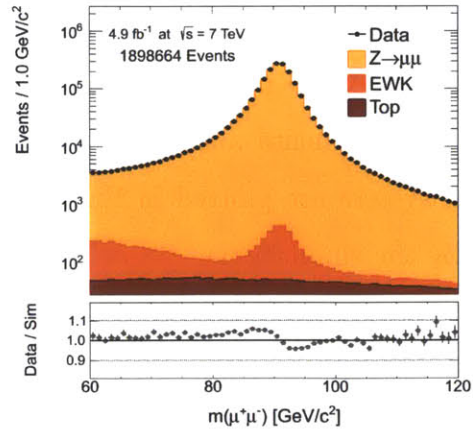
the very low and the very high  $p_T$  regions, which are consequences from the shortcomings of the POWHEG description. POWHEG does not have a treatment for the soft, non-perturbative QCD regime, and the perturbative calculation is only up to NLO. The transverse momenta and pseudorapidities of the leading and trailing muons in the dimuon system are plotted in Figure 4-13; the bulk of the distributions are well described by the simulation.

### **Electron channel**

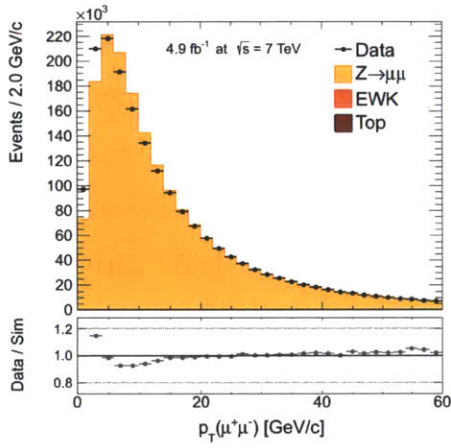
The mass, transverse momentum, and rapidity distributions for selected dielectron events are shown in Figure 4-14. In contrast to the muon channel, the mass distribution exhibits a relatively large discrepancy between data and simulation; the calibration of the ECAL is slightly worse in data and dilepton mass is the most sensitive quantity to miscalibration. In the context of the inclusive cross section measurement, this is a negligible effect. Effective electron energy scale and resolution corrections can be derived by comparing the observed  $Z$  line shape with simulation, since the simulation represents the ideal performance for calibration. Energy scale corrections are applied to events in data. Gaussian smearing is applied to simulated events to mimic the resolution in data. To first order, the severity of miscalibration depends on where the electron is measured in the ECAL, therefore the corrections are parametrised in  $|\eta|$  of the supercluster. Details of the procedure are described in Appendix A. A summary of the energy scale and resolution corrections are shown in Figure 4-15. A comparison of the mass distribution before and after these effective corrections is shown in Figure 4-16. As a result of the energy scale corrections, the observed yield increases by 0.04%, from 1 072 710 events to 1 073 123 events. As a result of the resolution correction to the simulation, the estimated signal drops by 0.006%, from 1 071 581 events to 1 071 515 events. Since the magnitude of these changes are so minute, these corrections are not incorporated into the cross section measurement.



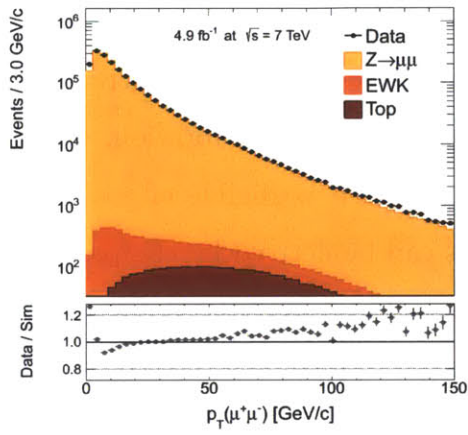
(a) dimuon mass (linear)



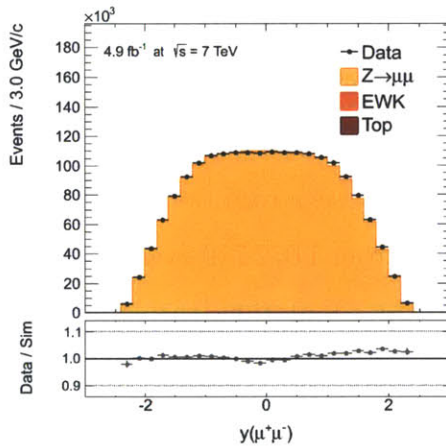
(b) dimuon mass (log)



(c) dimuon  $p_T$  (linear)

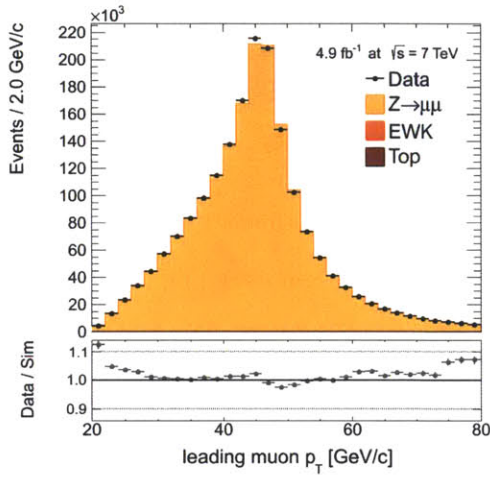


(d) dimuon  $p_T$  (log)

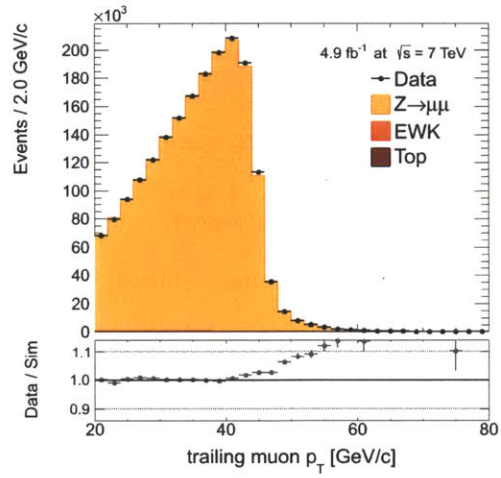


(e) dimuon rapidity

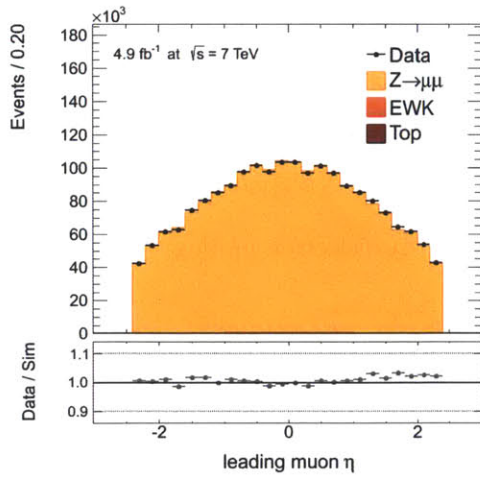
Figure 4-12: The dimuon mass, transverse momentum, and rapidity distributions.



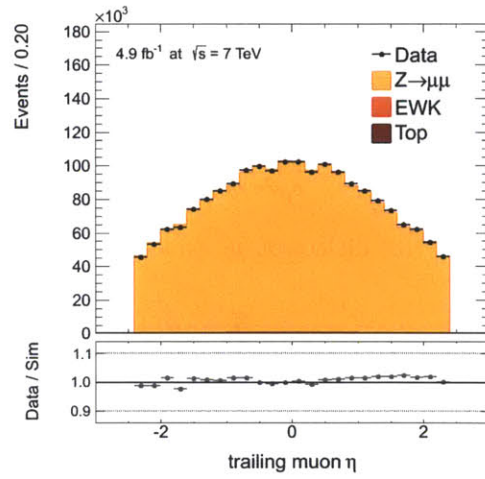
(a) leading muon  $p_T$



(b) trailing muon  $p_T$

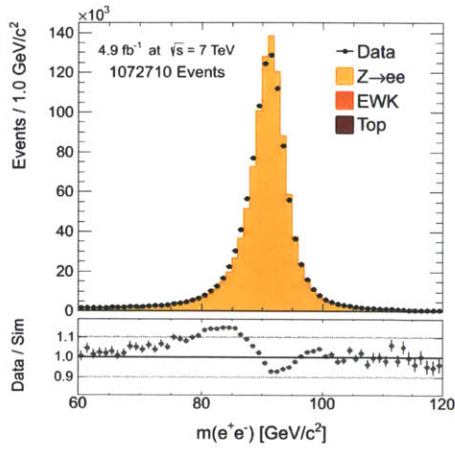


(c) leading muon  $\eta$

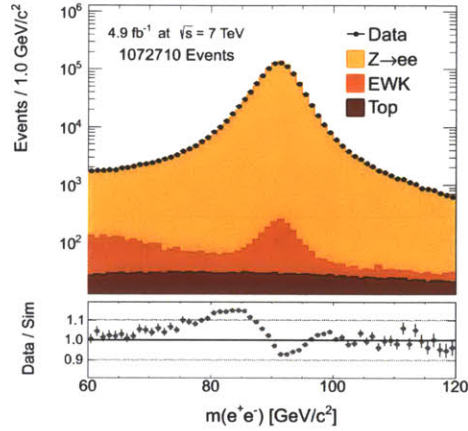


(d) trailing muon  $\eta$

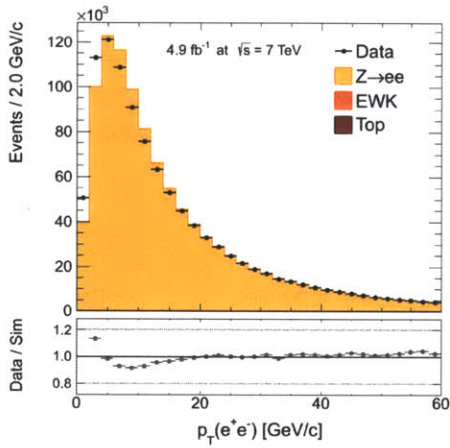
Figure 4-13: Transverse momentum and pseudorapidity distributions of leading and trailing muons.



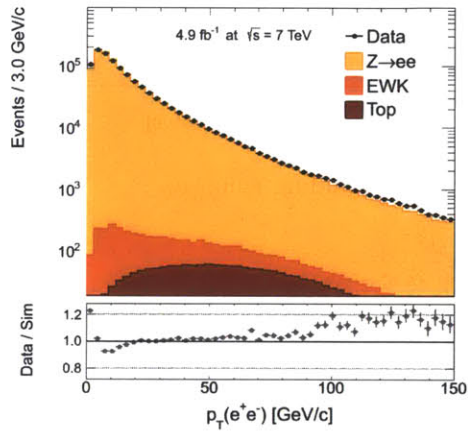
(a) dielectron mass (linear)



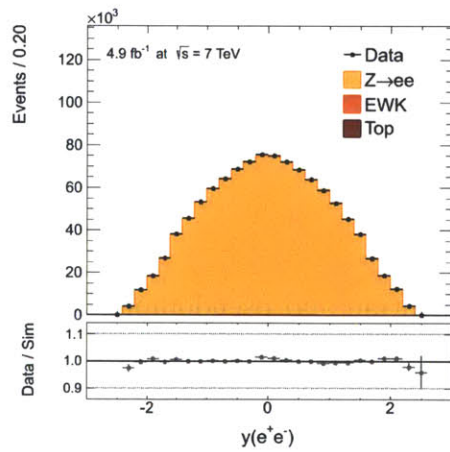
(b) dielectron mass (log)



(c) dielectron  $p_T$  (linear)



(d) dielectron  $p_T$  (log)



(e) dielectron rapidity

Figure 4-14: The dielectron mass, transverse momentum, and rapidity distributions.

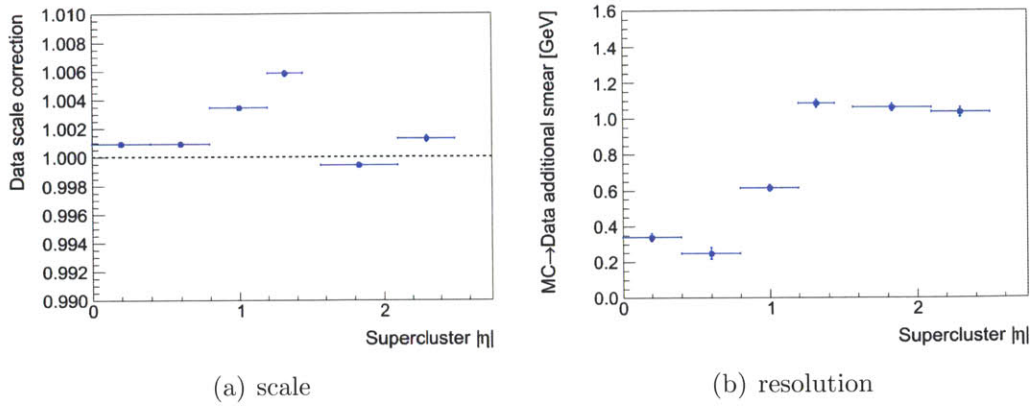


Figure 4-15: The electron energy scale corrections to be applied to data events and the additional Gaussian smearing to be applied to simulated events to achieve better agreement with observation.

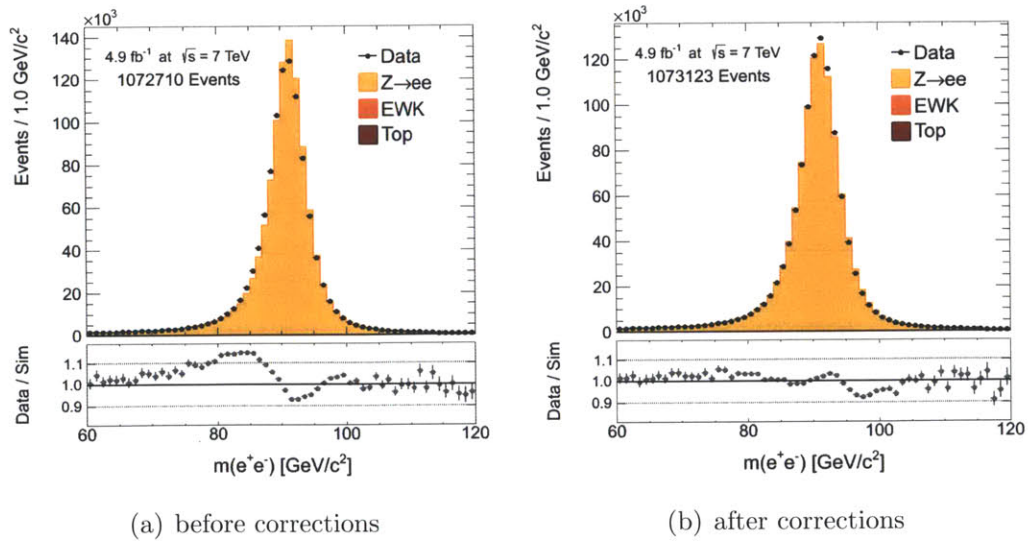
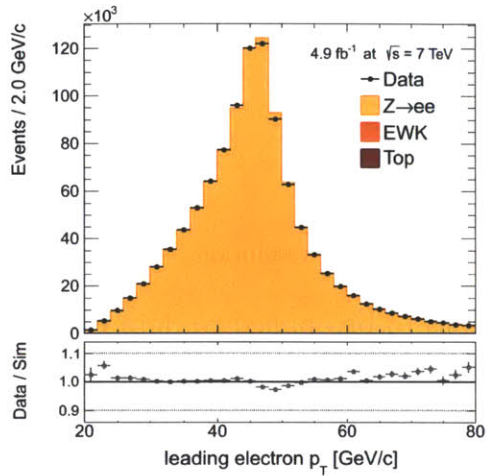
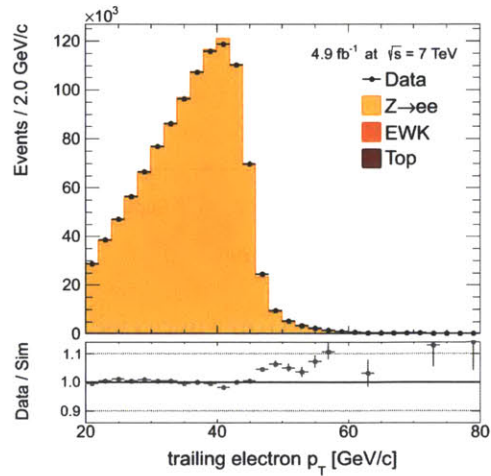


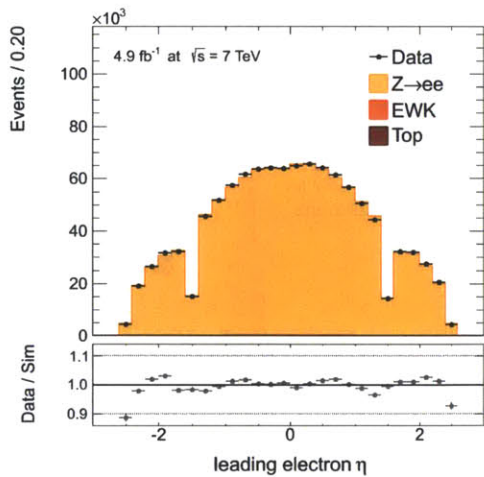
Figure 4-16: Dielectron mass distributions before and after effective energy scale and resolution corrections.



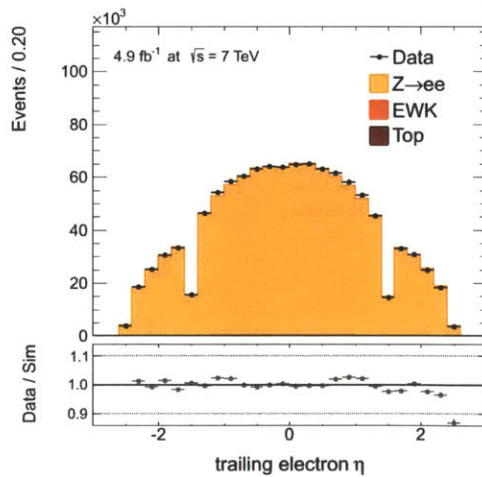
(a) leading electron  $p_T$



(b) trailing electron  $p_T$



(c) leading electron  $\eta$



(d) trailing electron  $\eta$

Figure 4-17: Transverse momentum and pseudorapidity distributions of leading and trailing electrons.

## 4.3 Acceptance

The acceptance is defined as the fraction of signal events that pass offline event selection computed from simulation. Unfortunately, there does not exist a collection of tools that can produce a signal sample to account for all QCD and electroweak effects with the best available theoretical calculations. Having chosen POWHEG with CT10 [47] combined with PYTHIA to be the baseline event description, systematic uncertainties are assigned to account for missing higher order corrections (QCD and electroweak) and uncertainties on the models themselves (parton distribution function). Sometimes, the estimation for a systematic uncertainty requires the use of a generator tool that does not provide the full kinematic information of the hard scattering process, so a simulated data sample complete with detector simulation and event reconstruction cannot be produced. In such cases, the acceptance is computed at “generator level”,  $A_{\text{GEN}}$ , which is the fraction of events with daughter leptons passing  $p_T$  and  $\eta$  cuts only. Table 4.7 lists the acceptances computed with the baseline signal sample.

Quantity	$ee$	$\mu\mu$
$A_{\text{GEN}}$	0.4786	0.4689
$A$	0.2407	0.4177

Table 4.7: Acceptance values at generator level and after full selection.

### 4.3.1 QCD resummation and NNLO corrections

The nominal model is accurate only to leading logarithmic order (LL) for soft, non-perturbative QCD effects – a limitation of the parton shower approximation implemented in PYTHIA. A description accurate to just beyond next-to-next-to-leading logarithms (NNLL) can be attained with a resummation procedure [48, 49]. Furthermore, POWHEG calculations are NLO in perturbative QCD while the most precise calculations available are NNLO. The RESBos [50] generator implements both the resummation and NNLO calculations which are missing in the baseline description. Of



the two corrections, the most significant one on the acceptance is due to the resummation treatment, which can be seen in the shapes of the  $Z$  boson  $p_T$  distributions from POWHEG and RESBos in Figure 4-18. Unfortunately, RESBos does not provide the kinematics for the scattered or radiated partons, so it cannot be used for full event generation. Generator level acceptances are compared to quantify the systematic uncertainty. RESBos incorporates PHOTOS [51] to simulate FSR and this feature is enabled. The systematic uncertainties from missing QCD resummation and NNLO corrections are estimated to be 0.84% and 1.08% in the electron and muon channels respectively.

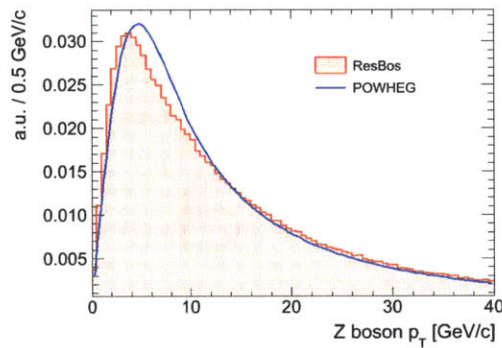


Figure 4-18: Comparison of the  $Z$  boson  $p_T$  distributions generated from POWHEG and ResBos.

### 4.3.2 Higher order QCD corrections

Fixed order cross section calculations depend on renormalization ( $\mu_R$ ) and factorization ( $\mu_F$ ) scales. Differences found when varying  $\mu_R$  and  $\mu_F$  are taken to indicate the uncertainty on the cross section from missing higher perturbative orders. RESBos cannot be used to evaluate this systematic uncertainty because it does not provide the option to vary  $\mu_R$  and  $\mu_F$ . These scale variation calculations can be performed using FEWZ [52, 53], an integrator that computes cross sections accurate to NNLO QCD. It is possible to impose requirements on the daughter lepton kinematics and boson mass range in FEWZ to compute a restricted cross section. Generator level

acceptances are considered and computed as follows,

$$A_\mu = \frac{\sigma'(\mu_R = \mu_F = \mu)}{\sigma(\mu_R = \mu_F = \mu)}, \quad \text{for } \mu = M, 2M, M/2, \quad (4.3)$$

where  $\sigma$  is the cross section with no requirements on decay daughters,  $\sigma'$  is the restricted cross section, and the acceptance,  $A_\mu$ , is computed at three different values for the  $\mu_R$  and  $\mu_F$  based on the boson mass,  $M$ . The variation from scale dependence is taken to be half of the largest pair-wise difference among  $\{A_M, A_{2M}, A_{M/2}\}$ , namely,

$$\delta_{\text{scale}} = \frac{1}{2} \cdot \max \{|A_M - A_{M/2}|, |A_M - A_{2M}|, |A_{2M} - A_{M/2}|\}. \quad (4.4)$$

The systematic uncertainties for missing beyond NNLO corrections in QCD are estimated to be 0.4% in both the electron and muon channels.

### 4.3.3 Parton distribution function

Parton distribution functions are obtained from global fits to experimental data from deep inelastic scattering, Drell-Yan, and jet processes. Hence, the parameters describing the proton structure have associated uncertainties. The CT10 PDF, provided by the CTEQ collaboration, contains 26 parameters derived with  $\alpha_s$  fixed to the world average of 0.118. The CTEQ group also provides two error PDFs per parameter corresponding to the upper and lower bounds of the 90% CL interval around the central value, as well as PDFs for  $\alpha_s$  fixed to different values. To propagate the uncertainty on the CT10 PDF to the acceptance, the uncertainty from  $\alpha_s$  and the uncertainty from the 26 parameters are estimated separately and then combined. Rather than produce fully simulated events samples for different error PDF sets, an event by event re-weighting procedure is done by scaling one PDF to another, and the acceptance calculation is repeated.

The  $\alpha_s$  uncertainty is propagated by considering PDFs for  $\alpha_s = 0.116, 0.120$ , as the uncertainty on  $\alpha_s$  is  $\pm 0.002$  at 90% CL. The result is a 0.1% uncertainty on the acceptance for both lepton channels.

To propagate the uncertainty on the CT10 parameters, the CTEQ group prescribes the following formula for deriving the uncertainty interval,

$$\begin{aligned}\Delta A^+ &= \sqrt{\sum_{i=1}^N [\max(A_i^+ - A_0, A_i^- - A_0, 0)]^2}, \\ \Delta A^- &= \sqrt{\sum_{i=1}^N [\max(A_0 - A_i^+, A_0 - A_i^-, 0)]^2},\end{aligned}\tag{4.5}$$

where  $A_0$  is computed with the central value PDF,  $i$  runs over the error sets,  $A_i^+$  and  $A_i^-$  denote respectively the acceptances with the positive and negative error PDFs for the  $i^{\text{th}}$  eigenvector in the parameter space, and  $N = 26$  for CT10 (52 error PDFs). The final PDF+ $\alpha_s$  uncertainty on the acceptance is obtained by adding the PDF uncertainty and the  $\alpha_s$  uncertainty in quadrature, taking the larger of  $\Delta A^+$  and  $\Delta A^-$  to form a symmetric uncertainty band, and dividing the uncertainty by 1.645 to obtain the more conventional 68% CL. The PDF+ $\alpha_s$  uncertainty is 1.04% in the electron channel and 0.92% in the muon channel.

#### 4.3.4 Electroweak corrections

At the energy scale of weak boson production,  $O(\alpha)$  electroweak effects (NLO EWK) are of similar order to  $O(\alpha_s^2)$  QCD effects (NNLO QCD). NLO EWK corrections are partially accounted for in the baseline sample since PYTHIA generates initial and final state photon radiation in the parton shower approximation. However, corrections involving loops with virtual bosons are missing. To quantify the effect of NLO EWK corrections, the HORACE [54–57] event generator is used. HORACE implements  $Z$  production and subsequent decay into leptons accurate to LO QCD and NLO EWK. It also uses the parton shower method to account for FSR beyond single photon emission. The output of HORACE contains all the kinematic information required to produce samples with detector simulation and reconstruction. Working with the acceptance after full selection rather than at generator level acceptance is particularly important for electrons where much of the energy lost to radiated photons can be

recovered in the ECAL or accounted for by the reconstruction.

There are two sources of systematic uncertainties associated with electroweak corrections. Firstly, there are the missing NLO EWK effects which can be quantified by comparing HORACE performing its full suite of corrections with HORACE accounting for just the FSR. Secondly, since FSR is described by PYTHIA in the baseline sample, the difference in acceptances between HORACE (FSR only) and PYTHIA is taken to indicate the systematic uncertainty in FSR modelling. The effect of QED ISR on the acceptance was found to be insignificant, by comparing PYTHIA samples with QED ISR enabled and disabled. Hence, the uncertainty on QED ISR modelling is omitted.

### 4.3.5 Summary of Systematic Uncertainties on the Acceptance

Table 4.8 lists all the systematic uncertainties on the acceptance. The uncertainties from various sources are summed in quadrature to obtain a total uncertainty for each lepton channel.

Quantity	ISR+NNLO	>NNLO	PDF	FSR	EWK	Total
$Z$ acceptance ( $e$ )	0.84%	0.40%	1.04%	0.54%	0.84%	1.72%
$Z$ acceptance ( $\mu$ )	1.08%	0.40%	0.92%	0.25%	1.08%	1.84%

Table 4.8: Summary of systematic uncertainties on the acceptance.

## 4.4 Lepton Efficiency

The baseline acceptance is computed assuming offline reconstruction and selection efficiencies in simulation and excludes inefficiencies from triggering. The efficiencies in data must be measured and then incorporated to obtain a more accurate value for the acceptance.

The offline electron reconstruction and selection efficiency is defined with respect

to a reconstructed ECAL supercluster and can be factorized into two components,

$$\varepsilon = \varepsilon_{\text{GSF}} \cdot \varepsilon_{\text{Sel}}, \quad (4.6)$$

where  $\varepsilon_{\text{GSF}}$  is the efficiency to reconstruct an ECAL-seeded GSF-electron given a supercluster, and  $\varepsilon_{\text{Sel}}$  is the efficiency for the GSF-electron to pass selection requirements.

The offline muon reconstruction and selection efficiency can be factorized as follows,

$$\varepsilon = \varepsilon_{\text{STA}} \cdot \varepsilon_{\text{TRK}} \cdot \varepsilon_{\text{Sel}}, \quad (4.7)$$

where  $\varepsilon_{\text{STA}}$  is the efficiency for a muon to be reconstructed as a standalone muon,  $\varepsilon_{\text{TRK}}$  is the efficiency for a muon to be reconstructed as a tracker track, and  $\varepsilon_{\text{Sel}}$  is the efficiency for a global muon to pass selection requirements. Note the global muon reconstruction efficiency is estimated by the product,  $\varepsilon_{\text{STA}} \cdot \varepsilon_{\text{TRK}}$ .

For both muons and electrons, the trigger efficiency is defined with respect to a lepton passing offline selection requirements.

#### 4.4.1 Tag-and-Probe Method

The efficiency of lepton reconstruction and selection are measured with  $Z \rightarrow \ell\ell$  events, using a method sometimes known as the “tag-and-probe” (T&P) technique. Dilepton events are selected, with mass between  $60 \text{ GeV}/c^2$  and  $120 \text{ GeV}/c^2$ , where one of the leptons must have  $p_T > 20 \text{ GeV}$  and satisfy the analysis selection requirements for electrons or muons; this lepton is denoted the “tag” lepton. The other lepton is denoted the “probe”. The efficiency of a particular selection requirement is the fraction of probes that pass the requirement. For the offline selection, the quantity of interest is the efficiency scale factor, defined as the ratio of the efficiency in data to the efficiency in simulation.

For a given set of selection criteria, the tag-probe candidates can be divided into two exclusive categories: the probe passes or it fails. The efficiency of this selection

is then given by,

$$\varepsilon = \frac{N_{\text{pass}}}{N_{\text{pass}} + N_{\text{fail}}}, \quad (4.8)$$

where  $N_{\text{pass}}$  and  $N_{\text{fail}}$  are the number of passing probes and the number of failing probes, respectively. When background contamination (namely, jet induced backgrounds) is not negligible,  $N_{\text{pass}}$  and  $N_{\text{fail}}$  are the probe counts in each category after background subtraction. In such cases, binned likelihood fits to the mass distributions are performed and the two categories are fitted simultaneously to facilitate the propagation of fit uncertainties from both categories to the uncertainty on the extracted efficiency. A description of the signal and background models used will be described in the next subsection.

Since the probe sample is dependent on the tag definition, a bias may be introduced if the efficiency in the phase space sampled by the probes is significantly different from that of the selected leptons for the  $Z$  cross section measurement. To reduce the impact of such biases, the efficiency scale factors are computed with efficiencies derived from simulated events also computed with the tag-and-probe method,

$$\rho = \frac{\varepsilon_{\text{T\&P-Data}}}{\varepsilon_{\text{T\&P-MC}}}. \quad (4.9)$$

The reasoning here is that tag-probe correlations are similar in data and simulation, so that  $\rho$  approximates the ratio of the true efficiencies in data to simulation better than  $\varepsilon_{\text{T\&P-Data}}$  estimates the true efficiency in data.

## Signal and Background Models

When background contamination is significant, the modelling of the tag-probe mass spectrum is important for an accurate estimate of the efficiency. In particular, the signal shape in the fail category may be afflicted by increased energy scale bias, degraded resolution, or greater amount of final state radiation (FSR) and bremsstrahlung compared to the mass spectrum of the pass category.

The signal model used is the convolution of the reconstructed mass spectrum from simulation with a Gaussian function ( $\text{Sim} * \text{Gaus}$ ): The T&P selection is ap-

plied in a simulated  $Z \rightarrow \ell\ell$  sample to produce the template. Consequently, FSR, bremsstrahlung, and detector effects are modelled by PYTHIA and GEANT4. Residual energy scale shifts and resolution degradation in data relative to the simulation are accounted for by the Gaussian. The model has two floating parameters.

The background model is a complementary error function multiplied with an exponential ( $erfc \times exp$ ):

$$b(m) = \text{erfc}(\beta(\alpha - m)) \cdot \exp(-\gamma(m - M_Z)), \quad (4.10)$$

where

$$\text{erfc}(x) = \frac{2}{\pi} \int_x^\infty e^{-t^2} dt. \quad (4.11)$$

This model can account for the kinematic turn-on in the mass spectrum due to high  $p_T$  cuts on the probe. This feature becomes important when measuring efficiencies binned in  $p_T$ . There are three floating parameters.

## Triggers for Efficiency Measurements

It is crucial to ensure the probe leg can be selected with minimal bias from the trigger. For muon efficiencies, single muon triggers are available and allow events to be collected for efficiency measurements that do not depend on any property of the probe muon. For electron efficiencies, dielectron triggers are available that are dedicated for efficiency measurements (and differ from the ones used for event selection in the cross section measurement). These triggers require at least one leg to be a high energy electron ( $E_T > 17\text{GeV}$ ) with stringent isolation and identification cuts, while the other leg is an electron or even just a supercluster passing a much lower energy threshold and without any isolation or identification requirements.

### 4.4.2 Measured Efficiencies and Scale Factors

This section will present the measured trigger efficiencies in data and  $\rho$  for offline selection. The efficiencies are extracted in bins of lepton  $p_T$  and  $|\eta|$ . A few examples

of fits to the mass distribution will be shown in the figures of this section; a complete documentation of the efficiency measurements can be found in Appendix B.

### Trigger efficiency

The efficiency of the trigger is measured with respect to an electron or muon passing offline selection. The background expectation is at the level of 0.45%. Hence, backgrounds are negligible and the signal yields in both the pass and fail categories are extracted by counting events in bins of  $p_T$  and  $|\eta|$ . The efficiencies for each leg of the dielectron trigger, averaging the various triggers employed over the whole run, are shown in Tables 4.9 and 4.10. This translates to an efficiency of better than 99% to trigger on a  $Z \rightarrow ee$  event occurring within acceptance. Results for the dimuon trigger are shown in Tables 4.9 and 4.10, for an efficiency of about 93% to trigger on  $Z \rightarrow \mu\mu$ . The single muon efficiency is listed in Table 4.13, and provides an additional 5% efficiency to triggering on dimuon events.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.9891 \pm 0.0007$	$0.9921 \pm 0.0007$	$0.9895 \pm 0.0011$	$0.9805 \pm 0.0023$
$30 < E_T < 40$	$0.9952 \pm 0.0002$	$0.9963 \pm 0.0003$	$0.9946 \pm 0.0004$	$0.9926 \pm 0.0008$
$40 < E_T < 50$	$0.9963 \pm 0.0002$	$0.9978 \pm 0.0002$	$0.9964 \pm 0.0003$	$0.9944 \pm 0.0007$
$50 < E_T < 60$	$0.9968 \pm 0.0004$	$0.9980 \pm 0.0004$	$0.9981 \pm 0.0005$	$0.9950 \pm 0.0016$
$E_T > 60$	$0.9970 \pm 0.0007$	$0.9979 \pm 0.0008$	$0.9988 \pm 0.0012$	$0.9935 \pm 0.0049$

Table 4.9: Efficiency to pass the leading electron requirement in the dielectron trigger, measured in data.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.9921 \pm 0.0006$	$0.9952 \pm 0.0006$	$0.9932 \pm 0.0011$	$0.9912 \pm 0.0019$
$30 < E_T < 40$	$0.9959 \pm 0.0002$	$0.9968 \pm 0.0003$	$0.9960 \pm 0.0004$	$0.9939 \pm 0.0008$
$40 < E_T < 50$	$0.9968 \pm 0.0002$	$0.9982 \pm 0.0002$	$0.9973 \pm 0.0003$	$0.9955 \pm 0.0006$
$50 < E_T < 60$	$0.9975 \pm 0.0003$	$0.9987 \pm 0.0003$	$0.9989 \pm 0.0005$	$0.9970 \pm 0.0014$
$E_T > 60$	$0.9977 \pm 0.0006$	$0.9989 \pm 0.0007$	$0.9993 \pm 0.0012$	$0.9955 \pm 0.0046$

Table 4.10: Efficiency to pass the trailing electron requirement in the dielectron trigger, measured in data.



	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9758 \pm 0.0008$	$0.9597 \pm 0.0015$	$0.9454 \pm 0.0013$	$0.8726 \pm 0.0033$
$30 < p_T < 40$	$0.9756 \pm 0.0005$	$0.9567 \pm 0.0010$	$0.9473 \pm 0.0008$	$0.8747 \pm 0.0024$
$40 < p_T < 50$	$0.9751 \pm 0.0004$	$0.9585 \pm 0.0008$	$0.9503 \pm 0.0007$	$0.8785 \pm 0.0026$
$50 < p_T < 60$	$0.9742 \pm 0.0011$	$0.9596 \pm 0.0018$	$0.9484 \pm 0.0017$	$0.8843 \pm 0.0052$
$p_T > 60$	$0.9745 \pm 0.0015$	$0.9575 \pm 0.0027$	$0.9482 \pm 0.0024$	$0.8680 \pm 0.0099$

Table 4.11: Efficiency to pass the leading muon requirement in the dimuon trigger, measured in data.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9786 \pm 0.0006$	$0.9651 \pm 0.0012$	$0.9622 \pm 0.0008$	$0.9159 \pm 0.0025$
$30 < p_T < 40$	$0.9787 \pm 0.0004$	$0.9635 \pm 0.0008$	$0.9609 \pm 0.0006$	$0.9123 \pm 0.0017$
$40 < p_T < 50$	$0.9781 \pm 0.0004$	$0.9650 \pm 0.0007$	$0.9620 \pm 0.0005$	$0.9153 \pm 0.0019$
$50 < p_T < 60$	$0.9773 \pm 0.0009$	$0.9652 \pm 0.0014$	$0.9606 \pm 0.0012$	$0.9194 \pm 0.0037$
$p_T > 60$	$0.9770 \pm 0.0013$	$0.9620 \pm 0.0024$	$0.9593 \pm 0.0017$	$0.9046 \pm 0.0078$

Table 4.12: Efficiency to pass the trailing muon requirement in the dimuon trigger, measured in data.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.3228 \pm 0.0007$	$0.3006 \pm 0.0012$	$0.3024 \pm 0.0009$	$0.2186 \pm 0.0014$
$30 < p_T < 40$	$0.8789 \pm 0.0005$	$0.8034 \pm 0.0010$	$0.7978 \pm 0.0008$	$0.2545 \pm 0.0011$
$40 < p_T < 50$	$0.8883 \pm 0.0005$	$0.8163 \pm 0.0009$	$0.8178 \pm 0.0007$	$0.2592 \pm 0.0011$
$50 < p_T < 60$	$0.8859 \pm 0.0011$	$0.8192 \pm 0.0019$	$0.8184 \pm 0.0014$	$0.2591 \pm 0.0023$
$p_T < 60$	$0.8742 \pm 0.0015$	$0.8026 \pm 0.0027$	$0.8084 \pm 0.0021$	$0.2612 \pm 0.0037$

Table 4.13: Efficiency to pass the single muon trigger, measured in data.

## GSF Reconstruction Efficiency

The efficiency of the GSF reconstruction algorithm is measured using supercluster probes. An example fit to the pass and fail categories for a particular  $|\eta| - p_T$  bin is shown in Figure 4-19. The efficiency scale factors are listed in Table 4.14, and shows that the efficiency in data is well described by the simulation.

## Electron Selection Efficiency

The electron selection efficiency is measured with GSF electron probes. An example fit is shown in Figure 4-20. The efficiency scale factors are listed in Table 4.15, and shows that the selection in data is a few percent worse than expected in the simulation.

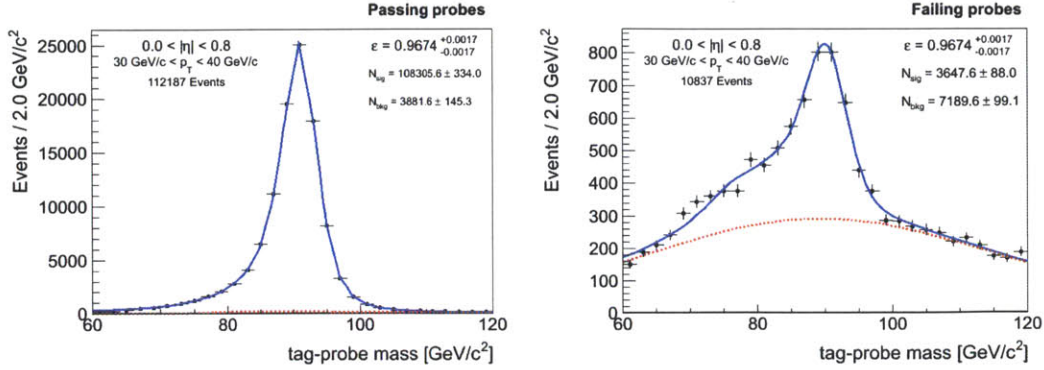


Figure 4-19: Fit for the GSF reconstruction efficiency for superclusters with  $30 \text{ GeV} < E_T < 40 \text{ GeV}$  and  $0 < |\eta| < 0.8$ .

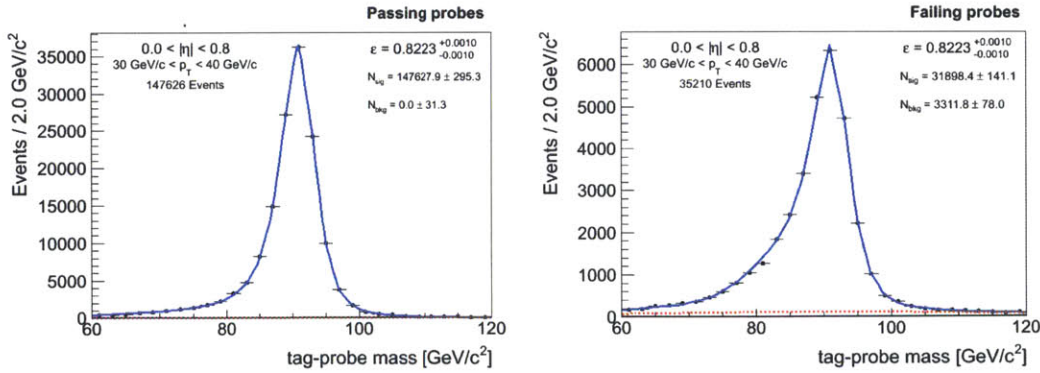


Figure 4-20: Fit for the selection efficiency for electrons with  $30 \text{ GeV} < E_T < 40 \text{ GeV}$  and  $0 < |\eta| < 0.8$ .

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.9995 \pm 0.0081$	$0.9796 \pm 0.0071$	$0.9977 \pm 0.0035$	$1.0008 \pm 0.0061$
$30 < E_T < 40$	$0.9944 \pm 0.0012$	$0.9966 \pm 0.0017$	$0.9963 \pm 0.0036$	$0.9926 \pm 0.0050$
$40 < E_T < 50$	$0.9958 \pm 0.0007$	$0.9969 \pm 0.0008$	$0.9964 \pm 0.0012$	$0.9924 \pm 0.0025$
$50 < E_T < 60$	$0.9922 \pm 0.0027$	$0.9937 \pm 0.0019$	$0.9962 \pm 0.0036$	$0.9903 \pm 0.0052$
$E_T > 60$	$0.9962 \pm 0.0054$	$0.9956 \pm 0.0025$	$1.0013 \pm 0.0056$	$0.9991 \pm 0.0193$

Table 4.14: Efficiency scale factors for GSF reconstruction efficiency.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.9728 \pm 0.0040$	$0.9494 \pm 0.0051$	$0.9908 \pm 0.0078$	$0.9604 \pm 0.0103$
$30 < E_T < 40$	$0.9785 \pm 0.0018$	$0.9614 \pm 0.0025$	$0.9752 \pm 0.0043$	$0.9463 \pm 0.0075$
$40 < E_T < 50$	$0.9769 \pm 0.0014$	$0.9734 \pm 0.0039$	$0.9789 \pm 0.0035$	$0.9466 \pm 0.0063$
$50 < E_T < 60$	$0.9779 \pm 0.0029$	$0.9785 \pm 0.0035$	$0.9885 \pm 0.0070$	$0.9390 \pm 0.0132$
$E_T > 60$	$0.9755 \pm 0.0040$	$0.9694 \pm 0.0052$	$0.9996 \pm 0.0110$	$0.9493 \pm 0.0212$

Table 4.15: Efficiency scale factors for electron selection efficiency.

## Standalone Muon Reconstruction Efficiency

The standalone muon reconstruction efficiency is measured with isolated tracker tracks for probes. The isolation requirement on the probe greatly improves the signal to background ratio, reducing the systematic uncertainties on the fit results, and induces negligible bias on the extracted efficiency. Figure 4-21 shows an example fit and the scale factors are listed in Table 4.16. While the efficiency in the barrel ( $|\eta| < 1.2$ ) is well described by the simulation, there is a difference of 1 – 2% in the endcap. The additional inefficiency in data is a result of a timing problem in the CSC subdetector during the last half of data taking.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9973 \pm 0.0007$	$0.9978 \pm 0.0010$	$0.9828 \pm 0.0019$	$0.9895 \pm 0.0017$
$30 < p_T < 40$	$0.9968 \pm 0.0004$	$0.9955 \pm 0.0006$	$0.9817 \pm 0.0005$	$0.9875 \pm 0.0012$
$40 < p_T < 50$	$0.9967 \pm 0.0003$	$0.9952 \pm 0.0005$	$0.9796 \pm 0.0004$	$0.9884 \pm 0.0010$
$50 < p_T < 60$	$0.9966 \pm 0.0009$	$0.9965 \pm 0.0011$	$0.9793 \pm 0.0008$	$0.9885 \pm 0.0015$
$p_T > 60$	$0.9951 \pm 0.0016$	$0.9921 \pm 0.0025$	$0.9805 \pm 0.0013$	$0.9978 \pm 0.0105$

Table 4.16: Efficiency scale factors for standalone muon reconstruction efficiency.

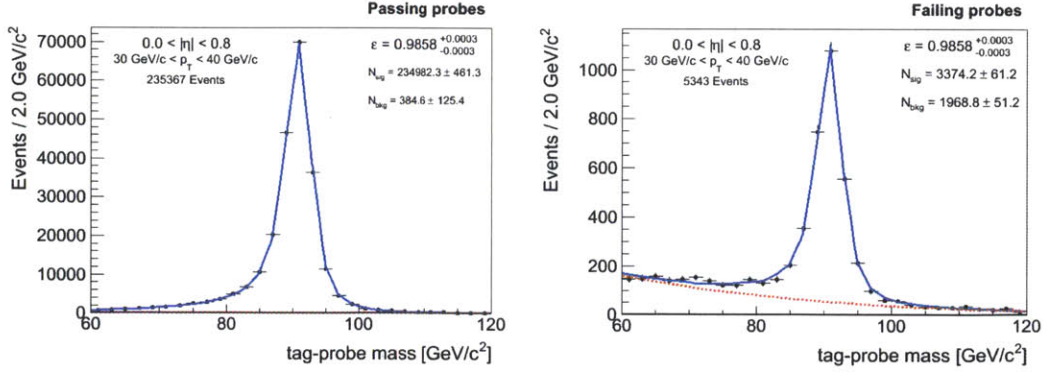


Figure 4-21: Fit for the standalone muon reconstruction efficiency for muons with  $30 \text{ GeV}/c < p_T < 40 \text{ GeV}/c$  and  $0 < |\eta| < 0.8$ .

## Tracking Efficiency

The tracking efficiency is measured with standalone muon probes. Figure 4-22 shows an example fit and the scale factors are listed in Table 4.17. The efficiency is well described by the simulation.

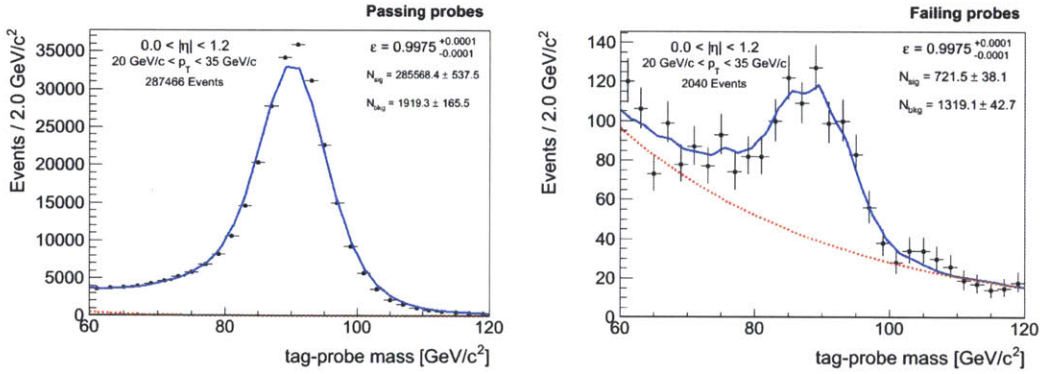


Figure 4-22: Fit for the track reconstruction efficiency for muons with  $20 \text{ GeV}/c < p_T < 35 \text{ GeV}/c$  and  $0 < |\eta| < 1.2$ .

	$0 <  \eta  < 1.2$	$1.2 <  \eta  < 2.4$
$20 < p_T < 35$	$0.9997 \pm 0.0003$	$0.9992 \pm 0.0003$
$35 < p_T < 50$	$0.9994 \pm 0.0001$	$0.9995 \pm 0.0001$
$p_T > 50$	$0.9990 \pm 0.0006$	$0.9990 \pm 0.0003$

Table 4.17: Efficiency scale factors for track reconstruction efficiency.

## Muon identification efficiency

The muon selection efficiency is measured with global muon probes. An example fit is shown in Figure 4-23. The scale factors are listed in Table 4.18, and shows that the selection efficiency is fairly well described by the simulation.

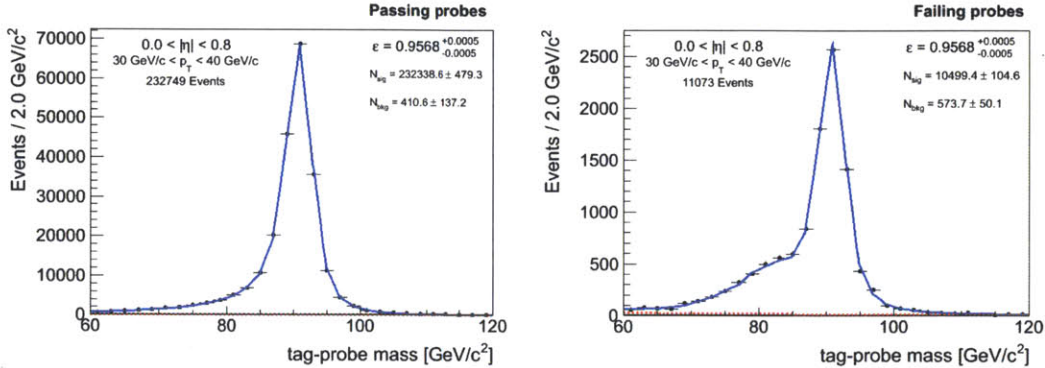


Figure 4-23: Fit for the selection efficiency for muons with  $30 \text{ GeV}/c < p_T < 40 \text{ GeV}/c$  and  $0 < |\eta| < 0.8$ .

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9900 \pm 0.0016$	$0.9833 \pm 0.0026$	$0.9861 \pm 0.0016$	$0.9839 \pm 0.0031$
$30 < p_T < 40$	$0.9963 \pm 0.0007$	$0.9886 \pm 0.0012$	$0.9902 \pm 0.0010$	$0.9842 \pm 0.0020$
$40 < p_T < 50$	$0.9970 \pm 0.0005$	$0.9893 \pm 0.0008$	$0.9916 \pm 0.0007$	$0.9823 \pm 0.0020$
$50 < p_T < 60$	$0.9952 \pm 0.0012$	$0.9879 \pm 0.0019$	$0.9911 \pm 0.0013$	$0.9801 \pm 0.0059$
$p_T > 60$	$0.9958 \pm 0.0017$	$0.9869 \pm 0.0029$	$0.9866 \pm 0.0019$	$0.9631 \pm 0.0123$

Table 4.18: Efficiency scale factors for muon selection efficiency.

### 4.4.3 Efficiency Corrections to the Acceptance

Separate treatments are used to account for the online trigger efficiency and the offline reconstruction and selection efficiency in data to correct the nominal acceptance. For the offline efficiency, the application is straightforward: scale every simulated signal event by the scale factors appropriate for each leg of the dilepton. For the trigger efficiency, the situation is more complicated since the dilepton triggers have asymmetric requirements on the two legs, and the single muon trigger is also used for the muon channel. The expected efficiency per  $Z$  event is calculated from considering

the compliment of all the possible ways to fail the dilepton trigger, and adding to that the efficiency recovered from the single lepton trigger. The resulting formula is,

$$\begin{aligned}
\varepsilon = & 1 - [1 - \varepsilon_{\text{lead}}(|\eta_1|, p_{T,1})][1 - \varepsilon_{\text{lead}}(|\eta_2|, p_{T,2})] \\
& - \varepsilon_{\text{lead}}(|\eta_1|, p_{T,1})[1 - \varepsilon_{\text{trail}}(|\eta_2|, p_{T,2})] - \varepsilon_{\text{lead}}(|\eta_2|, p_{T,2})[1 - \varepsilon_{\text{trail}}(|\eta_1|, p_{T,1})] \\
& + \varepsilon_S(|\eta_1|, p_{T,1})[1 - \varepsilon_{\text{trail}}(|\eta_2|, p_{T,2})] + \varepsilon_S(|\eta_2|, p_{T,2})[1 - \varepsilon_{\text{trail}}(|\eta_1|, p_{T,1})],
\end{aligned}
\tag{4.12}$$

where  $\eta_i$  and  $p_{T,i}$  are the kinematics of a lepton,  $\varepsilon_{\text{lead}}$  and  $\varepsilon_{\text{trail}}$  are respectively the efficiencies of the leading and trailing leg in the dilepton trigger, and  $\varepsilon_S$  is the efficiency of the single lepton trigger (zero for electrons). Equation 4.12 exploits two facts regarding the trigger requirements. First, if a lepton can pass the requirements of the leading leg of a dilepton trigger then it can also pass the trailing leg. Therefore, the possible ways to fail the dilepton trigger are: both leptons fail the leading leg requirements (second term on the right side of Equation 4.12), or one lepton satisfies the leading leg requirements but the other fails the trailing leg requirements (third and fourth terms). Secondly, the requirements on the single lepton trigger are more strict than the requirements on either leg of the dilepton trigger. Hence, efficiency is recovered by the single lepton trigger for events where one leg fails to be the trailing leg in the dilepton trigger, but the other lepton passes the single lepton requirements (fifth and sixth terms).

The baseline acceptance and the efficiency corrected acceptance (denoted  $A'$ ) are listed in Table 4.19.

Quantity	$ee$	$\mu\mu$
$A$	0.2407	0.4177
$A'$	0.2239	0.3949

Table 4.19: Acceptance before and after corrections with  $\rho$ .

#### 4.4.4 Systematic Uncertainties on the Efficiency

It is possible that the mass distributions in the pass and fail categories can be just as well described by other models for signal and background. Systematic uncertainties on the shape assumptions for the signal and background are estimated by considering reasonable alternative models to generate pseudo-data and evaluating the bias from fitting with the nominal shapes. To estimate the uncertainty on the signal model assumption, the background shape is fixed to the one obtained from the nominal fit and the alternative signal model is a Breit-Wigner distribution (fixed to the intrinsic  $Z$  boson mass and width) convoluted with the Crystal-Ball function. The template to generate pseudo-data is obtained by fitting the data with the alternative signal model and the fixed background shape, allowing the normalisation to float. The extracted efficiency becomes the “central value” in this exercise. Pseudo-data is generated by sampling the template and the efficiency is extracted by the nominal method. The average bias from 1000 such pseudo-experiments is taken as an estimate of the systematic uncertainty. For the uncertainty on the background model assumption, the just described procedure is repeated, but this time fixing the signal model while using a template obtain from simulated events for QCD and  $W$ +jets processes. Tables 4.20 and 4.21 lists the uncertainties for each efficiency step.

Quantity	Sig. Model	Bkg. Model
$\varepsilon_{\text{GSF}}$	0.64	0.70
$\varepsilon_{\text{Sel}}$	0.10	0.08

Table 4.20: Shape systematic uncertainties for electron efficiencies, in percent.

Quantity	Sig. Model	Bkg. Model
$\varepsilon_{\text{STA}}$	0.16	0.02
$\varepsilon_{\text{Trk}}$	0.07	0.29
$\varepsilon_{\text{Sel}}$	0.06	0.02

Table 4.21: Shape systematic uncertainties for muon efficiencies, in percent.

## 4.5 Luminosity

The basis of the luminosity calibration is that detector activity scales with instantaneous luminosity. Two techniques are employed: by considering the amount of energy deposited in the forward HCAL subdetector (HF), or by considering the number of pixel clusters [58,59]. The pixel cluster counting method has shown to be more robust than the method with HF, requiring fewer corrections and having smaller systematic uncertainties, and therefore provides the standard luminosity measurement.

The average number of observed pixel clusters is expected to be proportional to the instantaneous luminosity,

$$\frac{d\mathcal{L}}{dt} = \frac{\nu_{\text{orbit}}}{\sigma_{\text{pixel}}} \langle N_{\text{pixel}} \rangle, \quad (4.13)$$

where  $\nu_{\text{orbit}}$  is the frequency for a bunch to orbit the LHC (11 246 Hz),  $\langle N_{\text{pixel}} \rangle$  is the average number of pixel clusters, and  $\sigma_{\text{pixel}}$  denotes the “cross section” to form a pixel cluster. The quantity,  $\sigma_{\text{pixel}}$ , is given by,

$$\sigma_{\text{pixel}} = \frac{A_{\text{eff}}}{N_1 N_2 n_b}, \quad (4.14)$$

where  $N_i$  is the number of protons per bunch in a beam,  $n_b$  is the number of bunches in a beam, and  $A_{\text{eff}}$  is the effective transverse area of overlap of the two colliding beams. By measuring the electric current of the beam,  $N_i$  is determined, and the number of proton bunches in a beam is known, so this leaves only the determination of  $A_{\text{eff}}$  to complete the luminosity calibration. This is accomplished by systematically varying the positions of the colliding beams relative to each other to profile the shape of the beam in the two transverse directions, a technique known as the Van der Meer scan [60]. The beam profiles in both the  $x$  and  $y$  directions are well described by a sum of two Gaussians (with common central value). The effective widths,  $\sigma_x$  and  $\sigma_y$ , are about 50  $\mu\text{m}$ . The effective beam-beam overlap area is then computed as



$A_{\text{eff}} = 2\pi\sigma_x\sigma_y$ , and combined with Equations 4.13 and 4.14 gives,

$$\frac{d\mathcal{L}}{dt} = \frac{N_1 N_2 n_b \nu_{\text{orbit}}}{2\pi\sigma_x\sigma_y} \langle N_{\text{pixel}} \rangle. \quad (4.15)$$

The measured integrated luminosity for the data set is  $4.9 \text{ fb}^{-1}$  with a relative uncertainty of 2.2%. The uncertainty is dominated by the “afterglow” effect (1%) – essentially background noise from lingering or late arriving particles, and by the variation in  $\sigma_{\text{pixel}}$  measured from several scans (1.5%) [59].

## 4.6 Results

Now all the pieces are at hand to assemble the cross section according to Equation 4.1. The signal yields,  $N_S$ , are obtained by taking the observed yields of Table 4.6 and subtracting 0.45% for background. Since the expected background contamination is much smaller than the total systematic uncertainty, a conservative 100% uncertainty is assigned for simplicity without loss of precision: the background contribution is  $(0.45 \pm 0.45)\%$  of the total yield. The  $A \cdot \epsilon$  factor is given by the efficiency corrected acceptances listed in Table 4.19. The integrated luminosity,  $\mathcal{L}$ , is  $4.9 \text{ fb}^{-1}$  with an uncertainty of 2.2%. A summary of the systematic uncertainties are given in Table 4.22.

Source	$ee$	$\mu\mu$
Lepton Efficiency	1.10	1.80
Background	0.45	0.45
Acceptance	1.72	1.84
Luminosity	2.20	2.20

Table 4.22: Summary of systematic uncertainties in percent.

The cross sections for inclusive  $Z$  production decaying into electron or muon pairs are measured to be,

$$\begin{aligned} \sigma(pp \rightarrow ZX) \times \mathcal{B}(Z \rightarrow ee) &= 0.970 \pm 0.001_{(\text{stat.})} \pm 0.026_{(\text{syst.})} \pm 0.021_{(\text{lumi.})} \text{ nb} \\ \sigma(pp \rightarrow ZX) \times \mathcal{B}(Z \rightarrow \mu\mu) &= 0.973 \pm 0.001_{(\text{stat.})} \pm 0.021_{(\text{syst.})} \pm 0.021_{(\text{lumi.})} \text{ nb} \end{aligned}$$

The two results combine to give,

$$\sigma(pp \rightarrow ZX) \times \mathcal{B}(Z \rightarrow \ell\ell) = 0.972 \pm 0.001_{(\text{stat.})} \pm 0.017_{(\text{syst.})} \pm 0.021_{(\text{lumi.})} \text{ nb}$$

The prediction from theory at NNLO in perturbative QCD, computed using FEWZ with the MSTW 2008 parton distribution function, is  $0.97 \pm 0.03 \text{ nb}$  [39]. The cross section uncertainty includes the uncertainty on  $\alpha_S$  [61], the PDF model [62,63], and from missing higher order corrections beyond NNLO. The measurement is consistent with theory and dominated by systematic uncertainties..

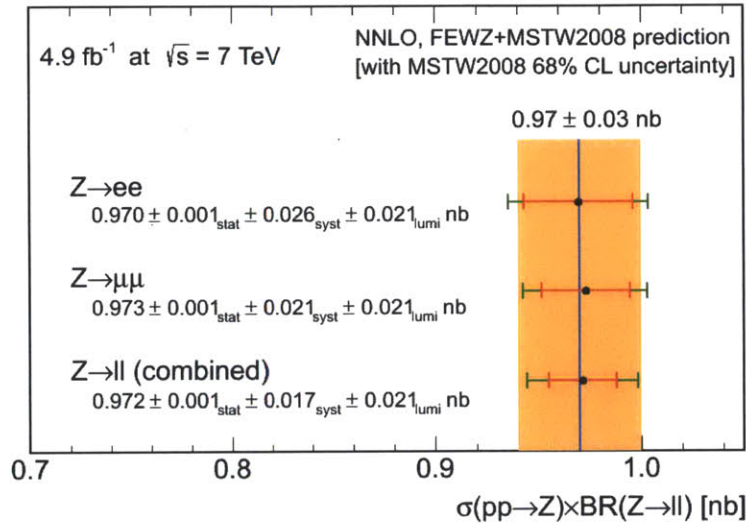


Figure 4-24: Comparison of the measured inclusive  $Z$  cross section with the theoretical prediction at NNLO in perturbative QCD. The yellow band represents the uncertainty on the prediction. The inner red error bars indicate the statistical and systematic uncertainties without the luminosity uncertainty, while the outer green error bars indicate the full uncertainty.



# Chapter 5

## Search for $H \rightarrow ZZ \rightarrow 2\ell 2\nu$

### 5.1 Introduction

The search for the Higgs boson in the  $ZZ \rightarrow 2\ell 2\nu$  channel examines events containing a lepton pair consistent with  $Z$  boson decay and significant amounts of missing transverse energy. This channel is particularly sensitive to a heavier Higgs boson, where high thresholds on  $\cancel{E}_T$  and dilepton  $p_T$  are imposed to considerably reduce backgrounds from known SM processes. The search is performed for Higgs mass hypotheses from  $250 \text{ GeV}/c^2$  to  $600 \text{ GeV}/c^2$ . The relevant backgrounds are weak diboson production ( $ZZ$ ,  $WZ$ ,  $WW$ ), top production, and single  $Z$  production.

While there are no neutrinos in  $Z \rightarrow \ell\ell$  events, the missing energy is induced from detector resolution effects. The total SM background is estimated from a combination of simulated events and control regions in data. Since the neutrinos escape detection, the mass of the candidate  $ZZ$  system cannot be reconstructed, but a useful transverse mass variable is defined and provides the primary discriminant of a Higgs signal from the background.

The model for the Higgs signal is the NLO description from POWHEG. The gluon fusion mode is modified by  $p_T$ -dependent  $k$ -factors computed from HqT [64–66], a tool that provides calculations up to NNLO and NNLL. Only the GF and VBF production modes are considered in this analysis, as the expected cross section for associated production modes  $VH$  and  $t\bar{t}H$  are negligible over the mass range of the

search.

## 5.2 Event Selection

The selection of  $H \rightarrow ZZ \rightarrow 2\ell 2\nu$  candidates begins with dielectron or dimuon events near the  $Z$  boson mass peak. The lepton trigger and offline selection requirements defined in Chapter 4 are applied and the efficiencies have been discussed in Section 4.4. Relative to the inclusive  $Z$  cross section measurement, a narrower mass window is chosen to reduce the tails of the dilepton mass distribution where top and  $WW$  contributions are more prominent. To be consistent with the  $H \rightarrow WW \rightarrow \ell\nu\ell'\nu'$  search in CMS [67], the mass window is taken to be within  $15 \text{ GeV}/c^2$  of the  $Z$  boson mass of  $m_Z = 91.1876 \text{ GeV}/c^2$  [68].

The following event vetoes are imposed to reduce background:

1. reject events containing a third electron or muon passing offline selection with  $p_T > 10 \text{ GeV}/c$ ,
2. reject events containing a well measured “soft” muon,  $p_T > 3 \text{ GeV}/c$ , that fails offline selection,
3. reject events containing a jet with  $p_T > 30 \text{ GeV}/c$  that contains more than one track with signed impact parameter significance larger than 2,
4. reject events with small azimuthal angle separation between a jet and  $\vec{E}_T$ ,  $\Delta\phi(\hat{p}_T^{\text{jet}}, \vec{E}_T) < 0.5$ , for any jet with  $p_T > 30 \text{ GeV}/c$ , and if there are no jets above  $30 \text{ GeV}/c$  then consider jets above  $15 \text{ GeV}/c$ .

The first criterion reduces  $WZ$  contributions. The second and third requirements serve to reduce the top background by exploiting properties of  $B$  hadrons, as top quarks almost always decay to bottom quarks; the second criterion targets semi-leptonic decays, while the third criterion takes advantage of the long lifetime. The fourth criterion reduces backgrounds from single  $Z$  production. In such events, the mismeasurement of the hadronic recoil is the cause for the energy imbalance and a jet

is expected to be closely aligned with the direction of  $\vec{\cancel{E}}_T$ . The transverse jet energy threshold of 30 GeV/c is chosen to avoid selecting random jets from the underlying event or pile-up interactions. However, for single  $Z$  events with considerable  $\cancel{E}_T$  but no jets above 30 GeV/c, this typically happens because the recoiling jet has been mismeasured to have momenta below the threshold, so then jets above 15 GeV/c are considered. From this point on, the single  $Z$  background will be referred to as the  $Z$ +jets background, to allude to the notion that  $\cancel{E}_T$  arises from jet energy mismeasurement.

For signal events, the average missing energy is correlated with the Higgs boson mass,  $m_H$ , so it makes sense to impose a  $\cancel{E}_T$  requirement that varies with the  $m_H$  hypothesis (see Table 5.2). A “transverse mass”,  $M_T$ , is defined from three observed quantities: the dilepton transverse momentum,  $\vec{p}_{T,\ell\ell}$ , the dilepton mass,  $m_{\ell\ell}$ , and the missing energy  $\vec{\cancel{E}}_T$ , according to the formula,

$$M_T^2 = \left( \sqrt{p_{T,\ell\ell}^2 + m_{\ell\ell}^2} + \sqrt{(\cancel{E}_T)^2 + m_{\ell\ell}^2} \right)^2 - \left( \vec{p}_{T,\ell\ell} + \vec{\cancel{E}}_T \right)^2. \quad (5.1)$$

The transverse mass provides a powerful discriminant because the distribution is peaked for the Higgs boson but is a falling spectrum for the SM background (see Figure 5-1). Transverse mass windows are imposed depending on the Higgs mass hypothesis and are listed in Table 5.2.

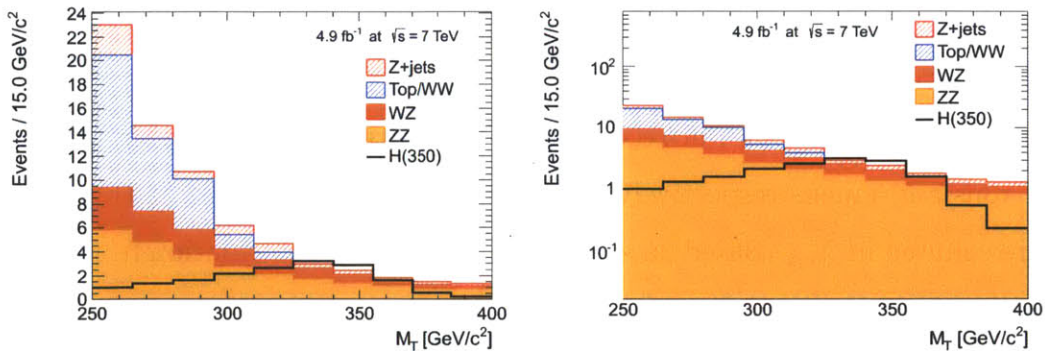


Figure 5-1: The expected transverse mass distribution for signal ( $m_H = 350 \text{ GeV}/c^2$ ) and background.

Mass	250	300	350	400	500	600
$\cancel{E}_T >$	70	80	80	80	80	80
$M_T$ region	[230, 300]	[250, 350]	[250, 400]	[250, 450]	[250, 600]	[250, 750]

Table 5.1: The  $\cancel{E}_T$  and  $M_T$  selections for different  $m_H$  hypotheses.

## 5.3 Background Estimation

A combination of detailed simulation studies and data-driven methods are used to estimate background contributions after event selection. Estimations for  $ZZ$  and  $WZ$  are taken from simulated events generated with MADGRAPH. The top,  $WW$ , and  $Z$ +jets backgrounds are estimated with control samples in data. The data-driven techniques are explained in the following sections.

### 5.3.1 Top and $WW$ backgrounds

Top and  $WW$  production are processes where the final states involving  $\mu^+\mu^-$ ,  $e^\pm\mu^\mp$  and  $e^+e^-$  occur at equal rates. The background is estimated by counting different-flavour ( $e\mu$ ) dilepton events passing analysis selection, and then correcting for efficiency differences between selecting electrons and muons. For example, the predicted dimuon yield is determined from the different-flavour yield as follows,

$$N_{\mu\mu} = \frac{1}{2} N_{e\mu} \cdot \frac{\varepsilon_\mu}{\varepsilon_e}, \quad (5.2)$$

where  $\varepsilon_e$  and  $\varepsilon_\mu$  are the overall reconstruction, selection, and triggering efficiencies for electrons and muons respectively. The factor of  $\frac{1}{2}$  arises because both  $e^+\mu^-$  and  $e^-\mu^+$  are counted in  $N_{e\mu}$ . Based on simulation studies, the contribution from  $Z \rightarrow \tau\tau$  is negligible for  $M_T > 150 \text{ GeV}/c^2$  – the analysis requires at least  $230 \text{ GeV}/c^2$  (see Table 5.2), but nevertheless would be covered by this method of counting different-flavour events.

A validation of the different-flavour counting method is performed by considering a control region dominated by the top and  $WW$  background. This control region

is defined by dileptons events passing all analysis cuts except that the mass lies outside the  $Z$  mass window (with a minimum of mass of  $12 \text{ GeV}/c^2$ ), and having  $\cancel{E}_T > 70 \text{ GeV}/c$  and  $M_T > 150 \text{ GeV}/c^2$ . Table 5.2 shows that observed and expected event counts agree. Distributions of the dilepton mass, the dilepton  $p_T$ , the missing energy, and transverse mass are plotted in Figures 5-2 to 5-5.

Process	$ee$	$\mu\mu$
$ZZ$	$2.6 \pm 0.1$	$4.3 \pm 0.1$
$WZ$	$4.2 \pm 0.1$	$5.3 \pm 0.1$
Top/ $WW$	$177.6 \pm 8.7$	$264.7 \pm 13.0$
Total Expected	$184.3 \pm 8.7$	$274.3 \pm 13.0$
Observed	187	272

Table 5.2: Observed and expected yields in the control region dominated by top and  $WW$  backgrounds. The  $ZZ$  and  $WZ$  yields are estimated in simulation, while the top and  $WW$  backgrounds are estimated using the method of counting different-flavour events.

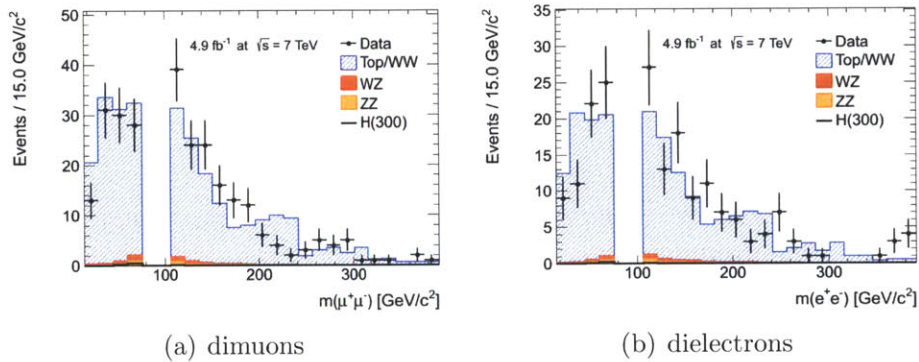


Figure 5-2: Dilepton mass distributions in the control region dominated by top and  $WW$  backgrounds.

### 5.3.2 Single $Z$ +jets background

Single  $Z$  production where the  $Z$  decays into electrons or muons does not involve neutrinos, therefore its contribution is significantly reduced by the tight  $\cancel{E}_T$  selection. The residual contribution in the signal region is mainly due to events with fake missing energy from mismeasurement of the hadronic recoil. The contribution of events with



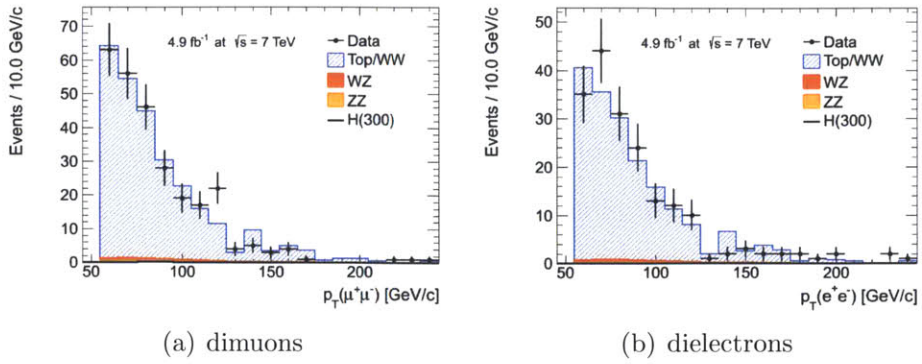


Figure 5-3: Dilepton  $p_T$  distributions in the control region dominated by top and  $WW$  backgrounds.

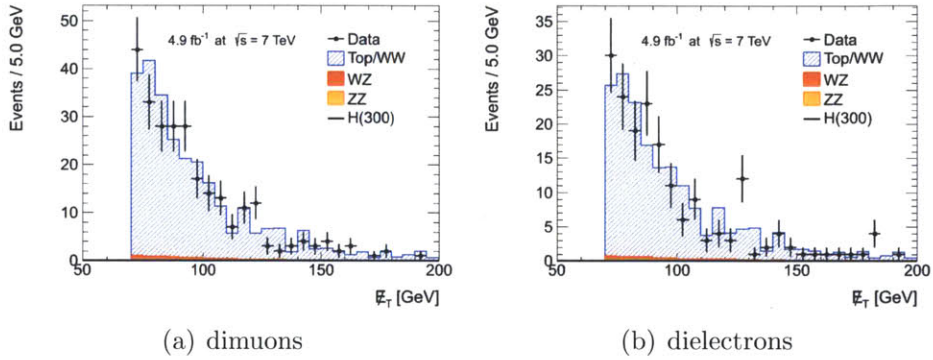


Figure 5-4: Missing transverse energy distributions in the control region dominated by top and  $WW$  backgrounds.

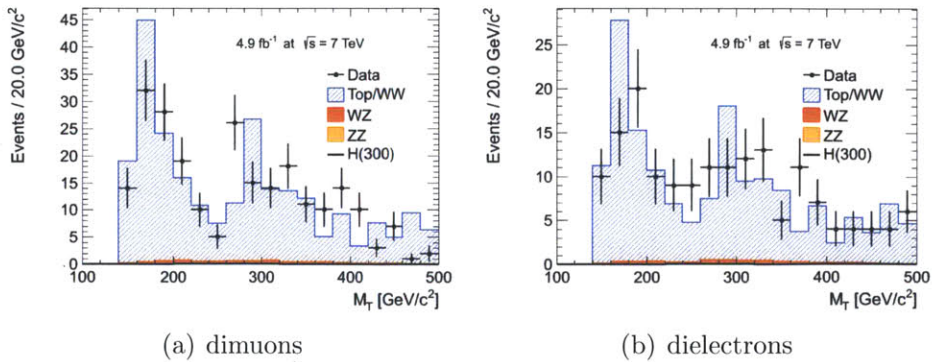


Figure 5-5: Transverse mass distributions in the control region dominated by top and  $WW$  backgrounds.

fake  $\cancel{E}_T$  from lepton momentum mismeasurement is expected to be small because selected dileptons fall within a fairly tight window around the  $Z$  boson mass.

An accurate description of the  $Z$ +jets background requires a good understanding of the  $\cancel{E}_T$  distribution tail, which is sensitive to many factors such as the jet energy corrections, the number of pileup interactions, and energy from out of time pileup contributions. These effects and the corresponding detector response are difficult to model precisely in the simulation. Instead, a template using the analogous  $\gamma$ +jets events in data provides a better model for the missing energy. The mechanism for generating missing energy in  $\gamma$ +jets events is expected to be similar to  $Z$ +jets events. Figure 5-6 shows the deficiency to reproduce the  $\cancel{E}_T$  distribution with simulation compared with the template derived from the  $\gamma$ +jets method, which is discussed in the following subsection.

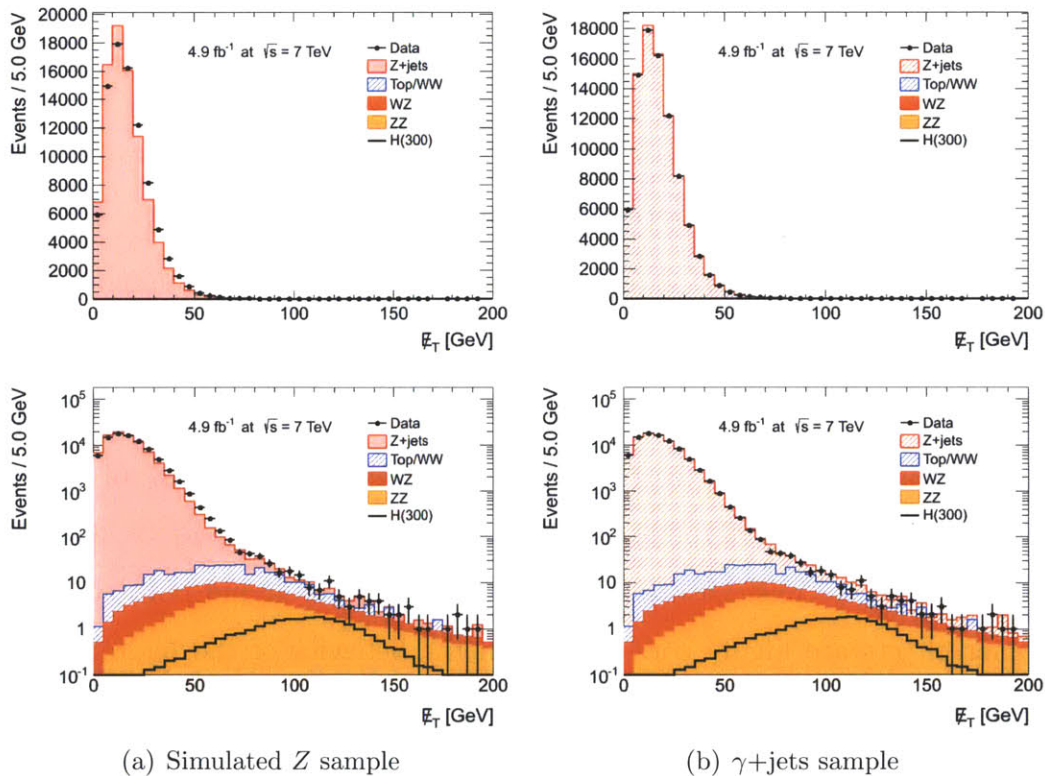


Figure 5-6: Comparison of  $\cancel{E}_T$  modelling using simulated  $Z$  events and using the  $\gamma$ +jets method. The plots on the top row are in linear scale for the  $y$ -axis, while the plots on the bottom row are in logarithmic scale.

## $\gamma$ +Jets Method

The photon data sample is selected from events obtained via a photon trigger by requiring one and only one photon passing tight identification ( $H/E$  and shower shape variables) and isolation requirements, summarized in Table 5.3, with  $p_T > 55 \text{ GeV}/c$  to match the dilepton  $p_T$  cut in the analysis. Only photons in the ECAL barrel are considered to minimize effects on the  $\cancel{E}_T$  distribution from photon energy resolution, which is worse in the endcaps. An electron veto is imposed on the photon, requiring that the photon is not matched to a pixel seed. To reduce contamination from  $W+\gamma$  or  $Z+\gamma$  events, events with one or more leptons passing selection requirements are rejected.

Cut	Requirement
Photon is not pixel seeded	
$\eta$	$ \eta  < 1.4442$
H/E	$H/E < 0.05$
$\sigma_{i\phi i\phi}$	$\sigma_{i\phi i\phi} > 0.001$
$\sigma_{i\eta i\eta}$	$0.001 < \sigma_{i\eta i\eta} < 0.013$
EcalIso	$\text{EcalIso} < 4.2 + 0.0060 \times p_T$
HcalIso	$\text{HcalIso} < 2.2 + 0.0025 \times p_T$
TrkIso	$\text{TrkIso} < 2.0 + 0.0010 \times p_T$

Table 5.3: Selection criteria for photons in the  $\gamma$ +jets sample. The minimum thresholds on  $\sigma_{i\phi i\phi}$  and  $\sigma_{i\eta i\eta}$  are imposed to avoid rare, pathological events where a single ECAL crystal registers a high energy “spike”.

The modelling of the  $\cancel{E}_T$  distribution is parametrised in the  $p_T$  of the photon/dilepton system and the number of counted jets (jets with  $p_T > 30 \text{ GeV}/c$ ). The fake  $\cancel{E}_T$  is due to mismeasurement of the hadronic recoil and both the  $p_T$  and the number of jets are highly correlated with the behavior of the hadronic recoil. Additionally, the photon triggers are prescaled so the photon events are sampling the pileup conditions differently than the dilepton events which pass selection. To address this, parametrisation in the primary vertex multiplicity is done. Reweighting factors are computed in bins of the number of counted jets and the dilepton  $p_T$ , and in the number of primary vertices, by dividing the corresponding distributions in the  $\gamma$ +jets data sample and the dilepton data sample:

$$w(p_T, n_{\text{jet}}) = \frac{N_{\gamma+\text{jets}}(p_T, n_{\text{jet}})}{N_{Z+\text{jets}}(p_T, n_{\text{jet}})} \times \frac{N_{\gamma+\text{jets}}(\text{npv})}{N_{Z+\text{jets}}(\text{npv})}. \quad (5.3)$$

In order for these reweighting factors to be insensitive to the presence of signal or backgrounds with real  $\cancel{E}_T$ , the  $\gamma$ +jets and  $Z$ +jets samples are restricted to  $\cancel{E}_T < 50 \text{ GeV}/c$ . To model the  $\cancel{E}_T$  distribution for estimating the  $Z$ +jets background, each  $\gamma$ +jets event selected is then associated with the corresponding weight factor  $w(p_T, n_{\text{jet}}, \text{npv})$ . To model the transverse mass distribution, a random value is drawn from the  $Z$  mass distribution in data with  $\cancel{E}_T < 50 \text{ GeV}/c$  to build the  $M_T$  observable by following Equation 5.1.

To define a control region enriched with  $Z$ +jets events in order to check the validity of the templates derived from the  $\gamma$ +jets method, the minimum  $\cancel{E}_T$  and  $M_T$  cuts are relaxed to  $50 \text{ GeV}/c$  and  $150 \text{ GeV}/c^2$  respectively, and the  $\Delta\phi(\hat{p}_T^{\text{jet}}, \vec{\cancel{E}}_T)$  cut is removed completely, while all other selection requirements remain in force. Table 5.4 shows that there agreement between observed and expected yields. Distributions of the dilepton mass, the dilepton  $p_T$ , the missing energy, and transverse mass are plotted in Figures 5-7 to 5-11.

Process	$ee$	$\mu\mu$
$ZZ$	$43.8 \pm 0.2$	$69.6 \pm 0.3$
$WZ$	$36.7 \pm 0.4$	$56.1 \pm 0.5$
Top/ $WW$	$92.0 \pm 6.2$	$137.5 \pm 9.3$
$Z$ +jets	$625.5 \pm 12.1$	$966.4 \pm 19.7$
Total Expected	$798.0 \pm 13.6$	$1229.6 \pm 21.8$
Observed	784	1183

Table 5.4: Observed and expected yields in the control region dominated by  $Z$ +jets backgrounds. The  $ZZ$  and  $WZ$  yields are estimated in simulation, top and  $WW$  background is estimated using the method of counting different flavour events, and the  $Z$ +jets yields are estimated with the  $\gamma$ +jets method.

### Contamination from processes with real $\cancel{E}_T$

The photon sample used to estimate the  $Z$ +jets background has contributions

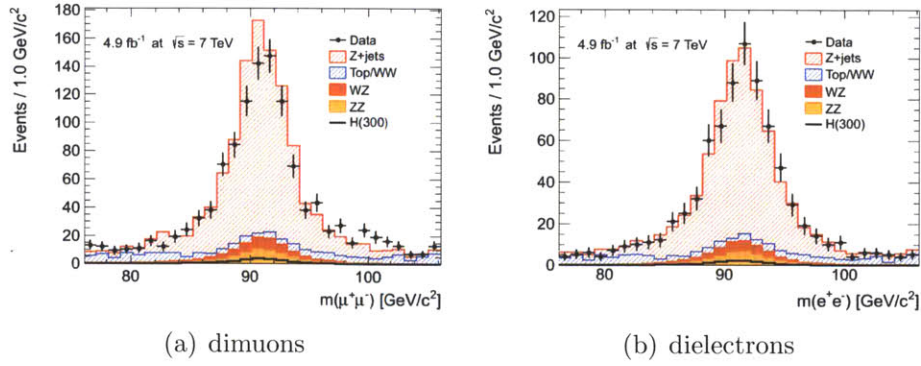


Figure 5-7: Dilepton mass distribution in the control region dominated by  $Z$ +jets background.

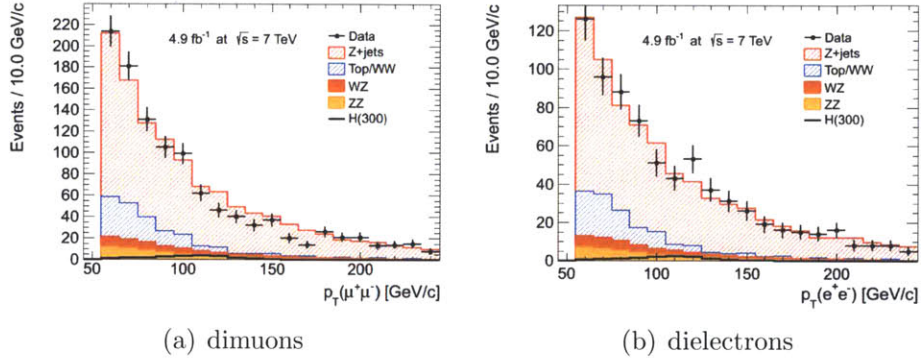


Figure 5-8: Dilepton  $p_T$  distribution in the control region dominated by  $Z$ +jets background.

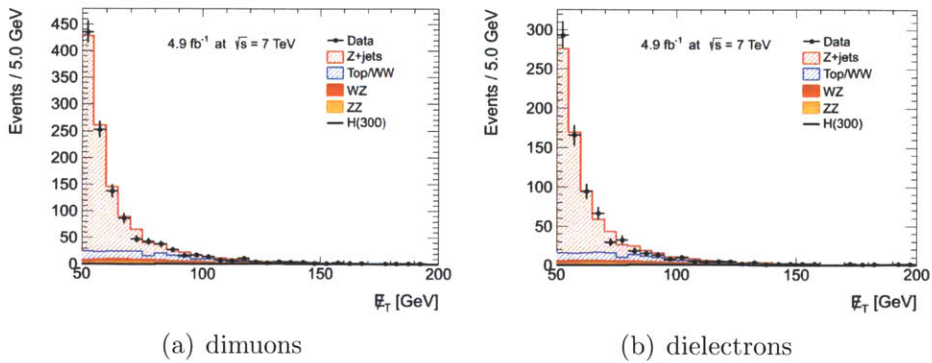


Figure 5-9: Missing transverse energy distribution in the control region dominated by  $Z$ +jets background.

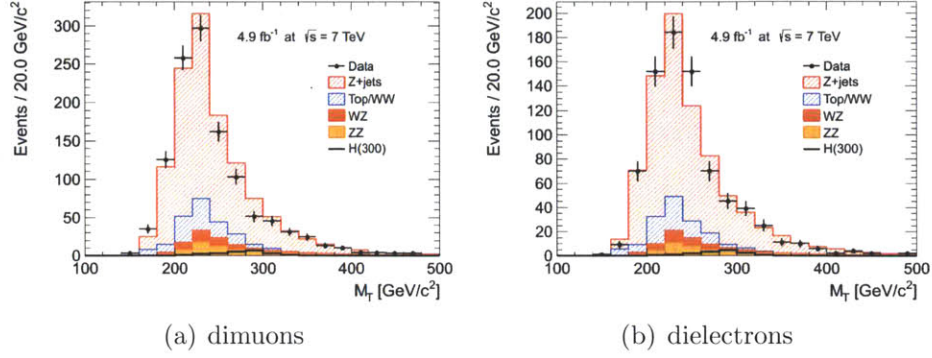


Figure 5-10: Transverse mass distribution in the control region dominated by  $Z$ +jets background.

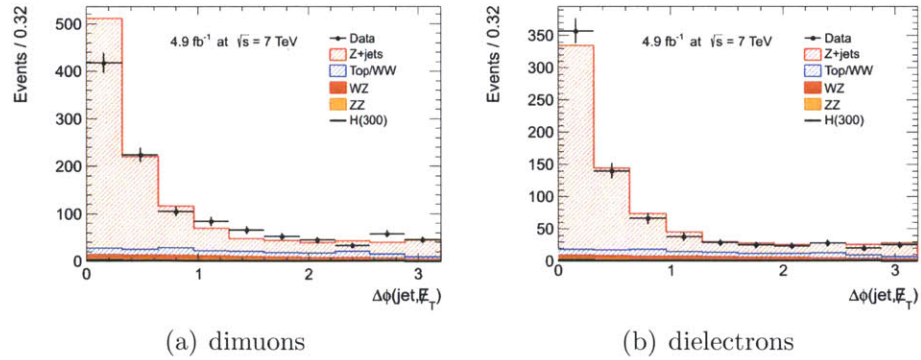


Figure 5-11:  $\Delta\phi(\hat{p}_T^{\text{jet}}, \vec{E}_T)$  distribution in the control region dominated by  $Z$ +jets background.



from physics processes other than  $\gamma$ +jets. Such processes include  $Z + \gamma$  events where the  $Z$  boson decays into neutrinos,  $W + \gamma$  events where the lepton from the  $W$  decay fails reconstruction or selection requirements, and  $W$ +jets events where the electron is misidentified as a photon. The missing energy from such events do not come from jet mismeasurement and as a consequence will bias the  $Z$ +jets estimate. Figure 5-12 illustrates where different processes are expected to contribute to the missing energy distribution predicted from photon events. The contamination is relatively more severe further down the tail of the distribution.

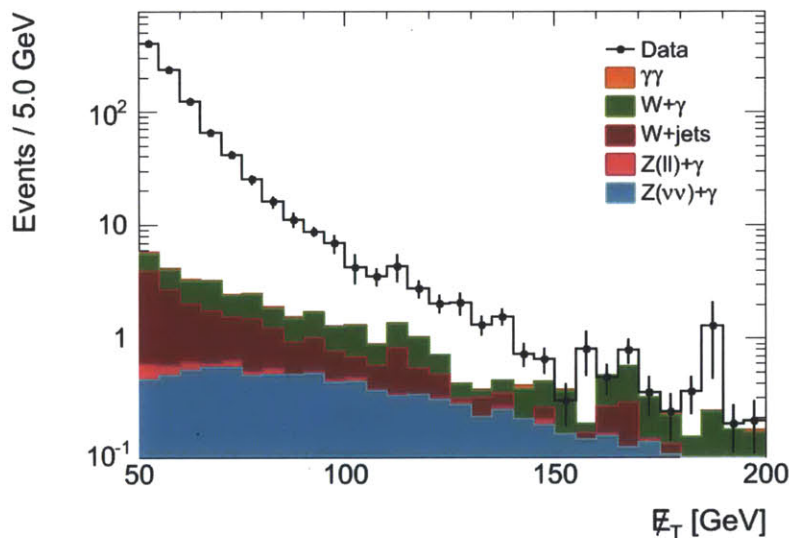


Figure 5-12: Predicted missing transverse energy distributions from the photon control region. The black histogram is the photon data. The contributions from various physics processes derived from simulations are stacked together and overlaid on top of the data histogram.

Consequently, using the photon sample in data to estimate the  $Z$ +jets background results in an over-prediction. Studies with simulation indicate that, after full selection, roughly half of the  $Z$ +jets estimate arise from real  $\gamma$ +jets events, while the rest comes mainly from  $Z(\nu\nu) + \gamma$ ,  $W + \gamma$ , and  $W$ +jets. However, the statistical uncertainties are large due to the small number of photon events in data, about 30% on the contamination subtracted estimate. The treatment for the  $Z$ +jets estimate is as follows: take the estimate to be half of the prediction provided by the photon

sample (without any explicit contamination subtraction), and assign a 100% relative uncertainty on this estimate. Effectively, this expresses the notion that the  $\gamma$ +jets prediction provides an upper bound on the  $Z$ +jets background estimate.

## 5.4 Systematic Uncertainties

The important sources of systematic uncertainties include:

- integrated luminosity,
- lepton efficiency and energy scale,
- theoretical uncertainties on expected yields derived from simulation,
- uncertainties on yields estimated from control regions in data,
- and uncertainties on the shape of the  $M_T$  distribution for each process.

The systematic uncertainties from leptons and luminosity have already been discussed in Chapter 4. The uncertainty assigned to the  $Z$ +jets estimate was discussed in Section 5.3. The other sources will be addressed here.

### 5.4.1 Theoretical Uncertainties on Yields from Simulation

Theoretical uncertainties need to be considered for the Higgs signal and the  $ZZ$  and  $WZ$  backgrounds are estimated with simulated events, . The dominant systematic uncertainties are on the parton distribution function and missing higher order corrections in perturbative QCD. For Higgs signal sample, there is also an additional uncertainty on the Higgs boson width, as this width is not accounted for in the simulation.

#### Uncertainty on PDF and Missing Higher Order QCD Corrections

Following the prescription recommended by the PDF4LHC working group [69], the yields are estimated using the CT10, MSTW2008, and NNPDF [70] sets, and an



uncertainty “envelope” is computed that covers the quantities computed with those three PDFs. The magnitude of the uncertainties range from 5% to 10%. To estimate the uncertainty from missing higher order corrections, yields at different renormalisation and factorisation scales are computed and compared, using the procedure already described in Section 4.3.2.

### Uncertainty Associated with Higgs Mass Width

For generating and simulating signal events, on-shell Higgs boson production is assumed and the subsequent decay is implemented with an ad-hoc Breit-Wigner. Off-shell contributions are neglected. A systematic uncertainty is assigned on the modelling of the Higgs width. The formula for this uncertainty is [71],

$$\frac{\Delta\sigma}{\sigma} = (150\%) \times \left( \frac{m_H}{1 \text{ TeV}/c^2} \right)^3. \quad (5.4)$$

## 5.4.2 Summary of Systematic Uncertainties on Normalisation

A summary of the systematic uncertainties on signal and background normalisation is given in Table 5.5. These are all the systematic uncertainties except for  $M_T$  shape uncertainties.

### 5.4.3 $M_T$ Shape Uncertainties

The uncertainty on the shape of the  $M_T$  distribution is accounted for by considering two bounding shapes on the central distribution. The bounding shapes need not lie completely above or completely below the the central shape, but simply provide uncertainties for each bin of the  $M_T$  histogram. For some sources of uncertainty, the bounding shapes do act as “upper” and “lower” bounds, as in the case of propagating lepton efficiency uncertainties where the bin-by-bin normalisation is correlated with the shape variation. The shape uncertainty propagated from the lepton efficiency uncertainties are typically bounded already by the efficiency uncertainty on the overall

Source	Uncertainty [%]
Luminosity	2.2
Muon efficiency	1.1
Muon momentum scale	0.5
Electron efficiency	1.8
Electron energy scale	1.0
Higgs cross section (GF)	12.8
Higgs cross section (VBF)	5.1
Higgs cross section due to width	2-32
$WZ$ cross section	6.2
$qq \rightarrow ZZ$	8.2
$gg \rightarrow ZZ$ cross section	20.0
Top and $WW$ background	10-13
$Z$ +jets background	100

Table 5.5: Summary of systematic uncertainties on normalisation.

normalisation, hence this source of shape uncertainty is negligible. A more relevant example is the bin-by-bin statistical uncertainty on the top and  $WW$  background shapes, which suffer from low event counts, and the bounding shapes are generated by varying together all bins of the central shape by  $\pm 1\sigma$ . For other sources, an  $M_T$  distribution is generated using an alternative method (such as a different event generator), and the bin-by-bin differences are “mirrored” across the central distribution to obtain the bounding shapes.

The shape uncertainties for each process are accounted for as follows:

- Higgs signal
  - No shape uncertainty is considered, as the shape variations from PDF uncertainties and from varying QCD scales are much smaller than the uncertainties on the normalisation.
- $ZZ$  and  $WZ$ 
  - Bin-by-bin statistical uncertainties and bounding shapes derived with PYTHIA as an alternative generator are considered.
- Top and  $WW$

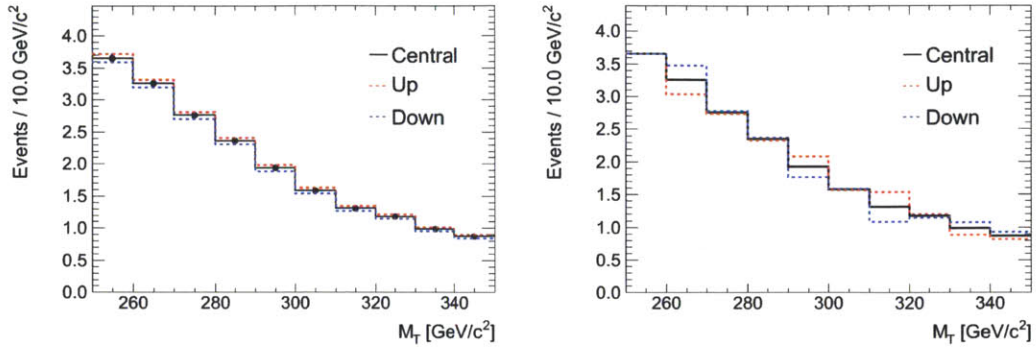


Figure 5-13: Shape uncertainties on the  $ZZ$  background for the  $m_H = 300 \text{ GeV}/c^2$  hypothesis. Left plot is the bin-by-bin statistical uncertainty and the right plot is the shape variation.

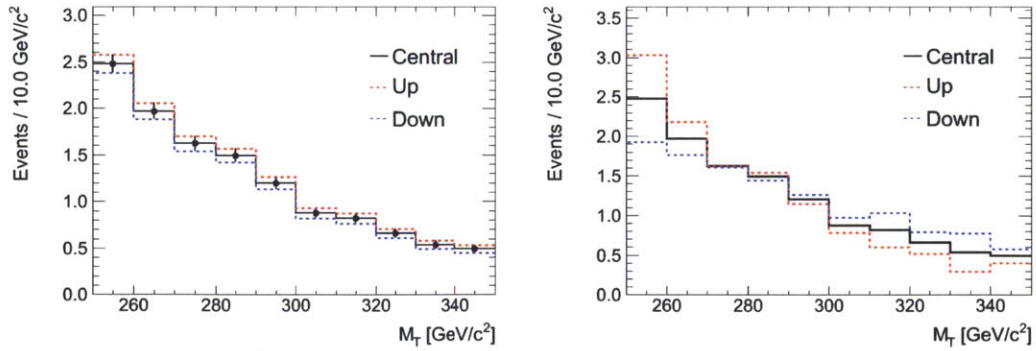


Figure 5-14: Shape uncertainties on the  $WZ$  background for the  $m_H = 300 \text{ GeV}/c^2$  hypothesis. Left plot is the bin-by-bin statistical uncertainty and the right plot is the shape variation.

- Shape uncertainty is dominated by the bin-by-bin statistical uncertainties of the central histogram.

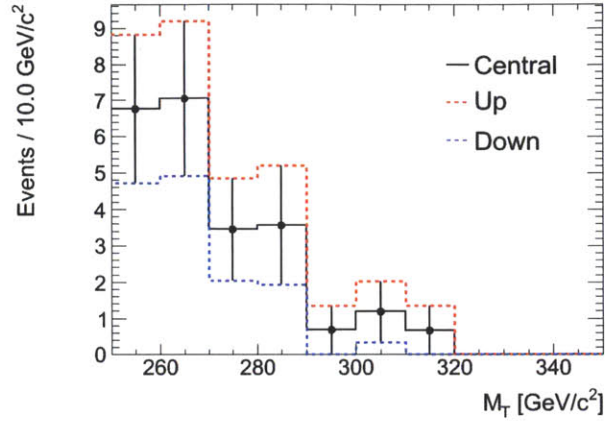


Figure 5-15: Shape uncertainty on the top and  $WW$  background for the  $m_H = 300 \text{ GeV}/c^2$  hypothesis.

- $Z$ +jets
  - No shape uncertainty is considered, as the large uncertainty on the normalisation is expected to dominate any reasonable shape variations.

## 5.5 Results

### 5.5.1 Yields

The expected and observed yields are tabulated in Table 5.6. In terms of event counts, there is no observed excess above the expected SM background. The observed distributions for  $M_T$  are also consistent with background expectations and are shown in Figures 5-16 to 5-21 for the each  $m_H$  hypothesis considered.

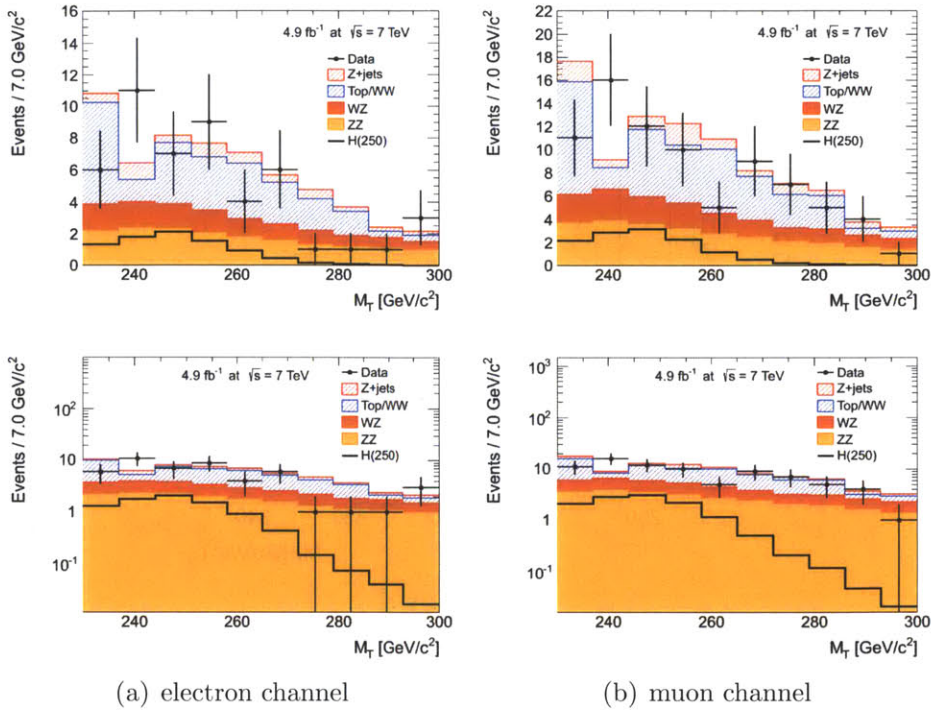


Figure 5-16:  $M_T$  distribution after selection for the  $m_H = 250 \text{ GeV}/c^2$  hypothesis.

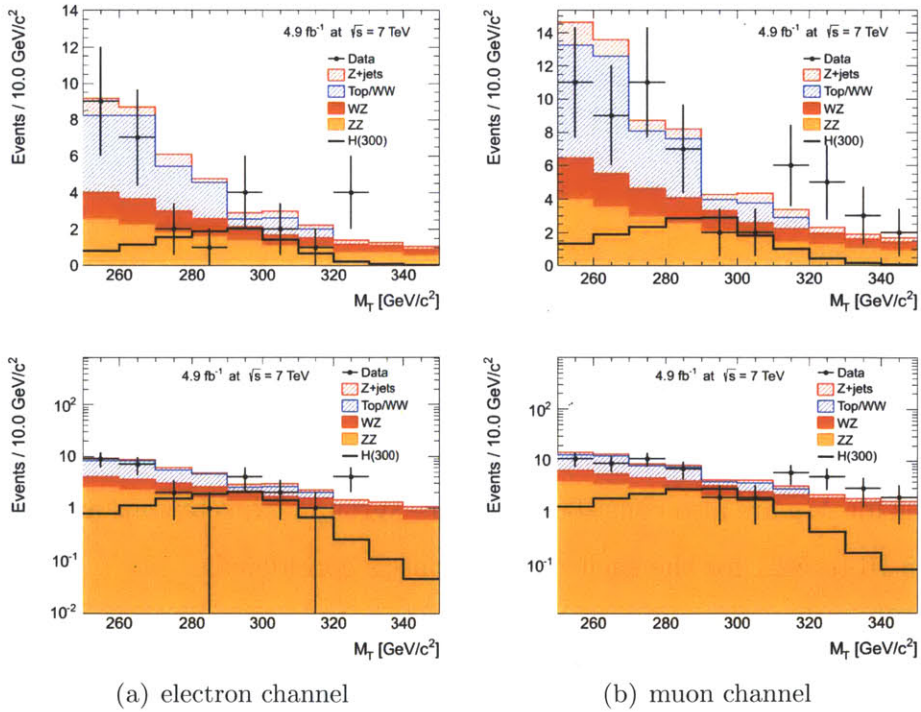


Figure 5-17:  $M_T$  distribution after selection for the  $m_H = 300 \text{ GeV}/c^2$  hypothesis.

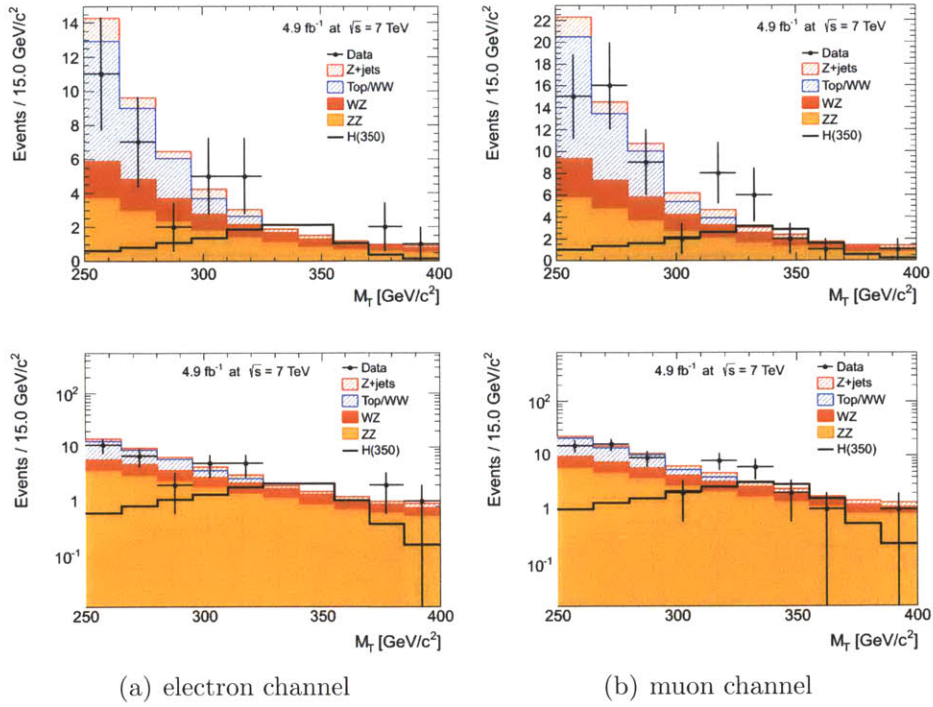


Figure 5-18:  $M_T$  distribution after selection for the  $m_H = 350 \text{ GeV}/c^2$  hypothesis.

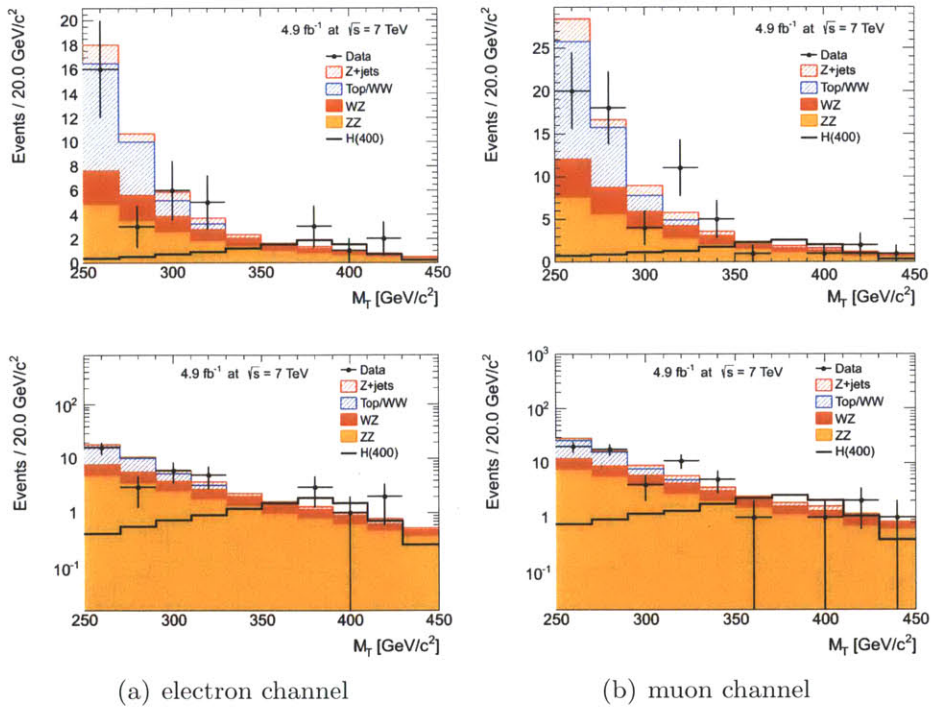


Figure 5-19:  $M_T$  distribution after selection for the  $m_H = 400 \text{ GeV}/c^2$  hypothesis.



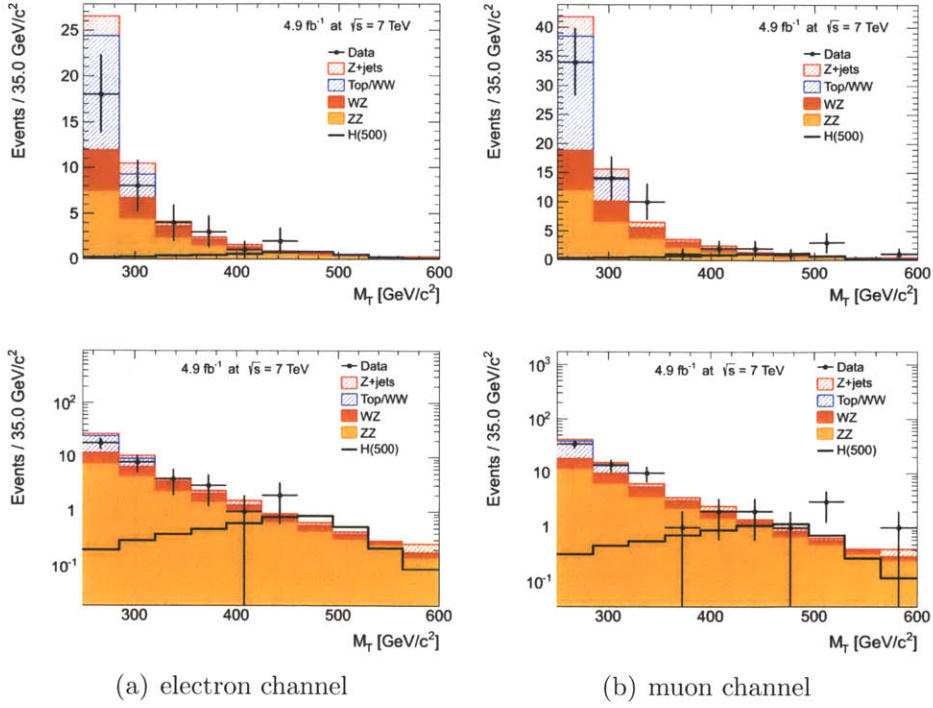


Figure 5-20:  $M_T$  distribution after selection for the  $m_H = 500 \text{ GeV}/c^2$  hypothesis.

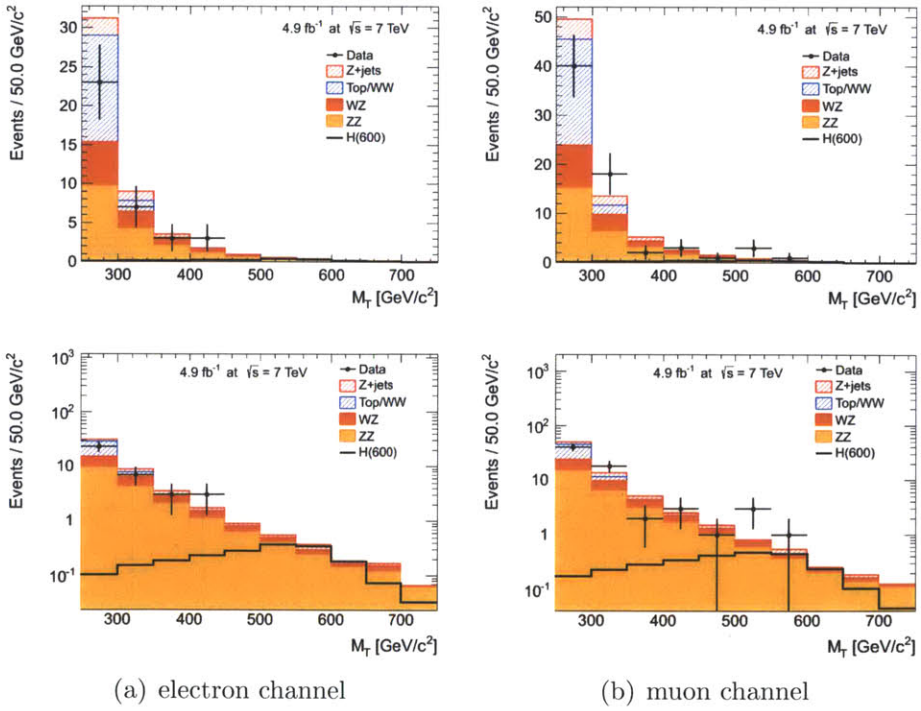


Figure 5-21:  $M_T$  distribution after selection for the  $m_H = 600 \text{ GeV}/c^2$  hypothesis.

$m_H$ [GeV/ $c^2$ ]	Signal	$ZZ$	$WZ$	Top/ $WW$	$Z$ +jets	$\sum$ Bkgd	Data
$ee$							
250	$8.4 \pm 0.1$	$16.8 \pm 0.2$	$11.1 \pm 0.2$	$25.4 \pm 3.2$	$5.5 \pm 5.5$	$58.8 \pm 6.4$	49
300	$9.9 \pm 0.1$	$13.9 \pm 0.1$	$7.9 \pm 0.2$	$15.0 \pm 2.5$	$3.8 \pm 3.8$	$40.6 \pm 4.5$	30
350	$11.7 \pm 0.1$	$16.0 \pm 0.1$	$8.8 \pm 0.2$	$15.0 \pm 2.5$	$4.1 \pm 4.1$	$43.9 \pm 4.8$	33
400	$9.8 \pm 0.1$	$17.2 \pm 0.2$	$9.2 \pm 0.2$	$15.0 \pm 2.5$	$4.3 \pm 4.3$	$45.7 \pm 5.0$	36
500	$4.5 \pm 0.1$	$18.4 \pm 0.2$	$9.5 \pm 0.2$	$15.0 \pm 2.5$	$4.6 \pm 4.6$	$47.6 \pm 5.3$	36
600	$2.0 \pm 0.1$	$18.8 \pm 0.2$	$9.6 \pm 0.2$	$15.0 \pm 2.5$	$4.2 \pm 4.2$	$47.6 \pm 4.9$	36
$\mu\mu$							
250	$12.5 \pm 0.2$	$26.7 \pm 0.2$	$17.0 \pm 0.2$	$13.7 \pm 4.9$	$9.1 \pm 9.1$	$91.6 \pm 10.4$	80
300	$14.8 \pm 0.2$	$12.6 \pm 0.2$	$12.2 \pm 0.2$	$23.3 \pm 3.9$	$5.8 \pm 5.8$	$62.9 \pm 7.0$	58
350	$17.2 \pm 0.2$	$24.8 \pm 0.2$	$13.4 \pm 0.2$	$23.3 \pm 3.9$	$6.3 \pm 6.3$	$67.9 \pm 7.4$	60
400	$14.3 \pm 0.1$	$26.6 \pm 0.2$	$14.0 \pm 0.2$	$23.3 \pm 3.9$	$6.9 \pm 6.9$	$70.8 \pm 7.9$	63
500	$6.4 \pm 0.1$	$28.6 \pm 0.2$	$14.5 \pm 0.2$	$23.3 \pm 3.9$	$7.2 \pm 7.2$	$73.6 \pm 8.2$	68
600	$2.8 \pm 0.1$	$29.0 \pm 0.2$	$14.6 \pm 0.2$	$23.3 \pm 3.9$	$7.3 \pm 7.3$	$74.3 \pm 8.3$	68

Table 5.6: Expected and observed yields for different  $m_H$  hypotheses.

## 5.5.2 Statistical Analysis

In the absence of a signal observation, the result of the search quantified by an upper limit on the rate of Standard Model Higgs production. For this purpose, a modified frequentist approach known as  $CL_s$  [72, 73] has been adopted by both the CMS and ATLAS collaborations [74]. To simplify the presentation of results, the limits are expressed in terms of the cross section divided by the predicted SM value, denoted by  $\mu$ . An upper limit below  $\mu = 1$  is interpreted as an exclusion of the SM Higgs. For each  $m_H$  hypothesis, a likelihood function for the observed data, given  $\mu$  and expectations for signal and background, are defined as the product of Poisson probabilities from each bin of the distribution in the discriminating variable (i.e.  $M_T$ ),

$$\mathcal{P}(\text{data}|\mu) = \prod_i \frac{(\mu \cdot s_i + b_i)^{n_i}}{n_i!} e^{-\mu \cdot s_i - b_i}, \quad (5.5)$$

where  $n_i$ ,  $s_i$ , and  $b_i$  denote respectively the observed yield, expected signal yield, and expected background yield in the  $i^{\text{th}}$  bin of the  $M_T$  histogram. To account for the uncertainties, each source of uncertainty is assigned a nuisance parameter. Given the measurements and uncertainties associated with each of these sources (such as lepton efficiency or cross section normalisation), a posterior probability function over the nuisance parameters are constructed. Applying Bayes' theorem and choosing



a flat prior, the likelihood function,  $p(\tilde{\theta}|\theta)$ , takes the same form as the posterior probability function, where  $\theta$  denotes the collection of nuisance parameters and  $\tilde{\theta}$  represents the information on the nuisances from supplementary measurements (such as the Tag-and-Probe procedure) or calculations (such as theoretical predictions on cross sections).

The nuisance parameters refer to quantities that affect normalisation and the expected signal and background yields are functions of the nuisances,  $s_i(\theta)$  and  $b_i(\theta)$ . The full likelihood function is,

$$\mathcal{L}(\text{data}|\mu, \theta) = \mathcal{P}(\text{data}|\mu, \theta) \cdot p(\tilde{\theta}|\theta), \quad (5.6)$$

For testing the Higgs signal hypothesis, a test statistic is defined by,

$$q_\mu = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}, \quad 0 \leq \hat{\mu} < \mu, \quad (5.7)$$

where  $\hat{\mu}$  and  $\hat{\theta}$  denote values of the parameters  $\mu$  and  $\theta$  that maximise  $\mathcal{L}$  with respect to the observed data, while  $\hat{\theta}_\mu$  maximises  $\mathcal{L}$  for fixed  $\mu$ . To avoid a negative signal strength, which is unphysical,  $\hat{\mu}$  is restricted to be non-negative. For setting an upper limit, the testing of  $\mu$  is restricted to  $\mu > \hat{\mu}$ . For a given  $\mu$  hypothesis, the test statistic for the observed data,  $q_\mu^{\text{obs}}$ , is computed, and the probability distribution for  $q_\mu$  is numerically evaluated by considering toy experiments. Two  $q_\mu$  distributions are constructed: one using pseudo-data generated with background contributions only, while the other includes a signal with strength  $\mu$ . These two probability distributions provide confidence levels for a signal and background scenario and a background only scenario,

$$CL_{s+b}(\mu) = \mathcal{P}(q_\mu \geq q_\mu^{\text{obs}} | \mu \cdot s + b), \quad (5.8)$$

$$CL_b(\mu) = \mathcal{P}(q_\mu \geq q_\mu^{\text{obs}} | b). \quad (5.9)$$

$CL_{s+b}$  and  $CL_b$  are the probabilities to obtain values for  $q_\mu$  larger than  $q_\mu^{\text{obs}}$  for the

signal and background hypothesis and the background only hypothesis, respectively. The disadvantage of using  $CL_{s+b}$  directly to derive an upper limit is that downward fluctuations of the background can easily lead to the exclusion of even zero signal at 95% confidence level [73]. The modification to have more robustness against such fluctuations, and therefore to provide more conservative limits, is to consider the ratio,

$$CL_s(\mu) = \frac{CL_{s+b}(\mu)}{CL_b(\mu)}, \quad (5.10)$$

then  $\mu$  is the upper limit at the  $1 - CL_s(\mu)$  confidence level.

An analytic approximation for the probability distribution of  $q_\mu$  is valid for a sufficient number of events [75]. This approximation, which is derived using results from references [76, 77], uses a single representative dataset (also known as the ‘‘Asimov’’ dataset) taken to be the expected signal and background yields with the measured nuisances. This asymptotic approximation is much quicker to compute. It is used to calculate the limits in this analysis.

### 5.5.3 Limits

Limits on the cross section of the SM Higgs boson decaying to  $ZZ \rightarrow 2\ell 2\nu$  are computed and expressed in terms of the ratio to the predicted SM cross section. The Higgs boson of a particular mass is considered excluded when the observed cross section ratio limit falls below 1. The limits are presented in Table 5.7 and plotted in Figure 5-22. Interpolating between data points, The expected exclusion region is  $290 \text{ GeV}/c^2 < m_H < 500 \text{ GeV}/c^2$ , while the observed exclusion is  $260 \text{ GeV}/c^2 < m_H < 460 \text{ GeV}/c^2$ .

Mass [GeV/c <sup>2</sup> ]	Observed	Median	68% C.L. band	95% C.L. band
250	1.1	1.36	[0.96, 1.96]	[0.71, 2.73]
300	0.6	0.88	[0.62, 1.28]	[0.45, 1.81]
350	0.5	0.59	[0.41, 0.87]	[0.30, 1.26]
400	0.5	0.58	[0.40, 0.87]	[0.29, 1.29]
500	1.3	1.04	[0.69, 1.63]	[0.49, 2.51]
600	2.1	2.12	[1.35, 3.52]	[0.94, 5.87]

Table 5.7: The median expected cross section ratio limits as a function of the Higgs mass, together with the  $1\sigma$  and  $2\sigma$  uncertainty bands.

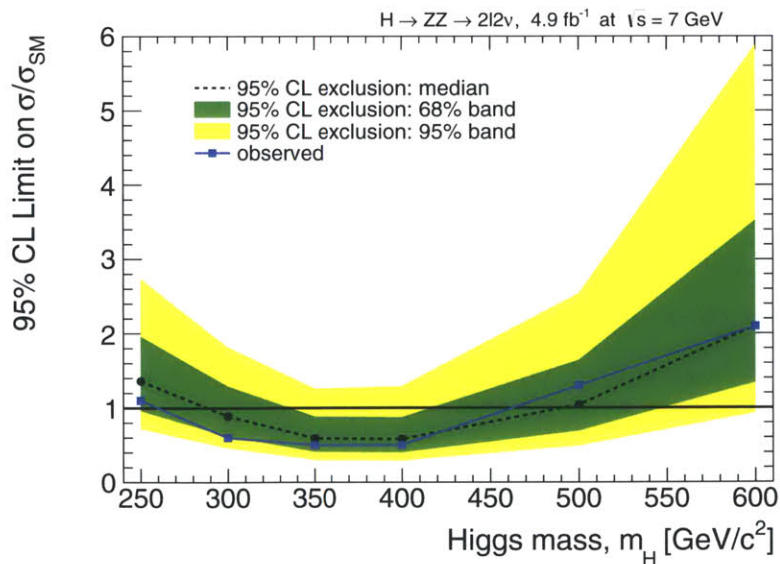


Figure 5-22: Upper limits on Standard Model Higgs production.

# Chapter 6

## Conclusion

The analysis of the 2011 data at the LHC in the  $ZZ \rightarrow 2\ell 2\nu$  channel revealed no evidence of the Standard Model Higgs boson in the mass range of  $250 \text{ GeV}/c^2$  to  $600 \text{ GeV}/c^2$ . Furthermore, a Higgs boson with mass between  $260 \text{ GeV}/c^2$  and  $460 \text{ GeV}/c^2$  is excluded. Combining all search channels analysed by the CMS collaboration, the SM Higgs boson is excluded to have a mass between  $127 \text{ GeV}/c^2$  to  $600 \text{ GeV}/c^2$  [78]. Of greater interest, however, are small excesses around  $120 \text{ GeV}/c^2$  to  $125 \text{ GeV}/c^2$  driven by observations in the  $\gamma\gamma$ ,  $ZZ \rightarrow 4\ell$ , and  $WW \rightarrow \ell\nu\ell'\nu'$  channels. By the summer of 2012, an excess around  $125 \text{ GeV}/c^2$  had attained enough significance to declare the observation of a new boson [79], in conjunction with an independent observation by the ATLAS collaboration [80]. Before anointing this newfound particle to be the Standard Model Higgs boson, it is necessary to establish whether its properties conform to expectations. Within a couple years of operation, the LHC project has delivered a new, tantalizing hint to further our understanding of Nature, and it is anticipated that with more data, more will be uncovered.



# Appendix A

## Electron Energy Scale and Resolution

Electron energy scale and resolution corrections to achieve better agreement between data and simulation are extracted in bins of supercluster  $|\eta|$ . The strategy assumes that a per electron energy scale shift and Gaussian smearing function is sufficient to modify the dielectron mass distribution in simulation to match the data. Six bins of  $|\eta|$  are considered, resulting in 21 combinations of electron pairs. Using the model of a template from simulated  $Z$  events convoluted with a Gaussian function, a simultaneous fit of all 21 categories is performed to extract the energy scale shifts and smearing constants. The results are tabulated in Table A.1. Figures A-1, A-2, and A-3 show plots of comparing data, simulation, and the fitted model.

Region	Scale	Smear (GeV)
$0 <  \eta  < 0.4$	$1.0009 \pm 0.0001$	$0.337 \pm 0.022$
$0.4 <  \eta  < 0.8$	$1.0009 \pm 0.0001$	$0.248 \pm 0.031$
$0.8 <  \eta  < 1.2$	$1.0035 \pm 0.0001$	$0.614 \pm 0.018$
$1.2 <  \eta  < 1.4442$	$1.0058 \pm 0.0002$	$1.082 \pm 0.022$
$1.566 <  \eta  < 2.1$	$0.9994 \pm 0.0002$	$1.063 \pm 0.021$
$2.1 <  \eta  < 2.5$	$1.0013 \pm 0.0002$	$1.035 \pm 0.028$

Table A.1: Electron energy scale and resolution corrections in each  $|\eta|$  bin. The scale correction is a multiplicative factor to be applied to electrons in data. The smearing constants are to be applied to simulated electrons.

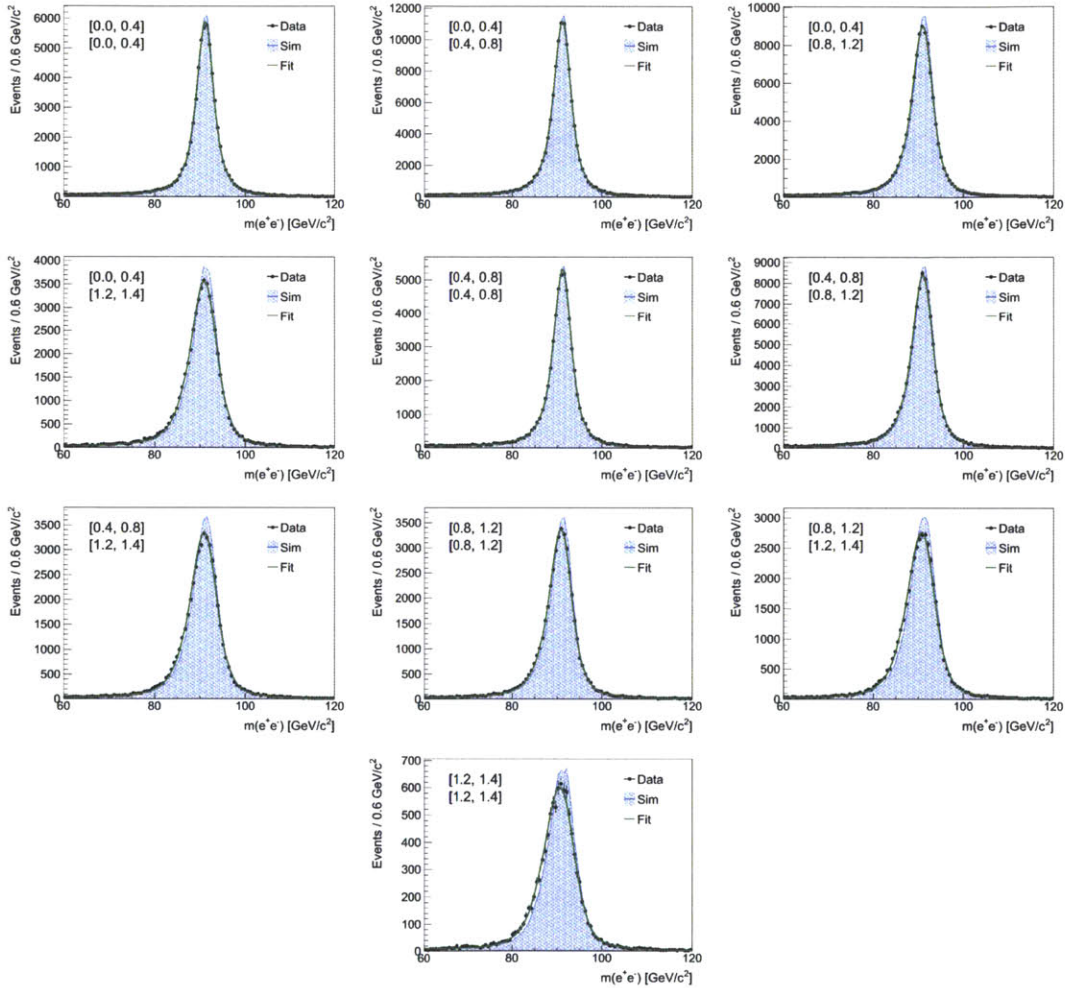


Figure A-1: Mass distributions for dielectrons with both electrons in the barrel comparing data, simulation, and simulation corrected for scale and resolution from the fit.

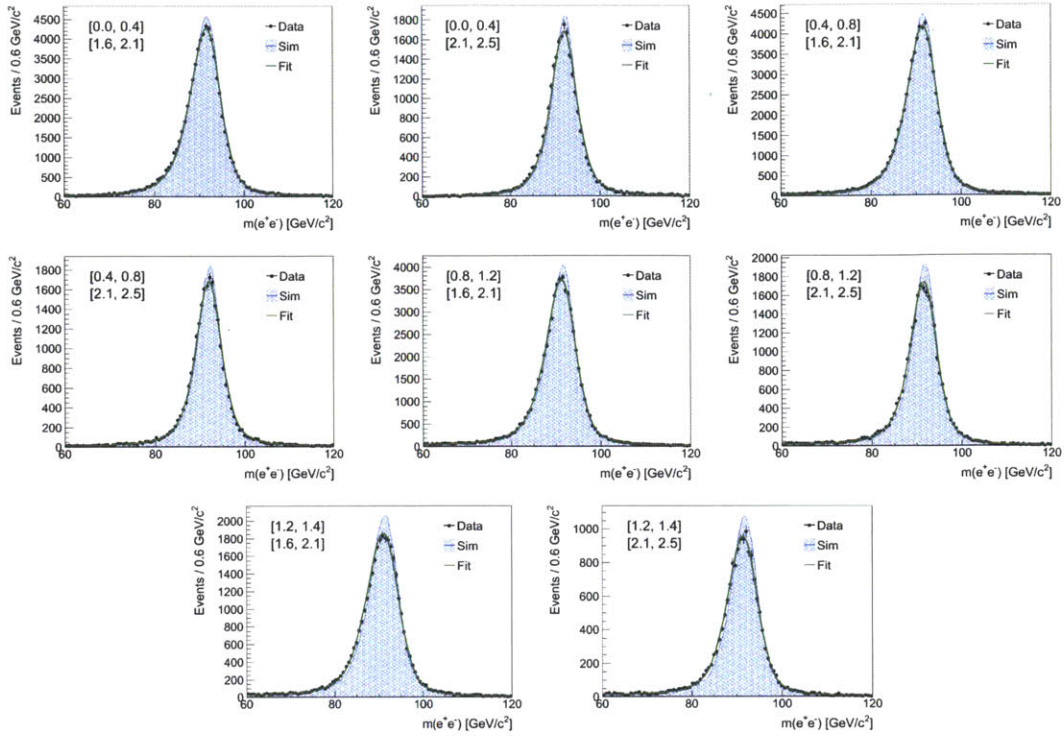


Figure A-2: Mass distributions for dielectrons with one electron in the barrel and one in the endcap comparing data, simulation, and simulation corrected for scale and resolution from the fit.

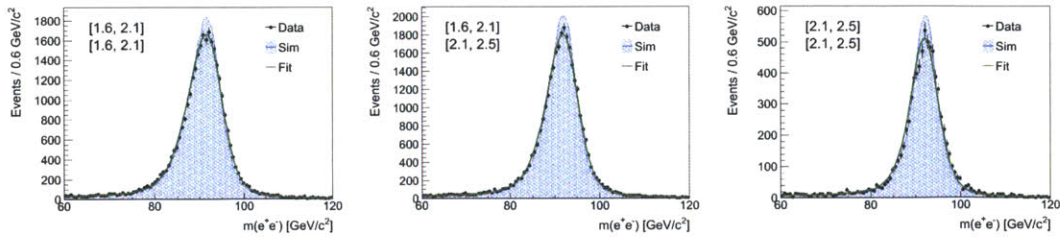


Figure A-3: Mass distributions for dielectrons with both electrons in the endcap comparing data, simulation, and simulation corrected for scale and resolution from the fit.





# Appendix B

## Lepton Efficiency Fits and Tables

This appendix documents all the fits performed to measure the various lepton efficiencies and provides tables for the efficiencies extracted from data and from simulation. In addition to partitioning the events in  $p_T$  and  $\eta$ , efficiencies in data are also measured separately in Run2011A and Run2011B, corresponding to  $2.2 \text{ fb}^{-1}$  and  $2.7 \text{ fb}^{-1}$  respectively, and averaged together to compute the efficiency scale factors. The data is split in this fashion because the transition from Run2011A to Run2011B represents the largest change in instantaneous luminosity and pile-up conditions. Furthermore, the triggers used to select events for efficiency measurements are prescaled more heavily in Run2011B, hence the separation of the two run eras reduces the bias towards Run2011A conditions in estimating overall efficiencies.

### B.1 Electron Efficiencies

#### B.1.1 GSF Reconstruction Efficiency

The GSF reconstruction efficiency is measured using supercluster probes. The efficiencies in simulation and data (averaged over Run2011A and Run2011B) are listed in Tables B.1 and B.2 respectively, and the scale factors are listed in Table B.3. Fits for Run2011A are shown in Figures B-1 through B-5 and for Run2011B are shown in Figures B-6 through B-10.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.9443 \pm 0.0012$	$0.9423 \pm 0.0015$	$0.9297 \pm 0.0018$	$0.8922 \pm 0.0027$
$30 < E_T < 40$	$0.9757 \pm 0.0005$	$0.9711 \pm 0.0007$	$0.9547 \pm 0.0011$	$0.9170 \pm 0.0019$
$40 < E_T < 50$	$0.9824 \pm 0.0004$	$0.9786 \pm 0.0006$	$0.9672 \pm 0.0009$	$0.9348 \pm 0.0018$
$50 < E_T < 60$	$0.9859 \pm 0.0008$	$0.9822 \pm 0.0011$	$0.9713 \pm 0.0019$	$0.9400 \pm 0.0037$
$E_T > 60$	$0.9874 \pm 0.0011$	$0.9841 \pm 0.0015$	$0.9740 \pm 0.0026$	$0.9405 \pm 0.0059$

Table B.1: GSF reconstruction efficiency in simulation.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.9438 \pm 0.0076$	$0.9231 \pm 0.0065$	$0.9275 \pm 0.0028$	$0.8929 \pm 0.0047$
$30 < E_T < 40$	$0.9703 \pm 0.0011$	$0.9678 \pm 0.0015$	$0.9512 \pm 0.0033$	$0.9102 \pm 0.0042$
$40 < E_T < 50$	$0.9783 \pm 0.0005$	$0.9756 \pm 0.0005$	$0.9637 \pm 0.0008$	$0.9278 \pm 0.0015$
$50 < E_T < 60$	$0.9782 \pm 0.0026$	$0.9760 \pm 0.0015$	$0.9677 \pm 0.0029$	$0.9309 \pm 0.0032$
$E_T > 60$	$0.9836 \pm 0.0053$	$0.9797 \pm 0.0020$	$0.9753 \pm 0.0049$	$0.9397 \pm 0.0172$

Table B.2: GSF reconstruction efficiency in data averaged over Run2011A and Run2011B.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.9995 \pm 0.0081$	$0.9796 \pm 0.0071$	$0.9977 \pm 0.0035$	$1.0008 \pm 0.0061$
$30 < E_T < 40$	$0.9944 \pm 0.0012$	$0.9966 \pm 0.0017$	$0.9963 \pm 0.0036$	$0.9926 \pm 0.0050$
$40 < E_T < 50$	$0.9958 \pm 0.0007$	$0.9969 \pm 0.0008$	$0.9964 \pm 0.0012$	$0.9924 \pm 0.0025$
$50 < E_T < 60$	$0.9922 \pm 0.0027$	$0.9937 \pm 0.0019$	$0.9962 \pm 0.0036$	$0.9903 \pm 0.0052$
$E_T < 60$	$0.9962 \pm 0.0054$	$0.9956 \pm 0.0025$	$1.0013 \pm 0.0056$	$0.9991 \pm 0.0193$

Table B.3: Data to simulation scale factors for GSF reconstruction efficiency.

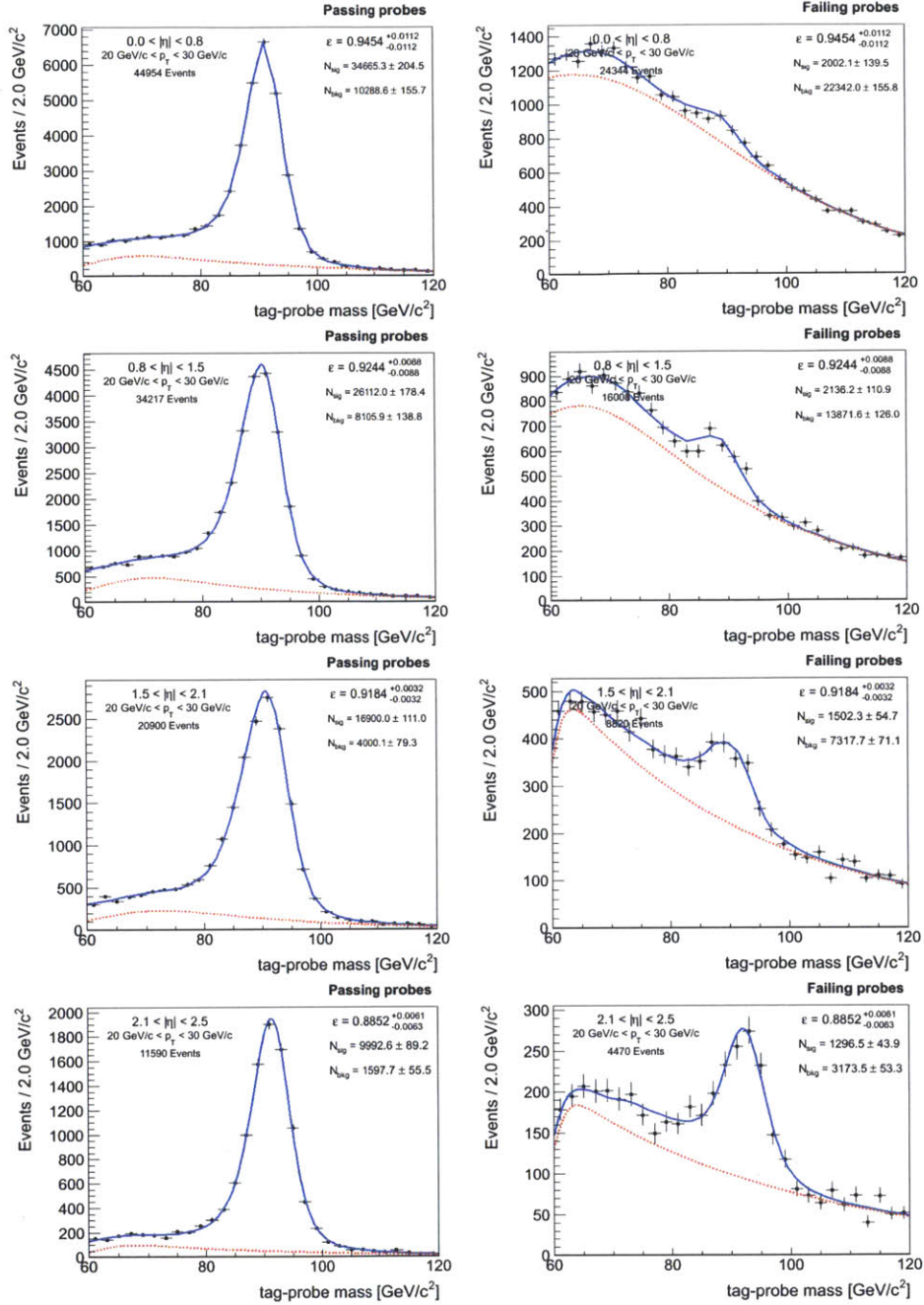


Figure B-1: Fits for GSF reconstruction efficiency for probes with  $20 \text{ GeV} < E_T < 30 \text{ GeV}$  in Run2011A data.

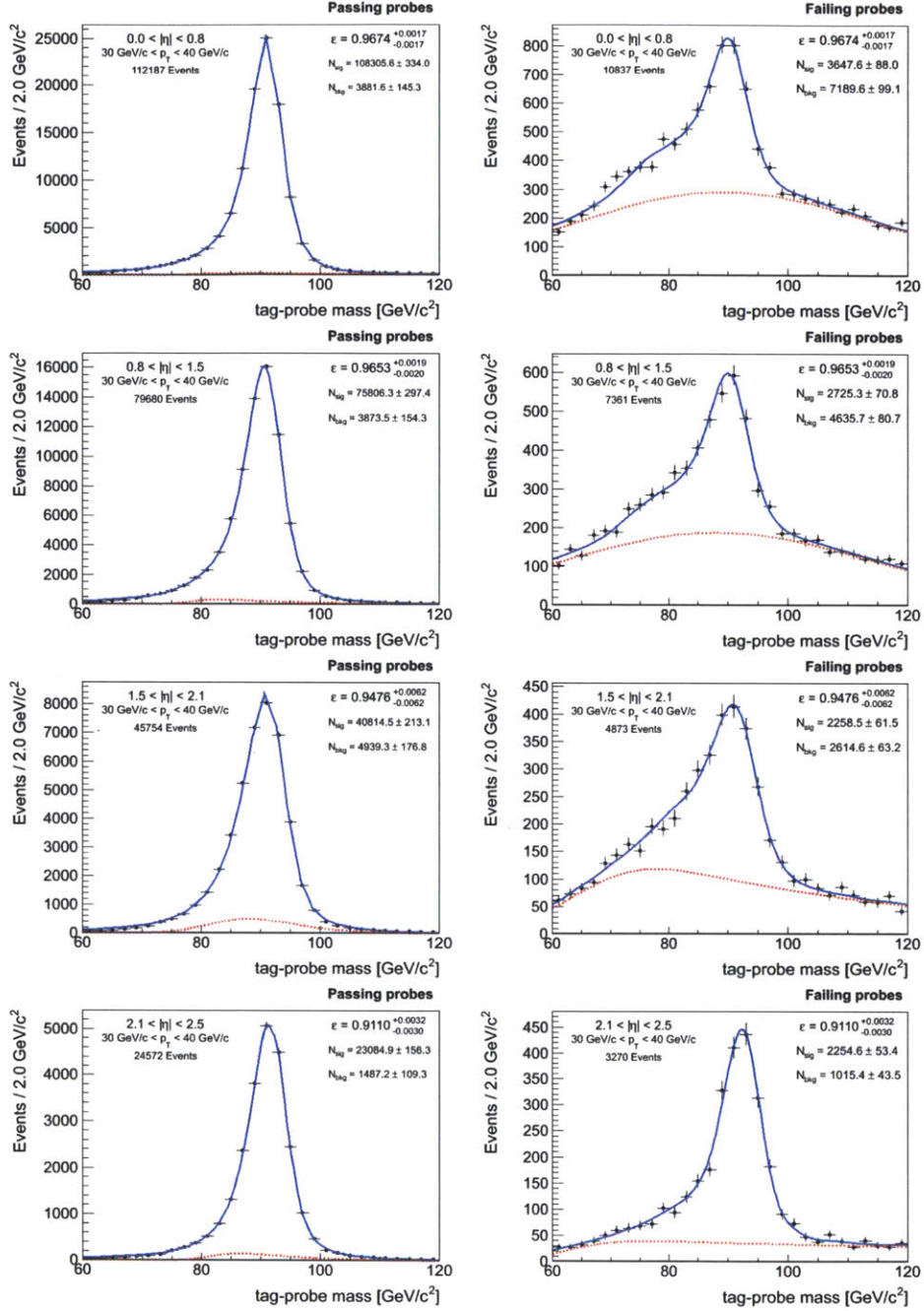


Figure B-2: Fits for GSF reconstruction efficiency for probes with  $30 \text{ GeV} < E_T < 40 \text{ GeV}$  in Run2011A data.

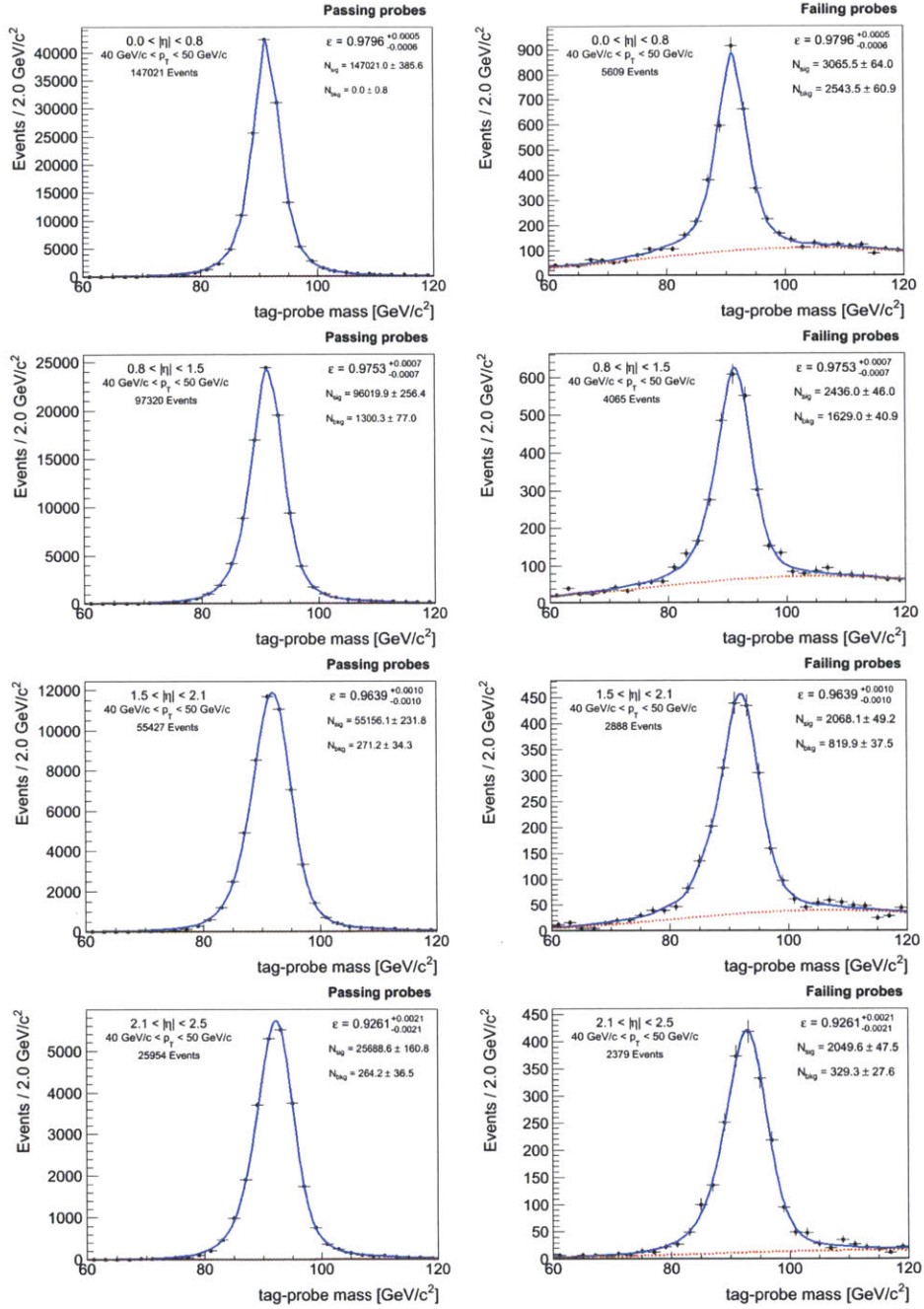


Figure B-3: Fits for GSF reconstruction efficiency for probes with  $40 \text{ GeV} < E_T < 50 \text{ GeV}$  in Run2011A data.

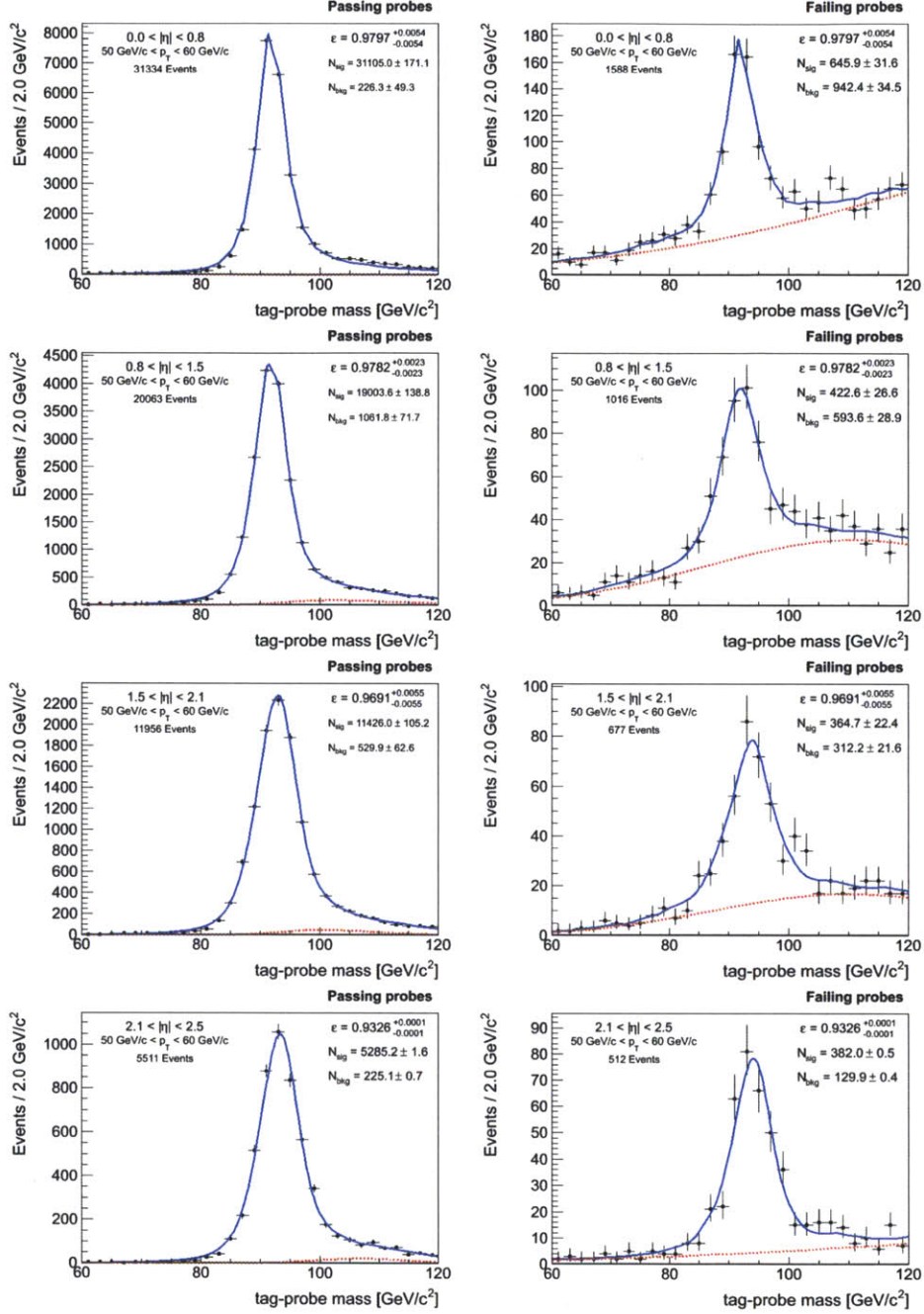


Figure B-4: Fits for GSF reconstruction efficiency for probes with  $50 \text{ GeV} < E_T < 60 \text{ GeV}$  in Run2011A data.



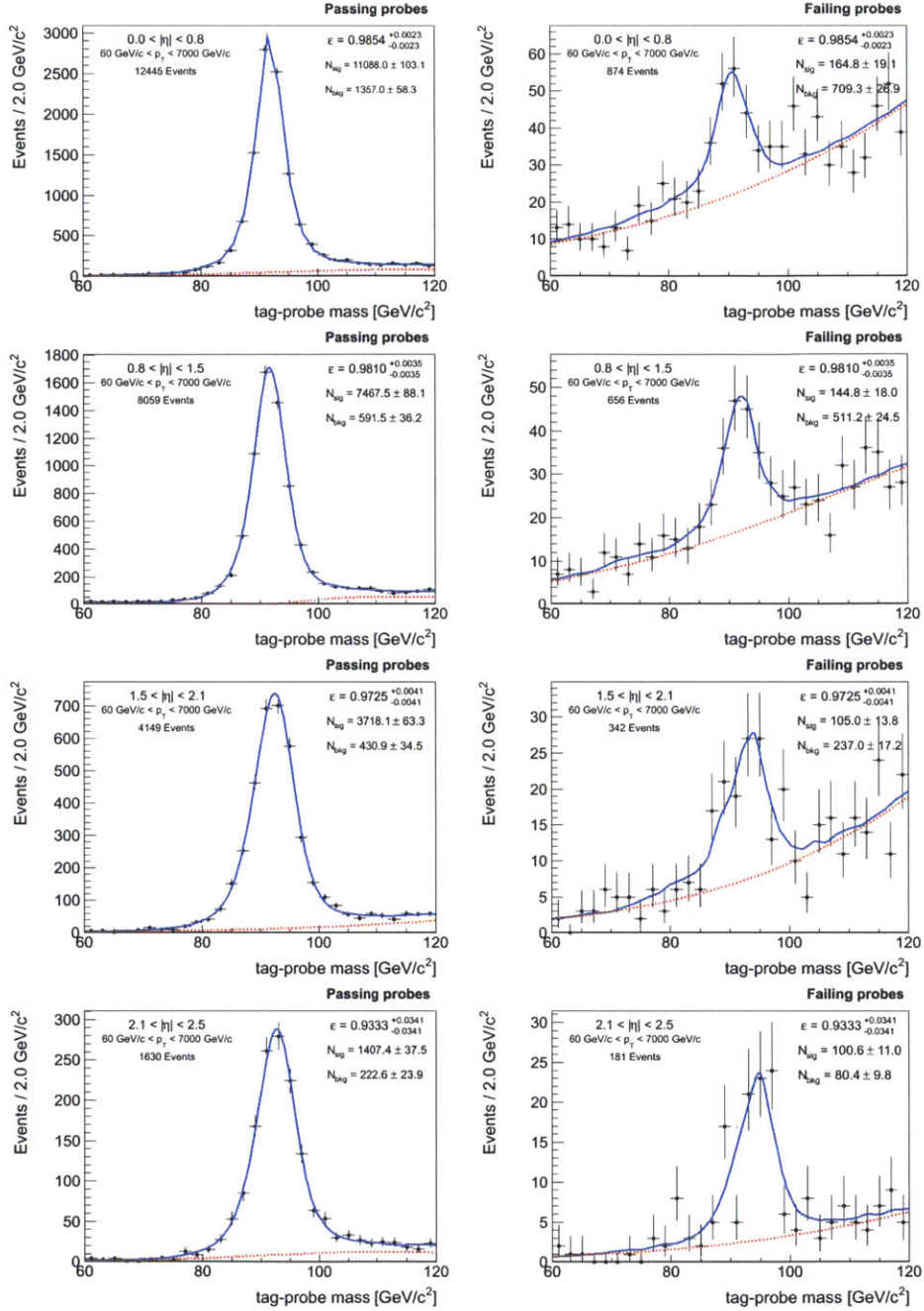


Figure B-5: Fits for GSF reconstruction efficiency for probes with  $E_T > 60$  GeV in Run2011A data.



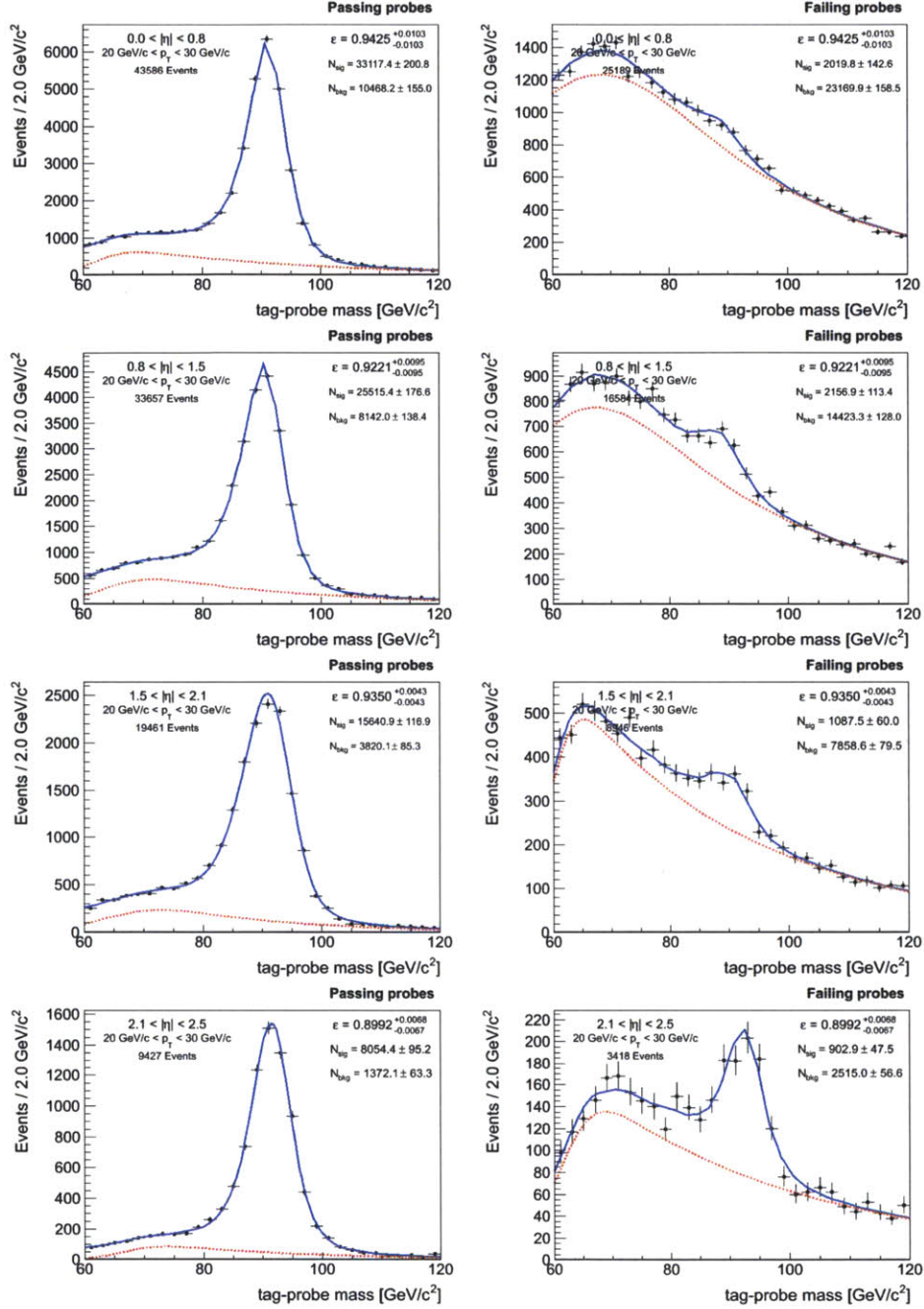


Figure B-6: Fits for GSF reconstruction efficiency for probes with  $20 \text{ GeV} < E_T < 30 \text{ GeV}$  in Run2011B data.

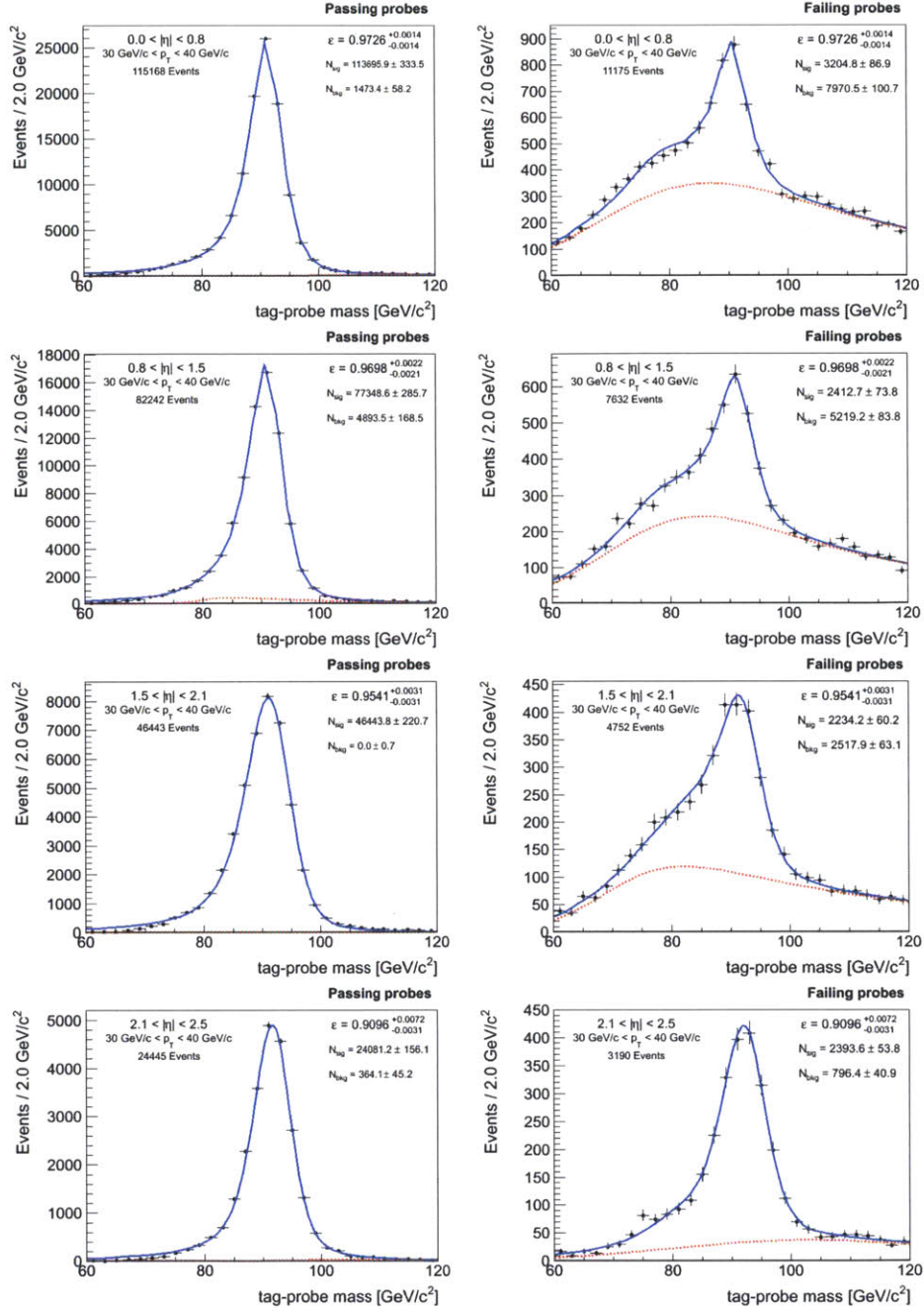


Figure B-7: Fits for GSF reconstruction efficiency for probes with  $30 \text{ GeV} < E_T < 40 \text{ GeV}$  in Run2011B data.

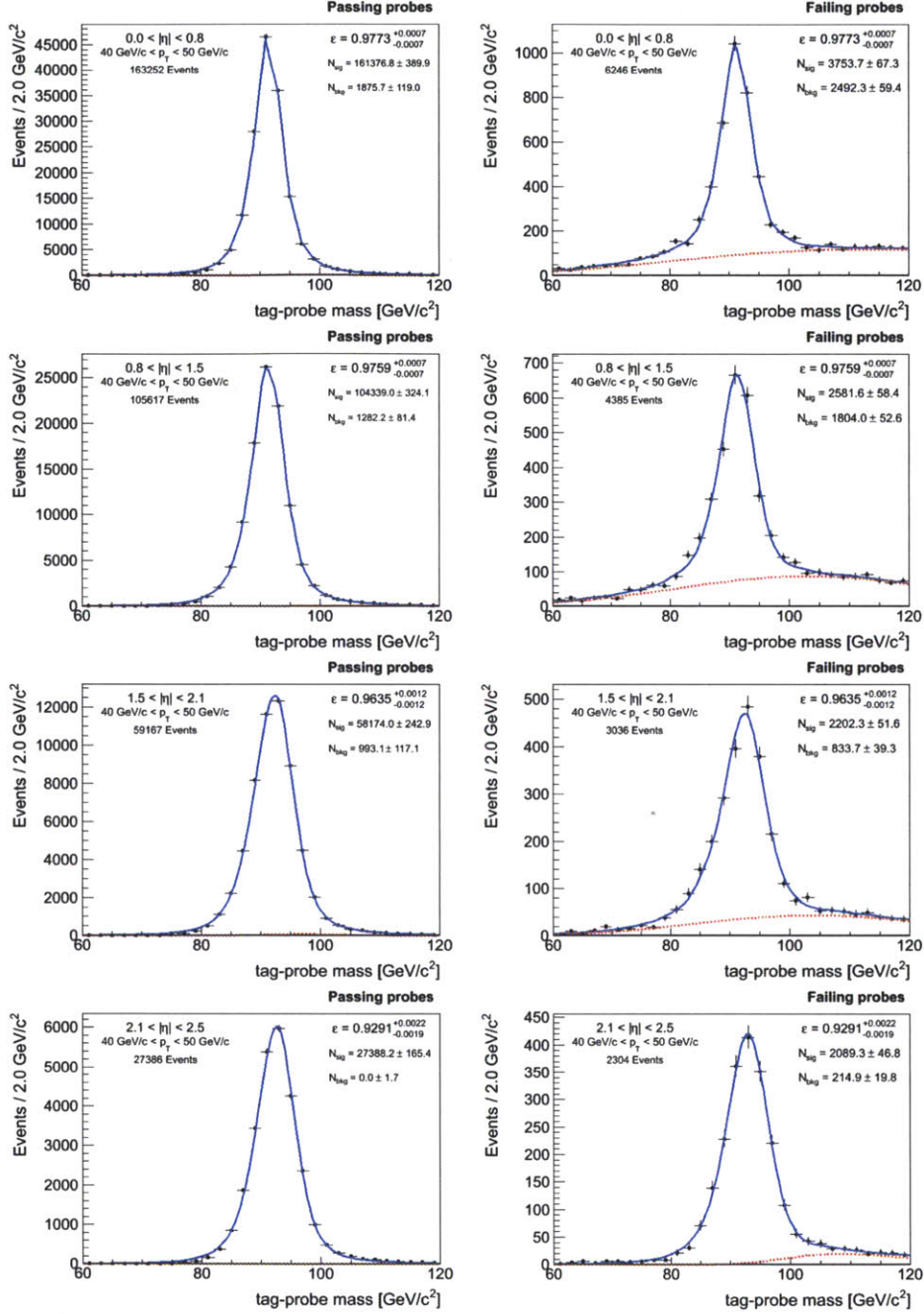


Figure B-8: Fits for GSF reconstruction efficiency for probes with  $40 \text{ GeV} < E_T < 50 \text{ GeV}$  in Run2011B data.

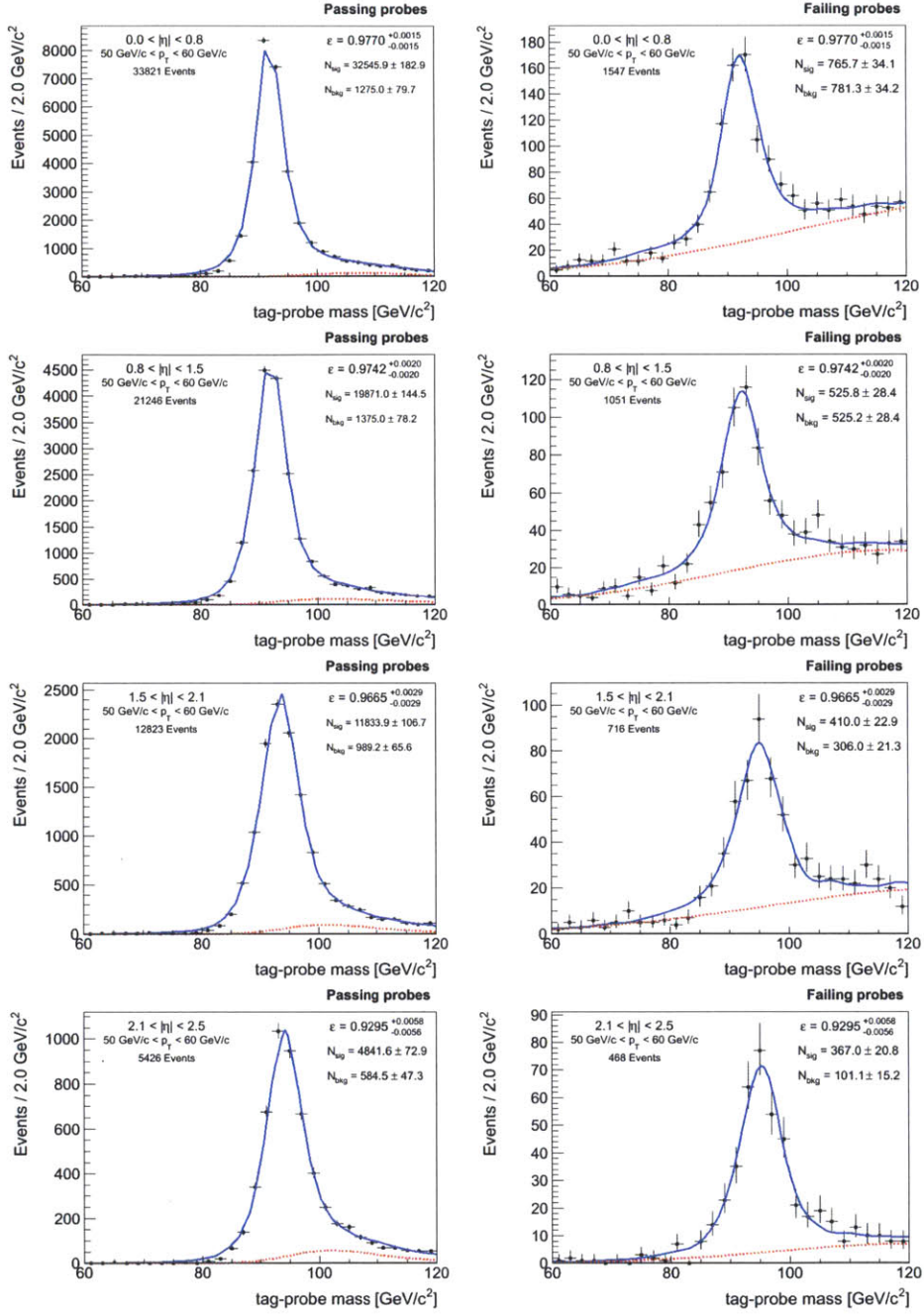


Figure B-9: Fits for GSF reconstruction efficiency for probes with  $50 \text{ GeV} < E_T < 60 \text{ GeV}$  in Run2011B data.



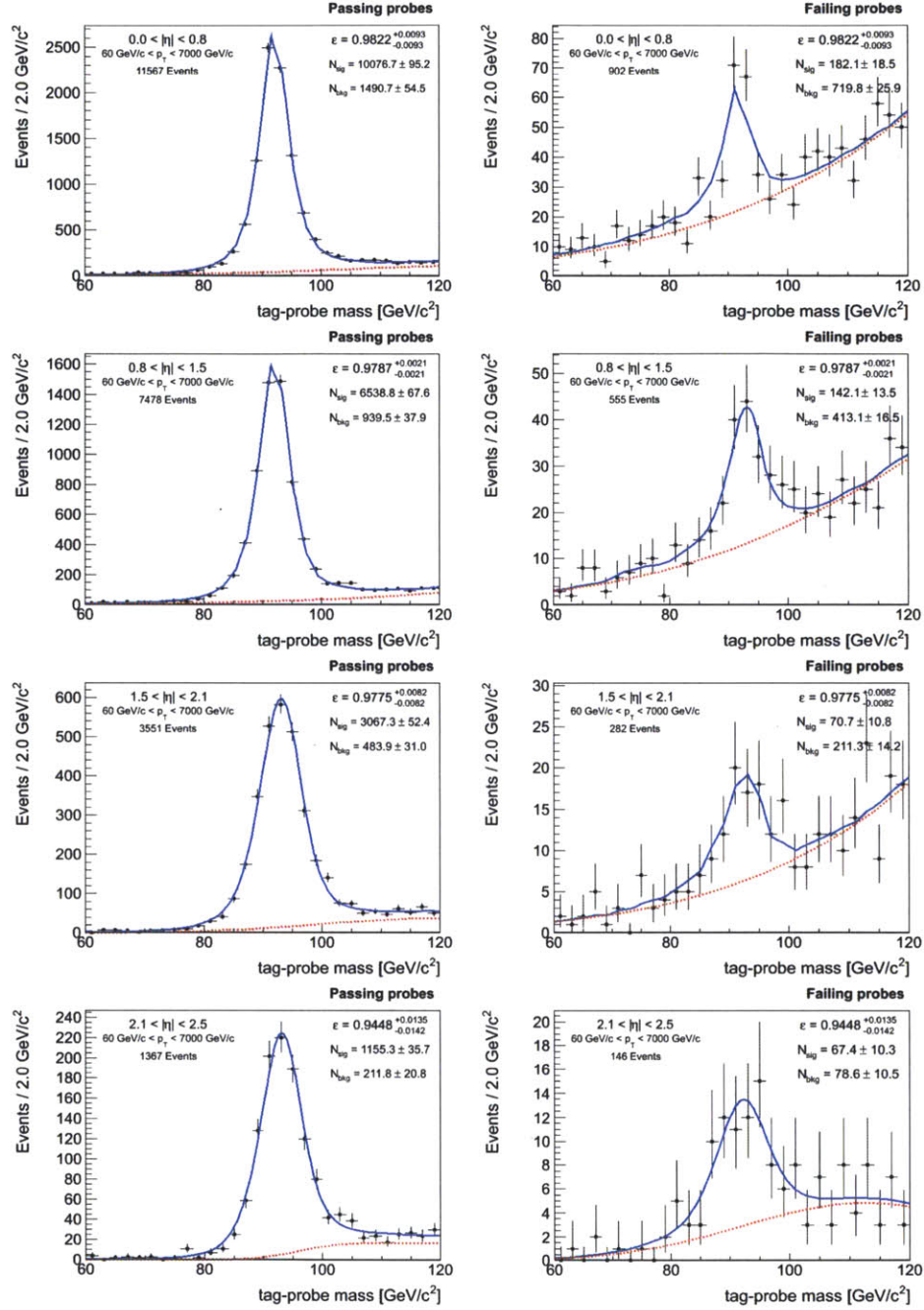


Figure B-10: Fits for GSF reconstruction efficiency for probes with  $E_T > 60$  GeV in Run2011B data.

## B.1.2 Electron Selection Efficiency

The electron selection efficiency is measured using GSF electron probes. The efficiencies in simulation and data (averaged over Run2011A and Run2011B) are listed in Tables B.4 and B.5 respectively, and the scale factors are listed in Table B.6. Fits for Run2011A are shown in Figures B-11 through B-15 and for Run2011B are shown in Figures B-16 through B-20.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.7524 \pm 0.0024$	$0.6921 \pm 0.0029$	$0.5652 \pm 0.0036$	$0.4997 \pm 0.0046$
$30 < E_T < 40$	$0.8297 \pm 0.0013$	$0.7950 \pm 0.0017$	$0.6589 \pm 0.0026$	$0.5499 \pm 0.0036$
$40 < E_T < 50$	$0.8680 \pm 0.0011$	$0.8540 \pm 0.0014$	$0.7192 \pm 0.0024$	$0.5922 \pm 0.0037$
$50 < E_T < 60$	$0.8757 \pm 0.0022$	$0.8732 \pm 0.0028$	$0.7422 \pm 0.0048$	$0.6207 \pm 0.0077$
$E_T > 60$	$0.8830 \pm 0.0030$	$0.8882 \pm 0.0037$	$0.7489 \pm 0.0070$	$0.6404 \pm 0.0118$

Table B.4: Electron selection efficiency in simulation.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.7320 \pm 0.0020$	$0.6571 \pm 0.0022$	$0.5600 \pm 0.0027$	$0.4800 \pm 0.0026$
$30 < E_T < 40$	$0.8119 \pm 0.0008$	$0.7644 \pm 0.0011$	$0.6426 \pm 0.0012$	$0.5204 \pm 0.0022$
$40 < E_T < 50$	$0.8479 \pm 0.0005$	$0.8313 \pm 0.0031$	$0.7041 \pm 0.0009$	$0.5605 \pm 0.0013$
$50 < E_T < 60$	$0.8563 \pm 0.0013$	$0.8544 \pm 0.0014$	$0.7337 \pm 0.0022$	$0.5828 \pm 0.0038$
$E_T > 60$	$0.8614 \pm 0.0020$	$0.8611 \pm 0.0029$	$0.7486 \pm 0.0043$	$0.6079 \pm 0.0076$

Table B.5: Electron selection efficiency in data averaged over Run2011A and Run2011B.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$1.5 <  \eta  < 2.1$	$2.1 <  \eta  < 2.5$
$20 < E_T < 30$	$0.9728 \pm 0.0040$	$0.9494 \pm 0.0051$	$0.9908 \pm 0.0078$	$0.9604 \pm 0.0103$
$30 < E_T < 40$	$0.9785 \pm 0.0018$	$0.9614 \pm 0.0025$	$0.9752 \pm 0.0043$	$0.9463 \pm 0.0075$
$40 < E_T < 50$	$0.9769 \pm 0.0014$	$0.9734 \pm 0.0039$	$0.9789 \pm 0.0035$	$0.9466 \pm 0.0063$
$50 < E_T < 60$	$0.9779 \pm 0.0029$	$0.9785 \pm 0.0035$	$0.9885 \pm 0.0070$	$0.9390 \pm 0.0132$
$E_T > 60$	$0.9755 \pm 0.0040$	$0.9694 \pm 0.0052$	$0.9996 \pm 0.0110$	$0.9493 \pm 0.0212$

Table B.6: Data to simulation scale factors for electron selection efficiency.

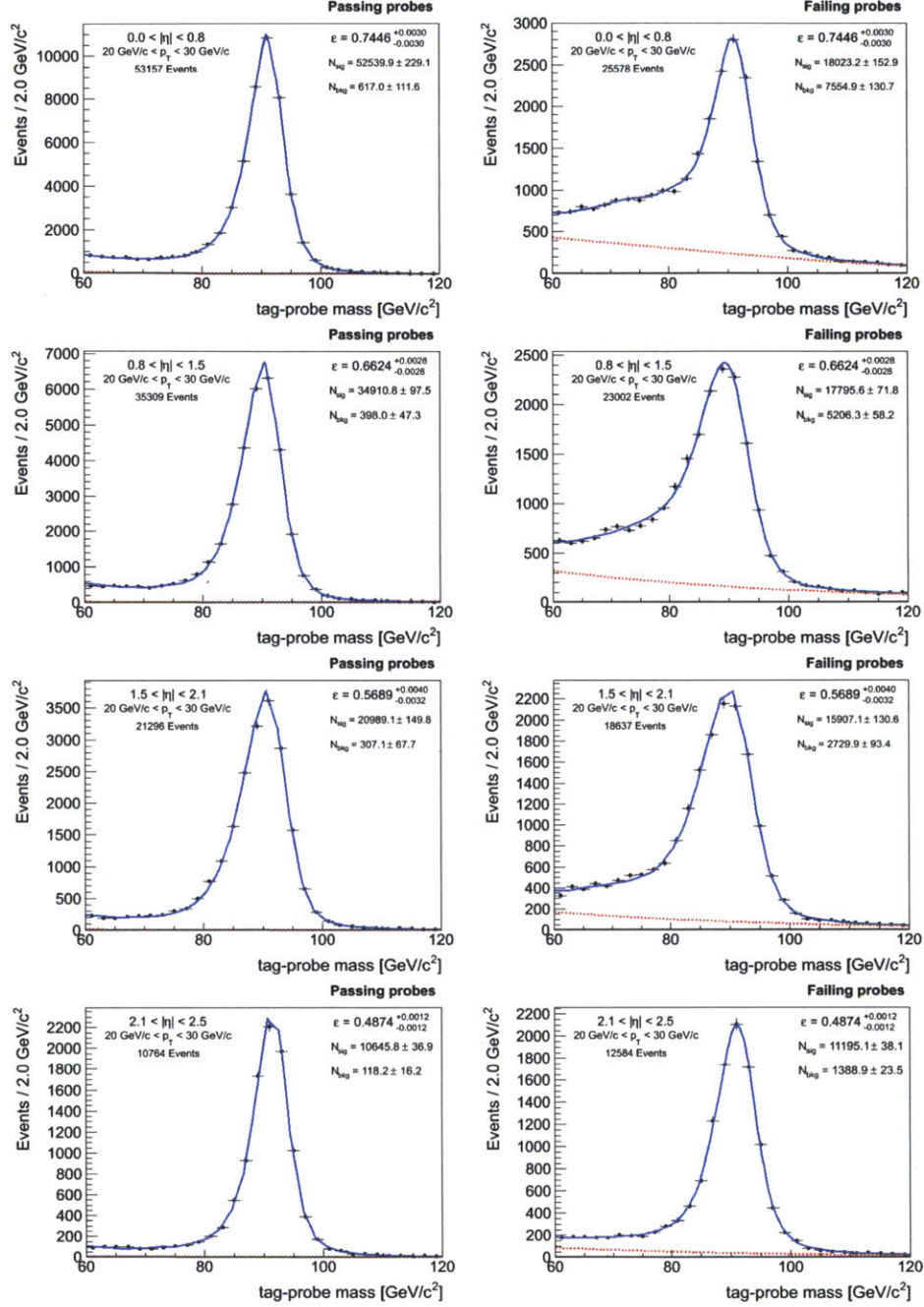


Figure B-11: Fits for electron selection efficiency for probes with  $20 \text{ GeV} < E_T < 30 \text{ GeV}$  in Run2011A data.

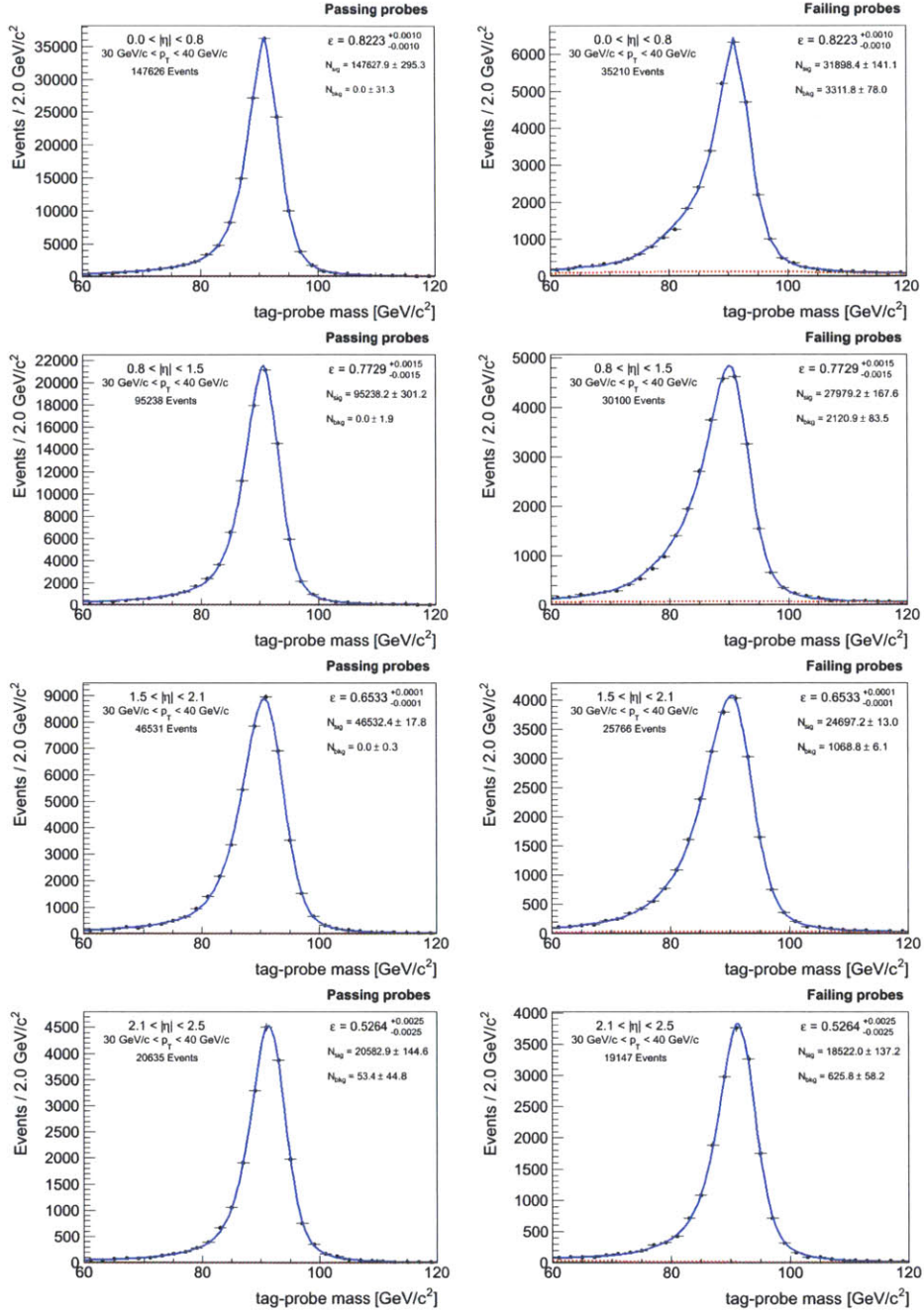


Figure B-12: Fits for electron selection efficiency for probes with  $30 \text{ GeV} < E_T < 40 \text{ GeV}$  in Run2011A data.



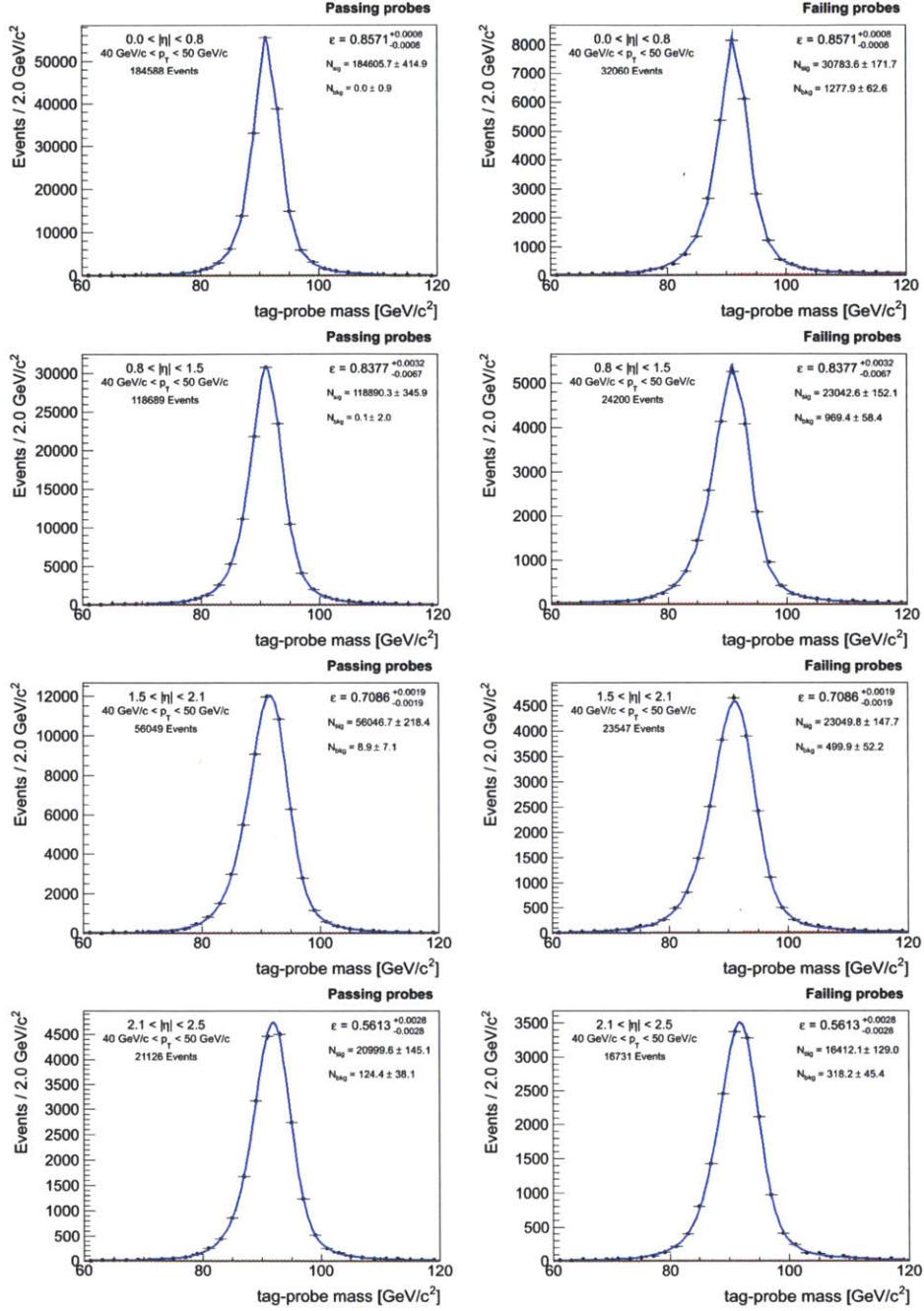


Figure B-13: Fits for electron selection efficiency for probes with  $40 \text{ GeV} < E_T < 50 \text{ GeV}$  in Run2011A data.

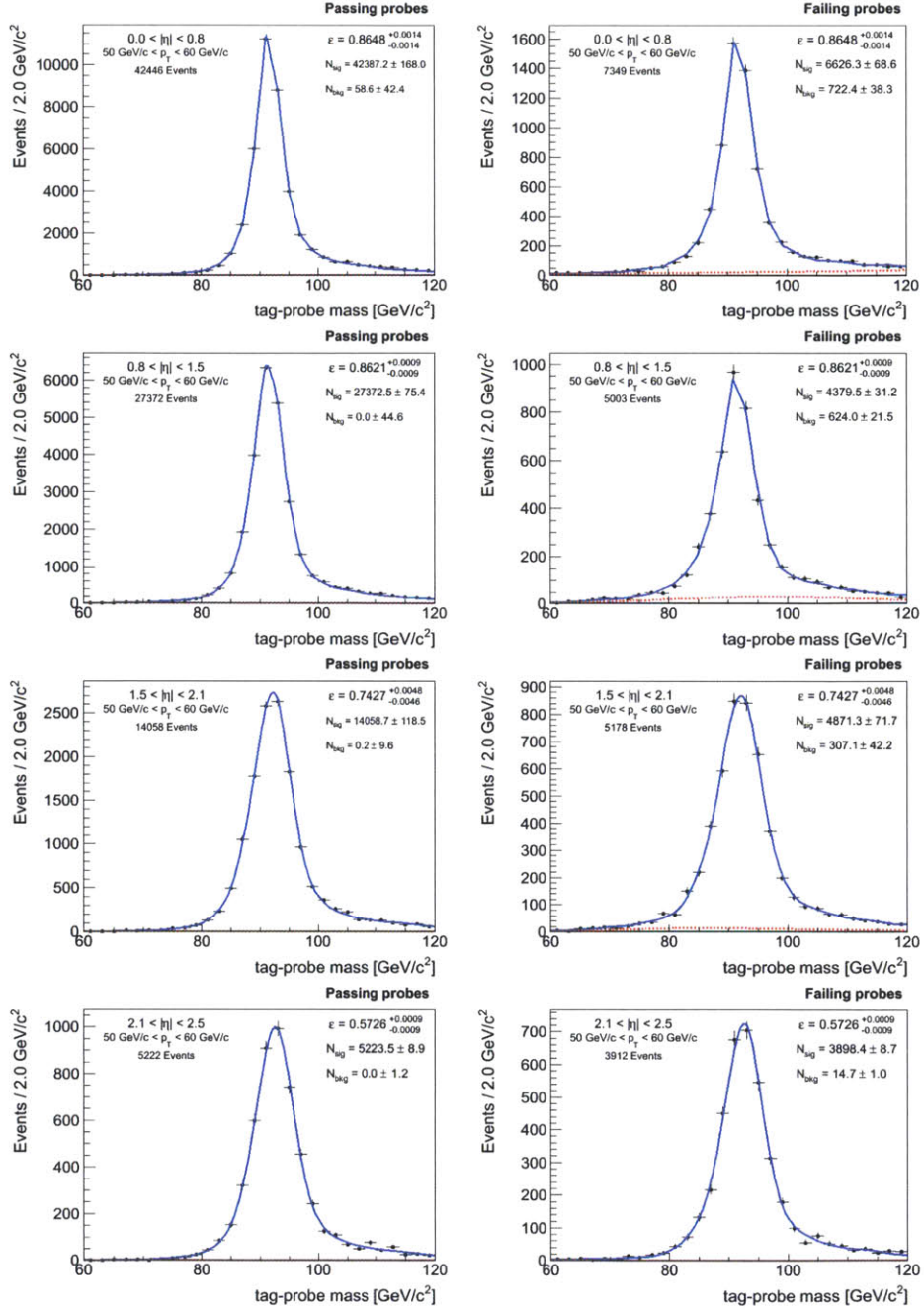


Figure B-14: Fits for electron selection efficiency for probes with  $50 \text{ GeV} < E_T < 60 \text{ GeV}$  in Run2011A data.

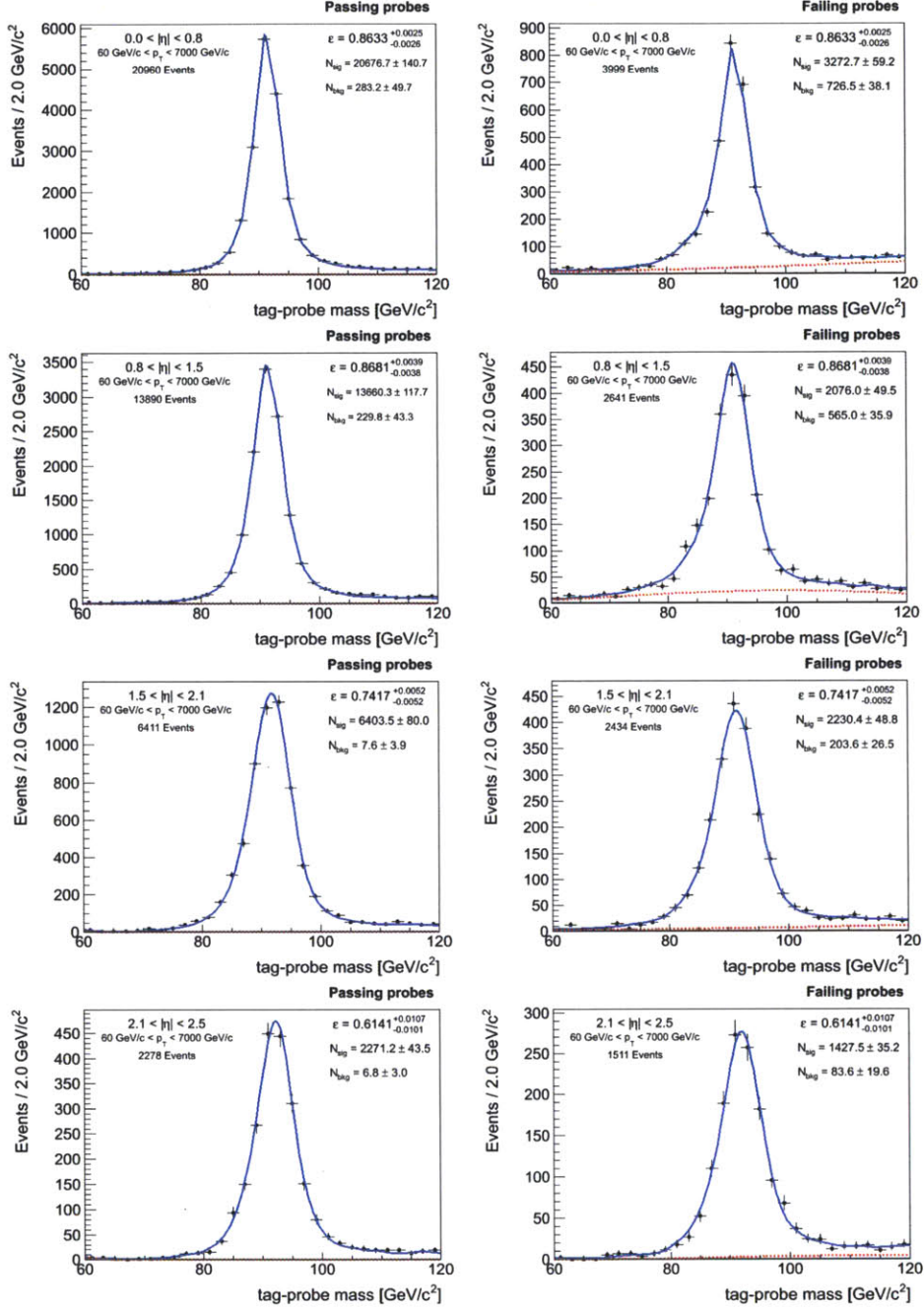


Figure B-15: Fits for electron selection efficiency for probes with  $E_T > 60$  GeV in Run2011A data.

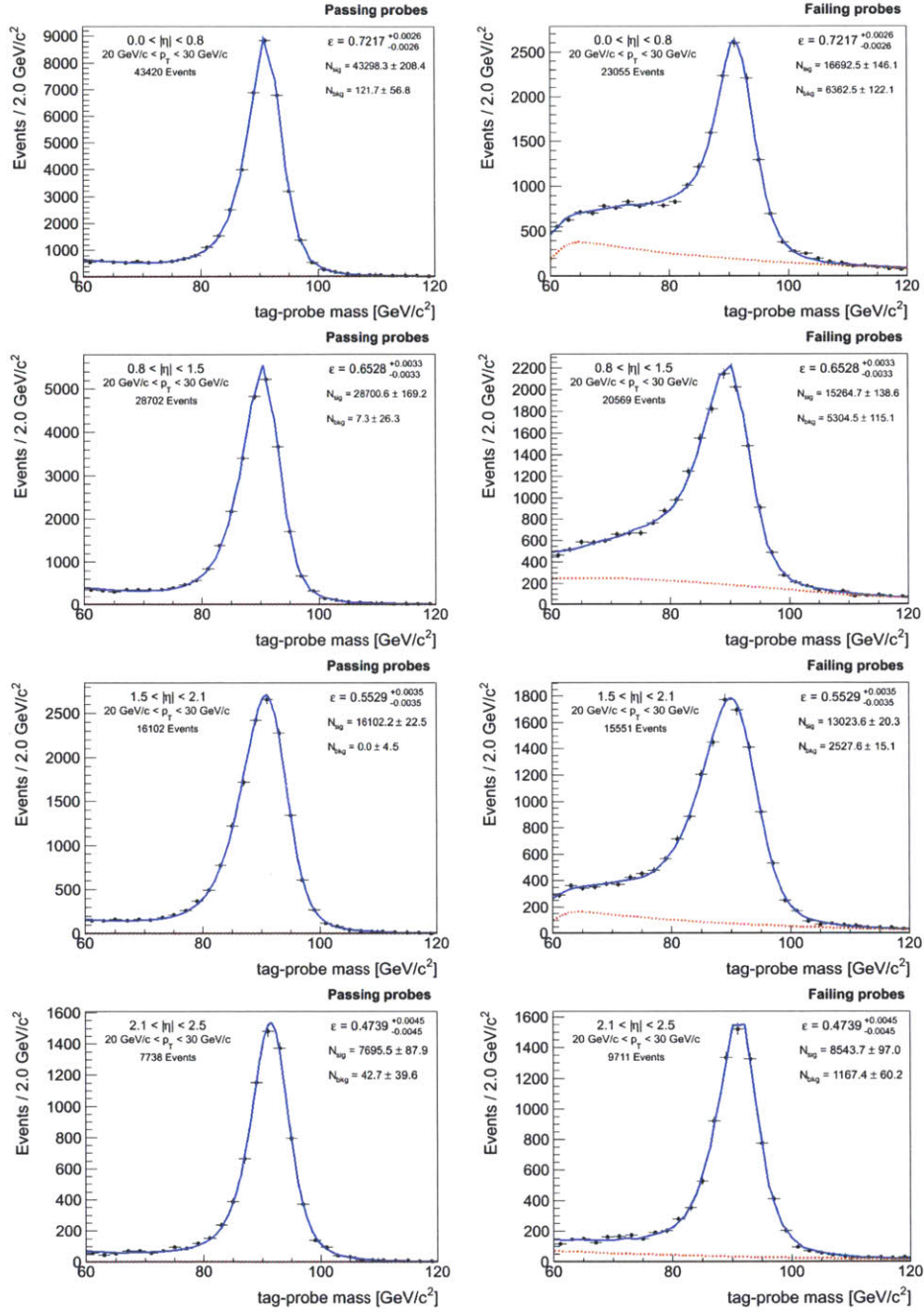


Figure B-16: Fits for electron selection efficiency for probes with  $20 \text{ GeV} < E_T < 30 \text{ GeV}$  in Run2011B data.

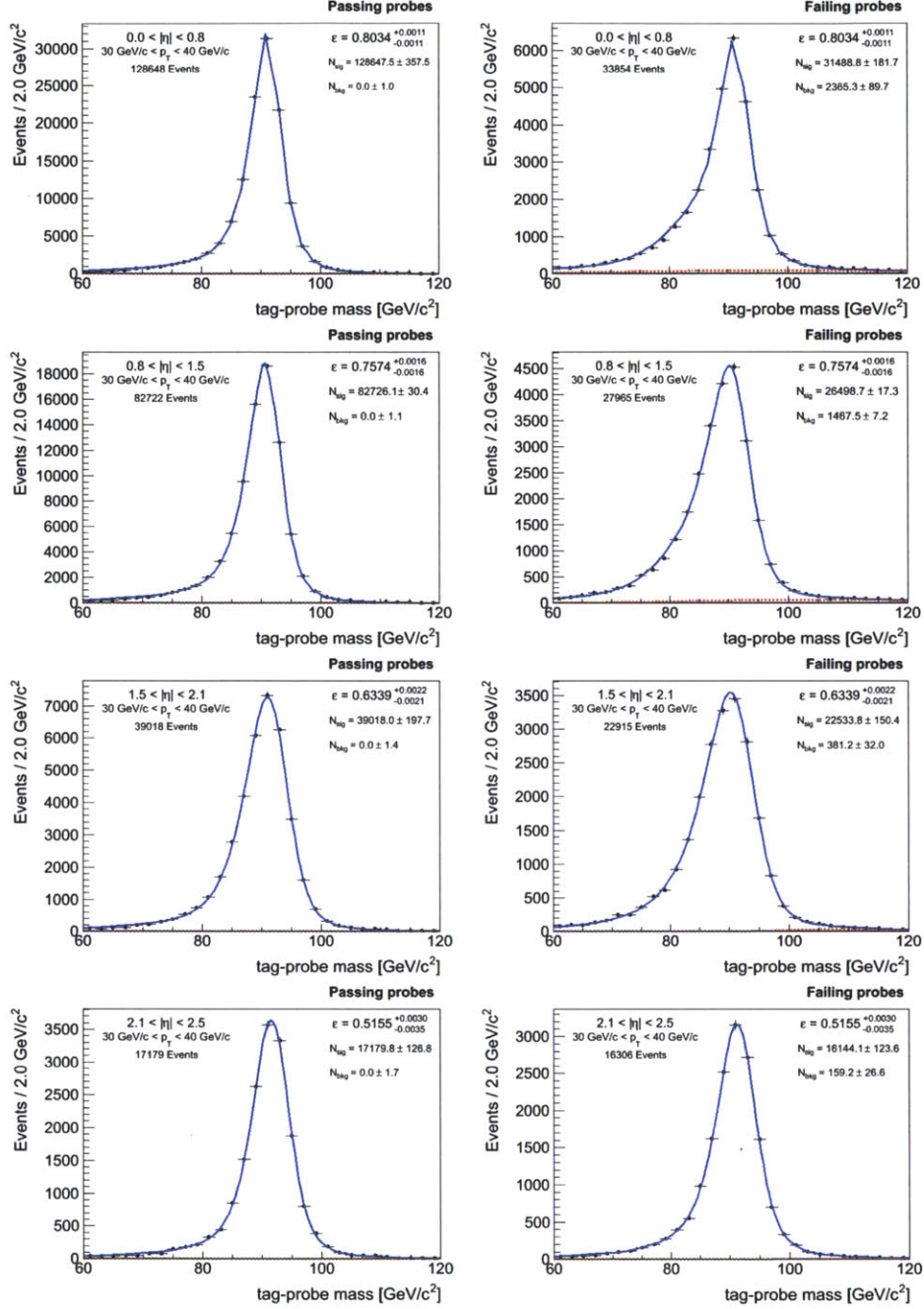


Figure B-17: Fits for electron selection efficiency for probes with  $30 \text{ GeV} < E_T < 40 \text{ GeV}$  in Run2011B data.



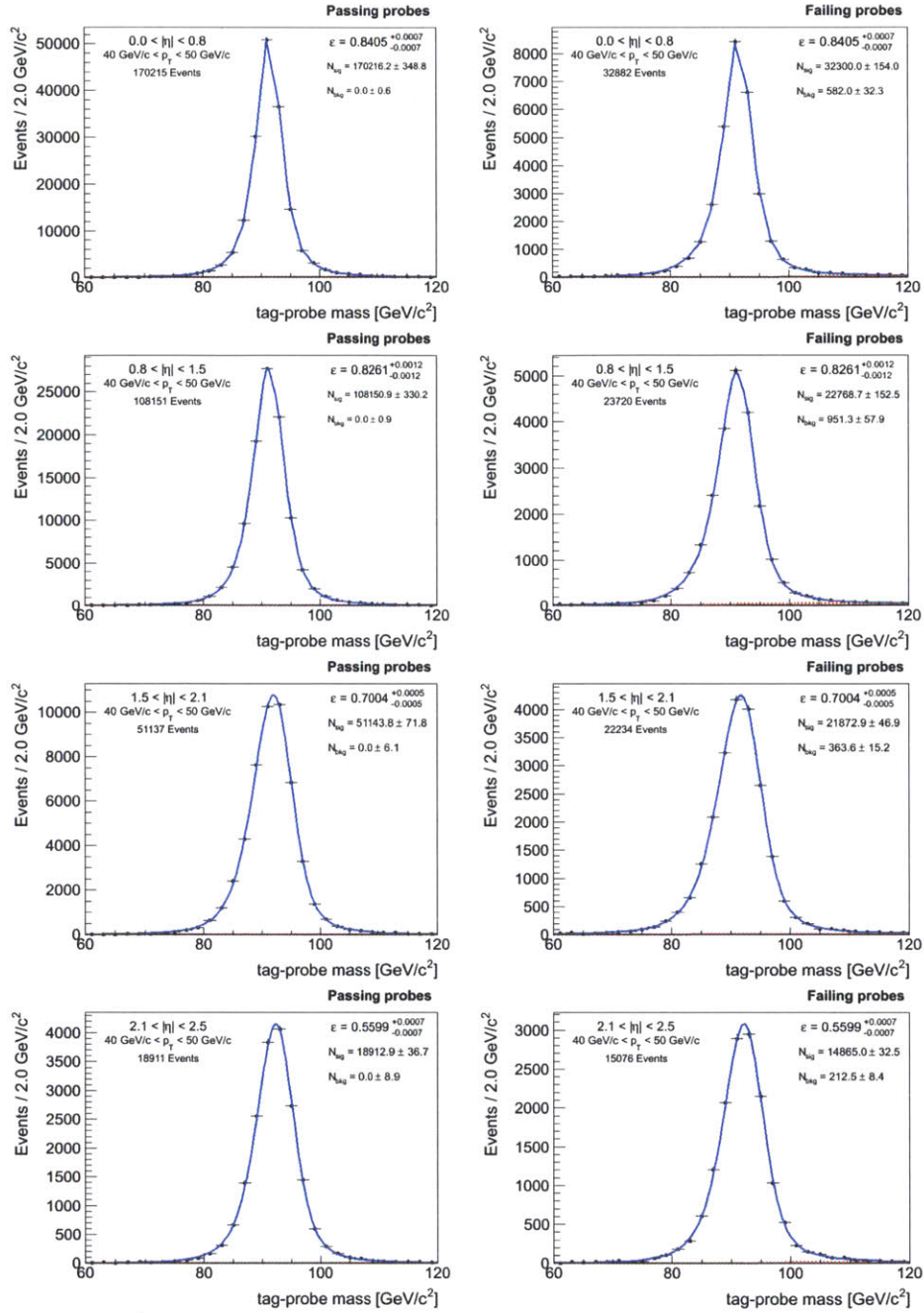


Figure B-18: Fits for electron selection efficiency for probes with  $40 \text{ GeV} < E_T < 50 \text{ GeV}$  in Run2011B data.

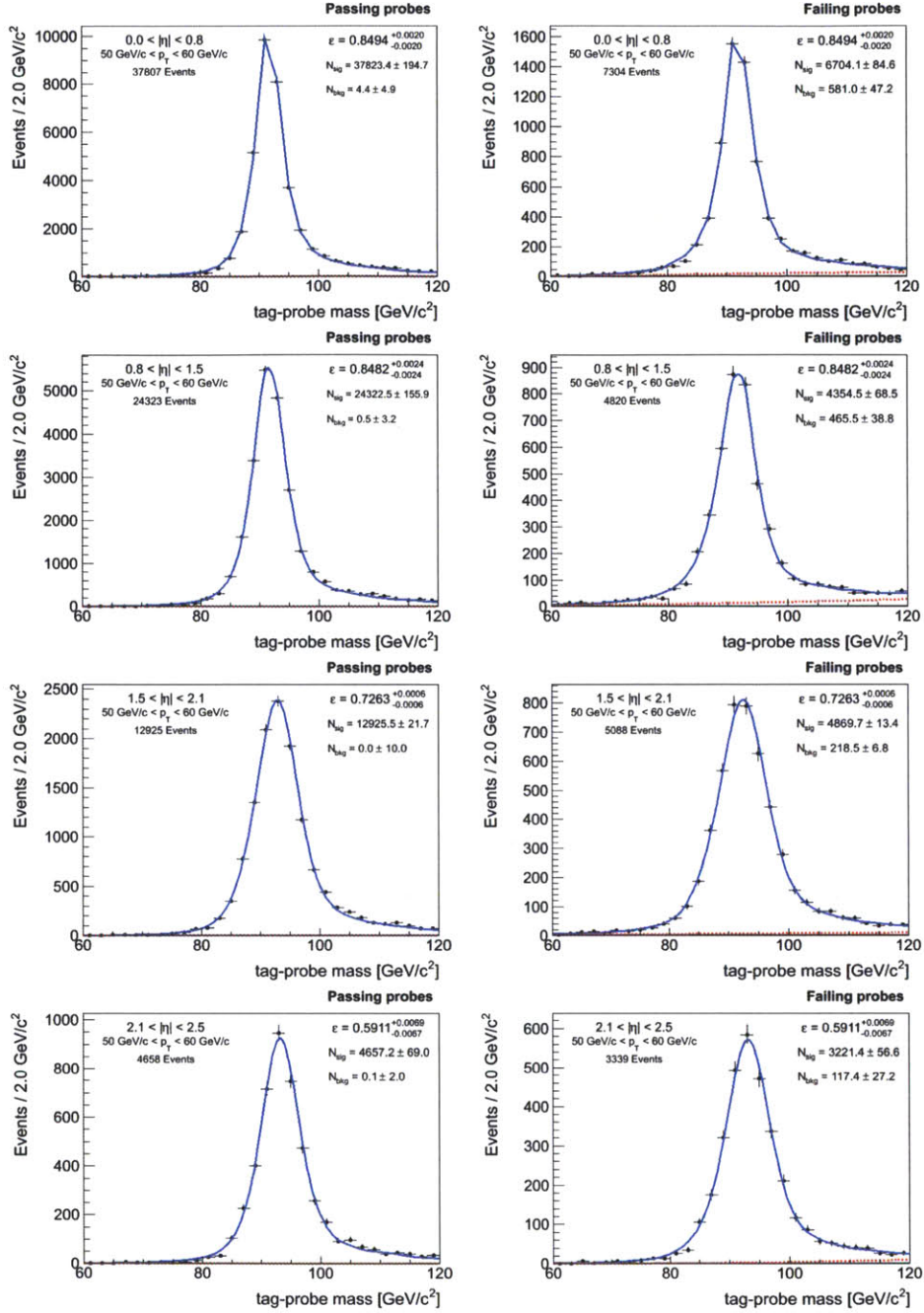


Figure B-19: Fits for electron selection efficiency for probes with  $50 \text{ GeV} < E_T < 60 \text{ GeV}$  in Run2011B data.

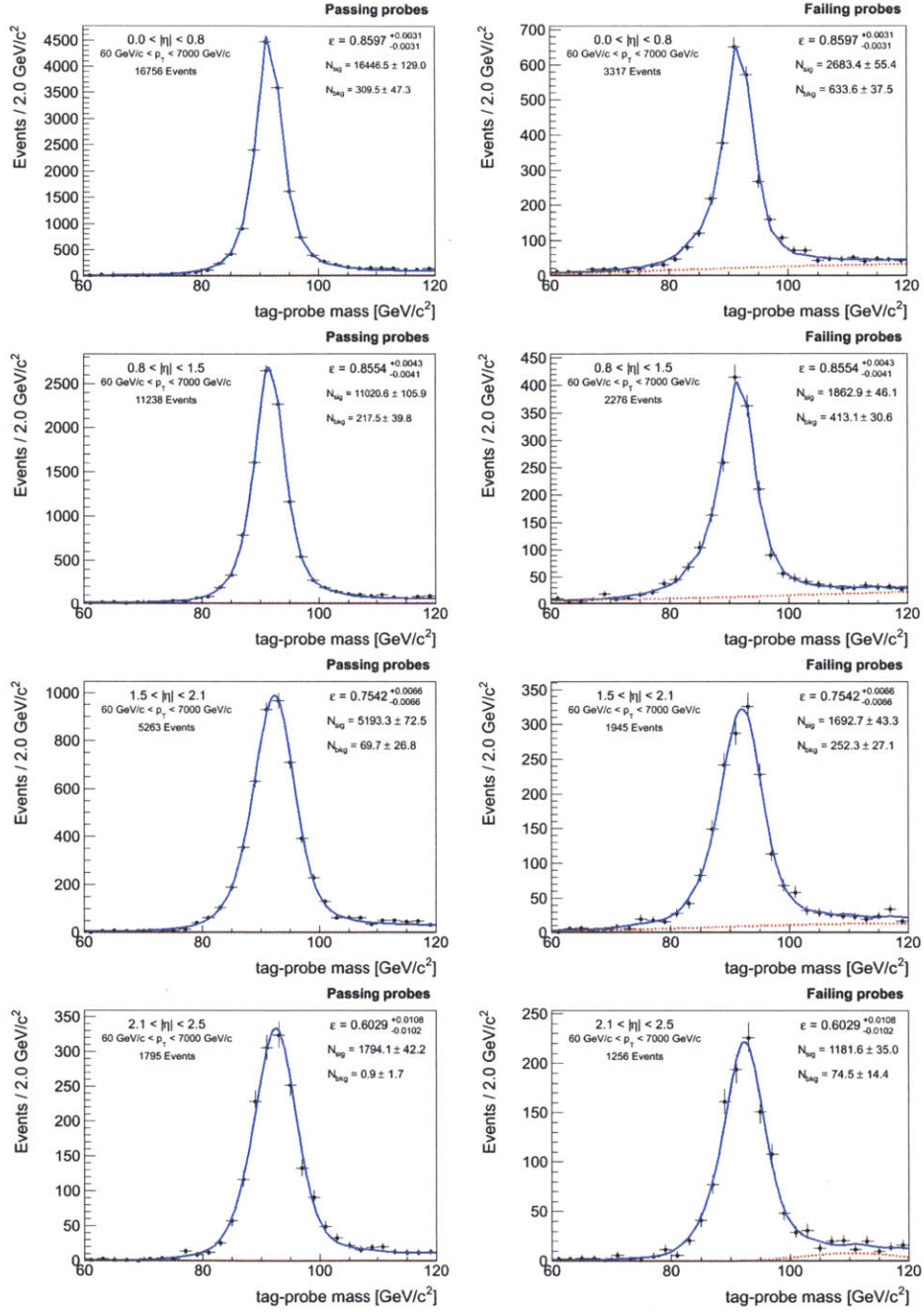


Figure B-20: Fits for electron selection efficiency for probes with  $E_T > 60$  GeV in Run2011B data.



## B.2 Muon Efficiencies

### B.2.1 Standalone Muon Efficiency

The standalone muon reconstruction efficiency is measured using isolated tracker track probes. The efficiencies in simulation and data (averaged over Run2011A and Run2011B) are listed in Tables B.7 and B.8 respectively, and the scale factors are listed in Table B.9. Fits for Run2011A are shown in Figures B-21 through B-25 and for Run2011B are shown in Figures B-26 through B-30.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9881 \pm 0.0005$	$0.9912 \pm 0.0007$	$0.9902 \pm 0.0005$	$0.9906 \pm 0.0010$
$30 < p_T < 40$	$0.9887 \pm 0.0003$	$0.9908 \pm 0.0005$	$0.9890 \pm 0.0004$	$0.9911 \pm 0.0007$
$40 < p_T < 50$	$0.9889 \pm 0.0003$	$0.9910 \pm 0.0004$	$0.9861 \pm 0.0004$	$0.9906 \pm 0.0007$
$50 < p_T < 60$	$0.9887 \pm 0.0007$	$0.9907 \pm 0.0010$	$0.9856 \pm 0.0008$	$0.9904 \pm 0.0014$
$p_T > 60$	$0.9888 \pm 0.0009$	$0.9900 \pm 0.0014$	$0.9855 \pm 0.0011$	$0.9899 \pm 0.0025$

Table B.7: Standalone muon reconstruction efficiency in simulation.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9854 \pm 0.0005$	$0.9890 \pm 0.0007$	$0.9731 \pm 0.0018$	$0.9801 \pm 0.0014$
$30 < p_T < 40$	$0.9855 \pm 0.0002$	$0.9863 \pm 0.0003$	$0.9709 \pm 0.0003$	$0.9787 \pm 0.0010$
$40 < p_T < 50$	$0.9856 \pm 0.0001$	$0.9863 \pm 0.0002$	$0.9660 \pm 0.0002$	$0.9791 \pm 0.0007$
$50 < p_T < 60$	$0.9853 \pm 0.0005$	$0.9872 \pm 0.0006$	$0.9652 \pm 0.0002$	$0.9790 \pm 0.0003$
$p_T > 60$	$0.9839 \pm 0.0013$	$0.9822 \pm 0.0020$	$0.9663 \pm 0.0007$	$0.9877 \pm 0.0100$

Table B.8: Standalone muon reconstruction efficiency in data averaged over Run2011A and Run2011B.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9973 \pm 0.0007$	$0.9978 \pm 0.0010$	$0.9828 \pm 0.0019$	$0.9895 \pm 0.0017$
$30 < p_T < 40$	$0.9968 \pm 0.0004$	$0.9955 \pm 0.0006$	$0.9817 \pm 0.0005$	$0.9875 \pm 0.0012$
$40 < p_T < 50$	$0.9967 \pm 0.0003$	$0.9952 \pm 0.0005$	$0.9796 \pm 0.0004$	$0.9884 \pm 0.0010$
$50 < p_T < 60$	$0.9966 \pm 0.0009$	$0.9965 \pm 0.0011$	$0.9793 \pm 0.0008$	$0.9885 \pm 0.0015$
$p_T > 60$	$0.9951 \pm 0.0016$	$0.9921 \pm 0.0025$	$0.9805 \pm 0.0013$	$0.9978 \pm 0.0105$

Table B.9: Data to simulation scale factors for standalone muon reconstruction efficiency.

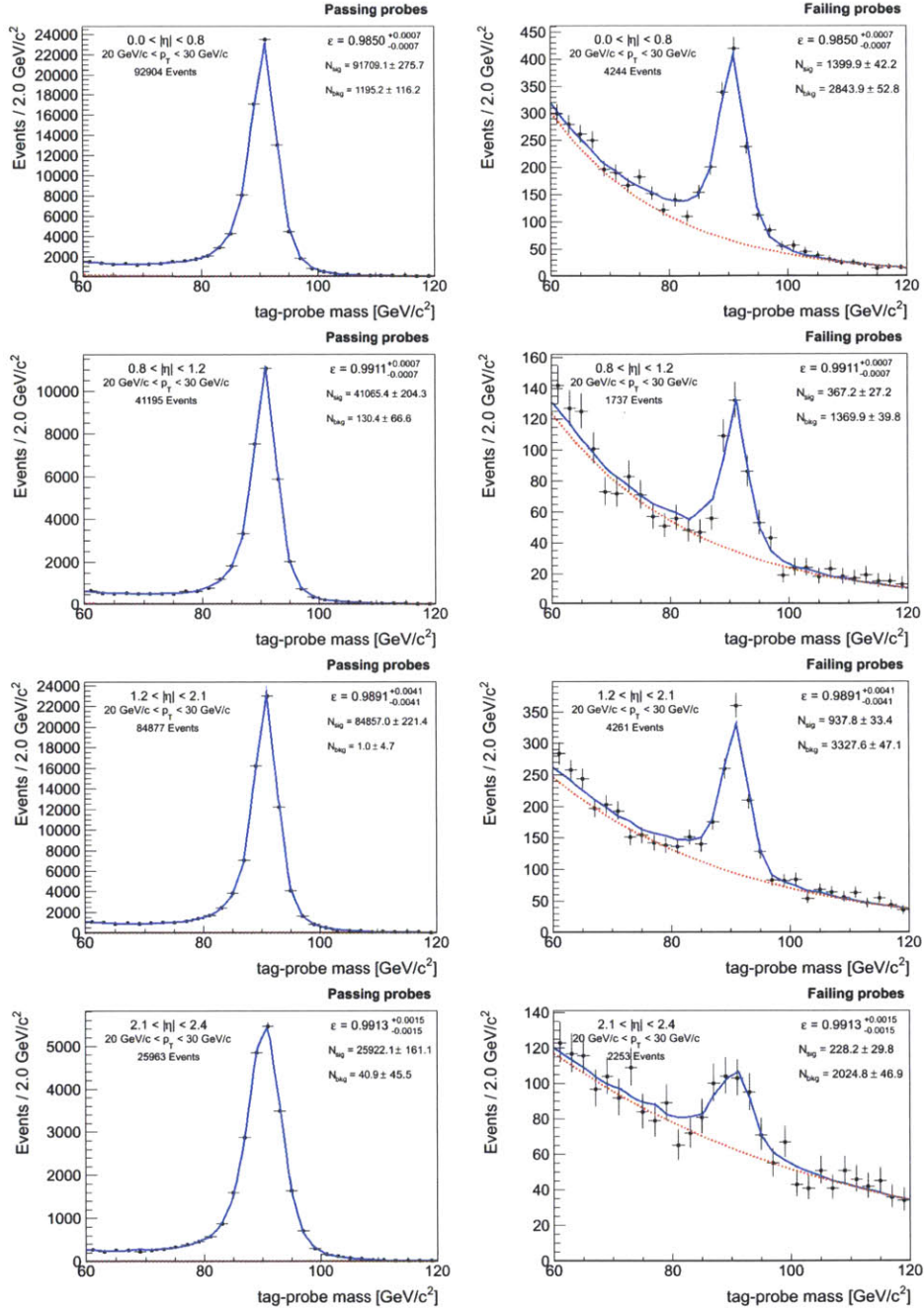


Figure B-21: Fits for standalone muon reconstruction efficiency for probes with  $20 \text{ GeV}/c < p_T < 30 \text{ GeV}/c$  in Run2011A data.

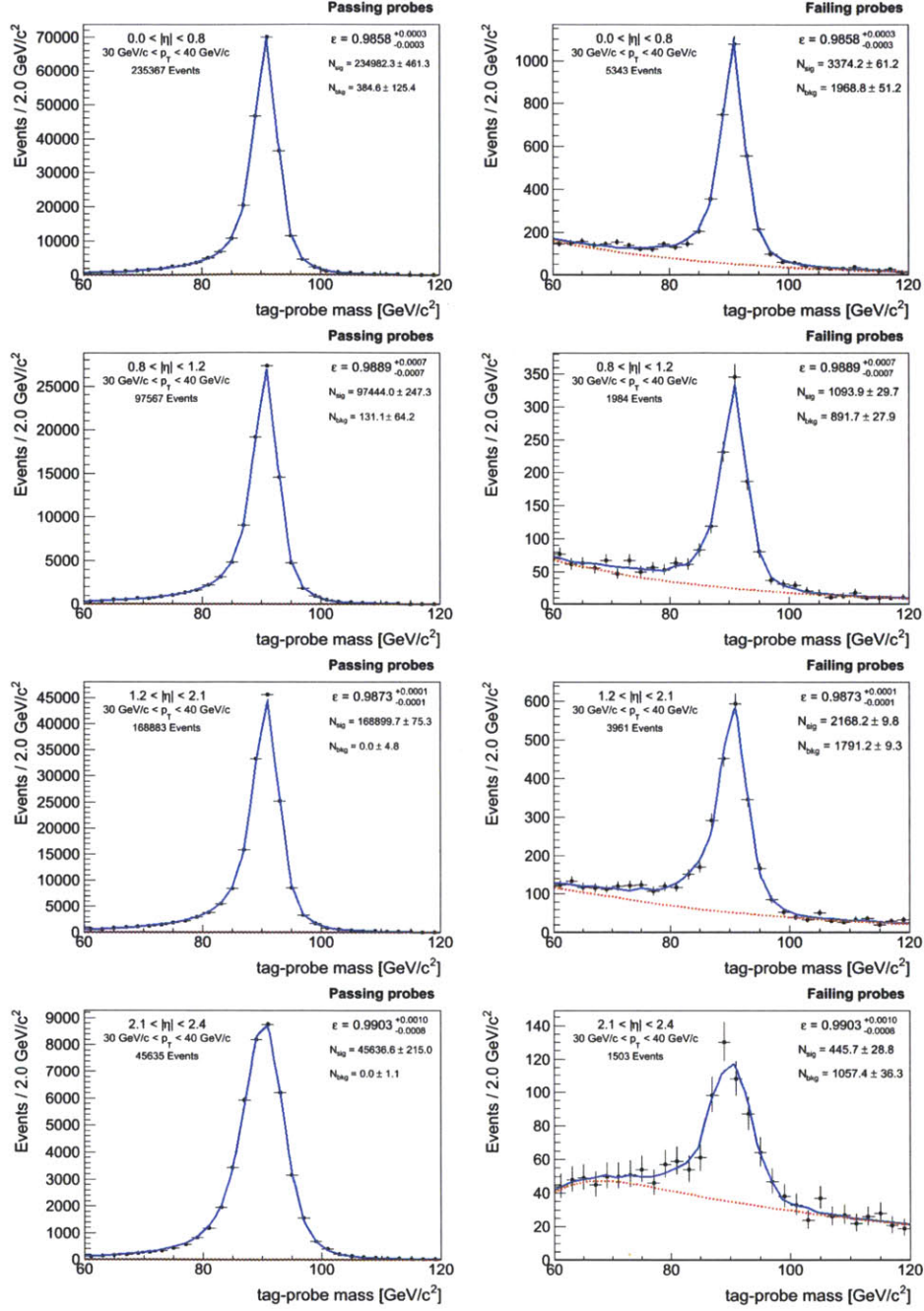


Figure B-22: Fits for standalone muon reconstruction efficiency for probes with  $30 \text{ GeV}/c < p_T < 40 \text{ GeV}/c$  in Run2011A data.

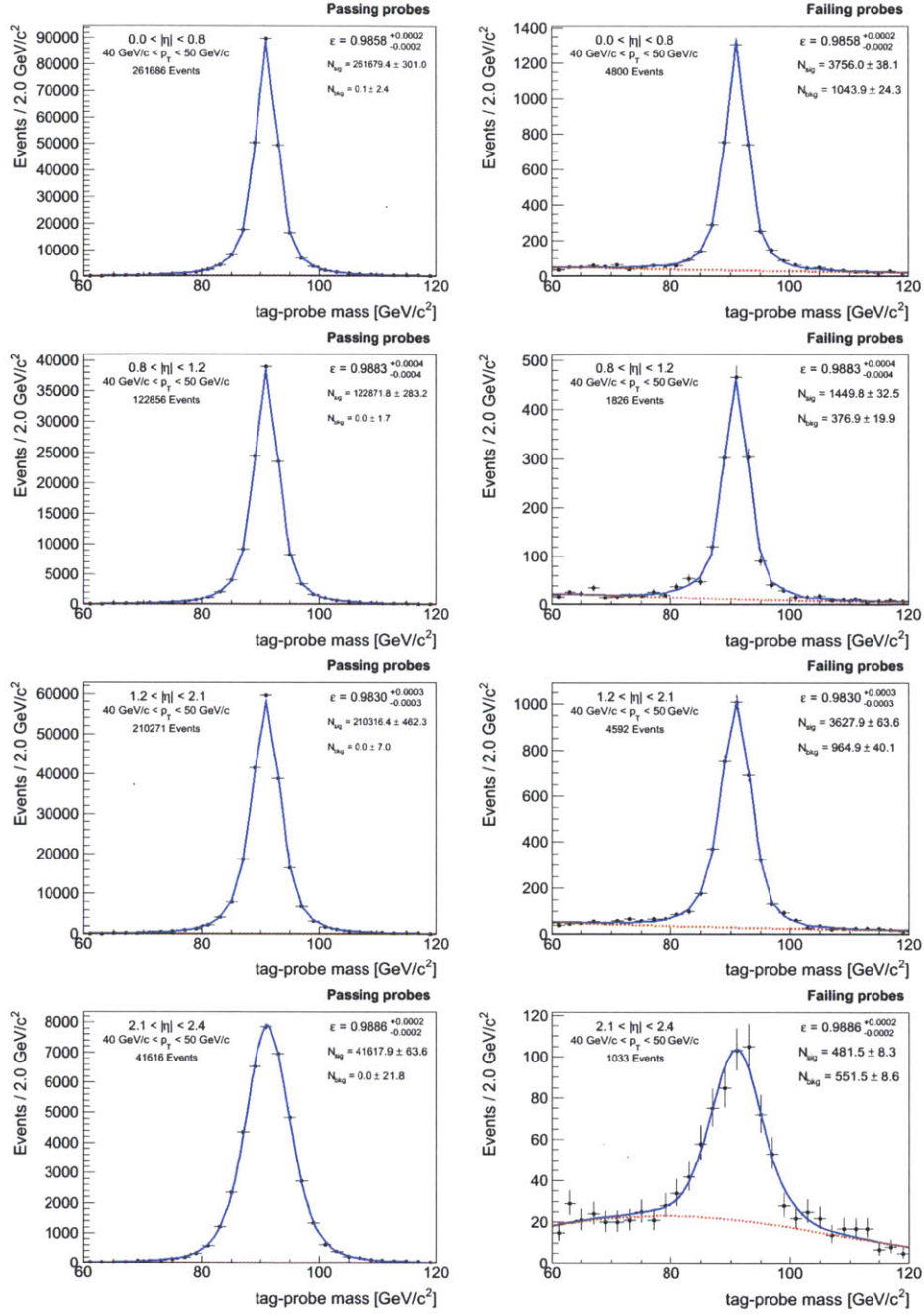


Figure B-23: Fits for standalone muon reconstruction efficiency for probes with  $40 \text{ GeV}/c < p_T < 50 \text{ GeV}/c$  in Run2011A data.

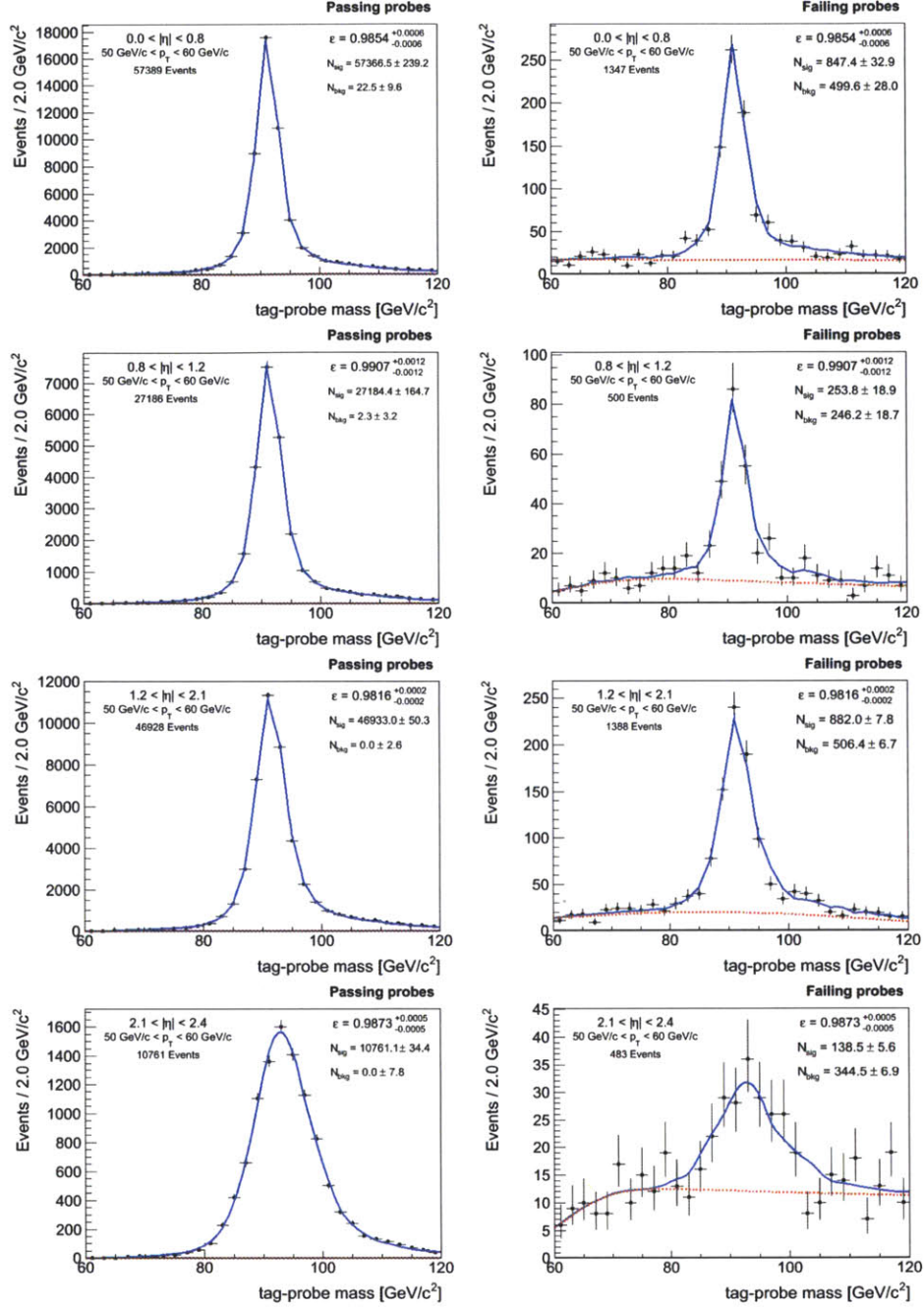


Figure B-24: Fits for standalone muon reconstruction efficiency for probes with  $50 \text{ GeV}/c < p_T < 60 \text{ GeV}/c$  in Run2011A data.



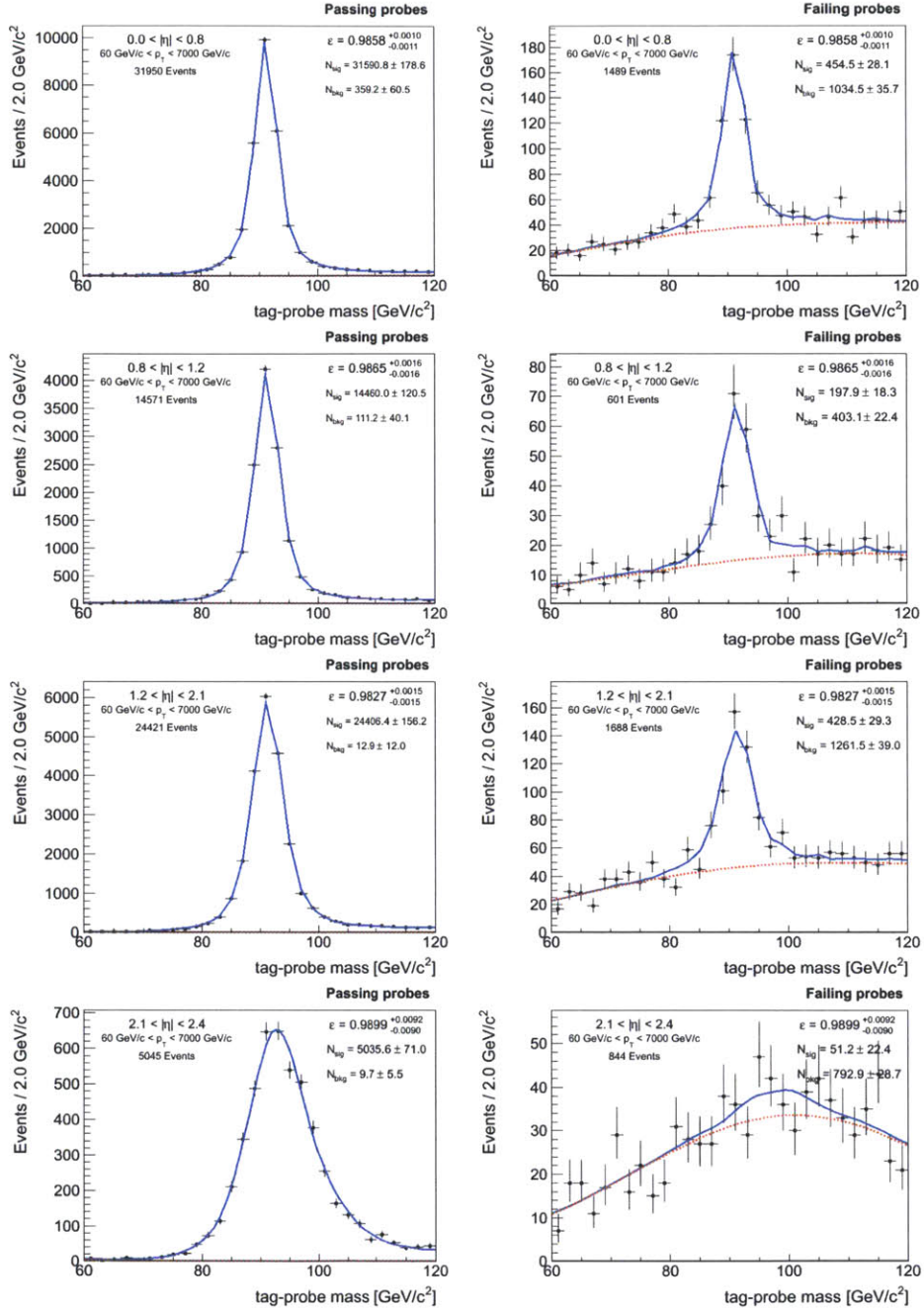


Figure B-25: Fits for standalone muon reconstruction efficiency for probes with  $p_T > 60 \text{ GeV}/c$  in Run2011A data.

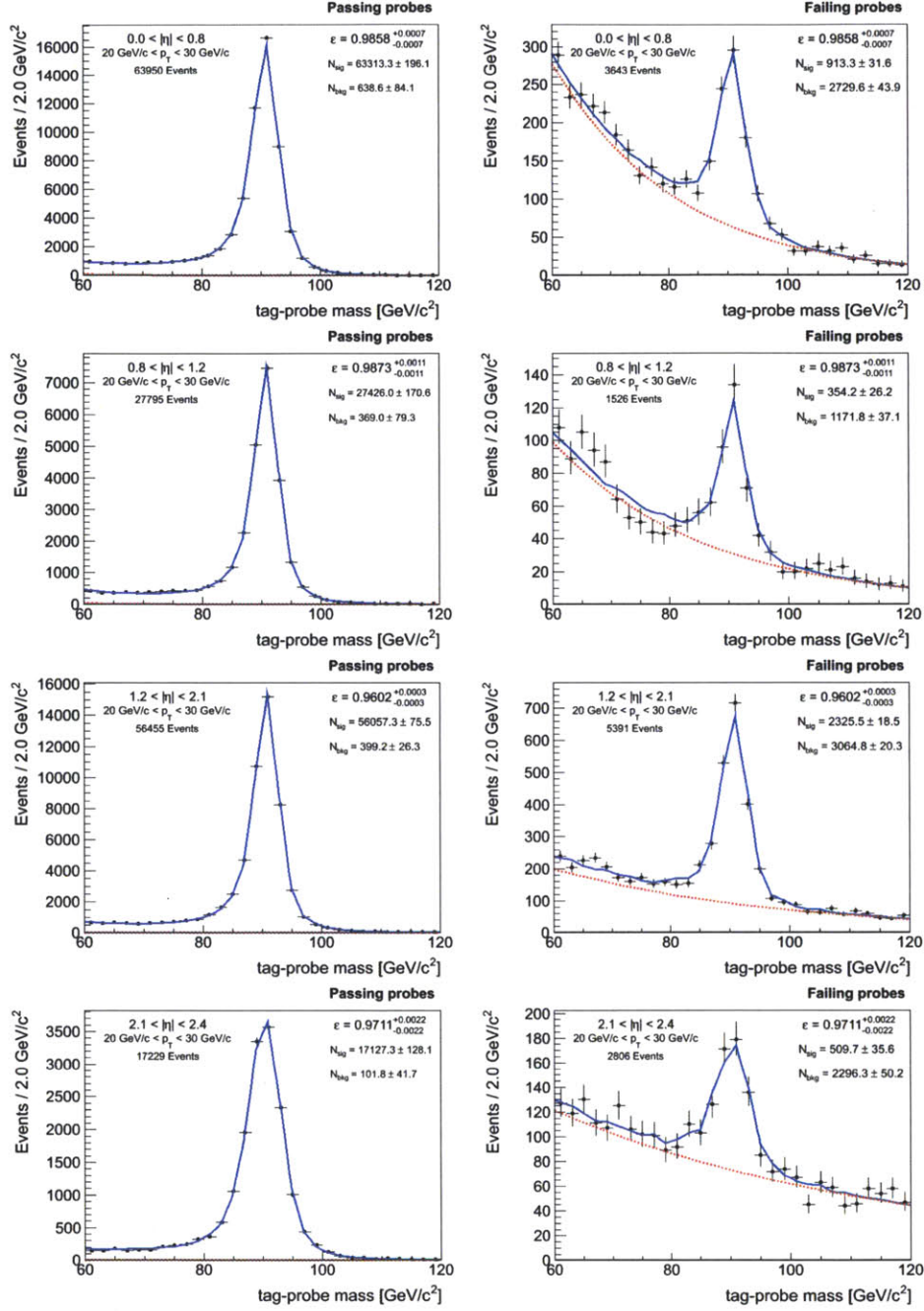


Figure B-26: Fits for standalone muon reconstruction efficiency for probes with  $20 \text{ GeV}/c < p_T < 30 \text{ GeV}/c$  in Run2011B data.



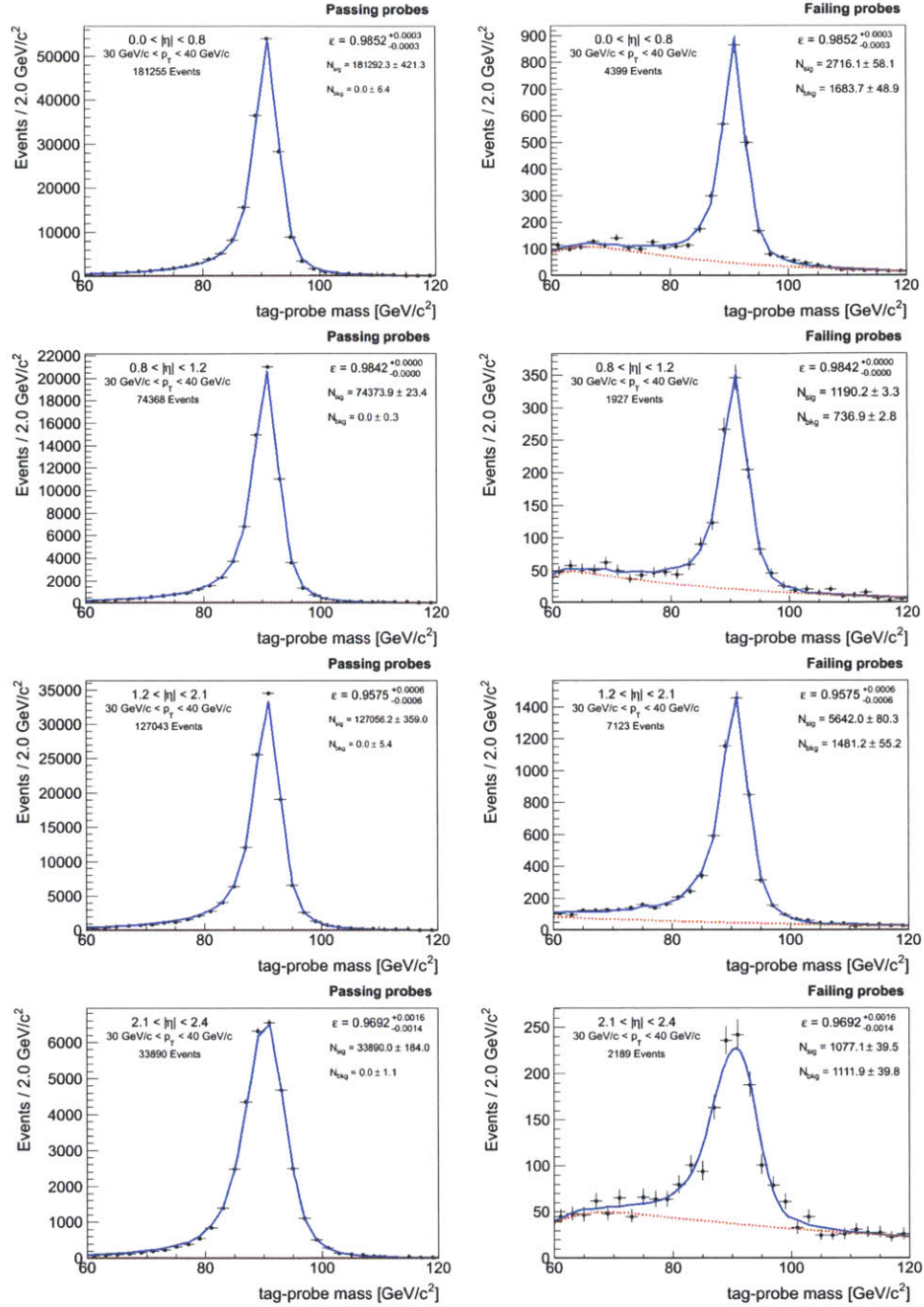


Figure B-27: Fits for standalone muon reconstruction efficiency for probes with  $30 \text{ GeV}/c < p_T < 40 \text{ GeV}/c$  in Run2011B data.

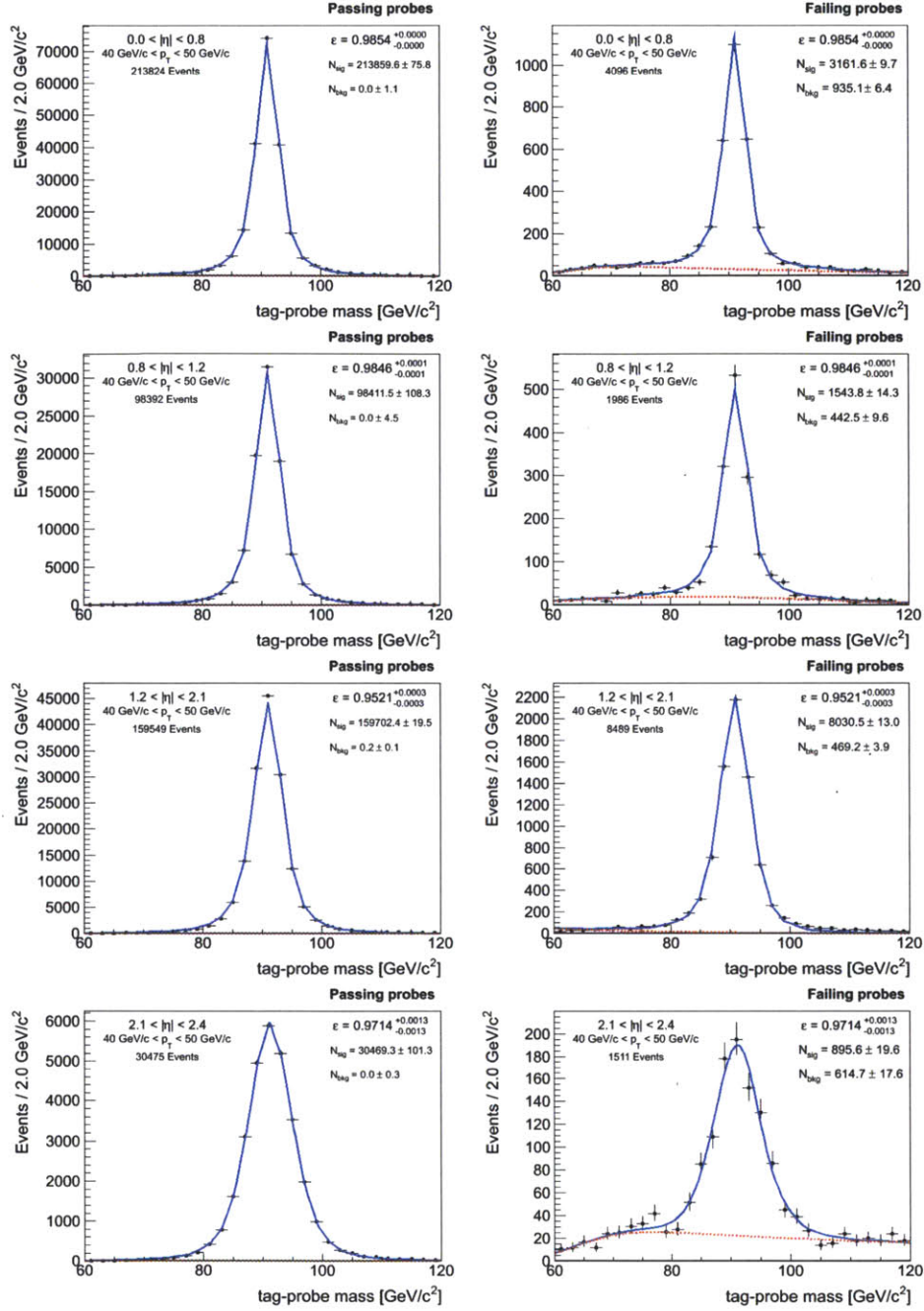


Figure B-28: Fits for standalone muon reconstruction efficiency for probes with  $40 \text{ GeV}/c < p_T < 50 \text{ GeV}/c$  in Run2011B data.

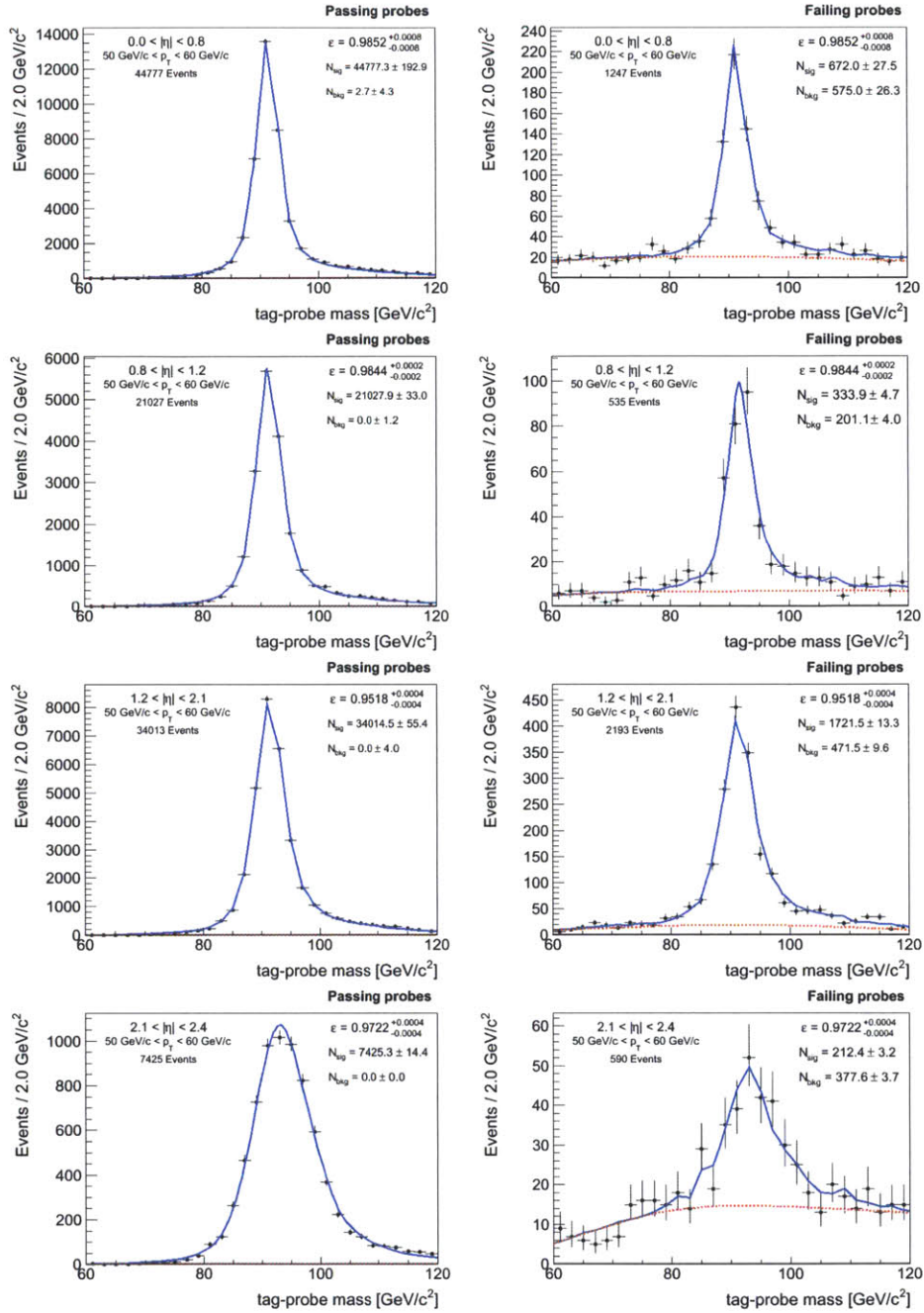


Figure B-29: Fits for standalone muon reconstruction efficiency for probes with  $50 \text{ GeV}/c < p_T < 60 \text{ GeV}/c$  in Run2011B data.

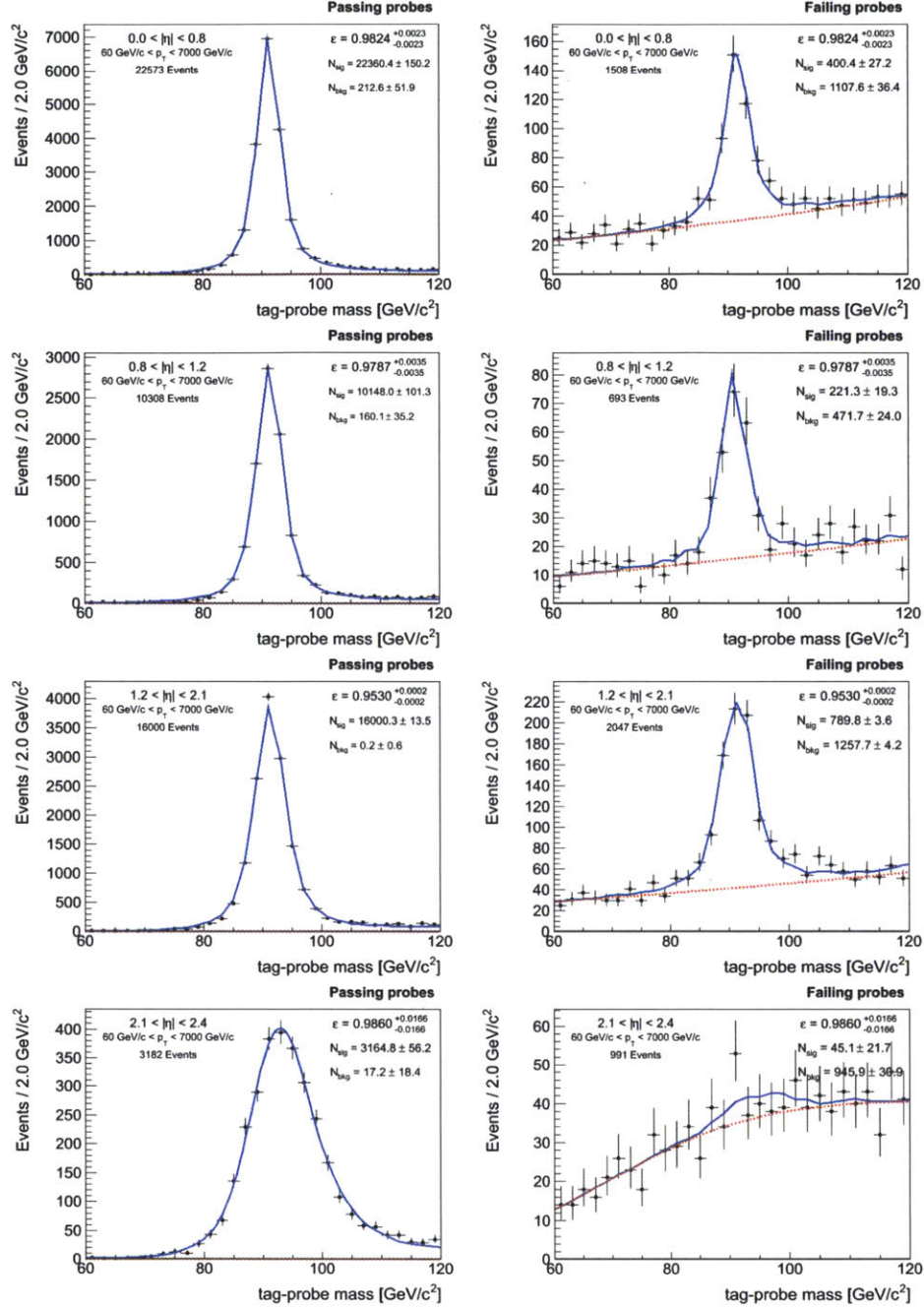


Figure B-30: Fits for standalone muon reconstruction efficiency for probes with  $p_T > 60 \text{ GeV}/c$  in Run2011B data.

## B.2.2 Muon Track Efficiency

The track reconstruction efficiency is measured using standalone muon probes. The efficiencies in simulation and data (averaged over Run2011A and Run2011B) are listed in Tables B.10 and B.11 respectively, and the scale factors are listed in Table B.12. Fits for Run2011A are shown in Figures B-31 and B-32, and fits for Run2011B are shown in Figures B-33 and B-34.

	$0 <  \eta  < 1.2$	$1.2 <  \eta  < 2.4$
$20 < p_T < 35$	$0.9980 \pm 0.0001$	$0.9986 \pm 0.0001$
$35 < p_T < 50$	$0.9989 \pm 0.0001$	$0.9992 \pm 0.0001$
$p_T > 50$	$0.9987 \pm 0.0002$	$0.9987 \pm 0.0002$

Table B.10: Track reconstruction efficiency in simulation.

	$0 <  \eta  < 1.2$	$1.2 <  \eta  < 2.4$
$20 < p_T < 35$	$0.9977 \pm 0.0002$	$0.9979 \pm 0.0002$
$35 < p_T < 50$	$0.9983 \pm 0.0001$	$0.9987 \pm 0.0000$
$p_T > 50$	$0.9977 \pm 0.0005$	$0.9977 \pm 0.0002$

Table B.11: Track reconstruction efficiency in data averaged over Run2011A and Run2011B.

	$0 <  \eta  < 1.2$	$1.2 <  \eta  < 2.4$
$20 < p_T < 35$	$0.9997 \pm 0.0003$	$0.9992 \pm 0.0003$
$35 < p_T < 50$	$0.9994 \pm 0.0001$	$0.9995 \pm 0.0001$
$p_T > 50$	$0.9990 \pm 0.0006$	$0.9990 \pm 0.0003$

Table B.12: Data to simulation scale factors for track reconstruction efficiency.



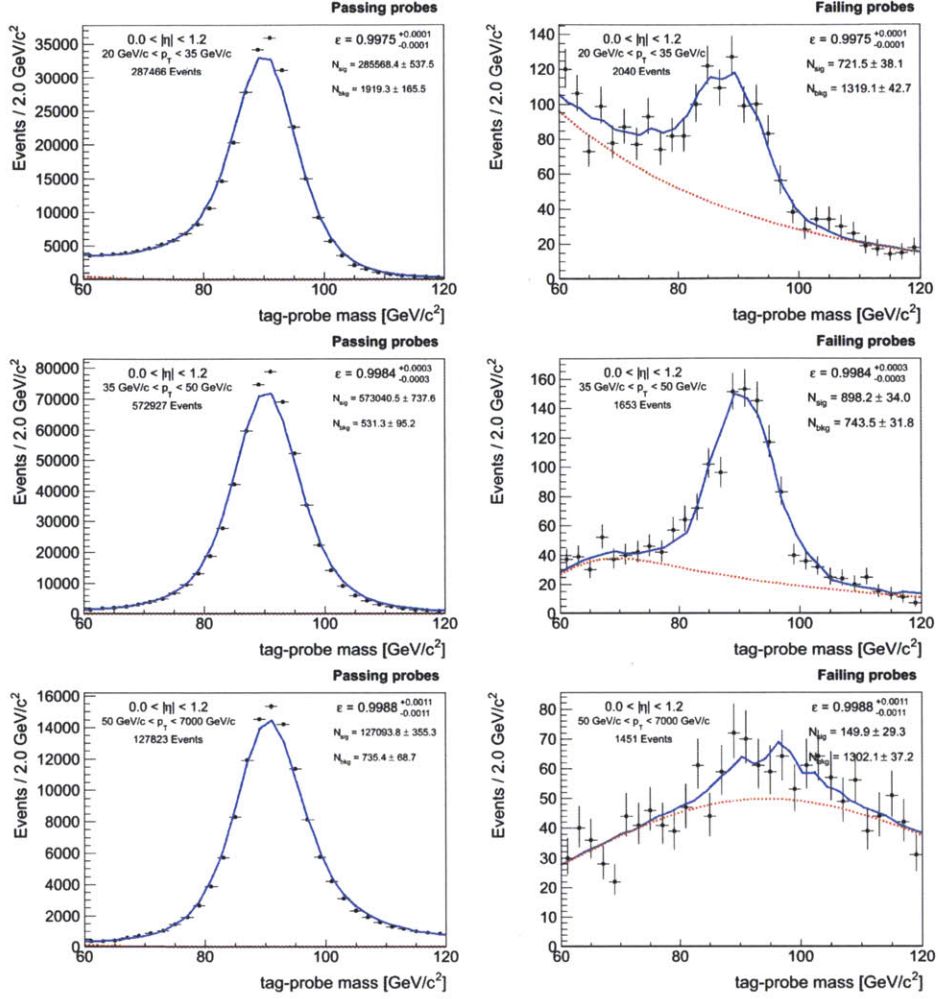


Figure B-31: Fits for track reconstruction efficiency for probes in the barrel ( $|\eta| < 1.2$ ) in Run2011A data.

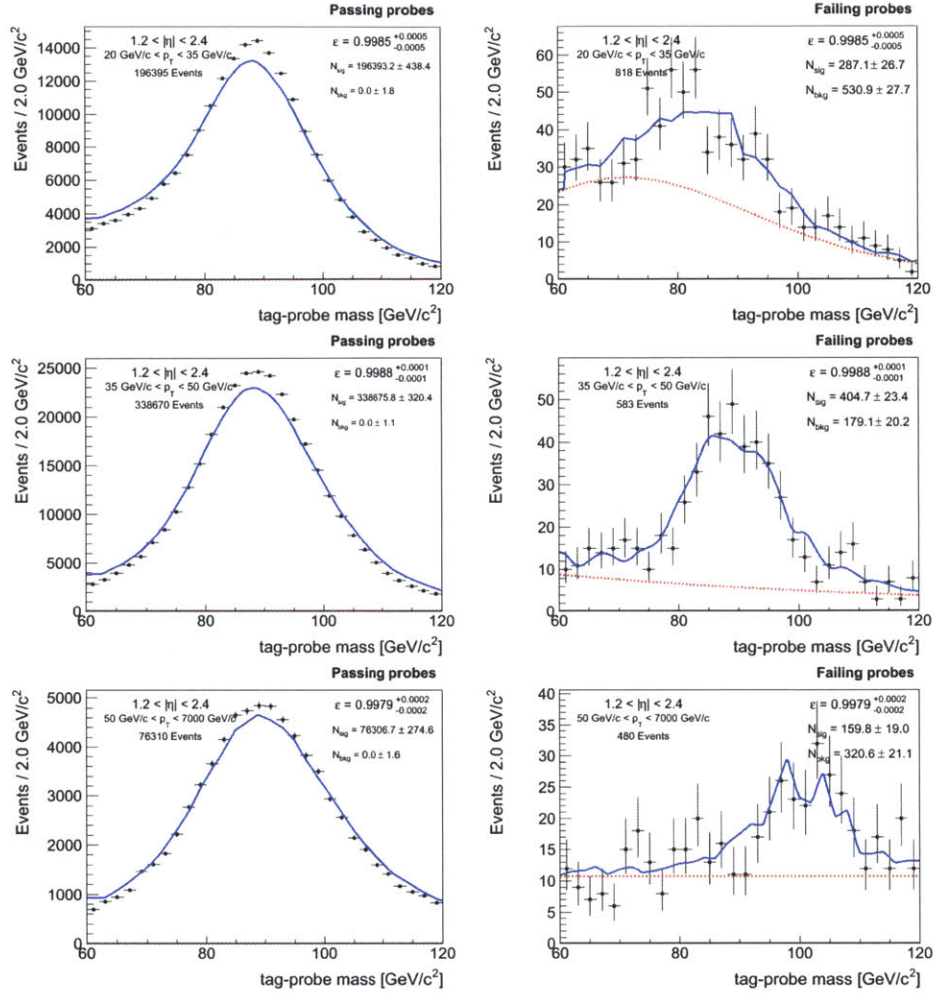


Figure B-32: Fits for track reconstruction efficiency for probes in the endcap ( $|\eta| > 1.2$ ) in Run2011A data.



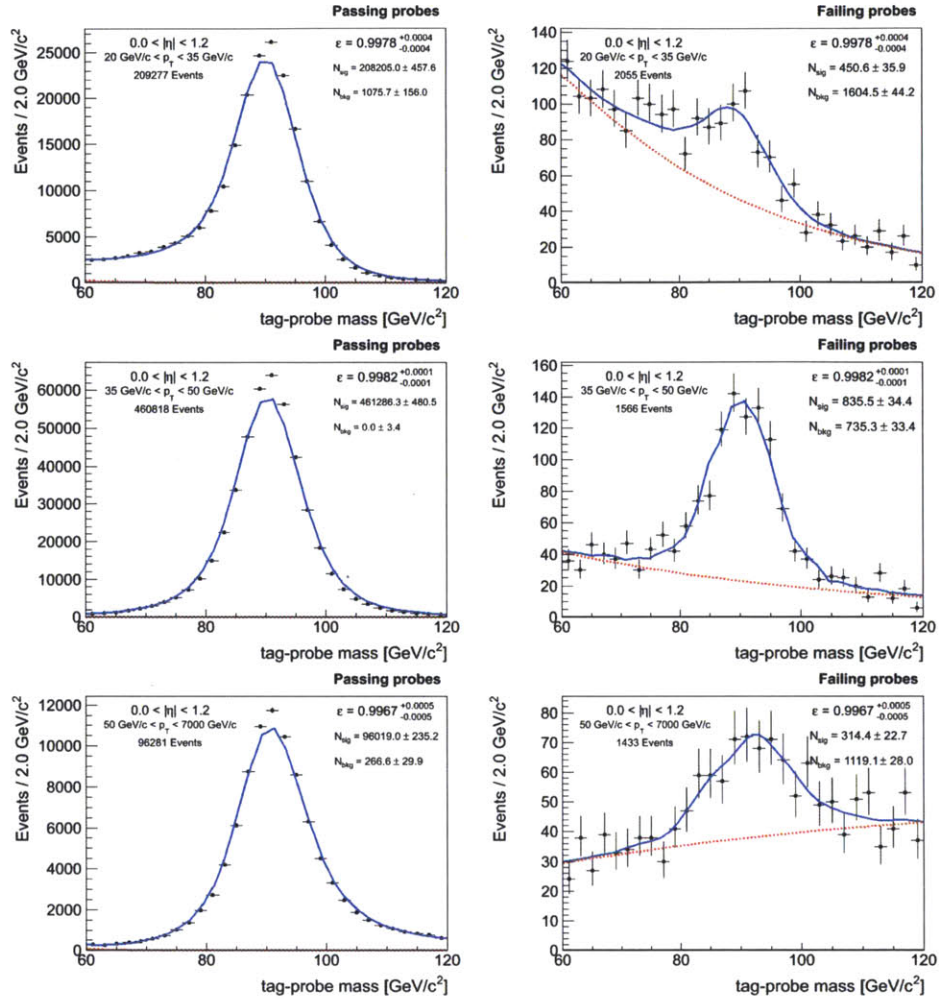


Figure B-33: Fits for track reconstruction efficiency for probes in the barrel ( $|\eta| < 1.2$ ) in Run2011B data.

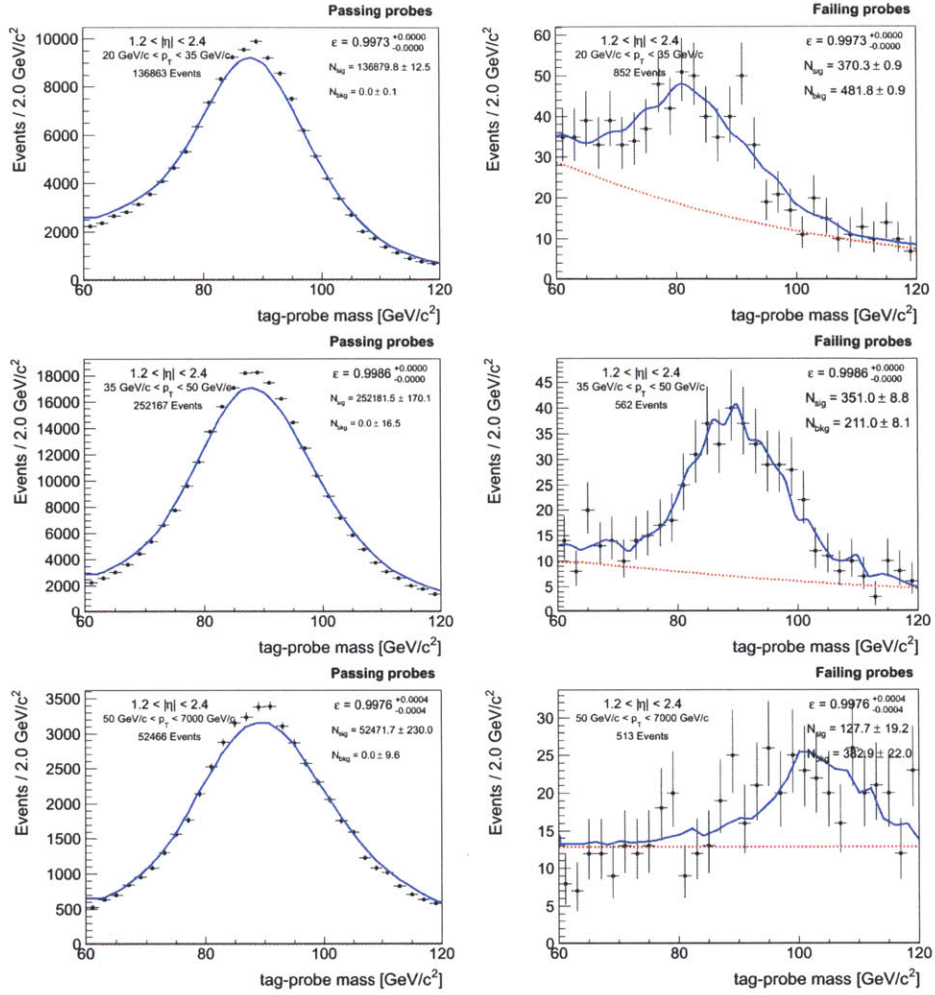


Figure B-34: Fits for track reconstruction efficiency for probes in the endcap ( $|\eta| > 1.2$ ) in Run2011B data.

### B.2.3 Muon Selection Efficiency

The electron selection efficiency is measured using GSF electron probes. The efficiencies in simulation and data (averaged over Run2011A and Run2011B) are listed in Tables B.13 and B.14 respectively, and the scale factors are listed in Table B.15. Fits for Run2011A are shown in Figures B-35 through B-39 and for Run2011B are shown in Figures B-40 through B-44.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9279 \pm 0.0011$	$0.9273 \pm 0.0018$	$0.9391 \pm 0.0012$	$0.9242 \pm 0.0023$
$30 < p_T < 40$	$0.9593 \pm 0.0006$	$0.9555 \pm 0.0010$	$0.9602 \pm 0.0007$	$0.9358 \pm 0.0017$
$40 < p_T < 50$	$0.9723 \pm 0.0005$	$0.9686 \pm 0.0007$	$0.9745 \pm 0.0005$	$0.9294 \pm 0.0018$
$50 < p_T < 60$	$0.9729 \pm 0.0010$	$0.9688 \pm 0.0016$	$0.9741 \pm 0.0011$	$0.8846 \pm 0.0045$
$p_T > 60$	$0.9714 \pm 0.0014$	$0.9662 \pm 0.0023$	$0.9729 \pm 0.0016$	$0.7447 \pm 0.0090$

Table B.13: Muon selection efficiency in simulation.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9186 \pm 0.0010$	$0.9118 \pm 0.0016$	$0.9260 \pm 0.0010$	$0.9093 \pm 0.0018$
$30 < p_T < 40$	$0.9558 \pm 0.0004$	$0.9446 \pm 0.0006$	$0.9508 \pm 0.0006$	$0.9210 \pm 0.0009$
$40 < p_T < 50$	$0.9694 \pm 0.0002$	$0.9583 \pm 0.0003$	$0.9663 \pm 0.0005$	$0.9129 \pm 0.0006$
$50 < p_T < 60$	$0.9682 \pm 0.0006$	$0.9572 \pm 0.0010$	$0.9655 \pm 0.0006$	$0.8670 \pm 0.0028$
$p_T > 60$	$0.9673 \pm 0.0008$	$0.9536 \pm 0.0015$	$0.9598 \pm 0.0010$	$0.7172 \pm 0.0029$

Table B.14: Muon selection efficiency in data averaged over Run2011A and Run2011B.

	$0 <  \eta  < 0.8$	$0.8 <  \eta  < 1.2$	$1.2 <  \eta  < 2.1$	$2.1 <  \eta  < 2.4$
$20 < p_T < 30$	$0.9900 \pm 0.0016$	$0.9833 \pm 0.0026$	$0.9861 \pm 0.0016$	$0.9839 \pm 0.0031$
$30 < p_T < 40$	$0.9963 \pm 0.0007$	$0.9886 \pm 0.0012$	$0.9902 \pm 0.0010$	$0.9842 \pm 0.0020$
$40 < p_T < 50$	$0.9970 \pm 0.0005$	$0.9893 \pm 0.0008$	$0.9916 \pm 0.0007$	$0.9823 \pm 0.0020$
$50 < p_T < 60$	$0.9952 \pm 0.0012$	$0.9879 \pm 0.0019$	$0.9911 \pm 0.0013$	$0.9801 \pm 0.0059$
$p_T > 60$	$0.9958 \pm 0.0017$	$0.9869 \pm 0.0029$	$0.9866 \pm 0.0019$	$0.9631 \pm 0.0123$

Table B.15: Data to simulation scale factors for muon selection efficiency.

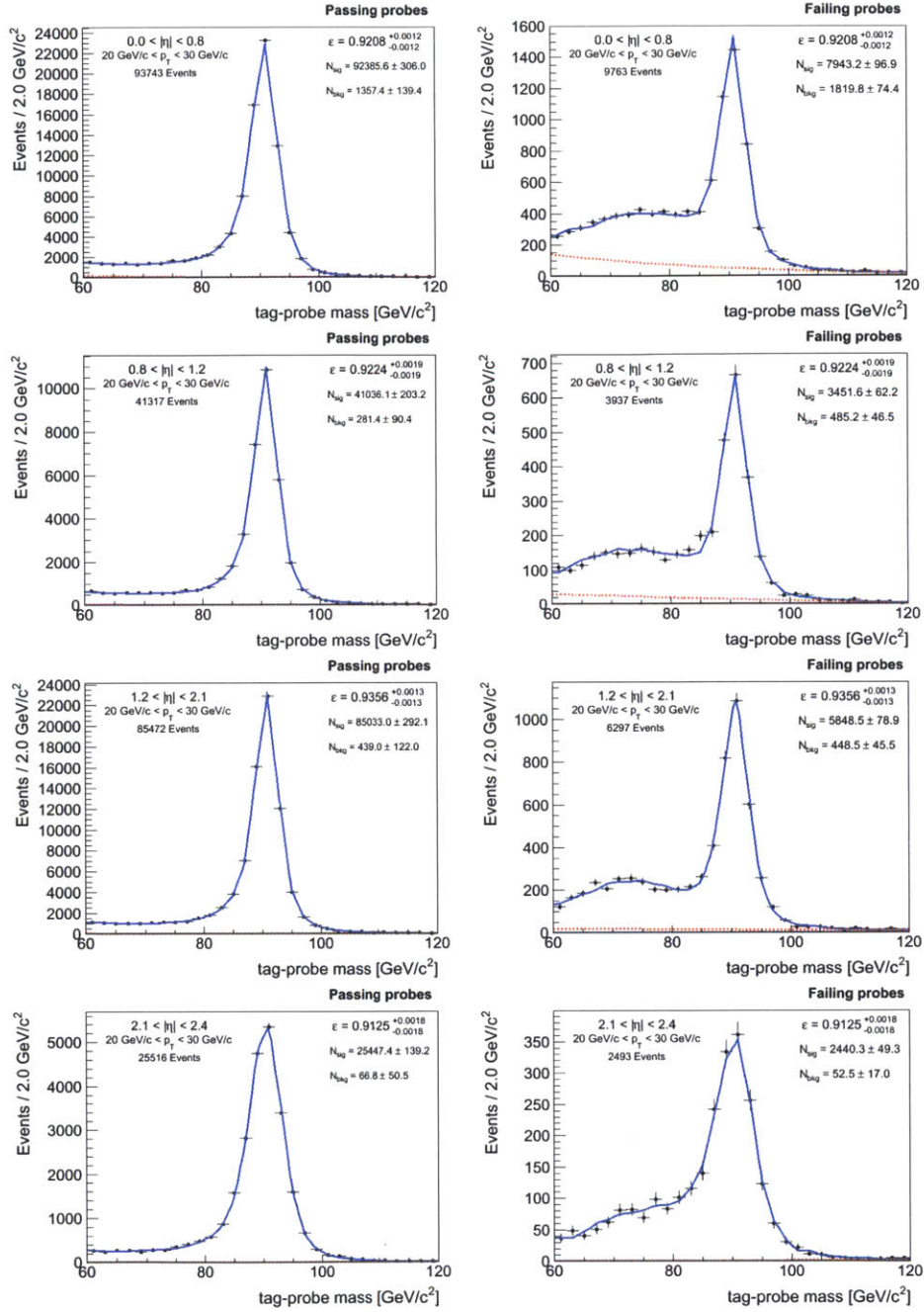


Figure B-35: Fits for muon selection efficiency for probes with  $20 \text{ GeV}/c < p_T < 30 \text{ GeV}/c$  in Run2011A data.

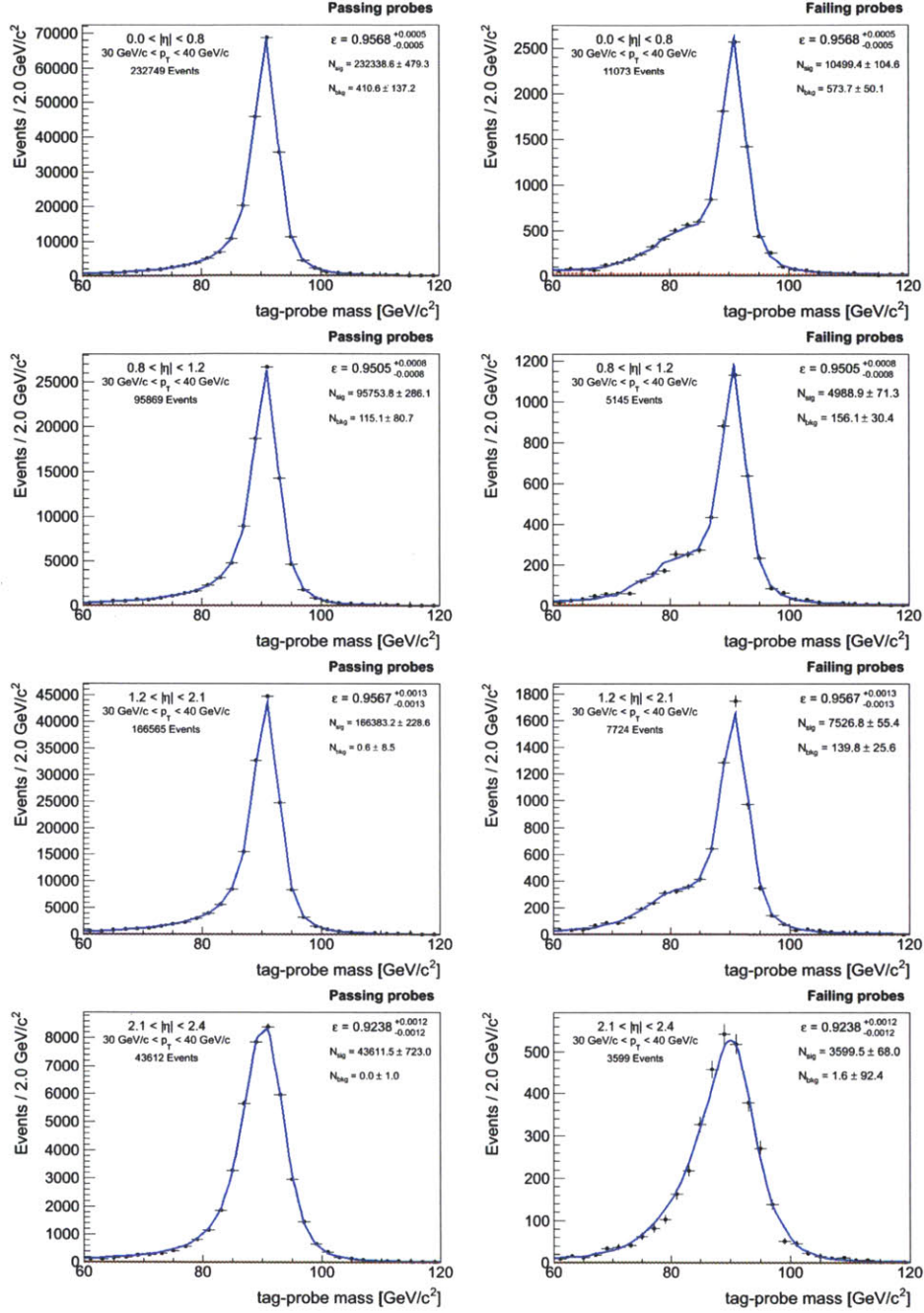


Figure B-36: Fits for muon selection efficiency for probes with  $30 \text{ GeV}/c < p_T < 40 \text{ GeV}/c$  in Run2011A data.

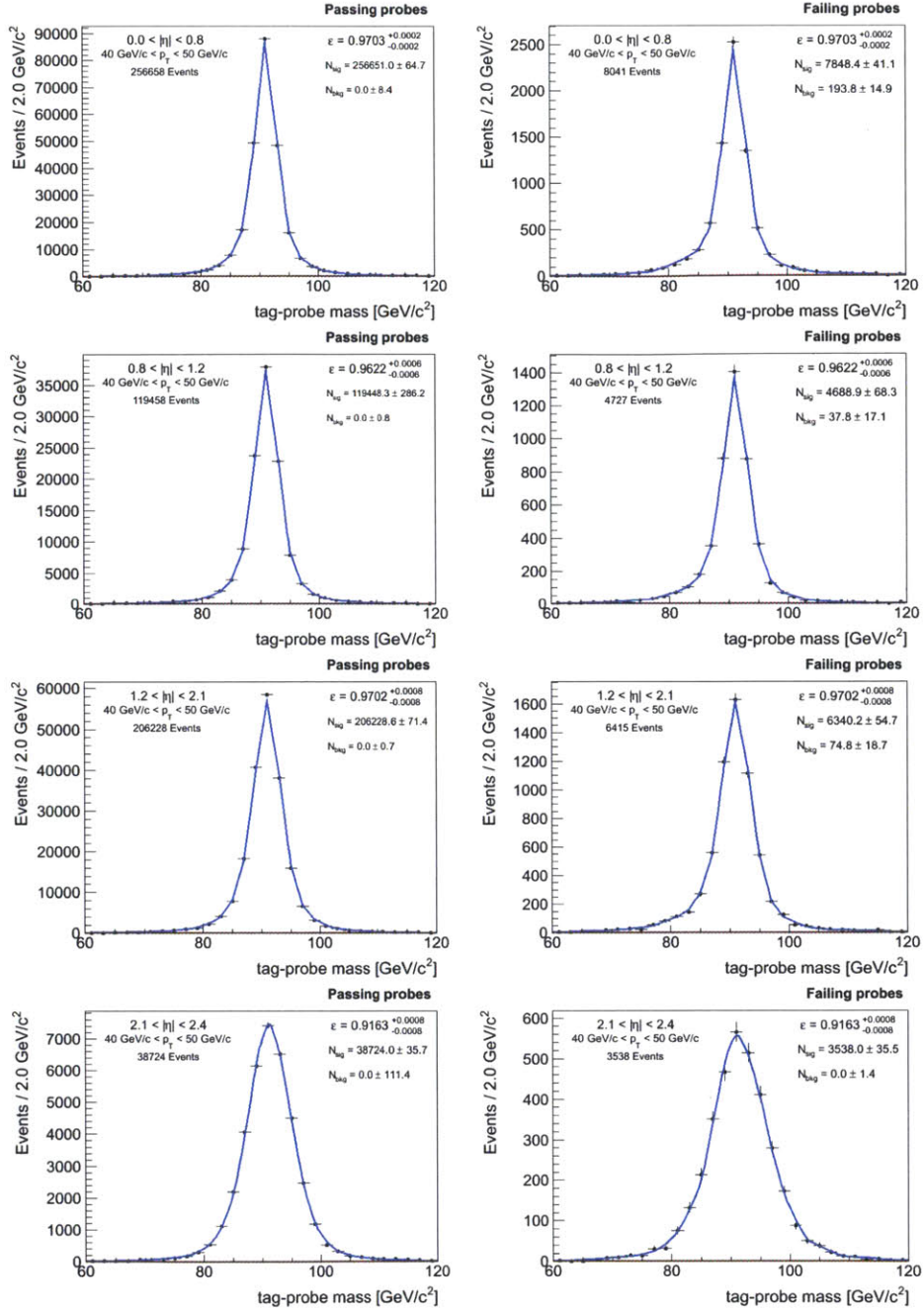


Figure B-37: Fits for muon selection efficiency for probes with  $40 \text{ GeV}/c < p_T < 50 \text{ GeV}/c$  in Run2011A data.



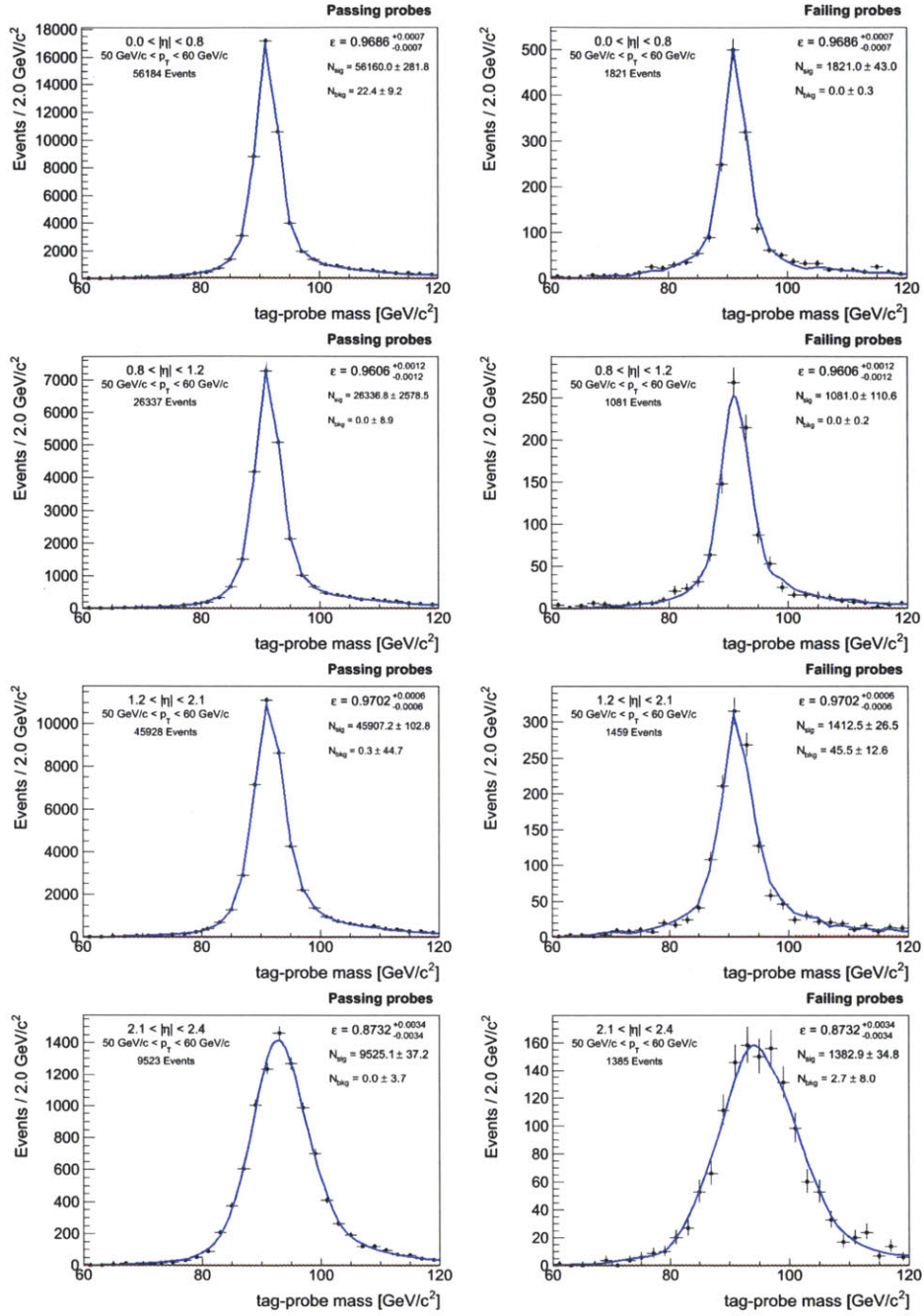


Figure B-38: Fits for muon selection efficiency for probes with  $50 \text{ GeV}/c < p_T < 60 \text{ GeV}/c$  in Run2011A data.



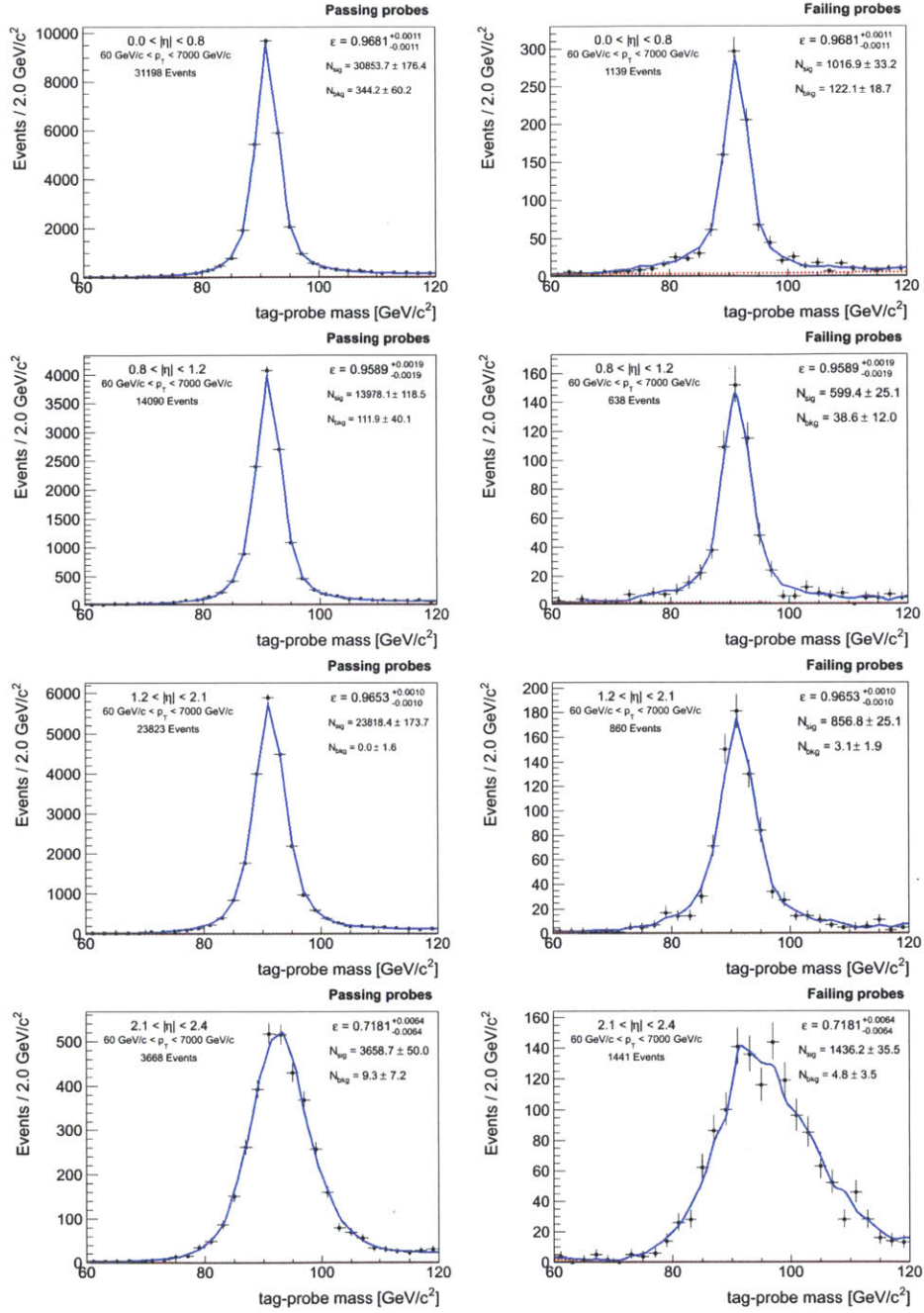


Figure B-39: Fits for muon selection efficiency for probes with  $p_T > 60$  GeV/c in Run2011A data.

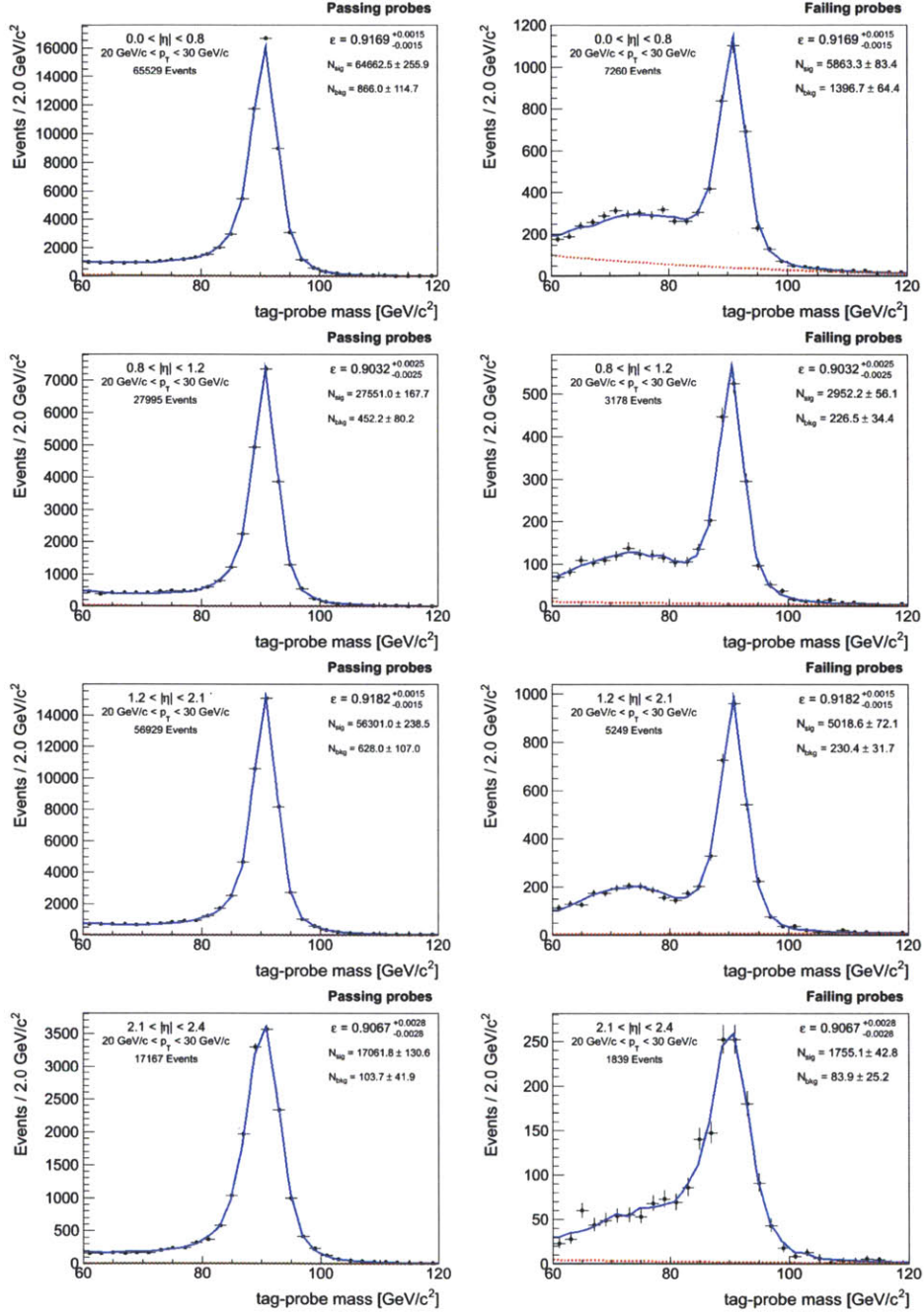


Figure B-40: Fits for muon selection efficiency for probes with  $20 \text{ GeV}/c < p_T < 30 \text{ GeV}/c$  in Run2011B data.

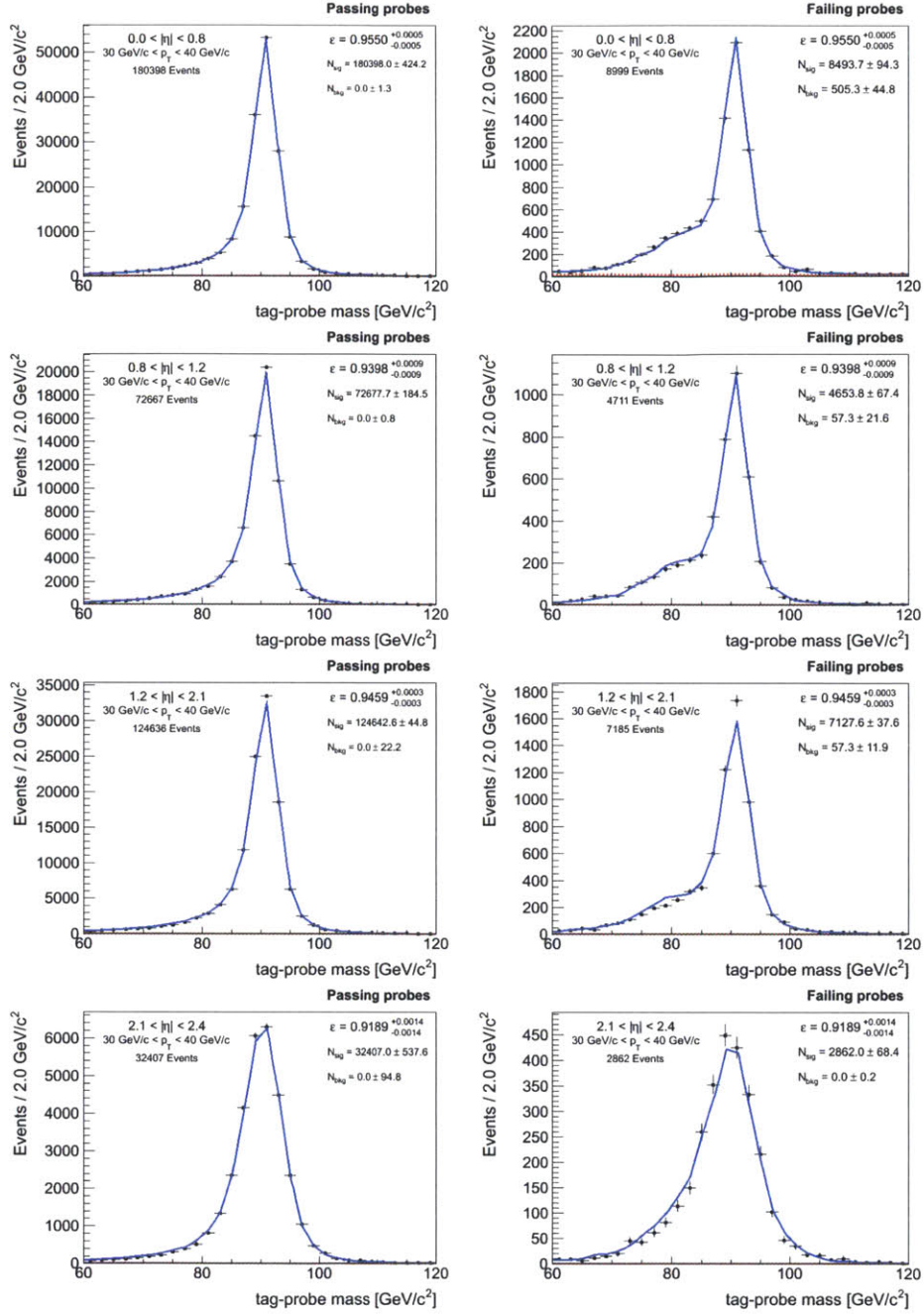


Figure B-41: Fits for muon selection efficiency for probes with  $30 \text{ GeV}/c < p_T < 40 \text{ GeV}/c$  in Run2011B data.

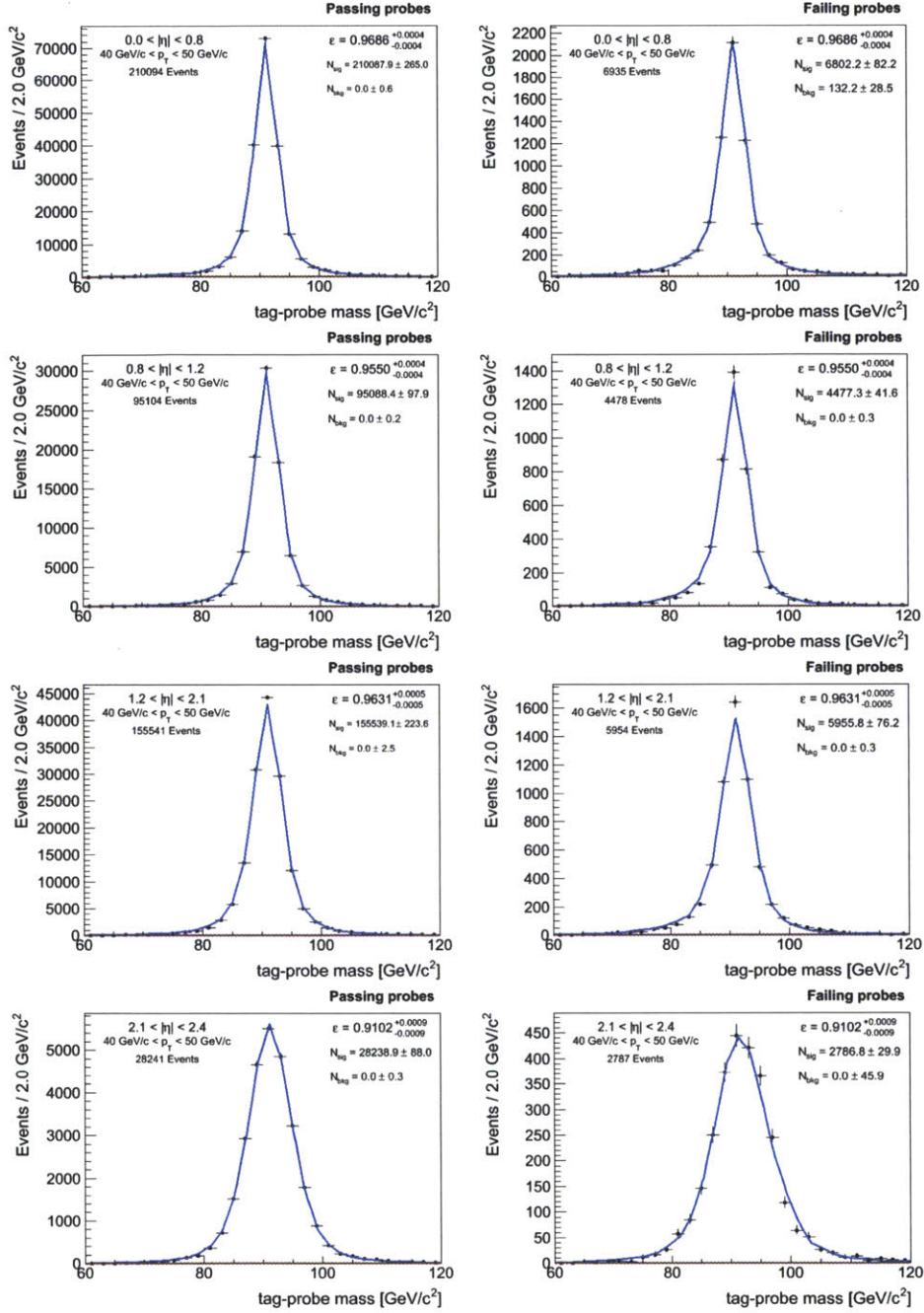


Figure B-42: Fits for muon selection efficiency for probes with  $40 \text{ GeV}/c < p_T < 50 \text{ GeV}/c$  in Run2011B data.

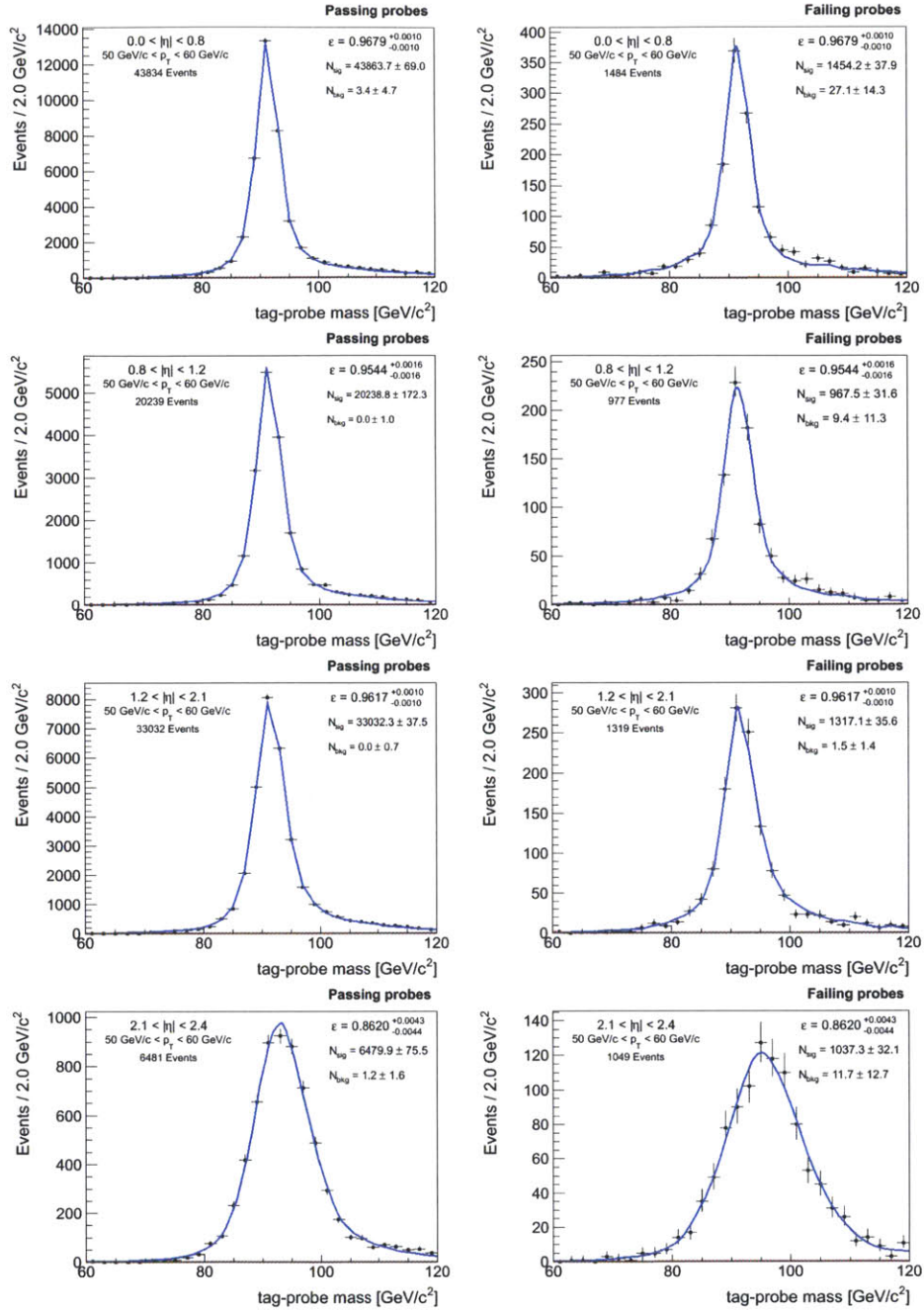


Figure B-43: Fits for muon selection efficiency for probes with  $50 \text{ GeV}/c < p_T < 60 \text{ GeV}/c$  in Run2011B data.



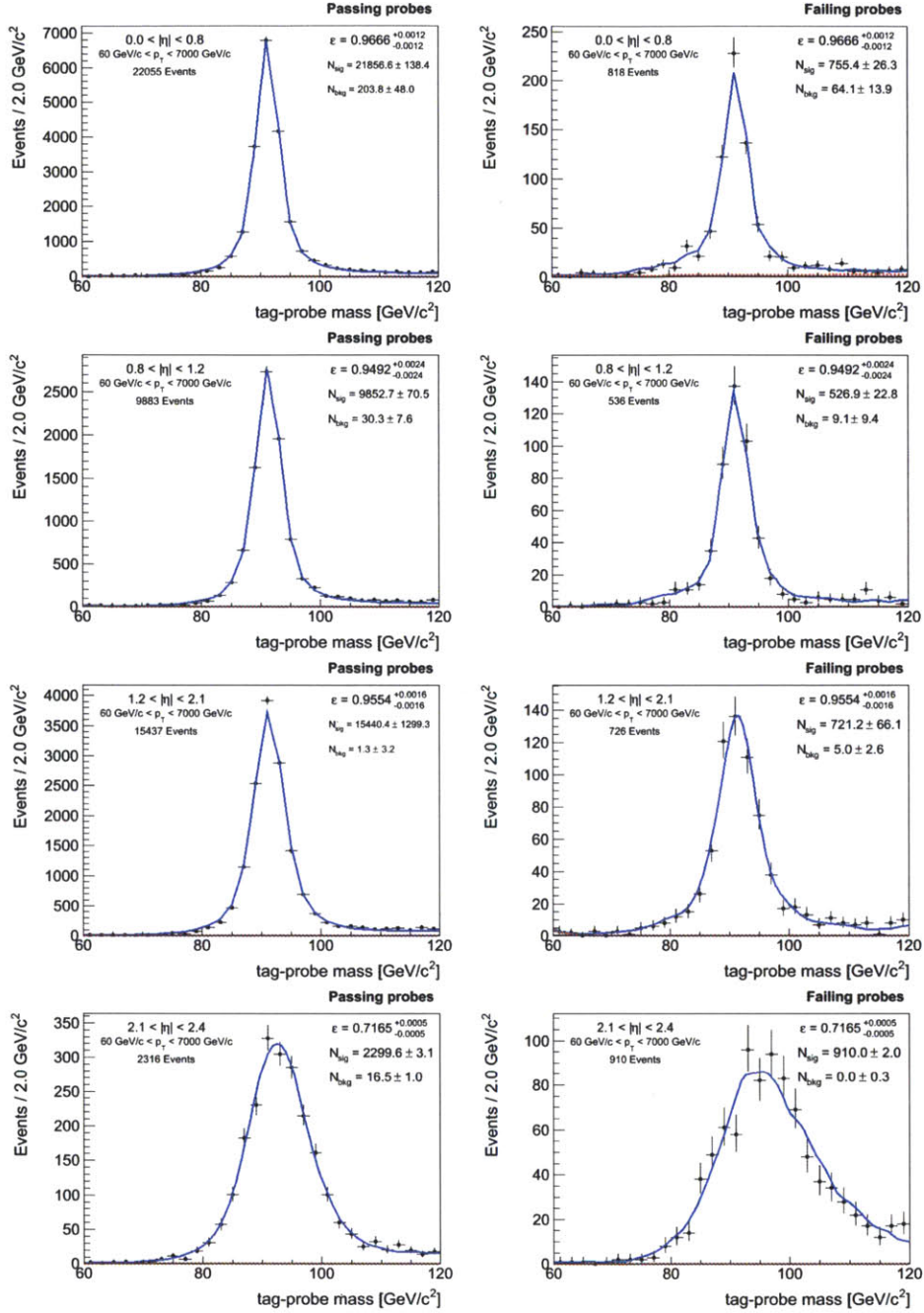


Figure B-44: Fits for muon selection efficiency for probes with  $p_T > 60$  GeV/c in Run2011B data.

# Bibliography

- [1] F. Halzen and A. D. Martin. *Quarks and Leptons*. John Wiley & Sons, New York, USA, 1985.
- [2] D. Griffiths. *Introduction to Elementary Particles*. John Wiley & Sons, New York, USA, 1987.
- [3] S. L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579 – 588, 1961.
- [4] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.
- [5] A. Salam. Weak and Electromagnetic Interactions. In N. Svartholm, editor, *Elementary particle theory*, pages 367–377. Almquist & Wiksell, 1968.
- [6] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, 1964.
- [7] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.
- [8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, 1964.
- [9] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *CERN-2011-002*, CERN, Geneva, 2011.
- [10] Y. A. Golfand and E. P. Likhtman. Extension of the Algebra of Poincare Group Generators and Violation of P Invariance. *JETP Lett.*, 13:323–326, 1971.
- [11] J. Wess and B. Zumino. Supergauge transformations in four dimensions. *Nuclear Physics B*, 70(1):39 – 50, 1974.
- [12] P. Fayet. Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino. *Nuclear Physics B*, 90(0):104 – 124, 1975.
- [13] P. Fayet. Spontaneously broken supersymmetric theories of weak, electromagnetic and strong interactions. *Physics Letters B*, 69(4):489 – 494, 1977.



- [14] N. G. Deshpande and E. Ma. Pattern of symmetry breaking with two Higgs doublets. *Phys. Rev. D*, 18:2574–2576, 1978.
- [15] R. Barbieri, L. J. Hall, and V. S. Rychkov. Improved naturalness with a heavy Higgs boson: An alternative road to CERN LHC physics. *Phys. Rev. D*, 74:015007, 2006.
- [16] E. Lundström, M. Gustafsson, and J. Edsjö. Inert doublet model and LEP II limits. *Phys. Rev. D*, 79:035013, 2009.
- [17] L. Evans and P. Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, 2008.
- [18] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, and The LEP Working Group for Higgs Boson Searches. Search for the Standard Model Higgs boson at LEP. *Physics Letters B*, 565(0):61–75, 2003.
- [19] TEVNPH (Tevatron New Phenomena, Higgs Working Group), CDF Collaboration, and D0 Collaboration. Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to  $8.6 \text{ fb}^{-1}$  of Data, 2011.
- [20] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008.
- [21] The CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006.
- [22] The CMS Collaboration. The CMS electromagnetic calorimeter project: Technical Design Report, 1997.
- [23] K. G. McKay. Avalanche Breakdown in Silicon. *Phys. Rev.*, 94:877–884, 1954.
- [24] R.-Y. Zhu. Radiation damage in scintillating crystals. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 413(23):297–311, 1998.
- [25] The CMS Collaboration. The CMS hadron calorimeter project: Technical Design Report, 1997.
- [26] M. Cacciari, G. P. Salam, and G. Soyez. The anti- $k_t$  jet clustering algorithm. *JHEP*, 0804:063, 2008.
- [27] M. Cacciari, G. P. Salam, and G. Soyez. The Catchment Area of Jets. *JHEP*, 0804:005, 2008.
- [28] Matteo C. and Gavin P. S. Pileup subtraction using jet areas. *Physics Letters B*, 659(12):119–126, 2008.

- [29] The CMS Collaboration. Jet Performance in  $pp$  collisions at 7 TeV, 2010. CMS-PAS-JME-10-003.
- [30] S. Alioli, P. Nason, C. Oleari, and E. Re. NLO vector-boson production matched with shower in POWHEG. *Journal of High Energy Physics*, 2008(07):060, 2008.
- [31] P. Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *Journal of High Energy Physics*, 2004(11):040, 2004.
- [32] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *Journal of High Energy Physics*, 2007(11):070, 2007.
- [33] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer. Madgraph 5: Going Beyond. *Journal of High Energy Physics*, 2011:1–40, 2011.
- [34] T. Sjöstrand, S. Mrenna, and P. Skands. Pythia 6.4 Physics and Manual. *Journal of High Energy Physics*, 2006(05):026, 2006.
- [35] R. Field. Min-Bias and the Underlying Event at the LHC. *ArXiv e-prints*, October 2011.
- [36] S. Agostinelli et al. Geant4a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.
- [37] J. et al. Allison. Geant4 developments and applications. *Nuclear Science, IEEE Transactions on*, 53(1):270 –278, 2006.
- [38] S. D. Drell and T.-M. Yan. Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies. *Phys. Rev. Lett.*, 25:316–320, Aug 1970.
- [39] The CMS Collaboration. Measurement of the inclusive  $W$  and  $Z$  production cross sections in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the CMS experiment. *Journal of High Energy Physics*, 2011:1–76, 2011.
- [40] P. J. Rijken and W. L. van Neerven. Order  $\alpha_s^2$  contributions to the Drell-Yan cross section at fixed target energies. *Phys. Rev. D*, 51:44–63, Jan 1995.
- [41] R. Hamberg, W. L. van Neerven, and T. Matsuura. A complete calculation of the order  $\alpha_s^2$  correction to the Drell-Yan K-factor. *Nuclear Physics B*, 359(23):343 – 405, 1991.
- [42] W. L. van Neerven and E. B. Zijlstra. The  $\mathcal{O}(\alpha_s^2)$  corrected Drell-Yan K-factor in the DIS and MS schemes. *Nuclear Physics B*, 382(1):11 – 62, 1992.
- [43] R. V. Harlander and W. B. Kilgore. Next-to-Next-to-Leading Order Higgs Production at Hadron Colliders. *Phys. Rev. Lett.*, 88:201801, May 2002.

- [44] C. Anastasiou, L. Dixon, K. Melnikov, and F. Petriello. High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order. *Phys. Rev. D*, 69:094008, May 2004.
- [45] M. Cacciari, G. P. Salam, and G. Soyez. FastJet user manual. *The European Physical Journal C*, 72:1–54, 2012.
- [46] M. Cacciari and G. P. Salam. Dispelling the myth for the jet-finder. *Physics Letters B*, 641(1):57 – 61, 2006.
- [47] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan. New parton distributions for collider physics. *Phys. Rev. D*, 82:074024, 2010.
- [48] J. C. Collins and D. E. Soper. Parton distribution and decay functions. *Nuclear Physics B*, 194(3):445 – 492, 1982.
- [49] J. C. Collins, D. E. Soper, and G. Sterman. Transverse momentum distribution in Drell-Yan pair and  $W$  and  $Z$  boson production. *Nuclear Physics B*, 250(14):199 – 224, 1985.
- [50] C. Balázs and C.-P. Yuan. Soft gluon effects on lepton pairs at hadron colliders. *Phys. Rev. D*, 56:5558–5583, Nov 1997.
- [51] E. Barberio, B. van Eijk, and Z. Waş. Photos - a universal Monte Carlo for QED radiative corrections in decays. *Computer Physics Communications*, 66(1):115 – 128, 1991.
- [52] K. Melnikov and F. Petriello. Electroweak gauge boson production at hadron colliders through  $\mathcal{O}(\alpha_s^2)$ . *Phys. Rev. D*, 74:114017, 2006.
- [53] K. Melnikov and F. Petriello.  $W$  boson production cross section at the large hadron collider with  $\mathcal{O}(\alpha_s^2)$  corrections. *Phys. Rev. Lett.*, 96:231803, Jun 2006.
- [54] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and A. Vicini. Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron colliders. *Journal of High Energy Physics*, 2007(10):109, 2007.
- [55] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and A. Vicini. Precision electroweak calculation of the charged current Drell-Yan process. *Journal of High Energy Physics*, 12:16, 2006.
- [56] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and M. Treccani. Multiple photon corrections to the neutral-current Drell-Yan process. *Journal of High Energy Physics*, 2005(05):019, 2005.
- [57] C.M. Carloni Calame, G. Montagna, O. Nicrosini, and M. Treccani. Higher-order QED corrections to  $W$ -boson mass determination at hadron colliders. *Phys. Rev. D*, 69:037301, 2004.

- [58] The CMS Collaboration. Absolute Calibration of Luminosity Measurement at CMS: Summer 2011 Update, 2011. CMS-PAS-EWK-11-001.
- [59] The CMS Collaboration. Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update, 2012. CMS-PAS-SMP-12-008.
- [60] S. van der Meer. Calibration of the effective beam height in the ISR. Technical Report CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968.
- [61] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Uncertainties on  $\alpha_s$  in global PDF analyses and implications for predicted hadronic cross sections. *The European Physical Journal C*, 64:653–680, 2009.
- [62] G. Watt. Parton distribution function dependence of benchmark Standard Model total cross sections at the 7 TeV LHC. *Journal of High Energy Physics*, 2011:1–41, 2011.
- [63] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Heavy-quark mass dependence in global PDF analyses and 3- and 4-flavour parton distributions. *The European Physical Journal C*, 70:51–72, 2010.
- [64] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini. The  $q_T$  spectrum of the Higgs boson at the LHC in QCD perturbation theory. *Physics Letters B*, 564(12):65 – 72, 2003.
- [65] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini. Transverse-momentum resummation and the spectrum of the higgs boson at the lhc. *Nuclear Physics B*, 737(12):73 – 120, 2006.
- [66] Daniel Florian, Giancarlo Ferrera, Massimiliano Grazzini, and Damiano Tommasini. Transverse-momentum resummation: Higgs boson production at the tevatron and the lhc. *Journal of High Energy Physics*, 2011:1–22, 2011.
- [67] The CMS Collaboration. Search for the Standard Model Higgs boson decaying to  $W^+W^-$  in the fully leptonic final state in  $pp$  collisions at  $\sqrt{s} = 7$  TeV. *Physics Letters B*, 710(1):91 – 113, 2012.
- [68] K. Nakamura et al. (Particle Data Group). The Review of Particle Physics. *J. Rev. G*, 37:075021, 2010.
- [69] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, S. Forte, A. Glazov, J. Huston, R. McNulty, T. Sjöstrand, and R. Thorne. The PDF4LHC Working Group Interim Recommendations. *ArXiv e-prints*, January 2011.
- [70] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, and M. Ubiali. Impact of heavy quark masses on parton distributions and LHC phenomenology. *Nuclear Physics B*, 849(2):296 – 363, 2011.

- [71] LHC Higgs Cross Section Working Group and S. et al. Dittmaier. Handbook of LHC Higgs Cross Sections: 2. Differential Distributions. *CERN-2012-002*, CERN, Geneva, 2012.
- [72] T. Junk. Confidence level computation for combining searches with small statistics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 434(23):435 – 443, 1999.
- [73] A. L. Read. Modified frequentist analysis of search results (the  $cl_s$  method), 2000.
- [74] ATLAS and LHC Higgs Combination Group CMS Collaborations. Procedure for the LHC Higgs boson search combination in Summer 2011. Technical Report CMS-NOTE-2011-005, CERN, Geneva, Aug 2011.
- [75] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71:1–19, 2011.
- [76] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9(1):pp. 60–62, 1938.
- [77] A. Wald. Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large. *Transactions of the American Mathematical Society*, 54(3):pp. 426–482, 1943.
- [78] The CMS Collaboration. Combined results of searches for the standard model Higgs boson in pp collisions at  $\sqrt{s} = 7$  TeV. *Physics Letters B*, 710(1):26 – 48, 2012.
- [79] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30 – 61, 2012.
- [80] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1 – 29, 2012.