Search for Higgs boson in the WW* channel in ATLAS and drift time measurement in the liquid argon calorimeter in ATLAS

Xifeng Ruan

Xifeng Ruan. Search for Higgs boson in the WW* channel in ATLAS and drift time measurement in the liquid argon calorimeter in ATLAS. Other [cond-mat.other]. Université Paris Sud - Paris XI; Institute of high energy physics (Chine), 2012. English. NNT : 2012PA112236 . tel-00755876

HAL Id: tel-00755876
https://theses.hal.science/tel-00755876

Submitted on 22 Nov 2012

UNIVERSITÉ
PARIS
SUD
Comprendre le monde,
construire l'avenir®

中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

UNIVERSITE PARIS-SUD

Ecole doctorale: 517 - Particules, Noyaux et Cosmos
Laboratoire de l'Accélérateur Linéaire

Discipline: Physique des particules

THÈSE DE DOCTORAT

Présentée par Xifeng RUAN

Search for Higgs boson in the $WW^{(*)}$ channel in ATLAS and drift time measurement in the liquid argon calorimeter in ATLAS

Recherche du boson de Higgs dans le canal $WW^{(*)}$ dans l'expérience ATLAS et mesure du temps de dérive du calorimètre à argon liquide dans l'expérience ATLAS

Soutenue le 19 Oct 2012 devant le jury composé de:

Directeur de thèse:	Zhiqing Zhang	LAL
Co-directeur de thèse:	Shan Jin	IHEP
Président et Rapporteur:	Yuanning Gao	Tsinghua Univ.
Rapporteur:	Samira Hassani	IRFU
Examinateur:	Congfeng Qiao	UCAS
Examinateur:	Guoming Chen	IHEP

Résumé

Une recherche du boson de Higgs est effectuée dans le canal $WW \rightarrow l\nu l\nu$ en utilisant l'ensemble des données de 2011 à une énergie dans le centre de masse de $\sqrt{s} = 7\,\mathrm{TeV}$ et une partie des données de 2012 à 8 TeV prises par l'expérience ATLAS auprès du LHC. Les luminosités intégrées correspondantes sont $4.7\,\mathrm{fb}^{-1}$ et $5.8\,\mathrm{fb}^{-1}$, respectivement. Plusieurs méthodes sont introduites pour estimer à partir des données la contribution de bruits de fond des différents processus afin de minimiser l'utilisation de la simulation. Pour la contribution du bruit de fond top dans le canal dominant avec zéro jet, elle est estimée avec une mthode que nous avons proposée. Une autre méthode pour corriger la forme de la distribution de l'énergie transverse manquante dans les événements Drell-Yan à partir des événements W+jets est également présentée. En 2011, le boson de Higgs du modèle standard avec la masse du Higgs de 133 à 261 GeV est exclue à 95% de niveau de confiance, tandis que la plage d'exclusion prévue est de 127 à 234 GeV. En 2012, un excès d'événements au-dessus du bruit de fond attendu est observé dans une plage de masse autour de 125 GeV. En combinant les deux échantillons, la probabilité minimale ("p-value") pour que l'hypothèse bruit de fond seul fournisse autant ou plus d'événements qu'observé dans les données est de 3×10^{-3}, ce qui correspond à une signifiance statistique de 2,8 écarts types. Le taux de production mesuré du signal par rapport au taux prédit pour le boson de Higgs du modèle standard à $m_H = 125\,\mathrm{GeV}$ est de $1,4 \pm 0,5$. La probabilité attendue pour un Higgs avec $m_H = 125\,\mathrm{GeV}$ est de 0,01, soit de 2,3 écarts types. La limite d'exclusion d'un Higgs dans un modèle avec une quatrième génération est également présentée en utilisant une partie de l'échantillon de données 2011, la gamme de masse entre 120 GeV et 600 GeV a été exclue à 95% de niveau de confiance. Enfin, l'étude sur le temps de dérive dans le calorimètre à argon liquide du détecteur ATLAS est effectuée en utilisant tous les échantillons de données du rayonnement cosmique, du faisceau splash et de collision. Les résultats ne montrent aucune non-uniformité significative sur la largeur de l'espace cellulaire mis à part un effet de "sagging" dans les régions de transition dû au poids du calorimètre.

Mots-clés: Modèle standard, Higgs, quatrième génération, WW, temps de dérive, calorimètre

Abstract

A Higgs search is performed in the $WW \rightarrow l\nu l\nu$ channel using the full 2011 data at a center-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ and part of 2012 data at 8 TeV taken by the ATLAS experiment at the LHC. The corresponding integrated luminosity values are $4.7\,\mathrm{fb}^{-1}$ and $5.8\,\mathrm{fb}^{-1}$, respectively. The cut based analysis is performed and several data-driven methods for background estimation are introduced. The jet veto survival probability method for top background estimation in 0-jet bin is proposed and used in the Higgs search. Another data-driven method to correct $E_{\mathrm{T}}^{\mathrm{miss}}$ shapes in the Drell-Yan process is also presented. In 2011, the standard model Higgs boson with the Higgs mass from 133 to 261 GeV is excluded at 95% CL, while the expected exclusion range is $127 - 234\,\mathrm{GeV}$. In 2012, an excess of events over expected background is observed at

$m_H = 125$ GeV. Combining both samples, the minimum observed p_0 value is 3×10^{-3}, corresponding to 2.8 standard deviations. The fitted signal strength at $m_H = 125$ GeV is $\mu = 1.4 \pm 0.5$. The expected p_0 for a Higgs with $m_H = 125$ GeV is 0.01, or 2.3 standard deviations. The exclusion limit for a Higgs in a fourth generation model is shown using part of the 2011 data sample, the mass range between 120 GeV and 600 GeV has been excluded at 95% CL. The study of the drift time in the liquid argon calorimeter in ALTAS is performed using all special data samples from cosmic muons, beam splash and beam collision data. The results show no significant non-uniformity on the cell gap width and a sagging effect due to gravity is observed.

Keywords : Standard model, Higgs, fourth generation, WW, drift time, calorimeter.

Contents

Introduction

The Higgs boson, responsible for giving masses to weak bosons, charged leptons and quarks, was the only undiscovered particle in the Standard Model(SM), which describes the constituents of all observed matters in our world and their interactions. Searches for the Higgs boson have been performed by many experiments over the last decades.

The Higgs search is one of the main research programmes at the Large Hadron Collider (LHC), the largest accelerator ever built in the world. It delivered pp collisions since 2010. This huge machine provides ultra high energy up to a nominal value of 7 TeV per beam and high luminosity, thereby giving excellent sensitivity for searching for Higgs as well as new physics. A Toroidal LHC ApparatuS(ATLAS) is one of the largest detectors at the LHC. It is a general purpose detector designed for discovering the Higgs boson and studying many other physics subjects. It provides very good tracking performance for charged particles such as muons and precise energy measurement for electrons and photons.

Several working groups are setup in ATLAS to search for the Higgs boson in different decay modes. Among these channels, the Higgs decay to WW^* to $l\nu l\nu$ is one of the most promising channels due to its large branching fraction over a large mass range. The final state with two isolated leptons and large missing transverse energy also provides a clean signature to suppress SM background contributions, in particular the huge quantum chromodynamics(QCD) related background.

In this thesis the analysis performed in this channel with the full 2011 data at a center-of-mass energy $\sqrt{s} = 7\,\mathrm{TeV}$ and part of the 2012 data at $\sqrt{s} = 8\,\mathrm{TeV}$ is described. The integrated luminosity values used are $4.7\,\mathrm{fb}^{-1}$ and $5.8\,\mathrm{fb}^{-1}$ for 2011 and 2012, respectively.

The major part of the thesis is devoted to the Higgs search both within the SM using all the data mentioned above and in a model with fourth generation heavy quarks using part of the 2011 data. One chapter is reserved for the drift time measurement performed during the thesis using dedicated beam splash and collision data samples.

The thesis is organised as follows:

Chapter 1: The Standard Model and Higgs mechanism are introduced. The current constraints from theoretical considerations and direct searches and indirect precision measurements are summarised.

Chapter 2: The LHC machine is briefly described and the ATLAS detector, in particular those components relevant for this analysis, is presented in more detail.

Chapter 3: The object (electrons, muons, and jets) reconstruction is introduced. Then follows the identification and selection criteria on these objects, which are used in the $H \rightarrow WW^{(*)}$ dilepton analysis.

Chapter 4: The cut-based Higgs search in the $l\nu l\nu$ channel is presented.

Chapter 5: The main purpose is to describe data-driven methods and the corresponding background estimations, including W+jets, Z+jets and top.

Chapter 6: The statistical analysis basis for a discovery or an exclusion is presented as well as the results for in the Higgs search in the fourth generation and in the SM.

Chapter 7: The analysis and results for the 2012 data are presented.

Chapter 8: The drift time measurement in the liquid Argon calorimeter in the ATLAS detector is shown.

Chapter 1

Introduction to the Standard Model and Higgs

1.1 The Standard Model

As a result of strong interplay between experimental discoveries and theoretical advances, a successful model to describe the constituents of our world and their interactions is founded. This is so called Standard Model (SM) of particle physics. The SM reveals that there are six types of leptons, six types of quarks and their antiparticles as well. All these together make up of everything in our world. There are four types of interaction which rule all of the materials in the world, propagated by a set of bosons. They are strong, electromagnetic, weak and gravity interaction. The gravity interaction is not incorporated in the SM. In the microscope scale, the gravity interaction is however so weak that it is negligible in the study of particle physics. In the SM, six leptons are arranged in 3 generations (from left to right in Fig. 1.1) with similar physical characteristics [4]). The particles in higher generations have heavier masses. In each generation, there is one particle with an integer charge (-1) and none zero mass, while the other lepton is charge neutral, which is called neutrino. Leptons only take part in electroweak interaction. Each lepton has its antiparticle with an opposite sign charge and same mass.

Quarks, the constituent of baryons and mesons, account for most of the masses in the world. Similarly six quarks are categorised to three generations, with heavier and heavier masses. Unlike the leptons, quarks have fractional charges. For example, the u, c, t quark have $2/3$ charge, and d, s, b have $-1/3$ charge. As quarks can only exist within the bounded state in the hadrons, only integer charges can be observed in the experiment. Quarks will take part in not only the electroweak interaction, but also the strong interaction indicated by another kind of charge, color. As the requirement of symmetry in hadrons, there are three colors in all, given the name red, blue and green. This assumption was verified by the R value measurement. R is the ratio of the cross section for e^+e^- annihilation into hadronic final states to the Quantum Electrodynamics (QED) cross section for muon-pair production [5].

In addition, 12 types of bosons propagate the interactions between particles, including 8 gluons, gamma, W^+, W^- and Z bosons. The interactions are described by several dynamics in the framework of quantum field theory. Table 1.1 shows the interactions between particles.

Interaction	Theory	Mediator	Strength	Range(m)
Strong	Quantum Chromodynamics (QCD)	Gluons	1	10^{-15}
Electromagnetic	Quantum Electrodynamics (QED), GWS	Photons	10^{-2}	∞
Weak	Glashow-Weinberg-Salam (GWS)	W^+, W^-, Z	10^{-7}	10^{-18}
Gravity	General Relativity	Gravitons	10^{-39}	∞

Table 1.1: Summary of four interactions and their underlying theory, mediator, interaction strength and range. The electromagnetic and weak interaction are unified in to the GWS theory [8, 9, 10]. The gravity interaction is not included in the SM.

1.2 Symmetry and Dynamics

In a theory, a principle usually requires that under certain operations, such as translation, rotation in space-time, mirror reflection or other transformation, the physical system is invariant. Each invariance under an operation leads to a conservation law. This is called symmetry. In the field theory, the dynamics of the theory is described by a Lagrangian function, which has an intrisic symmetry, namely the invariance of the Lagrangian, under a phase transformation on the complex field:

$$\phi(x) \rightarrow \phi'(x) = e^{ie\alpha}\phi(x), \tag{1.1}$$

in which $\phi(x)$ is the field, α is a scalar and e is the strength of the transformation. Such a transformation can be either global or local. In the local case the scalar α is a function of coordinates: $\alpha(x)$. A traditional view on this operation is that like the gravity theory, the metric $g_{\mu\nu}$ is changed under the intrinsic transformation. To keep the invariance of the Lagrangian, a covariant derivative is introduced:

$$D_\mu\phi(x) = (\partial_\mu - ieA_\mu(x))\phi(x)\,. \tag{1.2}$$

This formula gives a term quite like the connection in the tensor analysis and leads to define a vector field $A_\mu(x)$ like electromagnetic fields. It transforms as:

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu\alpha\,. \tag{1.3}$$

It is called gauge field and the invariance is called gauge invariance [6]. The form of such transformation is described by the group theory. A U(1) Abel transformation introduces the QED theory while a non-abelian group leads to more complicated cases [7]: in Eq. 1.4, the T^a represents the generators of the transformation group. This is the general expression of the symmetry transformation:

$$\phi(x) \rightarrow \phi'(x) = exp[-igT^a\alpha_a]\phi(x). \tag{1.4}$$

1.2.1 QED

If we consider a U(1) local transformation

$$\phi(x) \rightarrow \phi'(x) = e^{ie\alpha(x)}\phi(x)\,, \tag{1.5}$$

on Lagrangian:

$$L = i\overline{\phi}(x)\gamma^{\mu}\partial_{\mu}\phi(x) - m\overline{\phi}(x)\phi(x)\,, \tag{1.6}$$

it naturally gives the electromagnetic term A_μ in the covariant derivative as shown in Eq. 1.2. The constant e, which is called charge, appears in the interaction term:

$$e\overline{\phi}(x)\gamma^{\mu}A_{\mu}\phi(x)\,. \tag{1.7}$$

Such a Lagrangian describes the interaction between charged particles and photons. Meanwhile the kinematics term of electromagnetic fields can be introduced as:

$$L_{\mathrm{kin}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,. \tag{1.8}$$

The requirement of the gauge invariance gives a massless gauge boson: photon. This is called Quantum Electrodynamics (QED).

1.2.2 QCD

According to Yang Mills gauge theory, a non-abelian gauge theory, a $\mathrm{SU}_c(3)$ local transformation is introduced:

$$\phi(x) \rightarrow \phi'(x) = e^{-ig_s\frac{\lambda^a}{2}\alpha_a}\phi(x)\,. \tag{1.9}$$

Here $\frac{\lambda^a}{2}$ is the generator of the transformation group, g_s is the strong interaction strength and λ^a is called Gell-Mann Matrices, which satisfy the commutation relation:

$$[\frac{\lambda^a}{2}, \frac{\lambda^b}{2}] = if^{abc}\frac{\lambda^c}{2}, \tag{1.10}$$

where f^{abc} is called structure constant. The SU(3) group has 8 generators, each is basically represented by a 3×3 matrix. That means the quarks should be written in triplets corresponding to three colors. These 8 gauge fields give 8 vector bosons, which are called gluons, and have interaction with quarks. This dynamics describing the strong interaction is called QCD [11] [12] [13] [14], as shown in the Lagrangian:

$$L = -\frac{1}{4}(F^a_{\mu\nu})^2 + \sum_i \overline{q}_i(i\gamma^{\mu}D_{\mu} - m)q_i\,, \tag{1.11}$$

with

$$\begin{aligned} F^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc}A^b_\mu A^c_\nu \\ D_\mu &= \partial_\mu - ig_s A^a_\mu\frac{\lambda^a}{2}, \end{aligned} \tag{1.12}$$

where q is a triplet, the basic presentation of the SU(3) group, summed over all quark flavours. Due to the non-abelien configuration, the structure constant from the commutator of generators in the SU(3) group leads to the self interaction of the gauge field. When the higher order loop diagrams are considered, it is needed to subtract the divergences by renormalization. The equation of renormalization shows that in the QCD case the strength of interaction is stronger when the distance between particles increases. This is the reason why quarks are confined in the hadrons.

1.2.3 GWS theory

As only the left handed leptons and neutrinos take part in the weak interaction, the Lagrangian of leptons and quarks is written as follows:

$$\begin{aligned} L &= \bar{l}^i_L i\gamma^\mu\partial_\mu l^i_L + \bar{q}^i_L i\gamma^\mu\partial_\mu q^i_L \\ &\quad + \bar{l}^i_R i\gamma^\mu\partial_\mu l^i_R + \bar{u}^i_R i\gamma^\mu\partial_\mu u^i_R + \bar{d}^i_R i\gamma^\mu\partial_\mu d^i_R \,, \end{aligned} \tag{1.13}$$

in which l^i_L, q^i_L reperesent the left handed lepton, quark doublet, respectively:

$$l^i_L : \begin{pmatrix} \nu_{iL} \\ l_{iL} \end{pmatrix} = \frac{1-\gamma_5}{2}\begin{pmatrix} \nu_i \\ l_i \end{pmatrix} \,, \quad q^i_L : \begin{pmatrix} u_{iL} \\ d'_{iL} \end{pmatrix} = \frac{1-\gamma_5}{2}\begin{pmatrix} u_i \\ d'_i \end{pmatrix} \,, \tag{1.14}$$

where l_i represents the e, μ, τ, u_{iL} represents u, c, t quarks and d'_i represents the weak interaction eigenstates of d, s, b quarks. For the right handed parts, there is no right handed neutrinos: $l^i_R = \frac{1+\gamma_5}{2} l_i$. For quarks, $u_{iR} = \frac{1+\gamma_5}{2} u_i, d'_{iR} = \frac{1+\gamma_5}{2} d'_i$. If the left handed doublets are transformed under the SU(2) group:

$$l'^i_L = e^{ig\tau^j\beta^j(x)} l^i_L \,, \tag{1.15}$$

in which g is the coupling constant, τ is the Pauli matrix and β is a scalar as a function of coordinates. For the right handed part a $\mathrm{U}(1)_Y$ transformation is implemented:

$$l'^i_R = e^{i\frac{g'}{2}Y\alpha(x)} l^i_R \,. \tag{1.16}$$

Here Y is called hypercharge, which satisfies the relationship with charge and the third component of the weak isospin τ_3:

$$Q = \tau_3 + \frac{Y}{2} \,. \tag{1.17}$$

Meanwhile, if the SU(2) generator is applied on the right handed fields:

$$\tau^j l_{iR} = 0, \quad \tau^j u_{iR} = 0, \quad \tau^j d_{iR} = 0, \quad i, j = 1, 2, 3 \,. \tag{1.18}$$

Therefore if under the transformation of Eqs. 1.12 and 1.16 the Lagrangian is invariant, this variance is called the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry, a combined transformation of rotation in the weak isospin and weak hypercharge space. The gauge invariance gives the covariant derivatives:

$$D_\mu = \partial_\mu - \frac{i}{2}\begin{pmatrix} gA^3_\mu + Yg'B_\mu & \sqrt{2}gW^+_\mu \\ \sqrt{2}gW^-_\mu & -gA^3_\mu + Yg'B_\mu \end{pmatrix} \,, \tag{1.19}$$

where W represents the gauge field which only couples the left handed leptons. After a rotation, the four gauge bosons are reorganised as:

$$\begin{aligned}
A_\mu &= \cos\theta B_\mu + \sin\theta A^3_\mu \\
Z_\mu &= -\sin\theta B_\mu + \sin\theta A^3_\mu \\
\sin\theta &= \frac{g'}{\sqrt{g^2+g'^2}}\,, \\
\cos\theta &= \frac{g}{\sqrt{g^2+g'^2}}\,.
\end{aligned} \tag{1.20}$$

The covariant deravitives then become:

$$D_\mu = \partial_\mu - \frac{i}{2}\begin{pmatrix} g\sin\theta(1+Y)A_\mu + g\frac{\cos^2\theta - Y\sin^2\theta}{\cos\theta}Z_\mu & \sqrt{2}gW^+_\mu \\ \sqrt{2}gW^-_\mu & g\sin\theta(-1+Y)A_\mu - g\frac{\cos^2\theta + Y\sin^2\theta}{\cos\theta}Z_\mu \end{pmatrix}. \tag{1.21}$$

This leads to the interaction term:

$$\begin{aligned}
L_{\text{int}} &= \frac{g}{2\sqrt{2}}[\bar{l}_i\gamma^\mu(1-\gamma_5)\nu_i + \overline{d}'_i\gamma^\mu(1-\gamma_5)u_i]W^-_\mu \\
&+\frac{g}{2\sqrt{2}}[\overline{\nu}_i\gamma^\mu(1-\gamma_5)l_i + \overline{u}_i\gamma^\mu(1-\gamma_5)d'_i]W^+_\mu \\
&+\frac{g}{4\cos\theta}[\overline{\nu}_i\gamma^\mu(1-\gamma_5)\nu_i + \bar{l}_i\gamma^\mu(-(1-4\sin^2\theta)+\gamma_5)l_i \\
&+\overline{u}_i\gamma^\mu((1-\frac{8}{3}\sin^2\theta)-\gamma_5)u_i + \overline{d}'_i\gamma^\mu(-(1-\frac{4}{3}\sin^2\theta)+\gamma_5)d'_i]Z_\mu \\
&+g\sin\theta[-\bar{l}_i\gamma^\mu l_i + \frac{2}{3}\overline{u}_i\gamma^\mu u_i - \frac{1}{3}\overline{d}'_i\gamma^\mu d'_i]A_\mu\,.
\end{aligned} \tag{1.22}$$

Such rotation is called the Weinberg rotation and θ is called Weinberg angle. A_μ only couples to charged particles with a strength as the electric charge: $e = g\sin\theta$, which represents the electromagnetic field, while W couples to neutrinos and leptons and Z represents the neutral current of weak process. The W and Z bosons were discovered at UA1 in SPS in 1983 [1] [2], also the Weinberg angle was measured as $\sin^2\theta = 0.23$ [15]. Thus, the weak and electromagnetic interactions are unified under this framework [8, 9, 10].

1.3 Higgs Physics

1.3.1 The Higgs Mechanism

All of the fields, both bosons and fermions introduced above are massless according to the requirement on the gauge invariance, but quarks and leptons such as electrons do have masses. In the SM, it is the Higgs mechanism which explains how the particles gain their masses [4].

Assume there is a scalar field, called Higgs field, with a potential term in the Lagrangian described here:

$$V(\phi) = -\mu^2\phi^\star\phi + \lambda(\phi^\star\phi)^2\,. \tag{1.23}$$

The potential is expanded around the minimum value. As shown in Fig. 1.2, if the $\mu^2 > 0$ and $\lambda > 0$, it is possible to have multiple minimum potential positions which have the same vacuum expectation energy:

$$\phi_0 = \sqrt{\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}, \tag{1.24}$$

where the field ϕ can be replaced by $\phi_0 + \phi$. The original symmetry of the Lagrangian will be broken by expanding the field ϕ around one of the minimum. Take the U(1) transformation as an example, a complex scalar field $\phi(x)$ is defined as:

$$\phi(x) = \frac{1}{\sqrt{2}}[v + h(x) + i\chi(x)], \tag{1.25}$$

in which the h and χ are two real scalar fields, then add the kinematic term in the Lagrangian:

$$L = \frac{1}{2}\partial_\mu h \partial^\mu h + \frac{1}{2}\partial_\mu \chi \partial^\mu \chi - \frac{1}{2}(2\lambda v^2)h^2 - \lambda v h(h^2 + \chi^2) - \frac{1}{4}\lambda(h^2 + \chi^2)^2\,, \tag{1.26}$$

a mass term of h is generated: $m_h = \sqrt{2\lambda v^2}$ and the χ remains massless. In this case the U(1) symmetry is broken and a massless boson χ is generated, which is called the Goldstone boson, meanwhile the Higgs boson gains mass. By using a proper parameterisation on the field to eat up the Goldstone bosons, the gauge boson will also have a mass $m_A = ev/\sqrt{2}$. In general, if the transformation is performed by a group with N generators and the expanding around the vacuum point breaks M of N dimensions, there will be M massless Goldstone bosons generated and the Higgs boson gains mass. Under a proper parameterisation, the Goldstone bosons disappear and M gauge bosons have masses with N-M massless gauge bosons remaining. In the GWS theory, the SU(2)$_L$×U(1)$_Y$ symmetry breaks to the U(1) symmetry, in which case three dimensions are broken. So by repeating the above procedures 3 gauge bosons gain masses and the photon remains massless due to the unbroken symmetry:

$$m_{W^\pm} = \frac{1}{2}gv, \quad m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}, \quad m_A = 0\,. \tag{1.27}$$

By using the Weinberg angle, the electron charge and the vacuum expectation value v, the mass of W and Z can be calculated. The measured W and Z mass are $M_W = 80.399 \pm 0.023\,\text{GeV}$ and $M_Z = 91.1876 \pm 0.0021\,\text{GeV}$ [16].

To give mass to the fermions, the Yukawa coupling is introduced. Here shows the lepton part:

$$L_{\text{Yukawa lepton}} = -g_i^L \bar{l}_{iR} \Phi^+ l_L^i + h.c.. \tag{1.28}$$

Assume this scheme is also applied to all fermions, the fermion masses are given:

$$m_{li} \equiv \frac{g_i^L}{2}v, \quad m_{di} \equiv \frac{g_i^d}{2}v, \quad m_{ui} \equiv \frac{g_i^u}{2}v, \tag{1.29}$$

and the interaction strength:

$$-\frac{i}{v}m_f. \tag{1.30}$$

Here the coupling strength between fermions and the Higgs is proportional to the mass of fermions.

1.3.2 Higgs Production at the LHC

As one of the main physical purposes for the Large Hadron Collider (LHC) is the Higgs hunting, it is very important to know how the Higgs is produced through proton-proton process [17]. The Higgs boson is mainly produced from several processes such as gluon-gluon fusion, vector boson fusion(VBF), vector boson associated production $VH(V{=}W{,}Z)$, and $t\bar{t}H$ process. The leading Feynman diagrams are shown in Fig. 1.3, in which the gluon-gluon fusion process via a fermion loop has the largest production cross-section, then follows the pure electroweak process including the VBF production. The production cross sections as a function of Higgs mass for the SM at center-of-mass energies of $\sqrt{s} = 7\,\mathrm{TeV}$ and $\sqrt{s} = 8\,\mathrm{TeV}$ are shown in Fig. 1.4. For a model with a fourth generation heavy quarks, the gluon-gluon fusion production cross section is significantly enhanced. This is discussed in Chapter 6.

Figure 1.5 shows the SM Higgs decay branching ratio. The WW channel has the dominant branching fraction when the Higgs mass is over 160 GeV, therefore the Higgs to the WW channel has the largest sensitivity for the intermediate Higgs mass [21]. It also has comparable sensitivity as $H \to \gamma\gamma$ at the low Higgs mass region. The ZZ dominates in the high mass range and diphoton dominates in the low mass range. In the low Higgs mass region, the Higgs to $b\bar{b}$ or $\tau\tau$ has larger branching ratio but in these channels it is very hard to suppress the backgrounds.

1.3.3 Constraints on Higgs Mass

In the following, both the theoretical and direct and indirect experimental constraints on the SM Higgs boson mass are discussed.

Theoretical Constraints

The cross section from the WW elastic scattering in the SM is calculated with a number of Feynman diagrams shown in Fig. 1.6. If only the left three diagrams are used to calculate the WW scattering cross section, the amplitude for the scattering of longitudinal W and Z increases with the energy, which finally violates the unitarity at a typical energy of 1.2 TeV. However if the contribution from the Higgs intermediate state is introduced, shown in the right two diagrams in Fig. 1.6, the unitarity is restored. This gives constraint on Higgs mass $m_H < 1\,\mathrm{TeV}$ [37].

Another theoretical constraint on the Higgs mass is called triviality bound. By using the renormalization group equation, the running coupling constant λ is shown as:

$$\frac{d}{d\ln Q}\lambda(Q) = \frac{3\lambda^2}{2\pi^2} + \mathcal{O}(1)\,. \tag{1.31}$$

As m_H is a function of λ, if the Higgs mass is very large, the solution to this equation has a Landau pole. When the Higgs mass is very small, the λ may become negative at large energy scale. The former case is called triviality and the latter one the vacuum stability problem. To avoid this, a cut off on the Q should be applied: $Q = \Lambda$. If no new physics, in which the SM is embedded, exsits at the Planck scale $m_{p1} \simeq 10^{19} GeV$, then a smaller upper limits on Higgs mass is obtained: $m_H < 180\,\mathrm{GeV}$. The allowed range for the Higgs mass as a function of energy scale is shown in Fig. 1.7 [38, 39, 40].

Experimental Constraints

There are direct and indirect experimental constraints on Higgs mass. The direct search was performed by the LEP and Tevatron experiments. The LEP results are shown in Fig. 1.8. The lower bound for Higgs was set at 114.4 GeV at 95% CL [41]. The latest Tevatron results are shown in Fig. 1.9 [45]. Two Higgs mass regions $100 < m_H < 103\,\mathrm{GeV}$ and $147 < m_H < 180\,\mathrm{GeV}$ have been excluded at 95% CL. There is a significant excess of data events with respect to the background estimation in the mass range $115 < m_H < 140\,\mathrm{GeV}$. At $m_H = 120\,\mathrm{GeV}$, the p-value for a background fluctuation to produce this excess is $\sim 1.5 \times 10^{-3}$, corresponding to a local significance of 3.0 standard deviations.

In the indirect case, the Higgs mass can be constrained by performing a global fit to all precision measurements, with the most relevant measurements on the constraint of m_H being the Z, W bosons and top masses M_Z, M_W, m_t and the hadronic contribution to the electromagnetic coupling constant $\Delta\alpha_{\mathrm{had}}(M_Z^2)$. The result provided by the Gfitter group is $m_H = 95.6^{+30.5}_{-24.2}\,\mathrm{GeV}$ not including the direct Higgs searches in the fit (Fig. 1.10) [47].

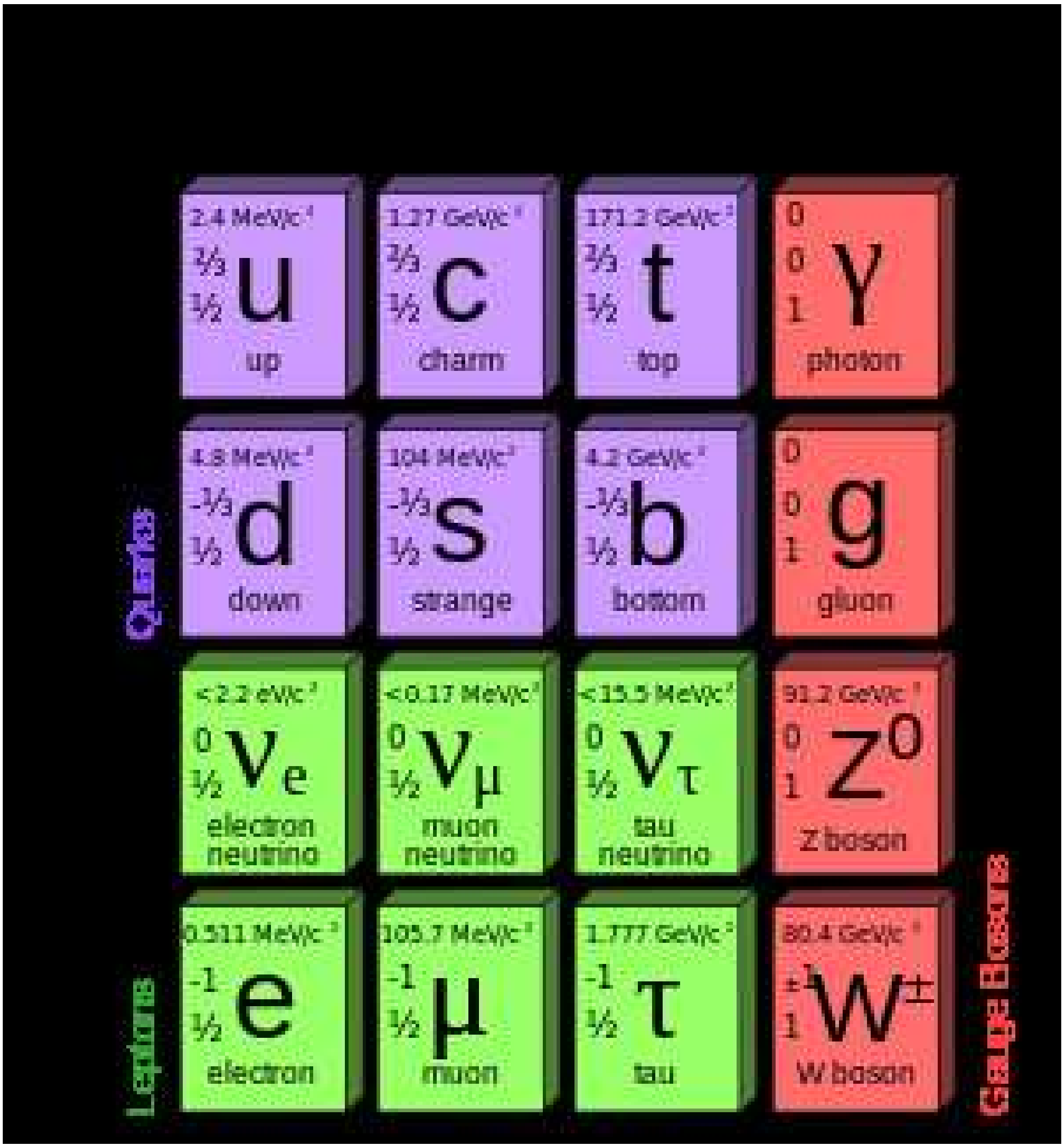


Figure 1.1: Elementary particles in the Standard Model.

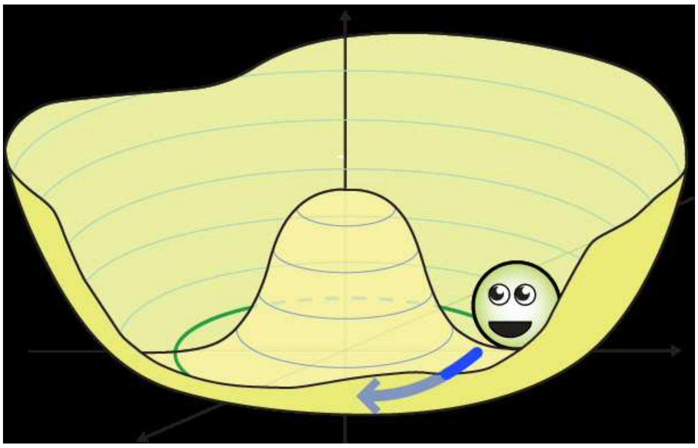
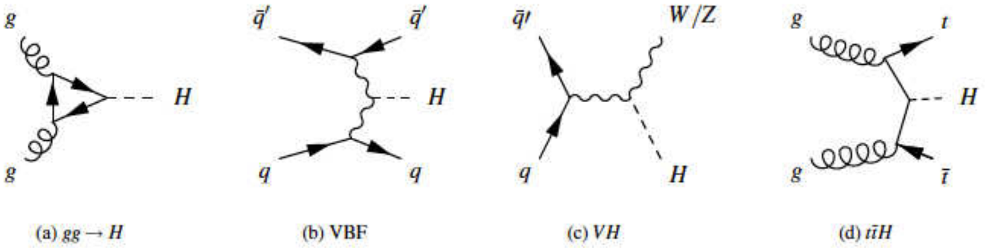


Figure 1.2: Illustration of the Higgs potential. Multiple minimum potentials exist in this field.



(a) $gg \to H$ (b) VBF (c) VH (d) $t\bar{t}H$

Figure 1.3: Feynman diagrams of the Higgs production.

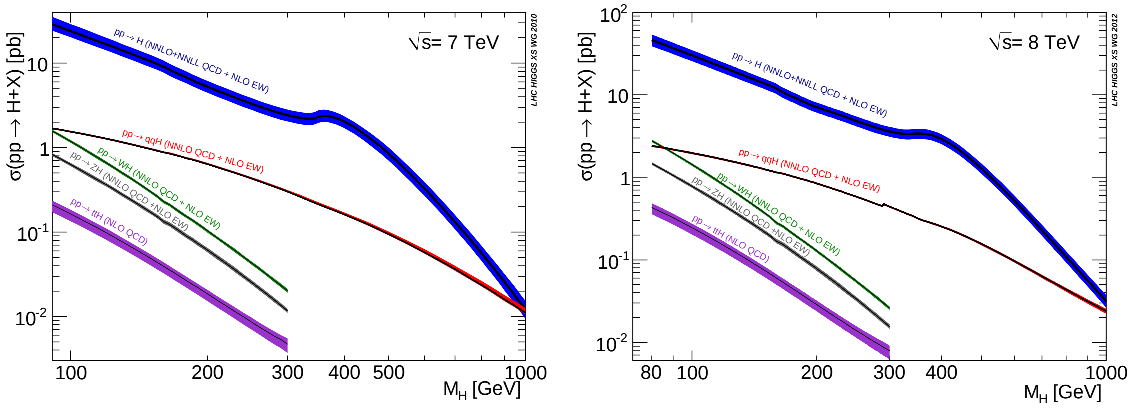


Figure 1.4: Production cross sections as a function of Higgs mass at the LHC at a center-of-mass energy of 7 TeV (left) and 8 TeV (right).

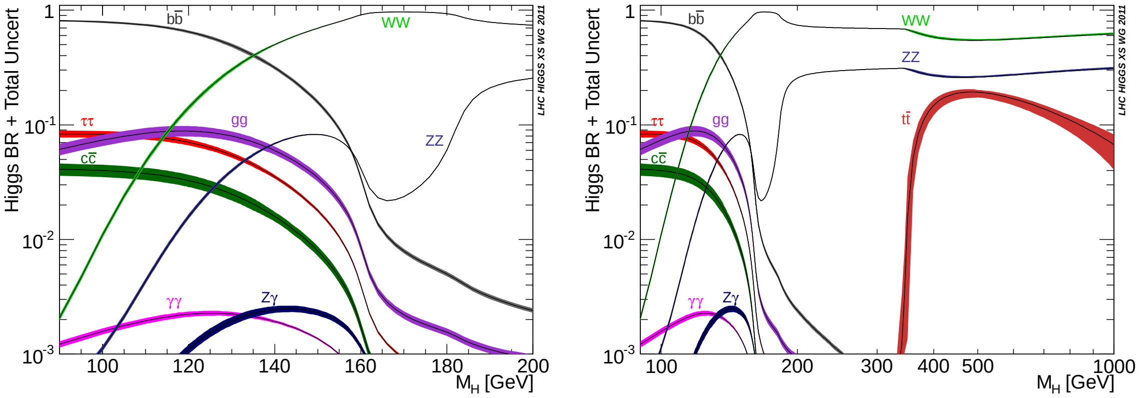


Figure 1.5: Higgs decay branching ratios and its error bands as a function of Higgs mass at the LHC. The left plot zooms to 90-200 GeV and the right one has a full range.

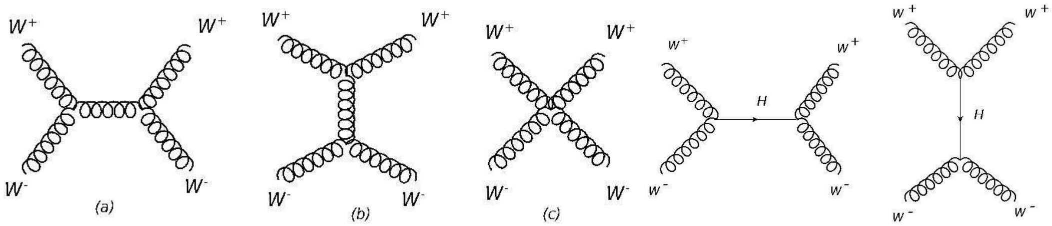


Figure 1.6: Feynman diagrams for WW elastic scattering. The left three plots are nominal WW interacton and the right two plots are the Higgs mediated contribution.

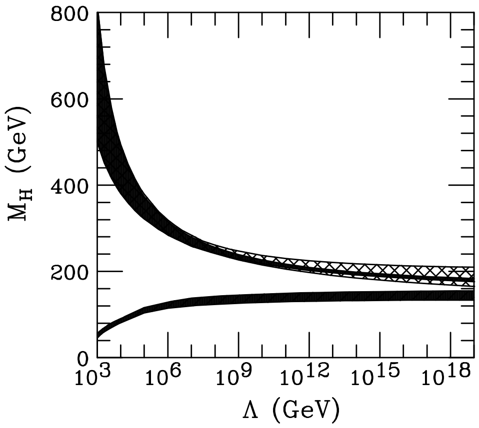


Figure 1.7: Upper and lower bounds on the Higgs mass as a function of cut off energy Λ. The bands indicate the theoretical uncertainties.

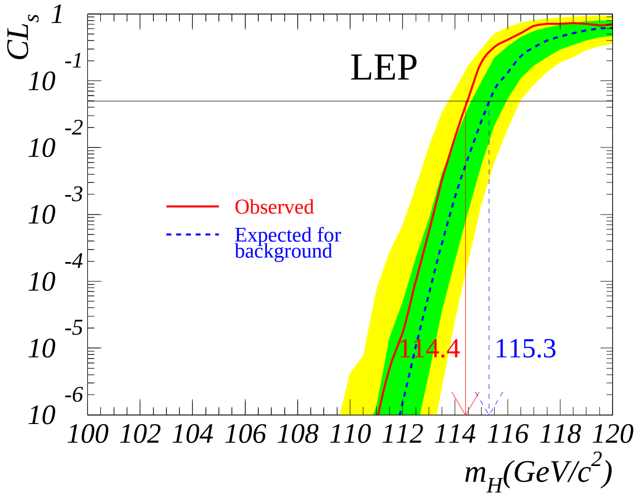


Figure 1.8: Observed exclusion limits (black line) in the Higgs search at LEP [41]. The Higgs boson with mass smaller than 114.4 GeV is excluded at 95% CL. The expected limits are also shown in the dash line.

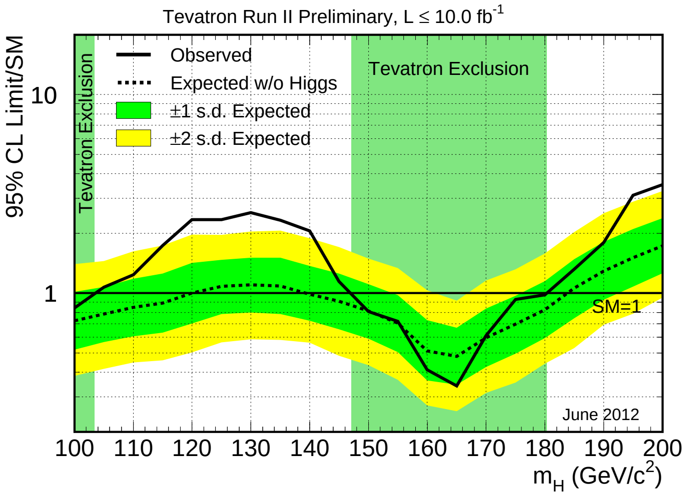


Figure 1.9: Observed and expected (median, for the background-only hypothesis) 95% CL upper limits on the ratios to the SM cross section, as functions of the Higgs boson mass for the combined CDF and D0 analyses. The bands indicate the 68% and 95% probability regions where the limits can fluctuate, in the absence of signal.

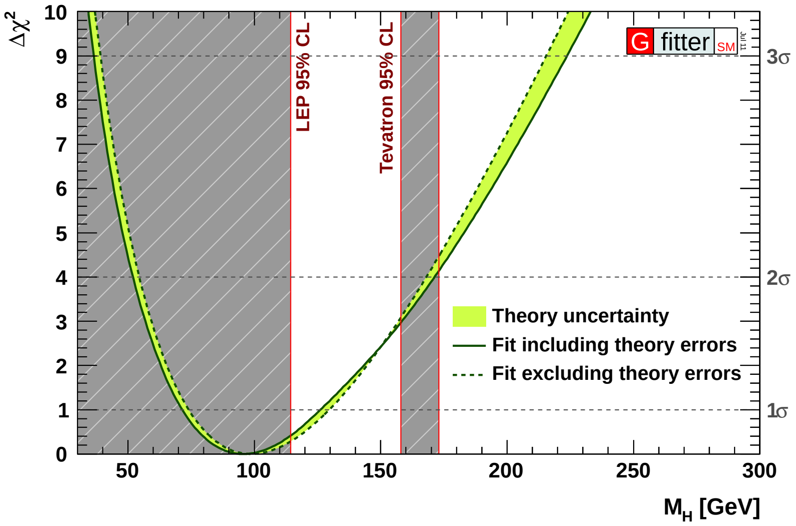


Figure 1.10: $\Delta\chi^2$ as a function of m_H. The grey bands are exluded by Tevatron and LEP direct searches.

Chapter 2

The Large Hadron Collider and ATLAS Detector

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [48] is the largest experimental device ever built in the world. It is located across the border of Swiss and France, mainly in a 90 meters depth tunnel with 27 km perimeter, which was used by the former LEP (Large Electron Positron Collider). Unlike an electron-positron collider or a proton-antiproton collider, the LHC is designed as a proton-proton collider to achieve a very high energy beam with relatively lower synchrotron radiation in the arcs. It is designed to have a nominal center-of-mass energy of 14 TeV with a peak instantaneous luminosity of $10^{34}\,\mathrm{cm^{-2}s^{-1}}$. To achieve such high energy in the circular tunnel, the high qualified bending superconducting magnets and the state of the art technology are required. The classical Nb-Ti superconducting magnets were used in a two-in-one mode in the narrow LEP tunnel. The magnets are cooled below 2 K, by using the superfluid helium, to provide up to 8.5 T central magnetic field. Besides the energy, the other important feature of the LHC is its high luminosity, which is defined as [58]:

$$\mathcal{L} = \frac{1}{4\pi}\frac{N_1 \cdot N_2 \cdot F}{\sigma_x \cdot \sigma_y \cdot t} \tag{2.1}$$

in which N_1, N_2 are the numbers of protons in the two colliding bunches, σ_x, σ_y characterize the widths of horizontal and vertical beam profile, t is the time interval between bunches and F is the geometric luminosity reduction factor due to the crossing angle at the interaction point. According to the LHC design, the typical number of protons per bunch is 1.15×10^{11}, the time interval between bunches is 25 ns. The number of interactions per bunch crossing is about 25 which is dominated by soft elastic pp collisions. In addition, the LHC can provide heavy ion collision, such as lead ion, with the design luminosity of $10^{27}\,\mathrm{cm^{-2}s^{-1}}$ at 2.8 TeV per beam energy.

Four main detectors and two small detectors are installed around the LHC (Fig. 2.1). They are: ALICE (A Large Ion Collider Experiment) [49], ATLAS (A Toroidal LHC ApparatuS) [50], CMS (Compact Muon Solenoid) [51], LHCb (Large Hadron Col-

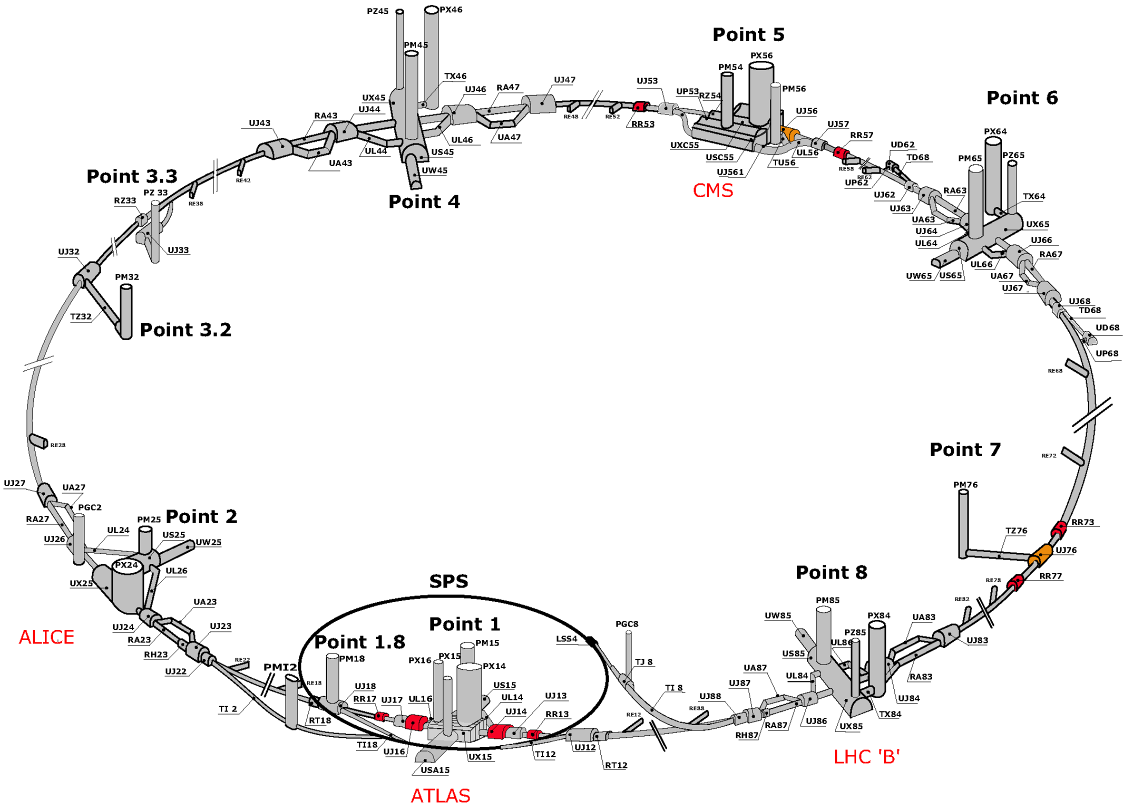


Figure 2.1: The LHC ring and its four main experiments.

lider beauty experiment) [52], LHCf [53] and TOTEM (TOTal cross section, Elastic scattering and diffraction dissociation Measurement at the LHC) [54]:

ATLAS and CMS: They are two general purpose detectors designed for testing the SM and search for the Higgs boson and new physics such as supersymmetry. The two independent detectors and measurements also allow valuable cross checks to be performed.

ALICE: ALICE is a detector concentrating on physics such as quark-gluon plasma in lead-ion collision.

LHCb: The main purpose of the LHCb experiment is the study of the matter and antimatter asymmetry in the Universe.

LHCf: The smallest LHCf experiment is intended to investigate the neutral particles such as pions, photons and neutrons produced in the very forward region of the proton-proton or nucleus-nucleus interactions. This will provide important information for the understanding the development of atmospheric showers induced by very high energy cosmic rays hitting the Earth atmosphere.

TOTEM: TOTEM is dedicated to the precise measurement of the proton-proton interaction cross section, as well as to the in-depth study of the proton structure.

On November 23, 2009, the LHC had seen its first collision at a center-of-energy of 900 GeV, one year after it circulated the first beam on September 10, 2008. Several months later, on March 30, 2010, it delivered its first high energy collision at 7 TeV, which was the highest energy ever achieved by human in history. The LHC maintained stable running in late 2010 and 2011 all year at 7 TeV with increasing luminosity. In the end of 2010, the LHC also proceeded the Pb-Pb ion run at 2.76 TeV per nucleon for heavy ion study. By the end of 2011, an integrated luminosity of about 5.6 fb^{-1} has been recorded by each of the ATLAS and CMS experiments. The integrated and peak luminosity as a function of time for four experiments are shown in Fig. 2.2. The luminosity increases swiftly in the 2011 all year as expectation. Moreover, in 2012, the LHC raised its energy to 8 TeV providing higher sensitivity for Higgs and new physics searches. By the end of June in 2012, about 5.8 fb^{-1} high quality data was accumulated by both experiments.

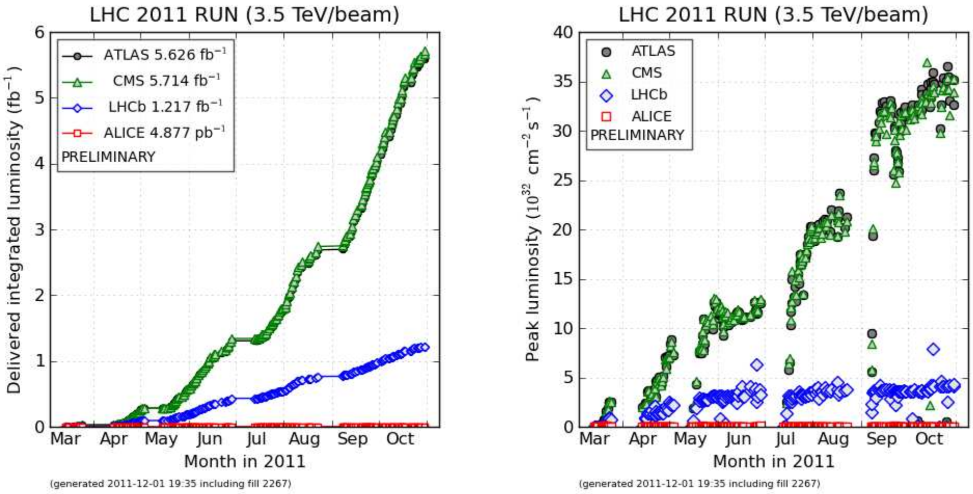


Figure 2.2: Integrated luminosity delivred by LHC as a function of time for different experiments (left). Peak luminosity as a function of time at different experiments (right).

2.2 ATLAS Detector

The ATLAS detector is the largest detector installed at the LHC, which is located at Meyrin, Swiss, 90 m deep underground. The ATLAS detector is used for general physical purpose, resulting in different considerations of complicity and accuracy. It should have a full coverage up to 4π and very good resolution on the energy and momentum measurement. Also to identify different particles, multiple subdetectors are built. The detector consists of the inner detector, electromagnetic calorimeter, hadronic calorimeter, muon spectrometer and magnet system, as shown in Fig. 2.3.

The coordinate system is defined as follows: the beam direction is defined as z axis, the x axis is pointing to the center of the LHC ring and the y axis is pointing

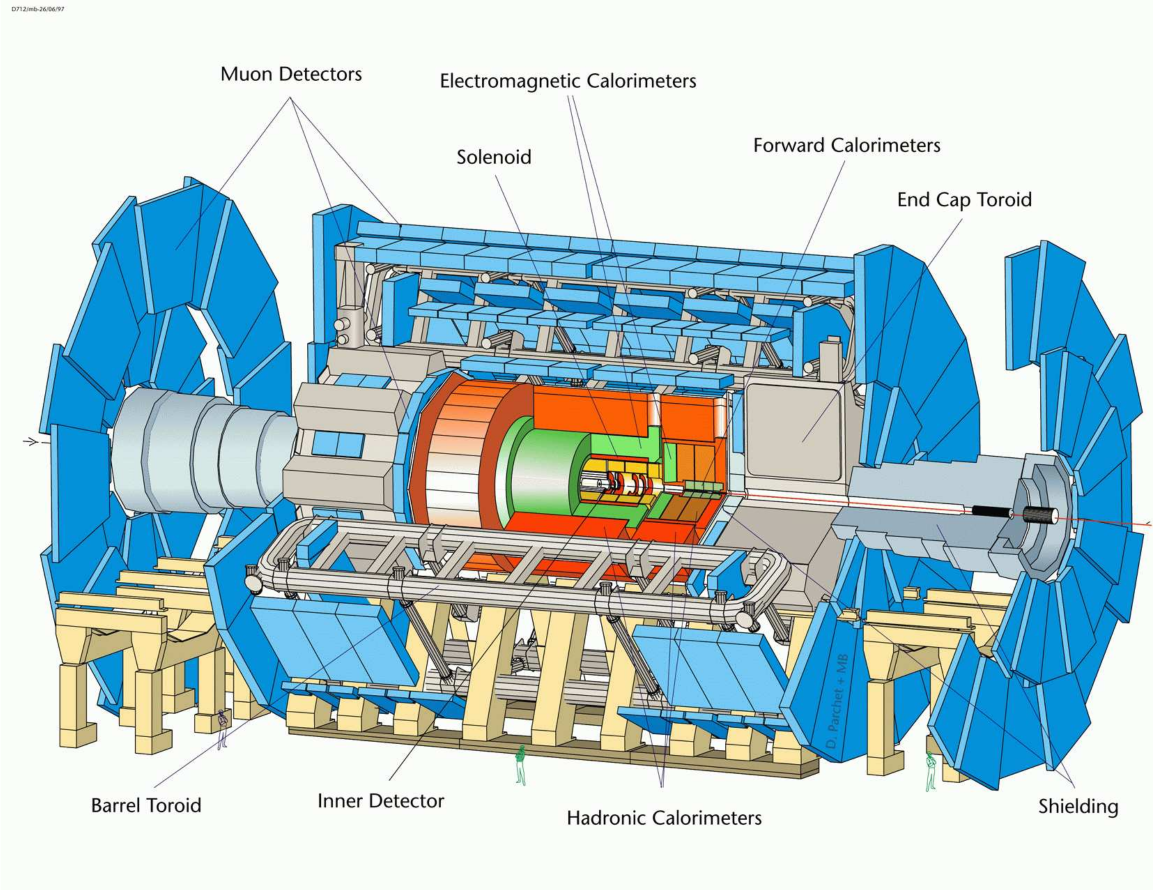


Figure 2.3: Schematic view of the ATLAS detector.

upward. The azimuthal angle ϕ is defined in the $x-y$ plane. As it is the pp collision, the pseudorapidity is defined as $\eta = -\ln\tan(\frac{\theta}{2})$. Five helix parameters measured at the point on the closest approach to the nominal beam axis $x = 0$, $y = 0$ are used to describe a charged particle:

- $1/p_{\mathrm{T}}$: Reciprocal of transverse momentum with respect to the beam axis.
- ϕ: Azimuthal angle $\tan\phi = p_y/p_x$, p_x and p_y being the x and y components of the momentum vector, respectively.
- d_0: Transverse impact parameter representing the transverse distance to the beam axis at the closest approach point. The sign is defined by the angular momentum of the track about the axis.
- z_0: Longitudinal impact parameter at the closest approach point.
- $\cot_0$: Cotangent of the polar angle, $\cot\theta = p_z/p_{\mathrm{T}}$ with p_z being the z component of the moment vector.

The detector is separated to two sides. The side A is the part of positive z and the side C is the other part, while B is the plane at $z = 0$.

2.2.1 Magnet system

Uniform, stable and strong magnetic field is very important for momentum measurement of particles. The ATLAS superconducting magnet system is composed of three parts: the central solenoid (CS) [55], the barrel toroid (BT) [56] and the end-cap toroids (ECT) [57]. The CS is implemented along the beam axis outside the inner detector and in front of the EM calorimeter, providing 2 T (2.6 T peak value) uniform magnetic field to the inner detector, sharing the vacuum vessel with the calorimeter to reduce material influences on the calorimeter by eliminating two vacuum walls. The radiation length given by materials of the total magnet part in CS is 0.66. The two ECT parts are arranged on both sides of the detector. Each of them consists of eight racetrack like coils symmetrically around the beam axis, by rotating 22.5° with respect to the BT coil system to provide radial overlap in the interface regions of both coil systems. The BT also has eight racetrack like coils in the out most region. The BT and ECT provide 3.9 T and 4.1 T peak field, respectively. The geometric property and the field strength of the ATLAS magnet system is shown in Tab. 2.1.

Property	Unit	Barrel Toroid	End-Cap Toroid(one)	Central Solenoid
Inner diameter	m	9.4	1.65	2.44
Outer diameter	m	20.1	10.7	2.63
Axial length	m	25.3	5	5.3
Number of coils	-	8	8	1
Peak field	T	3.9	4.1	2.6

Table 2.1: Geometric property and field strengh of the ATLAS magnet system.

2.2.2 Inner detector

The Inner Detector (ID) [58, 59] provides high granularity and resolution for vertex and track momentum measurement even in the condition with very large track density according to the high luminosity in the LHC. The ID consists of three parts: pixel detector, silicon strips (SemiConductor Tracker, SCT) and straw tube tracker (Transition Radiation Tracker, TRT). The resolution parameters and the acceptance in η of these detectors are shown in Tab. 2.2. The corresponding layout is shown in Fig. 2.4.

The pixel detector is the inner most detector, which provides very high granularity for position measurement. This layout mainly determines the impact parameter resolution and gives ability of identifying short-lived particles such as B and τ. There are three pixel layers in the barrel part, located at average radii of 4 cm, 10 cm and 13 cm and 5 disks on each side between 11 and 20 cm. Each pixel layer gives 1.7% radiation length.

Outside the pixel detector, it is the SCT layers, which provides 8 hits per track for momentum, impact parameter and vertex position measurements. In the barrel part, detectors are arranged in four modules in the radial direction, located at 30, 37.3, 44.7 and 52.0 cm. In the end-cap part, there are 9 wheels located from 85.4 to

System	**Position**	**Resolution** $\sigma(\mu m)$	η **coverage**
Pixels	1 removable barrel layer (B-layer)	$R\phi = 12,\ z = 66$	± 2.5
	2 barrel layer	$R\phi = 12,\ z = 66$	± 1.7
	5 end-cap disks on each side	$R\phi = 12,\ R = 77$	$1.7 - 2.5$
Silicon strips	4 barrel layers	$R\phi = 16,\ z = 580$	± 1.4
	9 end-cap wheels on each side	$R\phi = 16,\ R = 580$	$1.4 - 2.5$
TRT	Axial barrel straws	170 (per straw)	± 0.7
	Radial end-cap straws	170 (per straw)	$0.71.4 - 2.5$

Table 2.2: Resolution parameters and η acceptance of the ID subdetectors.

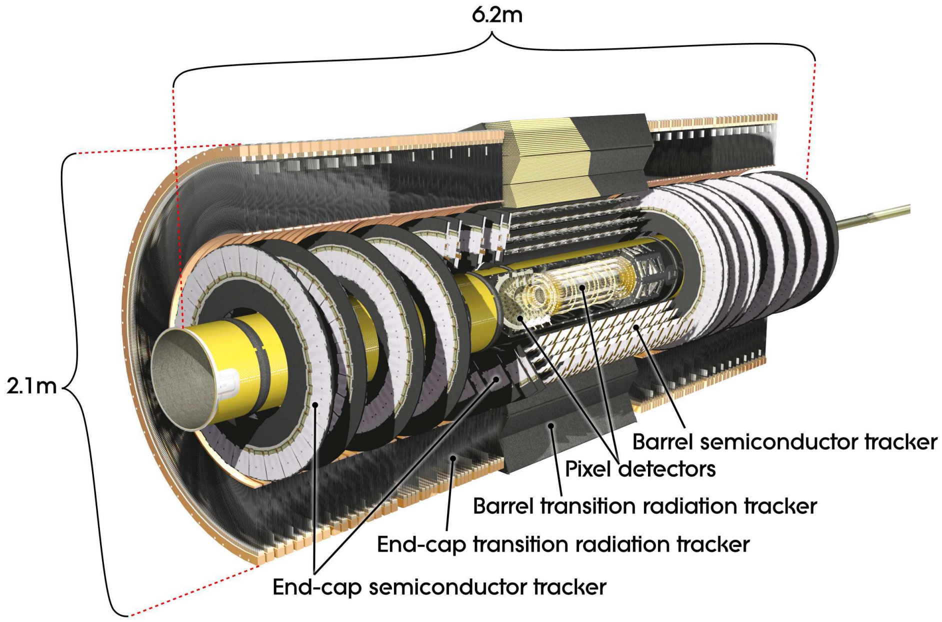


Figure 2.4: Structure of the inner detector.

272.0 cm in z direction. The total radiation length from all strips and other materials gives 2.16% at $\eta = 0$ and $1.58\% - 2.32\%$ for 9 wheels. There are 6.2 million readout channels in all.

Then follows the TRT, which is a set of straw tubes like capacitors placed within $30\,\mu m$ diameter gold-plated W-Re sense wires and filled with mixture gas: 70% Xe , 20% CO_2 and 10% CF_4. The maximum length of the straw is 144 cm and the diameter is 4 mm. The barrel part covers the radii from 56 to 107 cm, while the end-cap part covers the radii from 64 to 103 cm with 14 wheels nearest the interaction point and then extended to 48 cm in the last four wheels, which provides an acceptance in the region $|\eta| < 2.0$. There are 370 000 straw tubes in total. A typical track goes across 36 straws at least in the transverse plane. The TRT also provides the separation between electrons and pions by using the different behavior of emission of transition radiation photons [60]. For a several GeV case, electrons deposition in 7 straws under the TRT threshold is around 7 keV, while pions can achieve such energy in one or two straws.

The most relevant performance in the inner detector concerns the following aspects: measurements of vertices, the track parameters and the track reconstruction efficiency. The latter is related to the reconstruction of leptons, the isolation performance and the p_{T} calibration. The distributions of transverse impact parameter and longitudinal impact parameter multiplied by $\sin(\theta)$ are shown in Fig. 2.5 [61]. Figure 2.6 shows the expected track reconstruction efficiency measured as a function of η and p_{T} [61], in which the track reconstruction efficiency in the central region and with high p_{T} reaches 80%. The measured resolution for vertex position using 2011 data is shown in Fig. 2.7 [62], agreement between data and MC is obtained in these plots.

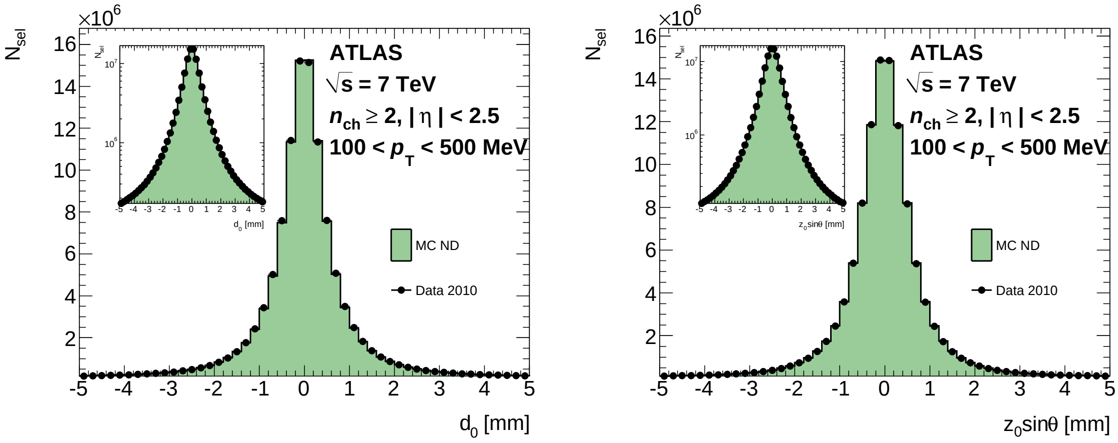


Figure 2.5: Comparison between data and simulation at $\sqrt{s} = 7\,\mathrm{TeV}$ for tracks with transverse momentum between 100 and 500 MeV: the transverse impact parameter (left) and longitudinal impact parameter multiplied by $\sin(\theta)$ (right). The inserts for the impact parameter plots show the log-scale plots. The p_{T} distribution of the tracks in non-diffractive (ND) MC is re-weighted to match the data and the number of events is scaled to the data.

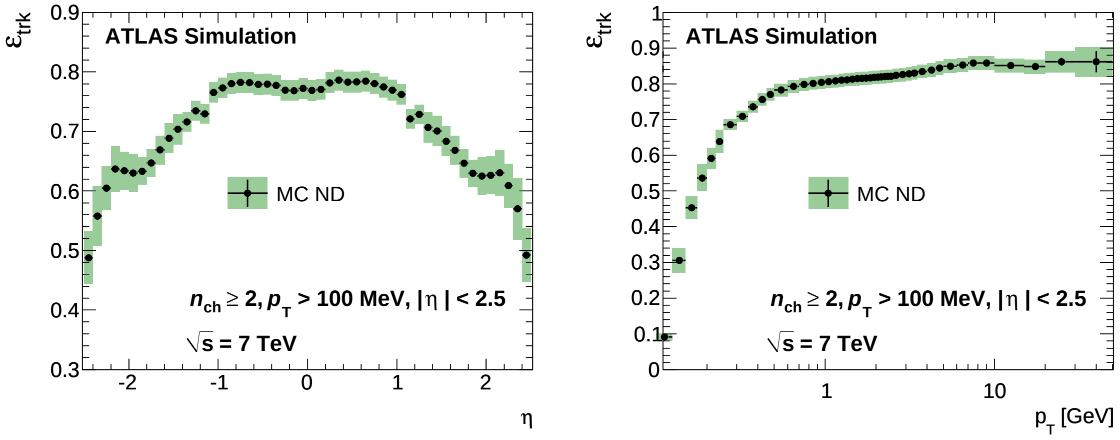


Figure 2.6: Track reconstruction efficiency based on non-diffractive (ND) MC shown as a function of η (left) and p_{T} (right). The statistical errors are shown as black lines, the total errors as green shaded areas. All distributions are shown at $\sqrt{s} = 7\,\mathrm{TeV}$ for number of charged particles greater than 2, $p_{\mathrm{T}} > 100\,\mathrm{MeV}$, $|\eta| < 2.5$.

2.2.3 Calorimeter

The ATLAS calorimeter [63, 64] consists of two parts, the electromagnetic (EM) calorimeter and the hadronic one. The EM calorimeter is very important for precise energy measurement of photons and electrons while the hadronic calorimeter will provide the jets energy information.

The EM calorimeter consists of the barrel part (EMB) ($|\eta| < 1.475$) and two end cap parts (EMEC) on each side of the barrel ($|\eta| < 3.2$). The EM calorimeters are sampling calorimeter using liquid argon as active material, filled in 2.09 mm gaps between the lead absorber. The calorimeter is designed as accordion shape to provide redundant coverage in the ϕ direction. It consists of three layers and a presampler outside the cryostat to correct energy lost in the material in front of calorimeter, as shown in Fig. 2.8. Each layer of the EM calorimeter has thousands of cells as the minimum units for energy measurement and providing high granularity ($\delta\eta \times \delta\phi$), as shown in Tab. 2.3. The front layer (sampling 1) has the best granularity in η while the middle layer (sampling 2) provides the best ϕ positioning. The radiation length in the barrel part is $24\,X_0$, dominated by the middle layer. In the end-cap region the radiation length is $26\,X_0$.

The energy in the EM calorimeter is calculated by measuring ADC signals in each cell and summed by layers after calibration, as shown:

$$E_{\mathrm{tot}} = w_{\mathrm{glob}}(w_{\mathrm{ps}} E_{\mathrm{ps}} + E_{\mathrm{front}} + E_{\mathrm{mid}} + E_{\mathrm{back}})\,. \tag{2.2}$$

The presampler weight w_{ps} is used to optimise the energy resolution. The corresponding energy resolution is generally expressed as:

$$\frac{\sigma}{E} = \frac{a}{E} \oplus \frac{b}{\sqrt{E}} \oplus c \tag{2.3}$$

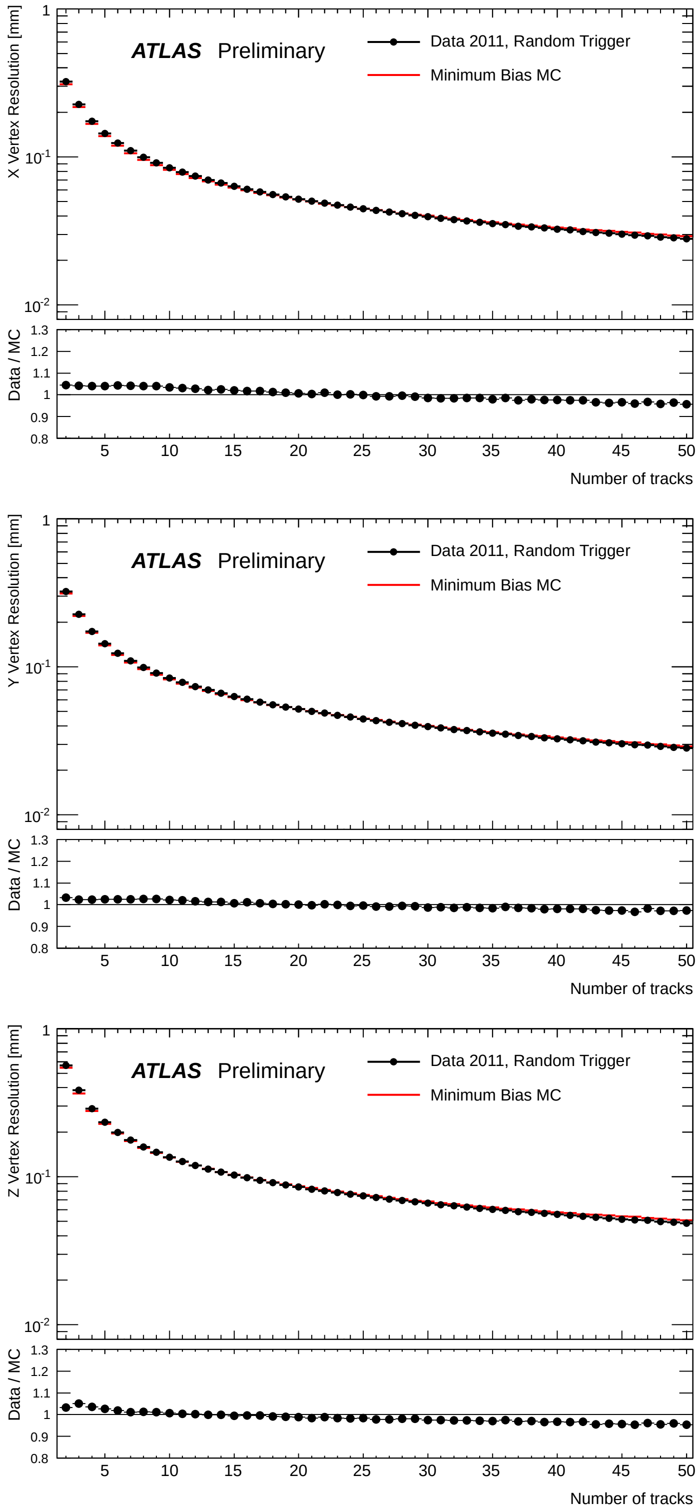


Figure 2.7: Measured resolution of vertex position as a function of number of tracks attached to this vertex. The resolution is shown in x (top), y (middle), z (bottom).

	Barrel		End-cap	
EM calorimeter				
Number of layers and $\lvert\eta\rvert$ coverage				
Presampler	1	$\lvert\eta\rvert < 1.52$	1	$1.5 < \lvert\eta\rvert < 1.8$
Calorimeter	3	$\lvert\eta\rvert < 1.35$	2	$1.375 < \lvert\eta\rvert < 1.5$
	2	$1.35 < \lvert\eta\rvert < 1.475$	3	$1.5 < \lvert\eta\rvert < 2.5$
			2	$2.5 < \lvert\eta\rvert < 3.2$
Granularity $\Delta\eta \times \Delta\phi$ versus $\lvert\eta\rvert$				
Presampler	0.025×0.1	$\lvert\eta\rvert < 1.52$	0.025×0.1	$1.5 < \lvert\eta\rvert < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$\lvert\eta\rvert < 1.40$	0.050×0.1	$1.375 < \lvert\eta\rvert < 1.425$
	0.025×0.025	$1.40 < \lvert\eta\rvert < 1.475$	0.025×0.1	$1.425 < \lvert\eta\rvert < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \lvert\eta\rvert < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \lvert\eta\rvert < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \lvert\eta\rvert < 2.4$
			0.025×0.1	$2.4 < \lvert\eta\rvert < 2.5$
			0.1×0.1	$2.5 < \lvert\eta\rvert < 3.2$
Calorimeter 2nd layer	0.025×0.025	$\lvert\eta\rvert < 1.40$	0.05×0.025	$1.375 < \lvert\eta\rvert < 1.425$
	0.075×0.025	$1.40 < \lvert\eta\rvert < 1.475$	0.025×0.025	$1.425 < \lvert\eta\rvert < 2.5$
			0.1×0.1	$2.5 < \lvert\eta\rvert < 3.2$
Calorimeter 3rd layer	0.05×0.025	$\lvert\eta\rvert < 1.35$	0.05×0.025	$1.5 < \lvert\eta\rvert < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	
LAr hadronic end-cap				
$\lvert\eta\rvert$ coverage			$1.5 < \lvert\eta\rvert < 3.2$	
Number of layers			4	
Granularity $\Delta\eta \times \Delta\phi$			0.1×0.1	$1.5 < \lvert\eta\rvert < 2.5$
			0.2×0.2	$2.5 < \lvert\eta\rvert < 3.2$
Readout channels			5632 (both sides)	
LAr forward calorimeter				
$\lvert\eta\rvert$ coverage			$3.1 < \lvert\eta\rvert < 4.9$	
Number of layers			3	
Granularity $\Delta x \times \Delta y$ (cm)			$FCal1$: 3.0×2.6	$3.15 < \lvert\eta\rvert < 4.30$
			FCal1:$\sim 4\times$ finer	$3.10 < \lvert\eta\rvert < 3.15$
				$4.30 < \lvert\eta\rvert < 4.83$
			FCal2: 3.3×4.2	$3.24 < \lvert\eta\rvert < 4.50$
			FCal2: $\sim 4\times$ finer	$3.20 < \lvert\eta\rvert < 3.24$
				$4.50 < \lvert\eta\rvert < 4.81$
			FCal3: 3.3×4.2	$3.32 < \lvert\eta\rvert < 4.60$
			FCal3: $\sim 4\times$ finer	$3.29 < \lvert\eta\rvert < 3.32$
				$4.60 < \lvert\eta\rvert < 4.75$
Readout channels			3524 (both sides)	
Scintillator tile calorimeter				
	Barrel		Extended barrel	
$\lvert\eta\rvert$ coverage	$\lvert\eta\rvert < 1.0$		$0.8 < \lvert\eta\rvert < 1.7$	
Number of layers	3		3	
Granularity $\Delta\eta \times \Delta\phi$	0.1×0.1		0.1×0.1	
Last layer	0.2×0.1		0.2×0.1	
Readout channels	5760		4092 (both sides)	

Table 2.3: Main parameters of the calorimeter system.

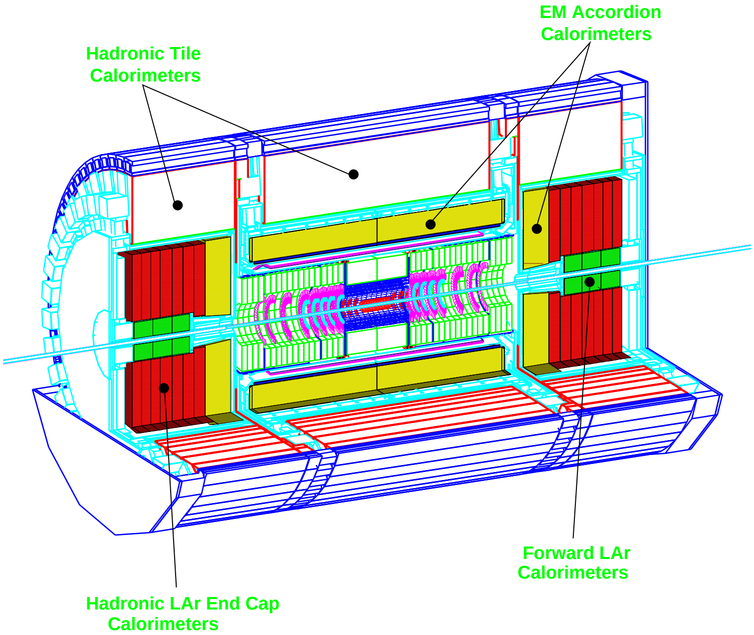


Figure 2.8: Schematic view of the ATLAS calorimeters.

in which a, b and c represent the noise term, the sampling term and the constant term, respectively. The expected resolution is $\sigma/E = 10\%/\sqrt{E[\mathrm{GeV}]} \oplus 0.7\%$ [65].

The hadronic calorimeter, which is used to complement the EM calorimeter for jet energy and $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement, consists in three parts: tile calorimeter, liquid Argon end-cap calorimeter and forward calorimeter. The tile calorimeter uses iron as absorber and scintillating tiles as the active material and consists of a central barrel and two extended barrels, providing η coverage to $|\eta| < 1.7$. The end-cap calorimeter and the forward region calorimeter use the same material as the EM calorimeter extending to $|\eta| < 3.2$ and $|\eta| < 4.9$. The resolution of the tile calorimeter for hadrons is $\sigma/E = 50\%/\sqrt{E[\mathrm{GeV}]} \oplus 3\%$ for the central region and $\sigma/E = 100\%/\sqrt{E[\mathrm{GeV}]} \oplus 10\%$ for the forward region.

2.2.4 Muon Spectrometer

The out most component of the ATLAS detector is the Muon spectrometer (MS) [66, 67] as shown in Fig. 2.9. There are eight superconducting barrel toroids around the beam axis and two end-cap toroid magnets in each side, providing the magnetic field circling the beam, which has different direction from magnetic field in the inner detector. Three layer barrel chambers and four layer end-cap wheels in each side provide the measurement of position, which is used for the muon track reconstruction. These chambers are partially overlapped near the boundary to avoid crack regions.

Monitored drift-tube chambers (MDT) performing the momentum measurement in the barrel part cover the pseudorapidity $|\eta| < 2$ in the inner most layer and cover $|\eta| < 2.7$ in the other layers. The Cathode strip chambers (CSCs) are used as complementary in the region $2 < |\eta| < 2.7$ due to their good time resolution and high rate capacity. The transverse momentum resolution combining ID and MS as a function of transverse momentum is shown in Fig. 2.10 [68]. It is about 3% in the interesting muon p_{T} region (not greater than 100 GeV). Moreover, to have a quick response against high rate interaction in the beam pipe, the Resistive Plate Chambers (RPC) and the Thin Gap Chambers (TGC) were chosen as the fast trigger system in the barrel ($|\eta| < 1.05$) and end-cap ($1.05 < |\eta| < 2.4$).

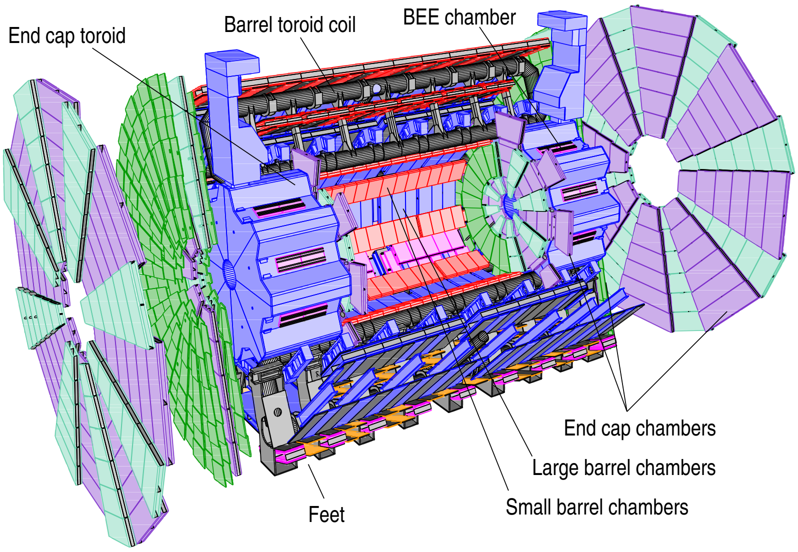


Figure 2.9: Structure of the Muon spectrometer in the ATLAS detector.

2.2.5 The ATLAS trigger system

Due to extremely high interaction rate at the LHC, fast and efficient trigger system is necessary for data taking [69]. There are three level trigger in ATLAS: level 1(L1), level 2(L2) and event filter (EF), the latter two together are also called High Level Trigger (HLT). Trigger criteria, combined for all purpose study, are arranged as the trigger menu.

The L1 trigger is hardware stage judgment to make fast decision at the very high collision rate at up to about 40 MHz. In the muon spectrometer, the RPC and TGC are used to achieve this task at very high speed but with limited accuracy. In EM calorimeter, a larger trigger tower $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ is used for E_T measurement. The L1 trigger provides the 'Region-of-Interest' (RoI) information, which includes some kinematic information of objects (electron, muon, photon etc) for further judgment. The L1 trigger finally reduces the event rate to < 75 kHz to meet the limitation on the ATLAS front-end system.

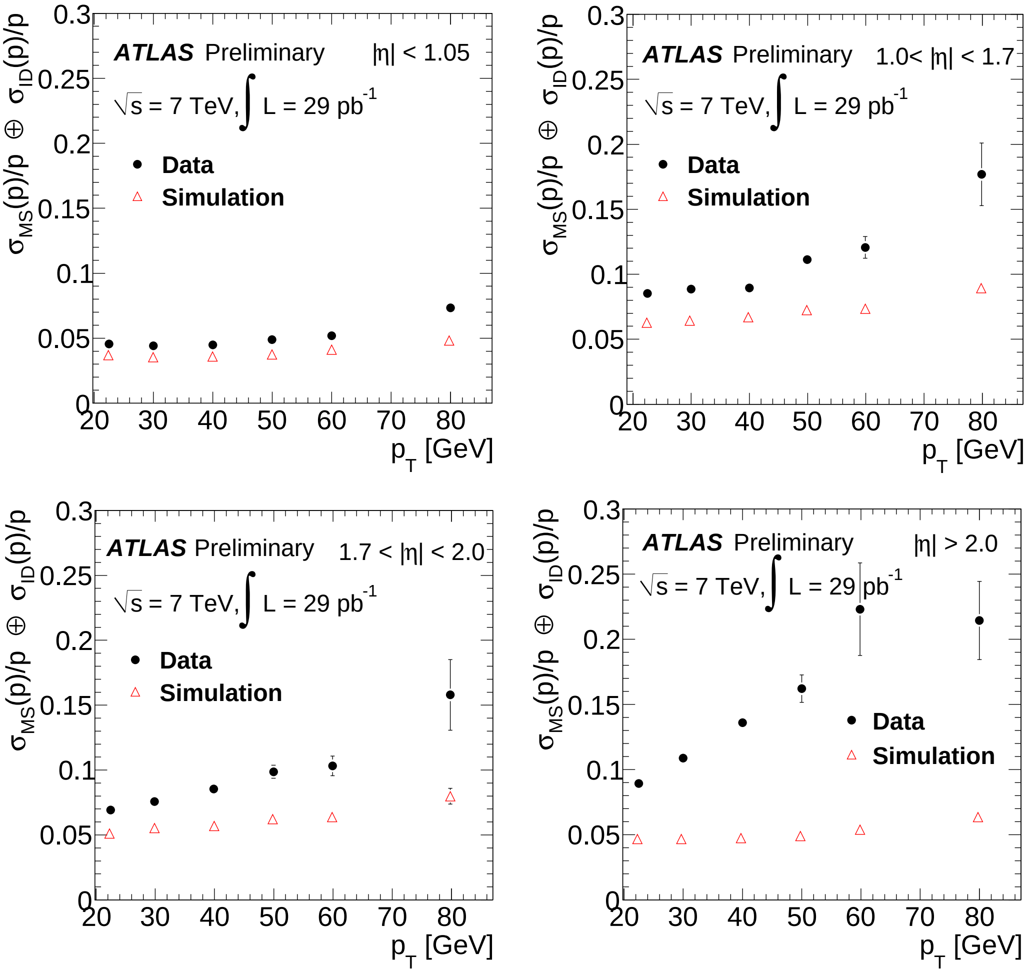


Figure 2.10: Muon p_T resolution combining contributions from ID and MS measured in data in different η region. Upper left: $|\eta| < 1.05$; Upper right: $1.05 < |\eta| < 1.70$; Lower left: $1.70 < |\eta| < 2$; Lower right: $2.0 < |\eta|$.

The L2 trigger makes use of the RoI information from the L1 trigger, with full precision and granularity. The L2 trigger reduces the event rate to 1 kHz then pass events to offline event filter. The EF trigger finally reduces the event rate to 400 Hz and at this stage the events are built and permanently stored.

2.2.6 The Luminosity Detectors

The Luminosity Cherenkov Integrating Detector (LUCID) and the ATLAS Beam Conditions Monitor (BCM) are two primary detectors used to measure the bunch-by-bunch luminosity in the ATLAS detector [70, 104]. The LUCID is located in both sides of the ATLAS detector, 17 m from the interaction point and surrounding the beam pipe. It has an array of 16 mechanically polished aluminum tubes filled with Cherenkov gas C_4F_{10} at the 1.1 bar presure. The photonmultipliers (PMT) situated at the back end of the tubes collect the Cherenkov photons generated in the tube and reflected by the tube walls. When the signal in the PMT is above some threshold, the detector records a hit. The LUCID can record event rate separately for each bunch crossing.

The Beam Conditions Monitor (BCM) is a fast device located 2 m away from the interaction point to monitor the collision condition in real time with a timing resolution of 0.7 ns and it can provide the coincidence rate per bunch crossing.

Chapter 3

Objects Reconstruction, Identification and Selection

In the ATLAS detector, the information including hits, tracks, energy deposition are reconstructed to physical objects for analysis. For instance, in the Higgs search the Higgs mass or transverse mass is calculated from energy of photon, electron, muon and/or transverse missing energy. In some cases, jets are useful for tagging events which have b quarks. In the high luminosity collision environment, thousands tracks fill the inner detector simultaneously and in the calorimeter as well, the energy deposition from different particles can be overlapped. The efficiency and the fake rate of object reconstruction/identification can significantly affect the final physical results. In the ATLAS detector, electrons, muons, photons, jets and τ are reconstructed from track, calorimeter, muon spectrometer and vertex information. Jets may be tagged as b jet. The total transverse missing energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) is calculated for study of neutrino related processes. In the Higgs searches, leptons and/or photons are required to be isolated to ensure they come directly from hard process so as to suppress background or pile up events. In this chapter, the reconstruction, identification, isolation and selection of electrons, muons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$, dedicated to Higgs to $WW^{(*)}$ analysis, is introduced.

3.1 Electrons

3.1.1 Electron Reconstruction

The electrons are reconstructed by combining tracking and calorimeter information. In the inner detector, tracks are reconstructed in the following steps [71]:

1. Hits in Pixels and SCTs are found and clustered into groups.

2. Space points are created by combining Pixel clusters and three SCT clusters from stereo-layers.

3. Using space points in a straight line, seeds for tracks are created. In this step ambiguities of close pixels are resolved. Tracks of poor quality are removed.

4. The silicon tracks are extrapolated to TRT straws, in which the timing information is interpreted into drift radius. The extended tracks are fitted again to find better track scores. A TRT seeded reconstruction is also used to find secondary tracks from long-lived particles via an out-inside procedure.

5. Finally a global χ^2 minimization, a Kalman filter and vertex fitters are implemented to finish the vertex reconstruction.

Then a cluster based algorithm are performed to reconstruct the electrons [72]. Energy deposits in the electromagnetic calorimeter are used to form energy clusters, using a sliding window clustering method. The cells of the calorimeter are shown in Fig. 3.1. In this algorithm the $\eta - \phi$ space is categorized into a grid of $N_\phi \times N_\eta$ elements ($\Delta\phi \times \Delta\eta = 0.025 \times 0.025$). In the EM calorimeter, 256 bins in ϕ and 200 bins in η from -2.5 to 2.5 are defined. In each bin, energies of all cells across the longitudinal layers are summed as the tower energy. Then a fixed size window (nominal $N_\phi \times N_\eta = 5 \times 5$) is used to define a pre-cluster: if the transverse energy deposited in the window is above the threshold $E_T^{\text{thresh}} = 3\,\text{GeV}$, to reject noise, a smaller size window (nominal $N_\phi \times N_\eta = 3 \times 3$) is then defined to locate the barycenter of the deposited energy as the seed. Finally, a duplicate removal algorithm is performed in the range $\Delta N_\phi \times \Delta N_\eta = 2 \times 2$ to reject overlapped pre-cluster with smaller transverse energy. The seed found in the pre-cluster is used to reconstruct the final EM clusters. All of the cells within a given $\eta - \phi$ range to the seed are filled into the final cluster. The size of the clusters for different egamma candidates is shown in Tab. 3.1. After reconstructing the clusters, the shower shapes are calculated. Both the non-TRT-only tracks and the TRT only tracks are extrapolated to the EM calorimeter to match with the clusters. Only one track matched cluster is stored as an electron. The energy in each cluster is corrected by taking into account the leakage outside the window and also the losses in the crack scintillators. The tracks are refitted by considering bremsstrahlung. Thus, in each electron object, two sets of 4-vectors are filled and the four momentum vectors are set by cluster/track combination. Usually the track 4-vectors are used to define the direction and the information in cluster 4-vectors is used to provide energy measurement.

Particle type	Barrel	End-cap
Electron	$N_\phi \times N_\eta = 3 \times 7$	$N_\phi \times N_\eta = 5 \times 5$
Photon-converted	$N_\phi \times N_\eta = 3 \times 5$	$N_\phi \times N_\eta = 5 \times 5$
Photon-unconverted	$N_\phi \times N_\eta = 3 \times 5$	$N_\phi \times N_\eta = 5 \times 5$

Table 3.1: Various cluster sizes for different particle types and calorimeter regions.

3.1.2 Electron Identification and Selection in Higgs Analyses

Electron Identification

To identify an electron, several levels of selection criteria are provided in the electron, called author and isEM criteria. The author value is obtained from different types of

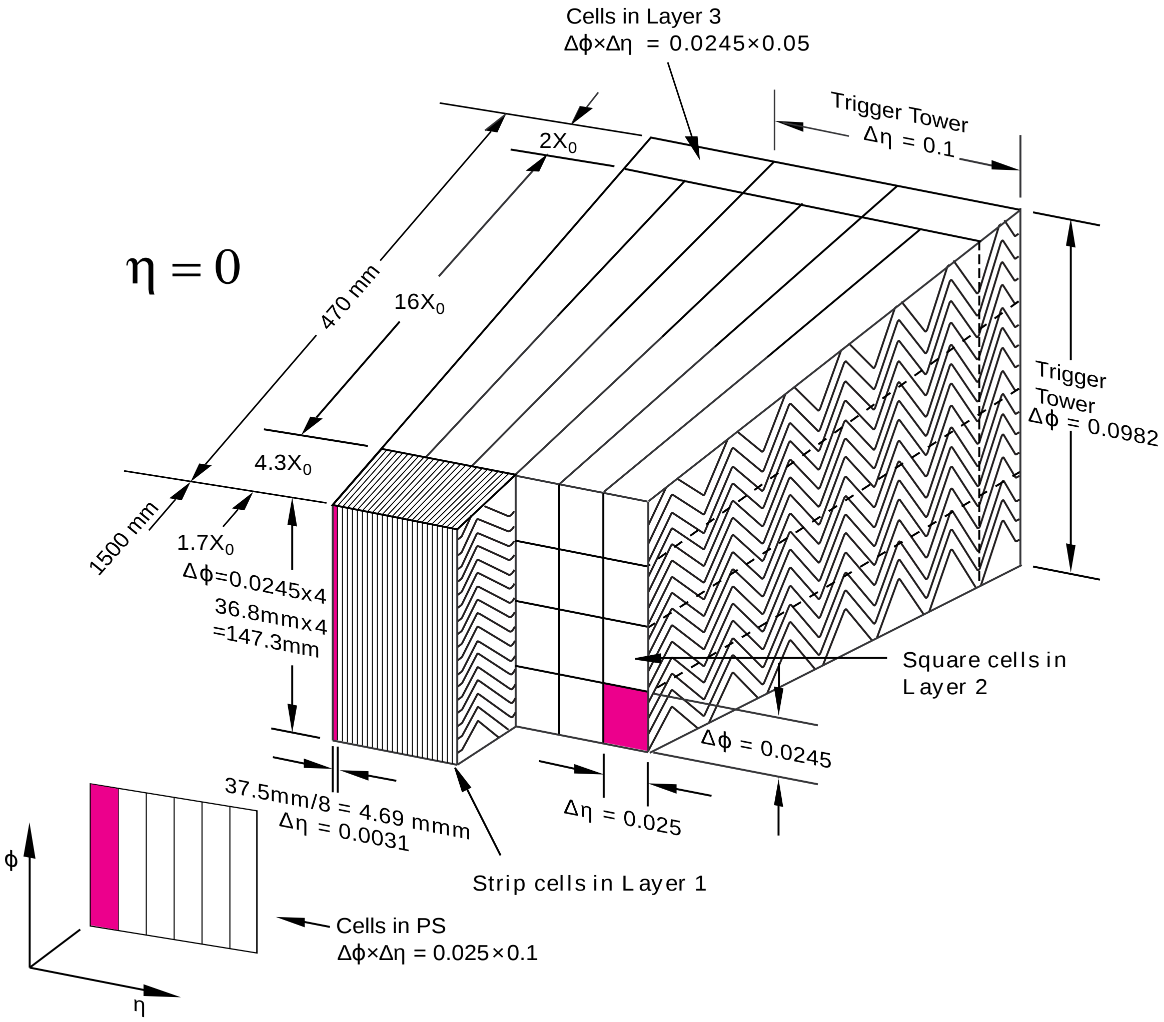


Figure 3.1: Structure and cell size in the EM calorimeter.

independent reconstruction algorithms in the egamma object to categorize the object into an electron, a soft electron, a forward electron, a photon and a converted photon. The isEM variable has three menus: loose, medium and tight (called tight++ for 2011 studies), which correspond to a combination of track and EM shower shape criteria. The loose (loose++) provides 95% efficiency measured by the tag-and-probe method in the Z mass window (typically ± 15 GeV around Z mass). Similarly, the efficiency for the medium (medium++) and tight (tight++) is 85% and 78%, respectively. The content of tight (tight++) menu is shown in Tab. 3.2.

Electron Selection in $H \to WW^{(*)}$ Analysis

In the Higgs to $WW^{(*)}$ analysis, the electron selection follows the recommendation of the egamma group. Electrons are required as prompt and isolated. The following selection criteria are applied as shown in Tab. 3.3.

In the 2011 data analysis the period dependent trigger requirements EF_e20_medium,

Kinematic variables: $e_T > 5\,\mathrm{GeV}$; $\lvert\eta\rvert < 2.47$
Calorimeter based cuts:
Hadronic leakage
Ratio e237/e277
Shower width in 2nd sampling
Total width in 1st sampling
Cut on $(E_{\max} - E_{\max 2})/(E_{\max} + E_{\max 2})$ in the 1st sampling
Cuts on track quality:
Pixel hits, SCT hits
B-Layer hits
Outliers for B-layer, Pix, SCT, TRT tracks
TRT hits and ratio with outliers
Transverse impact parameter <1 mm $E_{\mathrm{calo}}/E_{\mathrm{track}}$ cuts
Track cluster match

Table 3.2: Menu of the tight++ electron selection, cut values not shown here are $\eta - \phi$ or $\eta - p_{\mathrm{T}}$ dependent. e237 means the energy deposition in 3×7 cells area in layer 2. e277 means the energy deposition in 7×7 cells area in layer 2. $E_{\max}$ is the maximum energy deposition of a cell in the cluster in the 1st sampling and $E_{\max 2}$ is the second largest energy deposition.

EF_e22_medium and EF_e22vh medium1 are used to cope with the increasing luminosity and pile up. The periods for data are defined by different luminosity in data taking. The numbers after EF_e in the names represent the nominal p_{T} threshold values for these triggers. The suffix medium and medium1 indicate the tightness in the electron identification and vh means that the trigger has both η dependent p_{T} threshold and hadronic leakage cut at level 1. An electron trigger is required for the ee channel and either an electron or a muon trigger is required for the $e\mu$ channel. To suppress QCD contamination, the lepton transverse momentum has to be above a threshold: 25 GeV for the leading lepton and for 15 GeV for the subleading lepton. The higher lepton p_{T} threshold helps in rejecting more fake or non-prompt electrons but it also reduces the selection efficiency for low Higgs mass points below about 130 GeV. The η cut is made to avoid the crack region in the EM calorimeter. In the 2011 data taking, part of the LAr Front-End-Board opto-transmitter plug-ins (OTX) were dead and could not be replaced immediately so that events which have electrons in these regions are rejected. The tight++ identified electrons are selected. In such tight++ menu a series of optimised cuts are made to distinguish a genuine electron from a fake one by comparing the shower shape, track quality and hadronic leakage. Also the associated track of an electron candidate should have a primary vertex which has small impact parameters and their significance. The impact parameter significance is defined as the parameter such as d_0 and z_0 over their errors. To further isolate the electrons, a set of cluster and track isolation cuts are applied. The cluster

Trigger requirement:
B-J e20_medium K e22_medium L-M e22vh_medium1 Trigger match: e matched to EF_e20(22)_medium(1) and $p_{\mathrm{T}} > 23\,\mathrm{GeV}$
Kinematic cuts:
For all p_{T} apply electron energy correction in data and smearing in MC Author 1 or 3 p_{T}(leading) $> 25\,\mathrm{GeV}$, p_{T}(subleading) $> 15\,\mathrm{GeV}$ $\lvert\eta\rvert < 2.47$ and exclude crack region $1.37 < \lvert\eta\rvert < 1.52$ (cluster η)
Electron selection:
OTX cleaning cut, reject events if electron is in dead region Tight++ quality
Vertex requirements:
$\lvert z_0\rvert$(w.r.t.PV) $< 1\,\mathrm{mm}$ Impact parameter significance d_0(sig) < 10
Isolation requirements:
$p_{\mathrm{T}} > 15\,\mathrm{GeV}$, pile up corrected $E_{T,cone30}/p_{\mathrm{T}} < 0.14$, $P_{T,cone30}/p_{\mathrm{T}} < 0.13$

Table 3.3: Electron selection criteria in the Higgs to $WW^{(*)}$ analysis.

isolation is illustrated in the Fig. 3.2. The figure shows that the size of a cone around the fixed 5×7 window around an electron or a photon cluster. The energy deposit for an isolated electron is expected to distribute in the central of the window and have a very narrow width so that the energy deposit inside the cone but outside the window is expected to be small (etcone). The cone size used is indicated in the variable name: etcone30 and etcone40 corresponds to $\Delta R = 0.3, 0.4$, respectively. The isolation can be used to separate electrons from jets. Similarly, a track isolation can be defined by summing the transverse momentum of tracks around the electron track. These tracks are required to have the p_{T}over a given threshold and satisfy the impact parameter requirements with respect to the primary vertex so that the sensitivity to pile up effects is reduced.

Tag-and-Probe Method for Electron Identification and Trigger Efficiencies

The tag-and-probe method is widely used in determining trigger, particle identification, reconstruction and isolation efficiency in both data and MC. This method is normally performed in the $Z \to ll$ or $J/\psi \to ll$ mass windows with $l = e, \mu$. In these mass windows, the background contribution is relatively small, so that one can use one well identified lepton as the tagged lepton and the other one as the probe lepton. An example shown in [74] is presented here to explain the method for the determination of the electron identification efficiency.

In the $Z \to ee$ region, the tag electron is required to have $E_T > 20\,\mathrm{GeV}$ and to

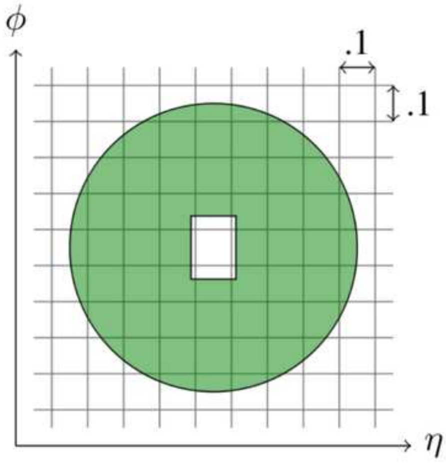


Figure 3.2: Illustration of the E_T cone in the EM calorimeter.

be matched with the corresponding trigger object. It should pass the tight identification as well. The probe electron should have opposite charge as the tagged electron and satisfy $E_T > 15\,\text{GeV}$ and $|\eta| < 2.47$. Also a $\Delta R > 0.1$ requirement between the two electrons is performed to avoid overlap. The tight requirement on the tag electron gives high electron purity. The reconstructed dielectron mass is constrained in $80 < m_{ee} < 100\,\text{GeV}$ in the $Z \to ee$ case. For electrons in $20 < E_T < 50\,\text{GeV}$, the background contamination is not large in the Z mass region and a same-sign background sample can be used to estimate the contamination in the Z mass region. In some cases, a template fit is performed to extract the number of Z events instead of direct background subtraction as shown above. The fit range is $40 < m_{ee} < 160\,\text{GeV}$ and the signal shape is modeled by a Breit-Wigner distribution convoluted by a Crystal Ball function, which describes the low-mass tail arising from material effects. Sometimes a template obtained from $Z \to ee$ MC simulation is also helpful to describe signal shape. For the other background, the shape is modeled by an exponential convoluted with a Gaussian.

One can apply such method in each p_{T}, η bin or in bins of the number of reconstructed vertices of the probe electrons and obtain the number of probe electrons before and after the electron identification. The corresponding efficiency for tight (2010) definition is shown in Fig. 3.3 [75] as a function of p_T and η. In Fig. 3.4 the identification efficiency of tight++ electron (2011) as a function of number of vertices is shown [76]. In 2012, the electron identification efficiency is improved by loosening pile up sensitive criteria and tightening pile up robust criteria in the re-optimisation. Similarly, the same method can be used to obtain the efficiencies of the electron trigger, reconstruction and isolation requirements. Figure 3.5 shows the trigger efficiency as a function of the transverse energy E_T and η of electrons [73].

The uncertainty of the efficiencies depends on the statistics in the tag-and-probe sample, background contamination, discriminating variables (e.g. invariant mass of the Z boson) and the bias introduced by the method itself. The uncertainties for reconstruction, identification and trigger are provided as a function of η and p_{T} in

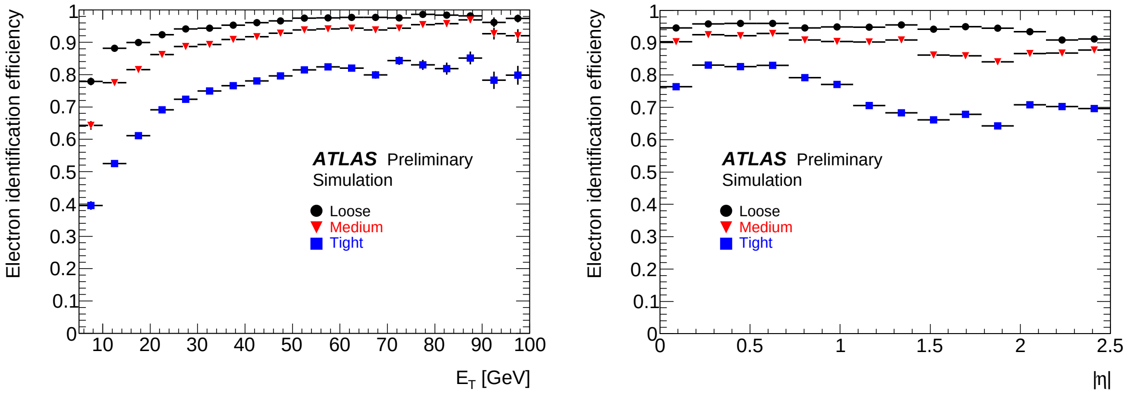


Figure 3.3: Efficiency of electron tight (2010 definition) shown as a function of p_T (left) and η (right) in MC simulation.

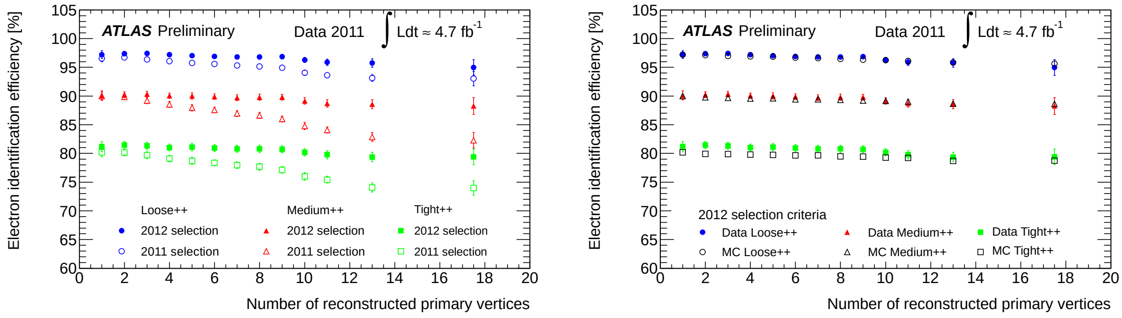


Figure 3.4: Efficiency of electron tight++ (2011, 2012 definition) shown as a function of number of reconstructed vertices. The left plot compares the efficiency between 2011 and 2012 data. The right one shows the comparison between 2011 Data and MC.

tool kit by Egamma group. The average systematic is concluded in Chapter 4.

3.2 Muons

3.2.1 Muon Reconstruction

Several muon reconstruction algorithms have been performed, named as STACO and MUID. In the Higgs to $WW^{(*)}$ analysis, the STACO muons are used. The STACO muon provides the following three algorithms [77]:

- MUONBOY: Use the hit information in the muon spectrometer and produce standalone segments and tracks which can be extrapolated to the vertex.

- STACO: Combine the track in the inner detector with the one in the muon spectrometer, named as combined muons.

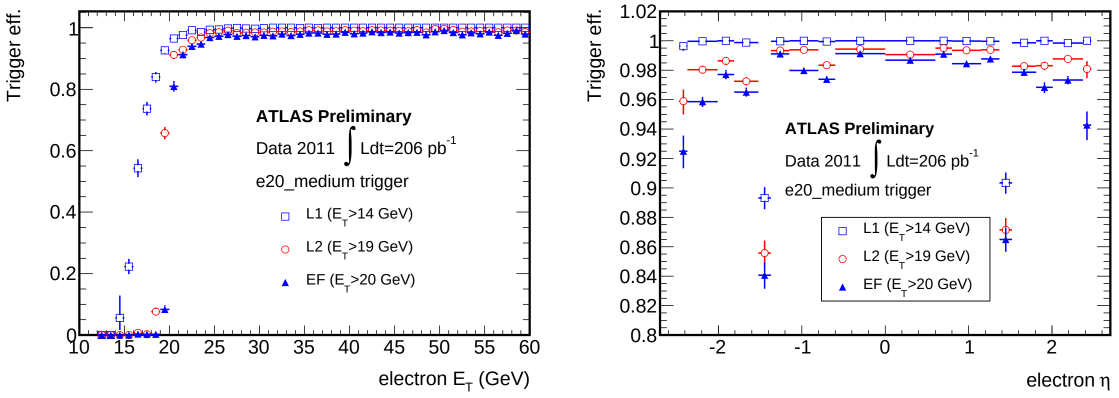


Figure 3.5: Trigger efficiency of electron e20_medium shown as a function of the transverse energy E_T (left) and η (right)

- Mutag: Combine the Muonboy segment not included in the Staco algorithm with an inner detector track.

The Staco algorithm loops in two muon track containers, MS and ID, compares the parameter vectors and their covariance matrices in the ID and MS to find the minimum χ^2 and to match one ID track and one MS track when the χ^2 is within a given cut value. Through this procedure the combined muons are obtained in the region $|\eta| < 2.5$.

3.2.2 Muons Selection in $H \to WW^{(*)}$ Analysis

In the Higgs to $WW^{(*)}$ analysis, a combined and isolated muon is required. According to the recommendation from the muon combined performance group, the selection criteria are shown as in Tab. 3.4. Similar to the selection of electrons, criteria such as p_{T} threshold, tracking requirements, primary vertex requirements and isolation cuts are applied on muons. Also the trigger and reconstruction efficiencies are determined using the tag-and-probe method as shown in Fig. 3.6 and Fig. 3.7 [78].

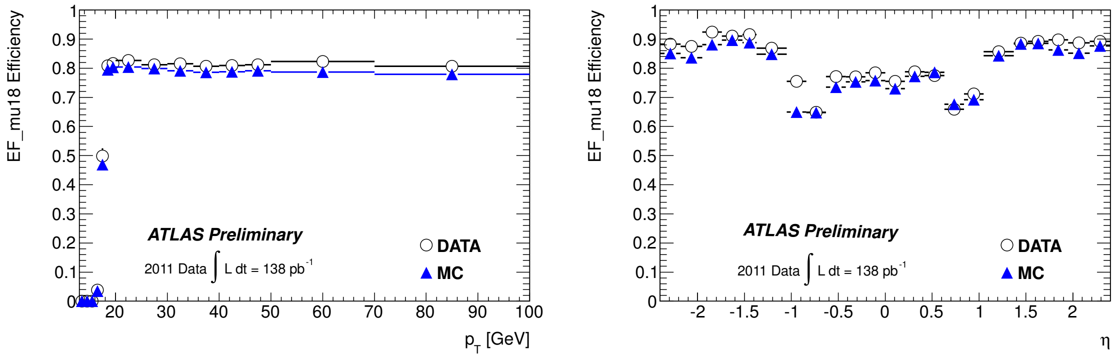


Figure 3.6: Trigger efficiency of muons shown as a function of p_T (left) and η (right).

Apply muon p_{T} smearing to MC only
Trigger requirement:
Periods B-I: mu18_MG Periods J-M: mu18_MG_medium
Staco combined muon
$P_T > 10\mathrm{GeV}$ $\lvert\eta\rvert < 2.4$
Tracking requirement:
Require a b-layer hit where expected Number of pixel hits + number of crossed dead pixel sensors > 1 Number of SCT hits + number of crossed dead SCT sensors $\geqslant 6$ $n = n\mathrm{TRThits} + n\mathrm{TRToutliers}$ $n > 5$ and $n\mathrm{TRToutliers} < 0.9 \times n$ for $\lvert\eta\rvert < 1.9$ If $n > 5$, then require $n\mathrm{TRToutliers} < 0.9 \times n$ for $\lvert\eta\rvert \geqslant 1.9$ $N = n\mathrm{TRTOutliers} + n\mathrm{TRTHits}$ If $\lvert\eta\rvert < 1.9$ then $N > 5$ and $n\mathrm{TRTOutliers}/N < 0.9$ If $\lvert\eta\rvert >= 1.9$ and $N > 5$ then $n\mathrm{TRTOutliers}/N < 0.9$ Number of pixel holes + number of SCT holes < 3
$\lvert z_0\rvert(\mathrm{w.r.t.PV}) < 1\,\mathrm{mm}$ Impact parameter significance $d_0(\mathrm{sig}) < 3$
Isolation:
$P_T > 15\,\mathrm{GeV}$ pile up corrected $E_{T,cone30}/p_{\mathrm{T}} < 0.14$ $P_{T,cone30}/p_{\mathrm{T}} < 0.15$

Table 3.4: Muon selection criteria in the Higgs to $WW^{(*)}$ analysis.

3.3 Jets

3.3.1 Jets Reconstruction and Calibration

The jets in the ATLAS detector are the experimental responses of hardrons coming from partons in the collision. It is presented as the energy deposit in a group of cells in the calorimeter as well as a bunch of tracks in the inner detector. Due to the large number of particles in a jet and the high luminosity at the LHC, which brings contamination from pile up and underlying events, it is difficult to reconstruct and calibrate jets in the detector. Mainly two kinds of algorithms have been performed in the reconstruction, the cone method, which uses seed, and the seedless method. The cone method is performed by collecting pre-clusters around a seed (usually the barycenter of energy in a given cone range) in the calorimeter. It is a simple method which has fast reconstruction speed but is also accompanied with two disadvantages, the inferred unsafe and collinear unsafe. The inferred safety means if there is a soft pre-cluster located between two hard pre-clusters, which may come from a gluon radiated in the collision, it may merge these two hard pre-clusters into one jet, while they should belong to different jets. The collinear safety reflects the fact that a pre-cluster with a high transverse momentum (e.g. 1 GeV) or two collinear pre-clusters

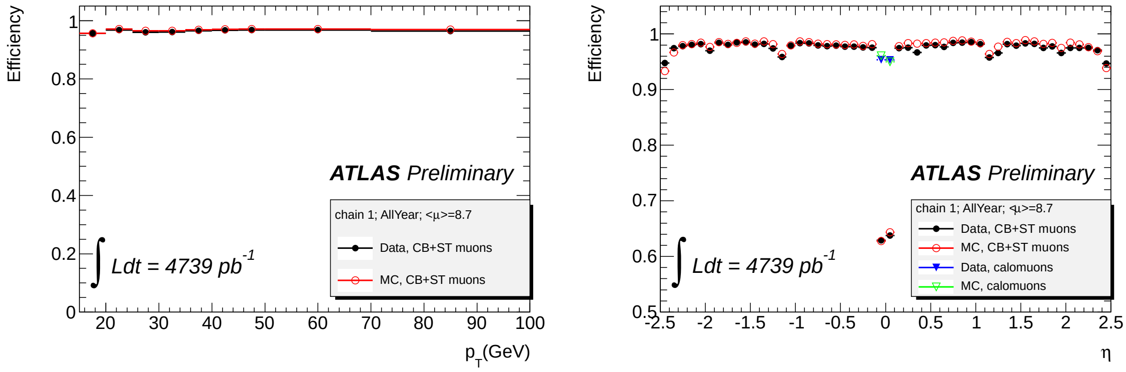


Figure 3.7: Efficiency of the muon reconstruction shown as a function of p_T (left) and η (right) using full 2011 data.

with lower transverse momentum (e.g. 0.5 GeV) can give quite different result in the reconstruction. Also, in a seed method the two collinear pre-clusters may change the sequence of the pre-cluster transverse momentum due to lower energy, so that the seed can be different from the one hard pre-cluster case. In addition, a cone method can cause dark tower, the large energy outside of any jets. To avoid these problems, the seedless method such as the Anti-kt method is proposed [79], though this method may consume longer computing time. Between two pre-clusters, a distance is defined as Eq. 3.1:

$$d_{i,j} = \min\left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2}\right)\frac{\Delta_{ij}^2}{R^2} \tag{3.1}$$

where p_{Ti} is the transverse momentum of a pre-cluster, i, j are the indices of pre-clusters, $\Delta_{ij}^2 = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ and R is a given range, it is chosen to 0.4 in this thesis. The two pre-clusters are merged if the $d_{i,j}$ of them has the minimum value. This process repeats until all the merged clusters are qualified as jets. In this procedure the pre-clusters with higher transverse momenta are merged before the lower E_T pre-clusters.

To calibrate the jet energy, the EM scale topological clusters are used as input pre-clusters in the jet finding algorithm. The clusters are categorized into two types, the electromagnetic type and hadronic type. For the EM type, the calibration is already done using the test beam and MC. It has the same calibration as for electrons. The hadronic type clusters are calibrated locally or globally by taking into account the signal weight and material effect in cells. In our analysis we are using the EM calibrated clusters with a jet energy scale (EM+JES). The jet energy scale is a function of jet E_T and η, derived using single isolated hadron response measured *in situ* with test beams data [80].

3.3.2 Jets Flavour Identification and Selection

For the 2011 data analysis, jets are reconstructed with the following selections:

- The jets are in the region $|\eta| < 4.5$ with $E_T > 25\,\mathrm{GeV}$.
- Overlap removal is performed between jets and any selected electrons.
- To identify jets coming from b quarks, b-tagging is performed. The jet b-flavour weight is calculated by the combined neural network method, which exploits the impact parameter significance and the topology of b and c hadron decays. The b-jets should have weight larger than -1.25 corresponding to a working point with 80% efficiency [81]. The b-tagging is only performed in $|\eta| < 2.5$ and $E_T > 20\,\mathrm{GeV}$.
- A cut on the jet vertex fraction (JVF) JVF$\geqslant 0.75$ is also applied for jets in $|\eta| < 2.1$. The JVF is the fraction of the summed p_T of tracks which belong to jets pointing back to the primary vertex. This is needed to reduce contributions from pile up effects.
- In the high pile up condition a significant excess has been observed in $2.5 \leq |\eta| \leq 3.5$ due to the in- and out-of-time pile up (pile up effect in the same or successive bunch crossing) as well as worse granularity and resolution of the calorimeter subdetectors in the transition region between the EMEC and FCAL detectors. Finally all the jets in the range $2.5 \leq |\eta| \leq 3.5$ are required to have a higher E_T threshold at $30\,\mathrm{GeV}$.

3.4 Transverse Missing Energy

The transverse missing energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed by summing energy over all of the cells in the calorimeter, taking into account the muon p_T, the energy loss in the cryostat, and cells not associated with any objects [82]:

$$\begin{aligned} E_{x(y)}^{\mathrm{miss}} = \quad & E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss},\gamma} + E_{x(y)}^{\mathrm{miss},\tau} + E_{x(y)}^{\mathrm{miss,jets}} \\ & + E_{x(y)}^{\mathrm{miss,soft\ jets}} + E_{x(y)}^{\mathrm{miss,CellOut}} + E_{x(y)}^{\mathrm{miss},\mu}, \end{aligned} \tag{3.2}$$

The $E_{\mathrm{T}}^{\mathrm{miss}}$ used in this analysis, which is called MET_RefFinal, is calibrated by a refined calibration: apply different weights to cells that belong to different reconstructed objects such as electrons and jets [83]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is an important component in reconstructing Higgs transverse mass, see Eq. 4.6 in Chapter 4. As there are missing neutrinos from decays of the $WW^{(*)}$ final state, the $E_{\mathrm{T}}^{\mathrm{miss}}$ gives a measurement of the transverse momentum of neutrinos and balances the visible transverse energy from leptons and jets. However, the pile up effect, underlying events, resolution of soft hardronic candidates or the high p_{T} candidates in the dead region, can cause the imbalance. The imbalance may result in longer $E_{\mathrm{T}}^{\mathrm{miss}}$ tail for the processes which have no missing energy, such as the Drell-Yan process. The tail can contribute to the Higgs signal region especially in the high pile up condition. To make further suppression on Drell-Yan and QCD events, the variable actually used in this thesis is defined as

$E_{\mathrm{T,rel}}^{\mathrm{miss}}$:

$$E_{\mathrm{T,rel}}^{\mathrm{miss}} = \begin{cases} E_{\mathrm{T}}^{\mathrm{miss}} & \text{if } \Delta\phi \geq \pi/2 \\ E_{\mathrm{T}}^{\mathrm{miss}} \times \sin\Delta\phi & \text{if } \Delta\phi \leq \pi/2 \end{cases} \tag{3.3}$$

where $\Delta\phi$ is the absolute difference in the azimuthal angle between the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector and the closest lepton or jet in the transverse plane.

3.5 Luminosity Determination

The luminosity in the ATLAS detector is measured using LUCID and BCM as described in Chapter 2. The integrated luminosity may be defined alternatively as:

$$\mathcal{L} = \frac{\mu^{\mathrm{vis}} n_b f_r}{\sigma_{\mathrm{vis}}} \tag{3.4}$$

where μ^{vis} is the average number of visible inelastic interactions per bunch crossing (BC) which can be calibrated by the van der Meer (vdM) scan [84], n_b and f_r are the number of bunches cross at the interaction point and the revolution frequency in the collider operation, respectively and $\sigma_{\mathrm{vis}} = \epsilon\sigma_{\mathrm{inel}}$ represents the inelastic interaction cross section multiplied with the efficiency of the measurement. Therefore the luminosity $\mathcal{L}$ is determined by the value of μ^{vis}, which is obtained from counting the number of the measured inelastic events N and the number of bunch crossings N_{BC}. When $\mu^{\mathrm{vis}} \ll 1$ the μ^{vis} is simply N/N_{BC}, otherwise it follows a Poisson distribution:

$$P_{\mathrm{Event,OR}}(\mu_{\mathrm{vis}}^{\mathrm{OR}}) = \frac{N_{\mathrm{OR}}}{N_{\mathrm{BC}}} = 1 - e^{-\mu_{\mathrm{vis}}^{\mathrm{OR}}} \tag{3.5}$$

$$\mu_{\mathrm{vis}}^{\mathrm{OR}} = -\ln\left(1 - \frac{N_{\mathrm{OR}}}{N_{\mathrm{BC}}}\right) . \tag{3.6}$$

This shows the relation between the average inelastic interaction in every bunch crossing and the number of observed events on either side of LUCID over the number of bunch crossings in a period. Also, one can get the probability that observing events on both sides of LUCID simultaneously, but it is more complicated. In the 2011 data taking, the peak μ^{vis} reached a relatively high value up to 24 and in 2012, this number can be as high as 40 (Fig. 3.8). The systematic uncertainty for the 2011 luminosity measurement is 3.9% [104] [1]. This number is 3.6% for the 2012 8 TeV measurement.

---

[1] The uncertainty is now improved to 1.8% [105].

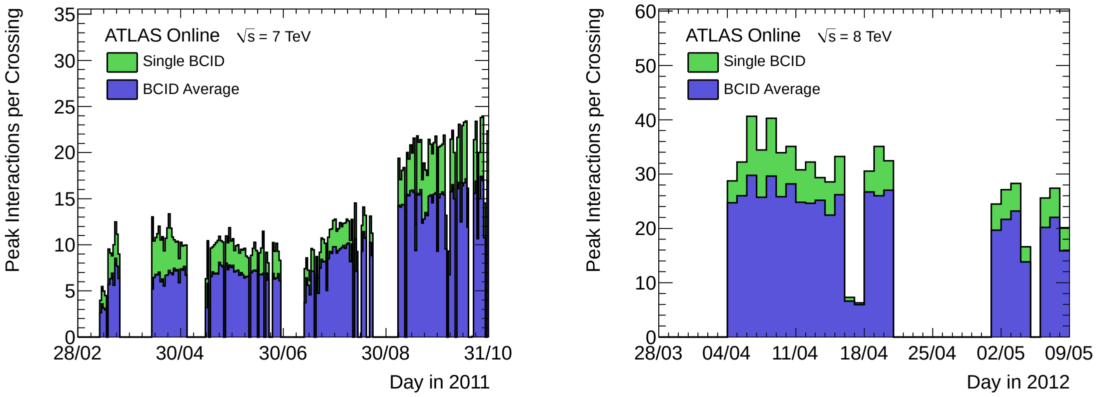


Figure 3.8: Maximum mean number of events per bunch crossing versus day. The online luminosity measurement is used for this calculation as for the luminosity plots. The maximum pile up for any bunch is shown in green and the maximum pile up averaged over all the colliding bunches is shown in blue. The left plot shows the result in 2011 and the right one shows part of the 2012 data sample.

Chapter 4

Event Selection in Higgs to $WW^{(*)}$ to $l\nu l\nu$ Analysis

In this study, the Standard Model Higgs boson search is performed in the $WW^{(*)}$ diboson channel, which has the largest decay branching fraction over all the decay modes in the intermediate Higgs mass region ($135 < m_H < 240\,\text{GeV}$) and has comparable sensitivity in the low Higgs mass region ($\sim 130\,\text{GeV}$) as the diphoton channel. Both of the W bosons are required to decay leptonically. In each event two prompt and isolated leptons are selected to suppress QCD background contamination in pp collisions. The dilepton final states have three lepton flavour combinations in the final state: ee, $\mu\mu$ and $e\mu$. The electrons and muons arise from both the direct decay in $W \to e\nu$, $W \to \mu\nu$ and a cascade decay $W \to \tau\nu$; $\tau \to e\nu$ or $\mu\nu$. The neutrinos in the decay cannot be detected in the detector, however, the missing neutrinos contribute to the missing energy in the transverse plane. Thus, the transverse mass of the Higgs boson is reconstructed instead of a full invariant mass. The two leptons are required to have opposite electric charges due to the neutral Higgs boson.

In addition to charged leptons, there may be jets in the final state. Depending on the number of jets reconstructed in the event, the background contribution can be very different. For this reason, events are categorized into 0-jet, 1-jet, more than 2 jets (named as 2-jets hereafter) bins. In the 0-jet bin, the Standard Model WW continuum represents the dominant irreducible background to the signal $gg \to H \to WW^{(*)}$, while other background events such as top events can be rejected efficiently with a jet veto requirement. In the 2-jets bin, the vector boson fusion (VBF) process with two forward going jets is expected to appear. Here, b-tagged jets are used to suppress the dominant top background contribution. Hence, the three flavour combinations and three jet bins result in nine sub-channels for the Higgs search in the $WW^{(*)} \to l\nu l\nu$ channel.

The search for Higgs boson in a SM-like model with a fourth generation is also performed with the same data analysis. The search sensitivity in the 4-generation model is higher than in 3-generation SM because of the larger production cross section, see Chapter 6. All analyses below are performed for the SM case unless explicitly stated otherwise.

The criteria to select leptons, jets and missing transverse energy have already been introduced in Chapter 3. The data and MC samples used in this analysis are presented in Sections 4.1 and 4.2, respectively. Trigger requirement and event selection cut flows for analyses in 0-, 1- and 2-jets are described in details in Sections 4.3 and 4.4. The systematic uncertainties are listed in Section 4.5. Some of the backgrounds are estimated using data-driven methods while others are based the MC predictions which are checked with data. The methods to estimate backgrounds will be covered in Chapter 5.

4.1 Data Sample

The analysis presented here uses the full data sample of 2011 taken by ATLAS at 7 TeV. The data sample corresponds to an integrated luminosity of $4.7\,\mathrm{fb}^{-1}$. The data is categorized into period and trigger in the data taking (egamma and muon streams). Different time periods can have different luminosity values, increasing with time. The single muon trigger requires events to have at least one muon with $p_{\mathrm{T}} > 18\,\mathrm{GeV}$. The single electron trigger sets the threshold at 20 or 22 GeV depending on the period, to cope with the increasing event rate. Events with both high p_{T} electrons and muons can be redundant, so that duplicate removal is performed during the analysis. Each event should have a primary vertex with at least three associated tracks of $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ to suppress non-collision background contributions. Quality cuts are also applied to reject events which have fake jets in the region where the detector is not functioning properly.

4.2 Monte Carlo Samples and Simulations

The Monte Carlo samples for different background processes are produced with various MC generators, which are listed in Tab. 4.1 together with the corresponding cross section values. The Higgs signal is generated by POWHEG [85, 86] for ggF and VBF and Pythia [87] for WH and ZH. The cross section and branching ratio values as a function of the Higgs mass are listed in Tab. 4.2 [88] and compared with the numbers of CERN Yellow Report 1.

Among all the background samples, the $W \to l\nu$, inclusive $Z/\gamma^* \to ll$ and $W\gamma$ are generated by ALPGEN [89], which describes well parton multiplicity distributions. The other backgrounds except $gg \to WW$, $b\bar{b}$, single top Wt, t channel are generated by MC@NLO [90], which provides next-to-leading-order (NLO) calculation. The use of leading-order (LO) AcerMC [91] in single top samples is due to the discrepancies in the comparison between MC@NLO and other generators in the b-jet rapidity distribution.

As the POWHEG and MC@NLO only described the p_T distribution of Higgs at leading-order and the p_T distribution is sensitive to lepton p_T and jet bin migration, it is necessary to reweight Higgs p_T distribution to NLO with large log resummation at NNLL (Next-to-Next-to-Leading-Log) [92].

In addition, the ALPGEN event generator uses Parton Distribution Functions (PDFs) CTEQ6L1 [93], a LO PDF set such that the description on lepton kinematics

Process	Generator	cross-section σ (pb) ($\times$ BR)
Inclusive $W \to \ell\nu$	ALPGEN	10.5×10^3 [95, 96]
Inclusive $Z/\gamma^* \to \ell\ell$ ($M_{\ell\ell} > 40$ GeV)	ALPGEN	10.7×10^2 [96, 97]
Inclusive $Z/\gamma^* \to \ell\ell$ ($10 < M_{\ell\ell} < 40$ GeV)	ALPGEN	3.9×10^3 [97]
$t\bar{t}$	MC@NLO	164.6
Single top t-channel	AcerMC	64.2 [98, 99]
Single top Wt	AcerMC	15.6 [98, 99]
Single top s-channel	MC@NLO	4.6 [98, 99]
WZ	MC@NLO	18.0
ZZ	MC@NLO	5.6
$qq/qg \to WW \to \ell\nu\ell\nu(\ell = e, \mu, \tau)$	MC@NLO	4.7
$gg \to WW \to \ell\nu\ell\nu(\ell = e, \mu, \tau)$	gg2WW	0.14
$W\gamma$	ALPGEN	289
$b\bar{b}$ (2-ℓ filter, $p_{\mathrm{T}} > 10GeV$)	PYTHIA	4270

Table 4.1: Cross section values at the center-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ for background processes. The $W \to \ell\nu$ and the $Z/\gamma^* \to \ell\ell$ cross sections include the branching ratio to a single lepton flavor, while the γW and two WW process include all three lepton flavors. The generators used for the simulation of the various processes are also indicated.

in the Z/DY sample is not perfect. Therefore a reweighting on the Z/DY events is performed using a PDF set based on NLO calculations, which is developed by the ATLAS SUSY working group. Moreover, a P_{Tll} reweighting is applied to correct the remaining mismodelling on the P_{Tll} shape in the Z/DY sample.

All of the MC samples are simulated by taking into account the time dependence on different data taking periods. After the event generation, the detector acceptances and efficiencies are obtained from full simulations of the ATLAS detector using GEANT4 [94]. This includes the time dependence on different data taking periods. The events are finally stored in NTUPLE format which can be easily accessed and conveniently analyzed. In this study, the pile up effect is modelled by Pythia 6 and the corresponding MC samples are called MC11c samples. In practice, the events are assigned randomly in each period to have the same fraction of the total number of events of the data. All of the MC are also reweighted to data using the number of primary vertices.

4.3 Trigger Requirements

The trigger criteria used in each period are listed in Tab. 4.3. In the event selection, each ee event should pass the electron trigger criteria and similarly for $\mu\mu$. The events in $e\mu$ channel should pass either of the trigger criteria. The number in the trigger name indicates the p_T threshold. The "medium" in some of the trigger names represents the tighter requirements to reduce the higher event rates when the number of pile up collisions in each bunch crossing increases. By studying on MC, the trigger

	σ for MC samples					σ from CERN YR				
m_H	POWHEG (fb)		Pythia (fb)		Br(H → WW)	ggF	VBF(u+t+s)	VBF(u+t)	WH	ZH
(GeV)	ggF	VBF	WH	ZH		(fb)	(fb)	(fb)	(fb)	(fb)
100	27.989	1.801	1.382	0.736	0.011	28.795	3.344	1.853	1.422	0.757
105	55.559	3.755	2.597	1.390	0.024	57.159	6.669	3.863	2.672	1.430
110	100.387	7.074	4.429	2.389	0.048	103.279	12.076	7.277	4.557	2.458
115	165.009	12.123	6.868	3.738	0.087	169.762	19.937	12.472	7.066	3.846
120	249.642	19.050	9.849	5.401	0.143	256.834	30.334	19.598	10.133	5.557
125	347.151	27.459	12.990	7.161	0.216	357.152	42.436	28.250	13.365	7.367
130	452.090	36.948	16.034	8.895	0.305	465.113	55.563	38.013	16.496	9.151
135	553.354	46.536	18.572	10.377	0.403	569.294	68.259	47.876	19.107	10.676
140	641.773	55.659	20.407	11.492	0.504	660.260	79.721	57.262	20.994	11.823
145	713.397	63.554	21.560	12.217	0.603	733.947	89.168	65.384	22.181	12.569
150	770.471	70.568	22.021	12.570	0.699	792.666	96.954	72.601	22.655	12.932
155	818.479	76.709	22.110	12.743	0.796	842.057	103.462	78.919	22.747	13.110
160	866.443	83.848	21.861	12.729	0.909	891.402	110.620	86.264	22.491	13.096
165	837.927	85.787	21.223	12.419	0.960	862.514	112.006	88.304	21.845	12.784
170	782.962	82.794	19.075	11.204	0.965	805.516	106.502	85.179	19.625	11.527
175	725.189	78.583	16.986	10.006	0.958	746.079	99.853	80.847	17.475	10.295
180	659.328	73.183	14.881	8.724	0.932	678.321	91.919	75.291	15.310	8.975
185	557.735	63.730	12.289	7.211	0.844	573.802	79.302	65.566	12.643	7.419
190	486.486	57.139	10.339	6.078	0.786	500.500	70.347	58.785	10.636	6.253
195	441.120	52.790	9.043	5.323	0.757	453.828	64.432	54.311	9.304	5.477
200	408.305	49.558	8.028	4.742	0.741	420.067	59.981	50.986	8.259	4.879
220	321.398	40.624	5.353	3.174	0.714	330.656	47.802	41.795	5.507	3.266
240	264.647	34.298	3.751	2.216	0.704	272.271	39.491	35.286	3.859	2.280
260	225.418	29.263	2.696	1.584	0.699	231.911	33.118	30.106	2.774	1.629
280	196.696	25.251	1.970	1.149	0.695	202.362	28.178	25.978	2.027	1.182
300	175.942	21.873	1.466	0.849	0.692	181.011	24.132	22.503	1.508	0.874
320	162.522	19.001	0.000	0.000	0.689	167.204	19.548	19.548		
340	158.516	16.486	0.000	0.000	0.687	163.083	16.961	16.961		
360	161.349	13.791	0.000	0.000	0.651	165.997	14.188	14.188		
380	144.866	11.559	0.000	0.000	0.609	149.039	11.892	11.892		
400	124.147	9.898	0.000	0.000	0.582	127.723	10.183	10.183		
420	103.967	8.591	0.000	0.000	0.564	106.961	8.838	8.838		
440	86.188	7.584	0.000	0.000	0.554	88.671	7.802	7.802		
460	71.291	6.749	0.000	0.000	0.549	73.344	6.943	6.943		
480	58.807	6.041	0.000	0.000	0.546	60.501	6.215	6.215		
500	48.668	5.443	0.000	0.000	0.546	50.070	5.602	5.600		
520	40.230	4.920	0.000	0.000	0.547	41.389	5.063	5.062		
540	33.323	4.464	0.000	0.000	0.549	34.283	4.595	4.593		
560	27.646	4.062	0.000	0.000	0.552	28.443	4.179	4.179		
580	22.978	3.701	0.000	0.000	0.555	23.640	3.806	3.808		
600	19.137	3.380	0.000	0.000	0.558	19.688	3.477	3.478		

Table 4.2: The $pp \to H \to WW \to l\nu l\nu$ cross sections used in the MC normalization and signal yield determination for the extraction of the upper limit. The columns form 7 to 11 contain the production cross section from the Yellow Report 1, multiplied with the branching ratio of Higgs to WW and $W \to l\nu$ for each W, the column 6 contains the Higgs branching fraction to WW and the columns 2-5 contains the effective cross section use to normalize each MC sample taking into account cross section and all branching fractions.

efficiency for signal events is above 98 % in the ee channel, 91 % in the $\mu\mu$ channel and 94 % in the $e\mu$ channel through the entire data taking period [100].

Period	ee channel	$\mu\mu$ channel	eμ channel
B - I	EF_e20_medium	EF_mu18_MG	EF_e20_medium || EF_mu18_MG
J	EF_e20_medium	EF_mu18_MG_medium	EF_e20_medium || EF_mu18_MG_medium
K	EF_e22_medium	EF_mu18_MG_medium	EF_e22_medium || EF_mu18_MG_medium
L - M	EF_e22vh_medium1	EF_mu18_MG_medium	EF_e22vh_medium1 || EF_mu18_MG_medium

Table 4.3: Period dependent trigger setup used in the analysis.

The trigger matching is performed for both leading and subleading leptons by requiring p_T larger than a given threshold. For electron triggers, the thresholds are $p_T > 21\,\text{GeV}$ for EF_e20_medium, $p_T > 23\,\text{GeV}$ for EF_e22_medium and EF_e22vh_medium1. For muon triggers, the thresholds are set at $p_T > 20\,\text{GeV}$ for both EF_mu18_MG and EF_mu18_MG_medium. For leptons passing the thresholds, it is matched with trigger object with a ΔR upper bound of 0.15. In each event at least one lepton must be matched to the trigger object. To correct for the small difference in trigger efficiencies between data and Monte Carlo, scale factors are applied on MC events in the analysis, which are calculated as follows:

$$\mathrm{SF} = \frac{1 - \left(1 - \epsilon_{\mathrm{MC}}^{\mathrm{lead}} \times \mathrm{SF}^{\mathrm{lead}}\right) \times \left(1 - \epsilon_{\mathrm{MC}}^{\mathrm{sub}} \times \mathrm{SF}^{\mathrm{sub}}\right)}{1 - \left(1 - \epsilon_{\mathrm{MC}}^{\mathrm{lead}}\right) \times \left(1 - \epsilon_{\mathrm{MC}}^{\mathrm{sub}}\right)} . \tag{4.1}$$

Here, $\epsilon_{\mathrm{MC}}^{\mathrm{lead}}$, $\epsilon_{\mathrm{MC}}^{\mathrm{sub}}$ are the per-lepton trigger efficiencies for the leading and subleading leptons, and $\mathrm{SF}^{\mathrm{lead}}$, $\mathrm{SF}^{\mathrm{sub}}$ are the per-lepton SFs for the leading and subleading leptons.

4.4 Event Selection

In the cut based analysis, various selection cuts are applied to suppress different background contributions. In the late 2011, the Higgs WW working group has decided to use a common analysis framework to provide the event selection cut flow tables and different kinematic distributions in both the background control regions and the signal region [102]. The common analysis code has been cross checked in detail by several independent analysis codes including mine.The event selection is presented in Tab. 4.4.

4.4.1 Common Event Selection

After selecting the objects as described in Chapter 3, event selections are performed to suppress background contributions in the Higgs search. First, each event in both MC and data has to have a primary vertex which has at least three tracks to suppress non-collision background events such as cosmic muon events. The data and MC are then required to have exactly two isolated leptons, with opposite charge. The leading lepton p_{T} is set at 25 GeV to have the good $E_{\mathrm{T}}^{\mathrm{miss}}$ agreement between data

Common Event Selection
Good Run List (Data only)
Primary Vertex Selection
Event Cleaning
Trigger requirement
Exactly two leptons
First leading lepton p_{T} >25 GeV
Two leptons with opposite charge
$m_{ll} > 12$ GeV, for same flavor, m_{ll} >10 GeV, for different flavor
Z mass veto: $\lvert m_{ll} - m_Z \rvert > 15$ GeV, for same flavor, m_Z=91.1876 GeV (PDG value)
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45$ GeV for same flavor and 25 GeV otherwise

0-jet	1-jet	2-jets
P_t^{ll} >45, 30 GeV for same/different flavor	b veto	central jets veto
	P_T^{Tot} <30 GeV	$\eta_{j1} \times \eta_{j2} <0$
	$Z \to \tau\tau$ veto	$\Delta\eta_{j1,j2} > 3.8$
		M_{jj} >500 GeV
		b veto
		P_T^{Tot} <30 GeV
		$Z \to \tau\tau$ veto

Common topological selection: Low mass region	High mass region
M_{ll} <50 GeV (80 GeV for 2 jets)	M_{ll} <150 GeV
$\Delta\phi_{ll} < 1.8$	
$0.75 \times m_H < m_T < m_H$	$0.60 \times m_H < m_T < m_H$

Table 4.4: The events selection criteria in Higgs search.

and MC. The subleading lepton p_{T} is set at 15 GeV for the nominal analysis. A cut on the dilepton invariant mass is further applied to suppress low mass resonances as well the on shell Z events which have invariant mass around m_Z: for same flavour channel, $m_{ll} > 12$ GeV and $\lvert m_{ll} - m_Z \rvert > 15$ GeV, for the different flavour channel, $m_{ll} > 10$ GeV.

Figure 4.1 shows the lepton p_T distributions for the ee (top row), $\mu\mu$ (middle row), and $e\mu$ (bottom row) channels with the minimum lepton p_T and m_{ll} requirements applied. For the top two rows, the plots on the left are for the leading lepton and the plots on the right are for the sub-leading lepton. For the bottom row, the plot on the left is for electrons and the right plot for muons. The lower part of each plot shows the ratio between the data and the background expectation from MC, with the yellow band indicating the total systematic uncertainty of the various components (see Section 4.5). The signal is shown for m_H=125 GeV. And Fig. 4.2 shows the lepton pseudorapidity distributions for the ee (top row), $\mu\mu$ (middle row), and $e\mu$ (bottom row) channels with the minimum lepton p_T and m_{ll} requirements applied. The position of all plots are the same as the ones in p_T distribution. Figure 4.3 shows the dilepton invariant mass distribution for the ee (top row), $\mu\mu$ (middle row), and $e\mu$ (bottom row) channels with the minimum lepton p_T and m_{ll} requirements applied. The left plots are in logarithmic scale and the right plots are in linear scale. These plots show good agreement between data and MC.

To supress off shell Z/DY events, the relative $E_{\mathrm{T}}^{\mathrm{miss}}$ ($E_{\mathrm{T,rel}}^{\mathrm{miss}}$) cut is applied: 45 GeV for the same flavour channels, 25 GeV for the $e\mu$ channel, respectively. The smaller cut for the $e\mu$ channel is used because the Z/DY events can only come from $Z \rightarrow \tau\tau$ decays, which is much smaller than the same flavour Z/DY contribution. Figure 4.4 shows the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ distributions of the ee (top row), $\mu\mu$ (middle row), and $e\mu$ (bottom row) channels with the minimum lepton p_T and m_{ll} requirements applied. The left plots are in logarithmic scale and right plots in linear scale.

After these common selection cuts, the events are categorized to 0-, 1- and 2-jet bins. The jet multiplicity is shown in Fig. 4.5. Table 4.5 shows the number of data and expected events in MC after the common selection. Some of the backgrounds are scaled with scale factors derived from data-driven methods, see Chapter 5.

4.4.2 Jet Bin Dependent Selections

In the 0-jet analysis, the transverse momentum of the dilepton system, p_T^{ll}, is required to be greater than 45 GeV for the $e\mu$ channel and 30 GeV for same flavour channels to suppress more Z/DY events. Also in this region the top contribution is strongly suppressed. In the 1-jet case, to reject top events, the jet should not be a b-tagged jet.

Another cut concerns $\mathbf{p}_T^{\mathrm{Tot}}$, which is defined as:

$$\mathbf{p}_T^{\mathrm{Tot}} = \mathbf{p}_T^{l1} + \mathbf{p}_T^{l2} + \mathbf{p}_T^{\mathrm{jets}} + \mathbf{p}_T^{\mathrm{miss}} , \tag{4.2}$$

and the cut is $\mathbf{p}_T^{\mathrm{Tot}} < 30\,\mathrm{GeV}$.

To suppress the $Z \rightarrow \tau\tau$ contribution in the $e\mu$ channel, a collinear approximation is made. This is based on the assumption that the lepton from the τ decay goes along the direction of τ so that if the mass of τ can be neglected the invariant mass of $m_{\tau\tau}$ is approximated as:

$$m_{\tau\tau} = \frac{m_{ll}}{\sqrt{x_1 x_2}} , \tag{4.3}$$

in which x_1, x_2 are the positive energy fraction carried by the visible particles in the τ decays. They are defined as:

$$x_1 = \frac{p_x^{l1} p_y^{l2} - p_y^{l1} p_x^{l2}}{p_y^{l2} E_x^{\mathrm{miss}} - p_x^{l2} E_y^{\mathrm{miss}} + p_x^{l1} p_y^{l2} - p_y^{l1} p_x^{l2}} , \tag{4.4}$$

$$x_2 = \frac{p_x^{l1} p_y^{l2} - p_y^{l1} p_x^{l2}}{p_x^{l1} E_y^{\mathrm{miss}} - p_y^{l1} E_x^{\mathrm{miss}} + p_x^{l1} p_y^{l2} - p_y^{l1} p_x^{l2}} . \tag{4.5}$$

The $|m_{\tau\tau} - m_Z|$ is required to be larger than 25 GeV. This cut is applied in 1- and 2-jet channels because in the 0-jet channel the two leptons are mostly back to back so that the approximation does not apply.

In the 2-jet analysis, in additon to the b-jet veto, $\mathbf{p}_T^{\mathrm{Tot}}$ cuts and $Z \rightarrow \tau\tau$ veto cuts mentioned above, the two leading jets are further required to be in different hemispheres and well separated in pseudorapidity: $\Delta\eta_{jj} \equiv |\eta_{\mathrm{jet1}} - \eta_{\mathrm{jet2}}| > 3.8$. Since

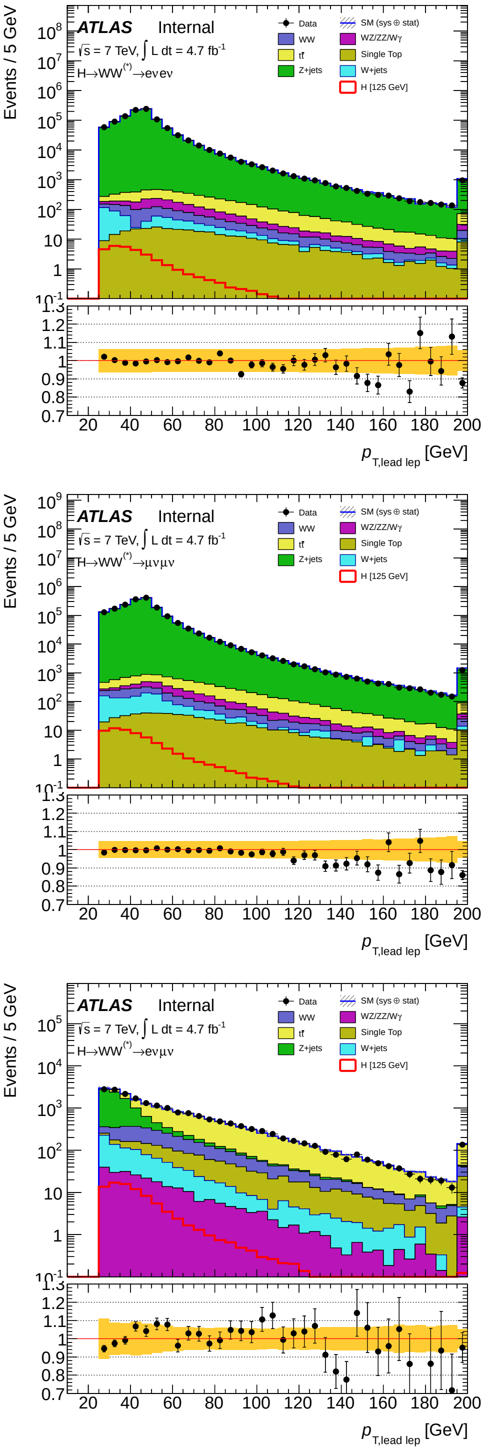


Figure 4.1: Lepton p_T distributions for the ee (top row), $\mu\mu$ (middle row), and $e\mu$ (bottom row) channels with the minimum lepton p_T and m_{ll} requirements applied. For the top two rows, the plots on the left are for the leading lepton and the plots on the right are for the sub-leading lepton. For the bottom row, the plot on the left is for electrons and the right plot for muons. The lower part of each plot shows the ratio between the data and the background expectation from MC, with the yellow band indicating the total systematic uncertainty in the normalization (but not the shape) of the various components. The signal is shown for $m_H = 125\,\mathrm{GeV}$. The final bin includes the overflow.

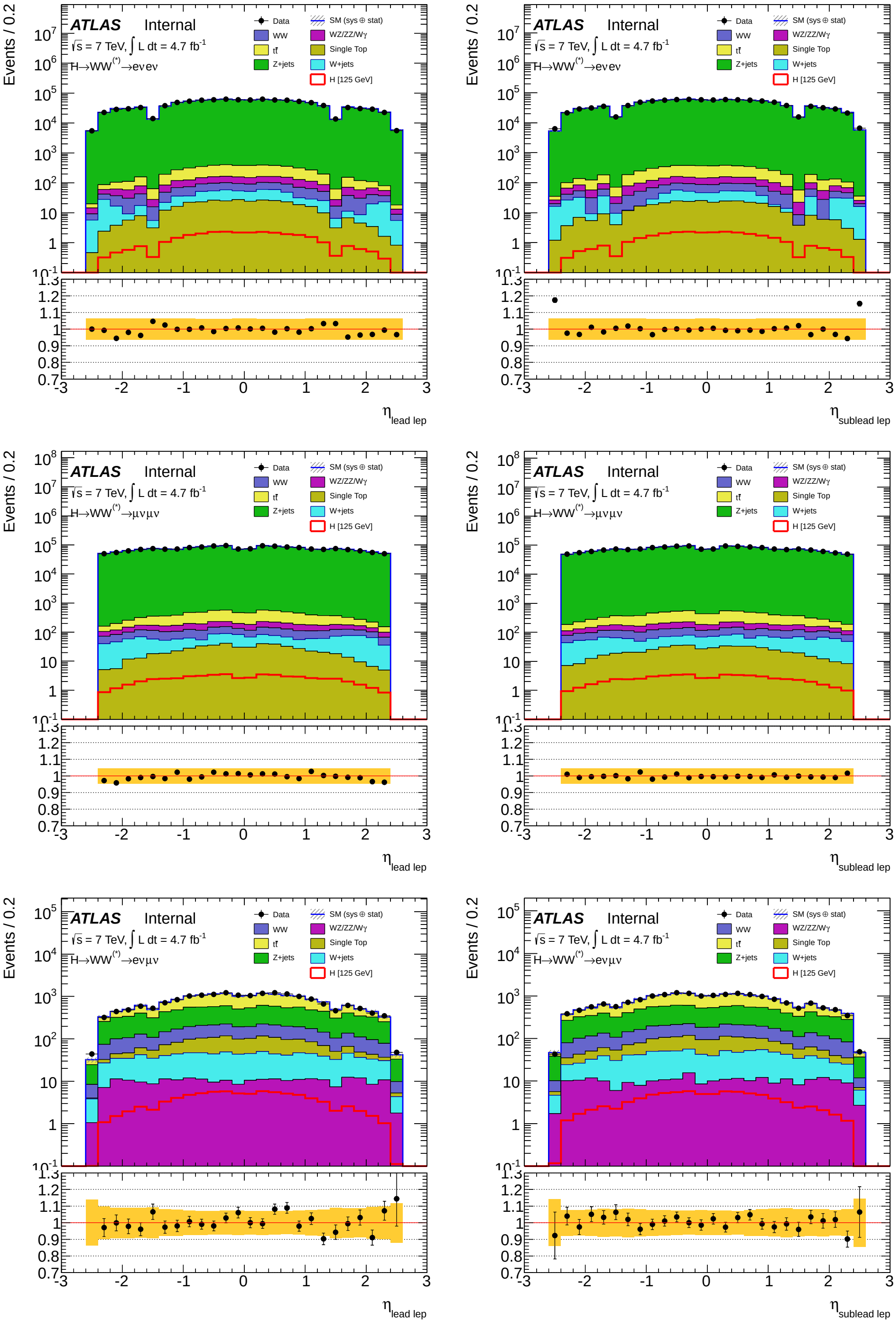


Figure 4.2: Lepton pseudorapidity distributions for the ee (top row), $\mu\mu$ (middle row), and $e\mu$ (bottom row) channels with the minimum lepton p_T and m_{ll} requirements applied. The layout of the plots is the same as in Fig. 4.1.

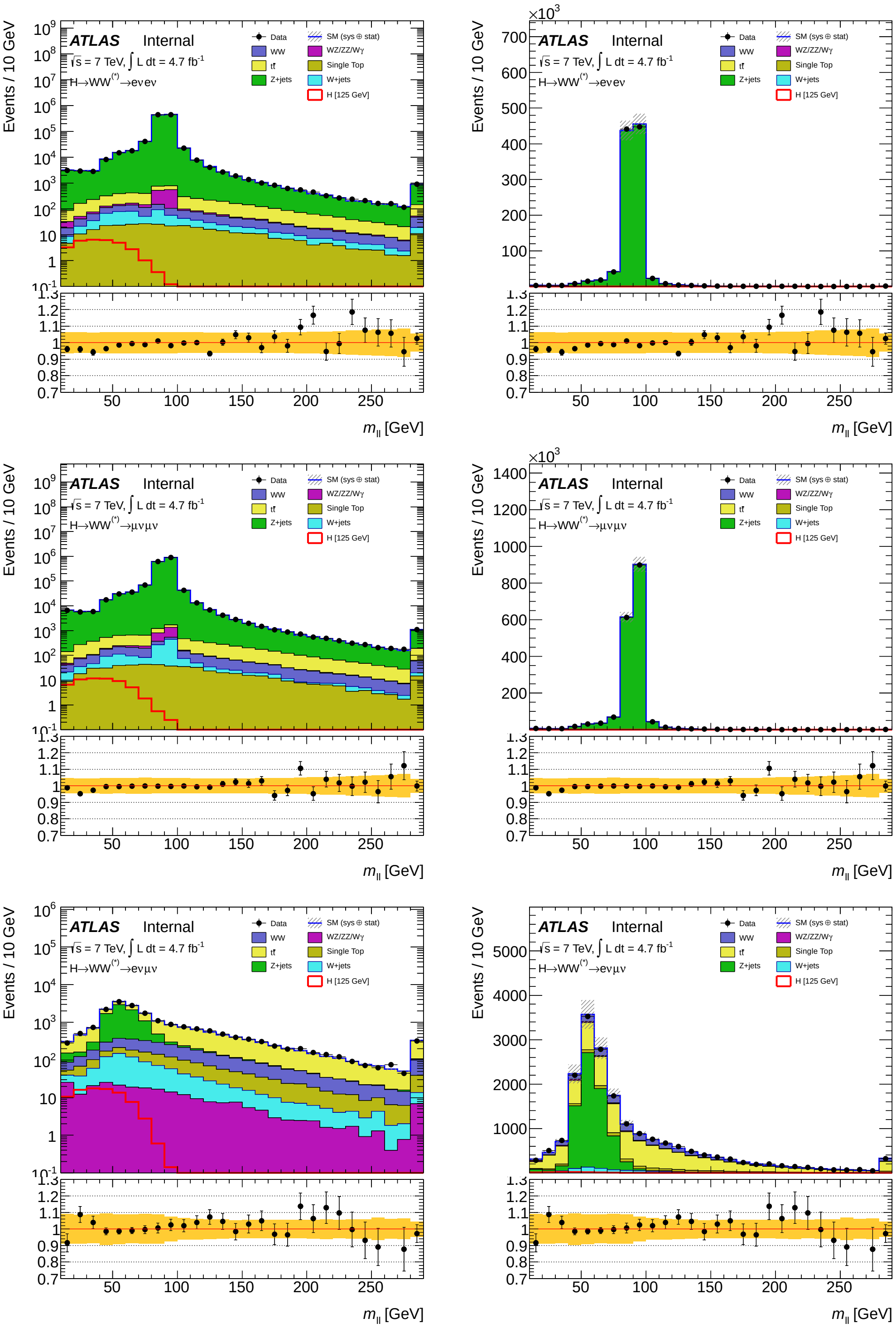


Figure 4.3: Dilepton invariant mass distribution for the ee (top row), $\mu\mu$ (middle row), and $e\mu$ (bottom row) channels with the minimum lepton p_T and m_{ll} requirements applied. The left plots are in logarithmic scale and the right plots are in linear scale. The lower part of each plot shows the ratio between the data and the background expectation from MC, with the yellow band indicating the total systematic uncertainty in the normalization (but not the shape) of the various components. The signal is shown for $m_H = 125\,\text{GeV}$. The final bin includes the overflow.

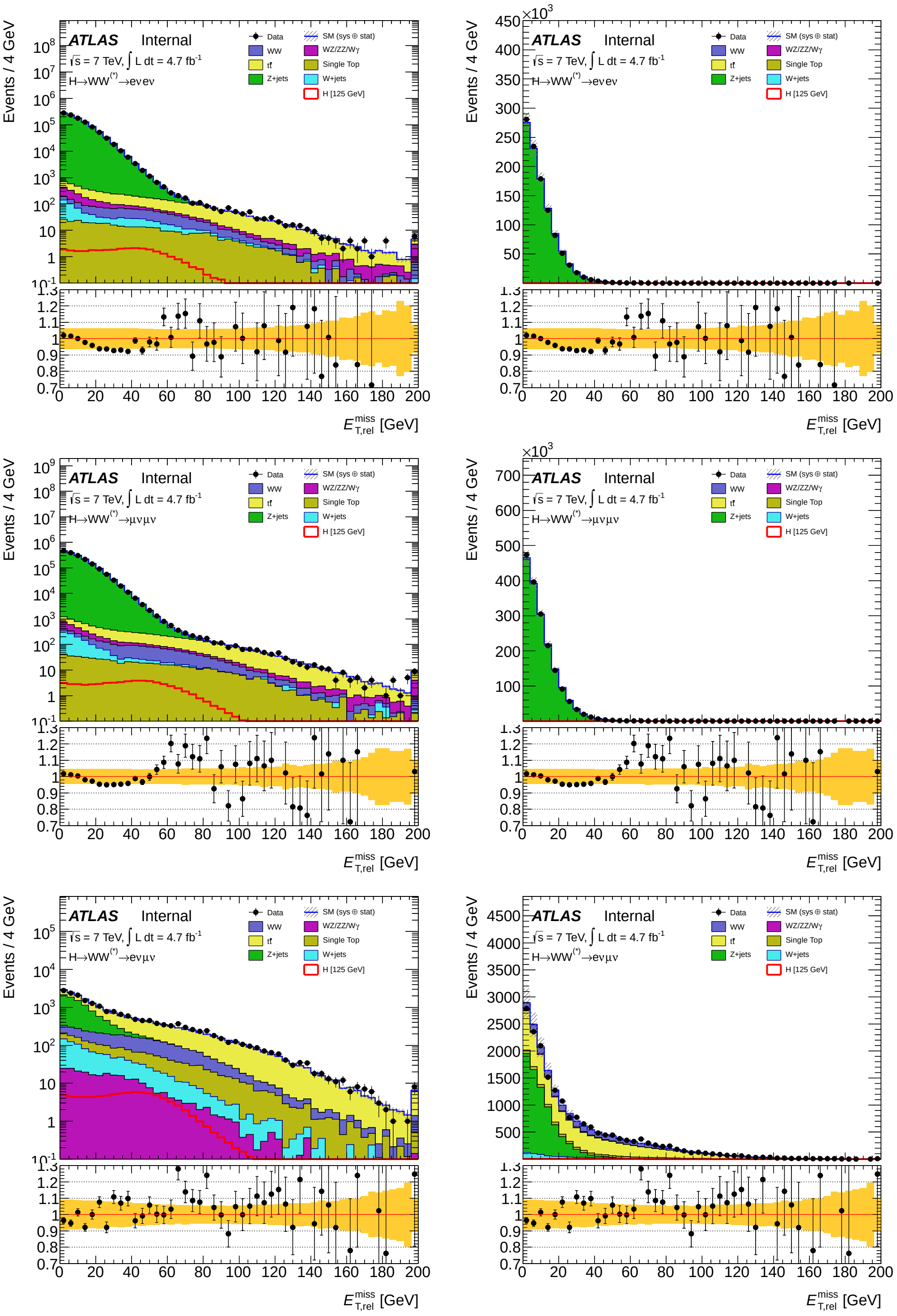


Figure 4.4: Distribution of $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ for the ee (top row), $\mu\mu$ (middle row), and $e\mu$ (bottom row) channels with the minimum lepton p_T and m_{ll} requirements applied. The left plots are in logarithmic scale and right plots in linear scale. The lower part of each plot shows the ratio between the data and the background expectation from MC, with the yellow band indicating the total systematic uncertainty in the normalization (but not the shape) of the various components. The signal is shown for $m_H = 125\,\mathrm{GeV}$. The final bin includes the overflow.

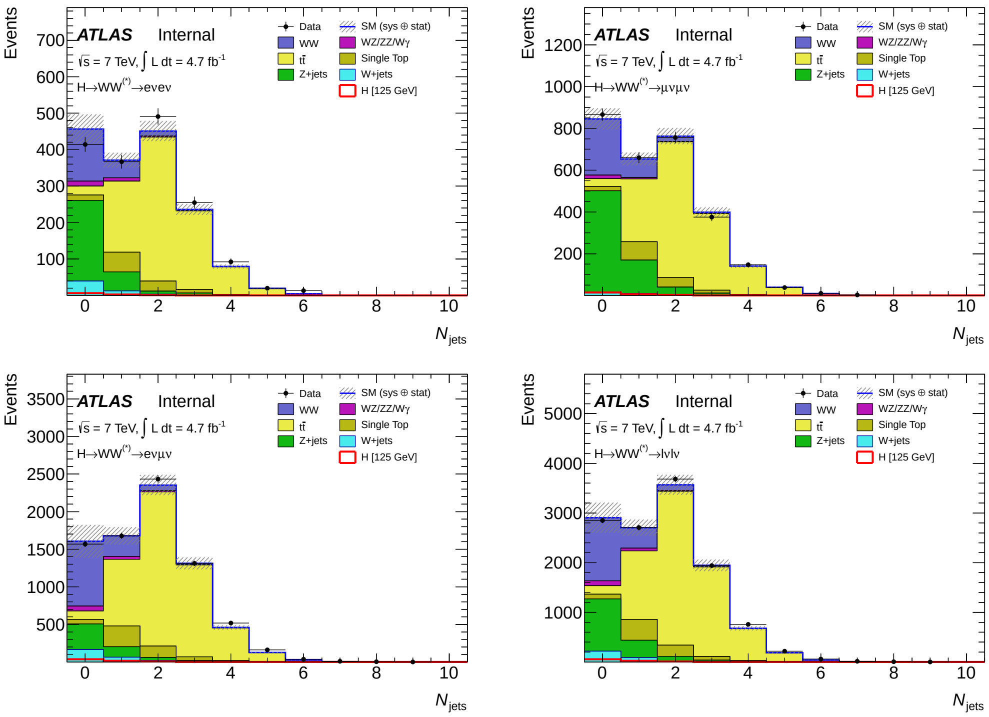


Figure 4.5: Multiplicity of jets with $p_T > 25\,\text{GeV}$ (for the transition region $p_T > 30\,\text{GeV}$ is required) after the cut on $E_{\text{T,rel}}^{\text{miss}}$ for the ee (top row left), $\mu\mu$ (top row right) and $e\mu$ (bottom row left) channel separated. The combined jet multiplicity is shown in the bottom right plot. The signal is shown for $m_H = 125\,\text{GeV}$. The hatched band around the sum of backgrounds indicates the total systematic uncertainty in the normalization (but not the shape) of the various components.

	Signal [125 GeV]	WW	WZ/ZZ/Wgamma	ttbar	Single Top	Z+jets	W+jets (MC)	W+jets (d-d)	Total Bkg. (d-d)	Observed	Data/MC (d-d)
lepton p_{T}	180.87 ± 0.40	3871.06 ± 9.55	3213.65 ± 20.73	16332.68 ± 27.65	1753.31 ± 11.35	2800788.09 ± 1309.25	1035.56 ± 55.09	3159.63 ± 37.12	2829118.43 ± 1310.32	2823123	1.00 ± 0.00
OS leptons	178.43 ± 0.39	3856.00 ± 9.53	2733.59 ± 15.79	16240.28 ± 27.57	1734.85 ± 11.30	2793617.99 ± 1307.64	694.48 ± 39.78	2373.89 ± 36.19	2820556.60 ± 1308.61	2816240	1.00 ± 0.00
$m_{\ell\ell} > 12, 10$ GeV	175.14 ± 0.39	3848.08 ± 9.52	2716.34 ± 15.68	16211.72 ± 27.55	1731.84 ± 11.29	2790757.50 ± 1307.32	686.22 ± 39.60	2194.45 ± 35.91	2817459.92 ± 1308.28	2806551	1.00 ± 0.00
Z veto (for $ee, \mu\mu$)	173.11 ± 0.39	3421.39 ± 8.97	559.69 ± 8.00	14456.35 ± 26.02	1542.36 ± 10.66	278756.95 ± 404.20	606.75 ± 37.50	1403.39 ± 20.67	300140.13 ± 405.88	298691	1.00 ± 0.00
$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	90.82 ± 0.28	1828.61 ± 113.91	177.88 ± 4.82	7322.28 ± 497.99	839.24 ± 57.60	1523.00 ± 33.78	251.62 ± 24.96	341.94 ± 5.92	12032.95 ± 688.35	12231	1.02 ± 0.06
ee:$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	10.24 ± 0.10	208.11 ± 13.14	26.70 ± 1.88	929.80 ± 63.55	107.98 ± 7.89	289.44 ± 14.91	38.36 ± 8.99	56.34 ± 1.10	1618.38 ± 91.86	1652	1.02 ± 0.06
$e\mu$:$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	58.28 ± 0.23	1230.87 ± 76.74	124.45 ± 4.35	4855.83 ± 330.36	558.06 ± 38.49	534.85 ± 13.88	193.18 ± 22.48	260.93 ± 5.05	7565.00 ± 445.09	7725	1.02 ± 0.06
$\mu\mu$:$E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45, 25$ GeV	22.31 ± 0.14	389.62 ± 24.42	26.72 ± 0.84	1536.65 ± 104.78	173.20 ± 12.31	698.71 ± 23.59	20.08 ± 6.08	24.67 ± 2.87	2849.57 ± 154.06	2854	1.00 ± 0.06

Table 4.5: Expected numbers of background events after selecting two leptons with opposite charge and $m_{\ell\ell} > 12\,\mathrm{GeV}$ ($m_{e\mu} > 10\,\mathrm{GeV}$), after applying the Z boson mass veto and the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ selections. The observed numbers of events in data are also given in the second last column and the ratio data over MC in the last column. The uncertainties shown are the statistical uncertainties.

the VBF process is expected to be the dominant Higgs signal source in the 2-jet analysis, events with additional jets in the central region ($|\eta| < 3.2$) are rejected. Furthermore, the invariant mass of the two leading jets should be larger than 500 GeV.

After all these 1- and 2-jet specific cuts, there are a few additional common topological cuts, which vary slightly depending on the probing Higgs mass region. For the mass region below ~ 195 GeV, the invariant mass of dilepton system should be smaller than 50 GeV to suppress the Standard Model WW background contribution in the 0- and 1-jet bins, while in 2-jet bin the cut value is increased to 80 GeV. The azimuthal angle difference between the two leptons, $\Delta\phi_{ll}$, is further required to be smaller than 1.8 exploiting the fact that due to spin correlations in the $WW^{(*)}$ system arising from the spin-0 nature of the SM Higgs boson and the $V-A$ structure of the W boson decay vertex, the charged leptons tend to emerge from the primary vertex pointing in the same direction. Finally the transverse mass variable, m_T, is used to test for the presence of a signal. This variable is defined as:

$$m_T = \sqrt{(E_T^{ll} + E_T^{\text{miss}})^2 - |\mathbf{p}_T^{ll} + \mathbf{p}_T^{\text{miss}}|^2}\,, \tag{4.6}$$

$$E_T^{ll} = \sqrt{|\mathbf{p}_T^{ll}|^2 + m_{ll}^2}\,, \tag{4.7}$$

in which $\mathbf{p}_T^{\text{miss}}$ is the missing transverse moment vector and ($\mathbf{p}_T^{ll}$ is the transverse momentum vector of the dilepton system. In an early analysis [101], a direct cut $0.75 \times m_H < m_T < m_H$ was applied. For the analysis of the full 2011 data, the search sensitivity is improved by exploiting the shapes of signal and background m_T distributions. In the high mass region, there is only a cut: $m_{ll} < 150$ GeV before the transverse mass cut. The transverse mass cut for this region in [101] was: $0.60 \times m_H < m_T < m_H$.

The distributions in each jet bin are shown in the following figures. Figure 4.6 shows kinematic distributions of m_{ll}, p_T^{ll}, $\Delta\phi_{ll}$ and m_T relevant to the 0-jet analysis after the jet veto cut. The expected signal is shown for m_H=125 GeV. In this region top is highly suppressed. The plots show clearly that p_T^{ll} and $\Delta\phi_{ll}$ can be used to suppress the Drell-Yan events.

Figure 4.7 shows distributions of the azimuthal opening angle $\Delta\phi_{ll}$ in the left figure and the transverse momentum of the dilepton system p_{T}^{ll} is shown in the right figure after the invariant mass cut. Figure 4.8 shows the transverse momentum p_{T}^{ll} of the two selected leptons in the left figure after the cut on the azimuthal opening angle $\Delta\phi_{ll}$ and the transverse mass m_T distribution in the right figure after all cuts. The data are in excellent agreement with the total background contributions indicating no significant excess of the Higgs signal in the 0-jet analysis.

Table 4.6 shows the expected numbers of signal (m_H=125 GeV) and background events for the 0-jet analysis in 4.7 fb^{-1} integrated luminosity as well as the observed numbers of events in data. The yields for the ee, $e\mu$, and $\mu\mu$ channels are added together. The composition in each of the lepton flavor channels is shown only for the final stage of the selections. The W+jets background is entirely determined from data, whereas for the other processes the expectations are taken from simulation with $Z/\gamma^* +$jets, $t\bar{t}$ and $tW/tb/tqb$, which are corrected by the normalisation factors from

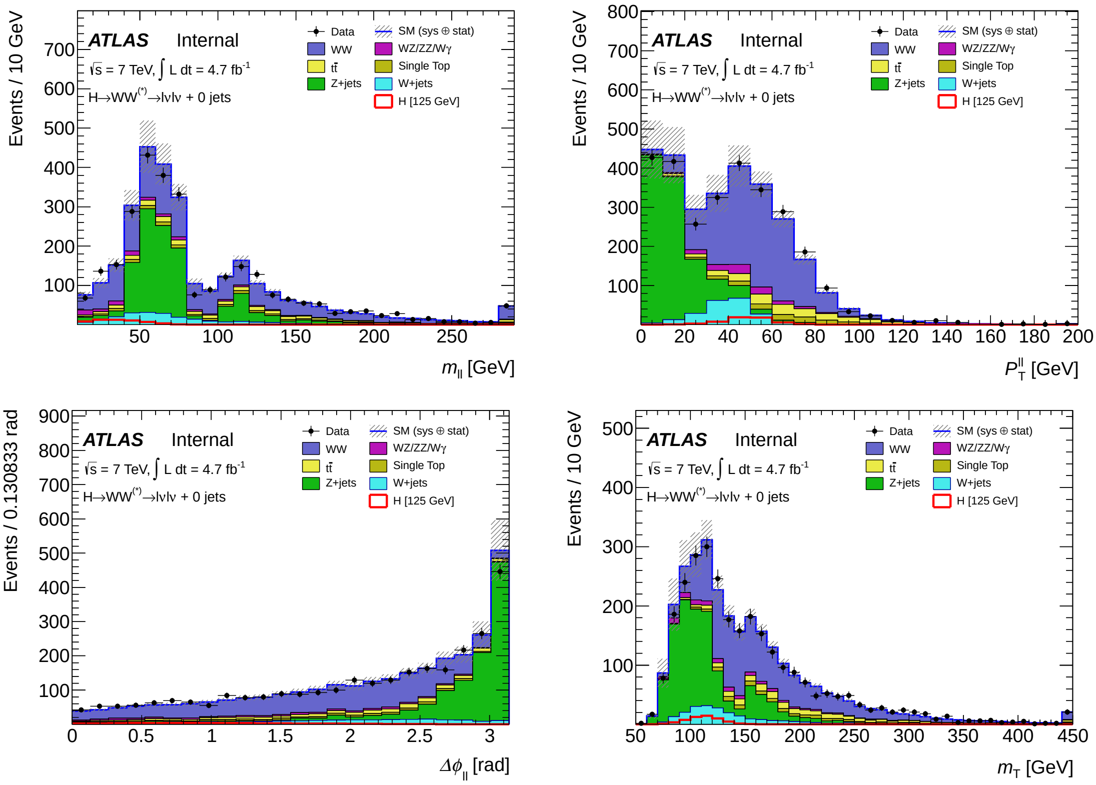


Figure 4.6: Kinematic distributions relevant to the 0-jet analysis after the jet veto cut. The invariant mass m_{ll} distribution is shown in the top left figure, the transverse momentum of the dilepton system p_T^{ll} is shown in the top right figure, the azimuthal opening angle $\Delta\phi_{ll}$ of the two selected leptons is shown in the bottom left figure and the transverse mass m_T distribution is shown in the bottom right figure. The expected signal is shown for $m_H = 125\,\mathrm{GeV}$. The hatched band around the sum of backgrounds indicates the total systematic uncertainty in the normalization (but not the shape) of the various components. The final bin includes the overflow.

the data driven control regions (See next chapter). The uncertainties shown are the statistical uncertainties.

Figure 4.9 shows kinematic distributions of m_{ll}, p_T^{Tot}, $\Delta\phi_{ll}$ and m_T relevant to the 1-jet analysis after the one jet requirement. Figure 4.10 shows the invariant mass $m_{\ell\ell}$ of the two selected leptons (left) and the azimuthal opening angle of the leptons after the $Z \rightarrow \tau\tau$ veto (right). Figure 4.11 shows the azimuthal opening angle $\Delta\phi_{\ell\ell}$ of the two selected leptons (left) after the cut on the invariant mass of the two selected leptons and the jet and the transverse mass m_T (right) after all cuts. Some excess(not greater than 2 standard deviation) in data over the background contributions is observed in the 1-jet analysis. But this is probably statistical fluctuation.

Table 4.7 shows the expected numbers of signal ($m_H = 125\,\mathrm{GeV}$) and background events for the 1-jet analysis in $4.7\,\mathrm{fb}^{-1}$ integrated luminosity as well as the observed numbers of events in data. The yields for the ee, $e\mu$, and $\mu\mu$ channels are added

	Signal [125 GeV]	WW	WZ/ZZ/Wgamma	ttbar	Single Top	Z+jets	W+jets (MC)	W+jets (d-d)	Total Bkg. (d-d)	Observed	Data/MC (d-d)
0j: jet veto	56.74 ± 0.23	1272.71 ± 79.34	96.62 ± 3.71	173.70 ± 12.16	94.92 ± 6.99	1038.66 ± 27.91	173.93 ± 20.94	216.98 ± 3.63	2893.59 ± 114.88	2849	0.98 ± 0.04
0j: $m_{\ell\ell} < 50$ GeV	45.22 ± 0.21	312.41 ± 19.64	40.81 ± 2.77	28.92 ± 2.34	19.23 ± 1.78	167.92 ± 9.60	56.42 ± 12.11	69.52 ± 2.00	638.83 ± 28.35	645	1.01 ± 0.06
0j: $p_{\mathrm{T},\ell\ell}$ >45,30 GeV	40.14 ± 0.19	282.12 ± 17.75	34.79 ± 2.51	27.27 ± 2.23	17.77 ± 1.68	27.98 ± 6.43	43.68 ± 10.60	49.16 ± 1.52	439.09 ± 26.44	443	1.01 ± 0.08
0j: $\Delta\phi_{\ell\ell} < 1.8$	39.02 ± 0.19	275.54 ± 17.34	33.32 ± 2.43	26.83 ± 2.20	17.63 ± 1.67	27.65 ± 6.43	38.94 ± 10.12	44.08 ± 1.42	425.04 ± 25.99	429	1.01 ± 0.08
ee:0j: $p_{\mathrm{T},\ell\ell}$ >45,30 GeV	4.01 ± 0.06	37.30 ± 2.52	4.08 ± 0.80	3.63 ± 0.56	3.29 ± 0.55	3.73 ± 1.90	5.46 ± 3.87	7.75 ± 0.38	59.77 ± 4.64	52	0.87 ± 0.14
$e\mu$:0j: $p_{\mathrm{T},\ell\ell}$ >45,30 GeV	26.76 ± 0.16	170.55 ± 10.81	24.50 ± 2.34	16.17 ± 1.45	9.65 ± 1.08	1.93 ± 0.80	37.19 ± 9.84	40.08 ± 1.37	262.89 ± 12.92	253	0.96 ± 0.08
$\mu\mu$:0j: $p_{\mathrm{T},\ell\ell}$ >45,30 GeV	9.37 ± 0.09	74.27 ± 4.82	6.21 ± 0.41	7.47 ± 0.80	4.83 ± 0.68	22.32 ± 5.13	1.03 ± 0.73	1.33 ± 0.54	116.43 ± 10.02	138	1.19 ± 0.14

Table 4.6: Expected numbers of signal ($m_H = 125$GeV) and background events for the 0-jet analysis in $4.7\,\mathrm{fb}^{-1}$ of integrated luminosity, as well as the observed numbers of events in data. The yields for the ee, $e\mu$, and $\mu\mu$ channels are added together, the composition in each of the lepton flavor channels is shown only for the final stage of the selections. The W+jets background is entirely determined from data, whereas for the other processes the expectations are taken from simulation with $Z/\gamma^* +$ jets, $t\bar{t}$, and $tW/tb/tqb$ corrected by the normalisation factors from the data driven control regions. The uncertainties shown are the statistical uncertainties.

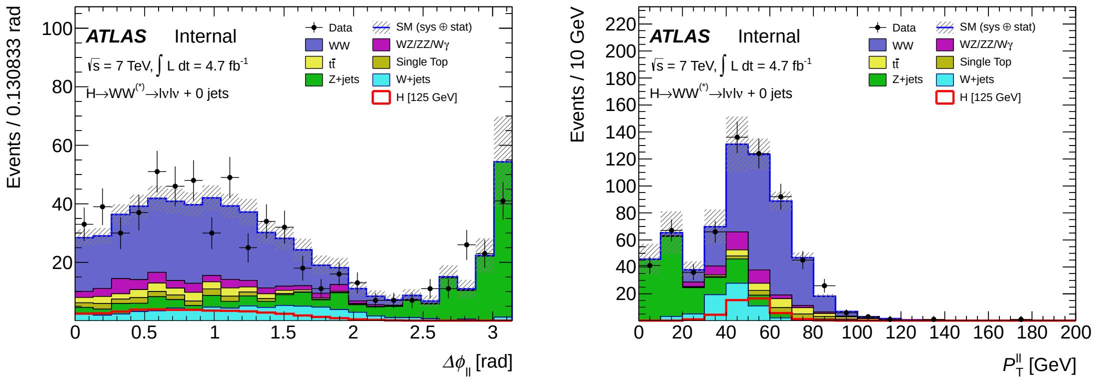


Figure 4.7: Distributions after the invariant mass cut of the azimuthal opening angle $\Delta\phi_{ll}$ left figure and the transverse momentum of the dilepton system p_{T}^{ll} is shown in the right figure. The expected signal is shown for $m_H = 125\,\mathrm{GeV}$. The lower part of each plot shows the ratio between the data and the background expectation from MC, with the yellow band indicating the total systematic uncertainty in the normalization (but not the shape) of the various components.

together, the composition in each of the lepton flavor channels is shown only for the final stage of the selections. The W+jets background is determined entirely from data whereas for the other processes the expectations are taken from simulation with Z/γ^* +jets, $t\bar{t}$, and $tW/tb/tqb$ corrected by the normalisation factors described in the next chapter. The systematics shown for WW, $t\bar{t}$, and $tW/tb/tqb$ are not used in the final limit setting. Instead the systematics are used to describe the uncertainty on the ratio of the yields in the control regions to the corresponding expectations in the signal regions. The uncertainties shown here are the statistical uncertainties. Also here, the number of observed data events (56) is higher than the total background contributions (42.35 ± 4.10) in agreement with the observation shown in Fig. 4.11.

Figure 4.12 shows kinematic distributions of $\Delta\eta_{jj}$, m_{jj}, $p_{\mathrm{T}}^{\mathrm{Tot}}$, m_{ll}, $\Delta\phi_{ll}$ and m_T relevant to the 2-jet analysis. Table 4.8 shows the results of cutflows in the 2-jet analysis.

4.5 Systematic Uncertainties

All of the detector related systematic sources are listed in Tab. 4.9. The experimental systematics mainly come from the objects energy resolution, scale and the selection efficiencies including jets, muons and electrons. The systematic of efficiencies of muons and electrons are introduced in Chapter 3 using the tag-and-probe method. The uncertainty of jets is determined from a combination of test beam simulation and in-situ measurements and no more than 14% for jets with p_T >25 GeV and $|\eta| < 4.5$ [103]. The JES (Jets Energy Scale uncertainty) is one of the most important uncertainties to this analysis especially in the 0-jet bin due to its significant influence on JVSP (Jet Veto Survival Probability), which is using to reject the top background against the signal. The uncertainty of missing transverse energy is determined by soft

	Signal [125 GeV]	WW	WZ/ZZ/Wgamma	ttbar	Single Top	Z+jets	W+jets (MC)	W+jets (d-d)	Total Bkg. (d-d)	Observed	Data/MC (d-d)
1j: exactly one jet	22.67 ± 0.13	343.24 ± 53.70	55.76 ± 2.66	1438.29 ± 59.74	435.91 ± 18.87	357.22 ± 16.63	64.23 ± 13.32	85.38 ± 2.90	2715.81 ± 142.44	2706	1.00 ± 0.06
1j: b-jet veto (25 GeV, 80% eff)	20.88 ± 0.13	319.10 ± 49.92	51.99 ± 2.55	411.70 ± 17.54	138.58 ± 6.60	332.40 ± 15.80	62.16 ± 13.22	75.95 ± 2.47	1329.72 ± 83.70	1369	1.03 ± 0.07
1j: $p_{\mathrm{T}}^{\mathrm{tot}} < 30$ GeV	14.03 ± 0.10	225.66 ± 35.32	34.06 ± 2.16	180.58 ± 8.02	80.16 ± 4.16	107.95 ± 7.72	46.59 ± 12.88	37.44 ± 1.69	665.85 ± 51.18	684	1.03 ± 0.09
1j: $Z \to \tau\tau$ veto	13.99 ± 0.10	220.20 ± 34.47	33.60 ± 2.16	173.39 ± 7.72	77.02 ± 4.03	84.59 ± 7.17	46.59 ± 12.88	37.71 ± 1.66	626.50 ± 49.65	644	1.03 ± 0.09
1j: $m_{\ell\ell} < 50$ GeV	10.93 ± 0.09	48.77 ± 7.71	13.54 ± 1.57	32.77 ± 1.92	17.62 ± 1.39	23.63 ± 3.05	23.09 ± 11.73	11.99 ± 0.96	148.33 ± 11.75	170	1.15 ± 0.13
1j: $\Delta\phi_{\ell\ell} < 1.8$	10.10 ± 0.09	44.46 ± 7.03	13.12 ± 1.57	31.06 ± 1.85	17.11 ± 1.37	10.01 ± 1.97	22.44 ± 11.71	9.88 ± 0.82	125.63 ± 10.39	145	1.15 ± 0.14
1j: $0.75 \cdot m_{\mathrm{H}} \leq m_{\mathrm{T}} \leq m_{\mathrm{H}}$	6.30 ± 0.07	15.98 ± 2.57	4.65 ± 0.97	7.86 ± 0.75	4.26 ± 0.61	4.58 ± 1.33	5.87 ± 3.23	5.03 ± 0.54	42.35 ± 4.10	56	1.32 ± 0.22
ee:1j: $\Delta\phi_{\ell\ell} < 1.8$	1.22 ± 0.03	6.56 ± 1.09	2.17 ± 0.63	4.32 ± 0.56	2.53 ± 0.47	1.93 ± 0.82	3.60 ± 3.02	1.64 ± 0.19	19.15 ± 2.08	19	0.99 ± 0.25
$e\mu$:1j: $\Delta\phi_{\ell\ell} < 1.8$	6.39 ± 0.07	25.16 ± 4.01	9.62 ± 1.42	17.69 ± 1.24	9.77 ± 0.98	2.88 ± 1.02	18.39 ± 11.31	7.27 ± 0.69	72.39 ± 5.62	90	1.24 ± 0.16
$\mu\mu$:1j: $\Delta\phi_{\ell\ell} < 1.8$	2.49 ± 0.04	12.74 ± 2.06	1.33 ± 0.20	9.05 ± 0.80	4.80 ± 0.65	5.20 ± 1.33	0.45 ± 0.45	0.97 ± 0.41	34.09 ± 3.58	36	1.06 ± 0.21

Table 4.7: Expected numbers of signal ($m_H = 125\,\mathrm{GeV}$) and background events for the 1-jet analysis in $4.7\,\mathrm{fb}^{-1}$ of integrated luminosity, as well as the observed numbers of events in data. The yields for the ee, $e\mu$, and $\mu\mu$ channels are added together, the composition in each of the lepton flavor channels is shown only for the final stage of the selections. The W+jets background is determined entirely from data, whereas for the other processes the expectations are taken from simulation with $Z/\gamma^* +$jets, $t\bar{t}$, and $tW/tb/tqb$ corrected by the normalisation factors described in chapter 5. The systematics shown for WW, $t\bar{t}$, and $tW/tb/tqb$ are not used in the final limit setting. Instead the systematics used describe the uncertainty on the ratio of the yields in the control regions to the corresponding expectations in the signal regions. The uncertainties shown are the statistical uncertainties.

	Signal [125 GeV]	WW	WZ/ZZ/Wgamma	ttbar	Single Top	Z+jets	W+jets (MC)	W+jets (d-d)	Total Bkg. (d-d)	Observed	Data/MC (d-d)
$\geq$ 2j: at least 2 jets	11.42 ± 0.09	142.27 ± 1.73	25.50 ± 1.53	5555.59 ± 16.11	313.60 ± 4.81	120.09 ± 6.68	13.46 ± 2.70	39.64 ± 3.66	6196.70 ± 18.60	6676	1.08 ± 0.01
$\geq$ 2j: central jet veto	8.95 ± 0.08	113.27 ± 1.55	19.54 ± 1.44	3278.61 ± 12.68	238.37 ± 4.33	89.49 ± 5.94	10.61 ± 2.38	25.46 ± 2.60	3764.75 ± 15.03	3811	1.01 ± 0.02
$\geq$ 2j: b-jet veto ($p_{\mathrm{T}} \geq 25$ GeV, 80% eff)	7.56 ± 0.07	97.84 ± 1.42	17.53 ± 1.38	353.21 ± 4.25	50.56 ± 2.00	77.36 ± 5.48	9.42 ± 2.23	18.86 ± 1.58	615.36 ± 7.65	667	1.08 ± 0.04
$\geq$ 2j: opp. hemispheres	4.17 ± 0.05	46.05 ± 0.97	7.27 ± 0.84	149.36 ± 2.73	20.83 ± 1.28	31.56 ± 3.39	3.59 ± 1.36	8.62 ± 1.06	263.69 ± 4.83	269	1.02 ± 0.06
$\geq$ 2j: $\Delta\eta_{\mathrm{jj}} > 3.8$	1.81 ± 0.02	8.39 ± 0.41	0.92 ± 0.17	23.19 ± 0.98	2.15 ± 0.37	5.79 ± 1.68	0.00 ± 0.00	1.72 ± 0.41	42.17 ± 2.07	40	0.95 ± 0.16
$\geq$ 2j: $m_{\mathrm{jj}} > 500$ GeV	1.28 ± 0.02	3.93 ± 0.28	0.38 ± 0.11	10.39 ± 0.61	1.04 ± 0.25	0.71 ± 0.35	0.00 ± 0.00	0.89 ± 0.32	17.34 ± 0.87	13	0.75 ± 0.21
$\geq$ 2j: $p_{\mathrm{T,tot}} < 30$ GeV	0.95 ± 0.01	2.51 ± 0.22	0.21 ± 0.08	4.29 ± 0.39	0.63 ± 0.20	0.27 ± 0.20	0.00 ± 0.00	0.38 ± 0.22	8.29 ± 0.58	6	0.72 ± 0.30
$\geq$ 2j: $Z \rightarrow \tau\tau$ veto	0.94 ± 0.01	2.45 ± 0.21	0.21 ± 0.08	4.24 ± 0.39	0.63 ± 0.20	0.16 ± 0.16	0.00 ± 0.00	0.32 ± 0.21	8.01 ± 0.56	6	0.75 ± 0.31
$\geq$ 2j: $m_{\ell\ell} < 80$ GeV	0.94 ± 0.01	1.08 ± 0.15	0.10 ± 0.06	1.36 ± 0.24	0.35 ± 0.14	0.16 ± 0.16	0.00 ± 0.00	0.18 ± 0.18	3.22 ± 0.40	2	0.62 ± 0.45
$\geq$ 2j: $\Delta\phi_{\ell\ell} < 1.8$	0.82 ± 0.01	0.75 ± 0.12	0.09 ± 0.06	0.85 ± 0.20	0.14 ± 0.10	0.00 ± 0.00	0.00 ± 0.00	-0.04 ± 0.03	1.79 ± 0.27	1	0.56 ± 0.56
ee:$\geq$ 2j: $\Delta\phi_{\ell\ell} < 1.8$	0.10 ± 0.00	0.10 ± 0.05	0.01 ± 0.01	0.03 ± 0.05	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	-0.01 ± 0.01	0.13 ± 0.07	0	0.00 ± 0.00
$e\mu$:$\geq$ 2j: $\Delta\phi_{\ell\ell} < 1.8$	0.53 ± 0.01	0.45 ± 0.09	0.08 ± 0.06	0.53 ± 0.17	0.07 ± 0.07	0.00 ± 0.00	0.00 ± 0.00	-0.02 ± 0.03	1.12 ± 0.21	1	0.89 ± 0.91
$\mu\mu$:$\geq$ 2j: $\Delta\phi_{\ell\ell} < 1.8$	0.19 ± 0.01	0.20 ± 0.07	0.00 ± 0.00	0.28 ± 0.11	0.07 ± 0.07	0.00 ± 0.00	0.00 ± 0.00	-0.01 ± 0.01	0.55 ± 0.15	0	0.00 ± 0.00

Table 4.8: Expected numbers of signal and background events for the 2-jet analysis in $4.7\,\mathrm{fb}^{-1}$ of integrated luminosity. The signal numbers correspond to the case of $m_H = 125\,\mathrm{GeV}$. The yields for the ee, $e\mu$, and $\mu\mu$ channels are added together; the composition in each of the lepton flavor channels is shown only for the final stage of the selections. The uncertainties shown are the statistical one.

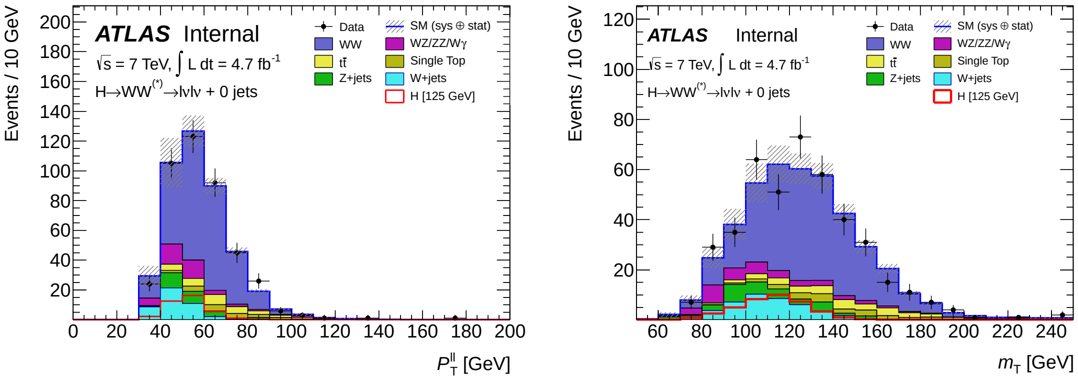


Figure 4.8: Transverse momentum p_{T}^{ll} of the two selected leptons is shown in the left figure after the cut on the azimuthal opening angle $\Delta\phi_{ll}$ and the transverse mass m_T distribution is shown in the right figure after all cuts. The expected signal is shown in both plots for $m_H = 125\,\mathrm{GeV}$. The hatched band around the sum of backgrounds indicates the total systematic uncertainty in the normalization (but not the shape) of the various components. The final bin includes the overflow.

calorimeter energy deposits and all of the uncertainties which are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. In addition, when a b-tagging is needed, the b-jet tagging efficiency uncertainty is introduced by varying the efficiency between 5% and 14% as a function of jet p_T [81]. The uncertainty on the integrated luminosity in 2011 is 3.9% [104].

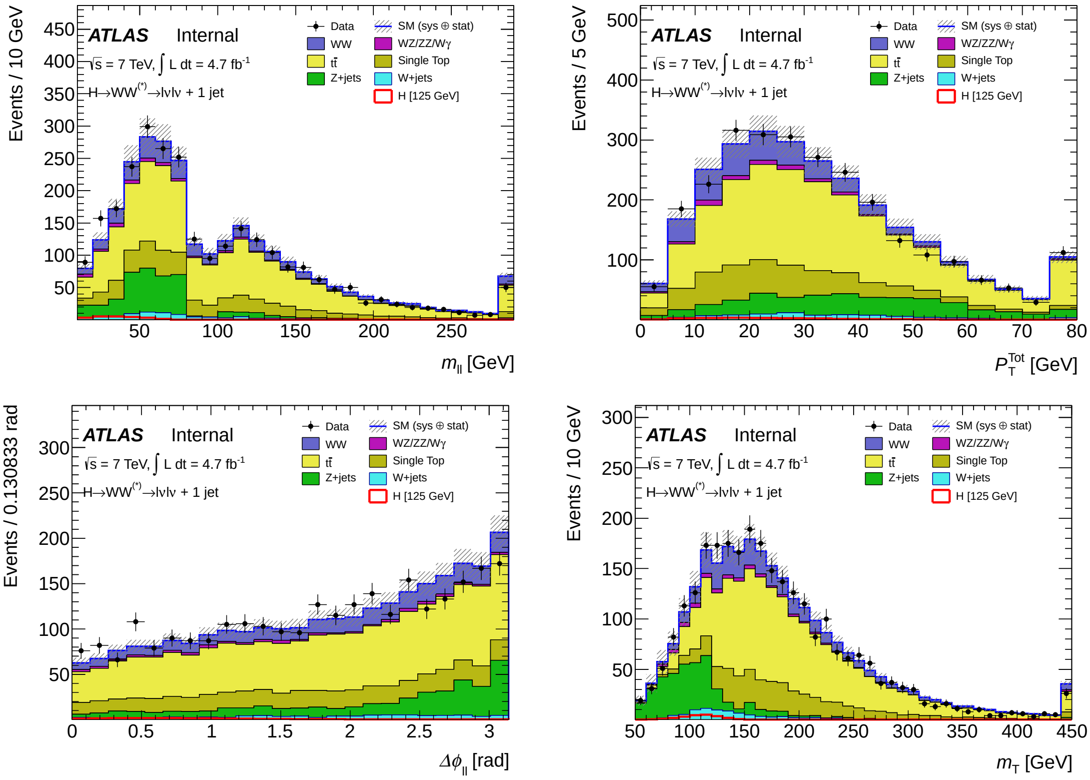


Figure 4.9: Kinematic distributions relevant to the 1-jet analysis after the one jet requirement has been applied: dilepton invariant mass (top left), total transverse momentum (top right), and azimuthal opening angle (bottom left) transverse mass (bottom left). The expected signal is shown for $m_H = 125$ GeV. The hatched band around the sum of backgrounds indicates the total systematic uncertainty in the normalization (but not the shape) of the various components. The final bin includes the overflow.

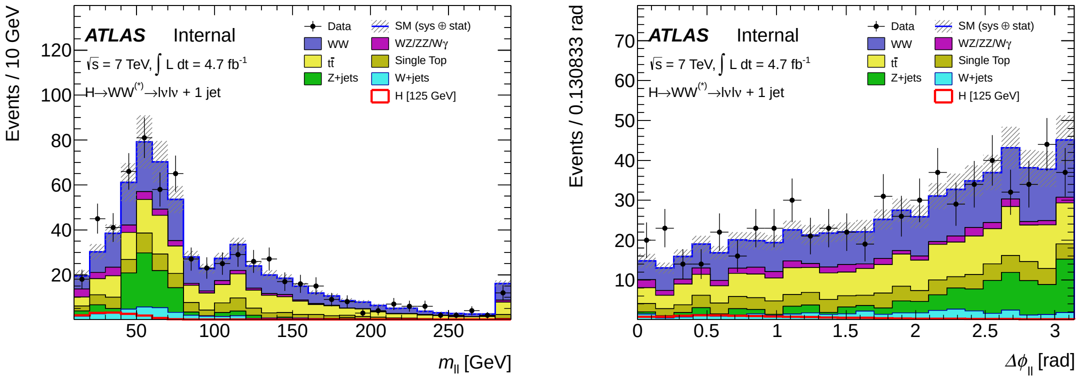


Figure 4.10: The invariant mass $m_{\ell\ell}$ of the two selected leptons (left) and the azimuthal opening angle of the leptons after the $Z \to \tau\tau$ veto (right). The expected signal is shown in both plots for $m_H = 125\,\text{GeV}$. The hatched band around the sum of backgrounds indicates the total systematic uncertainty in the normalization (but not the shape) of the various components. The final bin includes the overflow.

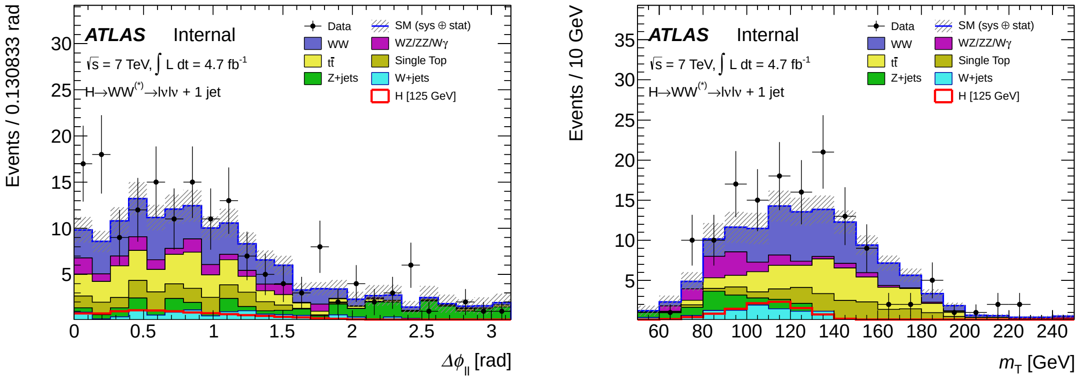


Figure 4.11: Azimuthal opening angle $\Delta\phi_{\ell\ell}$ of the two selected leptons (left) after the cut on the invariant mass of the two selected leptons and the jet and the transverse mass (right) after all cuts. The expected signal is shown for $m_H = 125\,\text{GeV}$. The hatched band around the sum of backgrounds indicates the total systematic uncertainty in the normalization (but not the shape) of the various components. The final bin includes the overflow.

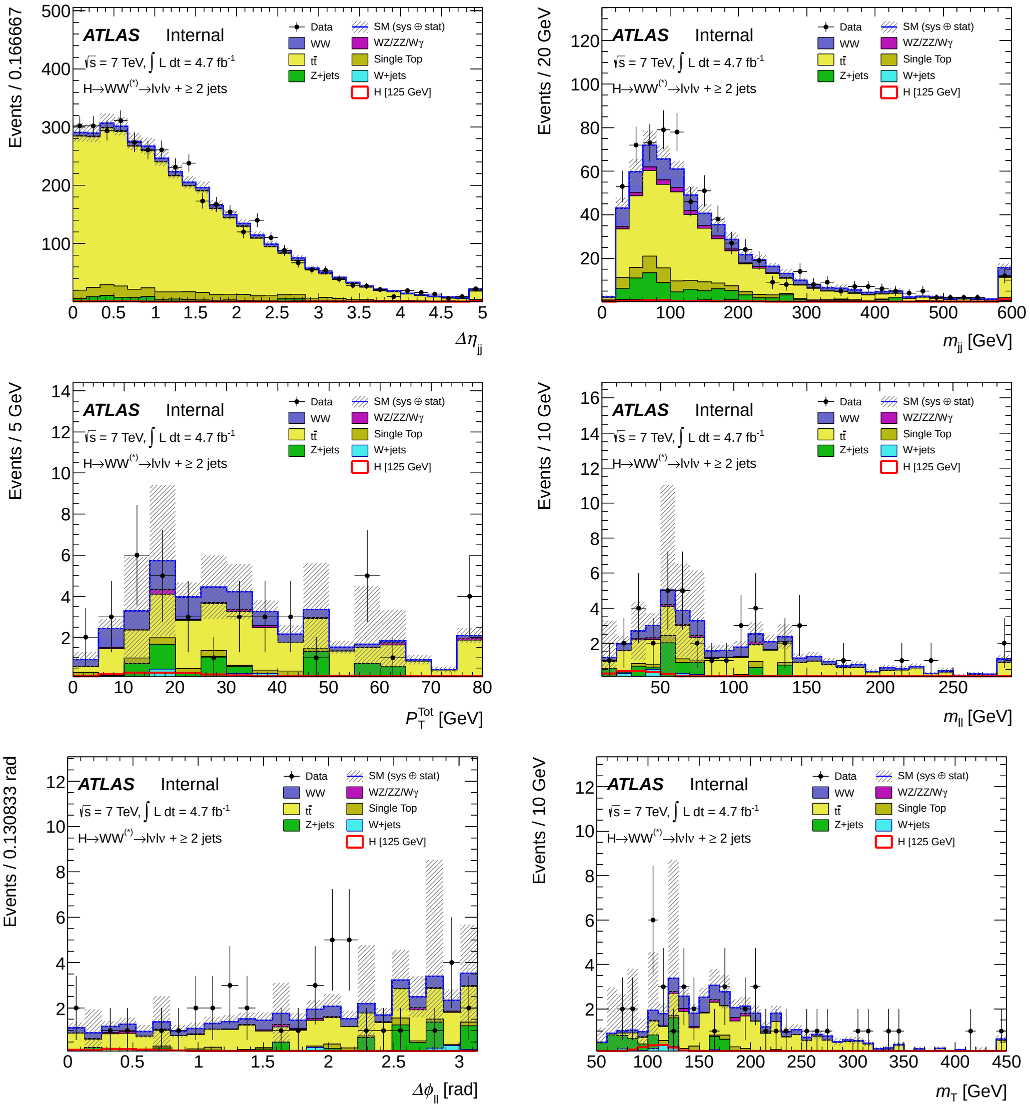


Figure 4.12: Kinematic distributions relevant to the 2-jet analysis. Top: $\Delta\eta_{jj}$ after requiring that the two leading jets are in opposite hemispheres. Upper right: m_{jj} after the cut on $\Delta\eta_{jj}$. The remaining plots are shown after the Central Jet Veto cut. Middle left: the distribution of p_T^{Tot}. Middle right: the distribution of $\Delta\phi_{ll}$. Lower left: the distribution of the dilepton invariant mass. Lower right: the distribution of m_{T}. The expected signal is shown for $m_H = 125\,\mathrm{GeV}$. The hatched band around the sum of backgrounds indicates the total systematic uncertainty in the normalization (but not the shape) of the various components. The final bin includes the overflow.

Source of Uncertainty	Treatment in the analysis
Jet Energy Resolution (JER)	MC jet resolution smeared using jet p_T, η-dependent parametrization
Jet Energy Scale (JES)	global JES: $< 14\%$ for jet $p_T > 25$ GeV and $\|\eta\| < 4.5$ pileup: $< 5\%$ for jet $p_T > 20$ GeV
Electron Selection Efficiency	Separate systematics for electron identification, reconstruction and isolation, added in quadrature Identification: 8% for $p_T < 15$ GeV, decreasing to 1% for $p_T > 30$ GeV in the central region Reconstruction: 0.6 - 1.2% for $p_T > 15$ GeV trigger: 1% uncertainty Total uncertainty of 2-5% depending on η and E_T
Electron Energy Scale	Uncertainty smaller than 1%, depending on η and E_T
Electron Energy Resolution	Energy varied within its uncertainty, 0.6% of the energy at most
Muon Selection Efficiency	0.3-1% as a function of η and p_T reconstruction smaller than 1%
Muon Momentum Scale and Resolution	Uncertainty smaller than 1%
b-tagging Efficiency	p_T dependent scale factor uncertainties, 4.8 - 13.7%
Missing Transverse Energy	Using METUtility package
Missing Transverse Energy PileUP	10% from JetTauEtMiss 2010 recommendations
Luminosity	3.9%

Table 4.9: Experimental sources of systematic uncertainty per object or event.

Chapter 5

Background Estimation

5.1 Introduction

In the $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$ analysis as in other Higgs search analyses, it is extremely important to estimate reliably the background contributions. Therefore, whenever it is possible, data-driven methods are designed and used either to replace the corresponding MC prediction or to check the MC prediction both on the normalisation and on the shape of the relevant kinematic distributions.

In the Higgs to $WW^{(*)}$ analysis, the main backgrounds are QCD(multijet or dijet), W+jets, Z+jets, top, and the Standard Model diboson processes: WW, WZ, ZZ and $W\gamma$. QCD is strongly suppressed to a negligible level by the requirement of two very tight isolated leptons, the 0-jet veto selection and the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut [107]. The W+jet events have a huge initial rate and can not be reliably simulated with sufficiently large MC sample. It is thus derived fully from data. The Z+jets can be suppressed by the Z mass peak veto and the $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{ll} cuts. However, the shape of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is very sensitive to the pile up effects in 2011 and especially in 2012 high luminosity data taking periods. Any mismodelling of the $E_{\mathrm{T}}^{\mathrm{miss}}$ affects the final estimate of the Z+jets contribution. A data-driven method is needed to correct such mismodelling of the $E_{\mathrm{T}}^{\mathrm{miss}}$. The top events have two isolated leptons and large $E_{\mathrm{T}}^{\mathrm{miss}}$ associated with two additional b jets. It can be suppressed by jet veto cut in the 0-jet analysis or b-jet veto in 1- and 2-jet bins. However, the jet energy scale uncertainty is the most significant experimental systematic uncertainty, which can vary up to 20% in rejecting the top background. So a data-driven method with reduced sensitivity to the JES variation is essential. The Standard Model WW is an irreducible background which has the same topology as the signal. The corresponding MC prediction, though fairly reliable, is verified using a dedicated control region selected from the dilepton invariant mass spectrum far away from the Higgs signal region. For the other backgrounds, either they are negligible or the MC predictions are checked to be reliable.

In this chapter, the data driven methods to estimate W+jets, top, Z+jets are described and discussed in detail in some cases to which we have directly contributed.

5.2 W+jets Background Estimation

As mentioned above, the W+jets background has huge cross section and can have an additional fake lepton coming from jet. The idea to evaluate the expected W+jets contribution is to estimate the number of events from a W+jets enriched data sample. The events of the sample are selected to have a real lepton defined as one identified (id) object and an anti-identified fake lepton. The number of events which pass the two lepton requirement is defined as:

$$\begin{aligned}
\mathrm{N}^{ee\text{-ch}}_{\text{one id + one fake}} &= \mathrm{f}_e \times \mathrm{N}_{\text{one id e + one anti-id e}} && (5.1)\\
\mathrm{N}^{\mu\mu\text{-ch}}_{\text{one id + one fake}} &= \mathrm{f}_\mu \times \mathrm{N}_{\text{one id }\mu\text{ + one anti-id }\mu} && (5.2)\\
\mathrm{N}^{e\mu\text{-ch}}_{\text{one id + one fake}} &= \mathrm{f}_e \times \mathrm{N}_{\text{one id }\mu\text{ + one anti-id e}} + \mathrm{f}_\mu \times \mathrm{N}_{\text{one id e + one anti-id }\mu}\,, && (5.3)
\end{aligned}$$

where f_e, f_μ are the electron and muon fake factors, which are obtained from jet enriched data samples after EF_g20_etcut trigger requirement:

$$f_e \equiv \frac{\mathrm{N}_{\text{id e}}}{\mathrm{N}_{\text{anti-id e}}}, \quad f_\mu \equiv \frac{\mathrm{N}_{\text{id }\mu}}{\mathrm{N}_{\text{anti-id }\mu}}. \tag{5.4}$$

Here the identified leptons are using the nominal selection as described in Chapter 3. The anti-identification number is defined as shown in Tab. 5.1.

Anti-id electron	Anti-id muon
Same p_T, η range as identified electrons	Same p_T, η range as identified muons
$N_{\text{hit}}(\text{SCT} + \text{Pixel}) \geq 4$	Same ID track requirement as id muons
$\lvert z_0\rvert < 1\,\text{mm}$	$\lvert z_0\rvert < 1\,\text{mm}$
Isolation	Isolation
Identified electrons removed	Identified muons removed

Table 5.1: Definitions of anti-identified electrons and muons for the fake factor determination.

The fake factor as a function of lepton transverse momentum is shown in Fig. 5.1. Therefore, the W+jets enriched sample using fake rates as event weights provide the final data-driven estimate of the W+jets contribution, after subtracting other background contaminations. This sample provide not only the number of events but also kinematic distributions of the W+jets background. Table 5.2 shows the number of W+jets events estimated from data in the Higgs to $WW^{(*)}$ cutflows.

The considered systematic sources of this method include the real lepton contribution in the fake rates, the sample and pile up dependence and the trigger bias, in which the difference of fake rate derived with or without trigger requirement was taken into account. By rejecting events which have an invariant mass in the Z mass region or a transverse mass greater than 30 GeV, the EW contribution reduces to 10% $\sim$ 20%, resulting in a 10% uncertainty in the remaining MC subtraction. The study shows that there is no trigger bias and pile up dependence on the fake rate.

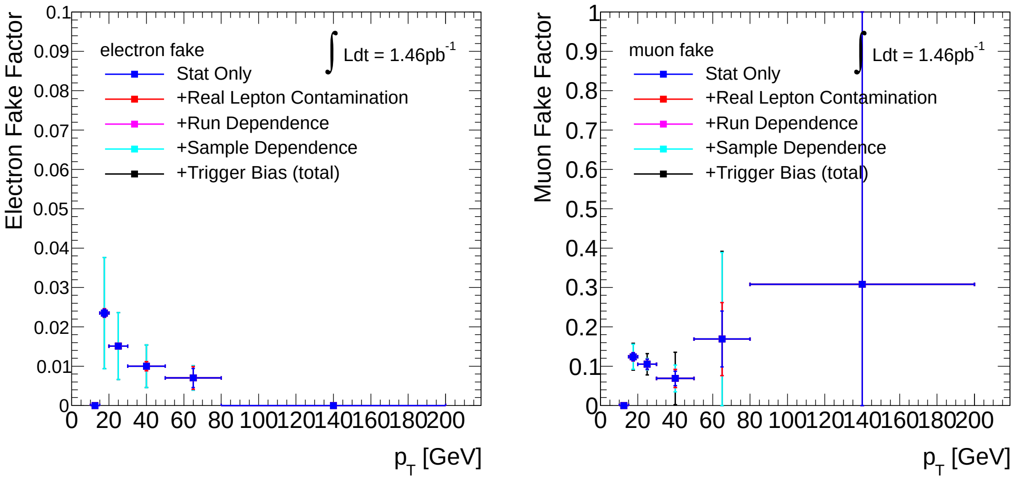


Figure 5.1: Fake factor as a function of p_T including all systematic uncertainties for electrons (left) and muons (right). This trigger is highly prescaled and the equivalent luminosity is 1.46 pb^{-1} using the full 2011 data set.

The difference of fake rate derived from different MC samples, dijet or W+jets, gives 20% $\sim$ 50% variation depending on p_T.

The result can be validated with the same sign event sample. In this same sign sample, the contamination from other non W+jets backgrounds is small so that W+jets have a high puirty. Good agreement is observed between data and MC in the shapes of kinematic distributions, as shown in Fig. 5.2. Using this method, the W+jet background estimate is not statistically limited contrary to the corresponding MC prediction.

5.3 Top Background Estimation

After the common event selection (Section 4.4.1), the top background is the dominant contribution in the inclusive event sample having two isolated leptons and large $E_{\mathrm{T}}^{\mathrm{miss}}$. It accounts for a small contribution in the total backgrounds in the 0-jet bin but dominates in the 1- and 2-jet bins. The number of top events in the 0-jet analysis, however, is directly proportional to the full Jet Veto Survival Probability (jet veto efficiency, JVSP). Studies based on top MC simulations show that JVSP is subject to various systematic variation such as JES, JER and pile up. Therefore, it is necessary to design a method to estimate reliably the top contribution in 0-jet channel with reduced sensitivities to different systematic variations [111].

5.3.1 Description of the Top Process

The inclusive process $pp \rightarrow W^+W^-b\bar{b}$ with subsequent W decays gives rise to the same leptonic final state with $E_{\mathrm{T}}^{\mathrm{miss}}$ as the Higgs signal events. The phase space of the inclusive process may be divided into three parts corresponding to double resonance

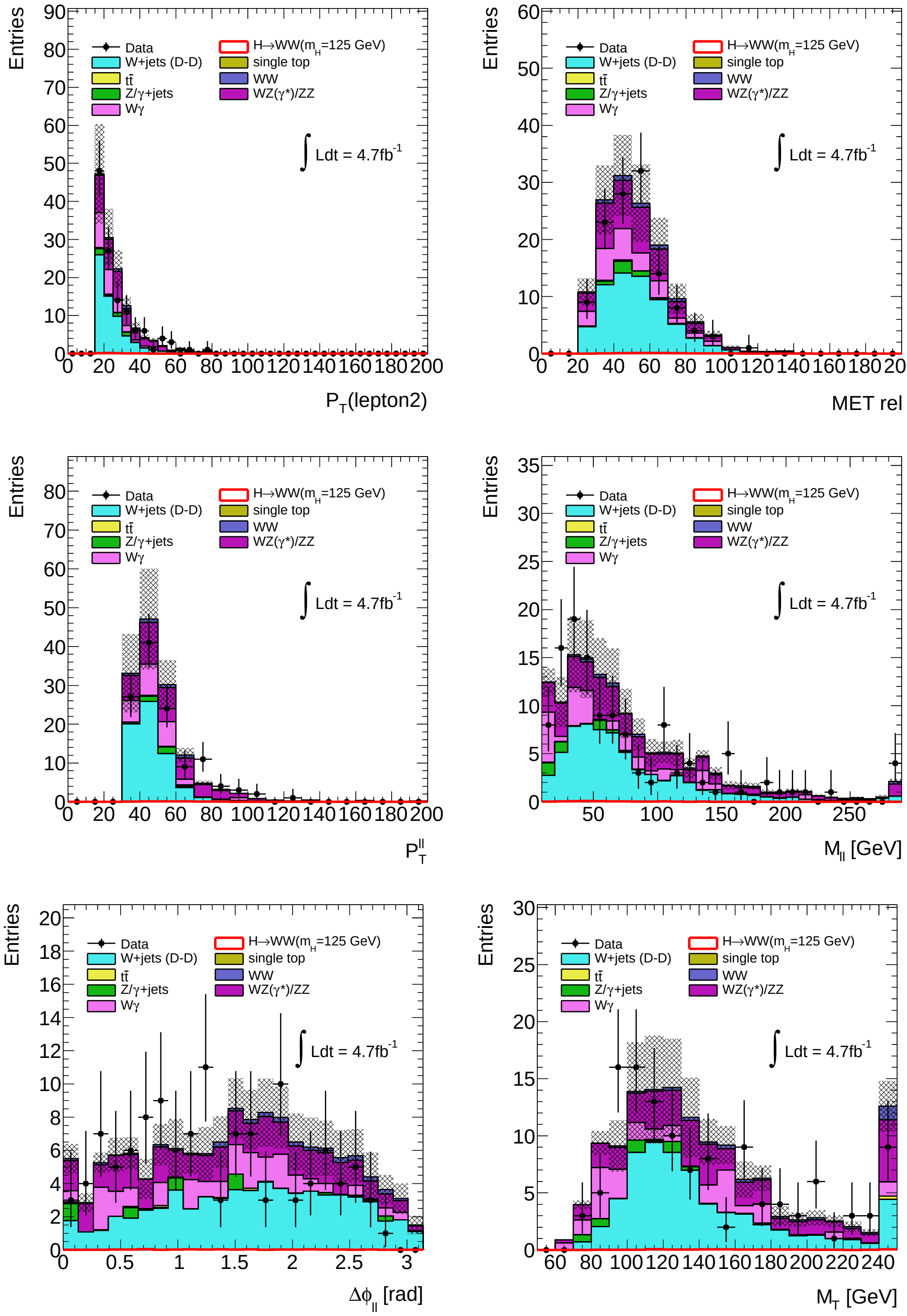


Figure 5.2: Various kinematic distributions in the same-sign W+jets control region before topological selections are applied. The W+jets shape is obtained using the data-driven method described in this section. Other background contributions are estimated using MC. The observed data agree well with the estimated W+jets contribution.

SR ($m_H = 125\,\text{GeV}$)	ee-channel	$\mu\mu$-channel	$e\mu$-channel	Total
0jet	39.3 ± 0.77 ± 21.8	12.2 ± 1.46 ± 3.62	164 ± 1.96 ± 89	216 ± 2.57 ± 114
pt(ll)	17.3 ± 0.48 ± 9.34	3.7 ± 0.79 ± 0.75	124 ± 1.59 ± 66.6	145 ± 1.84 ± 76.7
Mll	7.81 ± 0.36 ± 4.44	1.45 ± 0.50 ± 0.46	39.9 ± 0.99 ± 20.9	49.2 ± 1.16 ± 25.8
Dphi(ll)	7.76 ± 0.36 ± 4.42	1.45 ± 0.50 ± 0.46	34.9 ± 0.91 ± 18.4	44.1 ± 1.1 ± 23.3
MT(0jet)	4.83 ± 0.30 ± 2.79	1.19 ± 0.42 ± 0.39	21.4 ± 0.72 ± 11.2	27.4 ± 0.88 ± 14.4
1jet	12.1 ± 0.45 ± 6.54	8.61 ± 1.37 ± 7.96	62.2 ± 1.35 ± 34.9	82.9 ± 1.97 ± 49.4
bjet veto	11.2 ± 0.41 ± 6.03	6.12 ± 1.09 ± 6.07	56.3 ± 1.2 ± 30.4	73.6 ± 1.67 ± 42.5
pt(tot)	5.06 ± 0.29 ± 2.8	2.88 ± 0.71 ± 2.47	28.3 ± 0.89 ± 13	36.2 ± 1.17 ± 18.2
Ztt veto	5.03 ± 0.29 ± 2.79	2.94 ± 0.71 ± 2.6	28.6 ± 0.87 ± 13.1	36.5 ± 1.16 ± 18.4
Mll	1.98 ± 0.20 ± 1.14	1.25 ± 0.45 ± 0.55	8.41 ± 0.53 ± 4.23	11.6 ± 0.72 ± 5.92
Dphi(ll)	1.61 ± 0.18 ± 0.92	0.96 ± 0.39 ± 0.47	7.05 ± 0.45 ± 3.48	9.62 ± 0.62 ± 4.87
MT(1jet)	1.17 ± 0.15 ± 0.68	0.45 ± 0.28 ± 0.09	3.26 ± 0.30 ± 1.6	4.88 ± 0.44 ± 2.37

Table 5.2: Expected W+jets event rates in 2011 all year for the signal region, as estimated using the W+jets enriched control sample.

($t\bar{t}$), single resonance (single top) and non-resonance regions. In the MC simulations, one usually procedures several distinct processes corresponding to $t\bar{t}$ and single top in t- and s-channels and Wt final state (cf. Tab. 4.1). Therefore, it is not appropriate to use the sum of squared amplitudes of the $t\bar{t}$ and single top productions to describe the corresponding inclusive sample. In such a case, the non-resonant $pp \to W^+W^-b\bar{b}$ events and the effect of interference between single top Wt and $t\bar{t}$ can not be taken into account. This problem may be solved with the computation of QCD corrections at NLO to the complete set of $pp \to W^+W^-b\bar{b}$ diagrams with their implementation in MC@NLO or POWHEG. Before the appearance of such a tool it is necessary to resort to the LO description. The first results of a NLO QCD calculation for the full set of $pp \to W^+W^-b\bar{b}$ have been reported [108] and [109]. This represents significant progress towards a proper description of corners of the phase space of interest here.

The automated program MADGRAPH [110] is used to generate MC events corresponding to the various processes considered here. The impact of QCD higher order corrections and hadronization, evaluated with the package MC@NLO, is obtained with the Monte Carlo truth (hadronic) without the application of detector effects. In both cases default settings are used. The CTEQ6L1 and CTEQ6M parton density parameterizations are used, when appropriate. The renormalization and factorization scales are set to the top mass. No constraint is applied on the transverse momentum of the b-partons and these are required to be separated by $\Delta R > 0.4$. The decays of the W boson are treated with the narrow width approximation.

The $t\bar{t}$ process is by far the dominant contribution of the top background $pp \to W^+W^-b\bar{b}$. This is illustrated in Table 5.3, where the LO cross-sections in pb calculated using the MADGRAPH package by setting the factorization and renormalization scales to the top mass are shown for the processes $pp \to W^+W^-b\bar{b}$ and $pp \to t\bar{t} \to W^+W^-b\bar{b}$. The results in Table 5.3 do not include the branching ratios of the W bosons to leptons.

Process	7 TeV	14 TeV
$pp \to W^+W^-b\bar{b}$	103	625
$pp \to t\bar{t} \to W^+W^-b\bar{b}$	99.6	597

Table 5.3: Cross-sections (in pb) at LO for the processes $pp \to W^+W^-b\bar{b}$ and $pp \to t\bar{t} \to W^+W^-b\bar{b}$. Results are shown for two different values of the center-of-mass energy of the proton-proton collision. The MADGRAPH package is used to calculate the cross-sections.

5.3.2 Methodology and MC Truth Study

To extract the top contributions in the 0-jet analysis, two event samples are defined: one data sample with the requirement of two isolated leptons and large $E_{\mathrm{T}}^{\mathrm{miss}}$ and one data and/or MC control sample which has at least b-tagged jet in addtion to the dilepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement. From the first sample in data, which covers both the signal region and the control region, the number of expected top contribution can be estimated as:

$$N_{\mathrm{Top}}^{\mathrm{Exp}}(\ell\ell + E_{\mathrm{T}}^{\mathrm{miss}}) \approx N_{\mathrm{data}}^{\mathrm{all}}(\ell\ell + E_{\mathrm{T}}^{\mathrm{miss}}) - N_{\text{non-top}}^{\mathrm{MC}}(\ell\ell + E_{\mathrm{T}}^{\mathrm{miss}})\,, \tag{5.5}$$

where $N_{\mathrm{data}}^{\mathrm{all}}$ is the number of events observed in data and $N_{\text{non-top}}^{\mathrm{MC}}$ is the corresponding non-top contributions based on either a MC or a data-driven estimate. The b-jet requirement in the second sample allows a highly pure top sample to be selected, which can then be used to determine JVSP in both data and MC.

In the MC truth level, the study is performed using LO matrix elements. For a given region of phase space defined with a lower bound on the transverse momentum, p_T, and a maximum value of the parton pseudorapidity, η, the probability of a parton to miss such a region is defined as P_1:

$$P_1 = \left(\int_0^{p_T^v} \int_0^{|\eta^v|} \frac{d\sigma^{\mathrm{cuts}}}{dp_{Tb}\,d|\eta_b|} dp_{Tb}\,d|\eta_b| + \int_0^{\infty} \int_{|\eta^v|}^{\infty} \frac{d\sigma^{\mathrm{cuts}}}{dp_{Tb}\,d|\eta_b|} dp_{Tb}\,d|\eta_b| \right) / \sigma_{\mathrm{tot}}^{\mathrm{cuts}}\,, \tag{5.6}$$

where the integration over the p_T and $|\eta|$ of the $\bar{b}$ parton is assumed and

$$\sigma_{\mathrm{tot}}^{\mathrm{cuts}} = \int_0^{\infty} \int_0^{\infty} \frac{d\sigma^{\mathrm{cuts}}}{dp_{Tb}\,dp_{T\bar{b}}} dp_{Tb}\,dp_{T\bar{b}}\,, \tag{5.7}$$

σ^{cuts} is the cross-section of the top related backgrounds after the requirements. The expression for the full-jet veto survival probability at LO can be written as:

$$\begin{aligned} P_2 \;=\; & \left(\int \int_0^{p_T^v} \int \int_0^{|\eta^v|} \frac{d\sigma^{\mathrm{cuts}}}{dp_{Tb}\,dp_{T\bar{b}}\,d|\eta_b|\,d|\eta_{\bar{b}}|} dp_{Tb}\,dp_{T\bar{b}}\,d|\eta_b|\,d|\eta_{\bar{b}}| \right. \\ + \; & 2 \int_0^{\infty} \int_0^{|\eta^v|} \int_0^{p_T^v} \int_{|\eta^v|}^{\infty} \frac{d\sigma^{\mathrm{cuts}}}{dp_{Tb}\,dp_{T\bar{b}}\,d|\eta_b|\,d|\eta_{\bar{b}}|} dp_{Tb}\,d|\eta_b|\,dp_{T\bar{b}}\,d|\eta_{\bar{b}}| \\ + \; & \left. \int_0^{\infty} \int_{|\eta^v|}^{\infty} \int_0^{\infty} \int_{|\eta^v|}^{\infty} \frac{d\sigma^{\mathrm{cuts}}}{dp_{Tb}\,dp_{T\bar{b}}\,d|\eta_b|\,d|\eta_{\bar{b}}|} dp_{Tb}\,d|\eta_b|\,dp_{T\bar{b}}\,d|\eta_{\bar{b}}| \right) / \sigma_{\mathrm{tot}}^{\mathrm{cuts}}\,. \end{aligned} \tag{5.8}$$

It is important to note that the second term in Expression (5.6) and the last two terms in Expression (5.8) become significantly smaller than the first terms for values of $|\eta^v| > 3$. These terms become negligible for $|\eta^v| > 4$.

The survival probability relevant to the second control sample is defined as:

$$P_1^{\mathrm{Btag}} = \left(\int_0^{p_T^v} \int_0^{|\eta^v|} \frac{d\sigma^{\mathrm{Btag}}}{dp_T d|\eta|} dp_T d|\eta| + \int_0^{\infty} \int_{|\eta^v|}^{\infty} \frac{d\sigma^{\mathrm{Btag}}}{dp_T d|\eta|} dp_T d|\eta| \right) / \sigma_{\mathrm{tot}}^{\mathrm{Btag}}, \tag{5.9}$$

where $\frac{d\sigma^{\mathrm{Btag}}}{dp_T d|\eta|}$ and $\sigma_{\mathrm{tot}}^{\mathrm{Btag}}$ are the double differential and the total cross-section with the requirement that the event has at least one b-tagged parton. It is worth noting that the second term in Expression (5.9) is significantly smaller than the first.

The measured probability $P_1^{\mathrm{Btag,Exp}}$ can be obtained from the reconstructed probe jet multiplicity distribution in the b-tagged control sample by

$$P_1^{\mathrm{Btag,Exp}} = \frac{N_{\mathrm{Top}}^{\mathrm{Exp}}(\ell\ell + E_{\mathrm{T}}^{\mathrm{miss}} + b\text{-tagged}, 0j)}{N_{\mathrm{Top}}^{\mathrm{Exp}}(\ell\ell + E_{\mathrm{T}}^{\mathrm{miss}} + b\text{-tagged})} \tag{5.10}$$

where the denominator is the total number of events of the b-tagged samples and the numerator refers to the number of events having zero probe jet. The probe jet is defined as follows:

- Consider all jets reconstructed with the nominal jet selection in each event with at least one tagged b-jet,
- For events with more than one tagged b-jets, select the b-jet which has the largest b-tagged weight,
- Any remaining jet is counted as a probe jet if it has a ΔR with the b-tagged jet greater than 1.

For the LO parton level study performed in [111], the b parton is tagged using a random number as there are always two b parton in a top event.

The corresponding experimental full-jet veto survival probability P_2^{Exp} can be derived using the jet multiplicity distribution:

$$P_2^{\mathrm{Exp}} = \frac{N_{\mathrm{Top}}^{\mathrm{Exp}}(\ell\ell + E_{\mathrm{T}}^{\mathrm{miss}}, 0j)}{N_{\mathrm{Top}}^{\mathrm{Exp}}(\ell\ell + E_{\mathrm{T}}^{\mathrm{miss}})} \tag{5.11}$$

where the numerator and the denominator are the number of top background events after the application of a full-jet veto and the total number of top background events without the application of any requirements on the jet multiplicity in the final state, respectively.

When these probabilities are much smaller than one for low p_T thresholds, one can assume $P_2^{\mathrm{Exp}} \approx \epsilon_0^2 P_2$ and $P_1^{\mathrm{Btag,Exp}} \approx \epsilon_0 P_1^{\mathrm{Btag}}$ where ϵ_0 is an experimental efficiency of observing a low p_T jet in the region of interest. These relations may be rewritten as

$$\frac{P_2^{\mathrm{Exp}}}{\left(P_1^{\mathrm{Btag,Exp}}\right)^2} \approx \frac{P_2}{\left(P_1^{\mathrm{Btag}}\right)^2} . \tag{5.12}$$

The data-driven prediction will rely on the rate of top-related events observed in data:

$$\begin{aligned} N^{\mathrm{Exp}}_{\mathrm{Top}}(\ell\ell + E^{\mathrm{miss}}_{\mathrm{T}}, 0j) &\approx N^{\mathrm{Exp}}_{\mathrm{Top}}(\ell\ell + E^{\mathrm{miss}}_{\mathrm{T}}) \times P^{\mathrm{Exp}}_2 \\ &\approx N^{\mathrm{Exp}}_{\mathrm{Top}}(\ell\ell + E^{\mathrm{miss}}_{\mathrm{T}}) \left(P^{\mathrm{Btag,Exp}}_1\right)^2 \frac{P_2}{\left(P^{\mathrm{Btag}}_1\right)^2} . \end{aligned} \tag{5.13}$$

The term $N^{\mathrm{Exp}}_{\mathrm{Top}}(\ell\ell + E^{\mathrm{miss}}_{\mathrm{T}})$ is defined in Eq. 5.5.

Table 5.4 displays the results of jet veto survival probabilities for $pp \to W^+W^-b\bar{b}$ processes at 7 TeV center-of-mass energy in parton level. Results are shown for different values of the transverse momentum for the jet veto. The second, third columns show the values of P_1 and P_2, respectively, which correspond to the real JVSP. The fourth column displays the ratio P_1^2/P_2 and shows stability as a function of transverse momentum. The fifth and seventh column shows the quantities in the b-tagged region so that the correlation of angular distribution between the nominal region and b-tagged region and the influence from b-tag efficiency can be understood. A deviation from unity is indicative of the correlation. The fifth column quantifies the bias introduced by determining the JVSP from a part of the phase space where a b-parton is required. The bias seems moderate and constitutes 25% for $p_T^v = 30\,\mathrm{GeV}$. In the data, one cannot get the P_1 directly, so this is the reason why to choose a b-tagged control sample to evaluate the ratio of P_1^{btag} between data and MC instead of P_1. The difference between P_1^{btag} and P_1 is mostly canceled out between data and MC.

$p_T^v[GeV]$	P_1	P_2	P_1^2/P_2	P_1^{Btag}	P_1^{Btag}/P_1	$(P_1^{\mathrm{Btag}})^2/P_2$
20.0	0.07	0.005	0.89	0.09	1.29	1.50
22.5	0.08	0.007	0.89	0.10	1.27	1.45
25.0	0.10	0.010	0.94	0.12	1.27	1.51
27.5	0.11	0.013	0.95	0.14	1.27	1.53
30.0	0.13	0.018	0.93	0.16	1.25	1.46
32.5	0.15	0.025	0.90	0.18	1.21	1.32
35.0	0.17	0.032	0.94	0.21	1.18	1.31
37.5	0.20	0.040	0.99	0.23	1.16	1.32
40.0	0.23	0.051	1.01	0.26	1.14	1.30

Table 5.4: Results of jet veto survival probabilities for $pp \to W^+W^-b\bar{b}$ processes at 7 TeV center-of-mass energy. Results are shown for different values of the transverse momentum for the jet veto.

The results are stable within better than 15% over a large p_T^v range between 20 and 40 GeV including effect from correlations and interference. This effectively leads us to conclude that the effect of hadronic energy scale uncertainties will have little impact on $(P_1^{\mathrm{Btag}})^2/P_2$. It should be mentioned that the transverse momentum of parton is different from a reconstructed jet in the detector. The transverse energy of a reconstructed jet does not in general correspond to an exact value of the parton p_T. At LO, requirements on no parton with $p_T > 30\,\mathrm{GeV}$ has a similar efficiency as requiring that no jet with $p_T^v > 20\,\mathrm{GeV}$ is reconstructed after hadronization in the MC.

Table 5.5 shows the same results as in Tab. 5.4 but for $t\bar{t}$ production only. The values of the JVSP in Tab. 5.5 are significantly lower than for the full $pp \to W^+W^-b\bar{b}$ treatment. This is expected. What is particularly relevant for the robustness of the method is the fact that the ratio $(P_1^{\mathrm{Btag}})^2/P_2$ is stable well within typically 10%. The ratio P_1^2/P_2 shows a smaller bias with respect to that observed for the full $pp \to W^+W^-b\bar{b}$ processes.

$p_T^v[GeV]$	P_1	P_2	P_1^2/P_2	P_1^{Btag}	P_1^{Btag}/P_1	$(P_1^{\mathrm{Btag}})^2/P_2$
20.0	0.06	0.003	1.05	0.07	1.19	1.49
22.5	0.07	0.004	1.11	0.08	1.19	1.58
25.0	0.08	0.007	1.03	0.10	1.19	1.47
27.5	0.10	0.010	0.97	0.12	1.19	1.36
30.0	0.12	0.014	1.01	0.14	1.19	1.44
32.5	0.14	0.019	1.01	0.16	1.17	1.38
35.0	0.16	0.026	0.99	0.18	1.14	1.28
37.5	0.18	0.034	0.99	0.21	1.12	1.24
40.0	0.21	0.042	1.03	0.23	1.09	1.23

Table 5.5: Same as Tab. 5.4 for $t\bar{t}$ production only.

Table 5.6 shows the results for the $pp \to W^+W^-b\bar{b}$ at 14 TeV center-of-mass energy. The results for $(P_1^{\mathrm{Btag}})^2/P_2$ display a mild increase with respect to the results obtained for 7 TeV.

$p_T^v[GeV]$	P_1	P_2	P_1^2/P_2	P_1^{Btag}	P_1^{Btag}/P_1	$(P_1^{\mathrm{Btag}})^2/P_2$
20.0	0.07	0.005	1.12	0.09	1.23	1.71
22.5	0.09	0.007	1.04	0.10	1.20	1.51
25.0	0.10	0.010	1.01	0.12	1.22	1.50
27.5	0.12	0.013	1.07	0.14	1.21	1.55
30.0	0.14	0.018	1.05	0.17	1.22	1.55
32.5	0.15	0.022	1.08	0.18	1.19	1.51
35.0	0.18	0.030	1.04	0.20	1.15	1.37
37.5	0.20	0.038	1.04	0.45	1.14	1.34
40.0	0.45	0.047	1.04	0.25	1.11	1.28

Table 5.6: Results of jet veto survival probabilities for $pp \to W^+W^-b\bar{b}$ at 14 TeV center-of-mass energy. Results are shown for different values of the transverse momentum for the jet veto.

5.3.3 Data-Driven Top Estimate for 2011 Data Analysis

The method proposed based on the LO parton level study and published in [111] is applied in the $H \to WW^{(*)} \to l\nu l\nu$ analysis. The result for the 2011 data analysis is shown in Tab. 5.7 for the combined ee, $\mu\mu$ and $e\mu$ channels in B-K and L-M period. The first error is statistical error and the second one systematic. The ratio of the data-driven estimate over the corresponding MC prediction is very close to unity. Figures 5.3-5.9 show comparisons between data and MC for various kinematic variables of the b-tagged samples. Good agreement is observed in all the distributions.

Period	B-K	L-M	All
$P_1^{\text{Btag,Data}}$	0.217 ± 0.007	0.459 ± 0.007	0.453 ± 0.005
$P_1^{\text{Btag,MC}}$	0.459 ± 0.001	0.454 ± 0.001	0.456 ± 0.001
$P_2^{\text{MC}}(10^{-2})$	3.19 ± 0.06	3.39 ± 0.05	3.29 ± 0.04
$P_2^{\text{Estimated}}(10^{-2})$	2.86 ± 0.20	3.54 ± 0.23	3.20 ± 0.15
$N^{\text{Data}}(\ell\ell + E_T^{\text{miss}})$	5648	6583	12231
$N_{\text{non-top}}$	1490.3 ± 149.0	2333.2 ± 233.3	3823.5 ± 382.4
$N_{\text{Top}}^{\text{Estimated}}$	$119.0 \pm 9.6 \pm 19.0$	$150.7 \pm 13.3 \pm 24.1$	$268.7 \pm 18.1 \pm 43.0$
$N_{\text{Top}}^{\text{MC}}$	123.3	135.4	258.8
Ratio	$0.96 \pm 0.08 \pm 0.15$	$1.11 \pm 0.10 \pm 0.18$	$1.04 \pm 0.07 \pm 0.17$

Table 5.7: Results for data-driven top estimate for the 2011 data analysis. The last line shows the ratios of the data-driven estimates over the corresponding MC predictions.

5.3.4 Systematics Study

Detailed studies have been performed to study the systematic uncertainties of the data-driven top estimate. This is summarised in Tab. 5.8. The uncertainties of the first 12 error sources are derived using the common ntuples. In particular, the c and light-jet related b-tagging uncertainties are small because of the highly pure b-tagged control sample with small fake rate. For the non-top background substruction, we quote an overall uncertainty of 4% for the combined channel using 6%, 10%, 5% and 50% relative uncertainties for WW, $WZ/ZZ/W\gamma$, Z+jets and W+jets processes, respectively. In addition, the corresponding statistical uncertainty of the non-top background is included in the statistical uncertainties shown in table 5.7. The "MC generator/parton shower+had." uncertainty is obtained from a closure test.[1] In the closure test, the dependence on p_{T} threshold and η_{jet} cut value in this method has been studied. The variation between pseudo data and MC is shwon in Fig. 5.10.

---

[1] This is provided by S. Li for his SM WW analysis using several different samples.We quote the largest deviation as the systematic uncertainty for different top MC generators and parton shower and hadronization models.

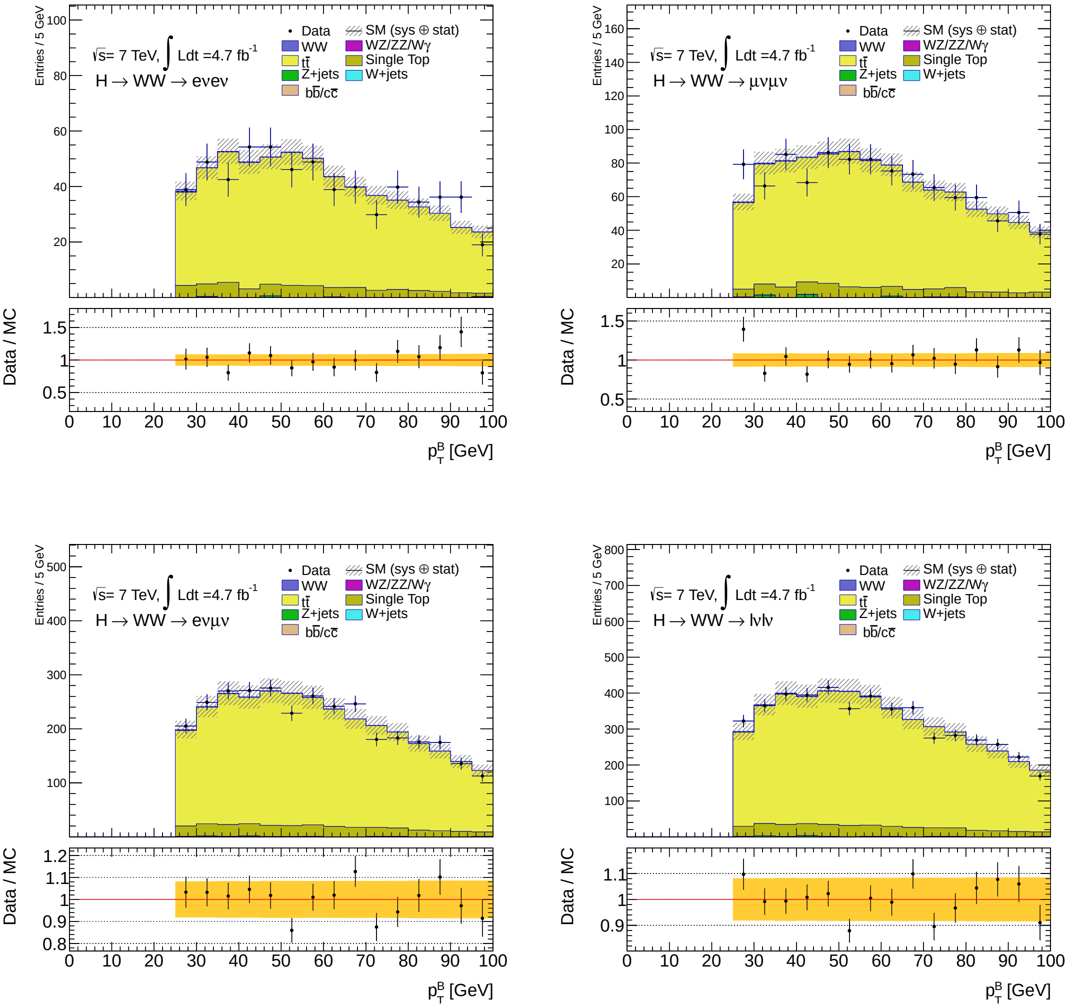


Figure 5.3: Tagged b-jet p_{T} distributions shown for the ee (upper left) $\mu\mu$ (upper right), $e\mu$ (lower left) and combined (lower right) channels.

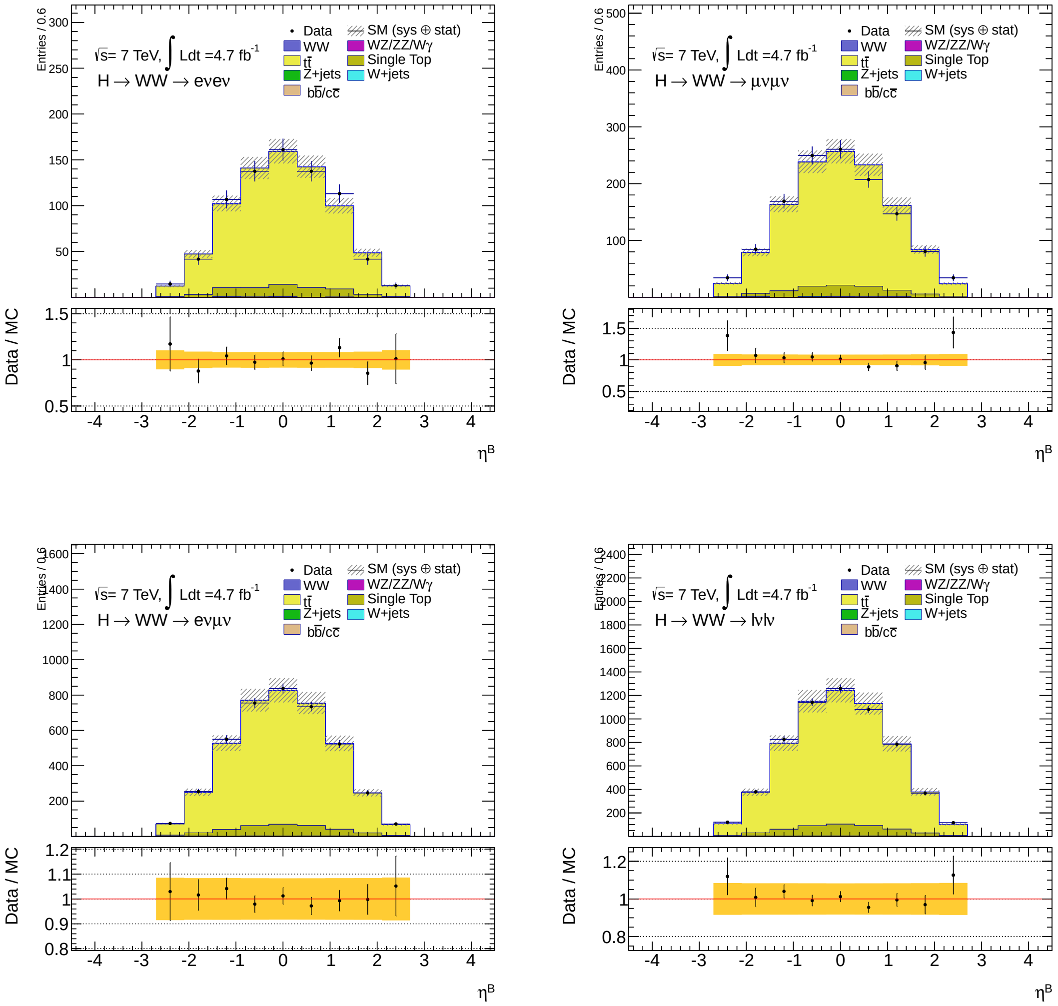


Figure 5.4: Tagged b-jet η distributions shown for the ee (upper left) $\mu\mu$ (upper right), $e\mu$ (lower left) and combined (lower right) channels.

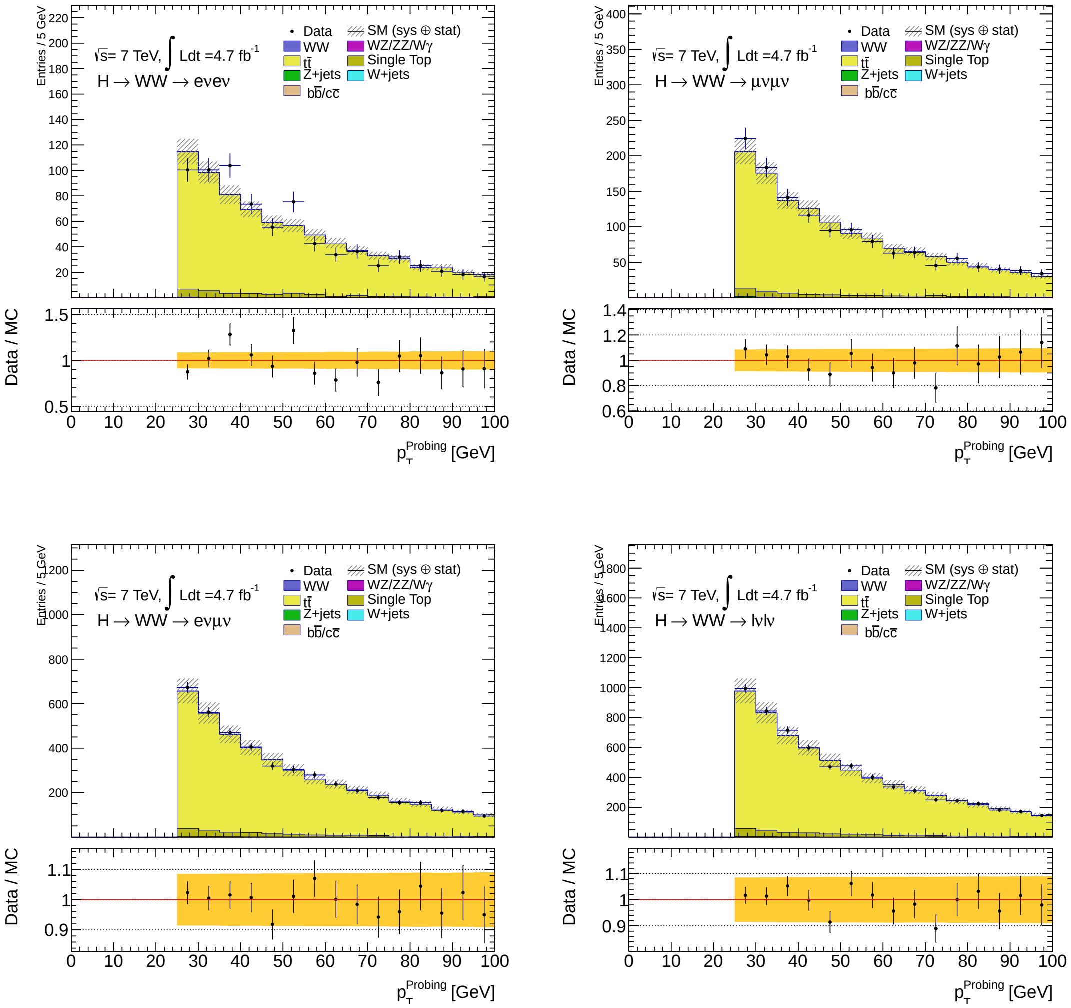


Figure 5.5: Probing jet p_{T} distributions in the b-tagged sample shown for the ee (upper left) $\mu\mu$ (upper right), $e\mu$ (lower left) and combined (lower right) channels.

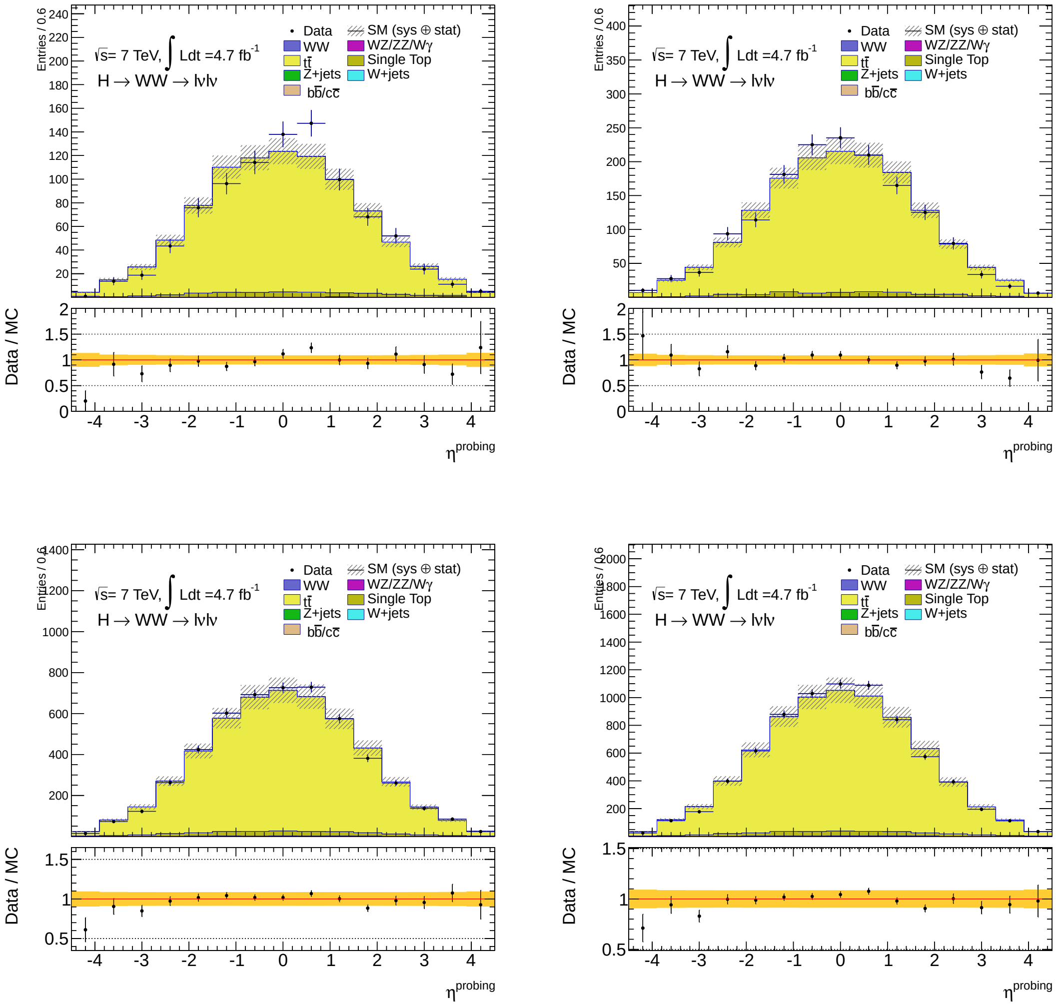


Figure 5.6: Probing jet η distributions in the b-tagged sample shown for the ee (upper left) $\mu\mu$ (upper right), $e\mu$ (lower left) and combined (lower right) channels.

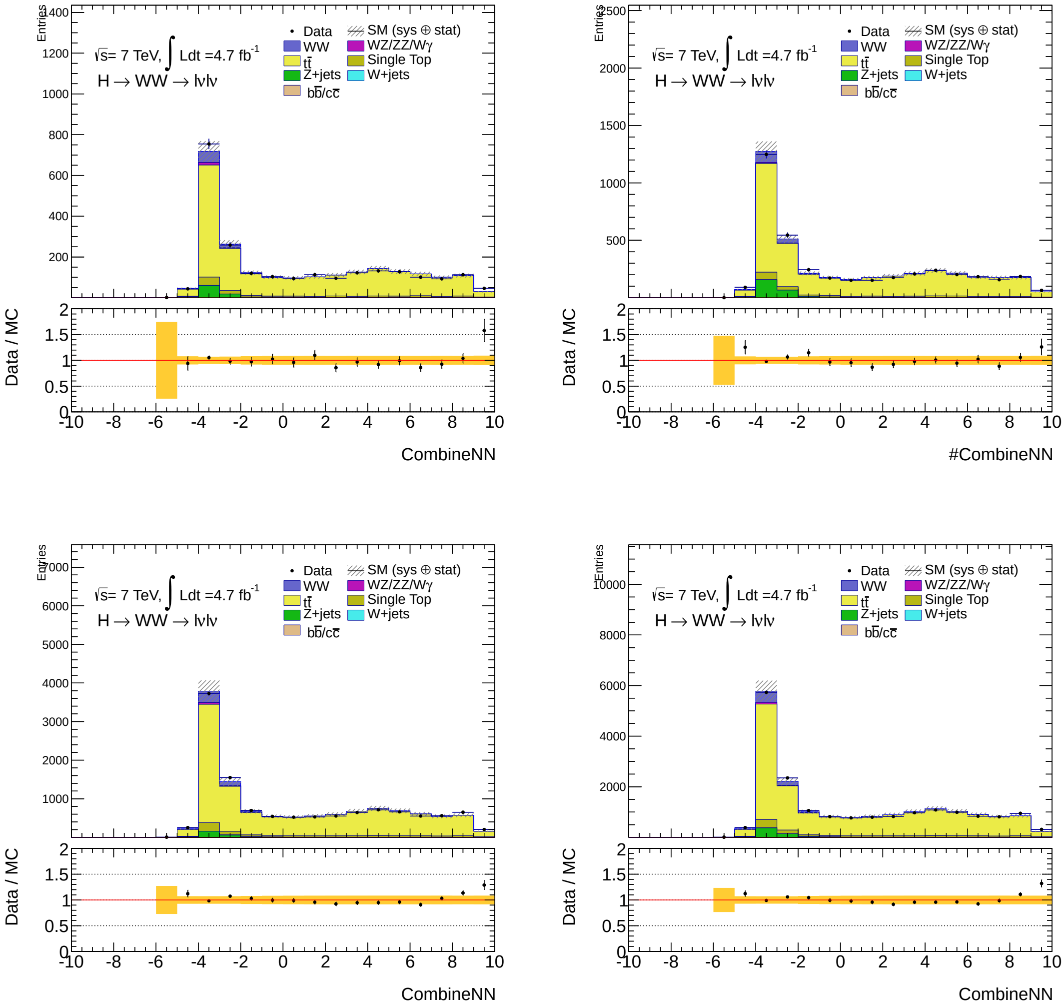


Figure 5.7: Combined NN value distributions for all the jets shown for the ee (upper left) $\mu\mu$ (upper right), $e\mu$ (lower left) and combined (lower right) channels. Jets with CombNN > -1.25 is used to provide a b-jet tagging efficiency of 80%.

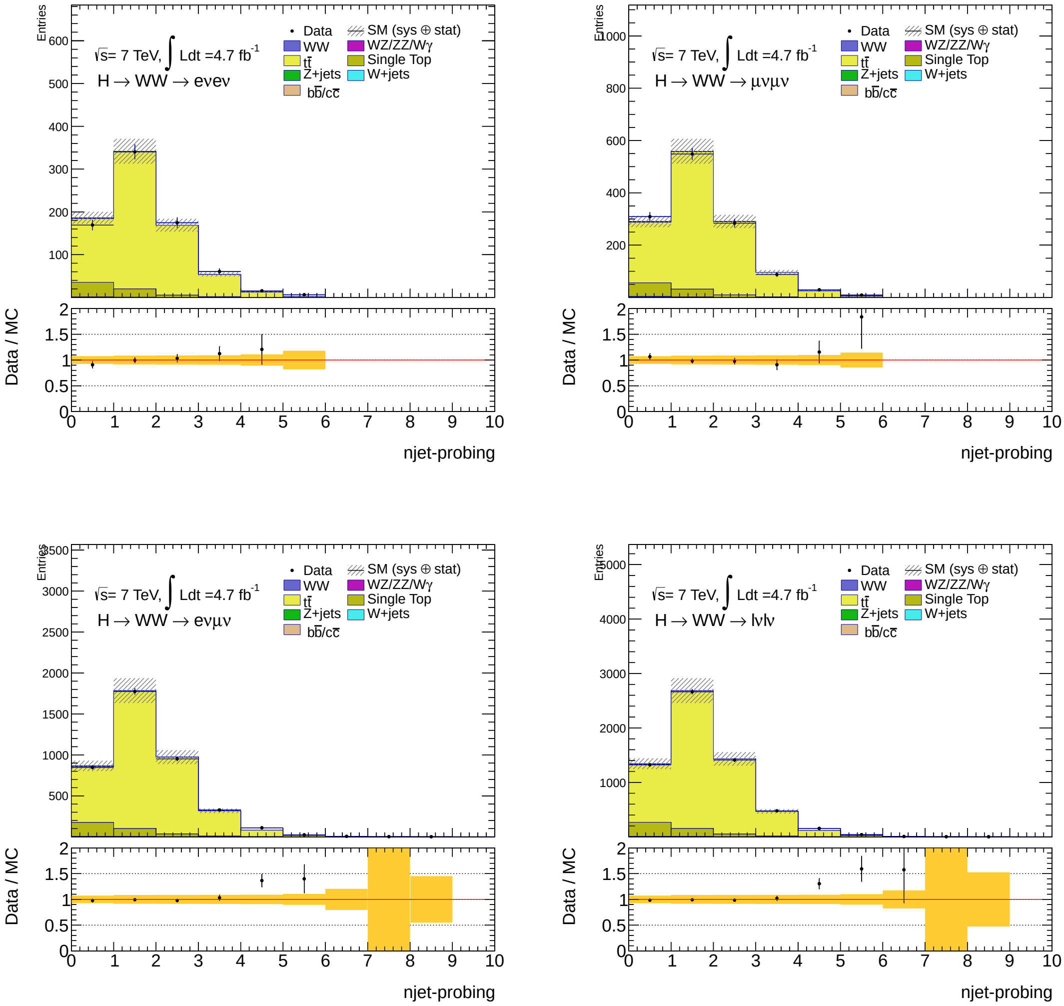


Figure 5.8: Number of probing jets in the b-tagged sample shown for the ee (upper left) $\mu\mu$ (upper right), $e\mu$ (lower left) and combined (lower right) channels.

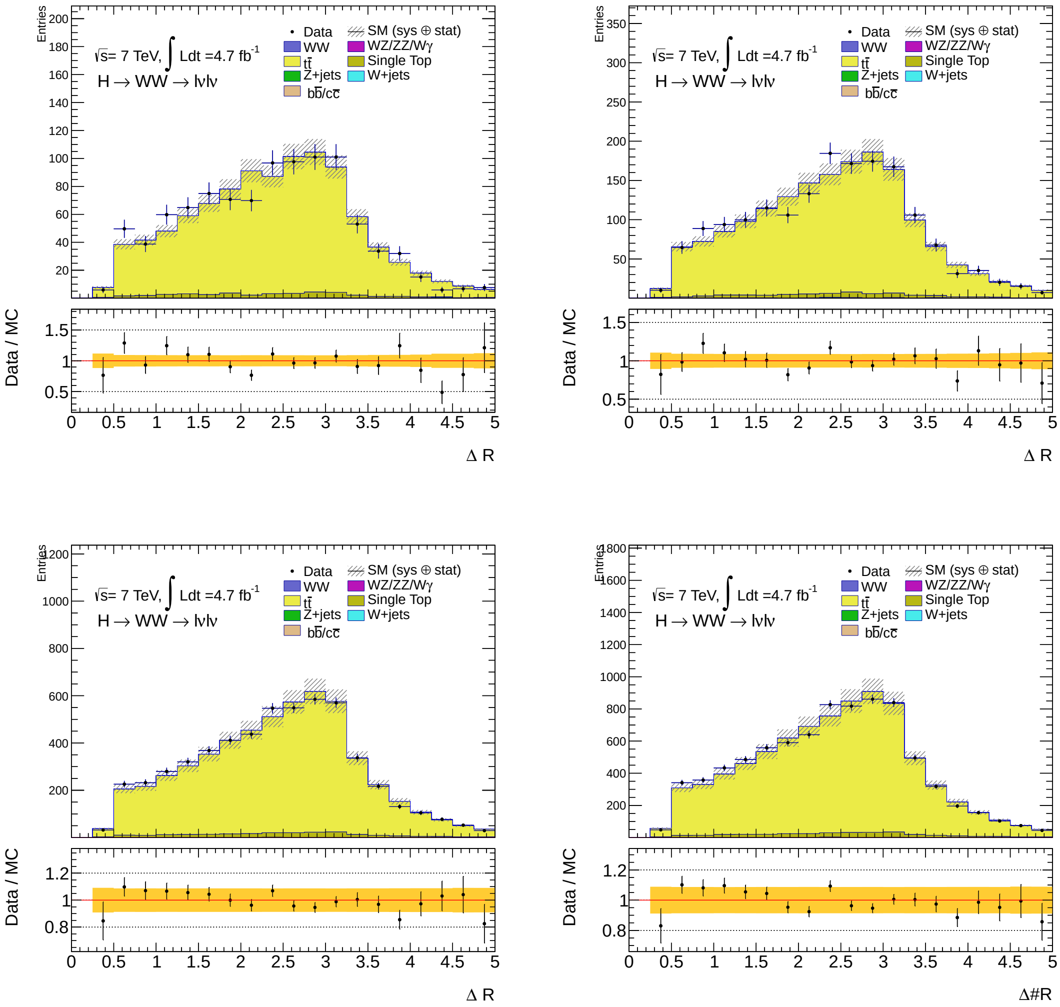


Figure 5.9: Distributions of ΔR between b-tagged jet and probing jet in the b-tagged sample shown for the ee (upper left) $\mu\mu$ (upper right), $e\mu$ (lower left) and combined (lower right) channels.

Another study shows the dependence on the p_{T} and η of the b tagged jet, see Fig 5.11.

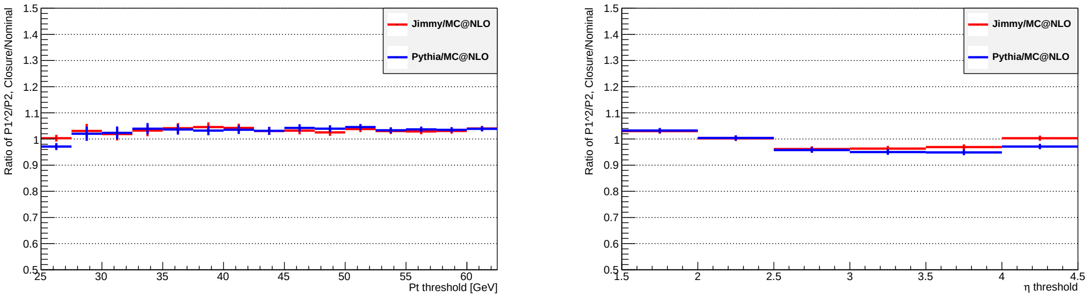


Figure 5.10: The variation on $P1^2/P2$ between pseudo data (two POWHEG MC samples) and MC (MC@NLO sample) as a function of p_{T} threshold (left) and η cut (right) on all of the jets.

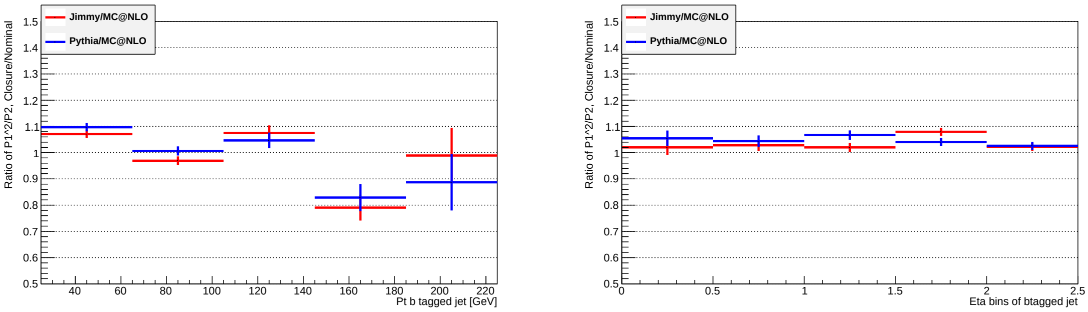


Figure 5.11: The variation on $P1^2/P2$ between pseudo data (two POWHEG MC samples) and MC (MC@NLO sample), when the b tagged jet is in different p_{T} (left) and η (right) bins.

The uncertainty for the single top and $t\bar{t}$ interference effets quoted in Tab.5.8 is derived based on AcerMC samples. Previously in [111], the uncertainty for the same error source was evaluated using the MADGRAPH package by comparing the inclusive $W^+W^-b\bar{b}$ sample with the $t\bar{t}$ sample for a given jet transverse energy threshold p_T varied in a large range from 20 to 40 GeV. For a given jet p_T threshold, the typical resulting variation on the top estimate is at the percent level with a largest variation of 12.5%. The largest variation was quoated in [111] as one of the theory uncertainties. The other component of the theory uncertainties of 7.5% corresponds to the variation of the jet p_T threshold which is taken as an estimate of the renormalization and factorization scale uncertainties. These theory uncertainties will be reevaluated once a NLO calculation will be available. The relative contribution of the single top background in the top sample is varied by ±30% based on the current cross section measurement, resulting in a variation on the top estimate of less than 1%. The initial and final state radiation uncertainty is estimated by using two special MC samples produced by the top working group. As these samples overestimate the actually observed difference between data and MC (see Fig. 5.12 [112]), we take half of the variation as an estimate of the uncertainty. The total experimental and theoretical uncertainty amounts to 11.1%. We decided to stay conservatively to a total uncertainty of 16% until the dominant theory uncertainty can be assessed by a NLO calculation.

Source [%]	Relative uncertainty (%)
Electron Energy Resolution	0.1/0.2
Electron Energy Scale	0.1/0.1
ID Muon Momentum Resolution	$-0.1/-0.1$
MS Muon Momentum Resolution	$-0.1/-0.1$
Muon Momentum Scale	< 0.1
Jet Energy Scale	$-0.3/1.0$
Jet Energy Resolution	3.0
Pileup	$0.4/ < 0.1$
Non-lepton terms to MET (cluster unc.)	$0.2/-0.1$
b-tagging (b-jet)	$3.5/-3.0$
b-tagging (c-jet)	< 0.1
b-tagging (light-jet)	0.4
Non-top bkg substruction	2.4
MC generator/paton shower+had.	< 2.2
Single top-$t\bar{t}$ interference	4.5
Theory (scale unc.)	7.5
Single top cross section variation (30%)	$0.9/-0.5$
Initial/final state radiation	4.0
Total systematics	11.1
Statistics	5.0
Total systematics $\oplus$ statistics	12.2

Table 5.8: A list of systematic uncertainties studied for the data-driven top estimate. When there are two given relative uncertainty values, they correspond to an upward and a downward variation of the error source.

The above mentioned top estimate is performed after the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut at the 0-jet veto cut level with the corresponding statistical and systematic uncertainties. This is equivalent to provide a top normalisation factor once the data-driven estimate is scaled by the top MC prediction. After the 0-jet veto cut, a few other topological cuts (m_{ll}, $P_{T,ll}$, $\Delta\phi_{ll}$ and M_T) are applied to further suppress the remaining background contributions. Given the low statistics of the remaining top background events after the 0-jet veto cut, it is not possible to define a meaningful top control sample to

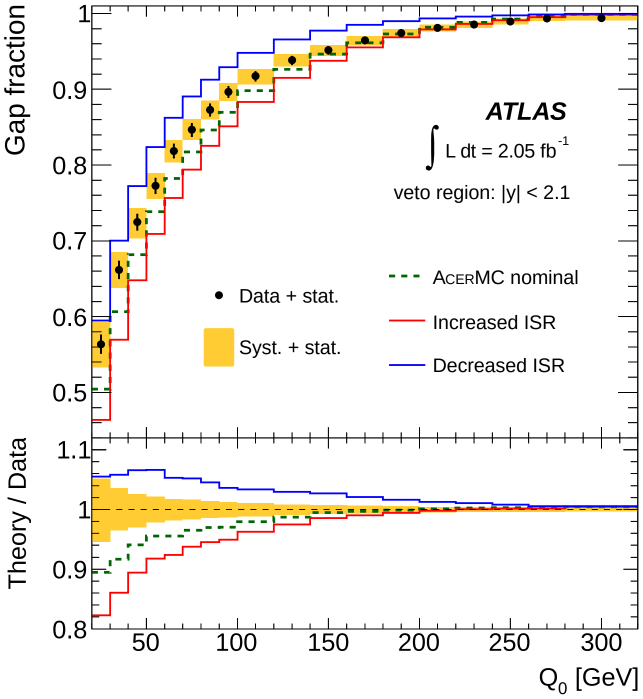


Figure 5.12: Gap fraction, defined as the fraction of events which have non additional jets other than $t\bar{t}$ as a function of Q_0. The Q_0 is the p_t threshold of the additional jet [112]

.

estimate the systematic uncertainties of these topological cuts. One possibility to avoid this is to move the cut under consideration before the 0-jet veto cut to get a new normalisation factor. The difference with the default normalisation factor may be used as an estimate of the uncertainty of the cut. The results are shown in Tab. 5.9. This uncertainty estimation method may be extended to the other channels with 1- and 2-jet analyses.

	Default	m_{ll}	$P_{T,ll}$	$\Delta\phi_{ll}$
Normalisation Factor (NF)	1.04	1.12	1.06	1.17
δ(SF)/SF (%)	–	8.1	2.3	12.8

Table 5.9: Relative systematic uncertainties of the topological cuts m_{ll}, $P_{T,ll}$ and $\Delta\phi_{ll}$.

For the statistical analysis (see next Chapter), we do not apply explicit cuts on m_T but use the shape of the distribution, a different treatment is thus needed here to evaluate the shape uncertainty of the m_T distribution. The whole question is whether one can use the b-tagged samples to get the shape uncertainty. To answer this, the m_T distribution of the b-tagged sample with that of untagged sample is compared. This is shown in Fig. 5.13. The shape of the m_T distributions is found to be fairly similar between the two samples. We can thus use the b-tagged samples to check the shape uncertainty between data and top MC. This is shown in Fig. 5.14. The four plots from left to right and top to bottom correspond to full b-tagged sample, 0-jet, 1-jet and 2 or more jets samples, respectively. For the 0-jet sample, the b-jet P_T threshold has been lowered from the nominal value of 25 GeV to 20 GeV. Since we are interested only in the shape uncertainty, the normalisation of the MC prediction has been fixed to that of data. From the ratio, one can conclude that there is no significant systematic tread in the relevant mass range of $60 - 240$ GeV used in the fit.

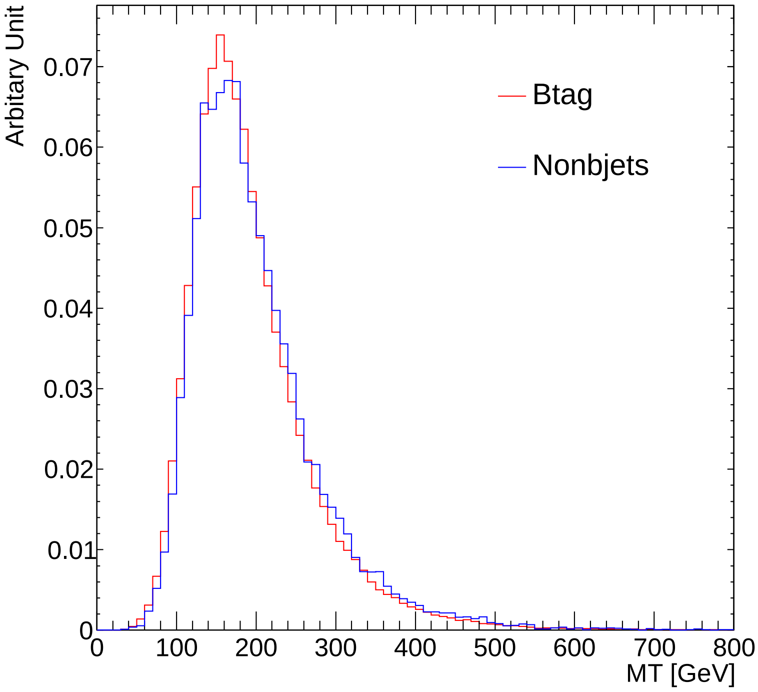


Figure 5.13: Distribution of M_T of the b-tagged sample (red histogram) in comparison with that of untagged sample (blue histogram).

5.3.5 Top Background Estimation in 1- and 2-Jet Bins

The top background estimation in 1- and 2-jet bins is much simpler. Using the top control region in 1- and 2-jet bins, defined by reversing the b-tagging requirement on the jets, the number of top events in these bins can be extracted using a normalisation factor from the control region to the signal region. However, this normalisation factor is not used to estimate the final significance for Higgs discovery or exclusion limits. Instead, these control regions are used directly in the likelihood fit in the statstical analysis.

From Tab. 4.8, it is observed that the data and MC ratio varies significantly in different steps of the cut flow, in particular after the m_{jj} cut. This hints a mismodelling in the m_{jj} distributions. To check this, a top control sample is defined with one tagged b-jet after the selection of two or more jets in the events and no additional jets

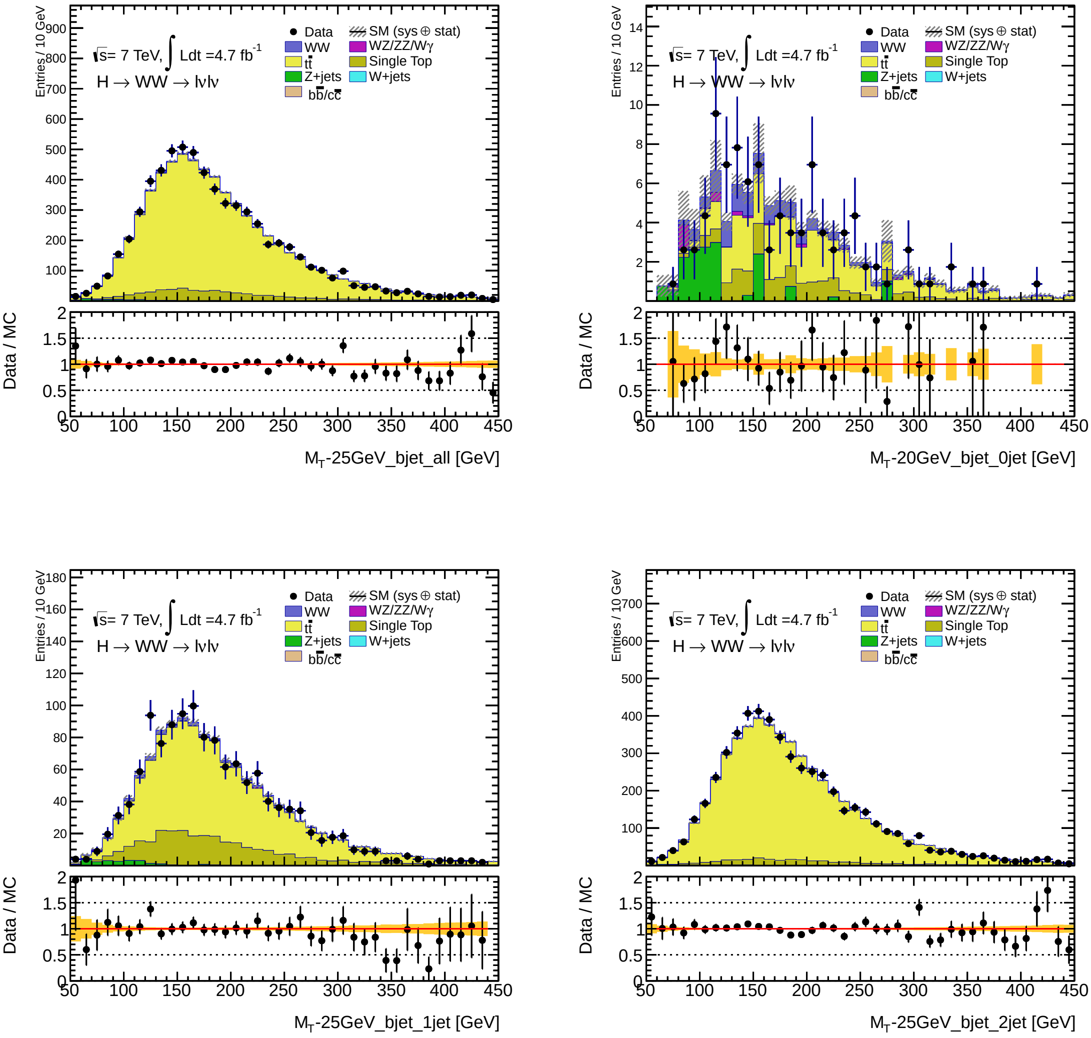


Figure 5.14: Distributions of M_T of the full b-tagged sample (top-left), 0-jet (top-right), 1-jet (bottom-left) and 2 or more jets (bottom-right) samples.

in the central pseudorapidity range $|\eta| < 3.2$ above a jet transverse energy threshold of 25 GeV. The invariant mass distribution of two leading jets is shown in Fig. 5.15. In the dominant low mass region, the data is higher than the MC prediction by about 5%, in agreement with the top scale factors determined in the 0-jet and 1-jet analyses. At the high mass tail, in particular in the signal region $m_{jj} > 500$ GeV, the data tends to be lower than the MC prediction by about 20% although the statistical significance is limited. In order to check the impact of the potential mismodelling of the top MC prediction, a polynomial function fit has been performed to the data and MC ratio as a function of m_{jj} in the mass range between 0 and 500 GeV. Among the three fitted polynomial functions P_1, P_2 and P_3 with 2, 3, and 4 free parameters respectively, the P_2 fit is found to give a slightly better χ^2 value per degree of freedom and is thus chosen to parameterise the top scale dependence as a function of m_{jj} (Fig. 5.16). The fit is performed up to 500 GeV only due to the limited statistics at higher masses. Above 500 GeV, a constant correction is applied, same as that of 500 GeV.

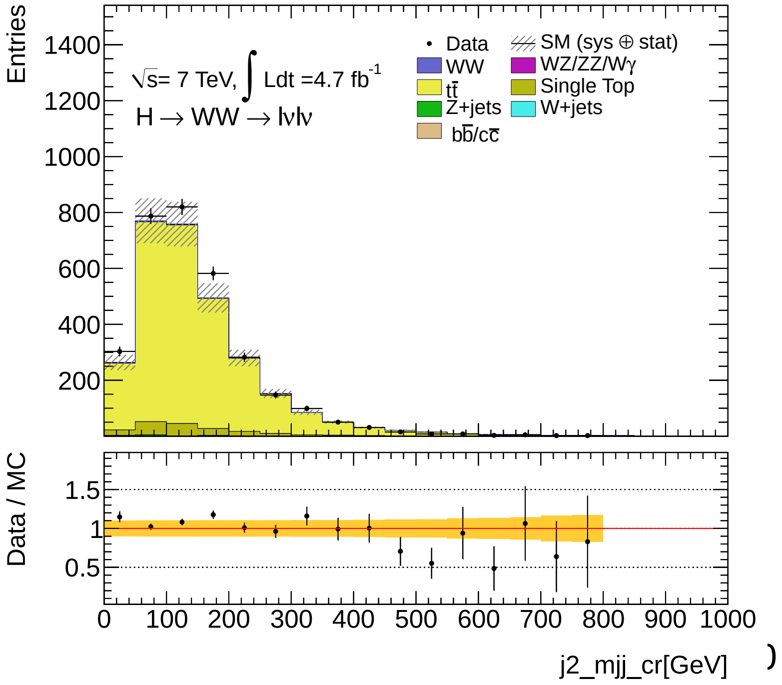


Figure 5.15: Dijet invariant mass distribution m_{jj} in data compared to MC in the b-tagged control sample in 2-jet analysis.

5.4 Z/Drell-Yan Background Estimation and Constraint on $E_{\mathrm{T}}^{\mathrm{miss}}$ Mismodelling Using W+jets Events

The Z/Drell-Yan background dominates the same flavour channels in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ region, which can be strongly suppressed by the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ and m_{ll} selection cuts. Due to its large cross section, it is still very important to estimate the remaining contribution. Two traditional methods are used in the estimation, the ABCD method and scale factor method.

5.4.1 ABCD Method

The ABCD method is based on the assumption that the correlation between m_{ll} and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ is small. The number of events in the signal region (high $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ and low m_{ll} region is extracted by using $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ and m_{ll} cut efficiencies from data directly to avoid

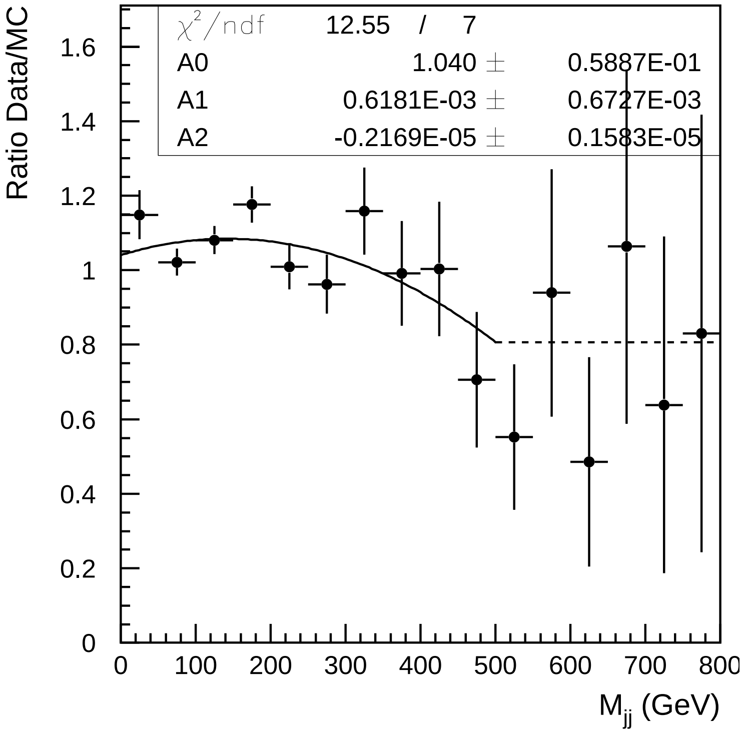


Figure 5.16: Fit with polynomial function P_2 to the data and MC ratio as a function of m_{jj} of the b-tagged sample in 2-jet analysis.

low statistics in MC in the signal region. The ABCD regions are defined as follows (Fig. 5.17):

- A: $12 < m_{ll} < 50\,\mathrm{GeV}$, $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$ (signal region),
- B: $12 < m_{ll} < 50\,\mathrm{GeV}$, $20 < E_{\mathrm{T,rel}}^{\mathrm{miss}} < 45\,\mathrm{GeV}$,
- C: $|m_Z - m_{ll}| < 15\,\mathrm{GeV}$, $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$,
- D: $|m_Z - m_{ll}| < 15\,\mathrm{GeV}$, $20 < E_{\mathrm{T,rel}}^{\mathrm{miss}} < 45\,\mathrm{GeV}$.

It is possible to extrapolate ratios from the control region to the signal region using the numbers of events in data through:

$$A_{\mathrm{est}} = B_{\mathrm{obs}} \times \frac{C_{\mathrm{obs}}}{D_{\mathrm{obs}}} \times \alpha\,, \tag{5.14}$$

where α is defined as:

$$\alpha = \frac{(A/B)}{(C/D)}\,, \tag{5.15}$$

where A, B, C and D are the Z+jets MC event yields in regions A, B, C and D, respectively. This factor is designed to correct for any differences in the ratios of Z+jets events between these regions in case they are not constant as assumed.

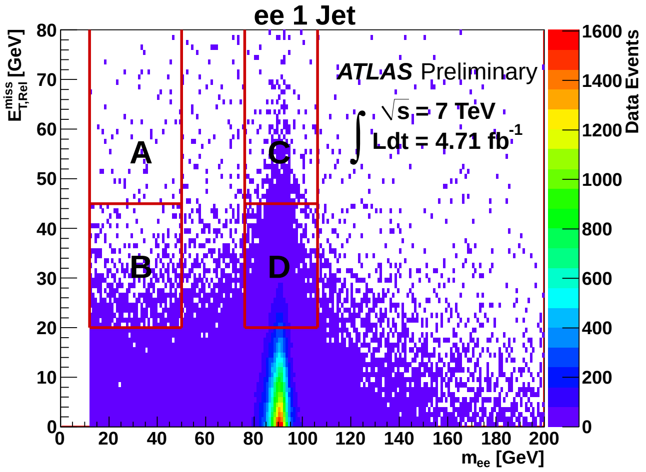


Figure 5.17: Regions of A, B, C and D in $m_{ll} - E_{\mathrm{T,rel}}^{\mathrm{miss}}$ plane used for the ABCD method and shown with data events in the *ee* 1-jet channel after the jet multiplicity cut.

The expected contamination from backgrounds such as diboson or top events is subtracted from the number of observed events in regions B, C and D using MC prediction of these backgrounds. The contamination is genrally small compared to the number of Z/DY events. The number of events estimated using the ABCD method is thus independent of the Z/DY MC.

The ABCD estimate can be calculated after each cut in the cutflow, but topological cuts with a dependence on either m_{ll} or $E_{\mathrm{T}}^{\mathrm{miss}}$ will have different efficiencies in the B/C/D regions and skew the resulting estimate. The systematic of this method varies from 15% to 34%, in different jet bins and lepton flavor combination.

5.4.2 Scale Factor Method

The scale factor method is another approach to estimate the Z/DY background in signal region A:

$$A_{\mathrm{est}} = \frac{C_{\mathrm{data}}}{C_{\mathrm{data}} + D_{\mathrm{data}}} \frac{C_{\mathrm{MC}} + D_{\mathrm{MC}}}{C_{\mathrm{MC}}} . \tag{5.16}$$

in which the region D is slightly different from ABCD method, which includes all low $E_{\mathrm{T}}^{\mathrm{miss}}$ area in Z mass window.

The result is shown in Tabs. 5.10 and 5.11 for the *ee* and $\mu\mu$ channels, respectively, in comparison with the MC prediction and the estimation using the ABCD method. This method also assumes that there is no dependence on $E_{\mathrm{T}}^{\mathrm{miss}}$ for Z/DY events inside and outside of the Z mass window. Therefore it has the similar problem as the ABCD method in that the normalisation factor (i.e. the ratio of the data-driven estimate over the MC prediction) may have a strong dependence in the cutflow. For this reason, an alternative method is studied and described in the next section.

Cut	Raw MC Est.	SF Est.	ABCD Est.
ee inclusive	48.04 ± 4.72	47.42 ± 4.80	46.95 ± 4.96
ee 0-jet	25.67 ± 3.80	24.93 ± 3.75	24.36 ± 3.82
ee 0-jet p_T^{ll}	3.57 ± 1.22	3.06 ± 1.52	3.71 ± 1.90
ee 0-jet $\Delta\phi_{ll}$	3.57 ± 1.22	3.06 ± 1.52	3.73 ± 1.91
ee 1-jet	15.85 ± 2.45	16.70 ± 2.80	16.18 ± 2.82
ee 1-jet b-tag veto	14.17 ± 2.28	15.05 ± 2.61	14.25 ± 2.57
ee 1-jet p_T^{Tot}	3.03 ± 0.99	3.18 ± 1.14	2.63 ± 0.96
ee 1-jet $\Delta\phi_{ll}$	2.20 ± 0.86	2.32 ± 0.97	1.92 ± 0.82

Table 5.10: Z+jets estimates in the ee channel using the scale factor method. For comparison, the estimate using the ABCD method is also shown.

Cut	Raw MC Est.	SF Est.	ABCD Est.
$\mu\mu$ inclusive	134.84 ± 7.63	140.21 ± 8.29	143.58 ± 8.90
$\mu\mu$ 0-jet	77.31 ± 6.15	78.87 ± 6.44	80.45 ± 6.89
$\mu\mu$ 0-jet p_T^{ll}	16.21 ± 2.51	20.02 ± 4.35	22.32 ± 5.13
$\mu\mu$ 0-jet $\Delta\phi_{ll}$	16.21 ± 2.51	20.02 ± 4.35	22.31 ± 5.13
$\mu\mu$ 1-jet	45.53 ± 4.11	48.02 ± 4.80	48.59 ± 5.08
$\mu\mu$ 1-jet b-tag veto	41.48 ± 3.89	43.71 ± 4.52	43.60 ± 4.72
$\mu\mu$ 1-jet p_T^{Tot}	11.10 ± 2.15	9.63 ± 2.13	9.28 ± 2.09
$\mu\mu$ 1-jet $\Delta\phi_{ll}$	6.22 ± 1.41	5.40 ± 1.35	5.20 ± 1.33

Table 5.11: Z+jets estimates in the $\mu\mu$ channel using the scale factor method. For comparison, the estimate using the ABCD method is also shown.

5.4.3 Constraint on $E_{\mathrm{T}}^{\mathrm{miss}}$ Mismodelling Using W+jets Events

The strong dependence of the normalisation factors of the ABCD and scale factor methods at different steps in the cutflow is believed to be due to at least partially a $E_{\mathrm{T}}^{\mathrm{miss}}$ mismodelling in the Z/DY samples. The measured $E_{\mathrm{T}}^{\mathrm{miss}}$ in the $Z \to ll(l = e, \mu)$ events is mostly a fake $E_{\mathrm{T}}^{\mathrm{miss}}$ due to detector resolution effects, mismeasurements in the detector or pile up contributions. The basic idea is to extract the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ from W+jets events in data, which have a much larger data sample than Z/DY, to correct the similar fake $E_{\mathrm{T}}^{\mathrm{miss}}$ in the latter sample. Therefore the basic assumption is that the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the W+jets and Z/DY samples are the same. However, the measured $E_{\mathrm{T}}^{\mathrm{miss}}$ in W+jets also contains a true $E_{\mathrm{T}}^{\mathrm{miss}}$ component from the neutrino in the W decay. To extract the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ in the W+jets events, the neutrino contribution must be subtracted. This is possible for W+jets MC events but needs approximation in data as the direction of the neutrino is not known.

The Method

As mentioned above, the true neutrino p_{T} and direction are not known in data, the only way to obtain the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ is to subtract the mean value of neutrino transverse momentum based on the MC prediction. The procedure is introduced as follows:

- Firstly, clean W+jets samples in data and MC are selected asking for one tight electron or combined muon only per event, which is isolated. Then a loose $E_{\mathrm{T}}^{\mathrm{miss}}$ cut $E_{\mathrm{T}}^{\mathrm{miss}} > 25\,\mathrm{GeV}$ is required to reject largely the QCD contamination. The transverse mass of the W boson is also required to be $m_T > 55\,\mathrm{GeV}$ and $m_T < 110\,\mathrm{GeV}$ to keep the W mass on shell. This is needed because when the neutrino transverse momentum is subtracted, the mean value of neutrino transverse momentum should not be diverged to two different peaks. In addition, all studies are made in 0-jet bin. The jet selection is the same as the nominal analysis.

- Based on the truth information in the W+jet MC, the p_{T} of neutrinos is obtained as a function of lepton p_{T} in fine bins of every 5 GeV. In each p_{T} bin, the neutrinos transverse momentum is decomposed along and perpendicular to the lepton transverse momentum, as shown in Fig. 5.18, called longitudinal neutrino p_{T} and transverse neutrino p_{T}, respectively.

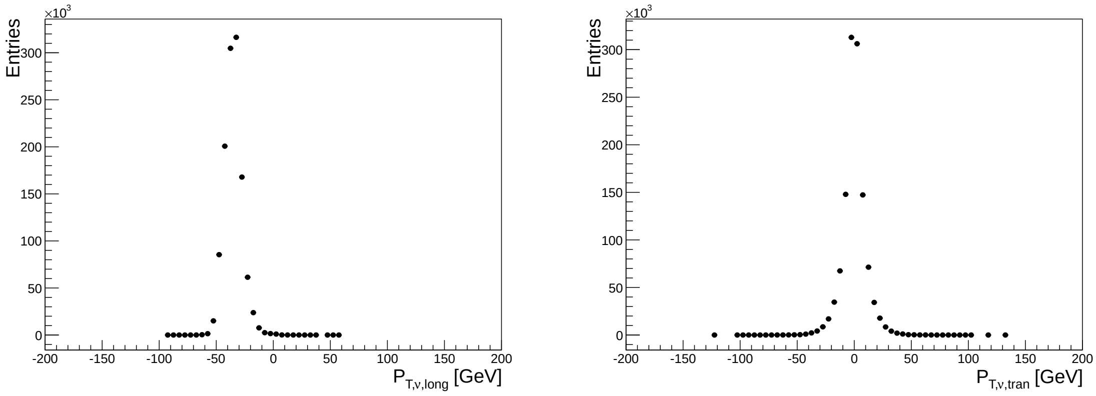


Figure 5.18: Example of neutrino p_{T} distribution in the W+jets MC events. The p_{T} of the only lepton is selected in the $30-35\,\mathrm{GeV}$ bin in the electron channel and BK period of the 2011 data analysis. The left plot shows the longitudinal p_{T} and the right plot shows the transverse p_{T}.

- In both data and MC, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is decomposed along and perpendicular to the lepton transverse momentum as well, as shown in Figs. 5.19 and 5.20.

- In the longitudinal direction, the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ in the W+jets event sample is obtained by subtracting the mean longitudinal p_{T} of the neutrinos in a given lepton

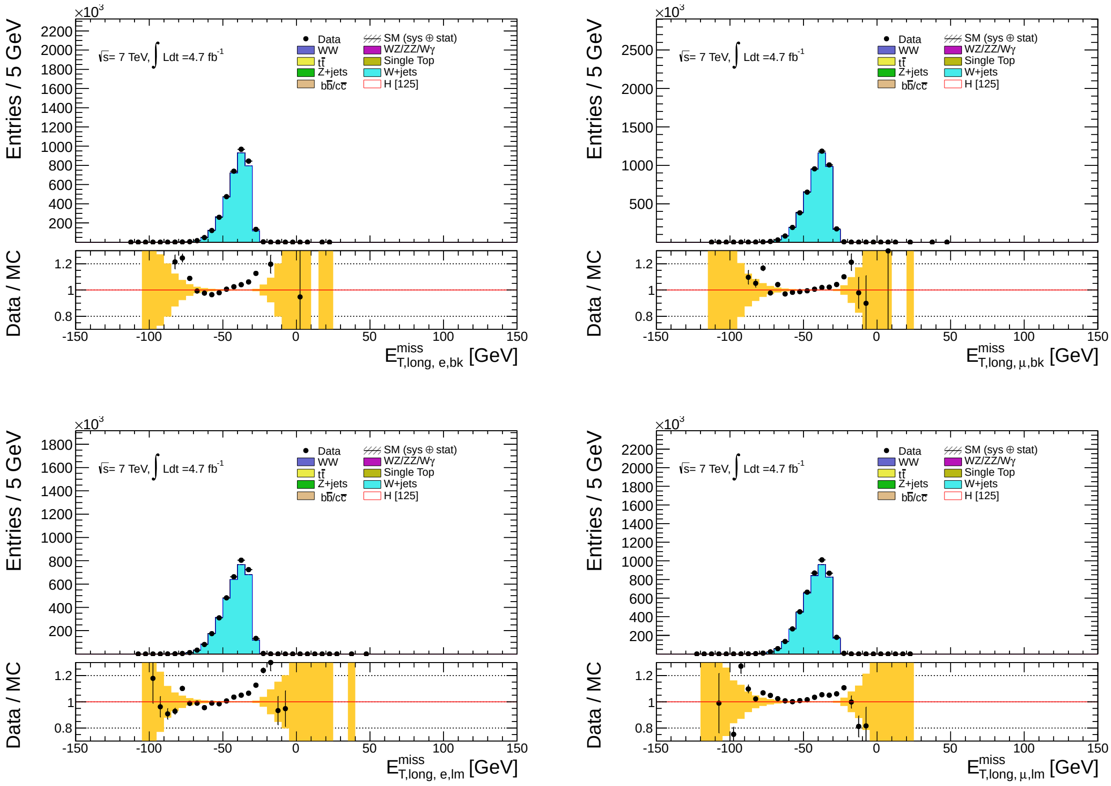


Figure 5.19: Longitudinal $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the W+jets events in data (dot) and MC (histogram). The upper left (right) plot is for the electron (muon) channel and BK period of the 2011 data analysis. The lower left (right) plot is for the electron (muon) channel and LM period.

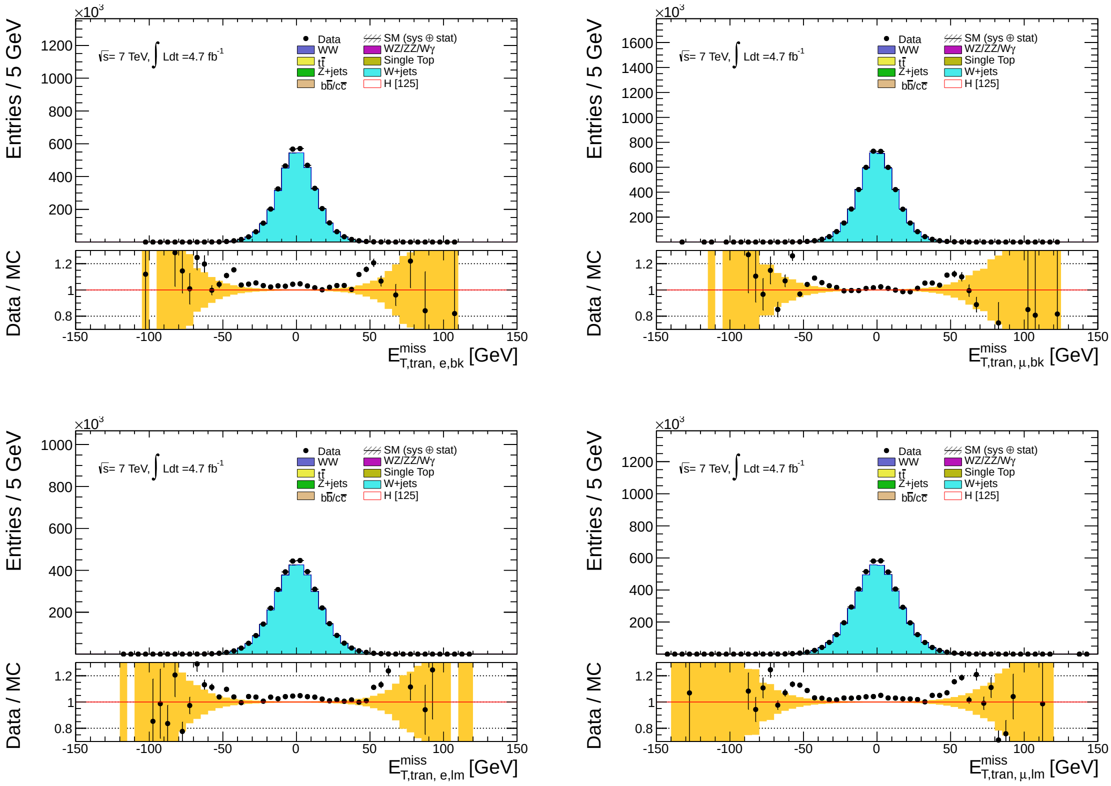


Figure 5.20: Same for Fig. 5.19 but for the transverse $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions.

p_{T} bin from the measured $E_{\mathrm{T}}^{\mathrm{miss}}$. In the transverse direction, the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ extraction is treated differently since the mean transverse p_{T} of the neutrinos is expected to be centered around zero (see the right panel of Fig. 5.18). Therefore no shift is needed here but only different widths need to be dealt with. The width of the measured transverse $E_{\mathrm{T}}^{\mathrm{miss}}$ distribtion in Fig. 5.20 is a convolution of the width of the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution with that of the neutrinos. Therefore a simple way to extract the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ is to generate the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution described by a Gaussian function with mean value 0 and standard deviation:

$$\sigma_{\text{fake met trans}} = \sqrt{\sigma^2_{\text{met, measured, trans}} - \sigma(\nu, \text{trans})^2}\,, \tag{5.17}$$

in which $\sigma(\nu, \text{trans})$ is obtained from MC.

The resulting fake $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in both longitudinal and transverse directions for data and MC after the subtraction of the neutrinos are shown in Figs. 5.21 and Fig. 5.22. The fake $E_{\mathrm{T}}^{\mathrm{miss}}$ amplitude can thus be easily obtained. This is shown in Fig. 5.23, in which only the negative part of the fake longitudinal $E_{\mathrm{T}}^{\mathrm{miss}}$ is used to avoid potential residual QCD contamination on the positive side. The ratio between data and MC in each bin of the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ amplitude represents the correction which is applied to the Z+jets events.

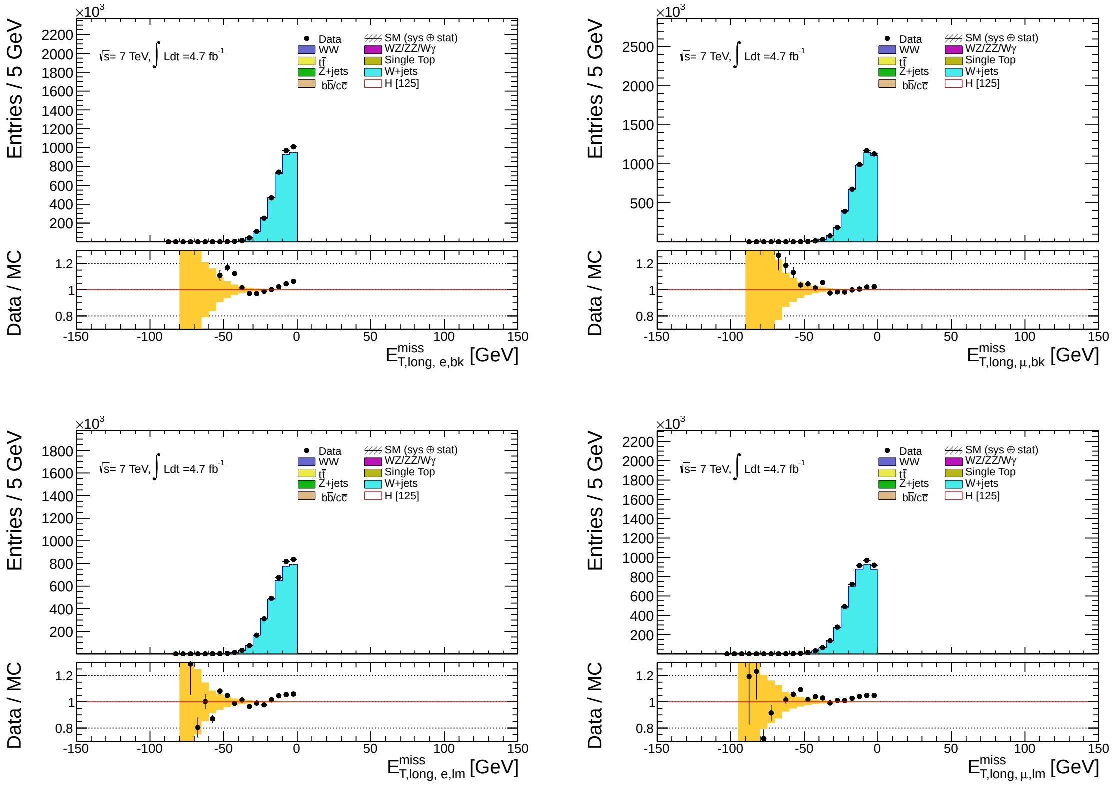


Figure 5.21: Same as Fig. 5.19 but after the neutrino subtraction.

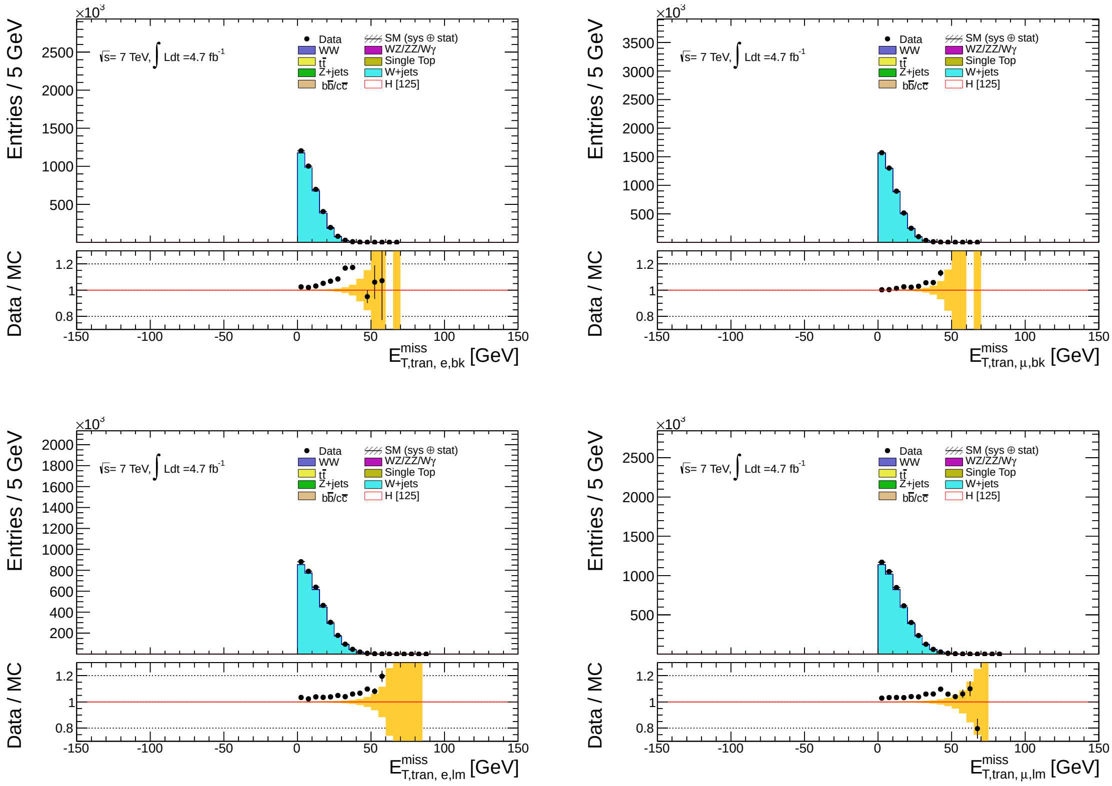


Figure 5.22: Same as Fig. 5.20 but after the neutrino subtraction.

A closure test is performed to validate whether such subtraction can reproduce the real difference between data and MC. For this test, a Pythia W+jet MC sample is used as the pseudo data. The real difference between data and MC is obtained by subtracting the true neutrino $E_{\rm T}^{\rm miss}$ instead of the mean $E_{\rm T}^{\rm miss}$ in different lepton $p_{\rm T}$ bins. The comparison of the two different subtrations is shown in Fig. 5.24. It shows when using the approximation described above, the result gives at most 7% difference from the true correction at a $E_{\rm T}^{\rm miss}$ cut value of 45 GeV. It is worth noticing that such a difference between data and MC only represents the mismodelling in the fake $E_{\rm T}^{\rm miss}$, the mismodelling due to the dilepton system in the Z/DY events can not be corrected for with the W+jets events.

To make sure that the correction of each bin is correctly applied in the Z+jets region, a rescaling on the correction factors of $E_{\rm T}^{\rm miss}$ is performed to correct kinematic differences between W and Z. The fake $E_{\rm T}^{\rm miss}$ distribution between W+jets and Z+jets are compared in Fig. 5.25 (before rescaling) and Fig. 5.26 (after rescaling). The bin position of corrections derived from the W+jets is rescaled by a factor 1.1 to match with the $E_{\rm T}^{\rm miss}$ in the Z+jets. To take into account the uncertainty of this scheme, the scale factor is varied by 10% and the result only changes by 2%. The scaled corrections in terms of a smooth function are then applied to the Z/DY events both in the Z mass window for validation and in the Higgs signal region to cross check the Z/DY estimate from the other methods.

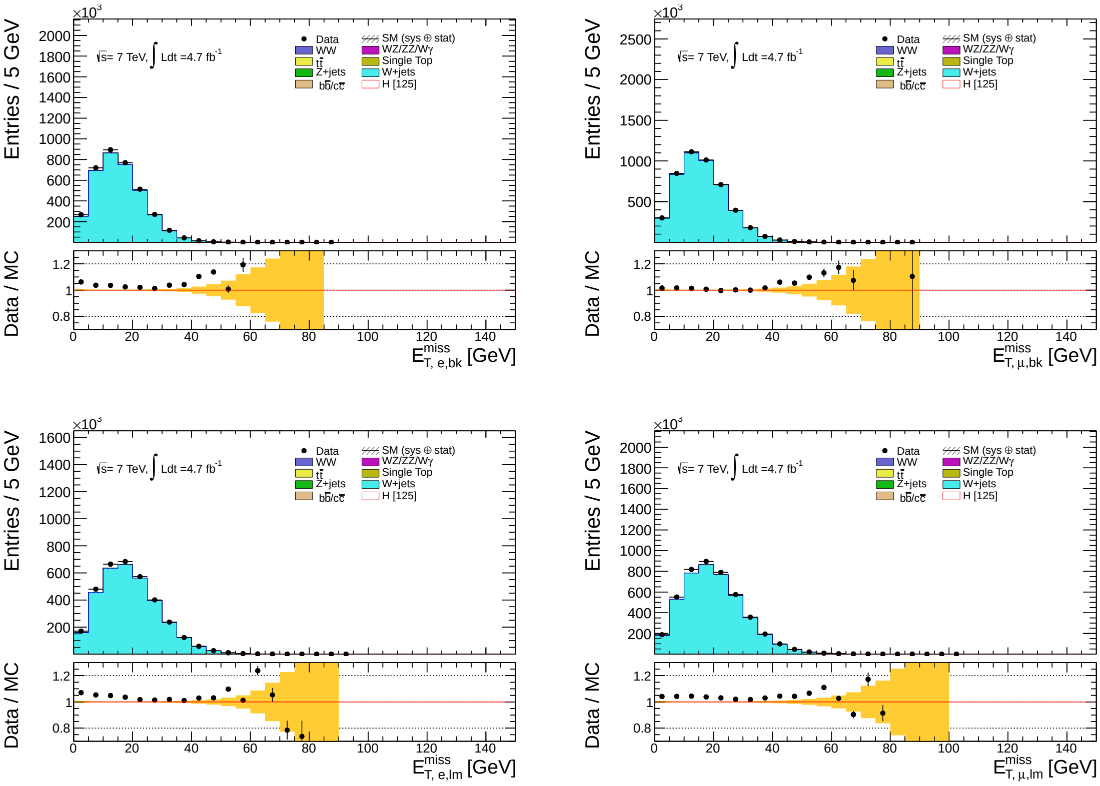


Figure 5.23: Distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ amplitude in the W+jets events in data (dot) and MC(histogram) after the neutrino subtraction. The plot layout is similar to Figs. 5.19-5.22.

Results and Systematics

The impact of the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ corrections for Z/DY events in the Z mass window is shown with the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions of Z+jets before and after the corrections in Fig. 5.27. The MC chosen is MC11b which has a much larger $E_{\mathrm{T}}^{\mathrm{miss}}$ mismodelling than MC11c due to a different pile up simulation. Therefore the correction effect can be better appreciated with MC11b. The result of MC11c is also shown later in Fig. 5.32. To check the influence of different kinematic cuts, similar comparison is shown in Fig. 5.28 after applying cuts $E_{\mathrm{T}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$ and $p_T^{lll} > 45\,\mathrm{GeV}$. The corresponding plots in the Higgs signal region are shown Figs. 5.29 and 5.30. These plots demonstrate that the mismodelling on the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ in the Z/DY events can be corrected with the mismodelling corrections derived from the W+jets events. The effect of the correction for the Z/DY background in the m_T distribution is shown in Fig. 5.31.

Similar results with MC11c are shown in Figs. 5.32 and 5.33. It is very hard to see significant corrections with MC11c because the $E_{\mathrm{T}}^{\mathrm{miss}}$ mismodelling effect in MC11c is much smaller. The remaining mismodelling is mainly due to a difference in the dilepton system between data and MC in the Z/DY events, which can not be corrected with the W+jets events due to the missing neutrinos.

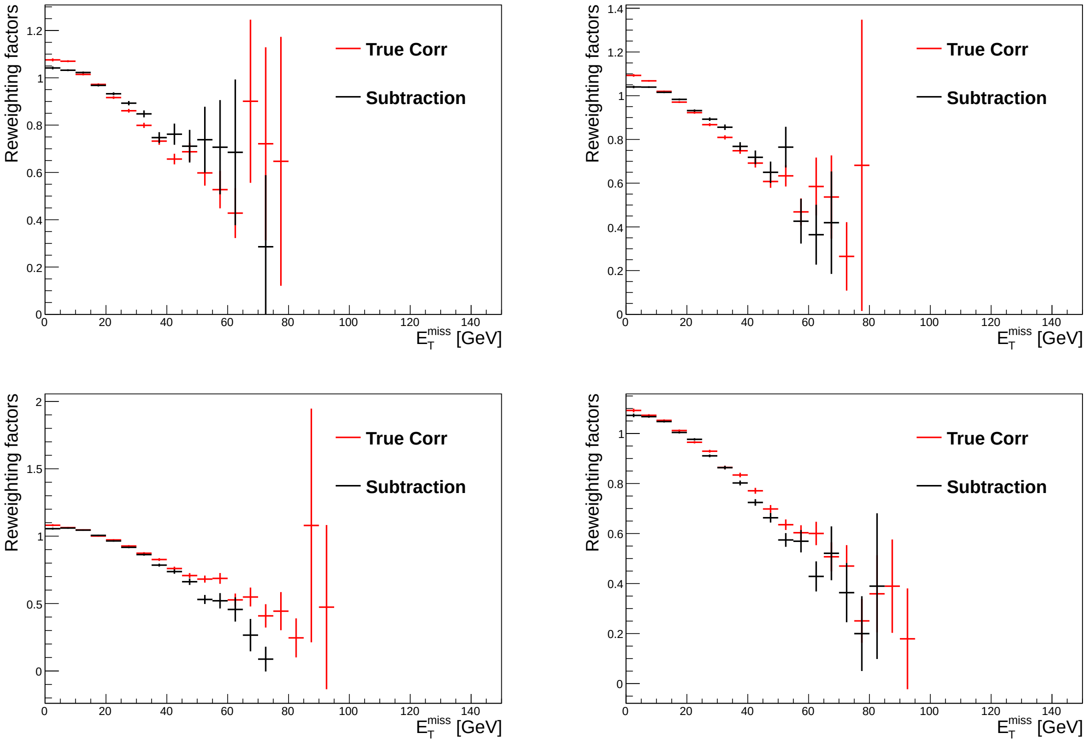


Figure 5.24: Correction factors as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ amplitude in the W+jets events in pseudo data and MC after the neutrino subtraction. The red corrections correspond to subtracting the true neutrino $E_{\mathrm{T}}^{\mathrm{miss}}$ event-by-event. The black ones correspond to subtracting the mean values of the neutrino p_{T} in lepton p_{T} bins. The error bar represents statistical error.

The numbers of the Z/DY events and the corresponding normalisation factors derived with the W+jets method are shown in Tab. 5.12, in comparison of that determined using the ABCD method. The normalisation factors from the two methods are in excellent agreement. In addition, the normalisation factors determined with the W+jets method are stable in the cutflow e.g. before and after the p_T^{ll} cut.

Systematic uncertainties have been studied, including the QCD contamination in the W+jets sample, the bias observed in the closure test, and the uncertainty on the scaling of the corrections from the W+jets sample to the Z/DY sample. The QCD contamination is negligible, based on a study using a loosen lepton identification sample. The systematic bias from the closure test amounts to 7% and the uncertainty of the scaling is 2%. However, it should be pointed out that this method can not correct the mismodelling related to the dilepton system in the Z/DY events. This mismodelling effect of the order of 5%-10% can be seen from Fig. 5.32.

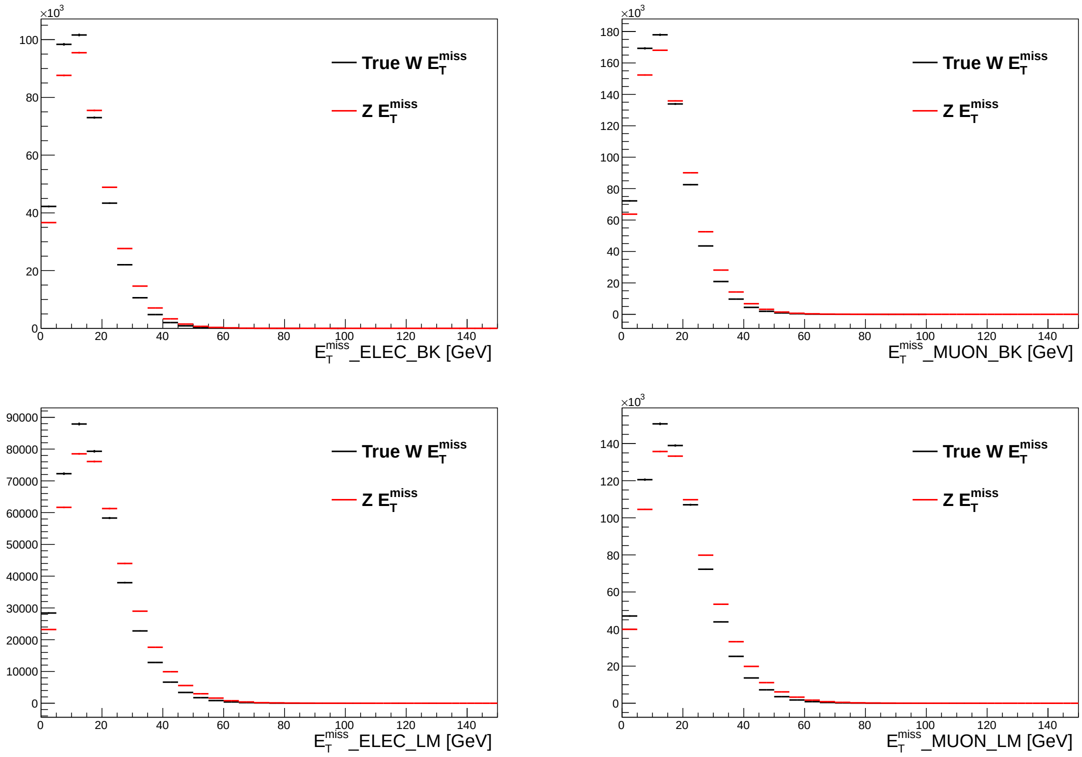


Figure 5.25: Fake $E_{\mathrm{T}}^{\mathrm{miss}}$ amplitude distribution in the W+jets events (black) and Z+jets (red). The difference represents the kinematic difference between W and Z.

5.5 SM WW backgrounds

When the Higgs mass is smaller than 200 GeV, the SM WW backgrounds can be easily estimated in the WW control region, which is defined by requiring $m_{\mathrm{ll}} > 80\,\mathrm{GeV}$, as shown in Fig. 5.34. Then a scale factor is obtained from WW control region by comparing data and MC. The scale factor may be directly applied to the WW MC in the signal region as a normalization. If the Higgs mass is larger than 200 GeV, it is very hard to define WW control region, the MC prediction is directly used. The contribution from the SM WW can also be extracted (fitted) in the limit setting, see Chapter 6.

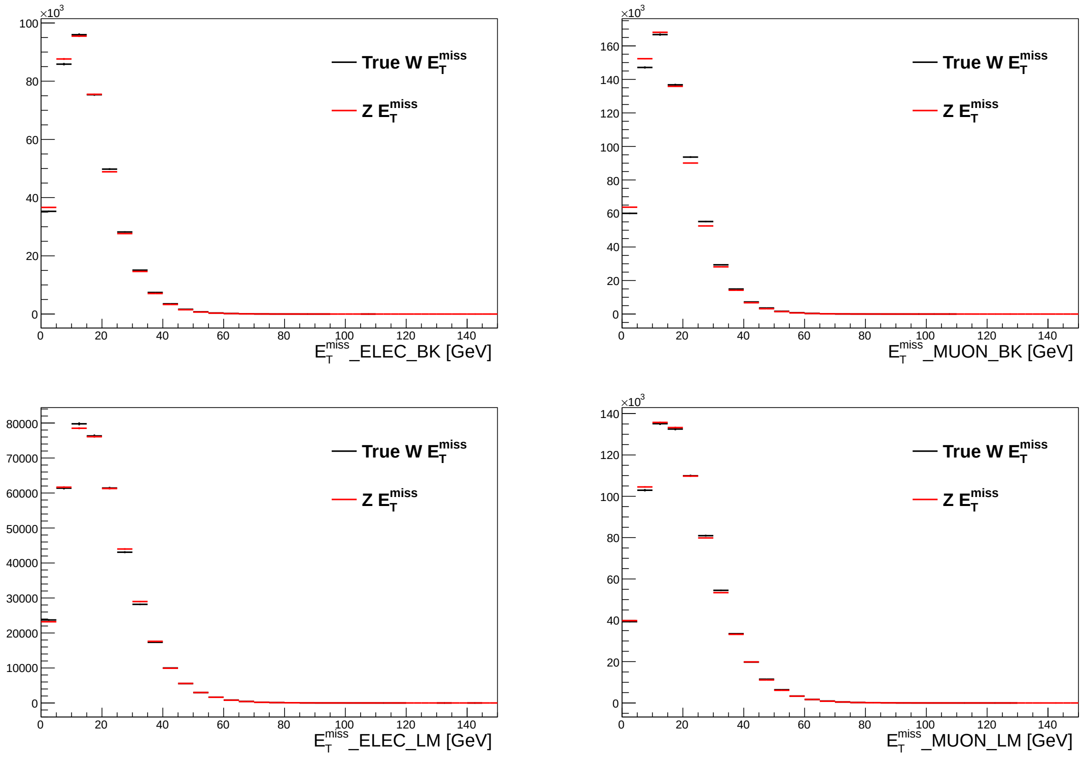


Figure 5.26: Same as Fig. 5.25 but after the 1.1 rescaling.

	ee B-K	$\mu\mu$ B-K	ee L-M	$\mu\mu$ L-M
Before correction $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$	46.3	83.4	224.9	442.9
After correction $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$	51.3	88.5	225.4	445.0
Before correction $p_T^{ll} > 45\,\mathrm{GeV}$, $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$	4.25	2.78	9.85	23.8
After correction $p_T^{ll} > 45\,\mathrm{GeV}$, $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$	4.90	2.98	9.90	23.9
Normalisation factor $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$	1.10	1.06	1.00	1.01
Normalisation factor $p_T^{ll} > 45\,\mathrm{GeV}$, $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$	1.15	1.08	1.01	1.00
Normalisation factor ABCD method $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$	1.05	1.18	0.98	1.01

Table 5.12: numbers of events and normalisation factors derived using the W+jets method in comparison with the ABCD results.

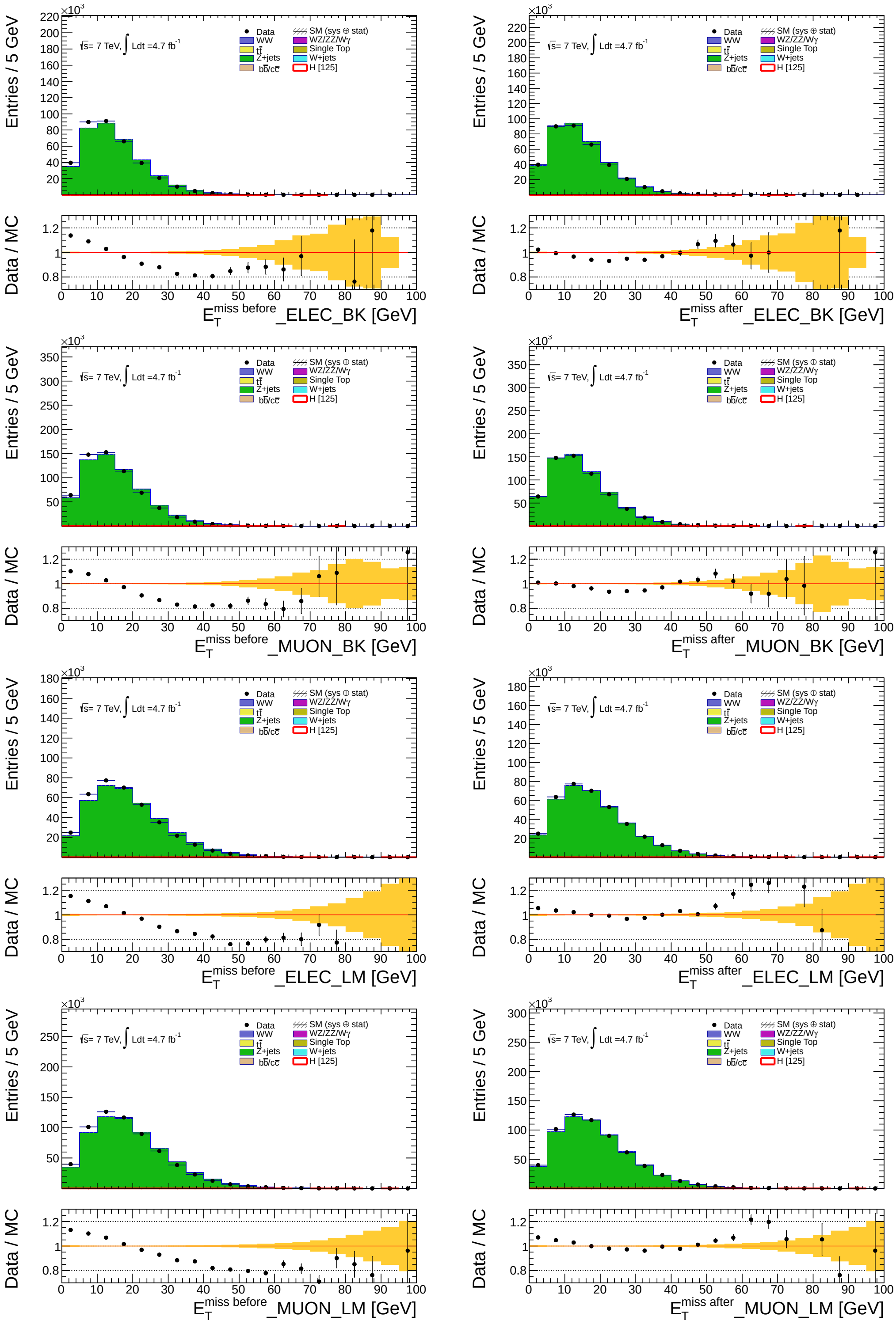


Figure 5.27: Distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ amplitude in the Z invariant mass window in data and MC (MC11b) before and after the correction. The left column is before correction and the right column is after correction. The first row is for the electron channel in the B-K period. The second row is for the muon channel in the B-K period. The third row is for the electron channel in the L-M period. The last row is for the muon channel in the L-M period.

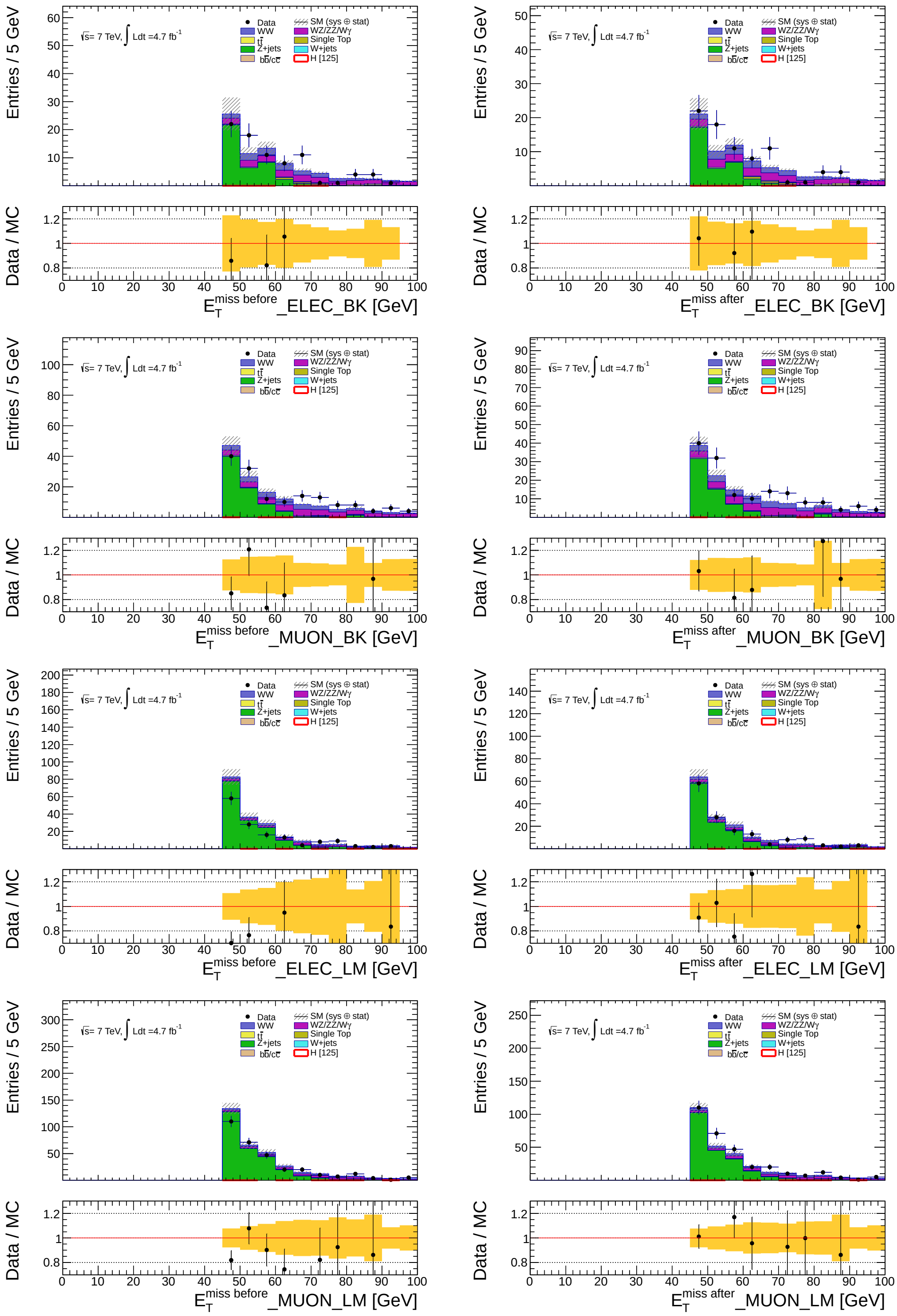


Figure 5.28: Similar plots as Fig. 5.27 but after additional cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{ll}.

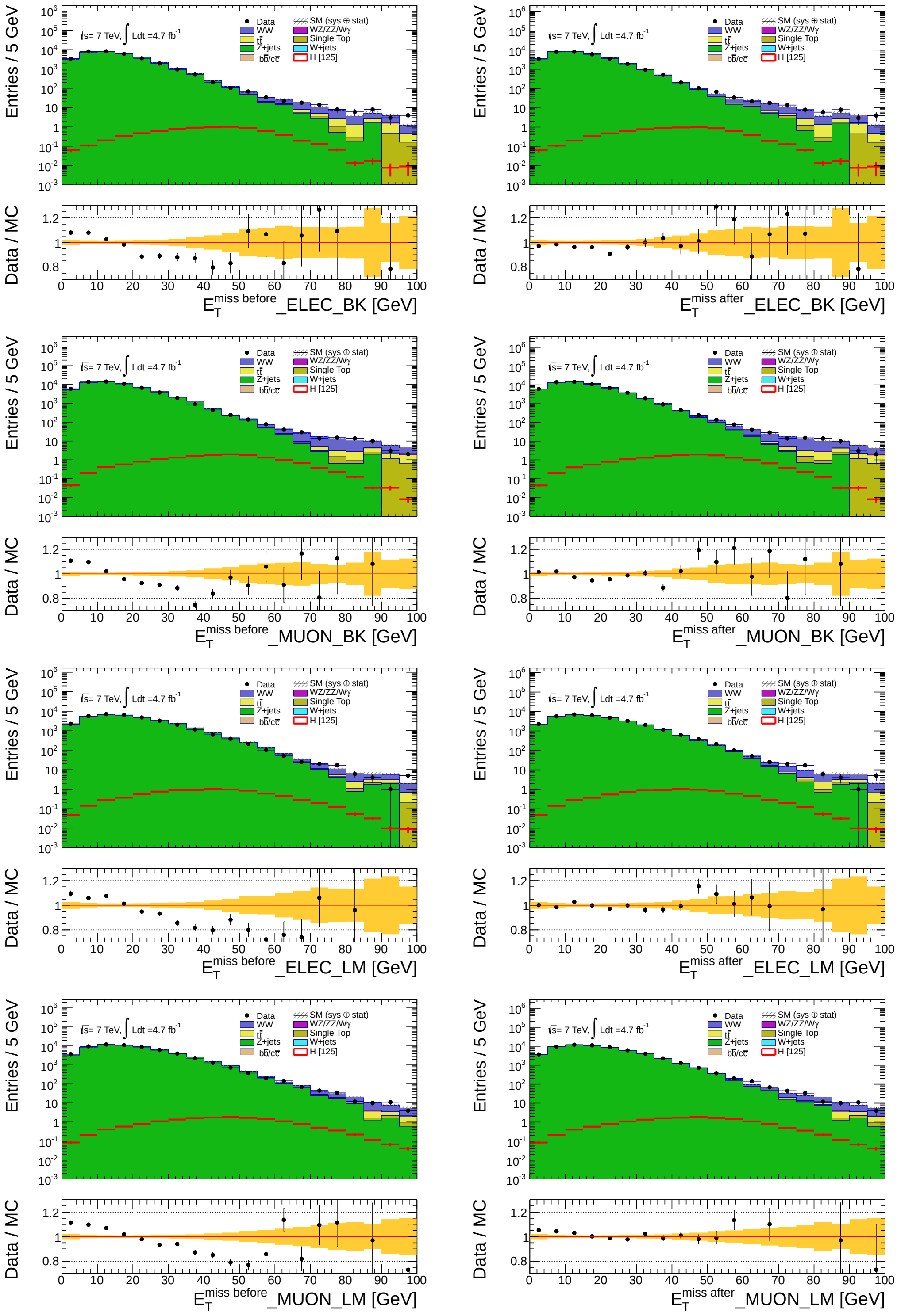


Figure 5.29: Similar plots as Fig. 5.27 but outside of the Z mass window (signal like).

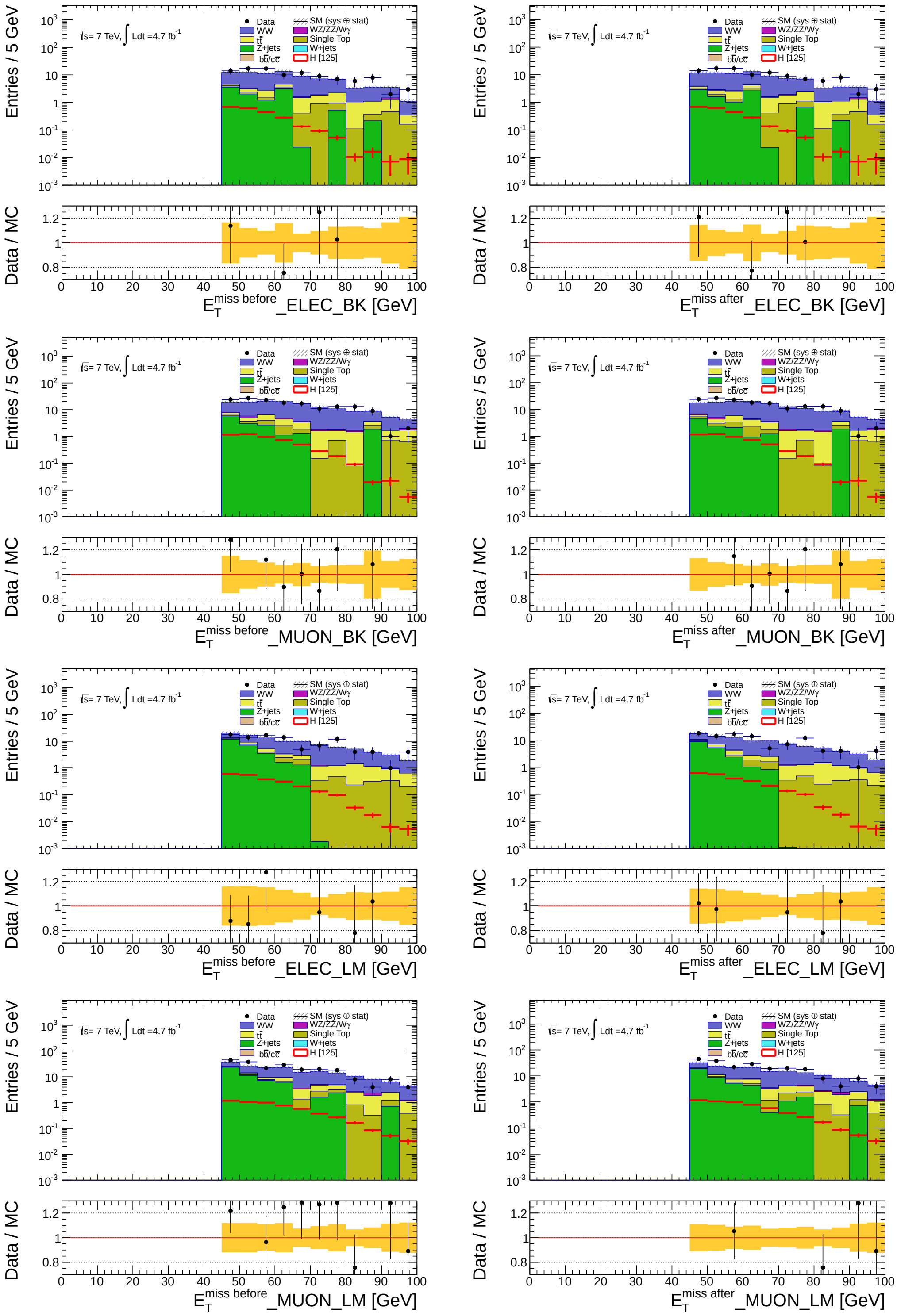


Figure 5.30: Similar plots as Fig. 5.27 but outside of the Z mass window and with addtional cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{ll}.

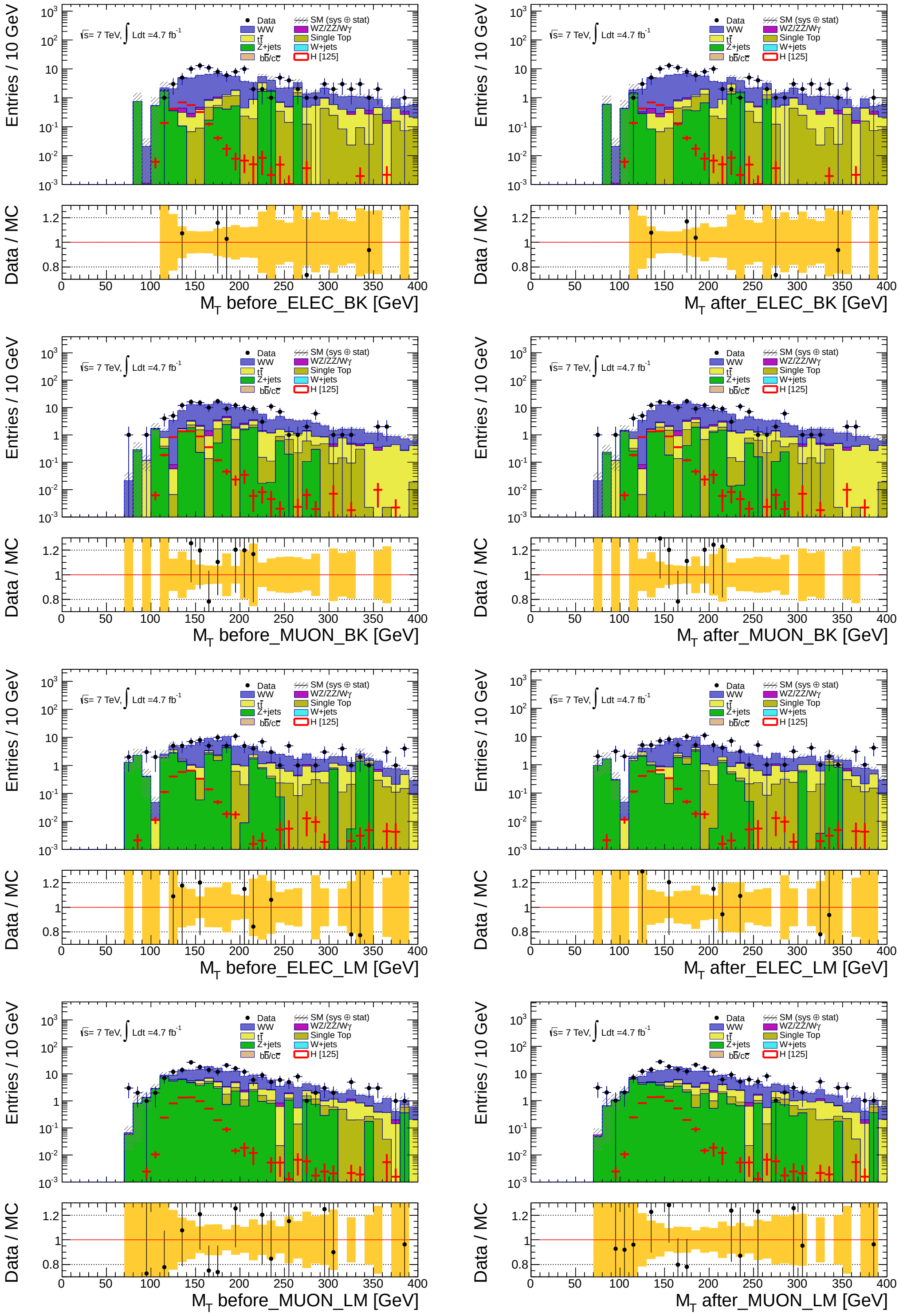


Figure 5.31: Distribution of m_T in the Higgs signal region (green) after $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_T^{ll} cuts in data and MC (MC11b) before and after the correction. No significant shape distortion is observed. The plot layout is the same as Fig. 5.27.

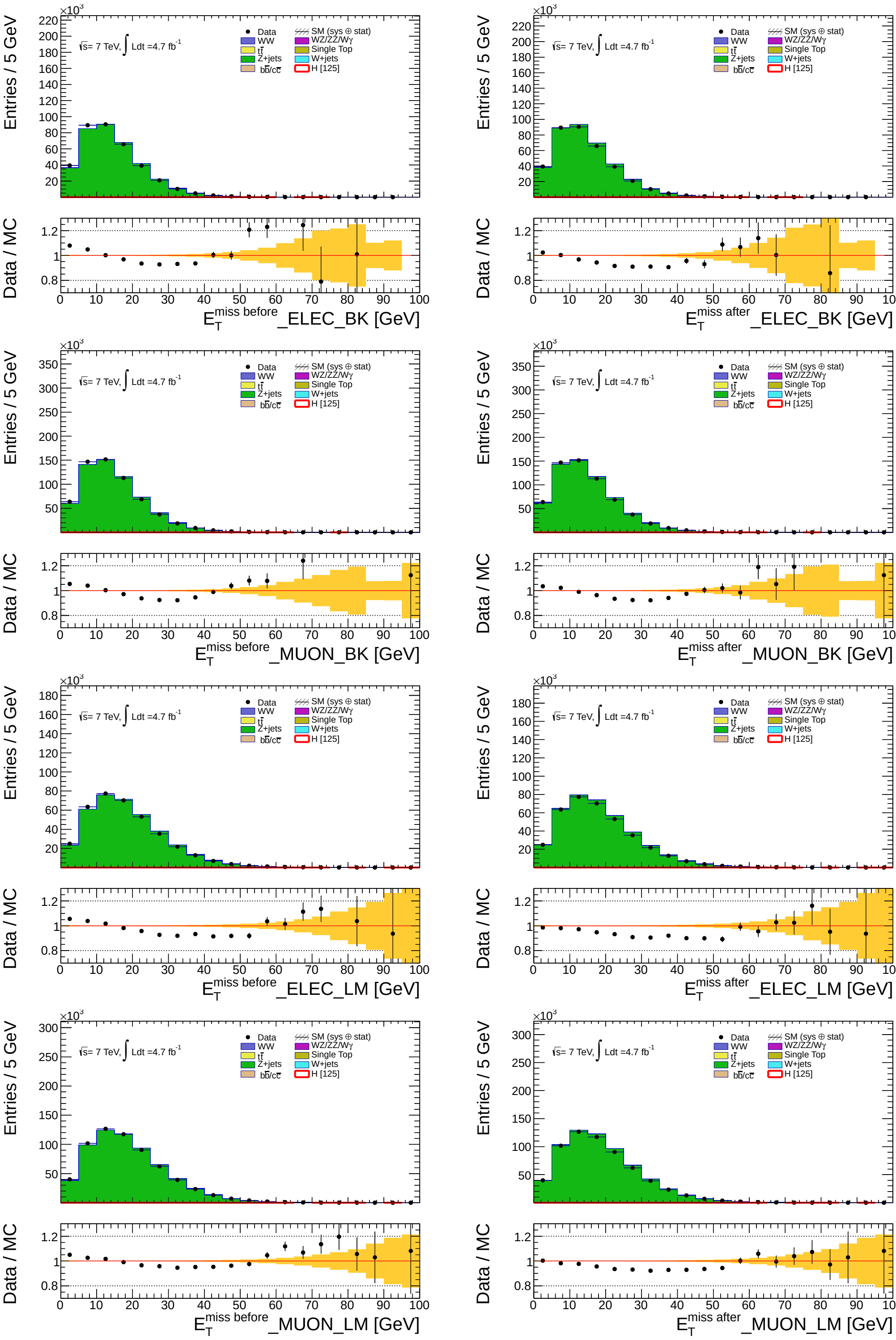


Figure 5.32: Similar plots as Fig. 5.27 but with MC11c instead of MC11b.

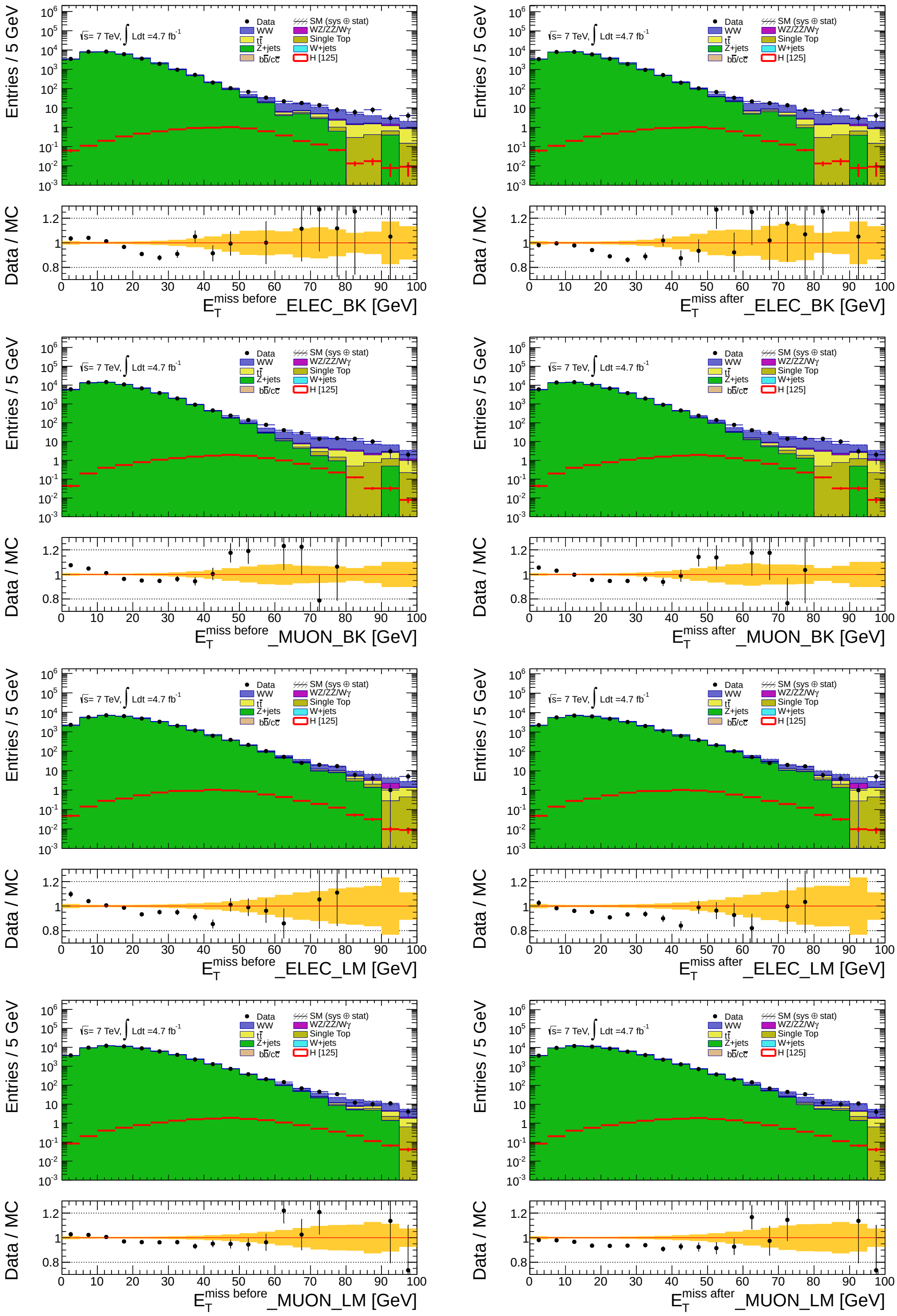


Figure 5.33: Similar plots as Fig. 5.27 but outside of the Z mass window and with MC11c.

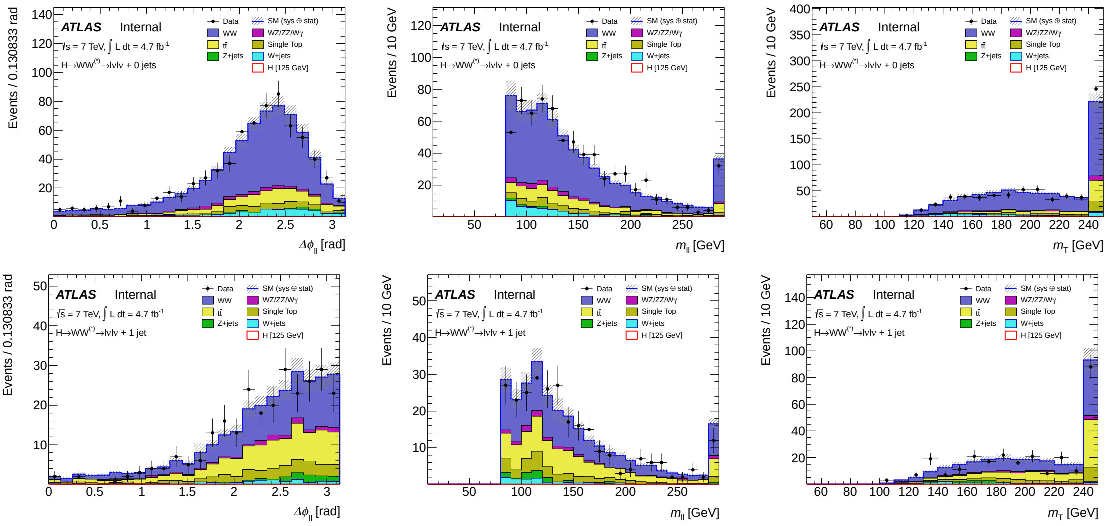


Figure 5.34: Kinematic distributions in the WW control region in the 0-jet bin (top) and 1-jet bin (bottom). All periods (B-M) are merged.

Chapter 6

Results of Higgs Search in 2011

In this chapter, the statistical analysis basis for the Higgs discovery or exclusion is first introduced in Section 6.1. This is then applied in Section 6.2 to a search for a Higgs boson in a model with a fourth generation using part of the 2011 data and in Section 6.3 for the SM Higgs boson with the full 2011 data.

6.1 Statistical Analysis Basis

Once the data analysis is performed with the number of observed events and expected background events, one needs to quantify for a signal (Higgs) discovery how large deviation from the non Higgs hypothesis is and in case with no significant deviation if a Higgs mass can be excluded at a given confidence level. The statistical analysis is performed using a binned likelihood function with the help of HistFactory framework and common RooStats tools [114, 115].

The likelihood function is constructed as the product of Poisson probability terms in each lepton flavor channel and jet bins, after all topological cuts except the m_T cut. In each region, the m_T shape is taken into account. In the H + 0-jet (H + 1-jet) signal regions, the m_T shapes are divided into five (three) bins. The control regions for WW and top (0-, 1-jet for WW and 1-, 2-jet for top) are introduced to extrapolate the background contribution. In the H + 2-jet signal region and control regions, the m_T is filled in only one bin due to the limited statistics. The bins in signal region are remapped to ensure enough statistics in each bin. In each channel (lepton or jet multiplicity), the likelihood function is defined as follows:

$$L(\mu, \dot{b}) = P(N|\mu s + \alpha \dot{b}) \times P(M|\dot{b}) \,, \tag{6.1}$$

where $\alpha = b_{\mathrm{SR}}^{\mathrm{exp}}/b_{\mathrm{CR}}^{\mathrm{exp}}$ is an extrapolation coefficient determined by MC, μ is the signal (s) strength parameter ($\mu = 1$ means the Standard Model Higgs prediction), and $\dot{b}$ is a free parameter representing the fitted background expectation in the control region for a given process (WW, top, etc), N is the measured number of events in the signal region, while M is an auxiliary control region measurement. The data-driven background estimation thus comes from maximising the likelihood over the parameters μ

and $\dot{b}$. The variation on α due to systematic uncertainties is parametrized as follows:

$$L(\mu, \dot{b}, \theta) = P(N|\mu s + \alpha\dot{b} \times \nu_\alpha(\theta)) \times P(M|\dot{b})\,, \tag{6.2}$$

in which $\nu_\alpha(\theta)$ is the parameterised response to α, θ is the nuisance parameter, which can be constrained by a Gaussian function: $N(\tilde{\theta}|\theta)$, representing an auxiliary measurement $\tilde{\theta}$ related to the nuisance parameter θ. Thus the whole model can be expressed as:

$$L(\mu, \vec{\theta}) = \{\prod_{l=\text{B-K,L-M}} \prod_{k=ee,e\mu,\mu\mu} \prod_{j=0}^{N_{\text{jets}}} \prod_{i=1}^{N_{\text{bins}}} P(N_{ijkl}|\mu s_{ijkl} + \sum_{m}^{N_{bg}} b_{ijklm})\} \times \{\prod_{i=1}^{N_\theta} N(\tilde{\theta}|\theta)\}\,, \tag{6.3}$$

in which the final likelihood function is the product of likelihood in two different data periods B-K and L-M, all lepton channels, jet bins and m_T bins.

The constraints on the numbers of background events for different processes are summarised here.

Standard Model $\boldsymbol{WW}$: For the low mass selections in the 0- and 1-jet channels, the normalization is constrained from a control region (obtained from the signal region by removing the $\Delta\phi_{ll}$ requirement and replacing the m_{ll} upper bound with a lower bound $m_{ll} > 80\,\text{GeV}$ ($m_{ll} > m_Z + 15\,\text{GeV}$) for the $e\mu$ (ee and $\mu\mu$) final states), two nuisance parameters are set for 0- and 1-jet bins. For 2-jet bin, the normalization is taken from Monte Carlo predictions.

$\boldsymbol{t\bar{t}}$ + single top: In the 0-jet channel, Monte Carlo is used along with a data-driven normalisation factor that corrects the jet veto efficiency with additional systematics on the normalisation factor. In the 1- and 2-jet channels, the normalization is constrained from a control region (obtained from the signal region definition by reversing the b-jet veto and removing the requirements on $\Delta\phi_{ll}$ and m_{ll}). There are two separate nuisance parameters for the 1- and 2-jet normalizations.

$\boldsymbol{Z/\gamma^*}$+jets: A data-driven method is used to determine the normalization in the signal region as well as the contamination in the control regions. In the signal region the shape is taken from Monte Carlo.

$\boldsymbol{W}$+jets: A data-driven method is used to determine both the normalization and the shape. There is a systematic uncertainty on the fake rate. Such uncertainty affects both the normalization and shape.

$\boldsymbol{WZ/ZZ/W\gamma}$: These backgrounds are taken from Monte Carlo.

Therefore the numbers of background events are constrained by two control regions. They are SM WW and top, presented in the Poisson distribution and a series Gaussian constraint terms.

When the likelihood function is constructed, a test statistic is defined as the ratio of log likelihood with the one side constraint:

$$\tilde{q}_\mu = \begin{cases} -2\ln\frac{\mathrm{L}(\mu,\hat{\hat{\theta}}(\mu))}{\mathrm{L}(0,\hat{\hat{\theta}}(0))} & \tilde{\mu} < 0, \\ -2\ln\frac{\mathrm{L}(\mu,\hat{\hat{\theta}}(\mu))}{\mathrm{L}(\hat{\mu},\hat{\theta})} & 0 \leq \tilde{\mu} \leq \mu, \\ 0 & \tilde{\mu} > \mu\,, \end{cases} \tag{6.4}$$

in which the $\hat{\mu}$ and $\hat{\theta}$ are the value obtained by maximizing the L and the $\hat{\hat{\theta}}$ comes from maximizing the L with a specific μ. The p-values of p_μ and p_b are derived from the sampling distributions $f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu)$ and $f(\tilde{q}_\mu|0,\hat{\theta}_0)$, respectively:

$$p_\mu = \int_{\tilde{q}_{\mu,\mathrm{obs}}}^{\infty} f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu)d\tilde{q}_\mu\,, \quad p_b = \int_{\infty}^{\tilde{q}_{\mu,\mathrm{obs}}} f(\tilde{q}_\mu|0,\hat{\theta}_0)d\tilde{q}_\mu\,. \tag{6.5}$$

In the Higgs discovery case, the p value is defined under the background-only hypothesis with a different constraint $\hat{\mu} > 0$:

$$p_0 = \int_{q_{0,\mathrm{obs}}}^{\infty} f(q_0|0,\hat{\theta}_0)dq_0\,. \tag{6.6}$$

And it can be interpreted as statistical significance Z:

$$Z = \Phi^{-1}(1-p_0)\,. \tag{6.7}$$

When Z is larger than 3 it is possible to claim an evidence or 5 for discovery. In the exclusion case, the CLs method is used and defined as:

$$\mathrm{CLs} = \frac{p_\mu}{1-p_b}. \tag{6.8}$$

The upper limit is usually set at 95% so that CLs=0.05. For a given μ, if the observed CLs is smaller than 0.05 and the μ is smaller than 1, it is said that the Higgs boson with such mass is excluded. It is CPU time consuming to calculate the p value to satisfy a given CLs, so an asymptotic approximation is introduced to simplify the computation [113].

6.2 Results for a Fourth generation Higgs Search

The SM is known to have three families of charged and neutral fermions. There is however no upper limit on the number of fermion famililities. A fourth family could in fact be the key to many unsolved puzzles, such as the hierarchies of the fermion mass spectrum including neutrino masses and mixing, electroweak symmetry breaking, baryogenesis, and a variety of interesting phenomena in CP and flavor physics [23, 24, 25, 26, 27].

As shown in Eq. 1.30, the interaction strength of the Higgs to a fermion is proportional to the fermion mass in Yukawa couplings. Thus, the production cross section is significantly enhanced by including the contribution of fourth generation quarks with heavy masses [22]. Fourth family leptons and quarks have been searched for previously by the LEP and Tevtran experiments and now by the LHC experiments. The most stringent lower mass limits at 95% CL of fourth family leptons and quarks are given in [16, 19, 20, 28, 29, 30, 31]:

$$m_{\nu_4} > 80.5 - 101.5\,\mathrm{GeV}, \tag{6.9}$$
$$m_{l_4} > 100.8\,\mathrm{GeV}\,, \tag{6.10}$$
$$m_{b_4} > 611\,\mathrm{GeV}\,, \tag{6.11}$$
$$m_{t_4} > 557\,\mathrm{GeV}\,. \tag{6.12}$$

In the SM, the dominant Higgs production process is the gluon-gluon fusion process where gluons from the colliding beams couple to a heavy quark loop from which the Higgs boson is emitted (see Fig. 1.3). The production cross section can be expressed as:

$$\sigma(pp \to HX) = \Gamma(H \to gg)\frac{\pi^2}{8m_H^3}\tau_H \int_{\tau_H}^{1} \frac{dx}{x} g(x, m_H^2) g(\tau_H/x, m_H^2)\,, \tag{6.13}$$

with

$$\Gamma(H \to gg) = \frac{G_F m_H^3}{36\sqrt{2}\pi}\left[\frac{\alpha_s(m_H^2)}{\pi}\right]^2 |I|^2\,, I = \sum_q I_q\,, \tag{6.14}$$

in which $g(x, Q^2)$ is the gluon distribution evaluated at x and Q^2, $G_F(\alpha_s)$ is the Fermi (strong) coupling constant. The quantity I_q is given in terms of $\lambda_q = m_q^2/m_H^2$:

$$I_q = 3[2\lambda_q + \lambda_q(4\lambda_q - 1)f(\lambda_q)]\,, \tag{6.15}$$

where

$$f(\lambda_q) = \begin{cases} -2\{\sin^{-1}\frac{1}{2\sqrt{\lambda_q}}\}^2, & \text{for } \lambda_q > \frac{1}{4}, \\ \frac{1}{2}(\ln\frac{\eta^+}{\eta_-})^2 - \frac{\pi^2}{2} - i\pi\frac{\eta^+}{\eta_-}, & \text{for } \lambda_q < \frac{1}{4}\,, \end{cases} \tag{6.16}$$

with $\eta^\pm = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \lambda_q}$. When the quark is very heavy, with $\lambda_q \gg 1, I_q \to 1$, otherwise if the quark is very light, $\lambda_q \ll 1, I_q \to 0$, so the contribution from the heaviest quark is dominant. When the fourth generation is included, the Higgs production corss section is enhanced over that of the SM by a factor:

$$R^{\mathrm{SM4/SM}}_{\sigma(gg\to H)} \equiv \frac{\sigma(gg \to H)_{\mathrm{SM4}}}{\sigma(gg \to H)_{\mathrm{SM}}} = \frac{|I_b + I_t + I_{t_4} + I_{d_4}|^2}{|I_b + I_t|^2}\,. \tag{6.17}$$

The dependence of the enhancement factor as a function of the Higgs boson mass m_H is shown in Fig. 6.1 [22] (left) for two different m_{d_4} values: infinite mass (in practice a value of 10 TeV is chosen) and 400 GeV. The t_4 mass is fixed as [32]:

$$m_{t_4} = m_{d_4} + 50 + 20 \times \ln(\frac{m_H}{115[\mathrm{GeV}]}), \tag{6.18}$$

to be consistent with the constraint of eletroweak precision data. The maximum enhancement factor of about 9 is reached at the small Higgs boson mass m_H value where $I_b \to 0$ and $I_t, I_{t_4}, I_{d_4} \to 1$. At low m_H, the enhancement is independent of the quark mass value of the fourth generation. At higher Higgs mass values, the heavier the quark mass, the smaller the enhancement factor. The heavy m_{q_4} scenario may not be physical as when it is beyond about 500 GeV the weak interaction among heavy particles becomes strong and perturbation theory breaks down. However, since the enhancement is the smallest, the resulting exclusion limits would be more conservative.

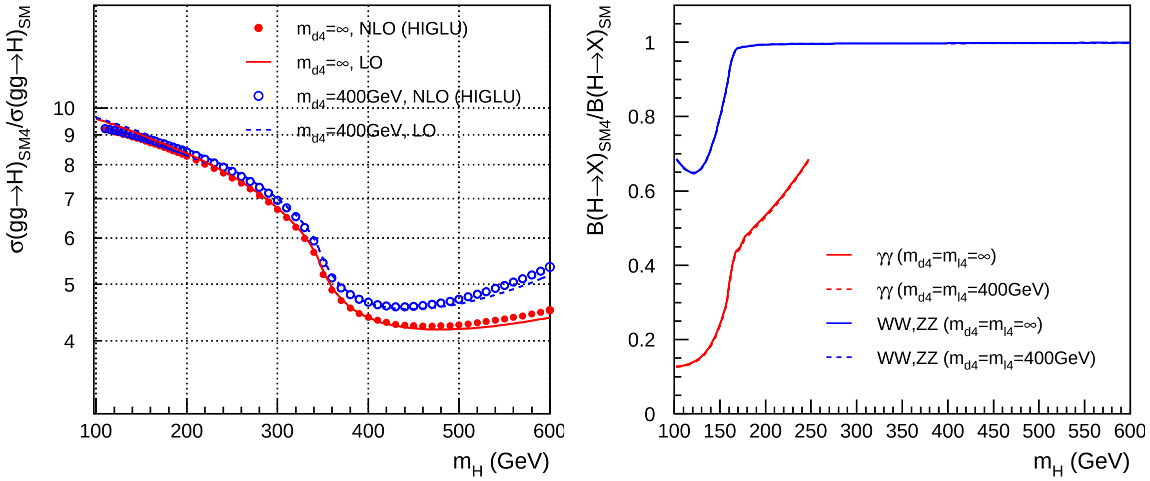


Figure 6.1: Enhancement factors of the Higgs production cross sections in a fourth generation over that of the SM as a function of m_H for two scenarios with $m_{d4} = 10$ TeV and 400 GeV(left) and branching ratios of Higgs decays in a fourth generation model over that in the SM for two mass scenarios (right).

In a similar way, the Higgs decay branching fractions are also affected in the presence of the fourth family. Indeed, the decay mode $H \to gg$ is the reverse of the gluon-gluon fusion process $gg \to H$. The partial decay width of the decay mode $H \to gg$ in a fourth generation model with respect to that of the SM is thus enhanced by the same factor as the corresponding Higgs production cross section in the gluon-gluon fusion process.

The decay mode $H \to \gamma\gamma$ is similar to $H \to gg$ except that charged leptons, the W-boson and charged Higgs bosons also contribute to the loop. The partial decay width is [33].

$$I = \sum_q Q_q^2 I_q + \sum_l Q_l^2 I_l + I_W + I_S \tag{6.19}$$

where Q_f denotes the charge of fermion f in units of e and

$$I_q = 3\left[2\lambda_q + \lambda_q(4\lambda_q - 1)f(\lambda_q)\right] , \tag{6.20}$$

$$I_l = 2\lambda_l + \lambda_l(4\lambda_l - 1)f(\lambda_l) , \tag{6.21}$$

$$I_W = 3\lambda_W(1 - 2\lambda_W)f(\lambda_W) - 3\lambda_W - \frac{1}{2} , \tag{6.22}$$

$$I_S = -\lambda_S\left[1 + 2\lambda_S f(\lambda_S)\right] , \tag{6.23}$$

and $\lambda_i = m_i^2/m_H^2$. In a fourth generation model, both I_Q and I_l terms receive additional contributions from fourth generation quarks q_4 and lepton l_4.

Finally, the partial decay width $\Gamma(H \to \gamma Z)$ is also affected by additional fourth generation quark loops. Therefore the branching fractions in a fourth generation model look different from that of the SM in particular for gg, $\gamma\gamma$ and γZ modes at low Higgs mass values [22].

For the relevant decay modes $\gamma\gamma$ and WW, ZZ, the ratio of the branching fractions in a fourth generation model over that in the SM

$$R^{\mathrm{SM4/SM}}_{B(H\to X)} \equiv \frac{B(H \to X)_{\mathrm{SM4}}}{B(H \to X)_{\mathrm{SM}}} \tag{6.24}$$

is compared in the two mass scenarios in Fig. 6.1 (right), which also shows that the effect of the mass scenarios is small.

The overall enhancement of the product of the Higgs production cross section and the Higgs branching fraction in a fourth family over that in the SM

$$R^{\mathrm{SM4/SM}} \equiv \frac{\left[\sigma(gg \to H) \times B(H \to X)\right]_{\mathrm{SM4}}}{\left[\sigma(gg \to H) \times B(H \to X)\right]_{\mathrm{SM}}} \tag{6.25}$$

is shown in Fig. 6.2 as a function of the Higgs mass for decay modes $X = \gamma\gamma$, WW and ZZ and for the two mass scenarios. The numerical values of $R^{\mathrm{SM4/SM}}$ for 59 Higgs mass points ranging from 110 GeV to 600 GeV are given in the linked web pages [36]. It is also checked that in the considered Higgs mass range, the effect on $R^{\mathrm{SM4/SM}}_{\sigma(gg\to H)}$ of using different PDFs for the cross section calculation is very small (0.2%) [22].

Based on the enhancement factors provided above, a search for a fourth generation Higgs has been performed by ATLAS combining the $WW^{(*)}$ channel with several other channels using part of the 2011 data sample [35, 34]. The fourth generation Higgs in the mass range between 120 GeV and 600 GeV is excluded at 95% CL as shown in Fig. 6.3.

6.3 Results for SM Higgs Search with 2011 Data

The results of the statistical analysis for the full 2011 data sample in the $H \to WW^{(*)} \to l\nu l\nu$ channel are summarised here. The p value as a function of m_H is displayed in Fig. 6.3. No significant excess of events over the expected background is observed over the entire mass range with the lowest p-value observed is 0.15. The

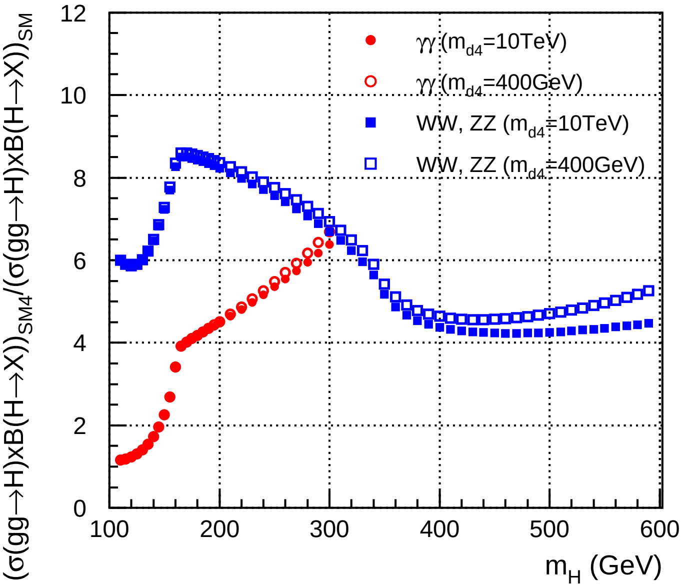


Figure 6.2: Final enhancement factor of the Higgs to WW(blue) and diphoton(red) final states in a fourth generation over that of the SM as a function of m_H for two scenarios with $m_{d4} = 10$ TeV and 400 GeV.

observed and expected cross section upper limits at 95% CL, as a function of m_H and normalised to the SM cross section, for the combined 0-, 1- and 2-jet analyses, are shown in Fig. 6.5. The limits exclude a SM Higgs boson with a mass in the range from 133 GeV to 261 GeV at 95% CL, while the expected exclusion range in the absence of a signal is $127 < m_H < 233$ GeV.

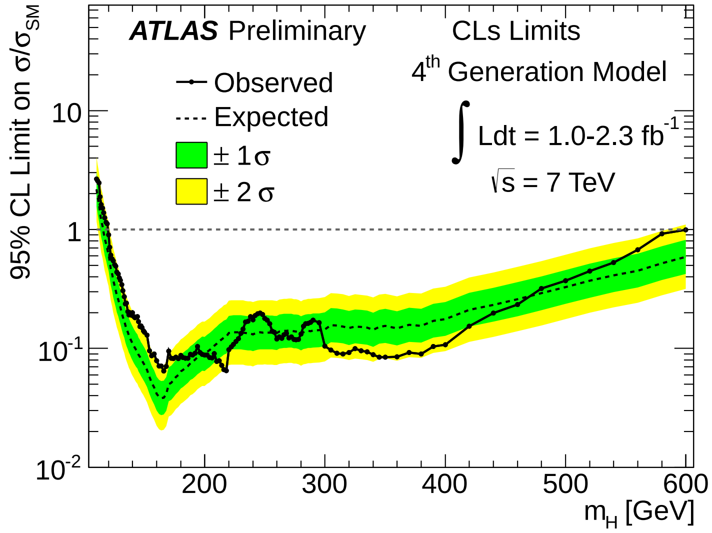


Figure 6.3: Observed (solid) and expected (dashed) 95% CL upper limits based on part of 2011 data on a fourth generation Higgs boson production cross section, normalised to the SM cross section, as a function of m_H. The inner (green) and outer (yellow) regions indicate the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands on the expected limits, respectively.

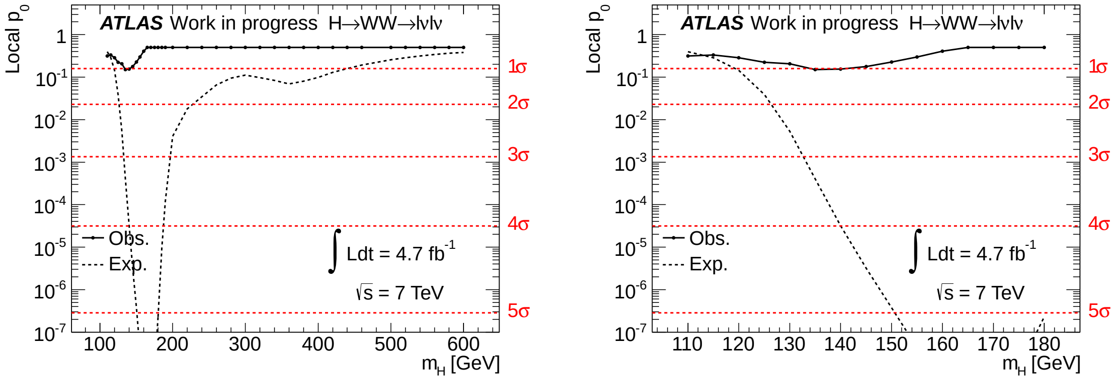
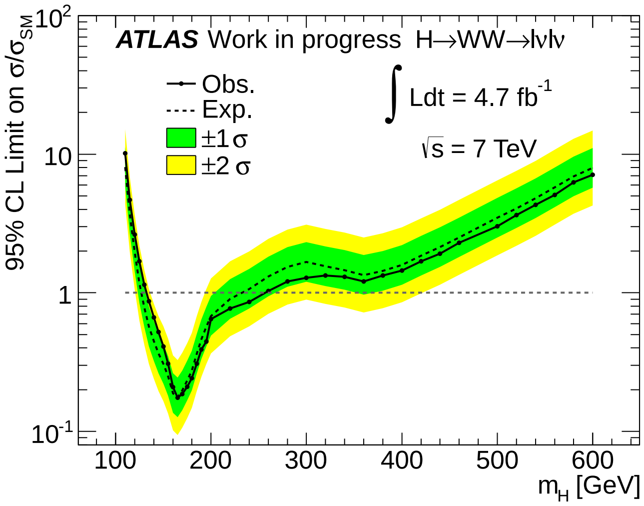


Figure 6.4: Background-only p_0 calculated using 2011 4.7 fb^{-1} data. The p_0 as a function of Higgs mass from 110 to 600 GeV is shown in the left plots, the dash line is the expected number from MC prediction and line with dot is the observed one. The right plot shows the same result zooming in 110 to 180 GeV.



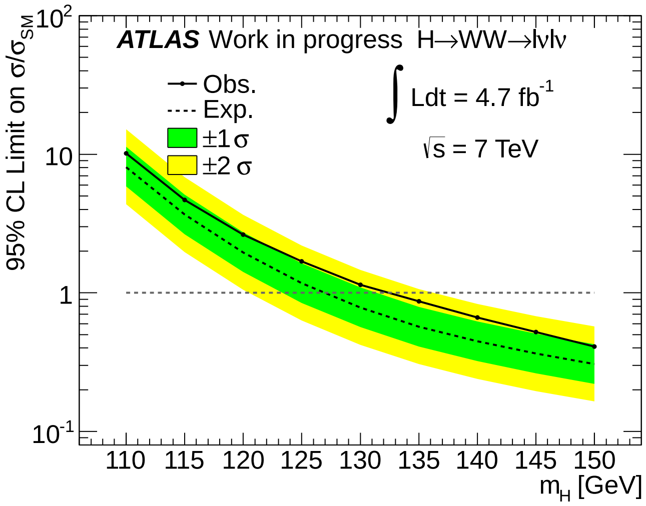


Figure 6.5: Observed (solid) and expected (dashed) 95% CL upper limits based on 2011 4.7 fb^{-1} data on the Higgs boson production cross section, normalised to the SM cross section, as a function of m_H, over the full mass range considered in the analysis (top) and restricted to the range $m_H < 150\,\text{GeV}$ (bottom). The inner (green) and outer (yellow) regions indicate the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands on the expected limits, respectively.

Chapter 7

Results of Higgs Search in 2012

The analysis of 2012 data (sample used for the ICHEP 2012 conference) corresponding to an integrated luminosity of $5.8\,\text{fb}^{-1}$ follows closely the analysis of the 2011 data [117]. Nevertheless there are a number of differences and new features which are described in Section 7.1.

7.1 Improvements and New Features in 2012

In 2012, the center-of-mass energy has been increased from 7 TeV to 8 TeV. By the end of June, a sample with an integrated luminosity exceeding that of 2011 has been taken by the ATLAS detector and analyzed [116] with the final results published [118]. In 2011, the average number of interactions per bunch crossing was around 10. In 2012, the average is increased to around 20 so that the analysis is more challenging and has to be improved to reduce its dependence on the pile up effects.

The first change is on the trigger, which has been updated to cope with the increasing interating rate, as described in Tab 7.1.

2011	ee channel	$\mu\mu$ channel	eμ channel
B - I	EF_e20_medium	EF_mu18_MG	EF_e20_medium || EF_mu18_MG
J	EF_e20_medium	EF_mu18_MG_medium	EF_e20_medium || EF_mu18_MG_medium
K	EF_e22_medium	EF_mu18_MG_medium	EF_e22_medium || EF_mu18_MG_medium
L - M	EF_e22vh_medium1	EF_mu18_MG_medium	EF_e22vh_medium1 || EF_mu18_MG_medium

2012 ee	EF_e24vhi_medium1 || EF_e60_medium1
2012 $\mu\mu$	EF_mu24i_tight || EF_mu36_tight
2012 $e\mu$ & μe	EF_e24vhi_medium1 || EF_e60_medium1 || EF_mu24i_tight || EF_mu36_tight

Table 7.1: Proposed trigger setup for 2011 periods (upper table) and 2012 run (lower table).

The next main modification in 2012 is on the lepton selection [135]:

- The topological cluster based calorimeter isolation on electrons is used instead of the cell based algorithm in 2011 to reduce the pile up dependence.

- The track isolation of electrons is improved due to the change of the lower threshold for each tracks which are used for the calculation of the track isolation from 900 MeV in 2011 to 400 MeV in 2012.
- The electron track impact parameter resolutions are significantly improved in 2012 due to the new Gaussian Sum Fitter (GSF) algorithm [106].
- Pile up correction is applied on muon isolation.
- The η range of muon is extended from 2.4 to to 2.5.

All of the impact parameter cuts are reoptimzed. The modified cuts for muons and electrons are shown in Tabs. 7.2 and 7.3.

2011 Cut	2012 Cut
pile up corrected etcone30/$p_{\mathrm{T}} < 0.14$	topo_etconecor30/$p_T < 0.16$
ptcone30/$p_{\mathrm{T}} < 0.13$	ptcone30/$p_T < 0.12$ for $15 - 25\,\mathrm{GeV}$ ptcone30/$p_T < 0.16$ for $> 25\,\mathrm{GeV}$
$d_0(\mathrm{sig}) < 10$	$d_0(\mathrm{sig}) < 3.0$
$\|z_0\|$(w.r.t.PV) $< 1\,\mathrm{mm}$	$\|z_0 \sin\theta\| < 0.4\,\mathrm{mm}$

Table 7.2: Electron isolation and impact parameter cuts.

2011 Cut	2012 Cut
mu_staco_etcone30/$p_{\mathrm{T}} < 0.14$	etconecor30/$p_T < 0.014 \times p_T - 0.15$ && etconecor30/$p_T < 0.20$
mu_staco_ptcone30/$p_{\mathrm{T}} < 0.15$	ptcone30/$p_T < 0.01 \times p_T - 0.105$ && ptcone30/$p_T < 0.15$
$d_0(\mathrm{sig}) < 3$	$d_0(\mathrm{sig}) < 3$
$\|z_0\|$(w.r.t.PV) $< 1\,\mathrm{mm}$	$\|z_0 \sin\theta\| < 1.0\,\mathrm{mm}$

Table 7.3: Muon isolation and impact parameter cuts.

Due to the increased pile up interactions, the calorimeter based $E_{\mathrm{T}}^{\mathrm{miss}}$ has seen it resolution degraded such that for a same $E_{\mathrm{T}}^{\mathrm{miss}}$ cut, the Z/DY background has significantly increased for the same flavour channels. For this reason, for the first results reported in [118], the analysis is limited to the $e\mu$ channel which has the dominant sensivity. To improve the search sensitivity in the analysis of the 2012 data, the channel has been splitted into two independent channels: $e\mu$ and μe where the first lepton stands for the leading lepton.

The jets selection is slightly changed. The p_{T} threshold remains 25 GeV for $|\eta| < 2.5$ and is raised to 30 GeV for $|\eta| > 2.5$. The JVF threshold is changed from 0.75 to 0.5 to reduce the pile up dependence, shown in Fig 7.1. The b-tagging algorithm has also been changed from a combined neutral network based one (JetFitterCombNN [119]) at 80% working point to a multivariate based one (MV1 [120]) at 85% working point.

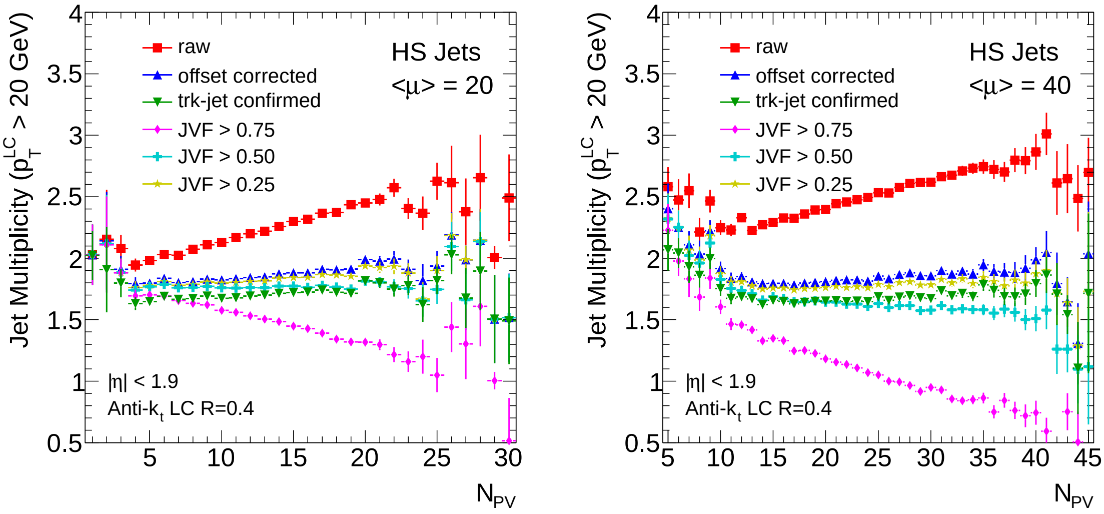


Figure 7.1: The JVF efficiency as a function of N_{PV} for calibrated LC(local cluster weighting)+JES jets for $<\mu>= 20$ (left) and $<\mu>= 40$ (right) running conditions.

The main new feature in the analysis of 2012 is that the signal region has been blinded with the following cuts:

- $\Delta\phi_{ll} < 1.8\,,$
- $m_{ll} < 50\,\mathrm{GeV}\,,$
- $82.5 < m_T < 140\,\mathrm{GeV}\,,$
- b-jet veto .

Thousands of kinematic distributions have been compared between data and the total SM background contributions derived either with the data-driven methods or with data-validated MC predictions in different background control regions. Good agreement has been observed in essentially all the distributions. One example is shown in Fig. 7.2 for the m_T distribution in the WW 0-jet control region for the $e\mu$ and μe channels. The KS prob shown on the plots is Kolmogorov-Smirnov probability, which describes the agreement between data and MC prediction. If the probability is close to one it means that these two distributions are compatible [121]. Figure 7.3 shows the data and MC comparison of m_T in top control region in which one of the jets must be b-tagged.

When the blinded analysis is approved by the Higgs working group, the signal region is unblinded. Figure 7.4 shows the m_T distribution in the signal region for 0-jet analysis, in which a significant excess over the background contribution observed in both the $e\mu$ and μe channels.

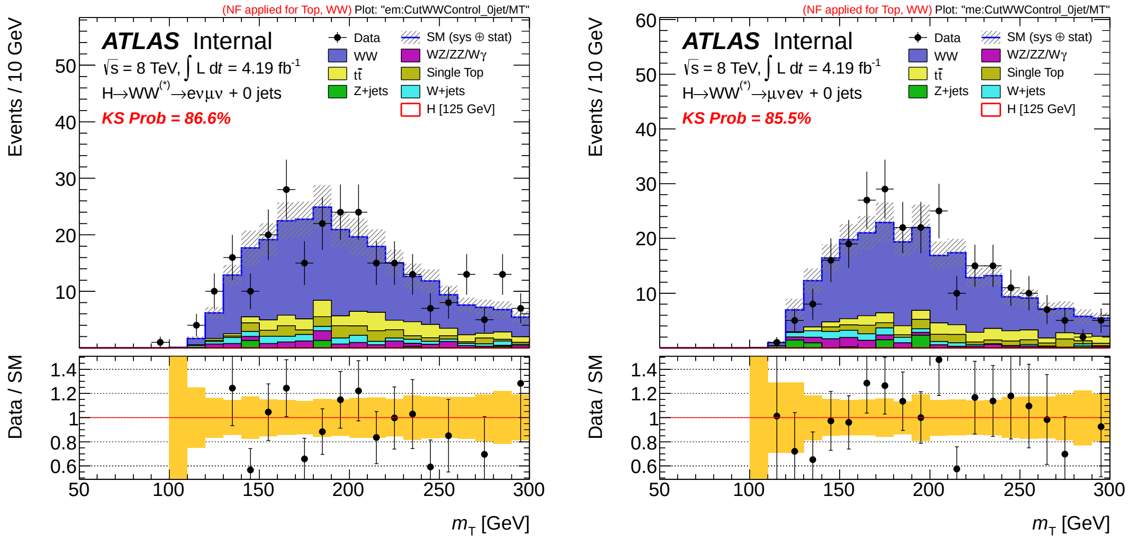


Figure 7.2: Transverse mass m_T distribution in the 0-jet analysis in the WW control region for the $e\mu$ (left) and μe (right) channels. The hatched band around the sum of backgrounds indicates the total systematic uncertainty in the normalization (but not the shape) of the various components.

7.2 Results for SM Higgs Search with 2012 Data

The result of the cut based analysis is shown in Tab. 7.4. Event excess is observed in all the channels: $e\mu$, μe, 0- and 1-jet. The statistical analysis is then repeated as for the 2011 analysis. Figure 7.5 shows the observed and expected p_0 value and the fitted signal strength μ over the range $110 < m_H < 190\,\mathrm{GeV}$, for all jets bins. An excess is observed over the expected background, reflected by a low observed p_0 and a fitted μ which deviates from zero. Due to the limited mass resolution for this analysis, the p_0 distribution is rather flat around $m_H = 125\,\mathrm{GeV}$. The value of p_0 at $m_H = 125\,\mathrm{GeV}$ is 8×10^{-4}, corresponding to 3.1 standard deviations. The minimum value of p_0, found at $m_H = 120\,\mathrm{GeV}$, is 6×10^{-4}, which corresponds to 3.2 standard deviations. The significance exceeds three standard deviations for a possible signal within the mass range $110 - 130\,\mathrm{GeV}$. The expected p_0 for a Higgs with $m_H = 125\,\mathrm{GeV}$ is 0.05 or 1.6 standard deviations. The fitted signal strength is also shown in Fig. 7.5 and amounts to $\mu = 2.1^{+0.8}_{-0.7}$ at $m_H = 125\,\mathrm{GeV}$. The increase of the fitted signal strength at lower m_H is due to the decreasing expected $\sigma \times Br$ for the signal.

The expected 95% CLs limit on $\sigma/\sigma_{\mathrm{SM}}$ excludes a SM Higgs boson with a mass down to 129 GeV. However, due to the observed excess in data, the observed exclusion lower limit is only at 145 GeV.

7.3 Combined 7 TeV and 8 TeV Results

The results obtained with the $5.8\,\mathrm{fb}^{-1}$ of 8 TeV data acquired in 2012 are combined with the published $4.7\,\mathrm{fb}^{-1}$ of 7 TeV results [102, 116]. The 7 TeV analysis resulted in

Cutflow evolution in the different signal regions

Zero Jet	Signal	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
Jet Veto	47.5 ± 0.4	1308 ± 9	125 ± 4	184 ± 4	109 ± 6	850 ± 32	138 ± 4	2714 ± 34	2691
$p_{\mathrm{T}}^{\ell\ell} > 30$ GeV	43.4 ± 0.4	1077 ± 8	99 ± 4	165 ± 4	98 ± 5	47 ± 8	102 ± 2	1589 ± 14	1664
$m_{\ell\ell} < 50$ GeV	34.9 ± 0.4	244 ± 4	33 ± 2	28 ± 2	17 ± 2	5 ± 2	29 ± 1	356 ± 6	421
$\Delta\phi_{\ell\ell} < 1.8$	33.6 ± 0.4	234 ± 4	32 ± 2	27 ± 2	17 ± 2	4 ± 2	25 ± 1	339 ± 6	407
One Jet	Signal	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
1 jet	24.9 ± 0.3	396 ± 5	74 ± 3	1652 ± 12	479 ± 12	283 ± 20	68 ± 3	2953 ± 27	2874
b-jet veto	21.1 ± 0.3	334 ± 4	56 ± 2	349 ± 6	115 ± 6	236 ± 18	53 ± 2	1144 ± 21	1115
$\lvert\mathbf{p}_{\mathrm{T}}^{\mathrm{tot}}\rvert < 30$ GeV	12.2 ± 0.2	210 ± 3	30 ± 2	139 ± 4	63 ± 5	124 ± 14	23 ± 2	590 ± 15	611
$Z \to \tau\tau$ veto	12.2 ± 0.2	204 ± 3	29 ± 2	133 ± 3	61 ± 5	98 ± 12	23 ± 2	547 ± 14	580
$m_{\ell\ell} < 50$ GeV	9.2 ± 0.2	37 ± 1	10 ± 1	21 ± 1	12 ± 2	16 ± 5	8.0 ± 0.9	104 ± 6	122
$\Delta\phi_{\ell\ell} < 1.8$	8.6 ± 0.2	34 ± 1	9 ± 1	20 ± 1	11 ± 2	3 ± 2	6.4 ± 0.7	84 ± 4	106
Two Jets	Signal	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
≥ 2 jets	14.5 ± 0.2	139 ± 3	30 ± 2	7039 ± 24	376 ± 11	104 ± 12	71 ± 4	7759 ± 29	7845
b-jet veto	9.6 ± 0.2	95 ± 2	19 ± 1	356 ± 6	44 ± 4	62 ± 9	21 ± 2	597 ± 12	667
$\lvert\Delta Y_{\mathrm{jj}}\rvert > 3.8$	2.0 ± 0.1	8.3 ± 0.6	2.0 ± 0.4	31 ± 2	5 ± 1	4 ± 2	1.4 ± 0.5	52 ± 3	44
Central jet veto (20 GeV)	1.6 ± 0.1	6.5 ± 0.5	1.3 ± 0.3	16 ± 1	4 ± 1	1 ± 1	0.5 ± 0.3	29 ± 2	22
$m_{\mathrm{jj}} > 500$ GeV	1.1 ± 0.0	3.2 ± 0.4	0.7 ± 0.2	6.2 ± 0.7	1.8 ± 0.6	0.0 ± 0.0	0.0 ± 0.2	12 ± 1	13
$\lvert\mathbf{p}_{\mathrm{T}}^{\mathrm{tot}}\rvert < 30$ GeV	0.8 ± 0.0	1.7 ± 0.3	0.3 ± 0.1	2.5 ± 0.5	0.8 ± 0.4	0.0 ± 0.0	0.0 ± 0.2	5.4 ± 0.7	6
$Z \to \tau\tau$ veto	0.7 ± 0.0	1.8 ± 0.3	0.3 ± 0.1	2.4 ± 0.4	0.8 ± 0.4	0.0 ± 0.0	0.0 ± 0.2	5.2 ± 0.7	6
$m_{\ell\ell} < 80$ GeV	0.7 ± 0.0	0.6 ± 0.2	0.1 ± 0.1	0.8 ± 0.3	0.3 ± 0.2	0.0 ± 0.0	0.0 ± 0.2	1.9 ± 0.5	3
$\Delta\phi_{\ell\ell} < 1.8$	0.6 ± 0.0	0.5 ± 0.2	0.1 ± 0.1	0.5 ± 0.3	0.3 ± 0.2	0.0 ± 0.0	0.0 ± 0.2	1.4 ± 0.4	2

Composition of main control regions

	Signal	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
WW 0-jet	0.3 ± 0.0	531 ± 5	43 ± 2	104 ± 3	62 ± 4	11 ± 4	38 ± 1	789 ± 9	820
WW 1-jet	0.1 ± 0.0	112 ± 3	13 ± 1	80 ± 3	34 ± 3	9 ± 4	7.7 ± 0.8	256 ± 6	255
Top 1-jet	2.2 ± 0.1	39 ± 2	10 ± 1	489 ± 6	195 ± 7	28 ± 7	7 ± 1	768 ± 12	840
Top 2-jet	4.9 ± 0.1	45 ± 2	11.7 ± 1.0	6371 ± 23	315 ± 10	45 ± 8	52 ± 3	6840 ± 26	7178

Signal region yield for $e\mu$ and μe channels separately

	0-jet $e\mu$	0-jet μe	1-jet $e\mu$	1-jet μe
Total bkg.	177 ± 4	162 ± 4	43 ± 2	40 ± 3
Signal	18.7 ± 0.3	14.9 ± 0.2	4.3 ± 0.1	4.2 ± 0.1
Observed	213	194	54	52

Table 7.4: Expected numbers of signal and background events after the requirements listed in the first column, as well as the observed numbers of events in the combined $e\mu$ and μe channels using 2012 8 TeV data. The signal is shown for $m_H = 125$ GeV. The W+jets background is estimated entirely from data, whereas MC predictions normalized to data in control regions are used for the WW, $t\bar{t}$, and $tW/tb/tqb$ processes in all the stages of the selection. Contributions from other background sources are taken entirely from MC predictions. In the middle part of the table, the expected numbers of signal and background events, and the observed numbers of events, are shown also in the control regions. For these rows, the W+jets contribution is still taken from the data-driven estimate but no normalization factors are applied, except that the top normalization factor is applied for the top background estimate in the WW control regions. The bottom part of the table lists the number of expected and observed events after the $\Delta\phi_{ll}$ cut separated by the flavor of the subleading lepton. Only statistical uncertainties associated with the number of events in the MC samples are shown.

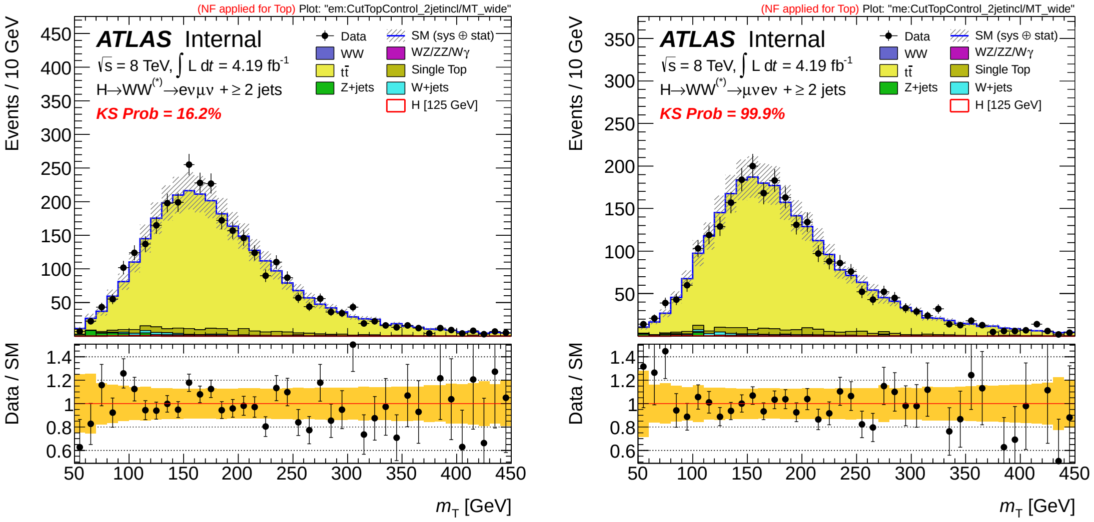


Figure 7.3: m_T distribution after the 2-jet requirement and requiring that at least one of the jets is b-tagged, in the 2012 data. Left: $e\mu$ channel. Right: μe channel.

a signal strength of $\mu = 0.5 \pm 0.7$ at $m_H = 125\,\text{GeV}$. The signal strengths measured with the 7 TeV and 8 TeV analyses separately are compatible within 1.5 standard deviations. Figure 7.6 shows the distribution of the transverse mass m_T after all selection criteria have been applied, summed for the 7 TeV and 8 TeV data, after subtracting the total estimated background. The 0- and 1-jet channels are added and the predicted $m_H = 125\,\text{GeV}$ signal is superimposed. No systematic uncertainties are included.

Figure. 7.7 shows the expected and observed p_0 value and the fitted signal strength for the all jets analyses with 7 TeV and 8 TeV data combined(left). The right plot shows the expected distribution in the presence of a Higgs boson with $m_H = 125\,\text{GeV}$. An excess of events is observed over the expected background, reflected by a low observed p_0 and a fitted μ which deviates from zero. The minimum value observed for p_0, found at $m_H = 125\,\text{GeV}$, is 3×10^{-3}, corresponding to 2.8 standard deviations. The fitted signal strength at $m_H = 125\,\text{GeV}$ is $\mu = 1.4 \pm 0.5$. The expected p_0 for a Higgs with $m_H = 125\,\text{GeV}$ is 0.01, or 2.3 standard deviations, for the combined 7 TeV and 8 TeV data.

Figure. 7.8 shows the two-dimensional likelihood contours for a simultaneous scan of μ and m_H, for this analysis and also for the $H \to ZZ^{(*)} \to 4\ell$ [122] and $H \to \gamma\gamma$ [134] analyses. The lack of mass resolution in the $H \to WW^{(*)} \to l\nu l\nu$ final state for low m_H can be seen clearly in contrast to the $H \to ZZ^{(*)} \to 4\ell$ and $H \to \gamma\gamma$ final states, but the best-fit values of μ and m_H are in reasonable agreement for all three analyses.

Figure. 7.9 shows the observed local p_0 from the combined 7 TeV and 8 TeV results, compared to the one expected in the presence of a signal at $m_H = 125\,\text{GeV}$. The shape and normalization of the p_0 curves as a function of m_H are in agreement.

The 95% CLs limit on $\sigma/\sigma_{\text{SM}}$ is expected to exclude a SM Higgs boson with a mass above 124 GeV with the combined 7 TeV and 8 TeV data in the absence of a

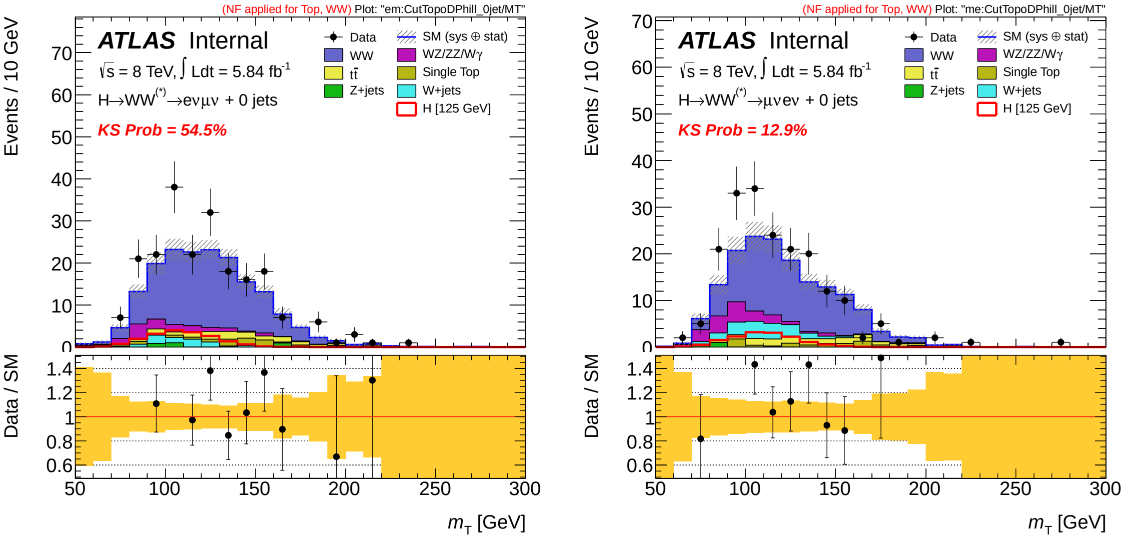


Figure 7.4: Transverse mass m_T distribution in the 0-jet analysis after all selection cuts in the $e\mu$ (left) and μe (right) channels after the unblinding. The expected signal is shown for $m_H = 125\,\text{GeV}$. The hatched band around the sum of backgrounds indicates the total systematic uncertainty in the normalization (but not the shape) of the various components.

signal. However, due to the observed excess of events the observed exclusion CLs lower limit is found at 137 GeV. The CLs limit is shown in Fig. 7.9.

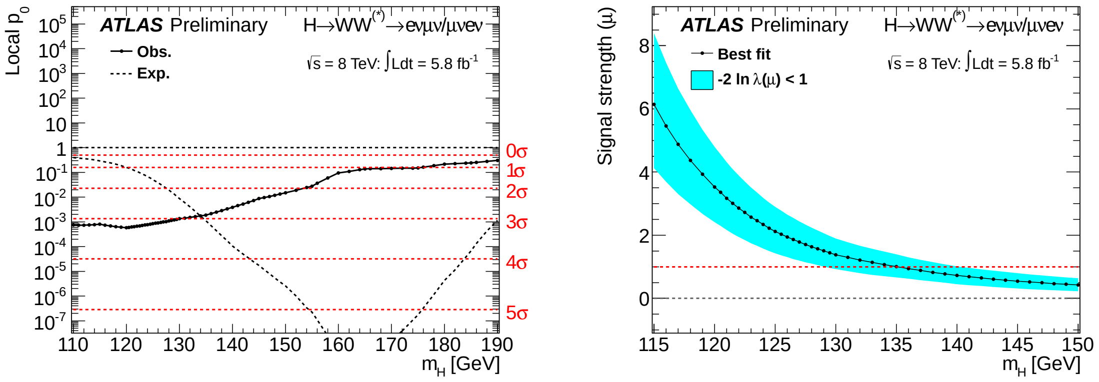


Figure 7.5: 2012 results, using 5.8 fb^{-1} of 8 TeV data. Left: observed (solid line) probability for the background-only scenario as a function of m_H. The dashed line shows the corresponding expectation for the signal + background hypothesis at the given value of m_H. Right: fitted signal strength parameter (μ) as a function of m_H for the low mass range.

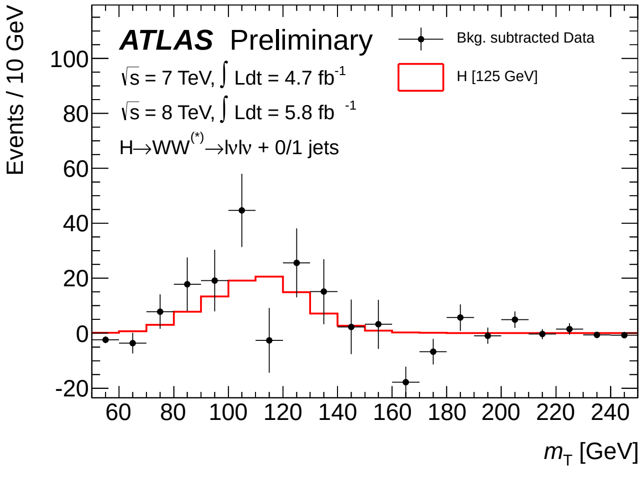


Figure 7.6: Distribution of m_T in data with the estimated background subtracted, overlaid with the predicted signal for $m_H = 125\,\text{GeV}$. The distributions are summed for the 0- and 1-jet analyses and the 7 TeV and 8 TeV data. The statistical errors of both the data and the subtracted background are reflected in the data points. The systematic uncertainty on the background estimate is not included.

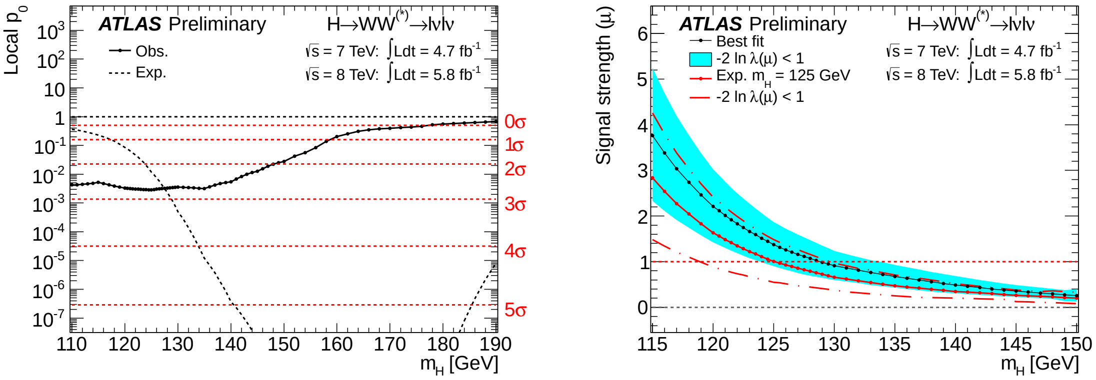


Figure 7.7: Combined 7 TeV and 8 TeV results. Left: observed (solid line) probability for the background-only scenario as a function of m_H. The dashed line shows the corresponding expectation for the signal + background hypothesis at the given value of m_H. Right: fitted signal strength parameter (μ) as a function of m_H for the low mass range (solid black line with cyan band). The expected result for a signal hypothesis of $m_H = 125\,\mathrm{GeV}$ (red line) is included for comparison.

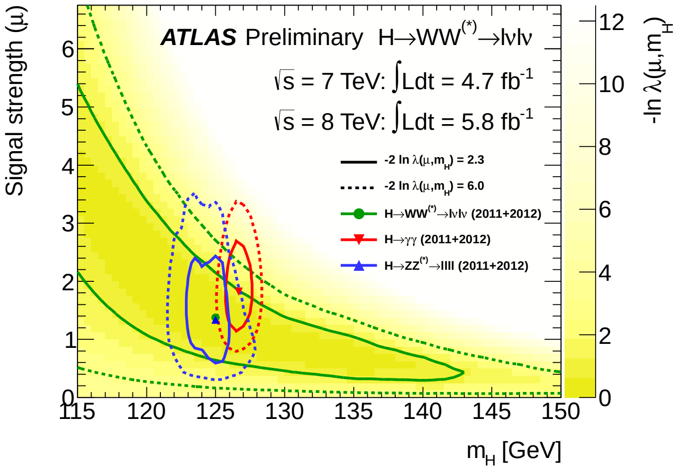


Figure 7.8: Approximate 68% and 95% two-dimensional likelihood ($\lambda(\mu, m_H)$) contours in the best-fit signal strength μ and m_H for the $WW^{(*)} \to l\nu l\nu$, $ZZ^{(*)} \to 4\ell$, and $\gamma\gamma$ analyses using the 2011 and 2012 data [123]. The yellow shading shows the $-\ln\lambda(\mu, m_H)$ values for $H \to WW^{(*)} \to l\nu l\nu$.

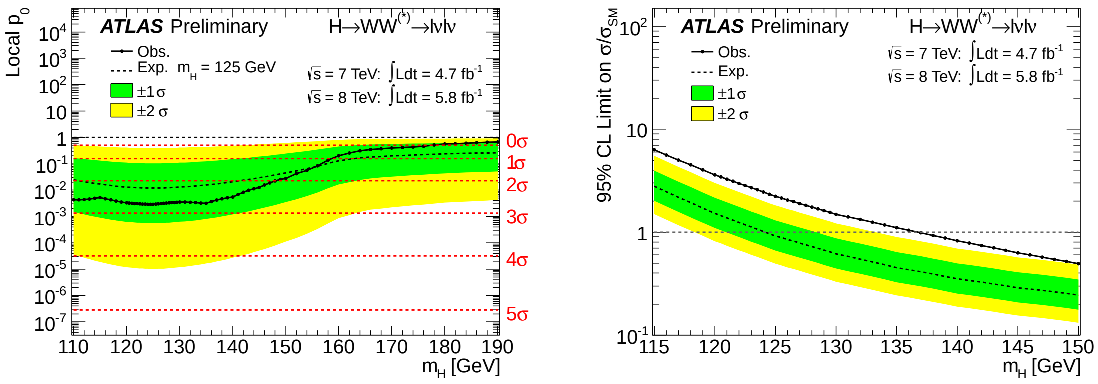


Figure 7.9: Left: Observed (solid line) probability for the background-only scenario, p_0, as a function of m_H, for the combined 7 TeV and 8 TeV data. The dashed line shows the corresponding expectation for the $m_H = 125$ GeV hypothesis. Right: Observed (solid) and expected (dashed) 95% CL upper limits on the cross section, normalized to the SM Higgs boson production cross section and as a function of m_H, over the full mass range considered in the 7 and 8 TeV combined data. Due to the excess of events observed in the low mass signal region, the corresponding mass points cannot be excluded as expected. The results at neighbouring mass points are highly correlated due to the limited mass resolution in this final state. The green and yellow regions indicate the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands on the expected p_0/limit, respectively.

Chapter 8

Drift Time Measurement in the ATLAS Liquid Argon Barrel Electromagnetic Calorimeter

8.1 Introduction

In two recent ATLAS notes [128, 125], the measurement of the drift time of ionization signals in the liquid argon gap in the barrel part of the EM calorimeter and the gap uniformity was performed by using cosmic muon events and splash events in 2008 and 2009, respectively. Similarly, in the summer of 2010 and March 2011, collision events were taken to be studied. The ionization pulses are sampled every 25 ns in 32 samples for the study, while normally in collision events only 5 samples around the peak of pulses are taken to measure the energy of each cell in the EM calorimeter. As described in [128, 125, 136], the drift time and the gap width are obtained from fitting a predicted ionization pulse shape onto data collected with 32 samples. By measuring the drift and gap width, one can access the intrinsic calorimeter uniformity which is directly related to the energy resolution constant term. The ionization pulse shape can be predicted with the Response Transformation Method (RTM) [127] which is available from the entire EM (EMB and EMEC) and is used for the energy reconstruction in ATLAS. An alternative prediction based on the "First Principle" Method (FPM) [128] was developed for the EMB and is used in here with the results reported also in [137]. As shown in [136], both predictions give compatible results.

The barrel part of the EM calorimeter covers the pseudorapidity region $|\eta| < 1.475$ whereas the two end cap parts cover the region $1.375 < |\eta| < 3.2$. The EM liquid argon calorimeter is a highly granular sampling detector with accordion-shaped electrodes and lead absorbers and is fully symmetrical in the transverse plane. Granularities differ from layers in the barrel part of the liquid argon calorimeter [58] (see Tab. 8.1).

The drift time and the gap uniformity are two intrinsic parameters of the calorimeter, their numerical values may be influenced by different methods and data samples, due to the particular structure of the detector and the different direction, multiplicity

Sampling layer	Granularity η	Granularity ϕ	Coverage η
PS	0.025	0.1	$\|\eta\| < 1.52$
Layer 1	0.003	0.1	$\|\eta\| < 1.475$
Layer 2	0.025	0.025	$\|\eta\| < 1.475$
Layer 3	0.05	0.025	$\|\eta\| < 1.475$

Table 8.1: Granularities and η coverage of the 4 layers in the barrel EM calorimeter with PS standing for presampler.

of incoming particles. The cosmic muons preferentially travel through the detector vertically and have in general a small multiplicity per triggered event. In the beam splash event the horizontal incoming particles, which have a very large multiplicity, trigger essentially all cells at the same time. As a result, previous studies of the beam splash data [125] showed evidence of a long range crosstalk effect between layers, which resulted in discrepancies between cosmic muons and beam splash events in layer 3. However, in collision events, incoming particles originating from collision vertices show different phenomena as well, which will be described in Section 8.4. This chapter is organized as follows: Section 8.2 gives the detail of data samples and the fit method. Section 8.3 shows the results of fits to individual summed pulses. In Section 8.4 a crosstalk correction between neighbour cells in pseudorapidity direction in layer 1 is introduced. The drift time and gap uniformity results are shown in Section 8.5. In Section 8.6, an estimate of the sagging effect in the barrel part of the calorimeter based on observed ϕ asymmetries is discussed.

8.2 Data Samples and Theoretical Model

The data were taken using L1Calo trigger in runs 160269 and 177960 in 2010 and 2011, respectively. In the barrel parts, 2 387 415 pulses in 979 364 events in 2010 data and 9 498 501 pulse in 700 224 events in 2011 high luminosity data were retrieved from the corresponding Event Summary Data (ESD) files with 1 GeV energy threshold on the pulses. Meanwhile, to study the crosstalk effect between neighbour cells in layer 1, a small sample of 50 236 events, corresponding to 408 931 pulses in 2010 data, was reconstructed with a lower energy threshold of 200 MeV.

For the analysis, pulses are selected using cuts on the maximum amplitude (S_{max}) and shapes of the pulses, as described in note [128]. Table 8.2 shows the number of pulses selected from the cosmic muons, beam splash and 2010, 2011 collision data, respectively. The 2010 collision data, which are used for fits to both individual and summed pulses, have comparable statistics as the cosmic muon data in layer 2 and more statistics in the presampler and layer 1. The 2011 collision data, which are 11 times larger than the 2010 data, cover essentially all barrel calorimeter region (in particular that in the back layer) and are used only for fits to summed pulses. In Table 8.2 a strong layer dependence is also shown. Layer 2 in collision data has more statistics because more energy is deposited there and layer 1 gets fewer pulses due

Sampling layer / No. pulses	collision 2011	collision 2010	splash	cosmic
PS	2.4 m	183 k	605.8 k	20.5 k
Layer 1	1.6 m	215 k	104.0 k	42.6 k
Layer 2	5.4 m	760 k	2 233.9 k	330.3 k
Layer 3	162 k	18.6 k	558.1 k	78.9 k
Total	9.5 m	1.2 m	3 501.8 k	472.3 k

Table 8.2: Number of selected pulses from 2010 and 2011 collision events in comparison with the corresponding numbers from cosmic muon data and beam splash data.

to the small cells and the energy threshold on S_{max}. Layer 3 is not well covered in collision data because it is hard for electrons and photons to go through all the layers. A comparison among the three data samples in S_{max} distribution is shown in Fig. 8.1 in different layers. The distribution of the maximum amplitude of the collision events looks similar to the one of the cosmic muons because in these two cases the trajectory of incoming particles in the barrel is similar. The two thresholds (100 ADC and 160 ADC counts for $|\eta| > 0.8$ and $|\eta| \leq 0.8$, respectively) on S_{max} in layer 2 come from two different gains of the same energy threshold in different regions. The small sample which includes the cells in layer 1 with a lower energy threshold at 200 MeV is not shown in Fig. 8.1.

A typical measured ionization signal shape is shown in Fig. 8.2 with the prediction. The signal pulse is recorded in 32 samples and the prediction, which is computed by FPM [128], is shown in 800 bins. A χ^2 fit is performed by minimizing the following χ^2 function

$$\chi^2 = \frac{1}{32 - N_p} \sum_{i=1}^{32} \left(\frac{S_i - S_i^{\mathrm{FPM}}}{\sigma} \right)^2 , \tag{8.1}$$

where N_p is the number of free parameters fitted, $S_i(S_i^{\mathrm{FPM}})$ are measured (predicted) signal amplitude in bin i and σ is the noise term, which is taken as the RMS value of the measured energy in pedestal runs and is about 8 ADC counts for the PS and layer 1, 5 ADC counts for layer 2 ($|\eta| \leq 0.8$) and layer 3 and 3.5 ADC counts for layer 2 ($|\eta| > 0.8$).

In the FPM method, the predicted pulse is obtained by convolution in frequency domain of a triangle like current signal with a transfer function of the electronic chain taking into account of path lengths, impendences, skin effect parameters of cables, feedthroughs, the shaper and preamplifier transfer function. Some model parameters are extracted from generated electronic calibration pulses [128].

In the fit four free parameters are determined. They are

- the drift time (T_{drift}) which is directly linked to the length of the undershoot of the pulse shape,
- the shift parameter ($|\delta_{\mathrm{gap}}|$) which affects the rise at the end of the pulse shape and quantifies any asymmetry in size between the two gaps with respect to the

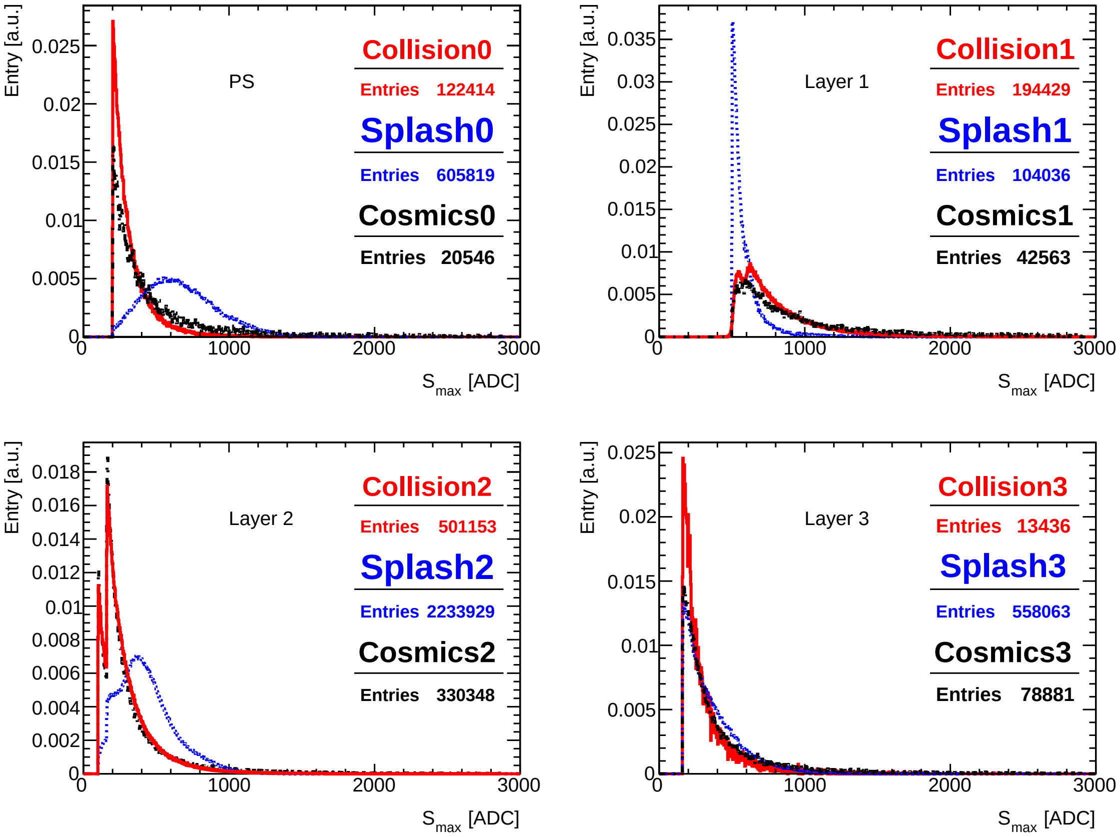


Figure 8.1: Maximum amplitude $S_{\max}$ of the 2010 collision events (red) and beam splash events (blue) and cosmic muon events (black) for four EM barrel sampling layers. The histograms are normalized to unity. Note in layer 2 the two peaks at low $S_{\max}$ correspond to two different gains for a same energy threshold.

nominal one (2.09 mm for layer 2 in the barrel),

- the maximum amplitude in terms of the number of ADC counts ($S_{\max}$) which measures the peak height of the pulse, and
- the timing (t_0) which corresponds to the starting position of the pulse.

It should be pointed out that this is a cell based parameter determination. Each electronic cell is built out of several gaps connected in parallel: for layers 2 and 3, there are 4 double-gaps (one gap on each side of an electrode) in parallel in the barrel and there are four times more gaps per cell in layer 1. The fitted parameters represent thus an average of the local gaps, both in depth along the cell, and in between the gaps forming a cell.

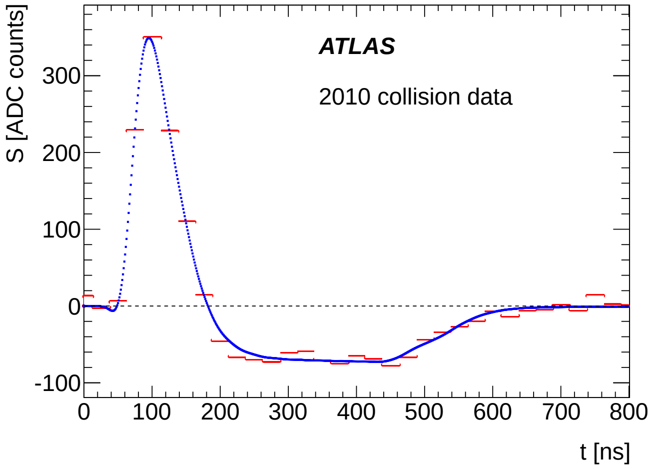


Figure 8.2: A typical pulse shape measured in ATLAS 2010 collision data (red histogram) is shown with the corresponding fit results (blue dots) based on the FPM prediction.

8.3 Extracted Drift Time from Individual Pulses and Summed Pulses

As it was done in the analysis of beam splash and cosmic muon events [125], fits are performed pulse by pulse in 2010 collision data, which are called individual fits. The fitted pulses are required to satisfy the following conditions. Only pulses recorded in nominal voltage (2000 V) are considered (except for cells in the presampler of the 2011 collision data, where the nominal voltage is reduced to 1600 V), the voltage should be symmetric in both sides of the gap and a series quality cuts are also applied on the pulse shapes [128]. What is different from the previous study is that the 2010 collision events have a large fraction of pulses with small amplitude. These pulses are likely to fluctuate in pulse shape so that they may be rejected by the quality requirement. The fluctuation in the tail of the pulse shapes also results in a large uncertainty in the fitted parameters. The fitted drift time (T_{drift}) and gap asymmetry ($|\delta_{\text{gap}}|$) are shown in Figs. 8.3 and 8.4 respectively, in comparison with similar results based on the cosmic muon and beam splash events. In the figure a χ^2_n cut is applied, where χ^2_n is a normalized χ^2 and it differs from that of Eq. 8.1 in that σ^2 is replaced by $\sigma^2 + (kS_{\text{max}})^2$, to take into account the uncertainty of the predicted amplitude. The layer dependent k-factors are chosen to have the values of $0.9\%, 1.1\%, 0.8\%$ and 0.75% [128] for the PS, layers 1, 2 and 3, respectively, such that the χ^2_n is in average independent of S_{max}. Figures 8.3 and 8.4 show large RMS values in both the drift time and the gap asymmetry, due to the fluctuation in the individual pulses.

To reduce the fluctuation, fits to summed pulses (summed fits) are performed. The summed pulses are the average of individual pulses. That means pulses in a same cell are added up event by event directly when they have the same timing. The 32 samples are converted to 800 samples to perform the fit. Figure 8.5 shows the

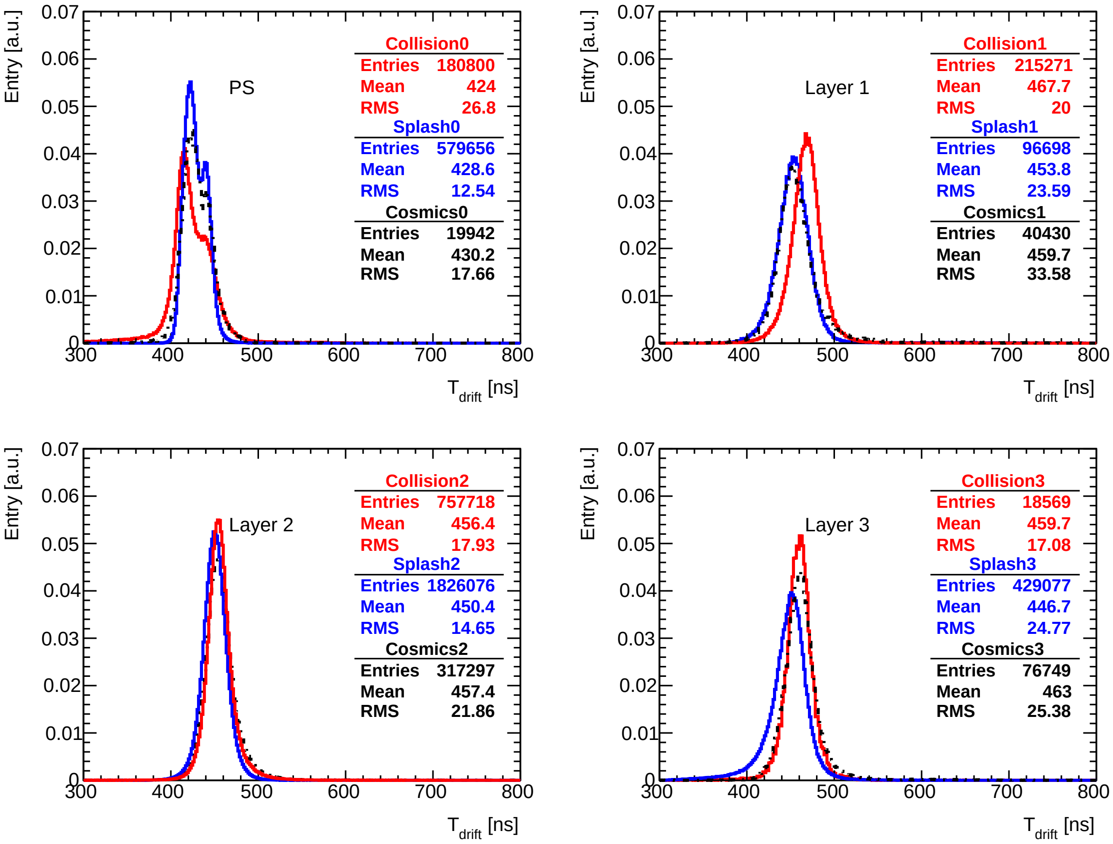


Figure 8.3: Extracted drift time (T_{drift}) from fits to individual pulses in the 2010 collision events, 2009 splash events and 2008 cosmic muon events, for four EM barrel calorimeter layers.

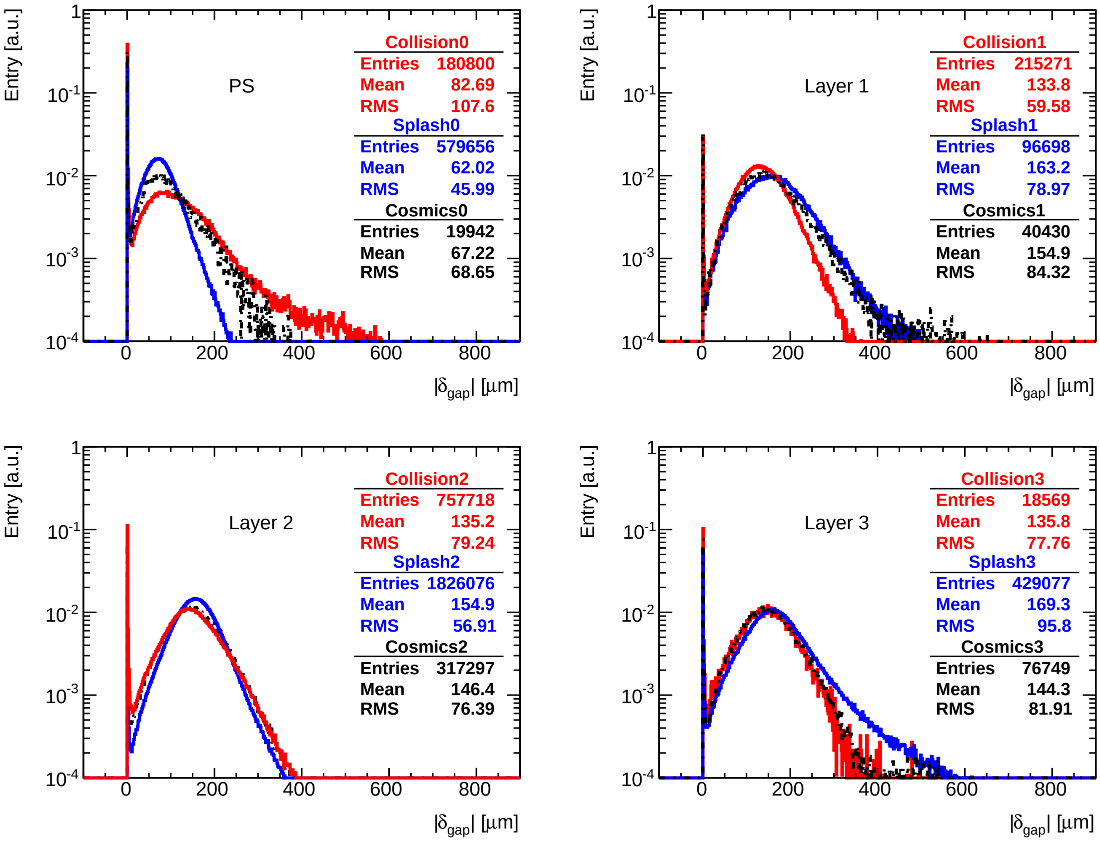


Figure 8.4: Extracted gap asymmetry ($|\delta_{\rm gap}|$) from fits to individual pulses in the 2010 collision events, 2009 splash events and 2008 cosmic muon events for four EM barrel calorimeter layers.

improvement in both the drift time and the gap asymmetry in layer 2, in comparison with the results in Figs. 8.3 and 8.4.

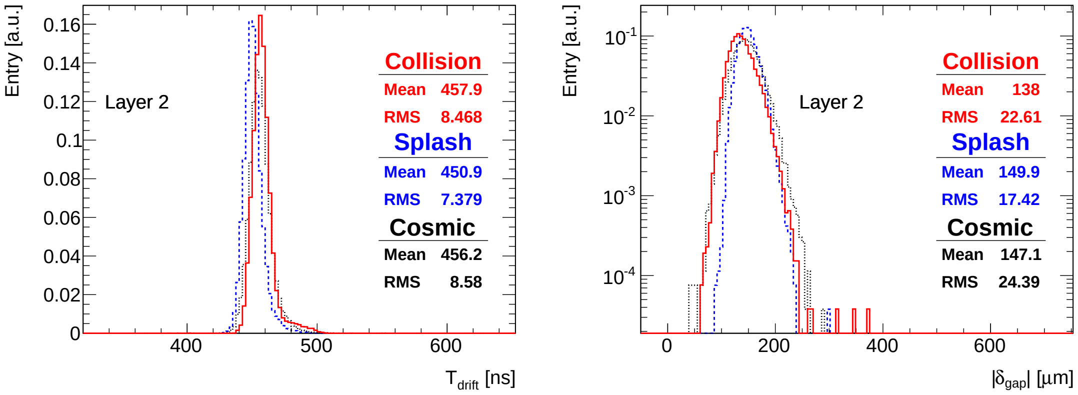


Figure 8.5: Improvement of T_{drift} and $|\delta_{\text{gap}}|$ based on summed fit in layer 2 in 2010 collision data is shown. The RMS values are reduced significantly in comparison with that of the individual fits.

The results of the summed fits are compared among the 2010, 2011 collision, cosmic muon and beam splash data in Figs. 8.6 and 8.7, by requiring more than two events in each cell. In layer 3 of the beam splash data the crosstalk effect is corrected as described in [125]. A global difference of the drift time and the gap asymmetry is observed in layer 1 as a function of the pseudorapidity (though the difference is slightly smaller with the 2011 collision data than with the 2010 collision data). An example is chosen in Fig. 8.8 to show that though there is no more fluctuation on the summed pulses, the tail of the pulse shape of the collision events is about 5 ns longer than the one in the cosmic muon events. It is visible from the difference in the region between 500 and 600 ns. This phenomenon is believed to be due to another type of crosstalk between neighbour cells in layer 1, because in pseudorapidity direction the cell density is 8 times larger than that in layer 2. It shows slightly better performance in 2011 collision data than the 2010 data.

8.4 Crosstalk Correction

As no difference is observed between beam splash and cosmic muon events in layer 1, the discrepancy of the drift time and the gap asymmetry in collision events is believed to be related to the crosstalk effect of neighbour cells due to the high granularity, which is also mentioned in [129]. Indeed, this crosstalk is expected to affect more collision events because of the pointing nature of the showers, which makes them narrower than those of cosmic muon events. For the beam splash events, the (horizontal) signals are about the same in all cells in layer 1 and thus a large net crosstalk effect is not expected. An energy cluster in the EM calorimeter in collision events is defined by a seed cell with the largest $S_{\max}$ and neighbour cells with smaller amplitudes around it. The crosstalk is supposed to occur between the seed and these neighbour cells,

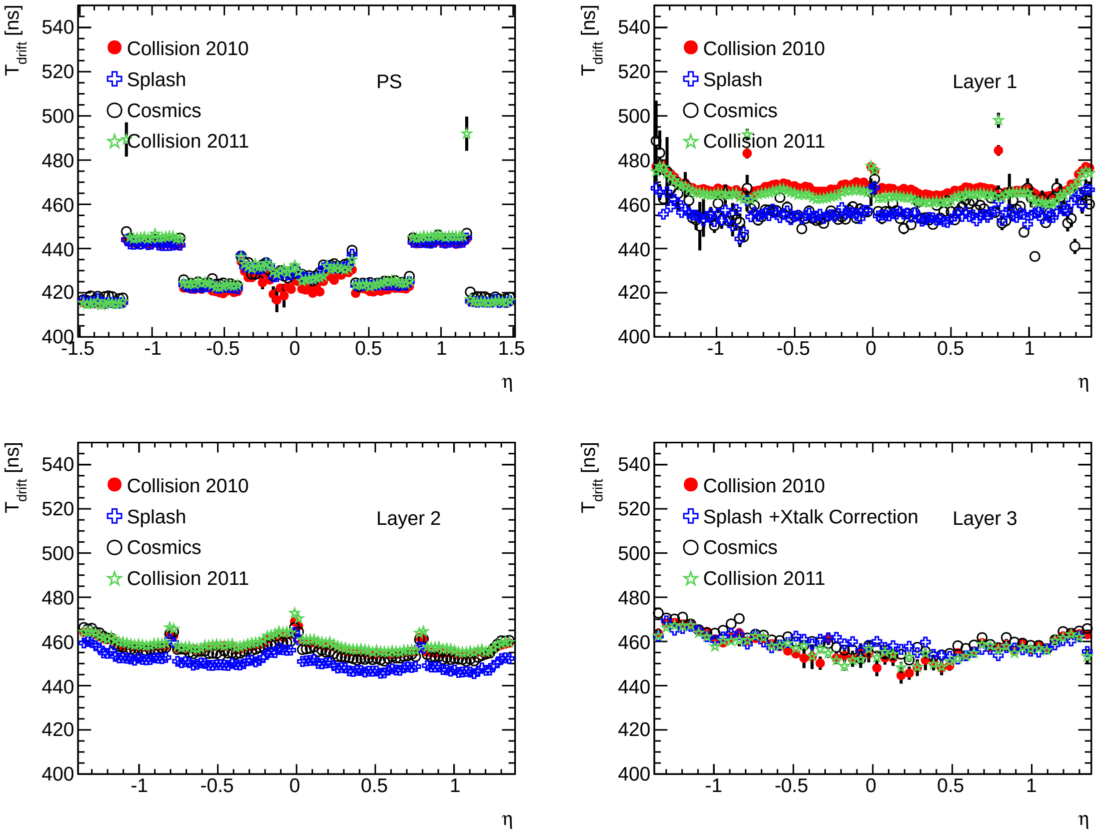


Figure 8.6: Comparison of T_{drift} vs. η based on the summed fits to 2010,2011 collision, 2009 splash and 2008 cosmic events.

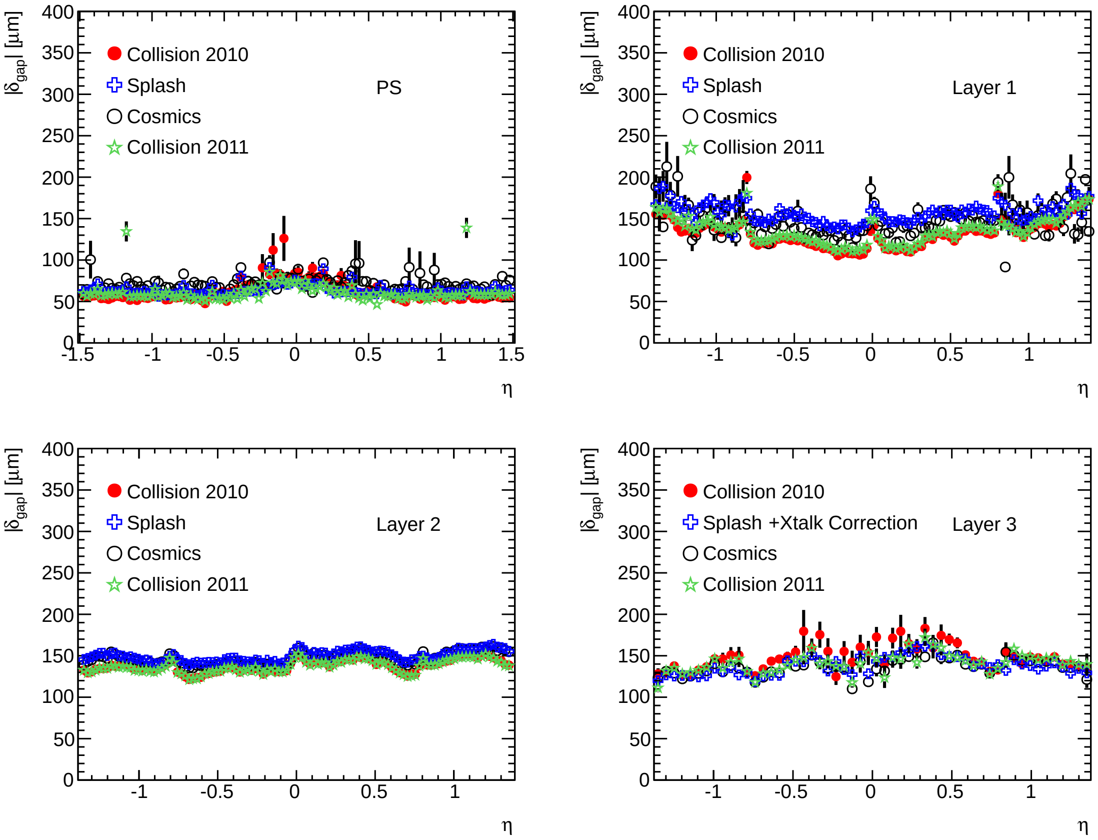


Figure 8.7: Comparison of $|\delta_{\text{gap}}|$ vs. η based on the summed fits to 2010,2011 collision, 2009 splash and 2008 cosmic events.

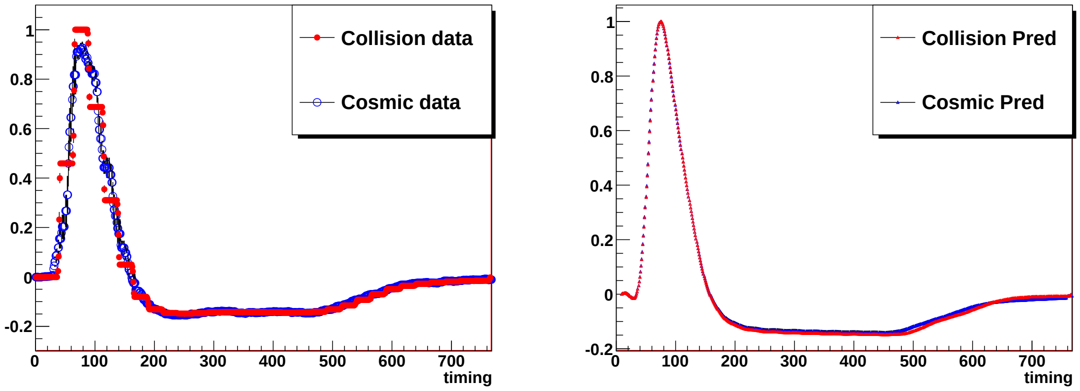


Figure 8.8: Comparison of pulse shapes between collision and cosmic events in $-1.34687 < \eta < -1.34375$ in layer 1 for the measured pulses (left) and the corresponding fits (right). Collision events have a longer tail around 5 ns (compare the region between 500 and 600 ns).

proportional to the difference in the maximum amplitudes. The study is performed using the small threshold sample mentioned earlier. It is assumed according to the discussion in [128] that when the pulses of neighbour cells are added to that of the seed to form a clustered pulse, the crosstalk effect of the clustered pulse is reduced. To verify the assumption, fits to clustered pulses are performed in the same way as the fits to summed pulses. In Fig. 8.9, an example is shown to illustrate that the two neighbour cells in pseudorapidity direction around the seed cell are added together to form a cluster. A total number of 5 509 pulses in layer 1 is fitted and the results are shown in Fig. 8.10. A fairly good agreement with beam splash and cosmic muon events is obtained in comparison with the summed fit. Additional trials with more neighbour cells (8 neighbour cells in all) are made and shown in Fig. 8.11. The differences on the drift time and the gap asymmetry are reduced by including more neighbour cells.

8.5 Results

The results of the drift time T_{drift} and the gap asymmetry $|\delta_{\mathrm{gap}}|$ obtained from the fits to the summed pulses are shown in Table 8.3. The numbers in the brackets are before crosstalk corrections. The projection distribution of the drift time T_{drift} and the gap asymmetry δ_{gap} of the first layer are shown in Fig. 8.12.

The 2010 and 2011 data show consistency in the drift time T_{drift} and the gap asymmetry $|\delta_{\mathrm{gap}}|$. In layer 1, the discrepancies of the drift time T_{drift} and the gap asymmetry $|\delta_{\mathrm{gap}}|$ are reduced by 59.8% and 27.3%, respectively, based on the crosstalk correction. In layer 2 the discrepancy of the gap asymmetry around 6.9% between cosmic and collision is left, as the difference is small and it is very hard to point out the major reason by comparing the pulse shapes. In the 2010 collision events, low statistics in $|\eta| < 0.6$ gives large fluctuation, while the 2011 data shows good

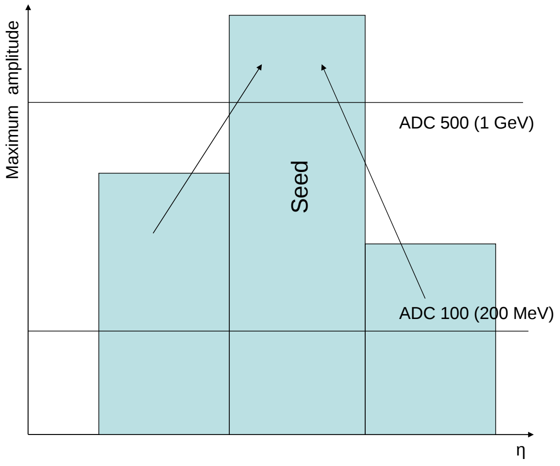


Figure 8.9: Lowering the energy threshold gives the possibility to recover and sum up the low energy neighbour cells with the seed in a cluster.

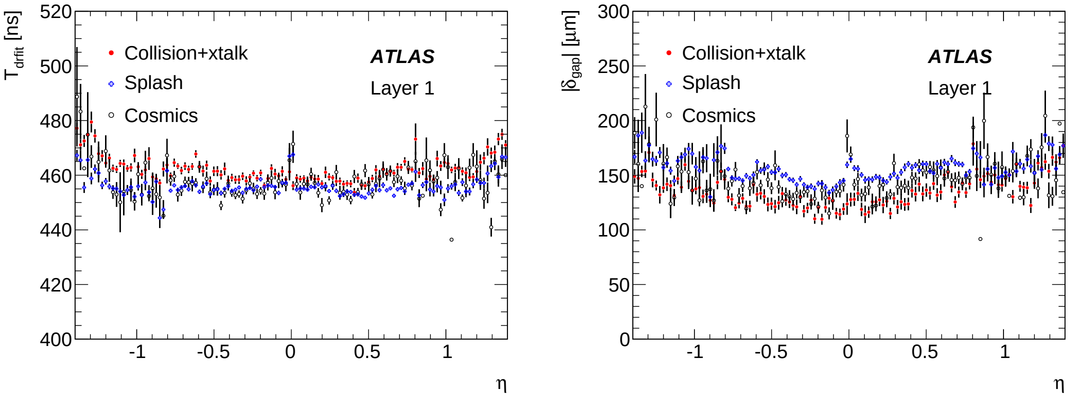


Figure 8.10: Drift time T_{drift} and gap asymmetry $|\delta_{\mathrm{gap}}|$ in layer 1. For the collision data, the results are based on the crosstalk correction with 2 closest neighbour cells.

agreement with cosmic muon events. This is shown in Fig. 8.13 for the $\eta-\phi$ coverage with the weight of T_{drift} in layer 3. A dependence on T_{drift} and ϕ is clearly shown in all data samples studies. Figure 8.14 shows the coverage with the weight of $|\delta_{\mathrm{gap}}|$ in layer 2, the collision events show a slightly lower value than the other two cases.

All in all, the results illustrate that the measurements from three different classes of events (cosmic muons, beam splash and collision events) are in excellent agreement both regarding the average values and the local structures in the $\eta-\phi$ plane, indicating not only the stability of the calorimeter but also the detailed understanding of pulse shapes in widely different conditions.

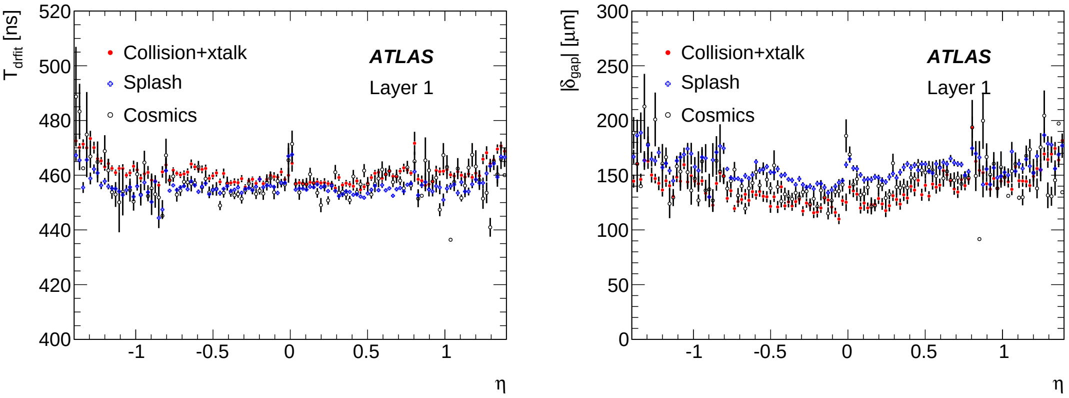


Figure 8.11: Drift time T_{drift} and gap asymmetry $|\delta_{\mathrm{gap}}|$ in layer 1. For the collision data, the results are based on the crosstalk correction with 8 closest neighbour cells.

Layer	Collision 2010	Collision 2011	Cosmic	Splash
		T_{drift}		
PS	426.8 ± 11.9	429.0 ± 12.6	429.7 ± 11.9	428.5 ± 10.1
Layer 1	$460.4(467.8) \pm 11.0(13.2)$	$(465.6) \pm (11.5)$	456.9 ± 15.7	455.5 ± 8.6
Layer 2	457.9 ± 8.5	459.0 ± 8.5	456.2 ± 8.6	450.9 ± 7.4
Layer 3	460.1 ± 14.2	459.0 ± 11.0	460.9 ± 11.5	$459.7(448.4) \pm 8.7(10.2)$
		$\lvert\delta_{\mathrm{gap}}\rvert$		
PS	60.3 ± 26.2	59.6 ± 17.1	69.2 ± 30.9	63.4 ± 8.6
Layer 1	$136.2(131.2) \pm 43.7(34.5)$	$(133.8) \pm (25.3)$	142.6 ± 40.8	152.7 ± 29.7
Layer 2	138.0 ± 22.6	137.4 ± 21.0	147.1 ± 24.4	149.9 ± 17.4
Layer 3	140.9 ± 48.7	138.6 ± 35.2	135.7 ± 34.8	$135.1(156.8) \pm 28.6(32.1)$

Table 8.3: Drift time T_{drift} and gap asymmetry $|\delta_{\mathrm{gap}}|$ and their RMS from 2010, 2011 collision events in comparison with the corresponding numbers from cosmic muon data and beam splash data. Numbers in the brackets are before the crosstalk corrections.

8.6 The ϕ Dependence of Gap Width and the Sagging Effect

The drift time T_{drift} is directly related to the liquid argon gap width w_{gap}

$$T_{\mathrm{drift}} = T^{0}_{\mathrm{drift}} \left(\frac{w_{\mathrm{gap}}}{w^{0}_{\mathrm{gap}}} \right)^{1+\alpha} \tag{8.2}$$

where $T^{0}_{\mathrm{drift}} = 457.9\,\mathrm{ns}$ and $w^{0}_{\mathrm{gap}} = 2.09\,\mathrm{mm}$ are the expected nominal drift time and the gap width, respectively, and $\alpha \simeq 1/3$ [126]. Using the drift time obtained from the fits to summed pulses in all 3 data sets in layer 2, the ϕ dependence of the resulting gap width is shown in Fig. 8.15. It is interesting to note that in each data set, the two ϕ sectors in negative or positive region are consistent but there is a systematic trend between the negative and positive ϕ regions (the difference in w_{gap} between the 3 data sets is discussed below). Namely the gap width is slightly lower in $\phi < 0$ than

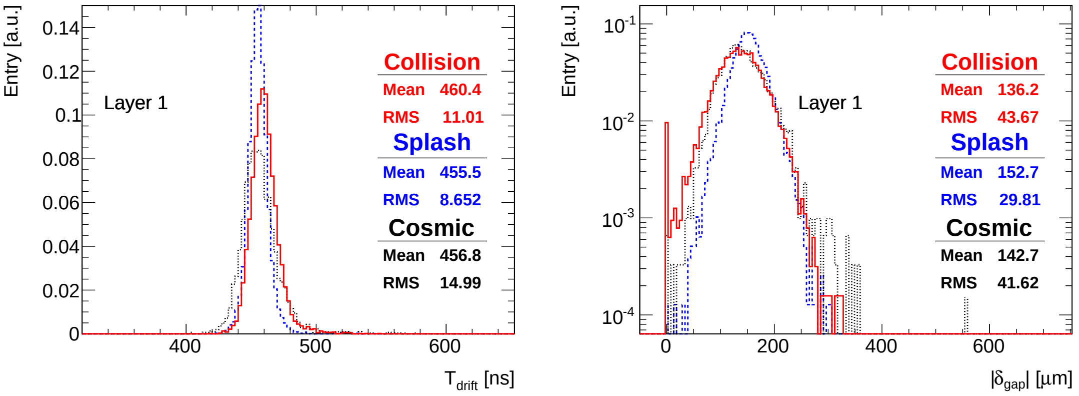


Figure 8.12: Projection distribution of the drift time T_{drift} and the gap asymmetry $|\delta_{\text{gap}}|$ in layer 1. For the collision data, the results are based on the crosstalk correction with 8 closest neighbour cells.

that in $\phi > 0$. This is shown in a more quantitative way in Table 8.4.

Data sample	Collision	Cosmics	Splash
$w_{\text{gap}}(\phi > 0)$ [μm]	2091.5 ± 0.3	2086.7 ± 0.3	2069.0 ± 0.2
$w_{\text{gap}}(\phi < 0)$ [μm]	2088.1 ± 0.2	2081.5 ± 0.2	2063.1 ± 0.2
Relative difference (%)	0.16 ± 0.02	0.25 ± 0.02	0.28 ± 0.01

Table 8.4: Gap width w_{gap} in positive ($\phi > 0$) and negative ($\phi < 0$) regions shown for 3 data sets in layer 2 and their relative difference. The quoted uncertainty is statistical.

The relative difference on the gap width (δw_{gap}) between the upper and lower ϕ region may be originated from the sagging effect. Namely due to the weight of the calorimeter, the effective gap size is modified such that it increases (decreases) with respect to the nominal one in the upper (lower) region. Since each calorimeter pitch contains 2 gaps of 2.09 mm each and there are 256 such pitches in one $\pi/2$ sector, the resulting sagitta effect is thus

$$\delta c = 2 \times 2.09 \times \sqrt{2} \times 256 \times \delta w_{\text{gap}}/2 \tag{8.3}$$

where the scale factor $\sqrt{2}$ accounts for an effective angle between the gap and the ϕ direction. This gives a numerical value of $\delta c = (1.2 \pm 0.1)$ mm, (1.9 ± 0.1) mm and (2.1 ± 0.1) mm based on the collision, cosmic muon and beam splash data, respectively.

The difference between the 3 data sets may be due to different average shower depths in the calorimeter which are smallest for showers coming from the interaction point in the collision data, largest for the beam splash data with many particle showers per event corresponding to an average position around the middle of layer 2 and in between for the cosmic muon data.

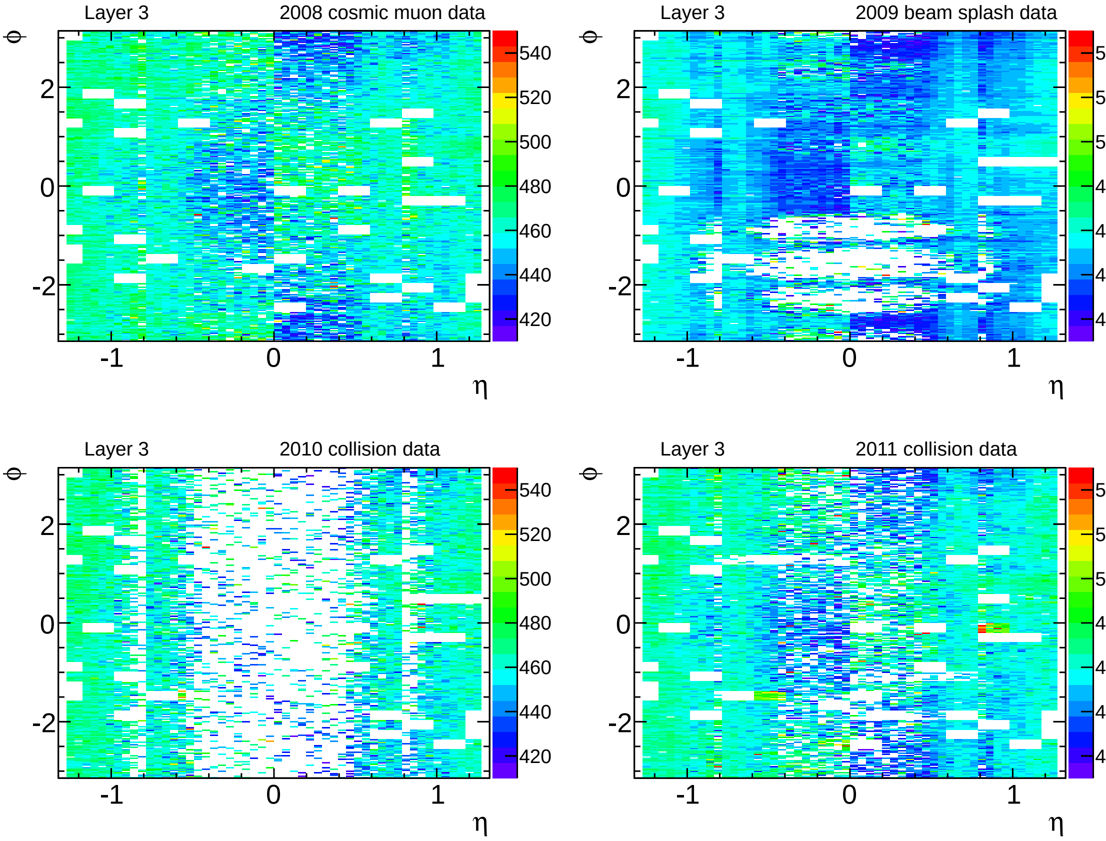


Figure 8.13: The $\eta - \phi$ coverage with the weight of T_{drift} in layer 3. A dependence on T_{drift} and ϕ is clearly shown.

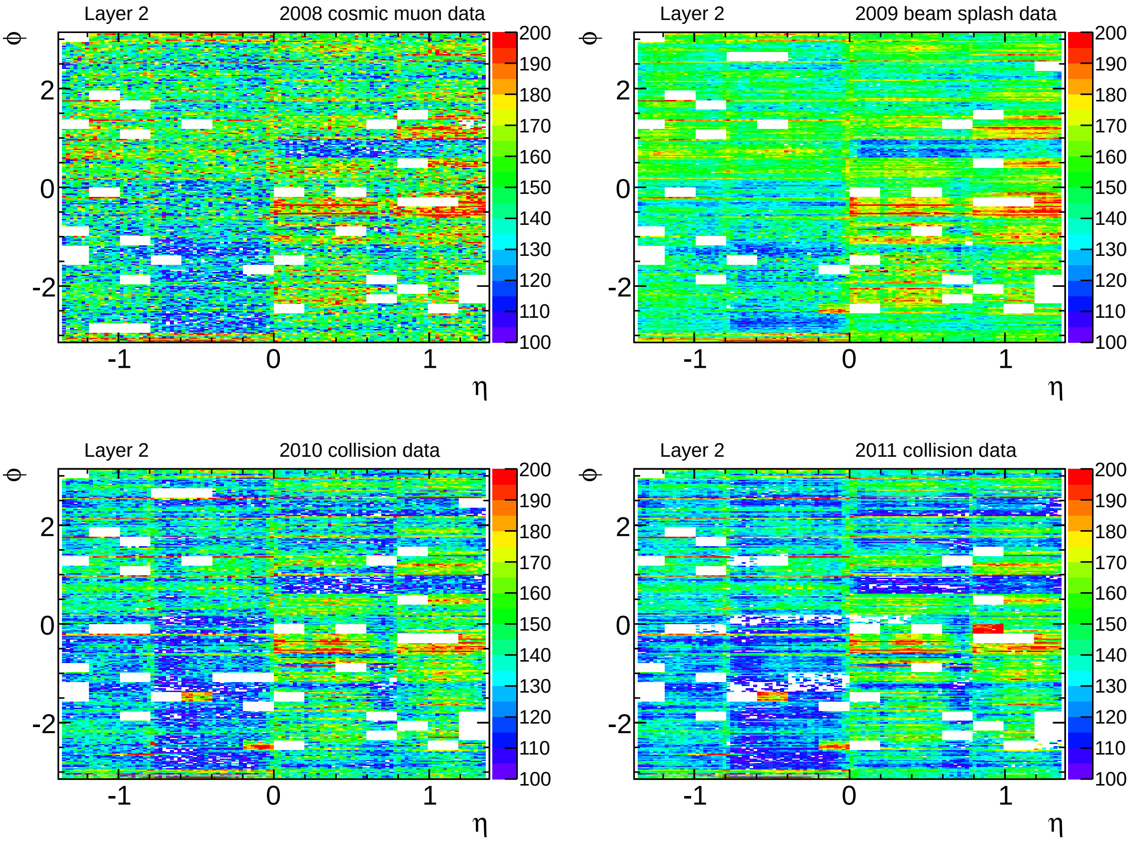


Figure 8.14: The $\eta - \phi$ coverage with the weight of $|\delta_{\text{gap}}|$.

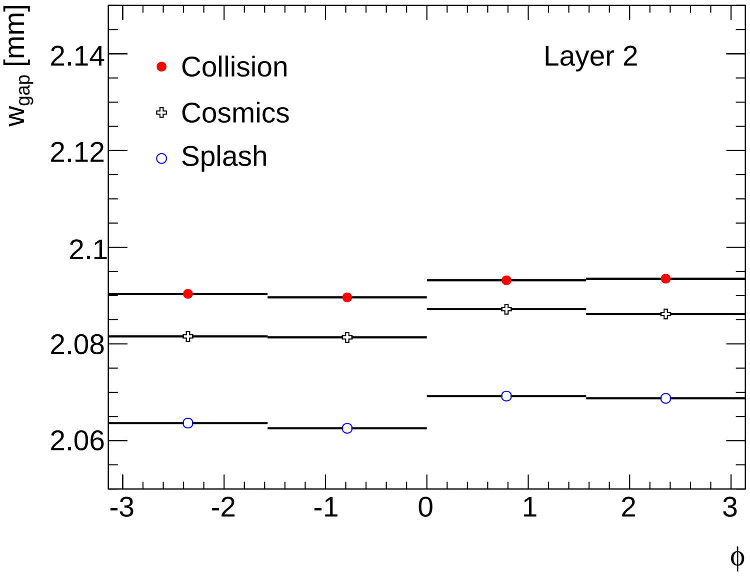


Figure 8.15: Gap width w_{gap} as a function of ϕ for 3 data sets in layer 2.

The drift time determined from fits to summed pulses in the collision data of 2010 and 2011 is further plotted in Fig. 8.16 (upper plot) with finer bins in ϕ corresponding to cell structures in layer 2. Larger drift time (thus larger gap size) is clearly visible (in particular for $\phi > 0$) in transition regions between ϕ modules which are indicated with dashed vertical lines. To reduce the fluctuation, the full ϕ range is subdivided into 4 regions (top: $[45^\circ < \phi < 135^\circ]$, left: $[\phi > 135^\circ$ or $\phi < -135^\circ]$, bottom: $[-135^\circ < \phi < -45^\circ]$ and right: $[-45^\circ < \phi < 45^\circ]$, the 4 modules in each region are then overlaid. The resulting drift time is shown in Fig. 8.16 (four middle plot). The largest deviation in drift time in the transition region (along the dashed line) appear in the top region, this coincides with an independent observation of deficit in electron calorimetry energy measurement from Z decays [138]. The relative difference in drift time between the transition regions and the baseline in the full ϕ range is presented in Fig. 8.16 (lower plot). The maximum relative variation on the drift time is about 3%. This variation, according to the relation $S_{\text{max}} \propto T_{\text{drift}}^{-0.3/1.3}$ [128], gives a corresponding variation of 0.7% on the pulse amplitude S_{max} or the energy measurement. The ϕ dependence seems to follow a cosine behaviour though with limited precision. This shape may eventually be used for relative energy corrections in the ϕ transition regions.

8.7 Conclusion

Collision events, taken in 2010 and 2011, are used to measure the drift time of ionization signals and the gap uniformity of the liquid argon electromagnetic calorimeter. The results are compared with the independent results based on 2009 beam splash events and 2008 cosmic muon events. Fits to summed pulses for any given cell are

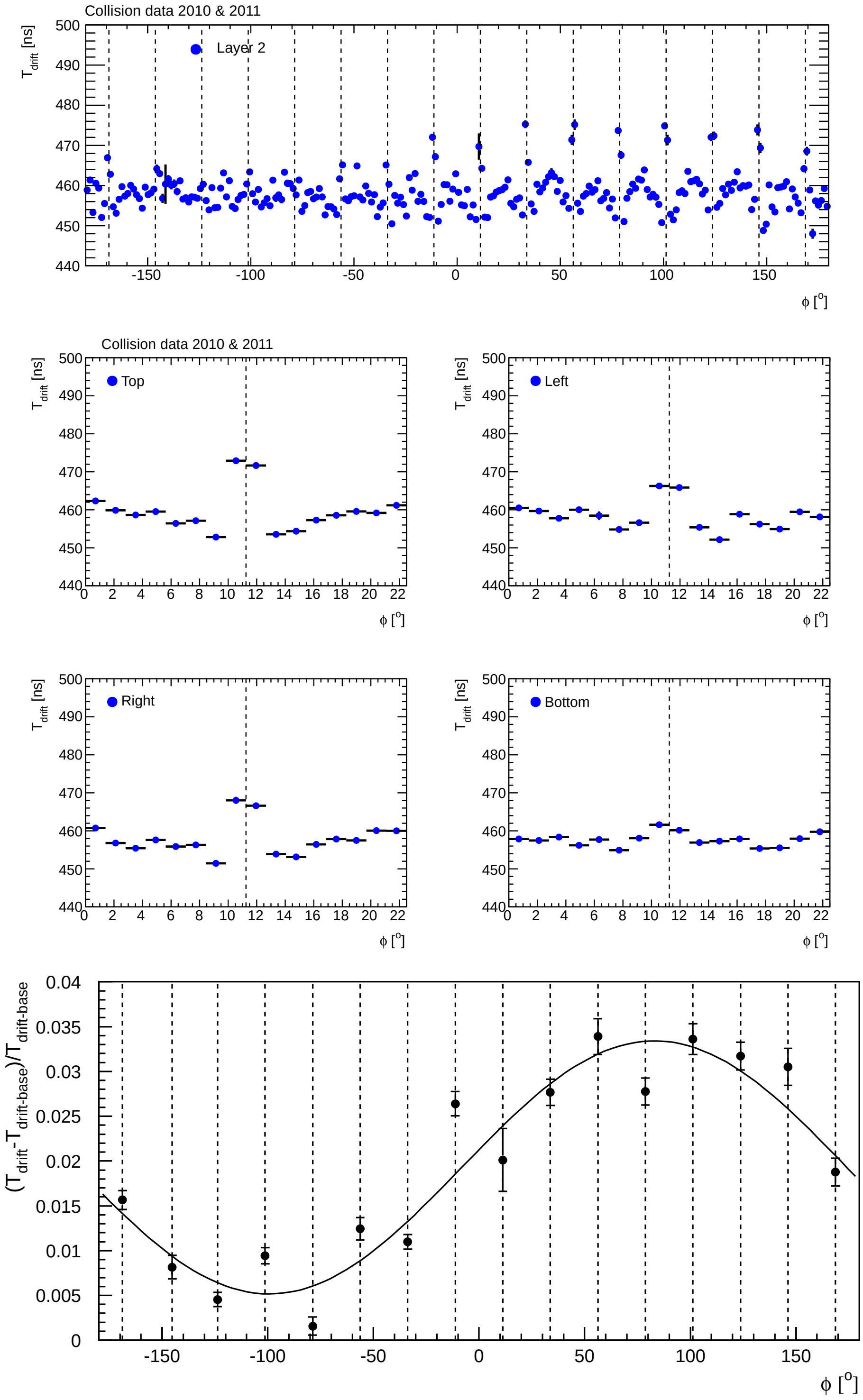


Figure 8.16: Dependence on ϕ of the drift time determined from fits to summed pulses in collision data of 2010 and 2011 in layer 2. The top plot shows the drift time measurement in the full ϕ range. The four middle plots show the average drift time in four ϕ regions (see text). The lower plot shows the relative variation of the drift time in the ϕ transition regions with respect to the baseline measurement outside of the transition regions.

performed to reduce the fluctuation in the pulse shapes. The crosstalk effect in electrical signals between neighbour cells in layer 1 is studied. The crosstalk effect is largely canceled when adding all cells in a cluster. Overall for both the drift time and the gap asymmetry an excellent agreement between data samples is observed, though a small difference in the gap asymmetry in layer 2 between collision events and cosmic muon events still remains, which cannot be easily corrected for without knowing the pulse shape of the crosstalk contribution. In the barrel part of the calorimeter, the sagging effect corresponding to numerical values of $1.2 - 2.1\,\mathrm{mm}$ is observed from the asymmetry measured in cells between the positive and negative ϕ sectors.

Conclusion

A Higgs boson search in the $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$ channel has been performed. The search in the SM uses all available data taken by the ATLAS experiment at the LHC in 2011 and 2012 at $\sqrt{s} = 7\,\mathrm{TeV}$ and $8\,\mathrm{TeV}$, respectively. The corresponding integrated luminosity values are $4.7\,\mathrm{fb}^{-1}$ in 2011 and $5.8\,\mathrm{fb}^{-1}$ in 2012. Due to the increased search sensitivity for a Higgs boson in a four generation model with additional heavy quarks, the analysis in the $WW^{(*)}$ channel based on the early data of 2011 was also combined with those in other channels to search for a fourth generation Higgs boson. In both cases, the $WW^{(*)}$ channel has the dominant sensitivity for a large part of the mass range $110 - 600\,\mathrm{GeV}$ covered by all combined channels.

The published analysis in the $WW^{(*)}$ channel is cut based, exploiting a clean final state signature with two isolated leptons and large missing transverse energy $E_{\mathrm{T,rel}}^{\mathrm{miss}}$. The transverse momenta of the leading and subleading leptons are required to be greater than $25\,\mathrm{GeV}$ and $15\,\mathrm{GeV}$, respectively. The larger transverse momentum value for the leading lepton guarantees a very efficient trigger of the event. The lepton can be either an electron or a muon, resulting in thus three flavour combinations: ee, $\mu\mu$ and $e\mu$. More stringent cut on the transverse missing energy in the same flavour channels is imposed to suppress the larger Drell-Yan background contributions there. For this reason, the first result of the 2012 data analysis is limited to the $e\mu$ channel (including μe) which has the dominant sensitivity due mainly to its smaller background contamination. The analyses are further categorized in 0-, 1- and 2-jet channels to cope with different background contributions. The dominant top background events are rejected using the b-jet veto. Several topological cuts (e.g. $\Delta\phi_{ll}$, p_T^{ll}, m_{ll}) are further applied to suppress the remaining Drell-Yan and the SM WW irreducible backgrounds.

In the analysis, several data-driven background methods are used to extract the expected background contamination in data. The W+jets events are extracted using fake rate determined in the QCD jet enriched samples. The Z+jets events are estimated using the ABCD method, extrapolating from the Z mass window to the signal region. The top background estimation in 0-jet bin is made using a data-driven method that we have proposed, which derives both the top normalisation and the jet veto survival probability from data. The method is robust against the jet energy scale and resolution variations which would have been the dominant systematic uncertainties. The systematic uncertainties of the method have been thoroughly studied including several closure tests. The method has been used in a number of ATLAS analyses including the SM WW cross section measurement and searches for supersymmetric particles [132, 133]. For the Higgs search in the $WW^{(*)}$ channel, it has been the baseline method since the very beginning of the data analysis.

To cross check the ABCD method, we have proposed an alternative data-driven method to estimate the Z+jets background contribution by correcting the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the Drell-Yan sample using larger statistics sample of W+jets events. The method exploits the fake $E_{\mathrm{T}}^{\mathrm{miss}}$ shape in the W+jets sample after subtracting the neutrino transverse momentum. The method has been tested in two different sets of MC samples used for the 2011 data analysis. The method works independently of larger or smaller $E_{\mathrm{T}}^{\mathrm{miss}}$ mismodelling. The results derived using this method are consistent with that of the ABCD method and have the advantage in that the estimation is more precise and the normalisation factor is not affected by further topological cuts.

Based on the enhanced Higgs production cross section and modified decay branching fractions that we evaluated for a fourth generation scenario, the search excludes such a fourth generation Higgs boson essentially in the full mass range $120-600$ GeV at 95% CL.

In 2011, the SM Higgs boson is excluded in the $H \to WW^{(*)} \to l\nu l\nu$ channel from 133 GeV to 261 GeV at 95% CL while the expected exclusion limits are $127-234$ GeV. In 2012, a significant excess of events is observed over the expected background. Combining 2011 and 2012 data, the minimum observed p_0 value, found at $m_H = 125$ GeV, is 3×10^{-3}, corresponding to 2.8 standard deviations (3.1σ in 2012 alone). The expected p_0 for a Higgs at the same mass is 0.01 or 2.3σ. The fitted signal strength at $m_H = 125$ GeV is $\mu = 1.4 \pm 0.5$.

Independent of the Higgs searches, an additional study is performed on the electron drift time in the liquid Argon (LAr) calorimeter in the ATLAS detector using dedicated data samples taken in 2008, 2009, 2010 and 2011 from cosmic rays, beam splash and beam collision events. The first principle method is used to analyze over more than 5 million single pulse shapes taken in each cell of the LAr calorimeter barrel part. Fits both to individual pulses and to summed pulses over all events are performed. The result shows no significant deviation on gap width in each cell in the calorimeter and the sagging effect due to the gravity effect on the calorimeter is observed.

Together with the other decay channels in particular the $\gamma\gamma$ and $ZZ^{(*)} \to 4l$ channels, both ATLAS and CMS experiments have observed a SM-like Higgs boson at mass around 125 GeV. Now the new focus is moved to measure the properties of the newly discovered particle to check whether it is indeed the SM Higgs boson. For this purpose, high statistical data sample is needed. The LHC is continuing the 8 TeV running till the end of 2012, extending by two more months of data taking originally foreseen. More than 15 fb^{-1} additional integrated luminosity is expected. In parallel, the analysis is being further improved by using multivariate techniques for object and event selections and by enlarging the phase space of the signal region e.g. to lower transverse momentum of 10 GeV for the subleading leptons. If the new particle is confirmed to be the SM Higgs boson, the SM will be complete. However, the existence of the Higgs boson in turn introduces the problem of gauge hierarchy, namely the fact that in the presence of new physics the Higgs mass and the weak scale itself are destabilized towards higher energy scale by quantum corrections. It is generally believed that the SM of particle physics is only an effective theory of a more fundamental underlying theory. Therefore search for new physics at the LHC will be the next major goal.

Bibliography

[1] M. Banner et al. (UA2 Collaboration), Phys. Lett. B. **122**, 476 (1983).

[2] P. Bagnaia et al. (UA2 Collaboration), Phys. Lett. B. **129**, 130 (1983).

[3] S. Abachi et al. (D0 Collaboration), Phys. Rev. Lett. **74**, 2632 (1995).

[4] Donald H. Perkins, “Introduction to High Energy Physics”, Cambridge University press 4th Edition.

[5] V.V. Ezhela, S.B. Lugovsky, O.V. Zenin, “Hadronic Part of the Muon g-2 Estimated on the $\sigma_{tot}^{2003}(e^+e^- \rightarrow hadrons)$ Evaluated Data Compilation”, arXiv: hep-ph/0312114. Jan 2004.

[6] L. O’Raifeartaigh, N. Straumann, “Early History of Gauge Theories and Kaluza-Klein Theories, with a Glance at Recent Developments”, arXiv: hep-ph/9810524. Apr 1999.

[7] C. N. Yang, R. Mills, Phys. Rev. **96**, 191 (1954).

[8] S. Glashow, Nucl. Phys. **22**, 579(1961).

[9] S. Weinberg, Phys. Rev. Lett. **19**, 1264(1967).

[10] A. Salam, “Elementary particle theory“, Almqvist and Wiksells, Stockholm, 1968.

[11] M. Gell-Mann, Phys. Lett. **8**, 214(1964).

[12] G. Zweig, CERN-Report 8182/TH40 (1964).

[13] H. Fritzsch, M. Gell-Mann and H. Leutwyler, Phys. Lett. B **47**, 365(1973).

[14] D. Gross and F. Wilczek, Phys. Rev. Lett. **30**, 1343(1973).

[15] The ALEPH Collaboration, The DELPHI Collaboration, The L3 Collaboration, The OPAL Collaboration, The SLD Collaboration, The LEP Electroweak Working Group, The SLD Electroweak and Heavy Flavour Groups, Phys. Rep. **427** 257(5-6) (2006).

[16] K. Nakamura et al. (Particle Data Group), J. Phys. G **37**, 075021 (2010).

[17] S. Dittmaier et al. LHC Higgs Cross Section Working Group, arXiv: hep-ph/1101.0593v3. May 2011.

[18] ATLAS Collaboration, ATL-PHYS-PUB-2011-001.

[19] T. Aaltonen et al. (The CDF Collaboration), "Search for heavy bottom-like quarks decaying to an electron or muon and jets in $p\overline{p}$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$", Phys. Rev. Lett. 106 (2011) 141803, arXiv: hep-ex/1101.5728.

[20] T. Aaltonen et al. (The CDF Collaboration), "Search for heavy top $t' \rightarrow Wq$ in lepton plus jet events", CDF conference note 10110 (2010).

[21] S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), LHC Higgs Cross Section Working Group, CERN-2011-002 (CERN, Geneva, 2011), `arXiv:1101.0593 [hep-ph]`.

[22] X. Ruan, Z. Zhang, "Impact on the Higgs Production Cross Section and Decay Branching Fractions of Heavy Quarks and Leptons in a Fourth Generation Model", arXiv:hep-ex/1105.1634v2, May 2011.

[23] H.-J. He, N. Polonsky and S.-f. Su, Phys. Rev. D **64** (2001) 053004, arXiv:[hep-ph/0102144].

[24] B. Holdom, W.S. Hou, T. Hurth, M.L. Mangano, S. Sultansoy and G. Unel, PMC Phys. A**3** (2009) 4, arXiv:0904.4698 [hep-ph].

[25] O. Eberhardt, A. Lenz and J. Rohrwild, Phys. Rev. D**82** (2010) 095006, arXiv:1005.3505 [hep-ph].

[26] M.S. Chanowitz, Phys. Rev. D**82** (2010) 035018, arXiv:1007.0043 [hep-ph].

[27] R.C. Cotta, J.L. Hewett, A.Ismail, M.-P. Le and T.G. Rizzo, "Higgs properties in the fourth generation MSSM: boosted signals over 3G plan", arXiv:1105.0039 [hep-ph].

[28] ATLAS Collaboration, Phys. Rev. Lett. **108** (2012) 261802, arXiv:1202.3076v2 [hep-ex].

[29] ATLAS Collaboration, Phys. Rev. Lett. **109** (2012) 032001, arXiv:1202.6540v2 [hep-ex].

[30] CMS Collaboration, "Search for heavy, top-like quark pair production in the dilepton final state in pp collisions at $\sqrt{s} = 7\,\text{TeV}$", arXiv:1203.5410v1 [hep-ex], CMS-EXO-11-050, CERN-PH-EP-2012-081.

[31] CMS Collaboration, "Search for heavy bottom-like quarks in 4.9 inverse femtobarns of pp collisions at $\sqrt{s} = 7\,\text{TeV}$", arXiv:1204.1088v1 [hep-ex], CMS-EXO-11-036, CERN-PH-EP-2012-100.

[32] G.D. Kribs, T. Plehn, M. Spannowsky and T.M.P. Tait, Phys. Rev. D**76** (2007) 075016, arXiv:0706.3718 [hep-ph].

[33] Vernon D. Barger and Roger J.N. Phillips, “Collider Physics (updated edition)”, Addison-Wesley Publishing Company, Inc. (1997).

[34] ATLAS Collaboration, Eur.Phys.J.C **71** (2011) 1728, arXiv:1106.2748v3 [hep-ex].

[35] ATLAS Collaboration, ATLAS-CONF-2011-135.

[36] Low mass scenario: `http://users.lal.in2p3.fr/zhangzq/atlas/4th/sm4-md-400GeV.txt.`

[37] B. W. Lee, C. Quigg and H. B. Thacker, Phys. Rev. D **16**, 1519(1977).

[38] T. Cheng, E. Eichten and L. Li, Phys. Rev. D **9**, 2259(1974).

[39] B. Pendleton and G. Ross, Phys. Lett. B **98**, 291(1981).

[40] C. Hill, Phys. Rev. D **24**, 691(1981).

[41] G. Abbiendi et al, “Search for the Standard Model Higgs Boson at LEP”, hep-ex/0306033, Jun 2003.

[42] The Tevatron Electroweak Working Group for the CDF and the DØ Collaborations, “Updated Combination of CDF and DØ results for the Mass of the W Boson”, Eprint arXiv:0908.1374 [hep-ex], 2009.

[43] The Tevatron Electroweak Working Group for the CDF and the DØ Collaborations, “Combination of CDF and DØ Results on the Width of the W boson”, Eprint arXiv:1003.2826 [hep-ex], 2010.

[44] The Tevatron Electroweak Working Group for the CDF and the DØ Collaborations, “Combination of CDF and DØ Results on the Mass of the Top Quark”, Eprint arXiv:1007:3178 [hep-ex], 2010.

[45] The CDF Collaboration, the D0 Collaboration, the Tevatron New Phenomena, Higgs Working Group, “Combined CDF and DØ Upper Limits on Standard Model Higgs Boson Production with up to 8.2 fb^{-1} of Data”, arXiv:1207.0449v2, Jul 2012.

[46] The CDF Collaboration, the D0 Collaboration, the Tevatron New Phenomena, Higgs Working Group, “Combined CDF and DØ Upper Limits on Standard Model Higgs Boson Production with up to 8.2 fb^{-1} of Data”, arXiv:1108.3331v2, Aug 2011.

[47] Gfitter group, “A Generic Fitter Project for HEP Model Testing”, “http://gfitter.desy.de/”.

[48] E. Lyndon (ed.) and B. Philip (ed.)", "LHC Machine", J. Instrum. **3**, S08001 (2008).

[49] ALICE Collaboration, "The ALICE experiment at the CERN LHC", J. Instrum. **3**, S08002 (2008).

[50] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", J. Instrum. **3**, S08003 (2008).

[51] CMS Collaboration, "The CMS experiment at the CERN LHC", J. Instrum. **3**, S08004 (2008).

[52] LHCb Collaboration, "The LHCb Detector at the LHC", J. Instrum. **3**, S08005 (2008).

[53] LHCf Collaboration, "The LHCf detector at the CERN Large Hadron Collider", J. Instrum. **3**, S08006 (2008).

[54] TOTEM Collaboration, "The TOTEM experiment at the CERN Large Hadron Collider", J. Instrum. **3**, S08007 (2008).

[55] A. Yamamoto et al, "The ATLAS central solenoid", Nucl. Instrum. Meth. A **584**, 53 (2008).

[56] A. Foussat et al, "Assembly concept and technology of the ATLAS barrel toroid", IEEE T. Appl. Supercond **16**, 565 (2006).

[57] D. E. Baynham et al, "Engineering status of the end cap toroid magnets for the ATLAS experiment at LHC", IEEE T. Appl. Supercond **10**, 357 (2000).

[58] ATLAS Collaboration, "ATLAS inner detector : Technical Design Report, 1", ATLAS-TDR-004,CERN-LHCC-97-016, CERN, Geneva, 1997.

[59] ATLAS Collaboration, "ATLAS inner detector : Technical Design Report, 2", ATLAS-TDR-005,CERN-LHCC-97-017, CERN, Geneva, 1997.

[60] S. Haywood et al, Nucl. Instrum. A **408** 242 (1998).

[61] ATLAS Collaboration, New J. Phys. **13** (2011) 053033.

[62] ATLAS Collaboration, "Vertexing Performance Data vs MC comparison for LPCC" ATL-COM-PHYS-2011-1312, CERN, Geneva, 2011.

[63] ATLAS Collaboration, "ATLAS liquid-argon calorimeter : Technical Design Report", ATLAS-TDR-002,CERN-LHCC-96-041, CERN, Geneva, 1996.

[64] ATLAS Collaboration, "ATLAS tile calorimeter : Technical Design Report", ATLAS-TDR-003,CERN-LHCC-96-042, CERN, Geneva, 1996.

[65] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", JINST 3 S08003 (2008).

[66] ATLAS Collaboration, “ATLAS muon spectrometer : Technical Design Report”, ATLAS-TDR-010,CERN-LHCC-97-022, CERN, Geneva, 1997.

[67] M. Deile et al, “Performance of the ATLAS precision muon chambers under LHC operating conditions”, Nucl. Instrum. Meth. A **518**, 65 (2004).

[68] ATLAS Collaboration, ATLAS-CONF-2011-046, CERN, Geneva, 2011.

[69] ATLAS Collaboration, “ATLAS Trigger Performance : Status Report”, CERN-LHCC-98-015, CERN, Geneva, 1998.

[70] ATLAS Collaboration, Eur. Phys. J. **C71**, 1630, (2011).

[71] T. Cornelissen, ATLAS Collaboration, ATL-SOFT-PUB-2007-007, CERN, Geneva, 2007.

[72] ATLAS Collaboration, ATL-LARG-PUB-2008-002, CERN, Geneva, 2008.

[73] ATLAS Collaboration, ATLAS-COM-DAQ-2011-032, CERN, Geneva, 2011.

[74] ATLAS Collaboration, Eur. Phys. J. **C72**, 1909, (2012).

[75] ATLAS Collaboration, ATLAS-CONF-2010-005, CERN, Geneva, 2010.

[76] ATLAS Collaboration, ATL-COM-PHYS-2012-260, CERN, Geneva, 2012.

[77] S. Hassini et al, NIM A, **572**, 77, (2007).

[78] ATLAS Collaboration, ATL-COM-PHYS-2011-1640, CERN, Geneva, 2011.

[79] M. Cacciari, G. P. Salam, G. Soyez, JHEP **04**, 063, (2008), arXiv: hep-ph/0802.1189.

[80] ATLAS Collaboration, “Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7\,\mathrm{TeV}$” CERN-PH-EP-2011-191, arXiv:hep-ex/1112.6426v1.

[81] ATLAS Collaboration, ATLAS-CONF-2011-089, CERN, Geneva, 2011.

[82] ATLAS Collaboration, ATLAS-COM-CONF-2012-105, CERN, Geneva, 2012.

[83] The ATLAS Collaboration, D. Cavalli et al., “The ATLAS TWiki”, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EtMissRefFinal.

[84] S. van der Meer, “Calibration of the ective beam height in the ISR”, CERN-ISR-PO-68-31, 1968.

[85] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG”, JHEP **0904** (2009) 002, arXiv:0812.0578 [hep-ph].

[86] P. Nason and C. Oleari, “NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG”, JHEP **1002** (2010) 037, arXiv:0911.5299 [hep-ph].

[87] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 physics and manual”, JHEP **0605** (2006) 026.

[88] A. Djouadi, J. Kalinowski, and M. Spira, “HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension”, Comput. Phys. Commun. **108** (1998) 56, arXiv:hep-ph/9704448.

[89] M. L. Mangano et al., “ALPGEN, a generator for hard multi-parton processes in hadronic collisions”, JHEP **0307** (2003) 001.

[90] S. Frixione and B. R. Webber, “ Matching NLO QCD computations and parton shower simulations”, JHEP **06** (2002) 029, arXiv:hep-ph/0204244.

[91] H.-L. Lai et al., “New parton distributions for collider physics”, Phys. Rev. D **82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[92] D. de Florian et al., “Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC”, JHEP **11** (2011) 064, arXiv:1109.2109 [hep-ph].

[93] P. M. Nadolsky et al., “Implications of CTEQ global analysis for collider observables”, Phys. Rev. D **78** (2008) 013004.

[94] S. Agostinelli et al., “GEANT 4, A Simulation Toolkit”, Nucl. Instrum. Meth. **A506** (2003) 250.

[95] K. Melnikov and F. Petriello, Phys. Rev. Lett. **96** (2006) 231803.

[96] K. Melnikov and F. Petriello, Phys. Rev. D **74** (2006) 114017.

[97] S. Catani et al., Phys. Rev. Lett. **103** (2009) 082001.

[98] N. Kidonakis, Phys. Lett. D **81** (2010) 054028.

[99] N. Kidonakis, Phys. Rev. D **83** (2011) 091503.

[100] The ATLAS Collaboration, ATL-COM-PHYS-2011-1714.

[101] ATLAS Collaboration, Phys. Rev. Lett. **108** (2012) 111802.

[102] The ATLAS Collaboration, ATLAS-CONF-2012-012.

[103] The ATLAS Collaboration, ATLAS-CONF-2010-056.

[104] The ATLAS Collaboration, ATLAS-CONF-2011-116.

[105] The ATLAS Collaboration, ATLAS-CONF-2012-080.

[106] R. Fruhwirth, Computer Physics Communications **100** (1997) no.1-2, 1-16.

[107] The ATLAS Collaboration, ATL-COM-PHYS-2011-1728.

[108] A. Denner, S. Dittmaier, S. Kallweit and S. Pozzorini, Phys. Rev. Lett. **106** (2011) 052001, arXiv:1012.3975 (2010).

[109] Giuseppe Bevilacqua et al., JHEP 1102, 083 (2011), arXiv:1012.4230v2 [hep-ph].

[110] J. Alwall et al., JHEP 0709, 028 (2007).

[111] B. Mellado, X. Ruan, Z. Zhang, Phys. Rev. D **84**, 096005 (2011).

[112] The ATLAS Collaboration, Eur. Phys. J. C **72** (2012) 2043.

[113] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. C**71** (2011) 1554.

[114] K. Cranmer et al., "HistFactory: A tool for creating statistical models for use with RooFit and RooStats", CERN-OPEN-2012-016.

[115] L. Moneta et al., "The RooStats Project", arXiv:1009.1003 [physics.data-an].

[116] The ATLAS Collaboration, ATLAS-CONF-2012-098.

[117] The ATLAS Collaboration, "Search for the Standard Model Higgs boson in the $H \rightarrow WW^* \rightarrow l\nu l\nu$ decay mode with $4.7\,\mathrm{fb}^-1$ of ATLAS data at $\sqrt{s} = 7\,\mathrm{TeV}$", CERN-PH-EP-2012-126, arXiv:1206.0756v1 [hep-ex].

[118] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", CERN-PH-EP-2012-218, arXiv:1207.7214v1 [hep-ex].

[119] The ATLAS Collaboration, ATLAS-CONF-2011-102.

[120] The ATLAS Collaboration, ATLAS-CONF-2012-043.

[121] A.N. Kolmogorov, "Sulla determinazione empirica di una legge di distribuzione", G. Ist. Ital. Attuari **4**, 83 (1933).

[122] The ATLAS Collaboration, ATLAS-CONF-2012-092.

[123] The ATLAS Collaboration, ATLAS-CONF-2012-093.

[124] P. Bernat, C. Collard, D. Fournier, and M. Kado, "Measurement of the drift time in the ATLAS barrel electromagnetic calorimeter using cosmic muon data - application to the calorimeter uniformity of response", ATL-LARG-PUB-2007-010.

[125] D. Fournier, X. Ruan, and Z. Zhang, Drift Time Measurement in the ATLAS Liquid Argon Barrel Electromagnetic Calorimeter using Beam Splash Events", ATL-COM-LARG-2010-034.

[126] ATLAS Collaboration, "Drift Time Measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using Cosmic Muons", arXiv:1002.4189.

[127] D. Banfi, M. Delmastro, and M. Fanti, "Cell response equalization of the ATLAS electromagnetic calorimeter without the direct knowledge of the ionization signals", JINST 1 (2006) P08002.

[128] C. Collard, D. Fournier, S. Henrot-Versill, and L. Serin, "Prediction of signal amplitude and shape for the ATLAS electromagnetic calorimeter", ATL-LARG-PUB-2007-010.

[129] J. Labb and R. Ishmukhametov, "Crosstalk Measurements in the Electromagnetic Calorimeter during ATLAS Final Installation", ATL-LARG-INT-2009-004.

[130] ATLAS Collaboration, Phys. Rev. Lett. **108** (2012) 111802, arXiv:1112.2577.

[131] ATLAS Collaboration, Phys. Rev. Lett. **107** (2011) 231801, arXiv:1109.3615.

[132] ATLAS Collaboration, "Measurement of the WW cross section in $\sqrt{s} = 7\,\mathrm{TeV}$ pp collisions with ATLAS", Phys. Rev. Lett. **107** (2011) 041802

[133] ATLAS Collaboration, "Measurement of the WW cross section in $\sqrt{s} = 7\,\mathrm{TeV}$ pp collisions with the ATLAS detector and limits on anomalous gauge couplings", Phys. Lett. B **712** (2012) 289-308.

[134] ATLAS Collaboration, ATLAS-CONF-2012-091.

[135] ATLAS Collaboration, ATL-COM-PHYS-2012-861.

[136] C. Gabaldon, L. March, J. del Peso, M. Delmastro, and J. Toth, "Measurement of the drift time in the ATLAS electromagnetic calorimeter using cosmic pulses", ATL-LARG-INT-2009-010 (2009) .

[137] X. Ruan, D. Fournier, S. Jin, Z. Zhang, "Drift Time Measurement in the ATLAS Liquid Argon Barrel Electromagnetic Calorimeter using Collision Events", ATLCOM-LARG-2011-039, ATL-LARG-INT-2012-001 (2012).

[138] N. Lorenzo-Martinez, M. Kado et al., "Inter-module widening effect correction", https://indico.cern.ch/getFile.py/access?contribId=3&sessionId=0&resId=0 &materialId=slides&confId=116227.

Acknowledgments

After two years study in IHEP, Beijing, China and another two and a half years in Orsay, France and CERN, Geneva, Switzerland, plentiful scenes with happiness, hardworking and sometimes a little frustration emerged in my mind. Time goes very fast and it's my turn to graduate. I would like to say that without my supervisors who are leading my way to the brightness or colleagues and friends who helped to solve the problems in my work and life, it is impossible for me to succeed, especially to join in the Higgs search and make contribution to the great discovery.

I would like to thank my several supervisors. Firstly it is Shan Jin, my Chinese supervisor who led me to the entrance of high energy physics and chose the topic of my thesis. His stringent requirements on me made me achieve in all aspects in high energy physics. He is quite talent, with precise sense of physics and plentiful of knowledge. Through discussion with him I formed the overview of high energy physics, from theory to statistics. He was also very patient and kind when he was teaching me. He was very willing to solve every difficulty on my work until I got the full understanding. I would like to thank him for his carefulness in the passed five years. Also, I would like to thank Qun Ouyang, my supervisor in the first two years of my Ph.D experience. He spent a lot of time and effort on my courses in the beginning of my study also he instructed me on my work of electronics, which is the basic requirement of the drift time study when I was in France later.

Also, I will express my respect to my French supervisor, Zhiqing Zhang. Though he is a very busy professor in two different experiments, he still spared every effort to direct my work in France. I had countless discussions with him on my work, strategic or concrete. His long run benefited me a lot, which is very important for a student to avoid inefficiency. He also encouraged me to attend many conferences to show my work to people so that I can have chances to practice my presentation.

Besides, though not my supervisors, I would like to thank Bruce Mellado in University of Wisconsin and Daniel Fournier in LAL. Bruce is a genius researcher who always has the magical approaches to solve very difficult problems. He also sponsored my last half years at CERN to attend the discovery of the new particle. Daniel is an experienced researcher who instructed me in the drift time analysis. He taught me every details in LAr calorimeter and be very patient until I fully understood. I will say thank again to you for your instruction though I am not a formal student of you.

I would like to express many thanks to my colleagues and friends in LAL as well. They helped me so much on my work and life, Jean-Baptiste Blanchard, Henso Abreu, Laurent Duflot, Marumi Kado, Genevieve Gilbert, Sylvie Prandt, Hengne Li, Chenghai Xu and every people that I didn't mention. They showed very generous hospitality and made my life in Orsay the unforgettable experience.

I would like to thank my colleagues and friends in IHEP. Yanping Huang, Hongwei Liu, each time when I was abroad, helped me to cope with things in China. Li Yuan always cooked dinners for me when I was busy with my thesis. Beijiang Liu, as an elder colleague helped me form a global understanding of computing and high energy physics. I would like to thank Liwen Yao, Yin Yang, Tianjue Min, Zhichao Zhan, Yu Bai, Lianyou Shan, Feng Lu, Kaiqi Ding, Huan Ren, Hongbang Liu, Kun Liu, Yao Qin, Lei Xu, Yadi Wang, Cui Li, Haisheng Zhao, Liqing Qin, Jinqing Zhang, Jieshen Yu, Zhenghao Zhang and every people that I didn't mention for their devotion.

I would like to thank Yesenia, Instituto de Fisica Corpuscular, Valencia; Shu Li, USTC, Hefei; Wenbin Qian, Tsinghua University, Beijing and every people at CERN who gave me advices.

Finally, I would like to thank my parents, and say sorry to them that I stayed far away from them for several years.