Search for Higgs bosons in the final state $ZZ^{(*)} \to \ell\ell qq$ with the ATLAS detector at the LHC

by

STYLIANOS ANGELIDAKIS

A dissertation submitted to the Department of Physics in fulfillment of the requirements for the degree

of

Doctor of Philosophy (PhD)

School of Science
National and Kapodistrian
UNIVERSITY OF ATHENS

Athens, January 2016

Search for Higgs bosons in the final state $ZZ^{(*)} \rightarrow \ell\ell qq$ with the ATLAS detector at the LHC

by
Stylianos Angelidakis

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

Abstract

This dissertation presents the search for Higgs-like bosons, in the decay mode $H \rightarrow ZZ \rightarrow \ell^+\ell^- q\bar{q}$, using data collected with the ATLAS detector, at the CERN Large Hadron Collider (LHC). Despite the fact that this channel is mostly sensitive in the Higgs boson mass range above 200 GeV, before the discovery of the Standard Model Higgs boson the search was extended over the mass interval 120-180 GeV. Reasonable sensitivity is observed above 140 GeV. The search uses 4.7 fb^{-1} of proton–proton collision data, at a centre-of-mass energy of 7 TeV.

After the discovery of the Standard Model Higgs boson, the search focused towards the detection of additional, heavy Higgs bosons, in the mass range $200 - 1000$ GeV. Proton–proton collision data, at a centre-of-mass energy of 8 TeV were used, corresponding to an integrated luminosity of 20.3 fb^{-1}. Since no significant excess of events is observed over the Standard Model prediction, upper limits are set, at 95% confidence level (CL), on the production cross-section of a heavy Higgs boson times the branching ratio of its decay to a pair of Z bosons. The results are also interpreted in the context of the Type-I and Type-II 2-Higgs-Doublet Models (2HDM). Finally, exclusion limits are estimated by combined statistical interpretation of the results of the $H \rightarrow ZZ$: $ZZ \rightarrow \ell^+\ell^- q\bar{q}$, $ZZ \rightarrow \ell\ell\ell\ell$, $ZZ \rightarrow \ell^+\ell^-\nu\bar{\nu}$ and $ZZ \rightarrow \nu\bar{\nu}q\bar{q}$ search channels, over the Higgs boson mass range $140 - 1000$ GeV.

The Advisory Committee:
Associate Professor Dimitrios Fassouliotis (Thesis Supervisor)
Professor Christine Kourkoumelis
Dr. Venetios Polychronakos

Στην μητέρα μου, Μαρία
και
στην ιερή σκιά του πατέρα μου.

This research has been co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALES. Investing in knowledge society through the European Social Fund.

Η παρούσα έρευνα έχει συγχρηματοδοτηθεί από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο - ΕΚΤ) και από εθνικούς πόρους μέσω του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» του Εθνικού Στρατηγικού Πλαισίου Αναφοράς (ΕΣΠΑ) - Ερευνητικό Χρηματοδοτούμενο Έργο: Θαλής. Επένδυση στην κοινωνία της γνώσης μέσω του Ευρωπαϊκού Κοινωνικού Ταμείου.

Acknowledgements

During the course of this research I received guidance, support and encouragement from many people, colleagues and friends, who I would like, given this opportunity, to heartfully thank. Thank you my mentor and supervisor, Professor Dimitris Fassouliotis, for your continuous encouragement and the persistent effort you put into training me in the scientific field. Thank you my advisor and team leader, Professor Christine Kourkoumelis, for your advice and support all these years. I am honored and delighted to have been part of this academic family.

I would like to express my sincere gratitude to my co-advisor, Dr. Venetis Polychronakos, for giving me the opportunity to spend time at CERN where I aqcuired valuable experience by physically working within the scientific community. Most of this work was performed under the guidance of Dr. Konstantinos Nikolopoulos, who has been continuously encouraging and supporting my career until today. For that I owe him a great debt of gratitude. Furthermore, I would like to express my appreciation to Professors Nikos Giokaris, Niki Saoulidou and Theodoros Alexopoulos who honor me by constituting my PhD committee. Finally, this dissertation would not have been as complete without the support of my colleague Drs. Kalliopi Iordanidou, Nikos Tsirintanis and Athina Kourkoumeli Charalambidi and Ariadni Antonaki. Many thanks to all!

I am honestly thankful to the team from the Università degli Studi di Roma "La Sapienza", Dr. Marco Rescigno and Dr. Francesco Lo Sterzo (the latter is now member of Academia Sinica), for our excellent cooperation throughout the whole research that embodies this dissertation. Also, many thanks to Dr. Michael Schernau for guiding me through my ATLAS qualification work and my engagement with the Cathode Strip Chambers.

Finally, I wish to express my deepest gratitude to my beloved wife, Sofia, who has been fortifying me with her love and support during the whole course of my postgraduate studies.

In conclusion, I would like to partially dedicate this dissertation to the memory of two dear professors, who unfortunately left us early; George Tzanakos and Pavlos Ioannou. Professor Tzanakos, was the primary responsible for my engagement in high energy physics; this dissertation marks the end of an academic path in which he initiated me. Professor Ioannou took interest in my work and career from the start and lavishly provided his advice and support. Both my professors are truly missed.

"Rearh what you cannot."

— Nikos Kazantzakis, *Report to Greco*

Περίληψη στα Ελληνικά

Το Καθιερωμένο Πρότυπο της Φυσικής Στοιχειωδών Σωματιδίων είναι η επικρατούσα θεωρία του υποατομικού κόσμου. Πρόκειται για μια θεωρία βαθμίδας η οποία εγκολπώνει την ισχυρή πυρηνική, την ασθενή και την ηλεκτρομαγνητική αλληλεπίδραση της φύσης στο πλαίσιο ομάδας συμμετρίας $SU(3) \otimes SU(2) \otimes U(1)$ από όπου οι νόμοι της κινηματικής και δυναμικής των σωματιδίων προκύπτουν μέσω απαιτήσεων συμμετρίας. Η κβαντική χρωμοδυναμική, η θεμελιώδης θεωρία των ισχυρών αλληλεπιδράσεων, συνιστά το $SU(3)$ τμήμα του Καθιερωμένου Προτύπου, με γεννήτορα τον κβαντικό αριθμό χρώμα. Η θεωρία των ασθενών αλληλεπιδράσεων συγκροτεί, μαζί με την κβαντική ηλεκτροδυναμική, την «ηλεκτρασθενή» αλληλεπίδραση η οποία διατυπώνεται ως ομάδα συμμετρίας $SU(2) \otimes U(1)$ με γεννήτορες τους κβαντικούς αριθμούς ασθενούς ισοσπίν και ασθενούς υπερφορτίου.

Η απαίτηση συμμετρίας της Λαγκραντζιανής συνάρτησης σε τοπικούς μετασχηματισμούς βαθμίδας αποκαλύπτει ότι οι δυνάμεις της φύσης διαδίδονται από τα γνωστά σήμερα μποζόνια βαθμίδας. Η ισχυρή πυρηνική δύναμη διαδίδεται από 8 γκλουόνια, η ασθενής από τα (έμμαζα) $\boldsymbol{W}^{\pm}$ και Z μποζόνια, ενώ διαδότης της ηλεκτρομαγνητικής αλληλεπίδρασης είναι το άμαζο φωτόνιο. Παρά την εντυπωσιακή αυτή πρόβλεψη του Καθιερωμένου Προτύπου, η οποία οδήγησε στην παρατήρηση των ασθενών αλληλεπιδράσεων ουδετέρου ρεύματος καθώς και στην ανακάλυψη του Z μποζονίου, είναι αδύνατο να προστεθούν στη Λαγκραντζιανή όροι μάζας των μποζονίων βαθμίδας χωρίς να καταλυθεί η $SU(2) \otimes U(1)$ συμμετρία. Στην προσπάθεια να διαφυλαχτεί η συμμετρία, το πρόβλημα της μάζας αντιμετωπίστηκε από τους Steven Weinberg και Abdus Salam με το μηχανισμό Brout–Englert–Higgs (BEH). Ο μηχανισμός BEH αποδίδει μάζα στα μποζόνια βαθμίδας, μέσω της αλληλεπίδρασης αυτών με το πεδίο BEH, χωρίς να καταλύεται η συμμετρία της Λαγκραντζιανής. Επιπρόσθετα όμως, ο μηχανισμός προβλέπει το μποζόνιο Higgs, η πειραματική ανακάλυψη του οποίου αποτέλεσε κεντρικό ζήτημα της Φυσικής των Στοιχειωδών Σωματιδίων κατά τις τελευταίες δεκαετίες.

Αναζήτηση του μποζονίου Higgs σύμφωνα με το Καθιερωμένο Πρότυπο:

Λόγω του βραχέος χρόνου ζωής του, το μποζόνιο Higgs δεν μπορεί να παρατηρηθεί άμεσα σε έναν ανιχνευτή σωματιδίων. Η ανίχνευσή του είναι εφικτή μόνο μέσα από τα κανάλια διάσπασης στις διάφορες τελικές καταστάσεις. Η αναζήτηση του μποζονίου Higgs με το κανάλι διάσπασης $H \to ZZ \to \ell^+\ell^- q\bar{q}$, χρησιμοποιεί τα δεδομένα του πειράματος ATLAS από την πρώτη περίοδο λειτουργίας του μεγάλου αδρονικού επιταχυντή (LHC), με δέσμες πρωτονίων ενέργειας 3.5 TeV. Λόγω του σχετικά μεγάλου ρυθμού παραγωγής της τελικής κατάστασης σε δύο λεπτόνια και δύο κουάρκ, αλλά και του γεγονότος ότι η μάζα του μποζονίου Higgs μπορεί να ανακατασκευαστεί πλήρως από τα προϊόντα της διάσπασης, το συγκεκριμένο κανάλι είναι ένα από τα σημαντικότερα για την ανίχνευση του μποζονίου, στην περιοχή μαζών $m_H > 150$ GeV.

Η ανάλυση των δεδομένων περιορίζεται στην περιοχή τιμών μάζας $120 < m_H < 180$ GeV και η επιλογή των γεγονότων σκοπεύει στην ανακατασκευή των δύο Z μποζονίων από δύο λεπτόνια (δύο μιόνια ή δύο ηλεκτρόνια) με αντίθετο φορτίο, και δύο αδρονικούς πίδακες με μεγάλη εγκάρσια ορμή. Το υπόβαθρο σε αυτή τη μελέτη προκύπτει κυρίως από παραγωγή Z μποζονίων σε συνδυασμό με αδρονικούς πίδακες (Z+jets), με το Z να διασπάται σε λεπτόνια, καθώς και από διαδικασίες όπου παράγονται ζεύγη *top* κουάρκ ($t\bar{t}$), πολλαπλοί αδρονικοί πίδακες (multijet) ή ζεύγη μποζονίων, κυρίως ZZ. Η εκτίμηση του υποβάθρου από διαδικασίες Z+jets και $t\bar{t}$ λαμβάνεται από προσομοίωση Monte Carlo αλλά ρυθμίζεται με βάση τα δεδομένα. Το υπόβαθρο από διαδικασίες πολλαπλών πιδάκων υπολογίζεται αποκλειστικά από τα δεδομένα, ενώ διαδικασίες με ζεύγη μποζονίων εκτιμώνται αποκλειστικά με προσομοίωση Monte Carlo.

Για τη διαχωρισμό υποψηφίων γεγονότων σήματος από γεγονότα υποβάθρου εφαρμόζονται κριτίρια επιλογής τα οποία απομονώνουν τις περιοχές των δεδομένων που είναι συμβατές με το μελετώμενο κανάλι διάσπασης. Στην περιοχή $120 < m_H < 180$ GeV όπου αναζητείται το μποζόνιο Higgs, τουλάχιστον ένα από τα δύο Z μποζόνια αναμένεται εκτός του κελύφους μάζας. Λόγω του υψηλού υποβάθρου από διαδικασίες $Z(\to \ell\ell)$+jets, η ανάλυση εστιάζει στην περιοχή όπου ένα πραγματικό Z μποζόνιο διασπάται αδρονικά και ένα δυνητικό, με μάζα μικρότερη από τη φυσική μάζα του Z μποζονίου, διασπάται σε λεπτόνια. Στη συνέχεια, απαιτείται η ελλείπουσα εγκάρσια ενέργεια να είναι μικρή, ώστε να απορριφθούν τελικές καταστάσεις με νετρίνα υψηλής ενέργειας, χαρακτηριστικές των διαδικασιών παραγωγής *top* κουάρκ. Τέλος, η ανάλυση κατατάσσει τα υποψήφια γεγονότα σε δύο κατηγορίες· εκείνα που περιλαμβάνουν δύο b πίδακες (δηλαδή πίδακες που φαίνονται να απορρέουν από *bottom* κουάρκ) και εκείνα με 0 ή 1 b πίδακες. Η διάκριση αυτή γίνεται προς εκμετάλλευση του γεγονότος ότι b πίδακες προκύπτουν σπάνια σε διαδικασίες Z+jets.

Τα τελικά συμπεράσματα από την παρούσα μελέτη εξάγονται αναζητώντας κορυφή στα δεδομένα, πάνω από το αναμενόμενο υπόβαθρο, στην κατανομή της μάζας του συστήματος των δύο λεπτονίων και των δύο αδρονικών πιδάκων. Δεδομένου ότι δεν παρατηρείται τέτοια κορυφή, υπολογίζονται άνω όρια αποκλεισμού, σε επίπεδο εμπιστοσύνης 95%, της ενεργού διατομής παραγωγής του μποζονίου Higgs, για τις διάφορες υποθέσεις μάζας που εξετάζονται.

Αναζήτηση πέρα από το Καθιερωμένο Πρότυπο

Με την ανίχνευση του μποζονίου Higgs, η οποία ανακοινώθηκε από τα πειράματα ATLAS και CMS τον Ιούλιο του 2012, ένα από τα σημαντικότερα ερωτήματα που γεννώνται είναι εάν πρόκειται για μοναδικό ή εάν υπάρχουν επιπλέον μποζόνια Higgs, όπως προβλέπουν θεωρητικά πρότυπα που επεκτείνουν το Καθιερωμένο Πρότυπο. Τέτοια πρότυπα αποτελούν ισχυρό κίνητρο για συνέχιση της αναζήτησης σε περιοχές ακόμα μεγαλύτερης μάζας.

Η αναζήτηση με το κανάλι διάσπασης $H \to ZZ \to \ell^+\ell^- q\bar{q}$ καλύπτει την περιοχή μαζών $200 < m_H < 1000$ GeV, χρησιμοποιώντας δεδομένα που συνελλέγησαν με τον ανιχνευτή ATLAS, κατά το έτος 2012, με ενέργεια δεσμών στον LHC 4 TeV. Εξετάζεται το σενάριο ενός βαρέος μποζονίου Higgs, με μικρό πλάτος συντονισμού, καθώς επίσης και μοντέλα που προβλέπουν ζεύγος διπλετών Higgs (2 Higgs Doublet Models). Και σε αυτή την περίπτωση, η ανάλυση επικεντρώνεται στην ανακατασκευή των δύο μποζονίων Z από δύο λεπτόνια (δύο μιόνια ή δύο ηλεκτρόνια) και δύο αδρονικούς πίδακες με μεγάλη εγκάρσια ορμή. Τα δύο ζεύγη αναμένονται να έχουν μάζα κοντά στη μάζα του μποζονίου Z ενώ η ελλείπουσα εγκάρσια ενέργεια πρέπει να είναι μικρή.

Η στατιστική που προσφέρουν τα δεδομένα επιτρέπει τη βελτιστοποίηση της ανάλυσης με διάκριση επιμέρους καναλιών σύμφωνα με το μηχανισμό παραγωγής του μποζονίου Higgs (ggF: συγχώνευση γκουονίων, VBF: συγχώνευση διανυσματικών μποζονίων) αλλά και σύμφωνα με των αριθμό των b αδρονικών πιδάκων στην τελική κατάσταση. Τέλος, για μεγάλες τιμές μάζας $m_H > 700$ GeV προστίθεται ένα επιπλέον κανάλι το οποίο επιτρέπει τον συνυπολογισμό περιπτώσεων όπου τα Z μποζόνια εκπέμπονται με μεγάλη ορμή, με αποτέλεσμα οι 2 αδρονικοί πίδακες να ανακατασκευάζονται ως ένας πίδακας με μεγάλη μάζα. Τα αποτελέσματα της μελέτης εξάγονται με την αναζήτηση κορυφής στα δεδομένα, πάνω από το αναμενόμενο υπόβαθρο, στην κατανομή της μάζας του συστήματος των δύο λεπτονίων και των δύο αδρονικών πιδάκων. Δεδομένου ότι δεν παρατηρείται τέτοια κορυφή, υπολογίζονται άνω όρια αποκλεισμού, σε επίπεδο εμπιστοσύνης 95%, της ενεργού διατομής των μηχανισμών ggF και VBF, καθώς και του παραμετρικού χώρου των 2HDM μοντέλων που εξετάζονται.

Συνδυασμός αποτελεσμάτων

Προκειμένου να προσεγγίσουν την απάντηση στο εάν υπάρχουν περισσότερα μποζόνια Higgs, ομάδες του πειράματος ATLAS που επικεντρώνονται σε διαφορετικές τελικές καταστάσεις της διάσπασης $H \to ZZ$, συνδύασαν τα αποτελέσματά τους για τον υπολογισμό των άνω ορίων αποκλεισμού με μεγάλη στατιστική δύναμη. Οι τελικές καταστάσεις που συμπεριλαμβάνονται είναι οι $ZZ \to \ell^+\ell^- q\bar{q}$, $ZZ \to \ell\ell\ell\ell$, $ZZ \to \ell^+\ell^-\nu\bar{\nu}$ και $ZZ \to \nu\bar{\nu}q\bar{q}$. Εξετάζεται τόσο το σενάριο ενός βαρέος μποζονίου Higgs, με μικρό πλάτος συντονισμού, όσο και 2HDMs, στην περιοχή μαζών $140 < m_H < 1000$ GeV.

Contents

Chapter 1

Introduction

The modern understanding of the subatomic world is embraced by the Standard Model (SM) of particle physics [1]. The SM is a gauge quantum field theory that accommodates the strong, weak and electromagnetic forces of nature in the framework of an $SU(3)\otimes SU(2)\otimes U(1)$ symmetry group [2], from which particle dynamics emerge through symmetry principles, on the basis of Noether's theorem [3]. For instance, symmetry of the Lagrangian under global phase rotations leads to the identification of a conserved current while local phase (gauge) invariance reveals that forces are mediated by gauge bosons which couple to conserved currents.

Quantum chromodynamics, the fundamental theory of strong interactions, constitutes the $SU(3)$ component of the SM. It is generated by the color quantum number which reflects a continuous symmetry of the strong interaction Lagrangian. Invariance under local transformations in the color space shows that the strong force is mediated by eight gluons which couple to color charge. The theory of weak interactions, due to their $V - A$ structure, forms a unified $SU(2) \otimes U(1)$ symmetry group, in conjunction with quantum electrodynamics, within the context of the electroweak (EW) theory. As described in the sections that follow, local gauge invariance dictates that the weak force is mediated by massive $W^{\pm}$ and Z bosons which couple to weak isospin while electromagnetic interactions are mediated by massless photons that couple to electric charge.

The classification of elementary particles in the SM is illustrated in Figure 1.1. Particles are distinguished into bosons (force-mediators of integer spin) and fermions (matter fields of spin-1/2). The twelve flavours of fermions are further classified according to how they interact (or equivalently, by what charges they carry) as quarks (strong, electromagnetic, weak), charged leptons (electromagnetic, weak) and neutrinos (weak). Each fermion has an associated antiparticle with identical mass but inverted internal quantum numbers. The Higgs boson is the latest addition to this list. As pointed out in the following sections, it is a by-product of electroweak symmetry breaking that couples to mass and its discovery has been a major subject of scientific interest in the last decades.

1.1 The Glashow–Weinberg–Salam Model

1.1.1 Electroweak Unification

In 1961 Sheldon Glashow showed that the weak and electromagnetic interactions can be accommodated by the gauge group $SU(2)_L \otimes U(1)_Y$ [4]. "L" denotes left-handed chiral states of fermion fields which form weak-isospin doublets $\chi_L = (\nu_\ell, \ell)_L^\top$, while right-handed components $\psi_R = \ell_R$ transform as singlets in the space of weak isospin. The weak hypercharge (Y) is related to the weak isospin (T) and the electric charge

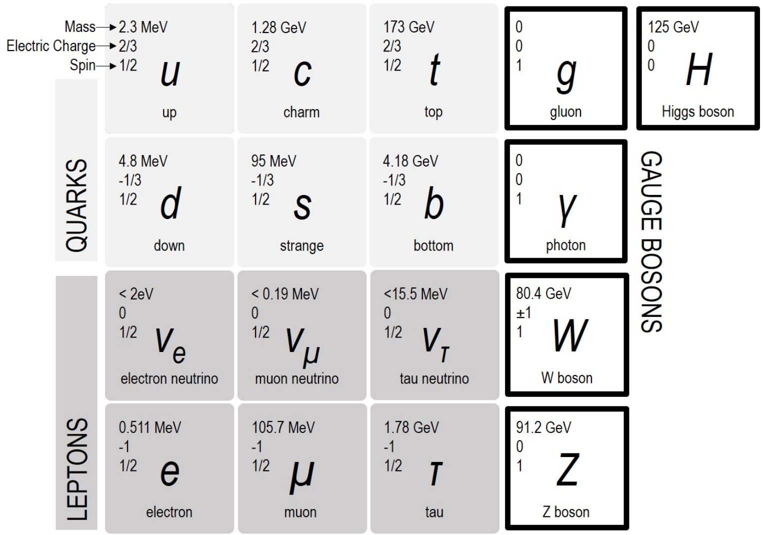


Figure 1.1: Schematic depiction of the SM particles.

(Q) through the Gell–Mann–Nishijima formula:

$$Q = T_3 + \frac{Y}{2}, \tag{1.1}$$

and acts on both left- and right-handed fermion components.

Local gauge invariance is established in the electroweak theory with the coupling of electroweak currents to gauge boson fields. Three vector fields $\boldsymbol{W}_\mu$ are expected to couple to weak isospin currents, with coupling strength g, whereas a neutral vector field B_μ couples to weak hypercharge with coupling strength $g^{'}/2$. These couplings are embedded into the free-fermion (Dirac) Lagrangian by substitution of the normal derivative operator ∂_μ with the covariant derivative:

$$\mathcal{D}_\mu = \partial_\mu + ig\boldsymbol{T} \cdot \boldsymbol{W}_\mu + \frac{ig^{'}}{2} Y B_\mu, \tag{1.2}$$

Terms such as $\mathcal{D}_\mu \phi$ then simply undergo the same transformation as the fields, namely:

$$\mathcal{D}_\mu \chi_L \to e^{i\boldsymbol{\alpha}(x)\cdot\boldsymbol{T}+i\beta(x)Y}\mathcal{D}_\mu \chi_L \tag{1.3}$$

$$\mathcal{D}_\mu \psi_R \to e^{i\beta(x)Y}\mathcal{D}_\mu \psi_R, \tag{1.4}$$

and consequently quantities such as $\bar{\psi}\mathcal{D}_\mu\psi$ are gauge invariant, provided that the gauge fields transform as:

$$\boldsymbol{W}_\mu \rightarrow \boldsymbol{W}_\mu - \frac{1}{g}\partial_\mu \boldsymbol{\alpha} - \boldsymbol{\alpha} \times \boldsymbol{W}_\mu \tag{1.5}$$

$$B_\mu \rightarrow B_\mu - \frac{1}{g^{'}}\partial_\mu \beta. \tag{1.6}$$

The local gauge invariant Lagrangian obtained with the above procedure is:

$$\begin{aligned}\mathcal{L}_{ewk} = \overline{\chi}_L(x) i\gamma^\mu(\partial_\mu + ig\boldsymbol{T}\cdot\boldsymbol{W}_\mu + i\frac{g^{'}}{2}YB_\mu)\chi_L(x) + \\ \overline{\psi}_R(x)i\gamma^\mu(\partial_\mu + i\frac{g^{'}}{2}YB_\mu)\psi_R(x) + \\ \frac{1}{4}\boldsymbol{W}_{\mu\nu}\boldsymbol{W}^{\mu\nu} + \frac{1}{4}B_{\mu\nu}B^{\mu\nu}.\end{aligned} \tag{1.7}$$

The last two terms are gauge invariant and describe the kinetic energy of the gauge fields, $\boldsymbol{W}_{\mu\nu} \equiv \partial_\mu \boldsymbol{W}_\nu - \partial_\nu \boldsymbol{W}_\mu - g\boldsymbol{W}_\mu \times \boldsymbol{W}_\nu$ and $B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu$. Finally, the Lagrangian (1.7) describes the case of a massless charged lepton. The lepton mass term has been deliberately omitted, and so are terms corresponding to the masses of the gauge bosons, since they would ruin local gauge invariance. This inability of the electroweak theory to describe massive bosons or fermions was lifted a few years later with the brilliant contributions of Steven Weinberg and Abdus Salam.

1.1.2 The Brout–Englert–Higgs Mechanism

The electroweak theory discussed in the previous section, cannot account for massive bosons or fermions as observed in nature. Terms describing massive bosons, of the form $\frac{1}{2}mA_\mu A^\mu$, are not gauge invariant and neither are massive fermion expressions, such as $m_e\overline{e}e$, since they mix left- and right-handed fermion components. This obstacle was overcome by Steven Weinberg [5] and Abdus Salam [6], who attributed the masses of gauge bosons and fermions to the Brout–Englert–Higgs (BEH) mechanism [7–9]. The simplest description of the mechanism introduces a complex scalar field, the “BEH field”, which acquires a non-zero vacuum expectation value through Spontaneous Symmetry Breaking. Mass terms are then revealed as couplings to the BEH field, in a way that respects the symmetry of the Lagrangian.

The SM assumes the simplest realization of a BEH field; a complex doublet of scalar fields:

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{1.8}$$

which transforms as an $SU(2)_L$ spinor and therefore must have weak hypercharge $Y = 1$, due to Equation (1.1). By definition, the BEH field is governed by the Lagrangian:

$$\mathcal{L}_{BEH} = (\mathcal{D}_\mu\phi)^\dagger(\mathcal{D}^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2, \tag{1.9}$$

where $V = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$ describes self-interaction of the field. For $\mu^2 < 0$ and $\lambda > 0$, the potential exhibits infinite, non-zero degenerate minima which lie symmetrically

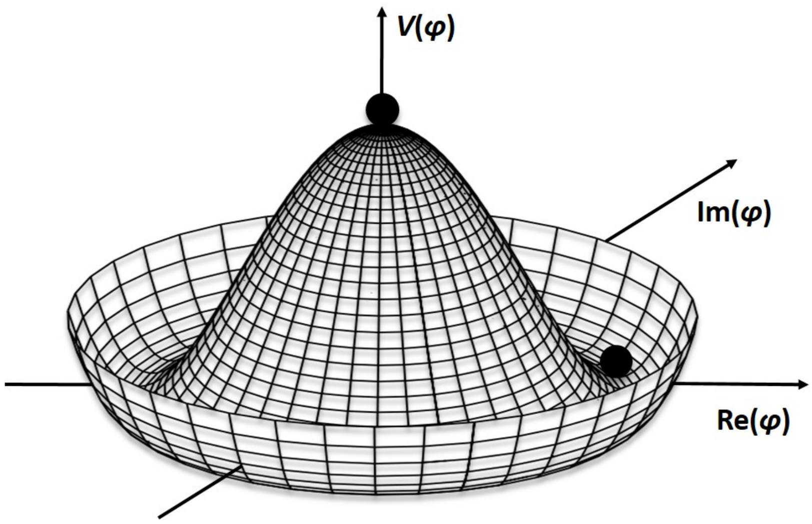


Figure 1.2: Visualization of the BEH potential as function of one of the doublet's components. As $\mu^2 < 0$ (for $\lambda > 0$), the potential acquires the shape of a "Mexican hat"; the center of symmetry $\phi = 0$ becomes unstable to small perturbations and the system falls into the lower energy state.

around $\phi = 0$, in the space defined by the real and complex components of ϕ_1,ϕ_2 (see Figure 1.2). Those vacuum states are defined by:

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}. \tag{1.10}$$

The system will naturally reside in a vacuum state, arbitrarily chosen, however displaced from the center of symmetry. Hence, although the Lagrangian (1.9) is kept invariant, gauge transformations will rotate the vacuum to a different physical state.

Since the BEH field (1.8) is a $Y = 1$ doublet, the appropriate vacuum state to consider is a T_3 eigenstate with $T_3 = -1/2$ [10]. In that case, although the symmetry is broken for all of the generators of the $SU(2)_L \otimes U(1)_Y$ group, the linear combination $Q = T_3 - Y/2$ yields 0; the vacuum is hence unaffected by transformations generated by the electric charge and consequently photons remain massless. Without loss of generality, the vacuum is chosen to lie on the real axis of ϕ_2:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{1.11}$$

where v is the vacuum expectation value and $h(x)$ represents radial excitations about the vacuum.[1]. By developing the partial derivatives from Equation (1.2), the La-

[1]Rotational excitations about the vacuum correspond to gauge transformations and have no physical interest. They can be canceled by a change of gauge due to Eq. 1.5

grangian (1.9) becomes:

$$\mathcal{L}_{BEH} = \frac{1}{4}g^2v^2W^+_\mu W^{-\mu} - \frac{1}{2}\left(\partial_\mu W^+_\nu - \partial_\nu W^+_\mu\right)\left(\partial^\mu W^{-\nu} - \partial^\nu W^{-\mu}\right) + \frac{1}{8}\left(g^2 + g^{''2}\right)v^2 Z_\mu Z^\mu - \frac{1}{4}\left(\partial_\mu Z_\nu - \partial_\nu Z_\mu\right)\left(\partial^\mu Z^\nu - \partial^\nu Z^\mu\right) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\partial_\mu h \partial^\mu h - \mu^2 h^2 + ... \tag{1.12}$$

where:

$$W^\pm = (W^1 \mp iW^2)/\sqrt{2} \tag{1.13}$$

appear as physical, charged bosons with mass $m_W = \frac{1}{2}vg$. A neutral vector boson Z has also emerged as linear combination of W^3 and B,

$$Z_\mu = \frac{-g^{'}B_\mu + gW^3_\mu}{\sqrt{g^2 + g^{''2}}} \tag{1.14}$$

with mass $m_Z = \frac{1}{2}v\sqrt{g^2 + g^{''2}}$. The orthogonal combination to Z_μ on the basis of W^3_μ and B_μ:

$$A_\mu = \frac{g^{'}B_\mu + gW^3_\mu}{\sqrt{g^2 + g^{''2}}} \tag{1.15}$$

is attributed to the photon, provided the condition:

$$gg^{'}/\sqrt{g^2 + g^{''2}} = e \tag{1.16}$$

for compliance with quantum electrodynamics. The absence of the photon term in (1.12) indicates that photons are massless. Finally, it is convenient to introduce a weak mixing angle θ_W to parametrize the mixing of the neutral gauge bosons. With the definition:

$$g^{'} = g\,tan\theta_W, \tag{1.17}$$

equations (1.14) and (1.15) can be rewritten as:

$$Z_\mu = -B_\mu sin\theta_W + W^3_\mu cos\theta_W \tag{1.18}$$

$$A_\mu = B_\mu sin\theta_W + W^3_\mu cos\theta_W \tag{1.19}$$

Due to spontaneous symmetry breaking, the gauge bosons acquire masses by interacting with the BEH field. Furthermore, the h-field appears as a physical boson, with mass $m_h^2 = -2\mu^2 = 2\lambda v^2 > 0$; it is identified as the physical Higgs boson. The values of v and g can be estimated from measured quantities such as the Fermi weak constant G_F, the electric charge e and $\sin^2\theta_W$. The parameter λ is a free parameter of the theory, related to the Higgs mass.

Fermion masses can also be incorporated into the electroweak theory, through gauge-invariant interaction terms, which involve Yukawa couplings of the fermions to the BEH field:

$$\mathcal{L}_{Yukawa} = -G_\ell[\overline{\psi}_R(\phi^\dagger\chi_L) + (\overline{\chi}_L\phi)\psi_R] \tag{1.20}$$

In the case of the electron, setting ϕ to the vacuum state (1.11) yields:

$$\mathcal{L}_{Yukawa} = -G_e \frac{v+h}{\sqrt{2}}(\bar{e}_R e_L + \bar{e}_L e_R) \tag{1.21}$$

$$= -G_e \frac{v+h}{\sqrt{2}}\bar{e}e. \tag{1.22}$$

where G_e is the electron Yukawa coupling. The electron has acquired a mass $m_e = G_e v/\sqrt{2}$ and the coupling of the Higgs boson to the electron is proportional to that mass (m_e/v).

The electroweak theory has been proved to be renormalizable, by Gerard t' Hooft and Tini Veltman [11]. For their contribution to the unification of the weak and electromagnetic interactions, Glashow, Weinberg and Salam shared the Nobel prize in 1979, while t' Hooft and Veltman were awarded the Nobel prize in 1999. Finally, the authors of the BEH mechanism, Peter Higgs and Francois Englert[2] shared the Nobel prize in 2013, one year after the discovery of the Higgs boson.

1.2 Experimental Evidence

The predictions of the electroweak theory lead to a streak of astounding experimental discoveries. In 1973, muonless events in deeply inelastic $\nu_\mu N \to \bar{\nu}_\mu N$ scattering were observed by the Gargamelle bubble chamber collaboration [12], working at the CERN Proton Synchrotron. This observation of weak neutral currents "marked the experimental beginning of the Standard Model of electroweak interactions and triggered huge activity at CERN and all over the world, both on the experimental and theoretical sides" [13].

The next step came a decade later, with the construction of the Super Proton Sychnotron proton-antiproton collider, at CERN. There, the UA1 and UA2 collaborations recorded $W \to e\nu$, $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ decays, making the first observation of the mediators of the weak force, the W and Z bosons [14–18]. This major discovery led to the Nobel Prize for physics being awarded to Carlo Rubbia and Simon van der Meer in 1984.

The continuously increasing amount of knowledge on weak interactions justified building the Large Electron Positron collider (LEP) for further study of the Z boson decay, W boson pair production and most importantly the precise measurement of the W and Z boson masses [19, 20]. In total, four detectors operated on the LEP; ALEPH (Appartus for LEP Physics), DELPHI (Detector with Lepton and Hadron Identification), L3 (Letter of Intent 3) and OPAL (Omnipurpose Apparatus for LEP). The results of those experiments established the SM with unprecedented precision, including the confirmation that the light neutrino species are three, in agreement with the observed generations of fundamental fermions [20].

Around 2001, the LEP finally gave way to the Large Hadron Collider (LHC) [21] for the search of the last missing piece of the electroweak puzzle, the Higgs boson. After years of searching, the collaborations of ATLAS and CMS, the two largest experiments operating at the LHC, announced the discovery of a heavy scalar boson

[2] Englert's co-author Robert Brout had passed away in 2011

on July 2012 [22, 23]. Subsequent data point strongly to properties as expected for the boson associated with the Brout—Englert—Higgs mechanism [24, 25].

1.3 Theories Beyond the Standard Model

The existence of a single Higgs boson is only the simplest theoretical scenario. Since the proposal of the BEH mechanism, numerous extensions Beyond the Standard Model (BSM) have been developed, treating the Higgs boson as part of an extended scalar sector and predicting the existence of additional bosons. The electroweak singlet (EWS) [26] model is the simplest extension of the SM Higgs sector. In this model, a heavy real singlet is introduced in addition to the SM scalar field of the original BEH mechanism. Spontaneous symmetry breaking then gives rise to two CP-even Higgs bosons.

The 2 Higgs Doublet Model (2HDM) [27] is another simple extension to the SM which assumes two complex, scalar SU(2) doublets, ϕ_1 and ϕ_2. Both fields acquire vacuum expectation values, v_1 and v_2, through spontaneous symmetry breaking, revealing five physical Higgs bosons; two CP-even particles (h and H), a neutral CP-odd one (A) and two charged ones ($H^{\pm}$). The Higgs sector of 2HDMs is described by seven free parameters:

- the masses m_h, m_H, m_A and $m_{H^{\pm}}$ of the Higgs bosons,
- the ratio $\tan\beta = v_1/v_2$, of the vacuum expectation values v_1 and v_2 of the two doublets,
- the mixing angle α between the two CP-even bosons,
- the potential parameter m_{12}^2 that mixes the two Higgs doublets.

Since the Lagrangian of the 2HDM has a very general form, another degree of freedom comes from the choice of symmetry for its Yukawa sector (which models the interaction of the two fields with fermions). Several types of 2HDM have been developed depending on this choice. In the Type-I model, ϕ_2 couples to all quarks and leptons, whereas in Type-II, ϕ_1 couples to down-type quarks and leptons and ϕ_2 couples to up-type quarks. The 'lepton-specific' model is similar to Type-I, except that leptons couple to ϕ_1 instead of ϕ_2, while the 'flipped' model is similar to Type-II, except that leptons couple to ϕ_2 instead of ϕ_1. In all of these models the coupling of the H boson to vector bosons is proportional to $cos(\beta - \alpha)$. In the limit $cos(\beta - \alpha) \to 0$ the light, CP-even Higgs boson h is indistinguishable from a SM Higgs boson with the same mass.

1.4 Production and Decay of the Higgs boson

The mass of the Higgs boson m_H is not predicted by SM. Production cross sections and decay branching ratios (BRs) are therefore calculated as functions of m_H. At a hadron collider such as the LHC, SM Higgs production mainly occurs through the following mechanisms, given in order of decreasing cross section:

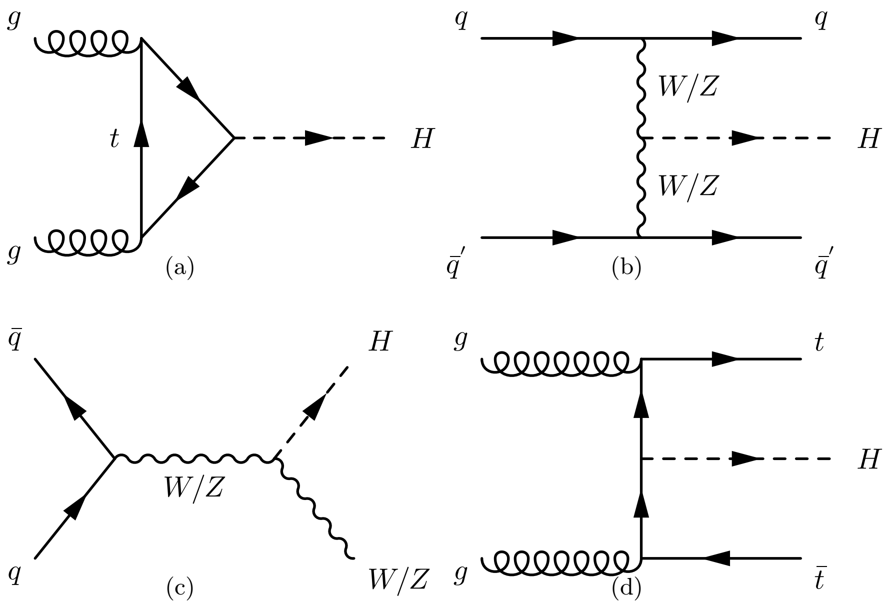
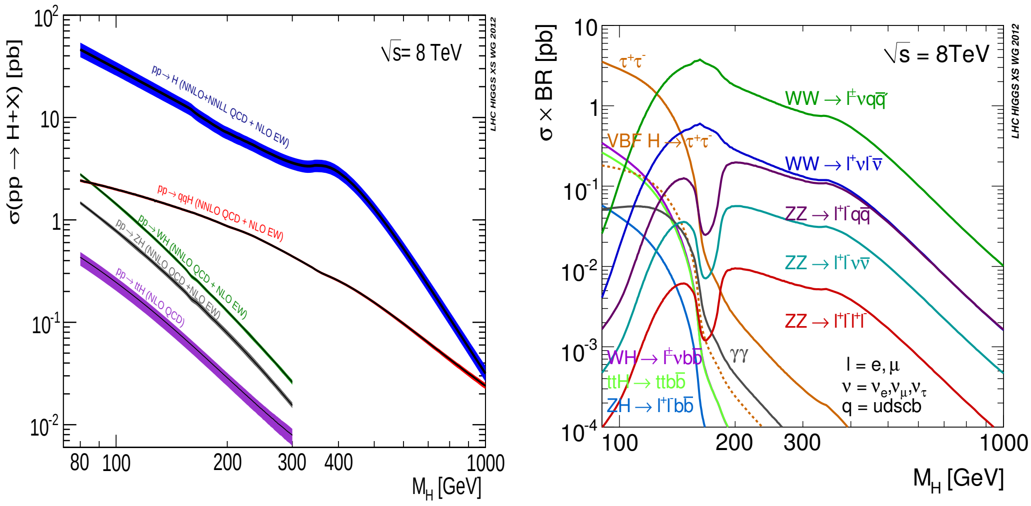


Figure 1.3: Leading order Feynmann diagrams for Higgs boson production via (a) gluon fusion (ggF) and (b) vector boson fusion (VBF). Examples of leading order diagrams are also shown for the (c) WH/ZH and (d) $t\bar{t}H$ associated production mechanisms.

- gluon fusion process $gg \rightarrow H$ (ggF), as in Figure 1.3(a),
- vector boson fusion $qq \rightarrow qqH$ (VBF), as in Figure 1.3(b),
- associated production with a W or Z boson, $qq, gg \rightarrow ZH$ (ZH), $qq, qg \rightarrow WH$ (WH), as in Figure 1.3(c),
- associated production with a pair of top quarks, $qq, gg \rightarrow ttH$ (ttH), as in Figure 1.3(d).

Since the Higgs boson does not couple to massless gluons, ggF proceeds via loops of massive coloured particles, predominantly top quark; contributions from lighter quarks are suppressed as a function of m_q^2. The cross sections for the different Higgs production modes are shown in Figure 1.4. Although ggF is the dominant mode up to ~ 1 TeV, VBF is a powerful discovery channel as well, as it gives rise to final states with distinctive signatures which can be exploited in order to reduce the background. Associated Higgs boson production modes are also important in the low mass range. Due to their distinctive signatures, these channels provide information on the top-Higgs Yukawa coupling ($t\bar{t}H$) and also give access to the Higgs decay into bottom quarks.

The Higgs boson itself cannot be directly observed in a detector, due to its short lifetime. Its discovery is therefore pursued with multiple search channels following its various decay modes. The decay branching ratios for the SM Higgs boson are shown in Figure 1.4(b). For a given m_H, the sensitivity of a search channel depends on the production cross section of the Higgs bosons, its decay branching fraction,



(a) (b)

Figure 1.4: (a) SM Higgs boson production cross sections at the LHC, at a center of mass energy of 8 TeV. The production modes ggF (blue), VBF (red), WH (green), ZH (gray) and ttH (purple) are shown. The bands correspond to theoretical uncertainties. (b) Branching ratios of important Higgs boson decay modes as functions of its mass.

reconstructed mass resolution, event selection efficiency and the level of background in the final state. For a low mass Higgs boson ($110 < m_H < 150$ GeV), where the natural width of the Higgs boson is only a few MeV, the most sensitive channels are $H \to \gamma\gamma$ and $H \to ZZ \to \ell^+\ell^-\ell^+\ell^-$, in which all final state particles can be precisely measured and the reconstructed m_H resolution is excellent. While the $H \to WW \to \ell\nu\ell\nu$ channel has relatively large branching fraction, the m_H resolution is poor due to the presence of neutrinos. The $H \to b\bar{b}$ and the $H \to \tau^-\tau^+$ channels suffer from large backgrounds and poor mass resolution. For $m_H > 150$ GeV, the sensitive channels are $H \to WW$ and $H \to ZZ$, where the W or Z boson decays into a variety of leptonic and hadronic final states. Among the channels with fully reconstructube final states, $H \to ZZ \to \ell^+\ell^- q\bar{q}$ is the dominant. The chapters that follow describe how this search channel has been exploited in the ATLAS experiment towards the discovery of the SM Higgs boson as well as the examination of BSM scenarios.

Chapter 2

The ATLAS experiment

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator, residing in a 27 km circular tunnel, 50-175 m beneath the surface of the earth, at CERN, in Switzerland [21]. Bunches of protons circulate the two LHC rings, in opposite directions, guided by superconducting magnets. The rings intersect at four points around which detectors have been assembled in order to record the products of the proton-proton collisions. During the first running period of the LHC (Run-I, 2010-2012), collisions took place at center-of-mass energies 7 and 8 TeV. In the second running period (Run-II, 2015-2018), the LHC will be operating at a center-of-mass energy of 13 TeV.

The largest of the experiments operating at the LHC are ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Solenoid), LHCb (The Large Hadron Collider beauty) and ALICE (A Large Ion Collider Experiment). ATLAS and CMS use general purpose detectors, optimized for precision measurements of Standard Model as well as new physics signatures. ALICE specializes in the study of the phenomenology in heavy ion (Pb-Pb) collisions, while LHCb is designed to make precision B-physics measurements.

The smaller experiments at the LHC are TOTEM and LHCf which focus on "forward particles", i.e. protons or heavy ions that brush past each other rather than meeting head on when the beams collide. TOTEM uses detectors positioned on either side of the CMS interaction point, while LHCf is made up of two detectors which sit along the LHC beamline, at 140 metres to either side of the ATLAS collision point. MoEDAL uses detectors deployed near LHCb in search of a hypothetical particle called the magnetic monopole.

2.1 The Accelerator Complex

A schematic of the accelerator complex at CERN is shown in figure 2.1. The proton bunches circulating the LHC rings, originate from a duoplasmatron source where hydrogen gas is ionized, breaking down into its constituent electrons and protons. The latter are then guided through a radio frequency quadrupole (QRF), that speeds up and focuses the particle beam, towards the linear accelerator (LINAC2). LINAC2 accepts proton bunches at an energy of 750 keV and injects them into the Proton Synchrotron Booster (PSB) at 50 MeV. The beam is further accelerated by the PSB, up to 1.4 GeV, followed by the Proton Synchrotron (PS) through which it reaches the 7 km in circumference Super Proton Synchrotron (SPS), at 25 GeV. The SPS feeds the beam pipes of the LHC with 450 GeV proton bunches of appropriate spacing. The proton bunches are accelerated by the LHC to the desired energy and continue to circulate the LHC rings for many hours. The LHC is designed to hold 2808 proton bunches per beam, corresponding to 25 ns bunch spacing, each containing

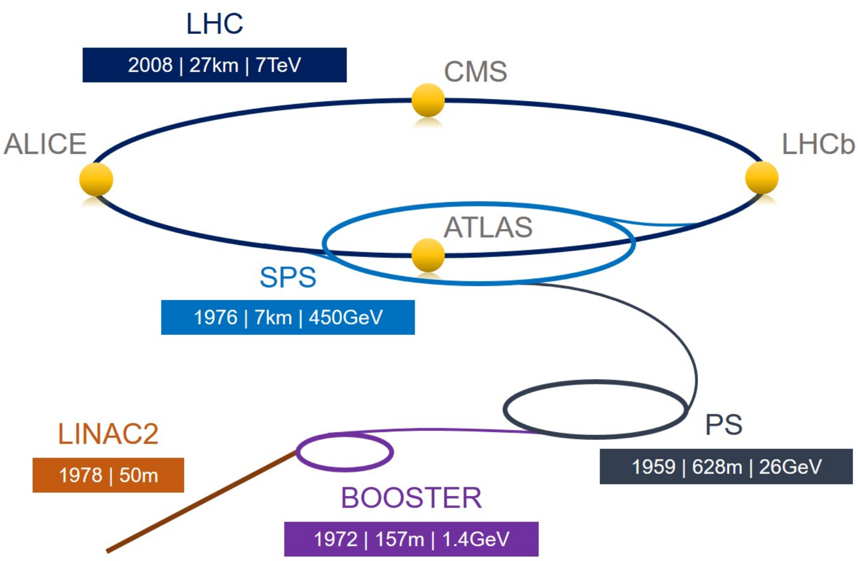


Figure 2.1: The LHC accelerator complex at CERN.

$\approx 1.15 \times 10^{11}$ protons.

2.2 The ATLAS Detector

ATLAS ("A Toroidal LHC ApparatuS") is a cylindrical, general purpose particle detector, designed to fully exploit the physics potential of the LHC [28–30]. It measures 46 m in length, 25 m in height and weighs approximately 7000 tn. The subdetector compartments of ATLAS are: the inner detector, which provides tracking very close to the interaction point, the calorimeters, which measure the energies of charged and neutral particles (other than muons and neutrinos) and the muon spectrometer, which identifies and measures the only charged particles that penetrate the detector, muons. A cut-away view of the ATLAS detector is shown in figure 2.2, while the main performance goals are listed in Table 2.1.

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\%\, p_T \oplus 1\%$	$\pm$ 2.5	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$\pm$ 3.2	$\pm$ 2.5
Hadronic calorimetry (jets)			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\pm$ 3.2	$\pm$ 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 0.05\%$, at $p_T = 1$ GeV	$\pm$ 2.7	$\pm$ 2.4

Table 2.1: General performance goals of the ATLAS detector [29]. For high-p_T muons, the muon-spectrometer performance is independent of the inner-detector system. The units for E and p_T are in GeV.

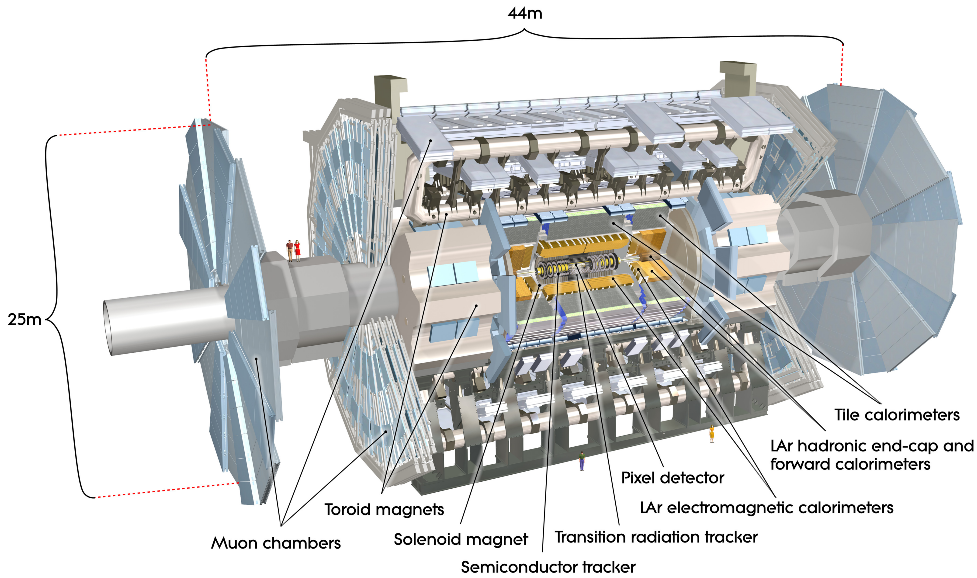


Figure 2.2: Cut-away view of the ATLAS detector. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tonnes.

2.2.1 Coordinate System of ATLAS

ATLAS uses a right-handed coordinate system, centered at the interaction point. The x-axis points towards the center of the LHC ring, the y-axis points upwards and the z-axis points along the beampipe. Quantities labelled as "transverse" refer to projections of observables on the $x-y$ plane. Cylindrical coordinates (r,ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis.

The pseudorapidity is defined in terms of the polar angle θ as:

$$\eta = \ln(\tan\frac{\theta}{2}) \tag{2.1}$$

in order to measure the polar separation of particle tracks. The angles of tracks originating from the interaction point are then described as coordinate pairs (η,ϕ) and the angular separation (for highly relativistic particles), expressed as ($\Delta\eta$,$\Delta\phi$), is a Lorentz invariant under boosts along the beam axis. Finally, the (dimensionless) distance ΔR in the $\eta-\phi$ plane is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.2}$$

2.2.2 Magnet System

ATLAS features a unique hybrid system of four large superconducting magnets. It consists of a central solenoid that surrounds the inner detector, and a system of three large, air-core toroids which generate the magnetic field in the muon spectrometer. A schematic of the magnet system is shown in Figure 2.3. The central solenoid is

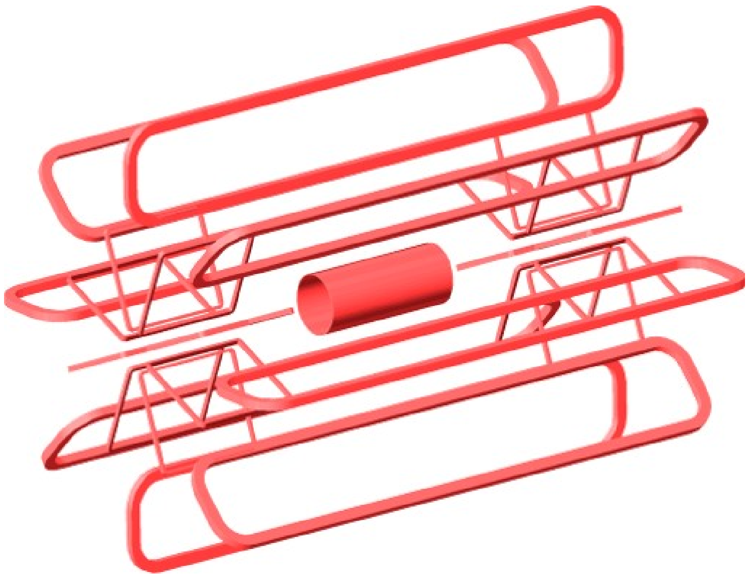


Figure 2.3: Schematic of the ATLAS magnet system.

aligned on the beam axis and provides a 2 T axial magnetic field for the inner detector, while minimising the radiative thickness in front of the barrel electromagnetic calorimeter. To achieve the desired calorimeter performance, the layout has been carefully optimised to keep the material thickness in front of the calorimeter as low as possible, resulting in the solenoid assembly contributing a total of ∼0.66 radiation lengths at normal incidence.

The cylindrical volume surrounding the calorimeters and both end-cap toroids is filled by the magnetic field (∼0.5 T) of the barrel toroid, a set of eight coils, encased in individual racetrack-shaped vacuum vessels. The overall size of the barrel toroid system, as installed, is 25.3 m in length, with inner and outer diameters of 9.4 m and 20.1 m, respectively. The bending power in end-cap regions of muon spectrometer system, is further optimized by end-cap toroids, which provide a 1 T magnetic field.

2.2.3 Inner Detector

The inner detector is the ATLAS subdetector closest to the interaction point [31]. In the barrel region, it comprises a set of concentric layers of fine granularity tracking systems, while in the end-caps the detection elements are installed on disks, perpendicular to the beam axis. The powerful axial magnetic field, generated by the ATLAS solenoid, bends the tracks of charged particles, allowing the measurement of their momentum and electric charge from the curvature. The solenoid is installed between the inner detector and the calorimeters. A cut-away view of the ATLAS inner detector is shown in figure 2.4.

Pixel Detector

The Pixel Detector is the innermost subsystem and therefore provides critical tracking information and determines the ability of the inner detector to measure

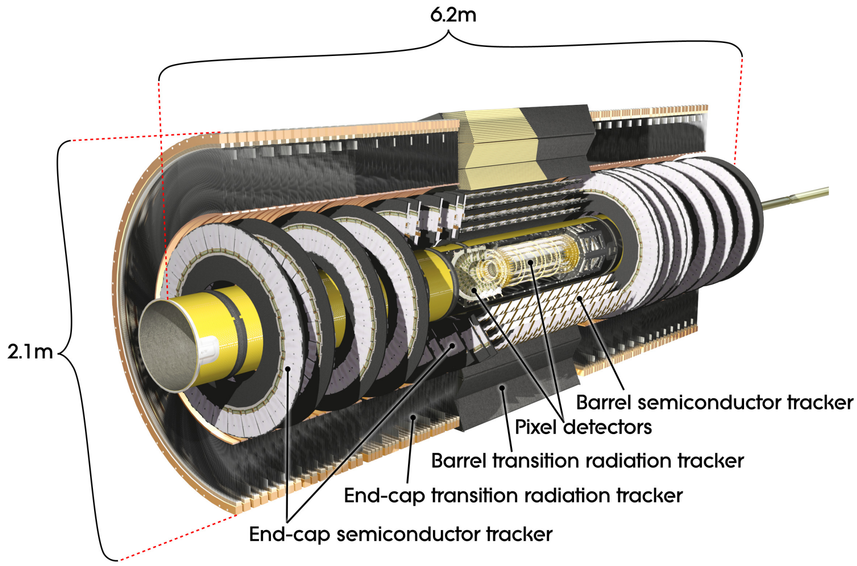


Figure 2.4: Cut-away view of the ATLAS Inner Detector.

vertices. Three cylindrical shells on the barrel (at radial positions 5.05, 8.85 and 12.25 cm) and four disks installed on each side (between radii of 9 and 15 cm) cover the full acceptance of the inner detector, $|\eta| < 2.5$, providing at least three precision measuremets per track. There are about 80.4×10^6 silicon pixels with intrinsic accuracy 10×115 μm^2 in $r\phi - z$ $(r\phi - r)$ in the barrel (end-caps) [32].

The innermost layer of the pixel detector is called the B-Layer since its close proximity to the interaction point determines the impact parameter resolution for B-physics and for b-tagging (i.e. the identification of high-p_{T} jets originating from b-quark decays). In 2014, the pixel detector was upgraded with a fourth pixel layer, called the Insertable B-Layer (IBL) [33], which was intalled at a radius of ≈ 3.3 cm between the existing Pixel Detector and a new (smaller radius) beam-pipe.

Semiconductor Tracker (SCT)

The SCT provides precision measurements in the intermediate radial range and contributes to the measurement of momentum, impact parameter and vertex position. The barrel compartment is divided into four layers of silicon strip modules, at radii of 30.0, 37.3, 44.7, and 52.0 cm, which measure the r, ϕ and z cylindrical coordinates of particle traces. In the end-cap regions, modules are installed on 2×9 disks with the strips running radially.

ATLAS contains 61 m^2 of silicon microstrip detectors, adding up to a total of 6.2 million readout channels. Each silicon detector spans 6.36×6.40 cm^2 with 768 readout strips of 80 μm. An SCT module consists of two pairs of silicon detectors, glued back-to-back, at a 40 mrad angle, in order to allow the estimation of the z coordinate. The intrinsic accuracy per module is 17 μm in

$r\phi$ and 580 μm in z (r) in the barrel (disks).

Transition Radiation Tracker (TRT)

The TRT enhances particle tracking by recording multiple hits per track, in the pseudorapidity region $|\eta| < 2.0$. It makes use of straw detectors, able to operate at very high rates, because of their small diameter (4 mm) and the isolation of the anode wires within individual gas volumes. Electron identification capability is also added by employing Xenon gas to detect transition radiation emmited in the radiator interleaved between the straw layers. Given that a particle track typically intersects 36 straws, the precision obtained from the collective information is ≈170 μm in the direction perpendicular to the straws.

The barrel section of the TRT covers the radial range from 56 to 107 cm. It contains about 50000 straws alligned in parallel with the beam-pipe, each divided in two at the center and read out at each end in order to reduce occupancy. The end-caps contain 320000 straws aligned radially, with readout at the outer radius. Each end-cap consists of 18 wheels. The innermost 14 cover the radial range from 64 to 103 cm, while the last four extend to an inner radius of 48 cm. Each readout channel provides a drift time measurement and two independent thresholds. The latter allow the detector to discriminate between tracking hits, which pass the lower threshold, and transition radiation hits, which pass the higher one.

2.2.4 Calorimeter System

The inner detector is succeeded by sampling calorimeters which stop particles in order to measure their energy. An inner electromagnetic calorimeter (ECal) and an outer hadronic calorimeter (HCal) cover the pseudorapidity regions $|\eta| < 3.2$ and $|\eta| < 4.9$ respectively. A cut-away view of the ATLAS calorimeters is shown in figure 2.5. The size and material of the calorimeters has been optimized in order to minimize punch-through into the muon system.

Electromagnetic calorimeter

The ECal is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$). Lead absorber and liquid argon (LAr) sampler layers are installed in an “accordion-shaped” arrangement (Figure 2.6) [34], providing full coverage in ϕ, without any cracks, and fast extraction of the signal at the rear or at the front of the electrodes. In the region ($|\eta| < 1.8$), a pre-sampler detector is used in order to measure the energy lost before electrons and photons encounter the ECal. The granularity of the ECal varies with η and has been optimized for powerful electron/photon identification, rejection of the hadronic background and distinction of single photons from $\pi^0 \to \gamma\gamma$ decays.

Hadronic calorimeter

The ATLAS HCal consists of a tile calorimeter, covering the region $|\eta| < 1.7$, a Hadronic End-Cap (HEC), covering $1.5 < |\eta| < 3.2$, and a Forward Calorimeter (FCal), covering $3.1 < |\eta| < 4.9$. The tile calorimeter is placed directly outside the ECal envelope. It is a large sampling calorimeter which uses

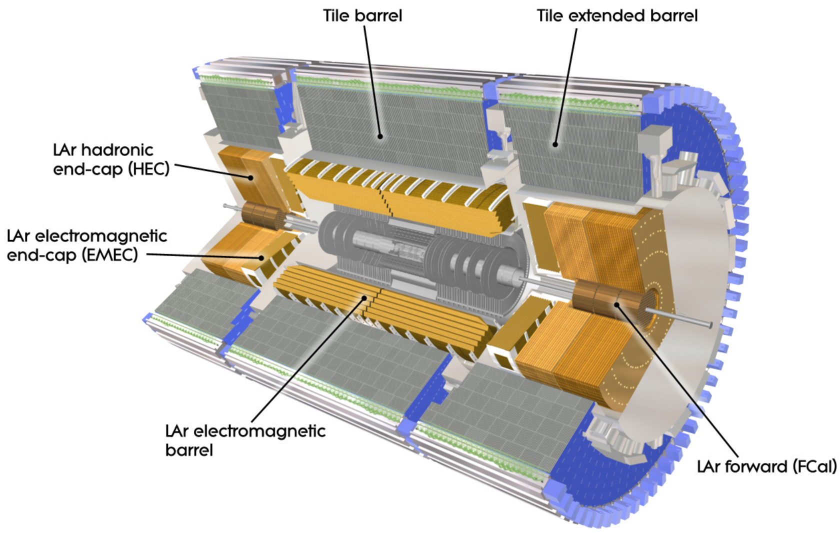


Figure 2.5: Cut-away view of the ATLAS calorimeter system.

steel as absorber and scintillating tiles, read out by wavelength shifting fibers, as active material. The HEC and FCal both use LAr for sampling, mainly because of its intrinsic radiation hardness. The HEC uses a copper absorber, while the FCal uses copper in its first layer, which is optimized for electromagnetic measurements, and tungsten in the two subsequent layers which predominantly measure the energy of hadronic interactions.

The calorimeters that make use of LAr are enclosed in three cryostats. The barrel cryostat contains the barrel ECal, whereas the two end-cap cryostats each contain an end-cap ECal, a HEC and an FCal. The total thickness of the electromagnetic calorimeter corresponds to $> 22(24)$ radiation lengths in the barrel (end-caps), whereas the total thickness of the entire calorimeter corresponds to > 10 interaction lengths. Figure (2.7) shows the material distribution, broken-down in detector elements.

2.2.5 Muon Spectrometer (MS)

The ATLAS MS is instrumented with both trigger and precise position measurement chambers [35]. Muon momentum is determined from the magnetic deflection of tracks by a system of three large air-core toroid magnets; a long barrel and two inserted end-cap magnets. Figure 2.8 shows an overview of the spectrometer.

Monitored Drift Tubes (MDT)

MDTs are the main muon detection elements in $|\eta| < 2.7$. They are pressurised, cylindrical drift tubes with a diameter of 29.970 mm, operating with Ar/CO2 gas (93/7) at 3 bar. The tube wall thickness is 400 μm and their length varies

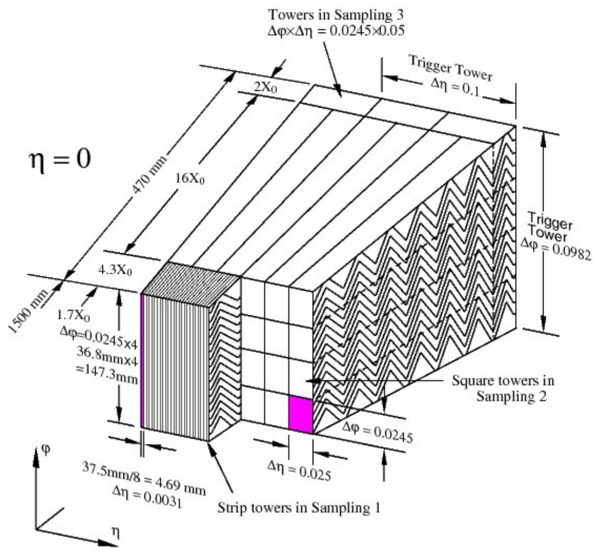


Figure 2.6: Sketch of the accordion structure of the ECal.

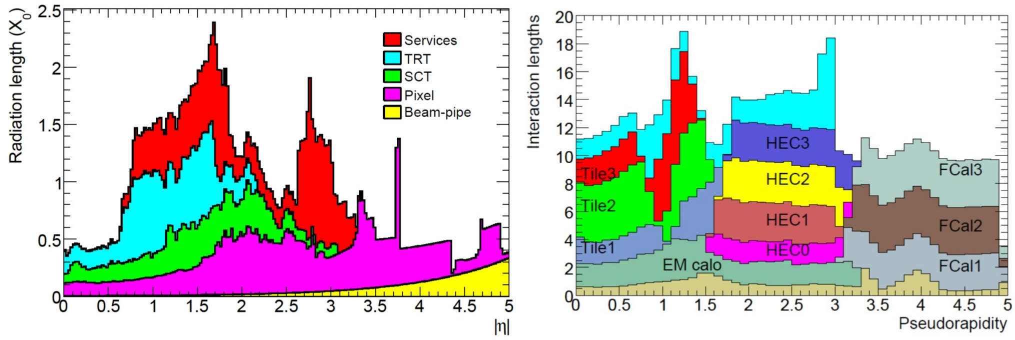
between 70 cm and 630 cm. The single-wire resolution is typically 80 μm, except very close to the anode wire.

MDT chambers are assembled with two sets (multi-layers) of three or four closely-spaced MDT layers (mono-layers). The multi-layers are arranged in parallel, on each side of a rigid support structure, as shown in figure 2.9. The arrangement in multi-layers improves the spatial resolution, beyond the single-wire limit, to $\approx$35 μm. Deformations of the chambers are monitored by built-in optical systems, which explains the "Monitored" of the MDTs.

Cathode Strip Chambers (CSC)

Althouth MDTs are well suited for precision measurements of muons in ATLAS, their large diameter and high operating pressure make them unsuitable for use in areas where high counting rates ($>$150 Hz/cm^2) are expected. Such high background rates are encountered in $|\eta| > 2.0$, in the first end-cap layer (Small Wheel), where CSCs operate. CSCs combine high spatial, time and double track resolution with high-rate capability and low neutron sensitivity. Their operation is considered safe up to counting rates of about 1000 Hz/cm^2 which is sufficient up to the forward boundary of the muon system, at $|\eta| = 2.7$.

CSCs are multi-wire proportional chambers with cathode readout and a small anode wire pitch (2.54 mm), equal to the distance between the cathode-anode plane. The gas used is Ar/CO_2 (80/20). A total of 32 four-layer chambers are arranged in the innermost end-caps (Small Wheels) of ATLAS, covering the pseudorapidity interval $2.04 \leq |\eta| \leq 2.7$. There are two chamber versions, large and small, which differ slightly in the active area. Chambers of both types are installed alternately and partially overlapping, as shown in figure 2.10. They are also inclined by an angle of 11.59° so that, on the average, tracks are normal



(a) (b)

Figure 2.7: (a) Material distribution, in units of radiation length, at the exit of the ID envelope. The distribution is shown as a function of η and averaged over ϕ. (b) Cumulative amount of material, in units of interaction length, in front of the electromagnetic calorimeters, in the electromagnetic calorimeters themselves, in each hadronic layer, and the total amount at the end of the active calorimetry. Also shown for completeness is the total amount of material in front of the first active layer of the muon spectrometer (up to $|\eta| < 3.0$)

to the chamber plane.

In each of the four gaps, the position-sensing cathode strips are lithographically etched. One of the cathodes has precision strips, parallel to the MDT anode wires, measuring the η-coordinate. The second cathode is segmented in coarser strips which are parallel to the CSC wires and provide the transverse, ϕ-coordinate as well as bunch-crossing timing.

Resistive Plate Chambers (RPC)

RPCs are gas chambers used for triggering in the barrel region. They provide rather coarse spatial resolution ($\sim$1 cm) but excellent time resolution $\sim$1 ns. Two layers of RPC modules sandwich the MDT's of the middle layer, while a third one is located close to the outer MDT layer. Each RPC module consists of two parallel bakelite plates, separated by a 2 mm gas-filled gap. Their outer surfaces are coated with graphite layers attached to a high-voltage power supply. The uniform electric field established between the plates ensures that RPC units operate in avalanche mode for fast readout and triggering.

Thin Gap Chambers (TGC)

TGCs provide muon trigger in the end-caps, covering $1.05 < |\eta| < 2.5$. They are multi-wire chambers operated in saturated mode. The cathode to anode gap is small (1.4 mm), smaller than the anode wire pitch (1.8 mm), hence the drift time of electrons emitted from ionization is very short, ensuring excellent time-resolution 2-3 ns.

Unlike RPCs, which are attached on MDT modules, TGCs are independent. One ring layer of TGC modules is located in front of the innermost tracking

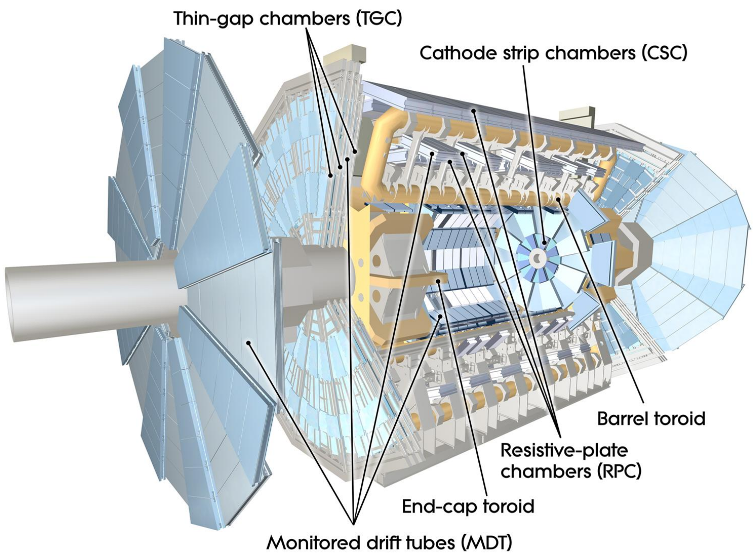


Figure 2.8: Schematic of the ATLAS muon spectrometer.

layer, at $|z| \approx 7$ m, covering the pseudorapidity interval $1.05 < |\eta| < 1.92$. Three more layers are installed at $|z| \approx 14$ m, the first one in front of the middle end-cap, covering $1.05 < |\eta| < 2.7$, and the other two behind the middle end-cap, covering $1.05 < |\eta| < 2.4$.

2.3 Luminosity Measurement

For the online calculation of luminosity, ATLAS uses two dedicated detectors, BCM (Beam Conditions Monitor) and LUCID (LUminosity measurement using a Cherenkov Integrating Detector). The BCM employs radiation hard diamond sensors, mounted around the beam pipe, at $z = \pm 184$ cm and $r = 55$ mm. Equipped with fast electronics (2 ns rise time), the BCM measures the difference in time-of-flight between the forward and backward stations in order to distinguish halo beam particles (machine losses) from normal proton-proton interactions. The BCM was initially designed to monitor the background levels and issue beam-abort requests when beam losses risk damaging the Inner Detector. Its fast readout also provides a bunch-by-bunch luminosity signal at $|\eta| = 4.2$, with a time resolution of 0.7 ns. LUCID is a Cherenkov detector located at a distance of ≈ 17 m from the interaction point. Its main purpose is to detect inelastic pp scattering in the forward direction in order to both measure the integrated luminosity and to provide online monitoring of the instantaneous luminosity and beam conditions. Both BCM and LUCID count the number of activated readout channels per bunch crossing and assume that the number of interactions in a bunch-crossing is proportianal to the measured interaction rate.

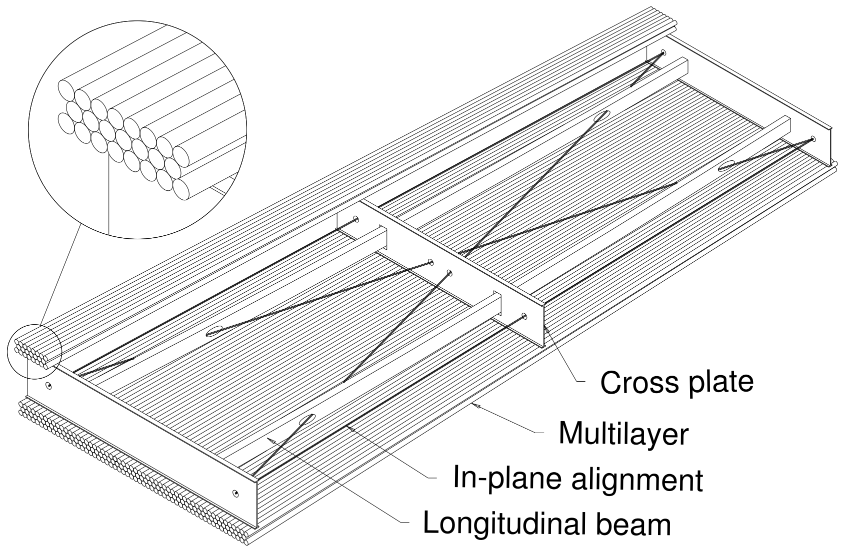


Figure 2.9: Schematic of an MDT chamber.

The luminosity of a *pp* collider can be expressed as:

$$L = \frac{R_{inel}}{\sigma_{inel}} = \frac{\mu n_b f_r}{\sigma_{inel}}, \tag{2.3}$$

where σ_{inel} is the inelastic *pp* cross-section and R_{inel} is the rate of inelastic collisions which depends on the revolution frequency f_r, the number of bunch pairs colliding per revolution n_b and the average number of inelastic interactions per bunch-crossing μ. Since the measurement of both μ and σ_{inel} is subject to the efficiency of the particular detector and algorithm used, the luminosity is formulated as:

$$L = \frac{\mu_{vis} n_b f_r}{\sigma_{vis}}, \tag{2.4}$$

where $\sigma_{vis} = \epsilon\sigma_{inel}$ is the total inelastic cross-section, multiplied by the effciency of a particular detector and algorithm, and similarly $\mu_{vis} = \epsilon\mu$. μ_{vis} is measured with the dedicated ATLAS detectors, however σ_{vis} has to be measured in a separate study.

For the measurement of σ_{vis}, it is convenient to express the luminosity as a function of the collider's parameters:

$$L = \frac{n_1 n_2 n_b f_r}{2\pi\Sigma_x\Sigma_y}, \tag{2.5}$$

where $n_{1,2}$ are the known numbers of protons in each of the colliding bunches (estimated by dedicated beam-charge monitors) and $\Sigma_{x,y}$ are the convolved beam sizes,

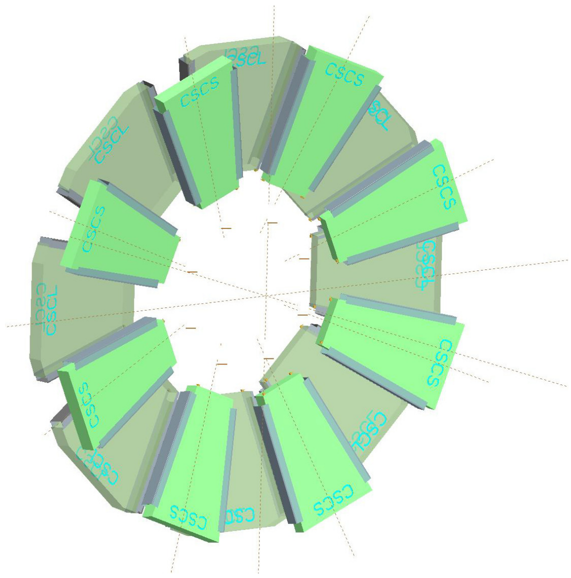


Figure 2.10: Layout of a CSC end-cap with eight small and eight large chambers.

defined by Simon van der Meer as:

$$\Sigma_u = \frac{\int R(\delta u) d(\delta u)}{R(0)}. \tag{2.6}$$

R is the interaction rate and u $(= x, y)$ measures the transverse displacement of the two crossing bunches. Σ_u are measured with dedicated "van der Meer scans" (vdM) [36], in which event rates R are measured for various, known transverse displacements in the x or y axis. The values of Σ_x and Σ_y are obtained numerically from the measurements.

In addition to the detectors listed above, further luminosity-sensitive methods have been developed using components of the ATLAS calorimeter system. These techniques do not identify particular events, but rather measure average particle rates over longer time scales and improve the luminosity estimation offline.

2.4 Trigger System

The LHC has been designed to produce an instantaneous luminosity of 10^{34} cm^{-2}s^{-1}, with 25 ns (40 MHz) bunch-crossing rate. With an average of more than 20 interactions per bunch crossing, the incoming data flow is predicted to be ≈60 TB/s, which greatly surpasses the offline computing power and storage capacity. Therefore, ATLAS employs a trigger and data acquisition system (TDAQ) [28,29], designed to capture the physics of interest with high efficiency while providing sufficient background rejection to reduce the event rate to as low as 200 Hz (≈300 MB/s).

The ATLAS trigger system is based on three levels of event selection; Level 1 (L1), which is hardware-based, Level 2 (L2) and Event Filter (EF). L2 and EF are collectively referred to as the High Level Trigger (HLT) and are based on software algorithms. The L1 trigger searches for high transverse-momentum muons, electrons, photons, jets and τ-leptons decaying into hadrons, as well as large missing and total transverse energy. Its selection is based on information from a subset of detectors, using a limited amount of information in order to make a decision in less than 2.5 μs. The L1 trigger reduces the event rate to $\approx$75 kHz.

For each event passing the L1 selection, the L1 trigger also defines one or more Regions-of-Interest (RoIs), i.e. geographical coordinates in η and ϕ where interesting features have been detected. This information is used by the HLT. L2 selections use all the available detector data within the RoI, at full granularity and precision. The L2 menus are designed to reduce the trigger rate to approximately 3.5 kHz, with an average event processing time of $\sim$10 ms. The final stage of the event selection is carried out by the EF which reduces the event rate to roughly 200 Hz. Its selections are implemented using offline analysis procedures within an average event processing time of $\sim$1 s, achieving a data rate of $\sim$1 GB/s.

Chapter 3

Calibration of Cathode Strip Chambers

As discussed in Section 2.2.5, CSCs are the ATLAS elements used for muon tracking close to the beam pipe, in the critical pseudorapidity region $2.04 \leq |\eta| \leq 2.7$. Each CSC plane consists of 192 precision strips, measuring the η coordinate, and 48 transverse strips, measuring the ϕ coordinate. The readout pitch of precision strips is 5.308 and 5.556 mm for large and small chambers respectively, whereas for transverse strips it is 12.922 and 21.004 mm. This segmentation of the readout cathode allows good spatial resolution, < 60 μm in the η coordinate.

The spatial resolution of the CSCs is not particularly sensitive to variations in the gas gain, since reconstruction algorithms scan the arrays of strips looking for charge peaks and determine a hit position from the ratio of the charge induced on strips adjacent to a peak. Of course, it is still important that variations in the gas gain among CSC layers are monitored and that they remain small, so that an effective charge threshold may be defined for distinction between signal and noise. On the other hand, variations in the electronics-gain between neighboring readout channels are crucial to the spatial resolution of the chambers. Such gain measurements were performed in 2012, using the first data collected with ATLAS, corresponding to ≈ 5.4 fb^{-1}.

3.1 High Voltage corrections

Although the gas gain in a CSC layer cannot be measured from the data, the gain ratio of two layers can be approximately calculated by the ratio of the charge induced, considering that the average amount of ionization is the same in all layers. Hence, charge measurements can be used to correct the high voltage applied to each CSC layer in order to establish uniform gain.

For each precision strip i, the induced charge (Q^i_{peak}) distribution is constructed from charge clusters in which the particular strip exhibits the maximum (peak) charge, compared to its neighbouring strips. Only clusters that belong to reconstructed muon tracks are considered. The most-probable value Q^i_{mpv} is then extracted by fitting a Landau line shape to the distribution, as shown in Figure 3.1 for one CSC strip[1]. Subsequently, the distribution of Q_{mpv} is constructed from the fit results on all CSC channels and the measured mean value is set as benchmark. The distribution of Q_{mpv} obtained with the 2012 ATLAS data set, is shown Figure 3.2. The mean value is $Q^{ref}_{mpv} = 197.8 \pm 0.3$ Ke.

Once the benchmark has been defined, the high voltage applied to each layer ℓ is corrected so that the most-probable charge induced within the layer (Q^{ℓ}_{mpv}) acquires the reference value. Q^{ℓ}_{mpv} is obtained from the distribution of Q_{peak} of all clusters

[1] The naming scheme for CSC chambers defines the wheel as “A” or “C”, followed by the chamber (sector) number. Small chambers have even numbers whereas large chambers have odd.

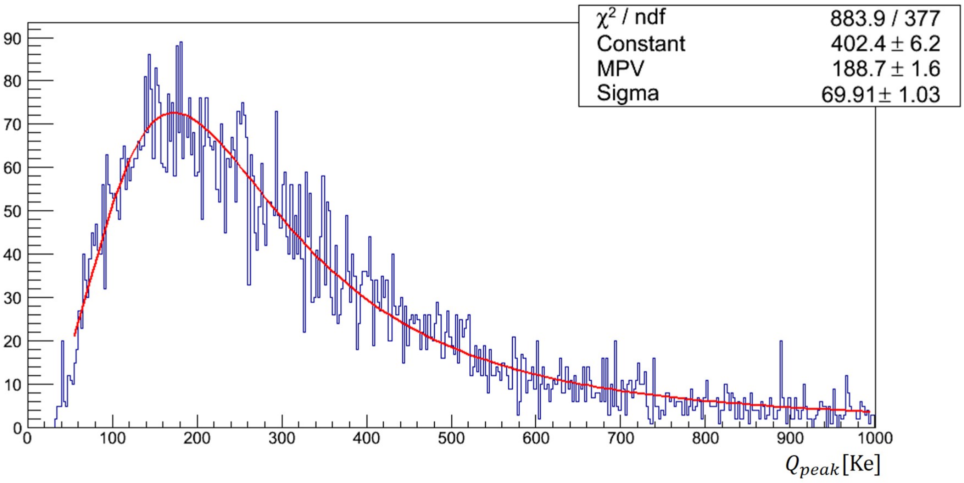


Figure 3.1: Distribution of induced charge (Q_{peak}) on sector C03, layer 2, channel 17. Only clusters in which the particular strip exhibits the maximum (peak) charge, compared to its neighbouring strips, are considered.

in the layer. The correction to the high voltage is derived from the relationship[2]:

$$Q^{ref}_{mpv} = Q^{\ell}_{mpv} e^{\Delta V/c}, \tag{3.1}$$

where $c = 50(\mathrm{V})/\ln 2$. The results obtained with this procedure, using the 2012 ATLAS data, are summarized in Tables 3.1 and 3.2. Those corrections were applied during a scheduled technical stop of the LHC.

3.2 Channel-to-channel gain variations

For the measurement of variations in the electronics-gain, the gain ratio $R_G = G^i/G^{i+1}$ between adjacent strips is estimated from the ratio of induced charges $R_Q = Q^i_{mpv}/Q^{i+1}_{mpv}$, as defined in the previous section. The Q^i_{mpv} values measured in the 2012 analysis, are presented in Figures 3.6–3.9, while the distribution of the R_Q ratios is shown in Figure 3.3. The latter has a mean value of ≈1, as expected, and its standard deviation is found 1.88%. Part of the uncertainty is contributed by errors in the determination of the Q^i_{mpv} values. The distribution of $\delta Q_{mpv}/Q_{mpv}$, formed from the fit results on all strips (Figure 3.4), shows a relative uncertainty of 0.64% (0.87%) for large (small) chambers. Hence, the corresponding contribution to the ratios R_Q is 0.91% (1.23%). By subtracting this component, an uncertainty of 1.65% (1.42%) remains as the upper limit of the contribution of gain variations.

An alternative approach, less prone to statistical errors, uses information from

[2]This relationship is derived from the fact that the amplitude increases roughly exponentially with the applied voltage and the knowledge that it doubles for each 50 V increase, for the particular chambers.

layer	Q^{ℓ}_{mpv}(Ke)	$Q^{\ell}_{mpv}/Q^{ref}_{mpv}$	HV_{corr}(V)	layer	Q^{ℓ}_{mpv}(Ke)	$Q^{\ell}_{mpv}/Q^{ref}_{mpv}$	HV_{corr}(V)
		C01				C02	
1	185.0	0.94	5	1	190.7	0.96	3
2	218.2	1.10	-7	2	185.3	0.94	5
3	146.8	0.74	21	3	177.2	0.90	8
4	177.4	0.90	8	4	182.3	0.92	6
		C03				C04	
1	202.8	1.00	-2	1	231.7	1.20	-11
2	0.0	0.00	0	2	156.8	0.79	17
3	193.4	0.98	2	3	195.5	0.99	1
4	227.0	1.10	-10	4	139.8	0.71	25
		C05				C06	
1	173.9	0.88	9	1	192.4	0.97	2
2	144.4	0.73	23	2	173.7	0.88	9
3	185.8	0.94	5	3	197.9	1.00	0
4	153.9	0.78	18	4	167.6	0.85	12
		C07				C08	
1	221.8	1.10	-8	1	179.7	0.91	7
2	211.0	1.10	-5	2	192.9	0.98	2
3	151.8	0.77	19	3	192.3	0.97	2
4	152.9	0.77	19	4	174.8	0.88	9
		C09				C10	
1	208.0	1.10	-4	1	206.2	1.00	-3
2	164.8	0.83	13	2	183.6	0.93	5
3	194.9	0.99	1	3	207.4	1.00	-3
4	200.7	1.00	-1	4	174.4	0.88	9
		C11				C12	
1	201.6	1.00	-1	1	168.8	0.85	11
2	179.3	0.91	7	2	195.5	0.99	1
3	196.5	0.99	0	3	178.1	0.90	8
4	157.6	0.80	16	4	186.6	0.94	4
		C13				C14	
1	180.9	0.91	6	1	190.2	0.96	3
2	190.8	0.96	3	2	191.9	0.97	2
3	185.4	0.94	5	3	193.6	0.98	2
4	171.0	0.86	11	4	188.3	0.95	4
		C15				C16	
1	182.1	0.92	6	1	195.7	0.99	1
2	199.4	1.00	-1	2	182.1	0.92	6
3	193.4	0.98	2	3	187.7	0.95	4
4	173.7	0.88	9	4	187.6	0.95	4

Table 3.1: HV corrections for CSC layers (sectors C16-C01), estimated using ATLAS data.

layer	Q^{ℓ}_{mpv}(Ke)	$Q^{\ell}_{mpv}/Q^{ref}_{mpv}$	HV_{corr}(V)	layer	Q^{ℓ}_{mpv}(Ke)	$Q^{\ell}_{mpv}/Q^{ref}_{mpv}$	HV_{corr}(V)
	A01				A02		
1	169.4	0.86	11	1	225.9	1.10	-10
2	201.8	1.00	-1	2	179.4	0.91	7
3	169.0	0.85	11	3	188.8	0.95	3
4	161.1	0.81	15	4	199.0	1.00	0
	A03				A04		
1	185.3	0.94	5	1	183.5	0.93	5
2	166.4	0.84	12	2	176.7	0.89	8
3	191.5	0.97	2	3	192.5	0.97	2
4	177.1	0.90	8	4	182.4	0.92	6
	A05				A06		
1	176.3	0.89	8	1	193.0	0.98	2
2	204.5	1.00	-2	2	193.3	0.98	2
3	0.0	0.00	0	3	197.1	1.00	0
4	174.2	0.88	9	4	170.7	0.86	11
	A07				A08		
1	222.4	1.10	-8	1	205.4	1.00	-3
2	202.4	1.00	-2	2	209.2	1.10	-4
3	168.4	0.85	12	3	194.1	0.98	1
4	199.2	1.00	-1	4	205.5	1.00	-3
	A09				A10		
1	0.0	0.00	0	1	171.6	0.87	10
2	212.7	1.10	-5	2	193.2	0.98	2
3	185.5	0.94	5	3	181.1	0.92	6
4	167.5	0.85	12	4	214.2	1.10	-6
	A11				A12		
1	251.9	1.30	-17	1	210.1	1.10	-4
2	154.9	0.78	18	2	181.6	0.92	6
3	183.5	0.93	5	3	204.5	1.00	-2
4	122.7	0.62	34	4	189.7	0.96	3
	A13				A14		
1	212.3	1.10	-5	1	171.1	0.86	10
2	170.0	0.86	11	2	227.7	1.20	-10
3	178.0	0.90	8	3	142.9	0.72	23
4	152.4	0.77	19	4	216.9	1.10	-7
	A15				A16		
1	168.4	0.85	12	1	183.0	0.93	6
2	168.2	0.85	12	2	186.1	0.94	4
3	198.2	1.00	0	3	165.3	0.84	13
4	190.1	0.96	3	4	208.3	1.10	-4

Table 3.2: HV corrections for CSC layers (sectors A01-A16), estimated using ATLAS data.

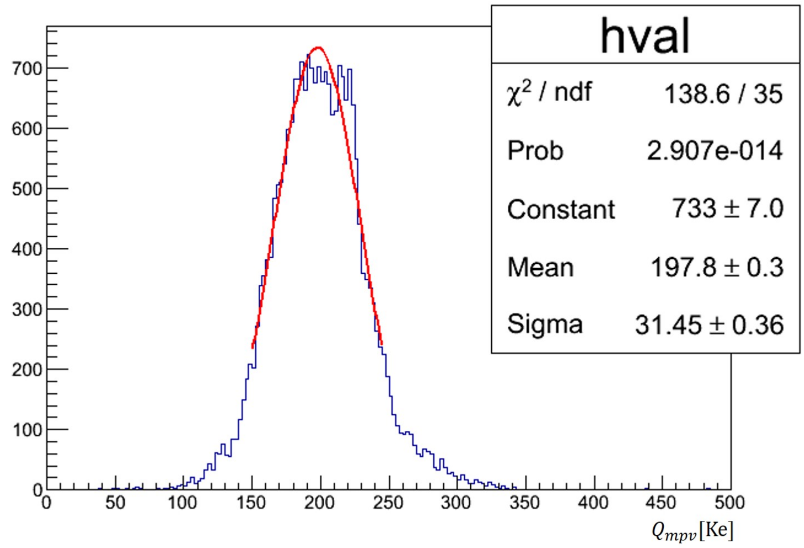


Figure 3.2: Distribution of Q_{mpv} obtained from all CSC channels. The distribution is fitted with a gaussian line shape.

matched clusters, i.e. pairs of clusters in η and ϕ, induced by the same avalanche. First, for each transverse strip j, the distribution of $R_Q^{ij} = Q_{peak}^i/Q_{peak}^j$ is constructed from clusters peaking at that strip j. Q_{peak}^i is the peak-strip charge of the matched clusters in η. Assumming that the peak-strip charge induced by the same avalanche is approximately equal in the two coordinates, then R_Q^{ij} is an estimate of the gain ratio G^i/G^j. Therefore, the mean value of each distribution provides:

$$R^j =< G_\eta > /G^j \tag{3.2}$$

where $< G_\eta >$ is an estimate of the average layer gain in precision strips. Although it is measured for a specific transverse strip j, it assummed to be the same for all transverse strips.

With the R^j values at hand, a second iteration follows, this time on precision strips. For each strip i, the distribution of $Q_{peak}^i/(Q_{peak}^j R^j)$ is constructed, using matched clusters, as in the previous step. From this distribution, the relative gain:

$$R^i = \frac{G^i}{< G >_\eta} \tag{3.3}$$

is evaluated by fitting a gaussian line shape. The ratios R^i/R^{i+1} are finally used as estimates of the gain ratios R_G (instead of the ratios in Q_{mpv} which are used in the first approach).

The R values measured in all precision strips, in the 2012 analysis, are graphically depicted in Figures 3.10-3.13, while the distribution of R^i/R^{i+1} is shown in Figure 3.5. The latter has a mean value of ≈1, as expected, and its standard deviation

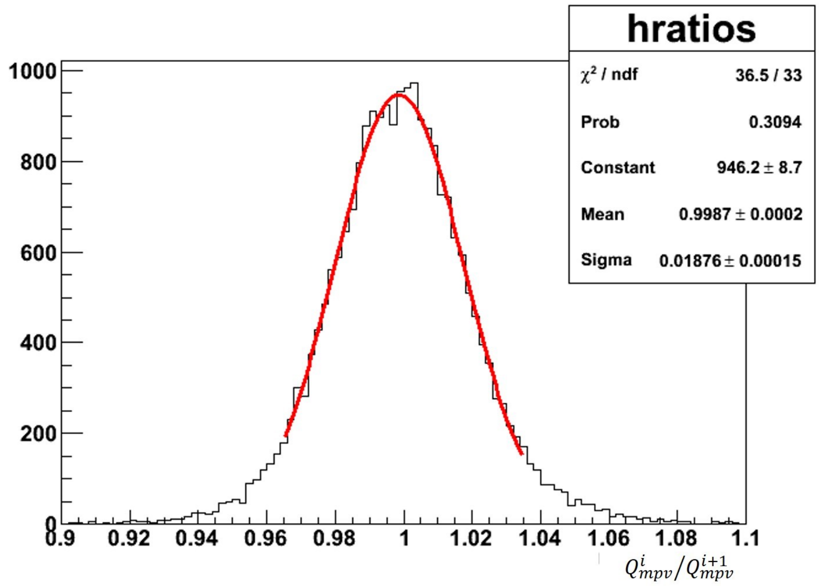


Figure 3.3: The distribution of induced charge (Q_{mpv}) ratios in adjacent strips.

is found 1.2%. Part of the uncertainty is contributed by errors in the determination of the R ratios. The distribution of $\delta R/R$, formed from the fit results on all strips, shows an average error of 0.36% (0.49%) for large (small) chambers. Hence, the contribution of these errors to the ratios R is 0.51% (0.69%). By subtracting this statistical component, an uncertainty of 1.09% (0.98%) remains as the upper limit of the contribution of gain variations. These results showed that channel-to-channel gain variations were not the dominant source of uncertainty in the position resolution of the CSCs. Furthermore, the measurements could be used in order to correct the position reconstruction algorithms but, given the size of the corrections and the statistical uncertainties that accompany the measurements, the results would be marginal.

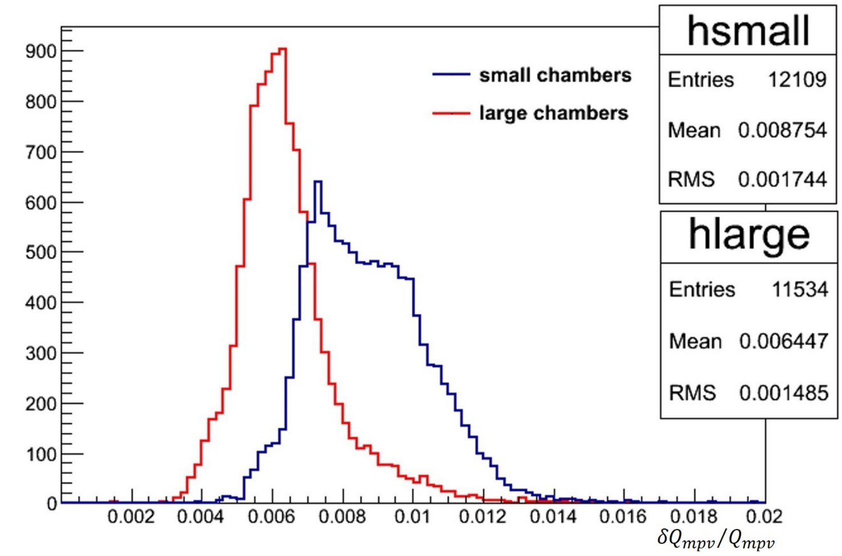


Figure 3.4: Distribution of the relative error in the measurement of the induced charge (Q_{mpv}) on each strip.

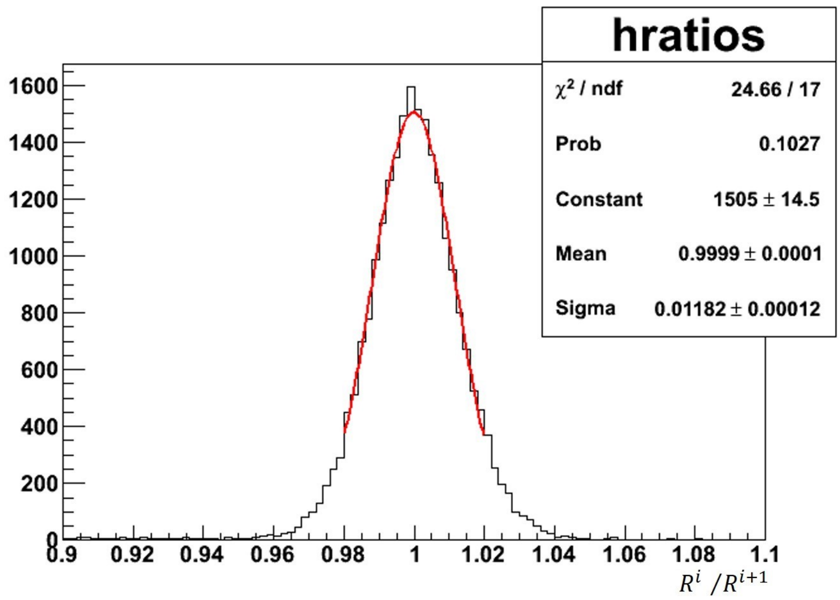


Figure 3.5: Distribution of R^i/R^{i+1}.

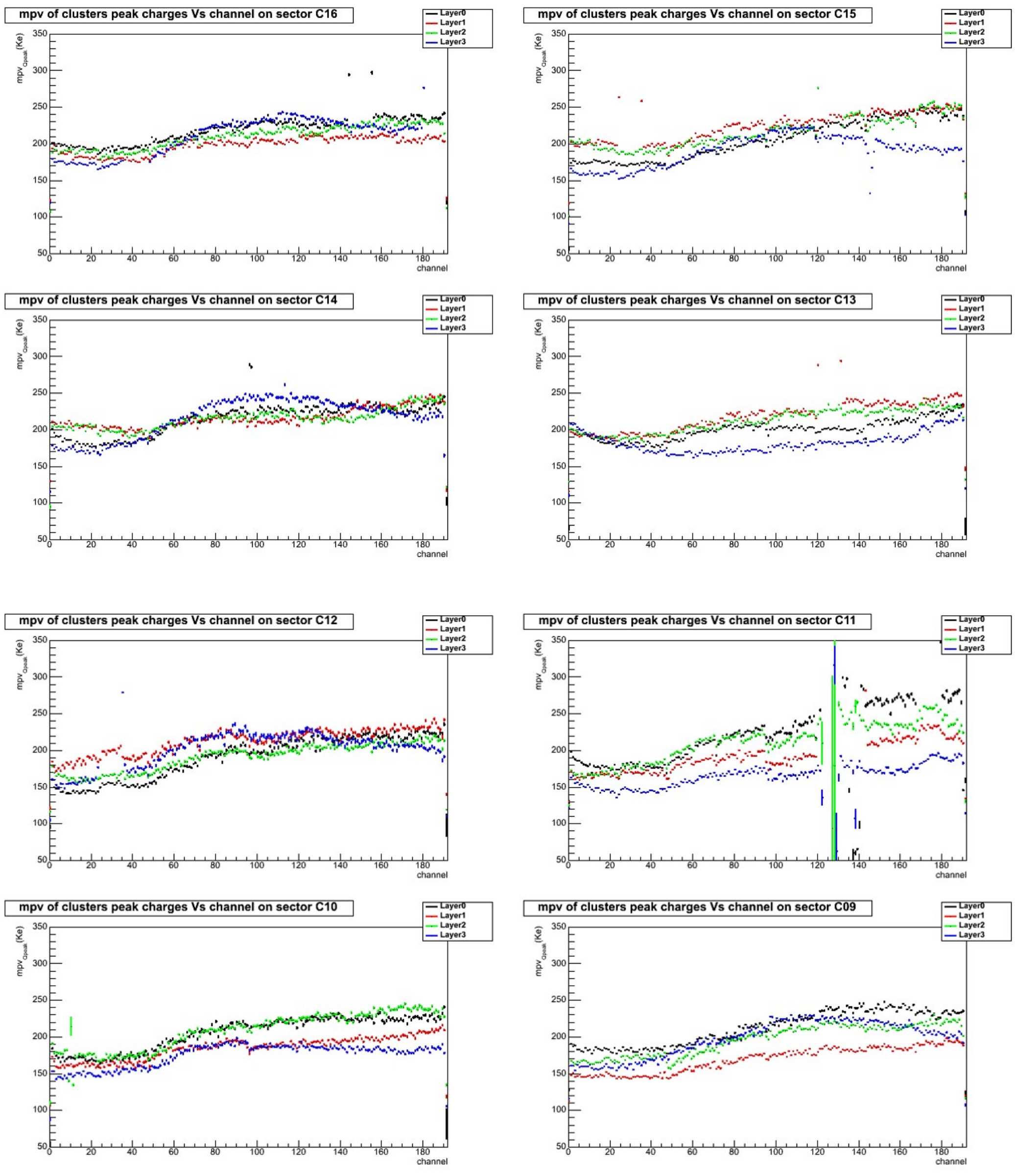


Figure 3.6: Induced charge (Q_{mpv}) on each precision strip (sectors C16 - C09). Dead or problematic channels are not used in the measurements.

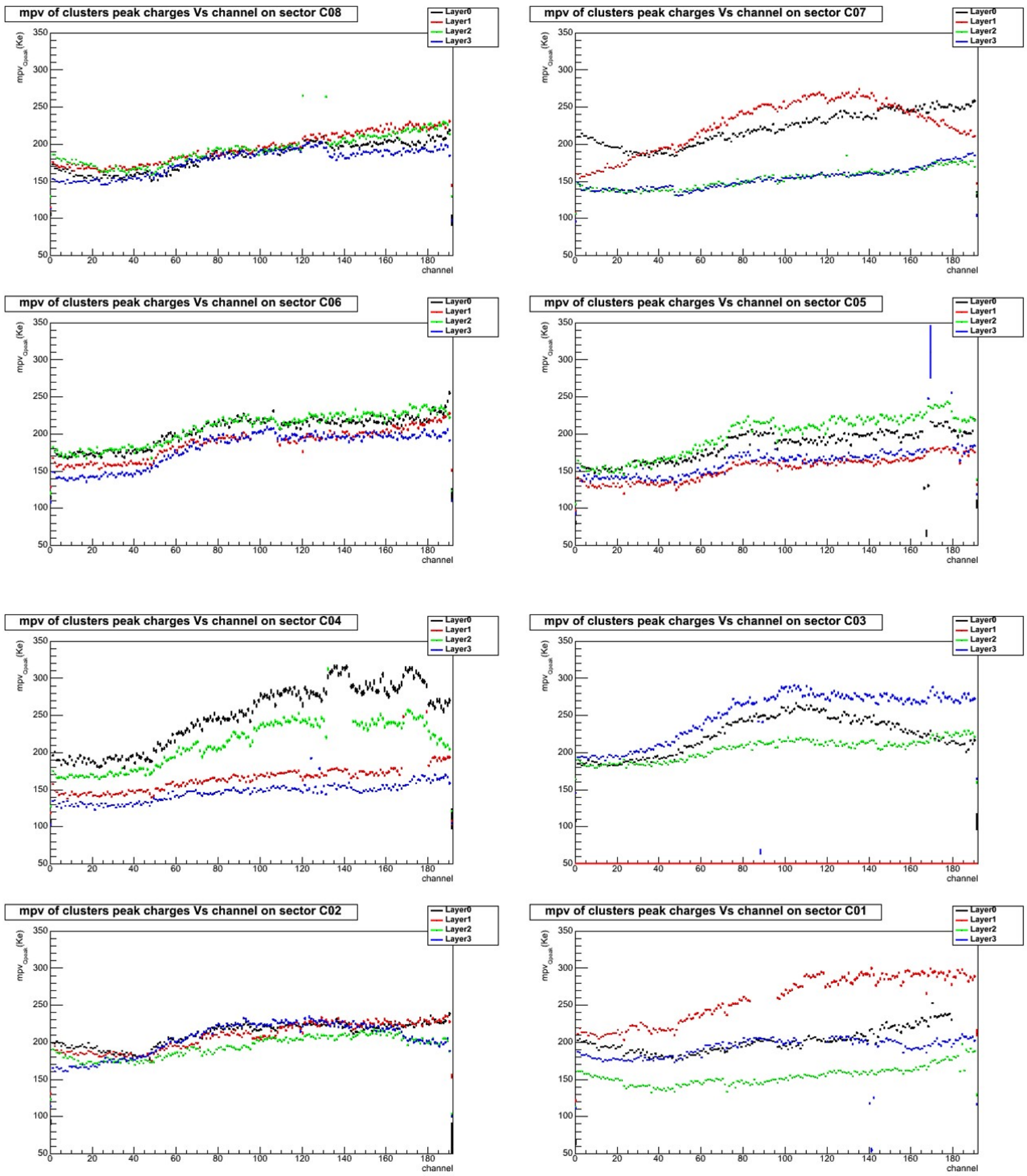


Figure 3.7: Induced charge (Q_{mpv}) on each precision strip (sectors C08 - C01). Dead or problematic channels are not used in the measurements.

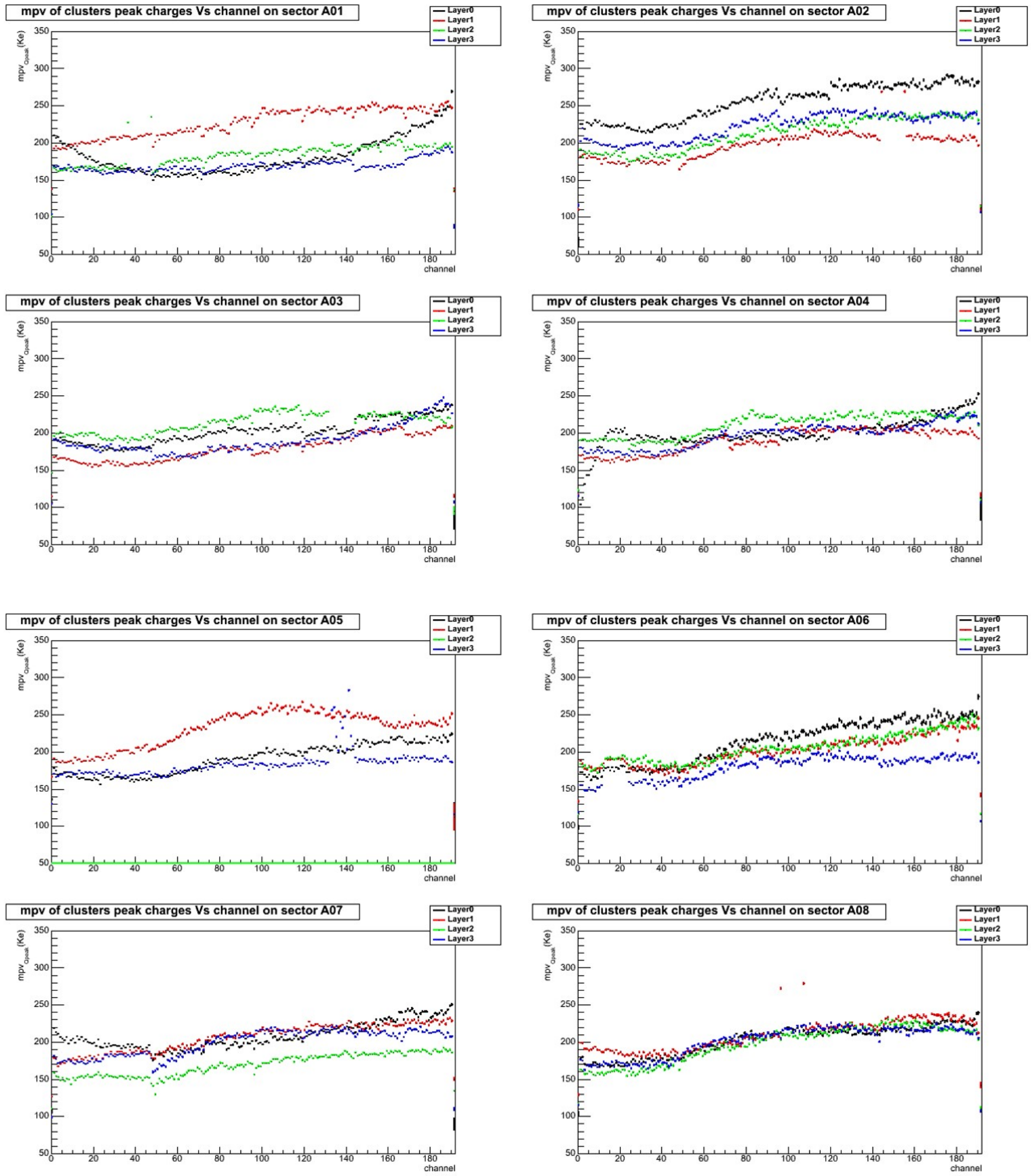


Figure 3.8: Induced charge (Q_{mpv}) on each precision strip (sectors A01 - A08). Dead or problematic channels are not used in the measurements.

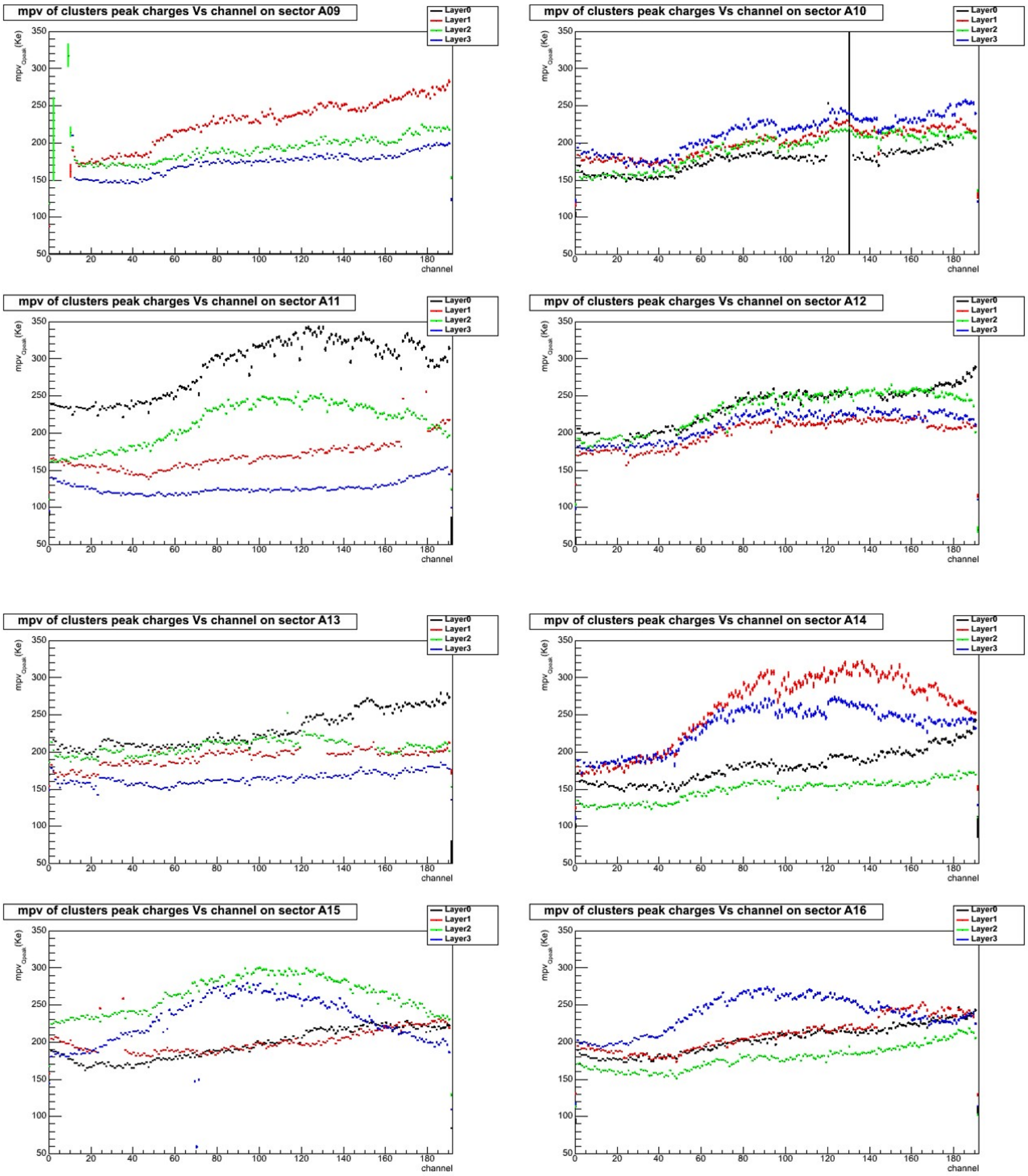


Figure 3.9: Induced charge (Q_{mpv}) on each precision strip (sectors A09 - A16). Dead or problematic channels are not used in the measurements.

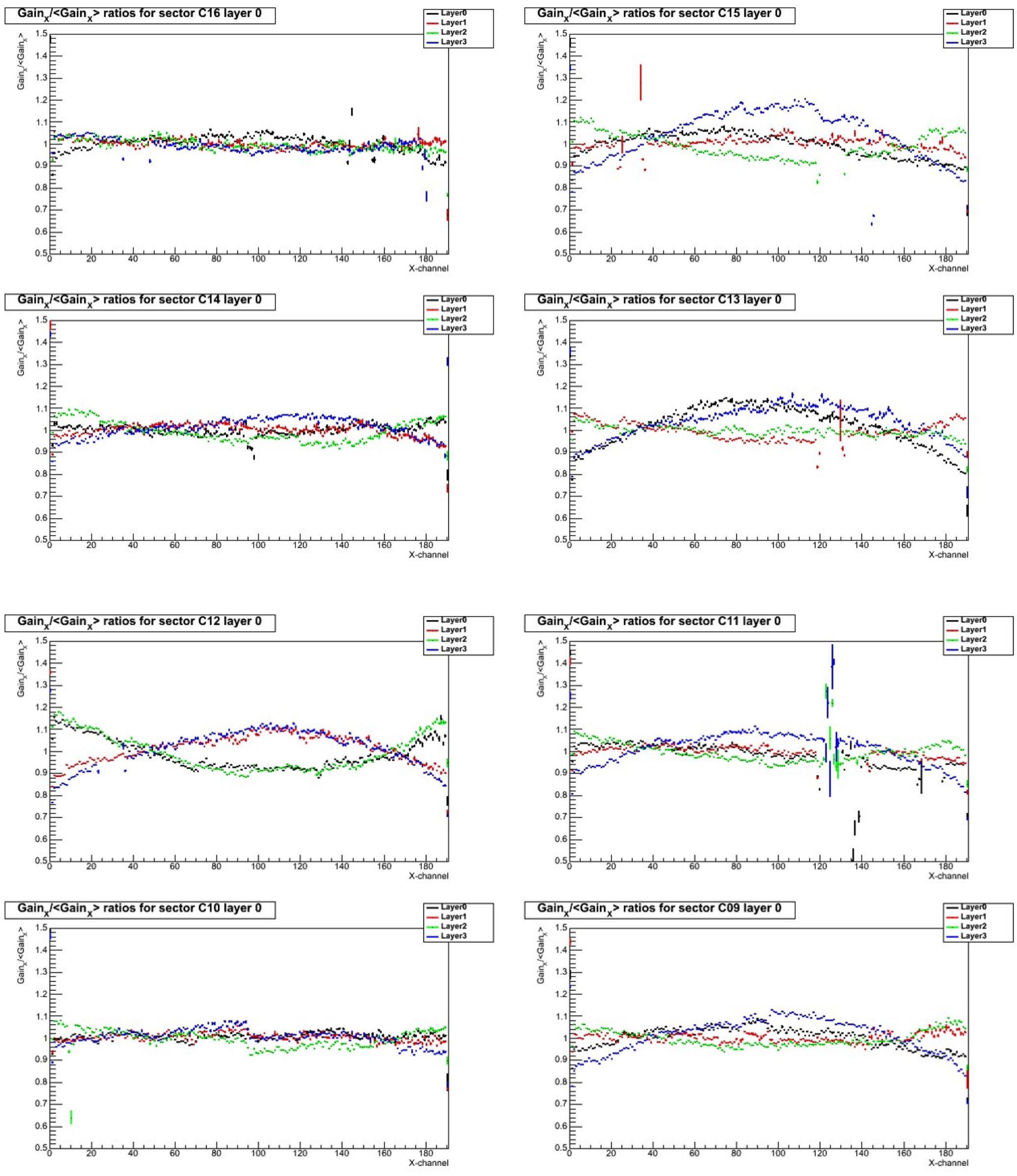


Figure 3.10: Relative gain estimate (R) for each precision strip (sectors C16 - C09). Dead or problematic channels are not used in the measurements.

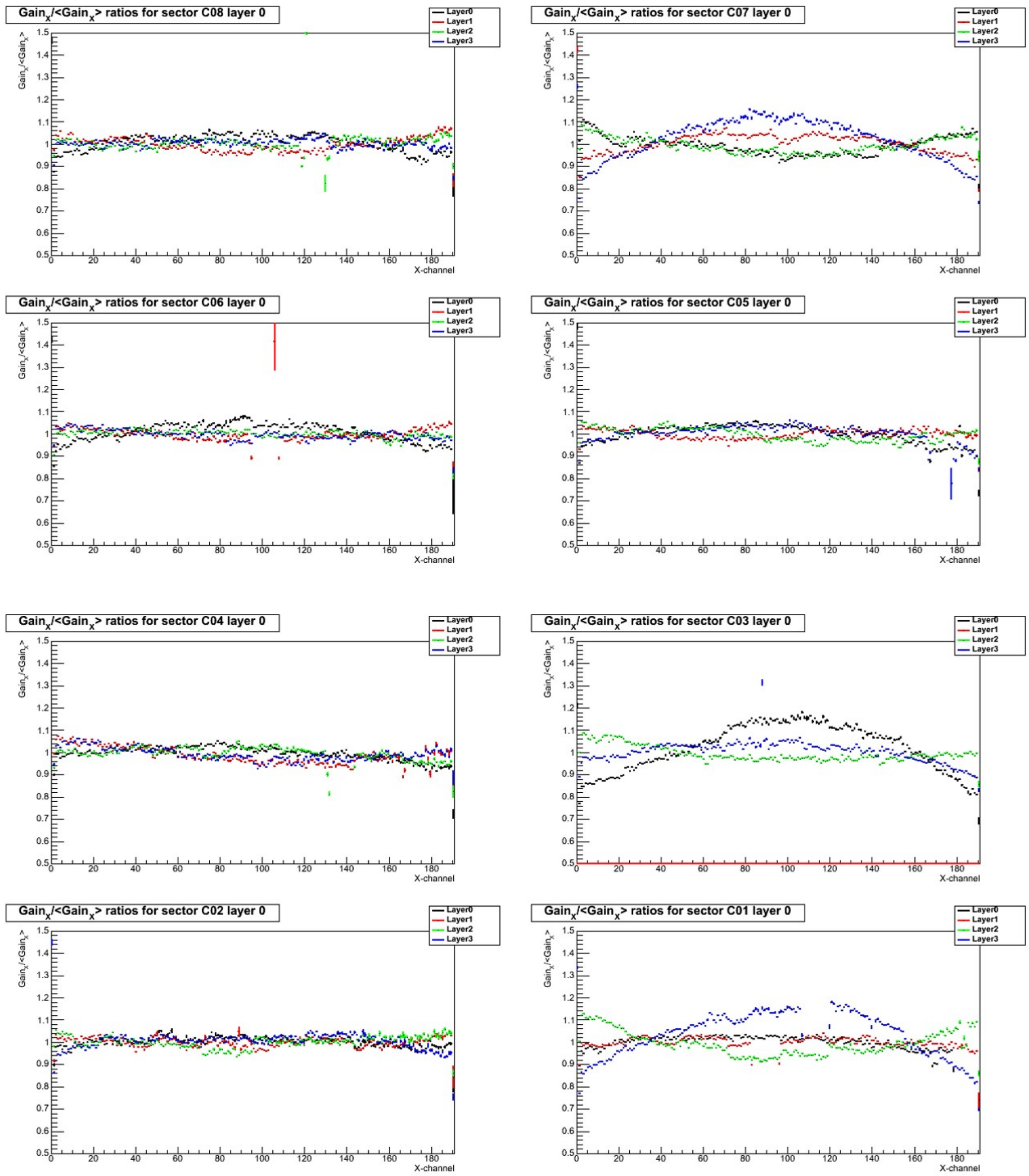


Figure 3.11: Relative gain estimate (R) for each precision strip (sectors C08 - C01). Dead or problematic channels are not used in the measurements.

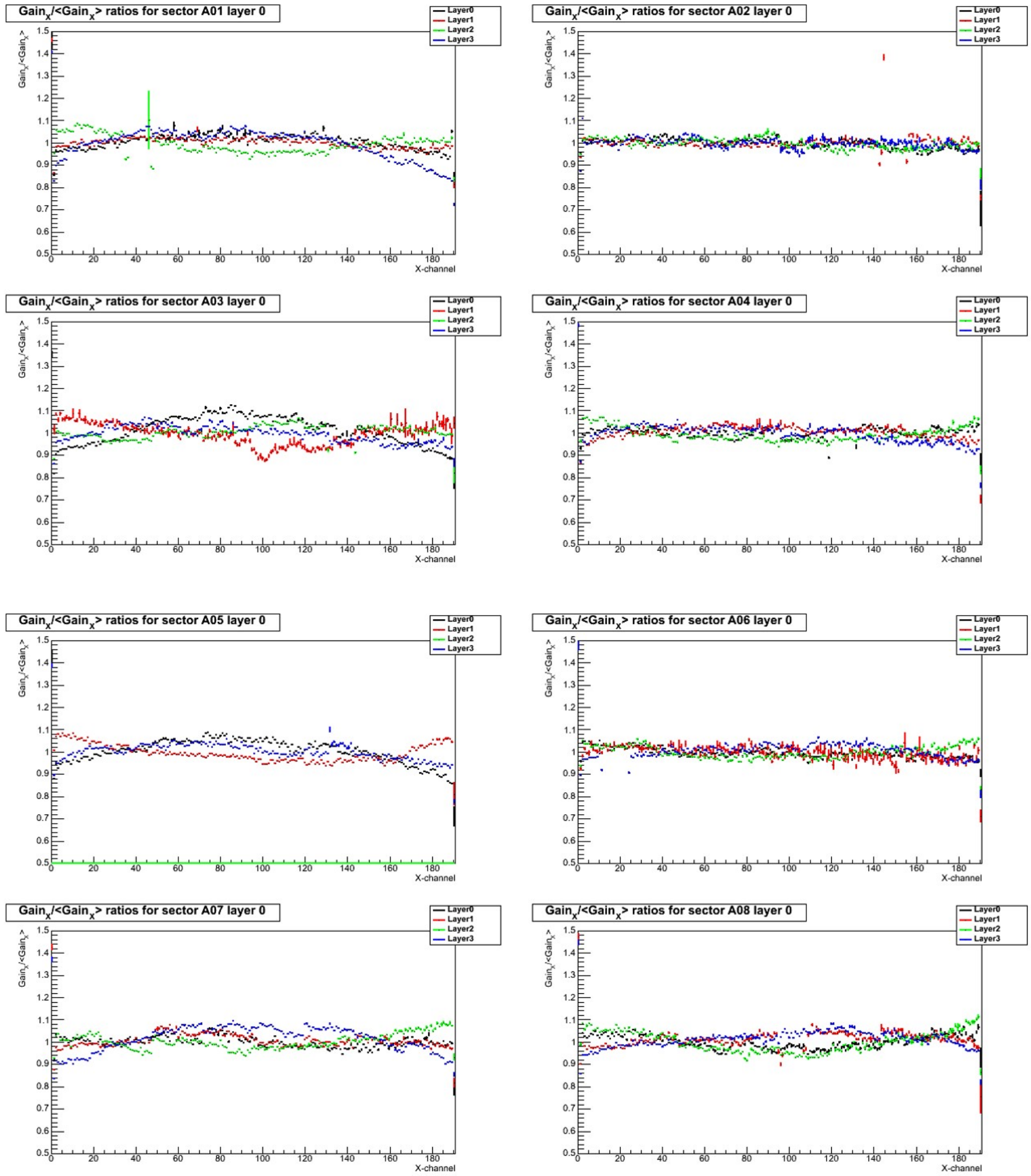


Figure 3.12: Relative gain estimate (R) for each precision strip (sectors A01 - A08). Dead or problematic channels are not used in the measurements.

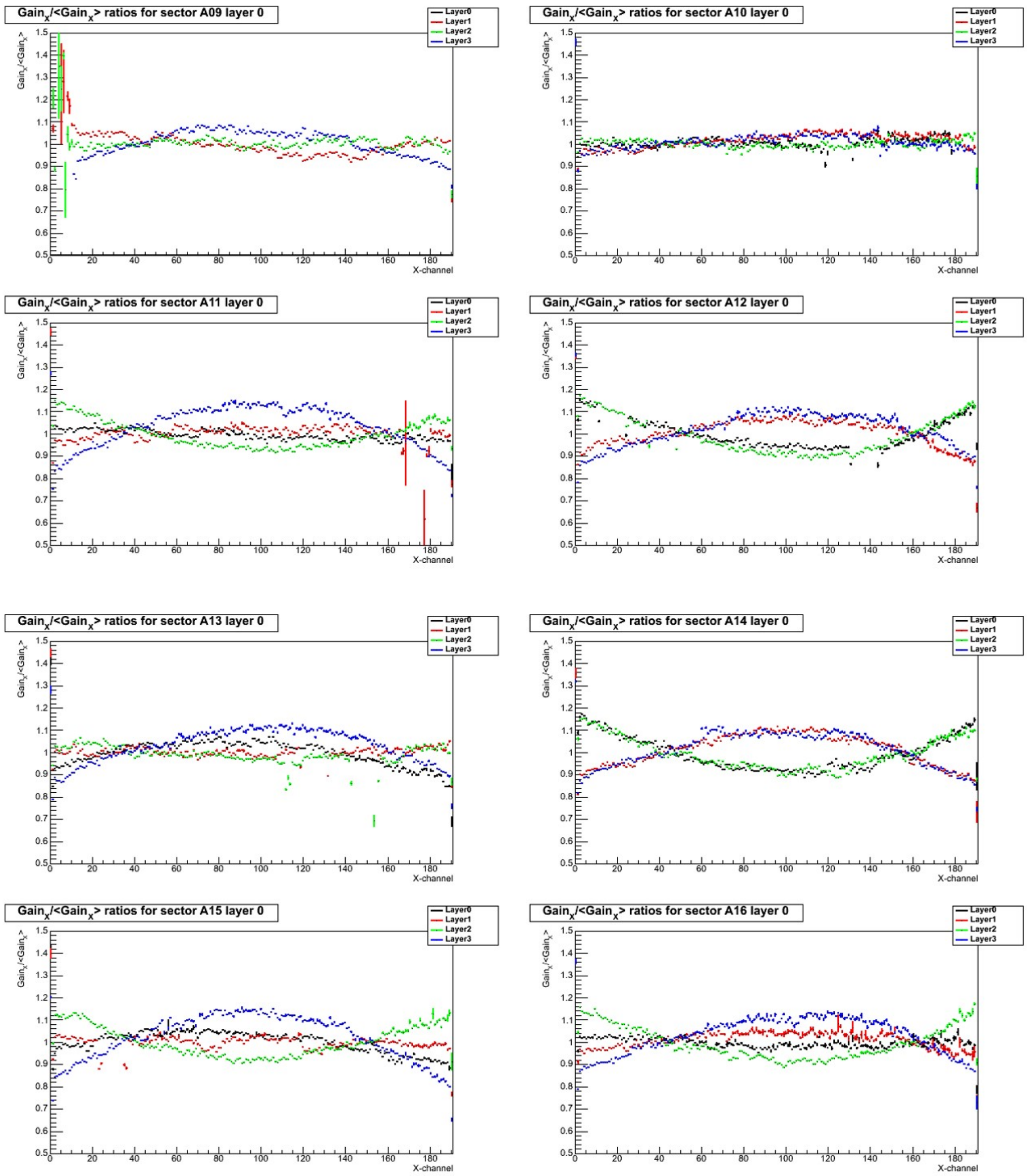


Figure 3.13: Relative gain estimate (R) for each precision strip (sectors A09 - A16). Dead or problematic channels are not used in the measurements.

Chapter 4

Search for the SM Higgs boson

The search for the SM Higgs boson, in the decay channel $H \to ZZ \to \ell^+\ell^- q\bar{q}$ ($\ell = e, \mu$), is well motivated since this final state has a large branching fraction compared to other interesting Higgs final states (see Section 1.4) and, in addition, it allows full reconstruction of the Higgs mass. On the other hand, it is a challenging study due to the high rates of background processes that give rise to signatures similar to the signal. These are mainly Z+jets processes, in which the Z boson decays leptonically, with secondary contributions from top quark and QCD multijet production.

The search in the Higgs boson mass range 120–180 GeV [37] uses data collected with the ATLAS detector during the 2011 LHC run, at a centre-of-mass energy of 7 TeV. The analysis strategy is to look for candidate events containing a pair of leptons and a pair of high-p_{T} jets, with each pair resembling a Z boson decay. Of course, it is taken into account that in the mass range under study, only one of the Z bosons may typically be on-shell. Due to the overwhelming background from real-Z+jets production, it is convenient to restrict the analysis to the region where the dilepton mass is well below the Z resonance while the dijet mass is consistent with the nominal Z boson mass. By imposing an upper limit on the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, the analysis also exploits the absence of neutrinos in the signal final state to reduce the background from $t\bar{t}$ processes. Events are finally classified into two categories; those with 2 b-tagged jets ("tagged") and those with 0 or 1 b-tagged jets ("untagged"), in order to benefit from the lower rate of Z+jets processes in the former. The conclusions of the search are drawn by looking for a peak above the background expectation in the invariant mass of the $\ell\ell jj$ system, $m_{\ell\ell jj}$.

4.1 Data and Monte Carlo samples

A number of quality criteria ensure the operation of the LHC with stable beams and that the ATLAS detector was operational with good efficiency while the data were being collected. These criteria are implemented using a "Good Runs List" (GRL) that defines the luminosity blocks in which the required quality conditions are met. The filtered data set corresponds to an integrated luminosity of 4.71 fb^{-1} [38].

Trigger requirements use both single-lepton and dilepton non-prescaled triggers with the lowest available p_{T} (muons) and E_{T} (electrons) thresholds. Thresholds of 18 GeV and 20 GeV were employed for the single muon and the single electron trigger respectively. The latter was increased to 22 GeV for the second half of the data (characterized by a higher instantaneous luminosity). The dilepton triggers require two same-flavour leptons with a threshold of 10 GeV for muons and 12 GeV for electrons.

Monte Carlo (MC) samples simulating the signal as well as background processes

m_H (GeV)	Gluon fusion σ (pb)	Vector-boson fusion σ (pb)	$\Gamma_{\ell\ell qq}/\Gamma_H$ (%)	Total $\sigma\cdot$ BR (fb)
120	16.65	1.2690	0.222	39.8
125	15.32	1.2110	0.370	61.2
130	14.16	1.1540	0.559	85.6
135	13.11	1.1000	0.789	112
140	12.18	1.0520	0.965	128
145	11.33	1.0040	1.110	137
150	10.58	0.9617	1.160	134
160	9.202	0.8787	0.583	59.8
170	7.786	0.8173	0.332	28.6
180	6.869	0.7480	0.847	64.5

Table 4.1: Cross sections in fb for the Monte Carlo samples modelling the $H \to ZZ \to \ell^+\ell^- q\bar{q}$ signal process. The cross sections have been evaluated from theoretical calculations [62].

are generated with 7 TeV center-of-mass energy and passed through the simulation of the ATLAS detector [39]. To account for any differences in the distribution of pile-up interactions (see Section 4.1.3) between MC simulation and data, the MC samples are reweighted accordingly.

4.1.1 Signal samples

Signal samples simulating $H \to ZZ \to \ell^+\ell^- q\bar{q}$, where $\ell = e, \mu, \tau$ and $q = d, u, s, c, b$, are generated using POWHEG [40–44], interfaced to PYTHIA [45] for showering and hadronization, which includes matrix elements up to next-to-leading order (NLO) in α_S. Both gluon fusion (ggF) and vector-boson fusion (VBF) production mechanisms are taken into account by the matrix elements. The generated signal samples correspond to luminosities that are much larger than the available luminosity in the data.

The SM signal cross-sections have been computed up to next-to-next-to-leading order (NNLO) [46–51] in α_S for the ggF process. NLO electroweak (EW) corrections are also applied [52, 53], as well as QCD soft-gluon resummations up to next-to-next-to-leading logs (NNLL) [54]. These calculations are detailed in Refs. [55–57] and assume factorization between QCD and EW corrections. Full NLO QCD and EW corrections [58–60] and approximate NNLO QCD corrections [61] are used to calculate the cross-sections for VBF signal production. [62].

4.1.2 Background samples

4.1.2.1 Z+jets production

Background samples for $Z \to \ell\ell$ ($\ell = e, \mu, \tau$), with $m_{\ell\ell} > 40$ GeV, are simulated using ALPGEN [63], which generates hard matrix elements for Z and $Zb\bar{b}$ production with a fixed number of $N_p(0 \to 5)$ additional partons in the final state. Dedi-

Process	σ(fb)
$Z + 0p, Z \to \ell\ell$	836 000
$Z + 1p, Z \to \ell\ell$	168 000
$Z + 2p, Z \to \ell\ell$	50 500
$Z + 3p, Z \to \ell\ell$	14 000
$Z + 4p, Z \to \ell\ell$	3 510
$Z + 5p, Z \to \ell\ell$	988
$Zb\bar{b} + 0p, Z \to \ell\ell$	8 208
$Zb\bar{b} + 1p, Z \to \ell\ell$	3 100
$Zb\bar{b} + 2p, Z \to \ell\ell$	1 113
$Zb\bar{b} + 3p, Z \to \ell\ell$	488

Table 4.2: Cross sections for Z + jets samples, generated using ALPGEN, with $m_{\ell\ell} > 40$ GeV. The cross sections are corrected to NLO in α_S and are assumed to be the same for $Z \to ee$, $Z \to \mu\mu$ and $Z \to \tau\tau$ production.

cated samples for $Z(\to \ell\ell)+b\bar{b}$ processes with $N_p(0 \to 3)$ additional partons are also available. In order to remove double counting of events between the inclusive and the heavy-flavour samples (overlap removal), events in which a pair of b-quarks is generated during parton showering are removed from the inclusive samples if their distance in the $\eta - \phi$ plane is $\Delta R > 0.4$; conversely, such events are removed from the heavy-flavour samples if $\Delta R < 0.4$.

The Z+jets samples are generated at LO in α_S but the production cross-sections are corrected to match NLO calculations [64, 65]. Specifically, a k-factor of 1.22 is applied to the inclusive Z+jets production, while an extra factor of 1.4 is applied on top of it in the case of heavy-flavour samples. The cross sections for the Z+jets samples are listed in Table 4.2.

4.1.2.2 Low mass Drell–Yan + jets production

Since the analysis requires a pair of same–flavour, opposite–charge leptons with invariant mass lower than the nominal Z-boson mass, a significant background arises from low mass Drell–Yan $\ell^+\ell^-$ hadro-production accompanied by multiple jets. ALPGEN is used for the simulation of these processes as well and is tuned to restrict the dilepton mass within $10 < m_{\ell\ell} < 40$ GeV, thereby extending the minimum of the Z+jets samples from 40 to 10 GeV. The region $m_{\ell\ell} < 10$ GeV is prohibitive for this analysis since it is contaminated by decays of Y and J/ψ mesons. As in the case of the Z+jets samples, possible overlaps between the inclusive and the heavy-flavour samples are removed using the overlap removal procedure. The cross sections for the DY+jets samples are listed in Table 4.3. The same k-factors applied for the case of Z+jets are also applied here.

Process	σ(fb)
DY+0p	3 723 000
DY+1p	107 000
DY+2p	50 500
DY+3p	10 200
DY+4p	2 260
DY+5p	561
DY$b\bar{b}+0p$	20 260
DY$b\bar{b}+1p$	3 160
DY$b\bar{b}+2p$	1 180
DY$b\bar{b}+3p$	566

Table 4.3: Cross sections for the low mass DY+jets samples generated using ALPGEN, with $10 < m_{\ell\ell} < 40$ GeV. The cross sections are corrected to NLO in α_S and are assumed to be the same for ee, $\mu\mu$, and $\tau\tau$ production.

4.1.2.3 Top pair and single top production

Samples of top-quark pair $t\bar{t}$, as well as single top and Wt processes are generated with MC@NLO [66], interfaced to JIMMY4.31 [67] for the simulation of the underlying event. The $t\bar{t}$ sample is filtered, requiring that at least one lepton with $p_T > 1$ GeV originates from a W boson. This ensures that only events with at least one leptonic (e, μ, τ) W boson decay are retained. The cross sections for the top-quark background samples are listed in Table 4.4. They are evaluated up to NLO in α_S [68–71] and convoluted with branching fractions taken from the Particle Data Book [72].

channel	σ (fb)	filter	σ_{filtered} (fb)
$t\bar{t}$	164 600	0.5562	91 551
single t (s-chan, $W \to \ell\nu$)	462	–	–
single t (s-chan, $W \to \mu\nu$)	455	–	–
single t (s-chan, $W \to \tau\nu$)	484	–	–
single t (t-chan, $W \to \ell\nu$)	7 117	–	–
single t (t-chan, $W \to \mu\nu$)	6 997	–	–
single t (t-chan, $W \to \tau\nu$)	7 448	–	–
Wt	14 600	–	–

Table 4.4: Cross sections for the $t\bar{t}$ sample in the lepton-hadron (ℓh) or lepton-lepton ($\ell\ell$) decay mode and for the single-top and Wt samples. All samples are generated with the MC@NLO event generator.

4.1.2.4 Diboson production

The background from ZZ production is largely irreducible, albeit very small, since it gives rise to the same final state as the signal process. The contributions from WZ and WW production are expected to be minimal. Diboson MC samples are generated using the MC@NLO event generator. The production cross-sections have been calculted up to NLO in α_S [73] and are listed in Table 4.5.

Channel	σ (fb)
$ZZ \to \ell\ell qq$	841.5
$ZZ \to \ell\ell\nu\nu$	160.4
$ZZ \to \ell\ell\ell\ell$	27.0
$ZZ \to \ell\ell\tau\tau$	27.0
$ZZ \to \tau\tau\tau\tau$	6.8
$ZZ \to \tau\tau\nu\nu$	80.3
$WW \to \ell\nu\ell\nu$	2012
$WW \to \ell\nu\tau\nu$	2012
$WW \to \tau\nu\tau\nu$	503
$W^+Z \to \ell\nu qq$	1688.9
$W^+Z \to \ell\nu\ell\ell$	159.2
$W^+Z \to qq\ell\ell$	489.4
$W^+Z \to \tau\nu\ell\ell$	79.6
$W^+Z \to \ell\nu\tau\tau$	79.6
$W^+Z \to \tau\nu\tau\tau$	39.8
$W^+Z \to qq\tau\tau$	249.2
$W^-Z \to \ell\nu qq$	912.6
$W^-Z \to \ell\nu\ell\ell$	86.1
$W^-Z \to qq\ell\ell$	269.3
$W^-Z \to \tau\nu\ell\ell$	43.0
$W^-Z \to \ell\nu\tau\tau$	43.0
$W^-Z \to \tau\nu\tau\tau$	21.5
$W^-Z \to qq\tau\tau$	134.7

Table 4.5: Cross sections for the diboson samples (where $\ell = e, \mu$), evaluated for the dilepton mass range $66 < m_{\ell\ell} < 116$ GeV from theoretical calculations [73] and convoluted with Z boson branching fractions from the Particle Data Book [72].

4.1.2.5 Inclusive W boson production

$W(\to \ell\nu)$+jets ($\ell = e, \mu, \tau$) processes are simulated with the Alpgen, which generates hard matrix elements for W, Wc and $Wb\bar{b}$ production with a fixed number of $N_p(0 \to 5)$ additional partons in the final state. Dedicated samples for $W(\to \ell\nu) + b\bar{b}$ and $W(\to \ell\nu) + c$ processes with $N_p(0 \to 3)$ additional partons are also available. As in the case of the Z+jets samples, possible overlaps between the inclusive and the heavy-flavour samples are removed using the overlap removal procedure.

Cross sections for the W+jets samples are listed in Table 4.6. A k-factor of 1.20 is applied to make the inclusive W boson production cross section agree with NLO calculations [64].

Process	σ(fb)
$W + 0p, W \to e\nu$	8 300 000
$W + 1p, W \to e\nu$	1 560 000
$W + 2p, W \to e\nu$	453 000
$W + 3p, W \to e\nu$	122 000
$W + 4p, W \to e\nu$	30 900
$W + 5p, W \to e\nu$	8 380
$Wb\bar{b} + 0p$	56 800
$Wb\bar{b} + 1p$	42 900
$Wb\bar{b} + 2p$	20 800
$Wb\bar{b} + 3p$	7 960
$Wc\bar{c} + 0p$	153 000
$Wc\bar{c} + 1p$	126 000
$Wc\bar{c} + 2p$	62 500
$Wc\bar{c} + 3p$	20 400
$Wc + 0p$	518 000
$Wc + 1p$	192 000
$Wc + 2p$	51 000
$Wc + 3p$	11 900

Table 4.6: Cross sections for the W + jets samples generated using ALPGEN. The cross sections listed include a k-factor of 1.20 and are assummed to be the same for $W \to e\nu$, $W \to \mu\nu$ and $W \to \tau\nu$ production.

4.1.3 Pile-up reweighting for Monte Carlo samples

Pile-up is an important effect to this analysis since spurious jets from pile-up interactions may be either mistaken as jets of the primary event or superimposed to them, thereby increasing the jet energy and the measured $E_{\mathrm{T}}^{\mathrm{miss}}$. There are two sources of pile-up; “in-time” and “out-of-time”. In-time pile-up is caused by additional proton–proton interactions occurring at the same bunch crossing as the hard interaction of interest. Such interactions produce extra soft particles which reduce the efficiency for selecting signal events. Out-of-time pile-up is generated from interactions that occur in preceding bunch-crossings and affect detector sub-systems whose latency is longer than the bunch spacing.

During 2011, the number of interactions per bunch crossing varied due to changes in the beam intensity and transverse size. Monte Carlo samples assume a 50 ns bunch spacing, as was the case for the vast majority of the data used in this search[1]. To model the effects of pile-up, a fixed distribution of additional minimum-bias interactions is used in the simulation which is then reweighted to the distribution observed in the data.

4.2 Lepton and Jet selection

The analysis uses reconstructed physics objects available in the ATLAS datasets. Necessary criteria are applied ensuring the quality of the objects reconstruction and additional requirements are imposed in order to focus around the kinematic region that describes the signal final state. Any discrepancies in the description of the data, related to the objects selection, are corrected for by applying appropriate smearing and efficiency corrections to the Monte Carlo.

4.2.1 Muons

Muons are identified using the STACO reconstruction chain [74]. Tracks, reconstructed in the muon spectrometer (MS) [28], are extrapolated to the beam pipe and an attempt is made to find a matching inner detector (ID) track. If such a match is found then a "Combined" (CB) muon is formed. Such muons have two independent momentum measurements and thus have the best momentum resolution. Remaining ID tracks are extrapolated to the MS and are "Segment-Tagged" (ST) as muons if they can be matched with a track segment in at least one muon station. Such muons have their momenta measured by the inner detector and are used to increase the acceptance in cases where a muon crossed only one layer of chambers in the MS, either due to having low p_{T} or due to being in a MS region with reduced acceptance. Both CB and ST muons are used in the analysis.

The muon selection criteria are listed in Table 4.7. Inner detector tracks associated to muons, are required to satisfy a series of quality criteria, based on the number of hits and absence thereof (holes), in the various layers of the inner detector[2]. Detector conditions, such as dead modules, are taken into account. Subsequently, muons from cosmic rays are filtered out by requiring that muon tracks pass close to the reconstructed primary event vertex, both longitudinally (z_0) and in the transverse plane (d_0). Furthermore, muons must be isolated; track-based isolation is defined as the ratio of the transverse track-momenta, within a cone $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.2$ of the muon (excluding the muon track itself), over the transverse momentum of the muon. Finally, since the background from QCD multijet processes (mostly $b\bar{b}$-pair production) is significant, even after track isolation has been imposed, further background rejection is required by imposing a cut on the significance of the transverse impact parameter, $|d_0|/|\sigma_{d_0}|$.

Muon momenta in Monte Carlo simulation are smeared to match the resolution measured in the data. Weights are also applied to account for differences in efficiency. The p_{T} and η distributions of muons which have been identified as the products of

[1] At the beginning of 2011, 12 pb^{-1}of the data were taken with a 75 ns bunch spacing.
[2] These criteria are recomended by the Muon Combined Performance groups of ATLAS.

Identification	Combined or segment-tagged STACO muons
Kinematic	$p_{\mathrm{T}} > 7$ GeV $\lvert\eta\rvert < 2.5$
Inner Detector	$N_{\mathrm{hits}}^{\text{b-layer}} > 0$ (except for muons passing uninstrumented/dead areas) $N_{\mathrm{hits}}^{\mathrm{pixel}} + N_{\mathrm{dead}}^{\mathrm{pixel}} > 1$ $N_{\mathrm{hits}}^{\mathrm{SCT}} + N_{\mathrm{dead}}^{\mathrm{SCT}} \geq 6$ $N_{\mathrm{holes}}^{\mathrm{pixel}} + N_{\mathrm{holes}}^{\mathrm{SCT}} < 3$ $\lvert\eta\rvert < 1.9$: $N_{\mathrm{tot}}^{\mathrm{TRT}} > 5$ and $N_{\mathrm{outliers}}^{\mathrm{TRT}} < 0.9 \times N_{\mathrm{tot}}^{\mathrm{TRT}}$ $\lvert\eta\rvert \geq 1.9$: If $N_{\mathrm{tot}}^{\mathrm{TRT}} > 5$, require $N_{\mathrm{outliers}}^{\mathrm{TRT}} < 0.9 \times N_{\mathrm{tot}}^{\mathrm{TRT}}$, where $N_{\mathrm{tot}}^{\mathrm{TRT}} = N_{\mathrm{hits}}^{\mathrm{TRT}} + N_{\mathrm{outliers}}^{\mathrm{TRT}}$.
Cosmic rejection	$\lvert d_0\rvert < 1$ mm $\lvert z_0\rvert < 10$ mm
Track isolation	$\sum_{\mathrm{tracks}} p_{\mathrm{T}}(\Delta R < 0.2)/p_{\mathrm{T}}^{\mu} < 0.1$
multijet suppression	$\lvert d_0\rvert/\lvert\sigma_{d_0}\rvert < 3.5$

Table 4.7: Summary of muon selection criteria. N_{hits} (N_{holes}) represents the number of hits (missing hits) in a particular subdetector of the inner tracker, while N_{dead} refers to the number of dead sensors crossed by the muon in a particular subdetector.

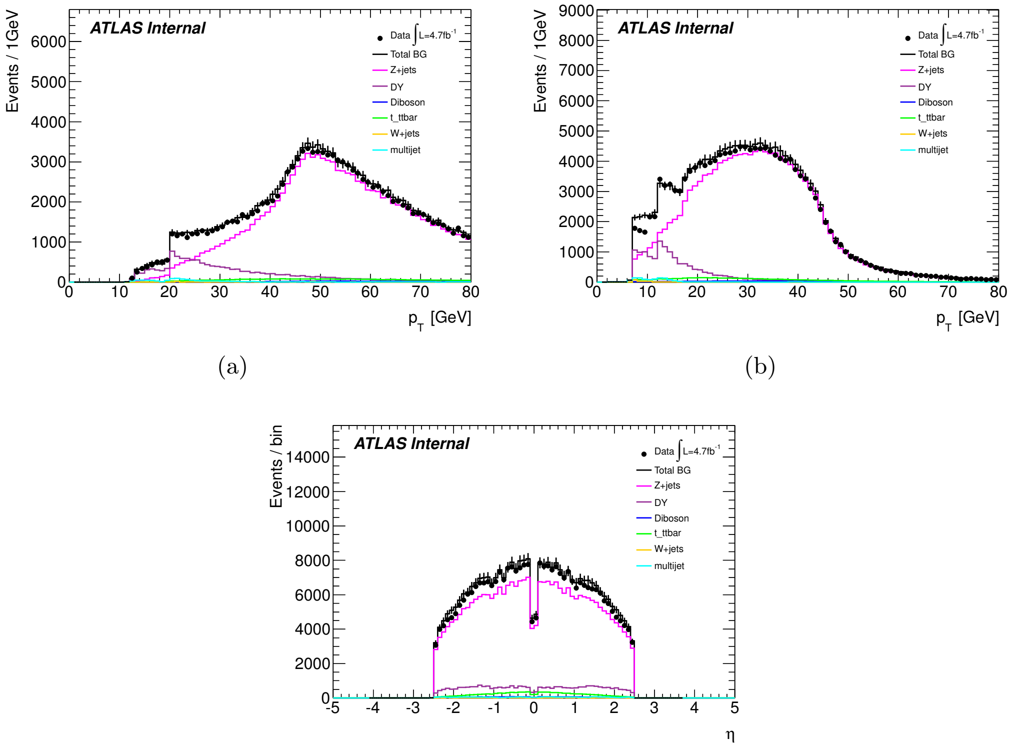
$Z \to \mu^+\mu^-$ decays are shown in Figure 4.1 for events containing at least two jets[3].

4.2.2 Electrons

Electrons are identified from electromagnetic calorimeter clusters [28] that are matched to tracks in the inner detector. Candidates are required to satisfy the tightest of the standard ATLAS identification criteria. These criteria are summarized in Table 4.8; they depend on the shape of the EM cluster, the quality of the track and the goodness of the match between the cluster and the track. Employing tighter identification criteria reduces significantly the multijet background. The energy of the electron is measured from the cluster in the calorimeter, while the direction is taken from the matching track. The pseudorapidity measurement from the cluster is considered only when the candidate's position with respect to the calorimeter is required; for instance, in order to achieve high reconstruction and trigger efficiency, candidates are required to lie within the region of precision EM measurement, $\lvert\eta_{\mathrm{cluster}}\rvert < 2.47$, and have transverse momentum $p_{\mathrm{T}} > 7$ GeV.

The electron selection criteria are listed in Table 4.9. Electron candidates must be isolated, using track-based isolation as defined for the case of muons. Since the

[3]In this and all subsequent plots, the background has been normalized to the luminosity of the data, unless otherwise specified. Modelling corrections have also been applied to the background, as described in Section 4.4.



(a) (b)

(c)

Figure 4.1: p_{T} of the (a) leading and (b) sub-leading muon and (c) η distribution of muons used to reconstruct Z boson candidates. The distributions are taken from events containing at least 2 jets.

background from multijet QCD processes (fake electrons, photon conversions and $b\bar{b}$-pair production) is significant even after the isolation criteria, further rejection is attempted by requiring small impact parameter significance. Finally, fake electrons from final state radiation of muons are removed by requiring that no muon (as defined in Section 4.2.1) lies within $\Delta R < 0.2$ of the electron track.

In order to achieve accurate description of the data, electron candidates from Monte Carlo samples have additional smearing applied. Differences in identification and reconstruction efficiencies between data and Monte Carlo are also parameterized as a set of p_{T} and η dependent scale factors used to weigh the Monte Carlo simulation. The p_{T} and η distributions of electrons which have been identified as the products of $Z \to ee$ decays are shown in Figure 4.2 for events containing at least two jets.

4.2.3 Jets

Jets are reconstructed from topological clusters [75] using the anti-k_T algorithm [76] with a radius parameter $R = 0.4$. The topological clusters are corrected from the electromagnetic scale, which is established using test beam measurements, to hadronic

Type	Description	Variable name
	loose++	
Shower shape	Shower width in the strips.	$w_{s,tot}$
	Difference in energy between the cells in the first and second maxima.	E_{ratio}
	Ratio of E_{T} in the first sampling of the hadronic calorimeter to E_{T} of the EM cluster.	R_{had1}
	Ratio in η of cell energies in 3×7 versus 7×7 cells in the 2^{nd} layer of the EM calorimeter.	R_{η}
	Lateral width of the shower in the 2^{nd} layer of the EM calorimeter.	w_2
Track matching	$\lvert\Delta\eta\rvert < 0.015$.	
	medium++	
Shower shape	Same variables as *loose++* but at tighter values.	
Track matching	$\lvert\Delta\eta\rvert < 0.005$.	
Track quality	Number of hits in the B-layer (≥ 1) for $\lvert\eta\rvert < 2.01$.	
	Number of hits in the pixel detector (> 1) for $\lvert\eta\rvert > 2.01$.	
	Ratio of high-threshold hits in the TRT over total hits in the TRT.	
	$\lvert d_0\rvert < 5$ mm.	
	tight++	
Shower shape	Same variables as *medium++* but at tighter values.	
Track matching	$\lvert\Delta\eta\rvert < 0.005$.	
	$\lvert\Delta\phi\rvert$ requirement.	
	E/p requirement.	
Track quality	Number of hits in the B-layer (≥ 1) for all $\lvert\eta\rvert$.	
	Number of hits in the pixel detector (> 1) for $\lvert\eta\rvert > 2.01$.	
	Tighter cuts on the ratio of high-threshold hits in the TRT.	
	Conversion bit (not matched to a conversion vertex).	

Table 4.8: Summary of electron identification criteria.

energy scale, using a p_{T} and η dependent jet energy scale (JES), determined from Monte Carlo simulation [77, 78]. Candidates are required to have $p_{\mathrm{T}} > 20$ GeV and are restricted within $|\eta| < 2.5$ that corresponds to the fiducial volume of the ATLAS tracker.

The jet selection criteria are listed in Table 4.10. The majority of jets originating from pile-up are removed by requiring that the fraction of the jet's transverse track-momentum contributed by tracks originating from the primary vertex, be at least 75%. This is implemented as a cut on the absolute value of the "jet vertex fraction" [79], $|\mathrm{JVF}| > 0.75$. To remove jets which do not originate from real, in-time energy deposits but rather from hardware problems, cosmic-ray showers or LHC beam conditions, it is demanded that candidates pass the "looser" standard ATLAS set of jet quality criteria. Finally, to avoid double-counting of objects in an event, a jet is removed if an electron (as defined in Section 4.2.2) is found within a cone $\Delta R = 0.4$ of the jet axis. The p_{T} and η distribution of jet candidates is shown in Figure 4.3.

4.2.3.1 Identification of *b*-jets

For the discrimination of jets containing decays of *b*-hadrons from those containing only light quarks, the ATLAS *b*-tagging algorithms exploit the long lifetime of *b*-hadrons ($c\tau \approx 450$ μm). In this analysis, several algorithms are used which accept

Identification	Author: Electron IsEM: *tight++*
Kinematic	$p_{\rm T} > 7$ GeV $\lvert\eta_{\rm cluster}\rvert < 2.47$
Track isolation	$\sum_{\rm tracks} p_{\rm T}(\Delta R < 0.2)/p_{\rm T}^{e} < 0.1$
multijet suppression	$\lvert d_0\rvert/\lvert\sigma_{d_0}\rvert < 6.5$

Table 4.9: Summary of electron selection criteria.

Identification	Anti-k_T $R = 0.4$ topological jets
Kinematic	$p_{\rm T} > 20$ GeV $\lvert\eta\rvert < 2.5$
Pile-up	$\lvert{\rm JVF}\rvert > 0.75$

Table 4.10: Summary of jet selection criteria.

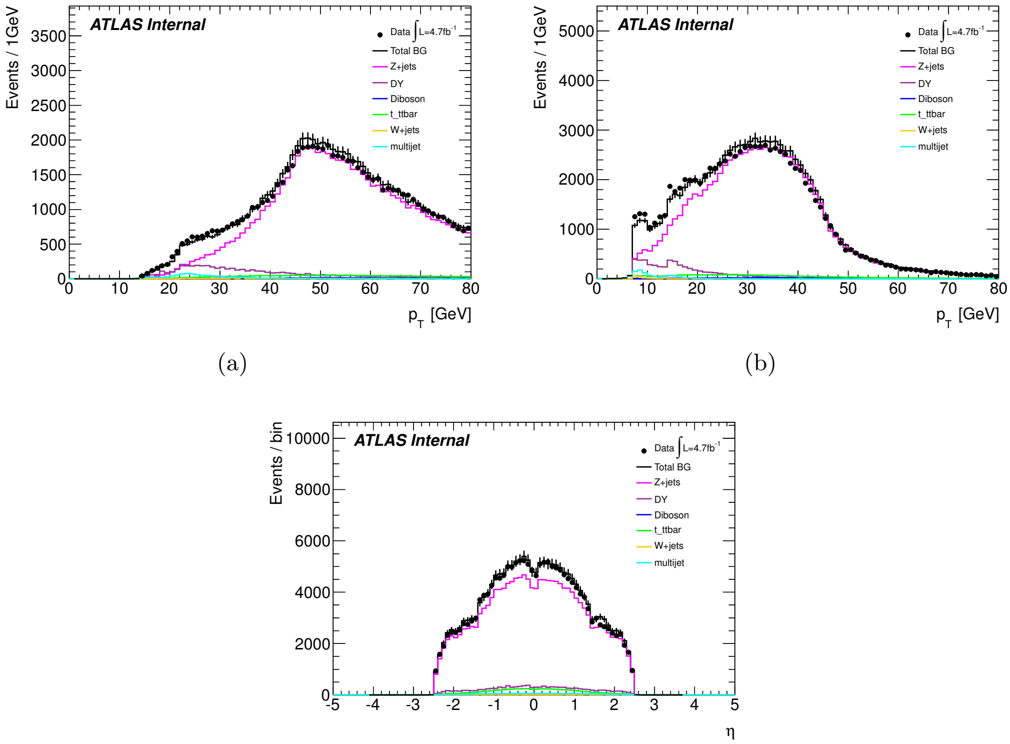
as input the set of tracks within a cone of $p_{\rm T}$-dependent ΔR around the jet axis:

- `IP3D` looks for tracks with a significant three-dimensional impact parameter with respect to the primary event vertex.
- `SV1` attempts to reconstruct secondary vertices, displaced from the primary vertex using the jet's tracks.
- `JetFitter` is an inclusive secondary vertex reconstruction algorithm that exploits the topology of the semi-leptonic b-and c-hadron decay cascade inside a jet by finding b- and c-vertices lying on a common line with the primary vertex.

The results of the three algorithms are combined using a multivariate method into a single discriminant henceforth referred to as `MV1` [80–82]. The selection is applied at the operating point that gives an efficiency of about 70%, on average, for identifying true b-jets, while the efficiencies for accepting c jets or light quark jets are 1/5 and 1/140 respectively. Figure 4.4 shows the light jet rejection as a function of the b jet tagging efficiency for different tagging algorithms, based on simulated top-antitop events. Finally, appropriate scale-factors are applied to the MC simulation in order to reproduce the b-tagging efficiency observed in the data.

4.2.4 Missing transverse energy

The missing transverse energy $E_{\rm T}^{\rm miss}$ is calculated from energy deposits in calorimeter cells within $|\eta| < 4.9$, and from muon measurements. Since $E_{\rm T}^{\rm miss}$ is used to indicate



(a) (b)

(c)

Figure 4.2: p_T of the (a) leading and (b) sub-leading electron and (c) η distribution of electrons used to reconstruct Z boson candidates. The distributions are taken from events containing at least 2 jets.

the presence of neutrinos in an event, it is an important quantity for the discrimination of signal from the top-quark background. To obtain a systematic uncertainty on the measured E_T^{miss} value, all systematic variations on object energies are propagated to the calculation of E_T^{miss}. Additional uncertainties on E_T^{miss} originate from variations of the energy scale and resolution of calorimeter clusters which are associated to reconstructed objects (the so-called "soft term").

4.3 Event selection

Triggered events are required to contain a reconstructed primary vertex formed by at least three tracks with $p_T > 150$ MeV. An "E_T^{miss}-cleaning" procedure follows in order to identify mismeasured jets (see Section 4.2.3) arising from hardware problems, cosmic-ray showers, and unstable LHC beam conditions. Such jets give rise to fake E_T^{miss} therefore, if an isolated low-quality jet with $p_T > 20$ GeV is found, the event is discarded. The selection then proceeds to the reconstruction of the $Z \to \ell\ell$ and $Z \to qq$ decays as elaborated in the sections that follow.

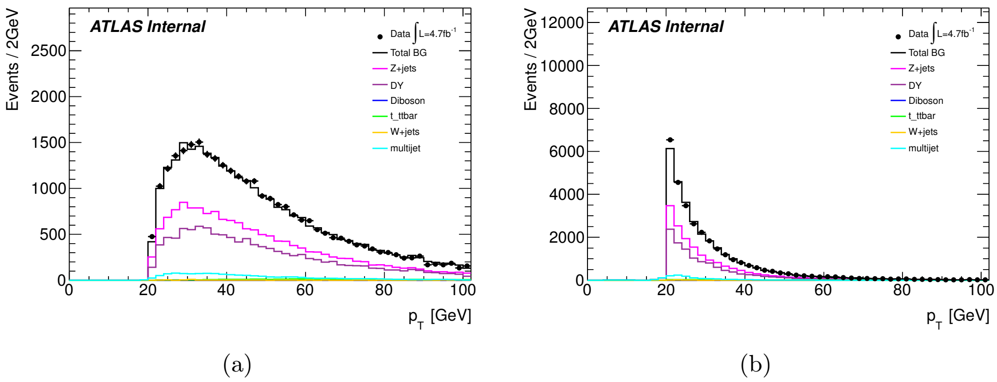


Figure 4.3: p_{T} of the leading (a) and sub-leading (b) jet, after the $Z \to \ell\ell$ boson mass selection and $E_{\mathrm{T}}^{\mathrm{miss}}$ cut, for the combined electron and muon channels.

4.3.1 $H \to ZZ \to \ell^+\ell^- q\bar{q}$ selection

Events are required to contain two electrons or muons with opposite charge sign. The opposite charge requirement reduces the multijet background, which produces both same charge and opposite charge lepton pairs. Events with additional leptons of either flavour are rejected in order to reduce WZ production.

The tranverse momenta of the selected leptons must agree with the actual trigger fired in each event. p_{T} greater than 12 (14) GeV is required for both muons (electrons) in cases where only the dilepton trigger is fired, while for the single lepton trigger, at least one lepton should have $p_{\mathrm{T}} > 20$ GeV. In case both triggers are fired, the union of the two conditions is used. The above cut values are chosen to lie on the triggers efficiency plateaus. Finally, the dilepton invariant mass $m_{\ell\ell}$ should lie within the range $20 < m_{\ell\ell} < 76$ GeV to suppress background events without a real Z boson. The distribution of $m_{\ell\ell}$ in the electron and muon channels is shown in Figure 4.5.

Since no high-p_{T} neutrinos are expected in the final state, the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be less than 30 GeV, for further reduction of the top quark background. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, displayed in Figure 4.6(b), shows that a missing transverse energy requirement reduces the top quark background significantly, therefore $E_{\mathrm{T}}^{\mathrm{miss}} < 30$ GeV is imposed. The event selection proceeds with the acquisition of jet candidates in order to reconstruct the $Z \to qq$ decay.

A significant fraction ($\sim$22%) of signal events are expected to contain b-jets, from $Z \to b\bar{b}$ decays [72]. On the other hand, b-jets are rare in the dominant $Z+$jets background (in which the Z boson decays to leptons). This feature is exploited in the analysis, by classifying candidate events into "tagged", with two b-tagged jets, and "untagged", with less than two b-tagged jets. Events with more than two b-tagged jets are discarded. The multiplicity distribution for b-tagged jets is shown in Figure 4.7(b).

Figure 4.6(a) shows the jet multiplicity distribution. In the tagged event category, the $Z \to q\bar{q}$ decay is reconstructed using the two b-tagged jets. For untagged events, the candidates are instead selected by using the kinematic fit described in

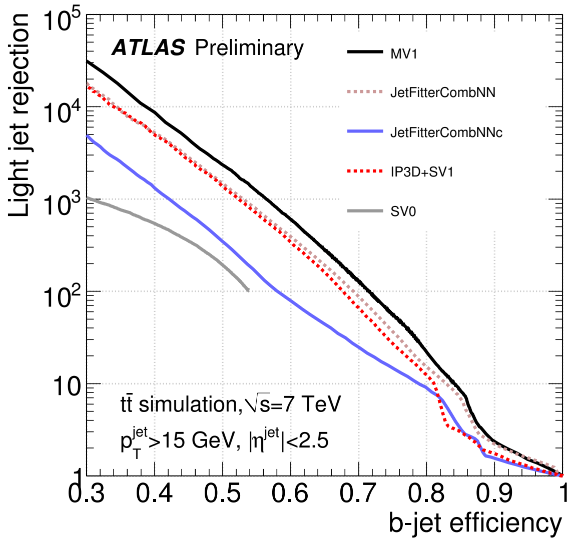


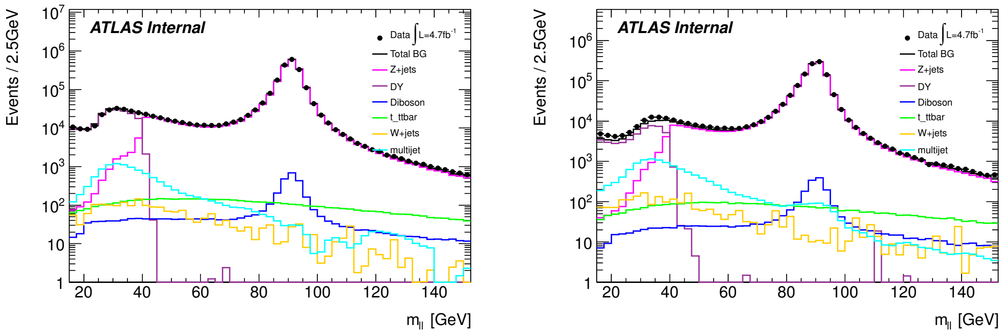
Figure 4.4: Expected performance of the various b-tagging algorithms in a simulated $t\bar{t}$ sample, for jets with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$ [81].

Section 4.3.2. In the case of muon events, additional reduction of the multijet background is achieved by requiring $\Delta R > 0.3$ between any of the muons and any of the selected jets. Finally, the dijet invariant mass m_{jj} is required to lie within the range $60 < m_{jj} < 115$ GeV. This criterion filters out most of the DY+jets background. The distribution of the dijet invariant mass m_{jj} is presented in Figure 4.8 for the untagged and tagged samples respectively. For the latter, the jet energies are scaled up by 5% to take into account the average energy scale difference between heavy- and light-quark jets [83].

The event selection criteria are summarized in Table 4.11. The efficiencies of the various criteria are listed in Appendix E. Summing up both muon and electron sample, the resulting total analysis selection efficiency for the $H \to ZZ \to \ell^+\ell^- q\bar{q}$ signal increases between 0.7% and 3% for the untagged selection and between 0.05% and 0.25% for the tagged. The efficiency is limited by the trigger selection at the low m_H side, reaching a maximum at $m_H \approx 170$ GeV and dropping somewhat at the high end, due to the requirement of an off-shell Z boson decaying to a lepton pair. These efficiencies include $Z \to \tau\tau$ decays. For the individual channels, taking into account only $Z \to \mu\mu$ and $Z \to ee$ decays, separately, the efficiencies are shown in Figure 4.9.

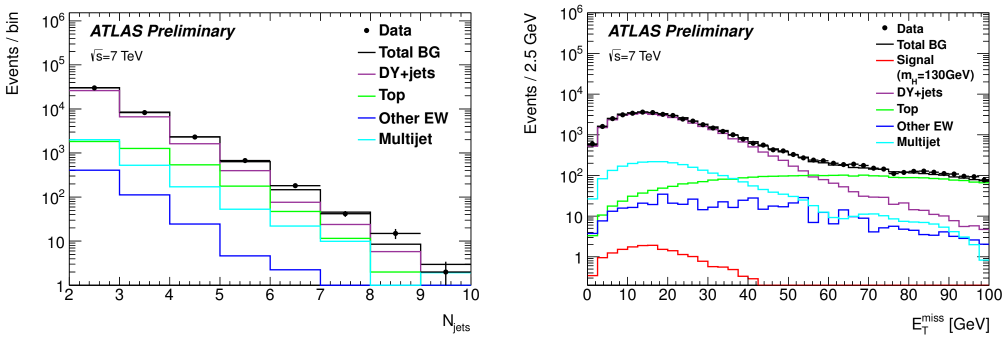
4.3.2 Kinematic Fit

The mass of the $H \to ZZ$ candidate is estimated from the invariant mass of the $\ell\ell jj$ system, $m_{\ell\ell jj}$. The resolution of $m_{\ell\ell jj}$ can be significantly improved by constraining the jet pair to the Z boson mass, since the reconstructed $Z \to jj$ mass resolution is worse than the intrinsic Z resonance width. A χ^2 is therefore constructed using the



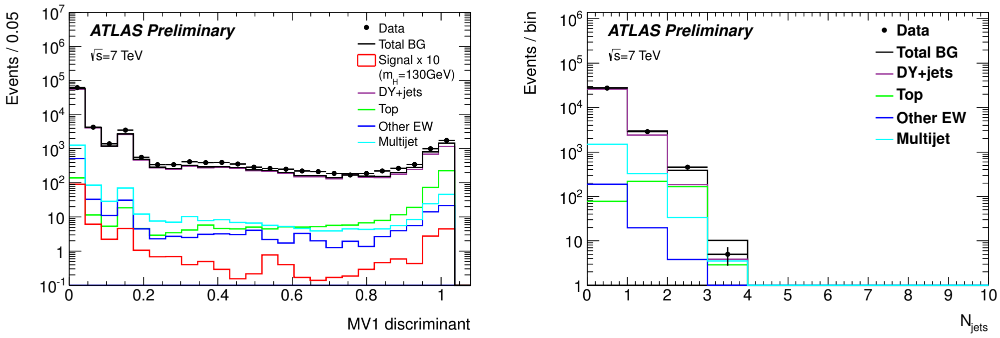
(a) (b)

Figure 4.5: $m_{\ell\ell}$ distribution in the (a) muon and (b) electron channel.



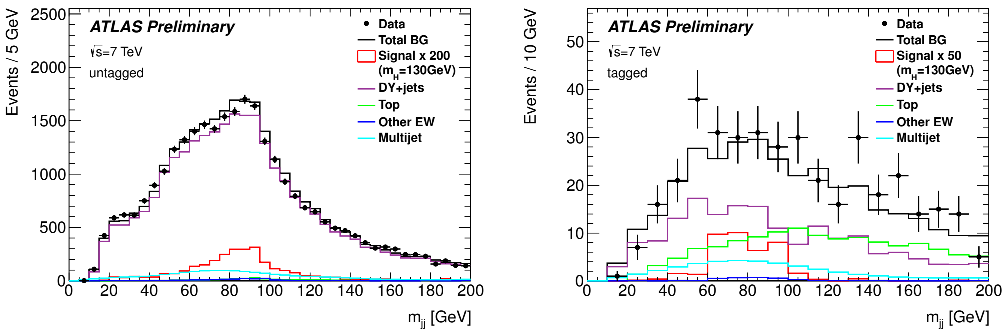
(a) (b)

Figure 4.6: (a) jet multiplicity and (b) the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in events with at least 2 selected jets, after the cut on the $Z \to \ell\ell$ boson mass. "Other EW", denotes all diboson and $W +$jets backgrounds.



(a) (b)

Figure 4.7: Distribution of (a) the b-tag multivariate discriminant for the selected jets and (b) the number of b-tagged jets.



(a) (b)

Figure 4.8: m_{jj} distribution after the missing energy requirement for the (a) untagged and (b) tagged samples for the combined electron and muon channels.

Criteria	Description
Pre-selection	Triggered event. Primary vertex formed by ≥ 3 tracks with $p_{\mathrm{T}} > 150$ MeV.
Leptons	Exactly 2 leptons (*tight*++ electrons/CB+ST muons). Opposite charge. Single-lepton trigger & $p_{\mathrm{T}}^{\ell 1} > 20$ GeV or double-lepton trigger & $p_{\mathrm{T}} > 12(14)$ GeV for both muons(electrons).
Jets *untagged* *tagged*	≥ 2 jet candidates < 2 b-tagged jets (selection by kinematic fit). 2 b-tagged jets (selection of b-tagged jets). Events with > 2 b-tagged jets are rejected.
	$E_{\mathrm{T}}^{\mathrm{miss}} < 30$ GeV. $20 < m_{\ell\ell} < 76$ GeV. $70 < m_{jj} < 105$ GeV.

Table 4.11: Summary of event selection criteria. The definitions of object candidates are given in Tables 4.7-4.10.

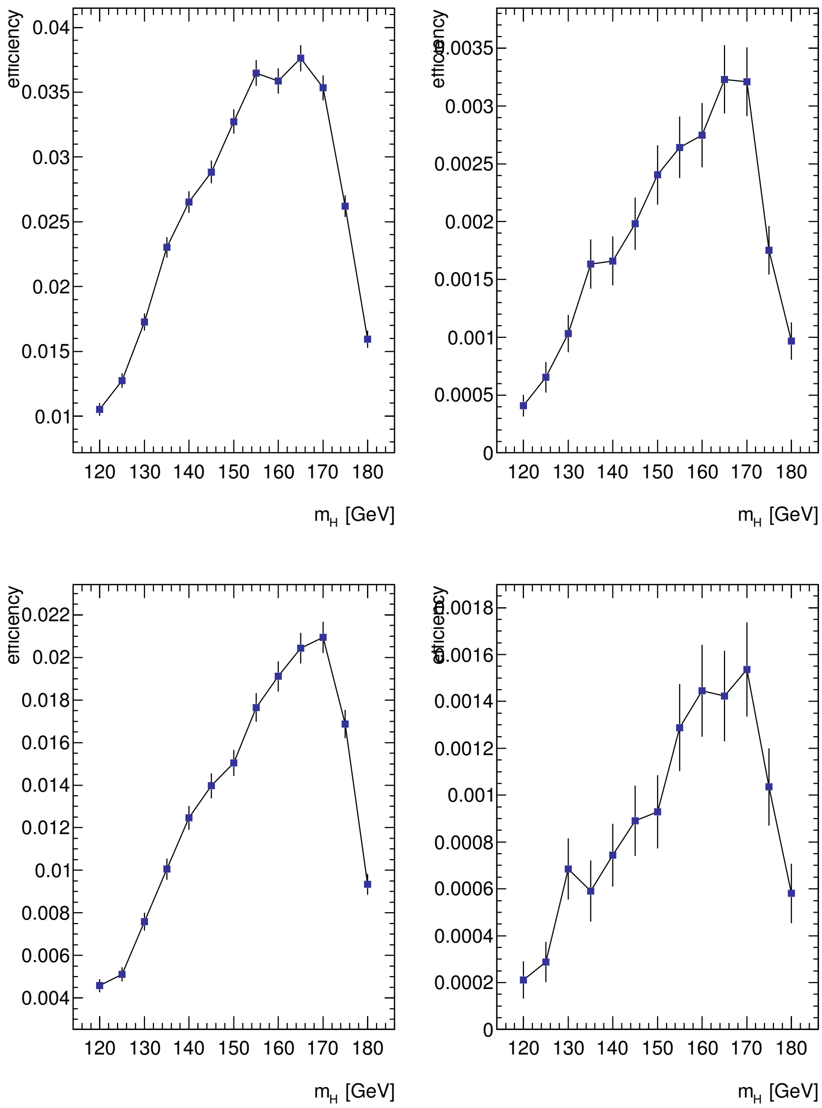


Figure 4.9: Total efficiency in the (top) muon and (bottom) electron channel for the untagged (left) and tagged (right) selection. The efficiency is limited by the trigger selection at the low m_H side, reaching a maximum at $m_H \approx 170$ GeV and dropping somewhat at the high end, due to the requirement of an off-shell Z boson decaying to a lepton pair.

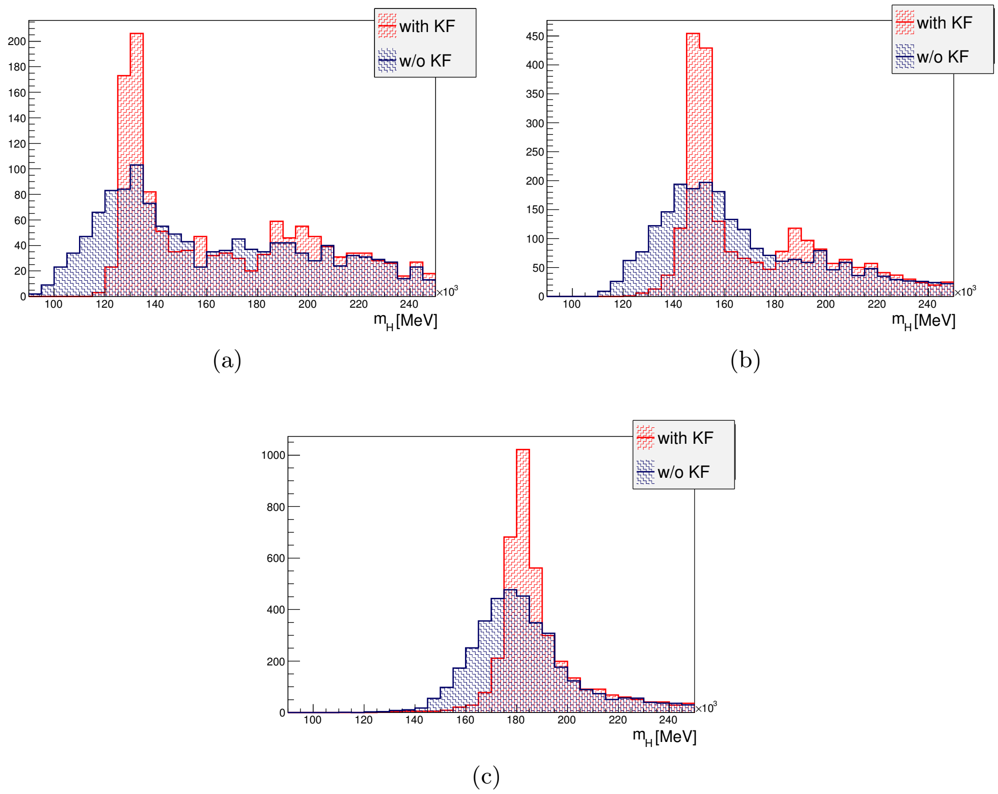


Figure 4.10: $m_{\ell\ell jj}$ distribution before and after the kinematic fit, for the Higgs boson masses (a) 130, (b) 150 and (c) 180 GeV

jets four-momenta, the jet energy resolution and the nominal Z-boson mass (m_Z) and width (Γ_Z):

$$\chi^2 = \left(\frac{m_Z - m_{jj}}{\Gamma_Z}\right)^2 + \sum_{i=1,2} \frac{(p_{\mathrm{Ti}} - p_{\mathrm{Ti}}^{fit})^2}{\sigma_{\mathrm{jet_i}}^2} \tag{4.1}$$

In the minimization of χ^2 the p_{T} of each jet is allowed to vary within the error $\sigma_{\mathrm{jet_i}}$ defined by the jet energy resolution.

This kinematic fit serves a dual purpose. First, it provides a criterion for chosing the jet pair that best matches a $Z \to q\bar{q}$ decay, i.e. the combination that gives the minimum χ^2. Furthermore, the fitted jet p_{T} are used in the calculation of $m_{\ell\ell jj}$ to improve the resolution of the reconstructed Higgs mass. The improvement of the $m_{\ell\ell jj}$ resolution observed in the signal is shown in Figure 4.10. The tail on the high mass end originates from events in which the dijet system is not produced by decays of on-shell Z bosons. For those events the kinematic fit is not effective and the resolution of $m_{\ell\ell jj}$ remains unchanged.

4.4 Background Modelling

The dominant background in this search is DY+jets production with multijet and top quark production contributing significantly. These backgrounds are determined using data driven methods. Smaller contributions are also expected from diboson (mainly ZZ) production as well as W+jets processes; those are taken directly from MC simulation.

4.4.1 DY + jets background

The modelling of the DY+jets background is studied using the sidebands (SBs) of the m_{jj} distribution which provide an almost pure sample of DY+jets events. The control region (CR) is defined by following the nominal selection, but replacing the nominal m_{jj} criterion, which defines the signal region (SR), with the union of $40 < m_{jj} < 60$ GeV (low-mass SB) and $115 < m_{jj} < 160$ GeV (high-mass SB). The SBs are statistically independent from the SR and chosen to be close to it, so that kinematics are kept similar, while the mass ranges are defined so that the low-mass and high-mass SBs provide similar event statistics.

The normalization of the DY+jets background is corrected by comparing the MC prediction to the data, in the CR, after subtracting the small contributions of the other background processes from the data. The measured scale factor is then extrapolated to the signal region. Since the DY+jets, the top quark and the multijet backgrounds are all normalized from the data, a correction to one of them affects the estimation of the other two. Therefore, the normalization of the three backgrounds is performed iteratively until the results stabilize.

The final scale factors for the DY + jets background are:

Untagged electron:	1.02 ± 0.03(stat) $\pm$ 0.02(syst)
Untagged muon:	0.99 ± 0.02(stat) $\pm$ 0.04(syst)
Tagged:	1.22 ± 0.13(stat) $\pm$ 0.12(syst)

The systematic uncertainty in the untagged sample receives contributions both from the subtraction of the multijet background from the data as well as from uncertainties measured by varying the SB windows. In the tagged sample the systematic uncertainty is dominated by the uncertainty on the subrtracted contributions of top quark and multijet backgrounds. The description of the data before and after the correction is shown in Figures 4.11 and 4.13 respectively.

For the verification of the robustness of this method and the evaluation of uncertainties induced by the sideband description, estimates have been obtained both by using the low-mass and high-mass SBs separately and also by altering the sizes and positions of the mass windows. Examples of such tests are presented in appendix A.

4.4.2 Top quark background

Top quark production constitutes a significant background only in the tagged event category. This background is dominated by $t\bar{t}$ events, with two leptons originating from either the W bosons or the b quarks produced in top quark decays.

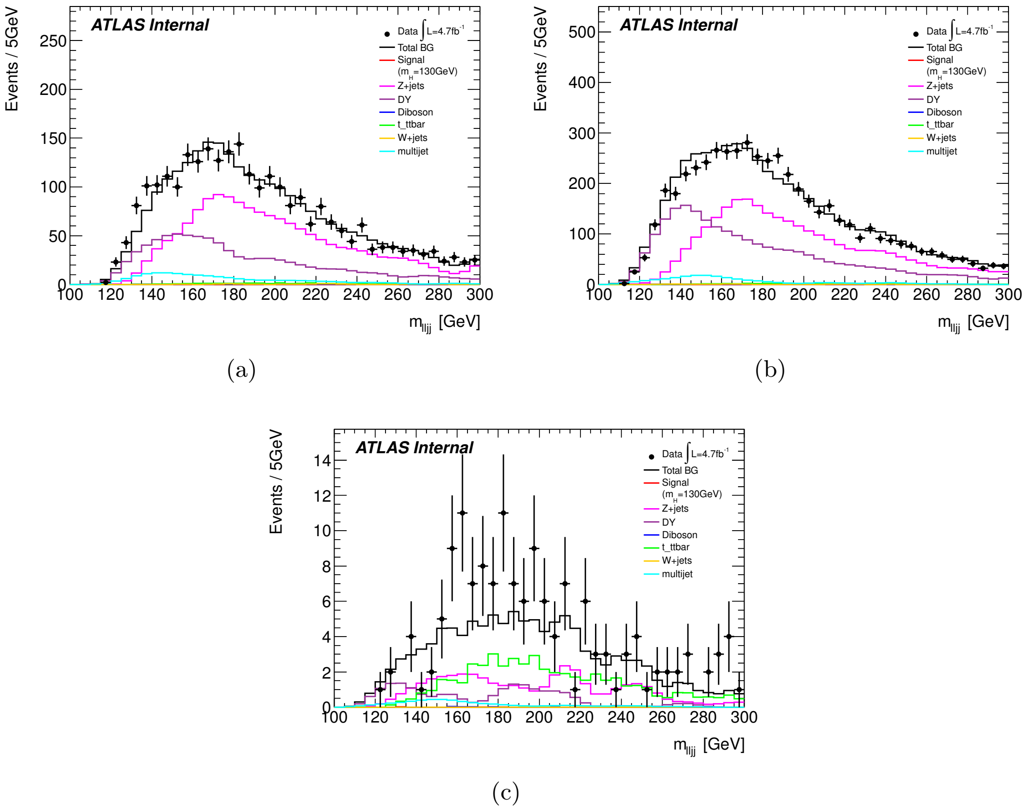


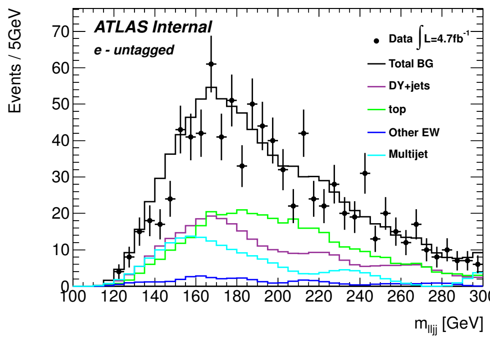
Figure 4.11: m_{lljj} distribution in the DY + jets CR, before any corrections to background modelling, in the (a) muon and (b) electron untagged category and in (c) the tagged category.

The normalization of the top quark background is estimated from the data. The CR is defined by inverting the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement, i.e. requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV. In the tagged sample, this region is dominated by top quark decays and receives small contribution (≈4%) from multijet events. In the untagged sample, beside top quark production, the CR receives almost equal contribution from DY+jets and small contribution from multijet events. The comparison between data and background prediction, in the inverted-$E_{\mathrm{T}}^{\mathrm{miss}}$ CR, is presented in Figure 4.12. An alternative CR, defined by changing the nominal lepton selection to require opposite-flavour ($e\mu$), opposite charge leptons is also used in order to check for systematic effects.

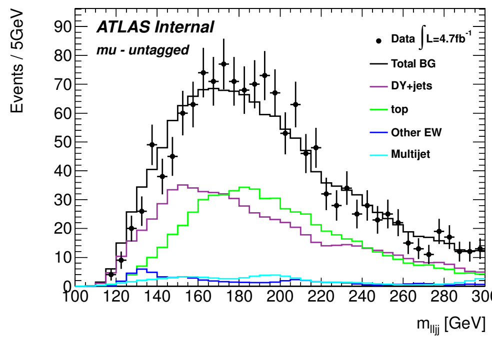
The final scale factors for the top quark background are:

Untagged electron:	$1.03 \pm 0.09(\mathrm{stat}) \pm 0.07(\mathrm{syst})$
Untagged muon:	$1.04 \pm 0.08(\mathrm{stat}) \pm 0.06(\mathrm{syst})$
Tagged:	$1.09 \pm 0.06(\mathrm{stat}) \pm 0.04(\mathrm{syst})$

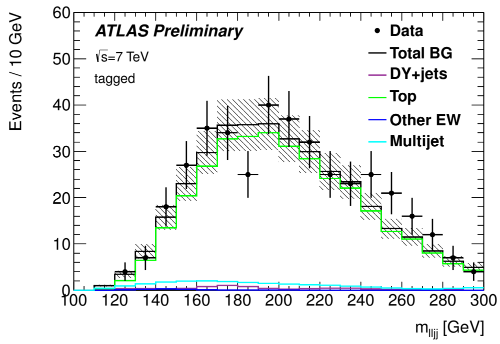
In all cases, the dominant systematic contribution arises from processes which are subtracted from the data; the DY+jets in the untagged category and the multijet background in the tagged category.



(a)



(b)



(c)

Figure 4.12: m_{lljj} distribution in the inverted-$E_{\mathrm{T}}^{\mathrm{miss}}$ top CR, before any corrections to background modelling, in the (a) electron and (b) muon untagged category and in (c) the tagged category.

4.4.3 Diboson background

Since no suitable control region can be defined for the diboson background, it is estimated directly from MC simulation, as discussed in Section 4.1.2.4.

4.4.4 Multijet background

The multijet background is significant in the $m_{\ell\ell}$ range studied in this analysis. Since this background is notoriously difficult to simulate, it is estimated from the data. The shape of the background is obtained from regions dominated by multijet processes, without signal, after subtracting any contribution from the known simulated backgrounds. In the muon channel, the shape is obtained from the region of same charge, isolated muon pairs, while, for electrons, opposite charge candidates, with one of them failing the track-isolation requirement, are chosen instead. The particular regions have been found to provide the best description of the $m_{\ell\ell jj}$ distribution, both in the m_{jj} SBs and in the signal region. Regarding the latter, possible presence of signal would peak around a particular mass and is not expected to affect the overall agreement between data and background. Due to lack of statistics in the tagged sample, common shapes are used for both the untagged and tagged cases. These shapes are obtained from an inclusive sample in which the jet pair is always selected by kinematic fit.

Once the multijet background shape has been obtained, it is subsequently normalized using the two independent methods, described in the paragraphs that follow.

4.4.4.1 Background estimation using a template fit method

In the template fit approach, the multijet background yield is estimated by fitting the dilepton invariant mass distribution to the data, at the selection point where ≥ 2 jets are required, right before imposing the $m_{\ell\ell}$ mass window. The fit is performed over the extended mass range $15 < m_{\ell\ell} < 120$ GeV, using two fitted components; the multijet template and the sum of all the other (background and signal) contributions. From the fit results, the fraction of multijet events to the data is measured, within the dilepton mass window ($20 < m_{\ell\ell} < 76$ GeV), and it is assummed to be constant in all subsequent stages of the event selection. The estimated yields, as fractions of the data, are:

Untagged electron:	$[12.0 \pm 1.4(\text{stat}) \pm 2.3(\text{syst})]\%$
Untagged muon:	$[4.3 \pm 1.0(\text{stat}) \pm 0.0(\text{syst})]\%$
Tagged electron:	$[11.9 \pm 2.9(\text{stat}) \pm 1.6(\text{syst})]\%$
Tagged muon:	$[10.9 \pm 2.8(\text{stat}) \pm 0.0(\text{syst})]\%$

The considerably higher contribution of multijet events in the tagged event category is due to leptonic decays of b quarks. The associated systematic uncertainties are estimated by using different templates. In the muon channel, the modified shape is obtained by requiring opposite charge muons, but with inverted track-isolation requirement for one of them. The difference has been found to be negligible. In the case of electrons, alternative shapes are obtained by using same charge electrons or using electrons which pass only loose identification quality criteria (specifically one *loose*++ and one *medium*++ electron, see Section 4.2.2), but fail the tighter (*tight*++) ones.

4.4.4.2 Background estimation using an ABCD method

The ABCD method provides data-driven estimation of the multijet background yield in the "signal region" (A), using three control regions (B,C,D), dominated by multijet processes. The four regions are defined as:

***A*:** events with leptons of opposite charge, both isolated.

***B*:** events with leptons of opposite charge, one isolated one non isolated.

***C*:** events with leptons of same charge, both isolated.

***D*:** events with leptons of same charge, one isolated one non isolated.

Anti-isolation is required exclusively from only on one of the leptons, in order to keep the selection as close to the original one (i.e. of opposite charge, isolated leptons) as possible. Considering that there is no correlation between the charge sign and isolation requirements, we accept that the ABCD relation:

$$n_A = n_B \frac{n_C}{n_D} \tag{4.2}$$

is true, within statistical uncertainty, allowing the estimation of the number of multijet events n_A, expected in signal region A, from the number of multijet events n_B, n_C, n_D, measured in regions B, C and D respectively.

In order to perform the above calculation, a profile-likelihood approach is followed [84]. By denoting the unknown number of multijet events in region A as μ^U and introducing two nuisance parameters τ_B, τ_C, the number of multijet events in each of the four regions is expressed as:

$$\begin{aligned} n_A &= \mu^U \\ n_B &= \mu^U \tau_B \\ n_C &= \mu^U \tau_C \\ n_D &= \mu^U \tau_B \tau_C \end{aligned} \tag{4.3}$$

and the corresponding total events are then:

$$\begin{aligned} \mu_A &= s_A + b_A + \mu^U \\ \mu_B &= s_B + b_B + \mu^U \tau_B \\ \mu_C &= s_C + b_C + \mu^U \tau_C \\ \mu_D &= s_D + b_D + \mu^U \tau_B \tau_C \end{aligned} \tag{4.4}$$

where $s_{A,B,C,D}$ and $b_{A,B,C,D}$ are known contributions from signal and other background processes respectively. The likelihood function is finally defined as the product of likelihoods corresponding to the counting experiments in the four regions, namely:

$$L(n_A, n_B, n_C, n_D \mid \mu^U, \tau_B, \tau_C) = \prod_{i=A,B,C,D} \frac{e^{-\mu_i} \mu_i^{n_i}}{n_i!}. \tag{4.5}$$

The parameter of interest μ^U is estimated from the minimization of $\log L$, along with the nuisance parameters τ_B and τ_C, thereby providing the normalization of the multijet background in the signal region.

The final estimation for the multijet background percentage over the total background is:

Untagged electron:	[6.6 ± 0.6(stat) ± 2.3(syst)]%
Untagged muon:	[3.5 ± 0.2(stat) ± 0.3(syst)]%
Tagged electron:	[11.5 ± 5.1(stat) ± 2.6(syst)]%
Tagged muon:	[11.0 ± 4.4(stat) ± 3.7(syst)]%

Statistical uncertainties are determinded from the number of events in the control regions. Normalization systematic uncertainties are evaluated by repeating the calculation at different stages of the event selection and also by using different control regions with inverted track-isolation requirements for the selected muons. Shape systematics are studied using distributions from the m_{jj} SBs, instead of the signal region. The overall systematic uncertainty is finally obtained by adding in quadrature normalization and shape uncertainties. Further studies on the estimation of the multijet background are discussed in Appendix B.

4.4.4.3 Final multijet background estimation

As mentioned above, the two methods of background estimation are based on different control regions and techniques; their weighted average is used as the final estimate of the multijet contribution and the maximum of either

- the difference between the estimations and the weighted average or
- the systematic uncertainty estimated for both methods

is conservatively assigned as a systematic uncertainty on the multijet background normalization. The uncertainty is estimated ≈15% for the muon untagged case, ≈35% in the electron untagged sample and ≈50% for the tagged sample. The shape uncertainty is studied by substituting the templates used in each method by the corresponding ones from the m_{jj} sidebands; the impact is found to be negligible.

4.4.5 Summary of backgrounds

The comparison between data and Monte Carlo simulation for the m_{jj} side band region, using the data driven corrections described in the previous sections, is presented in Figure 4.13.

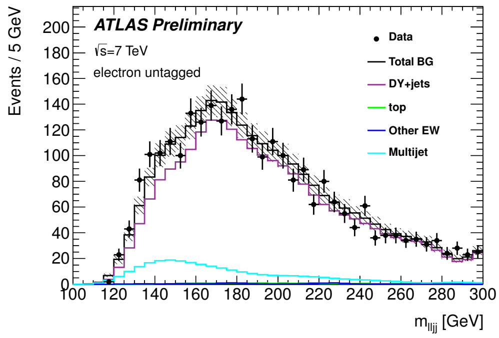
4.5 Systematic uncertainties

Theoretical and experimental uncertainties are applied to signal and background samples which are based on MC simulation. Uncertainties associated to the largest backgrounds, namely DY+jets, multijet and top quark production, are evaluated with data driven methods, as discussed in Section 4.4.

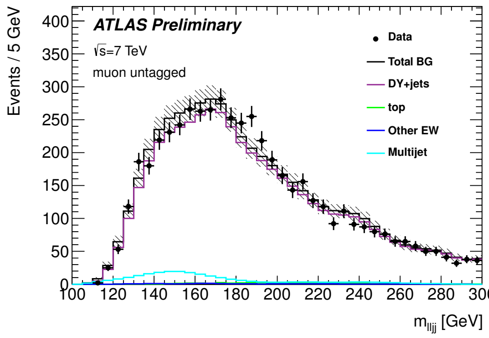
4.5.1 Experimental systematic uncertainties

The uncertainty on the integrated luminosity has been measured ±3.9% [85, 86]. This uncertainty is considered for all MC simulated processes which are not normalized from the data (which is everything except the Z+jets, top quark and multijet background), and is correlated accross all samples.

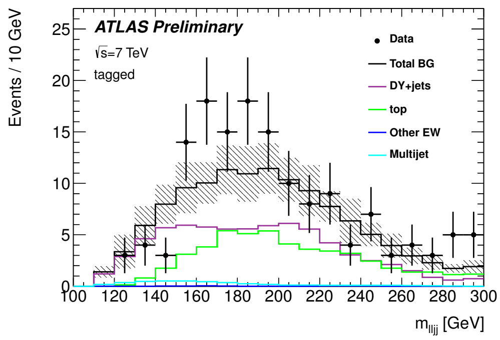
The uncertainties regarding efficiency corrections and calibrations of simulated objects are summarized in Table 4.12. The majority of these uncertainties are provided by the respective combined performance groups of ATLAS and include uncertainties on lepton and jet trigger and identification efficiencies, the energy or momentum calibration and resolution of leptons and jets, and the b-tagging efficiency and



(a)



(b)



(c)

Figure 4.13: m_{lljj} distribution in the DY + jets CR in the (a) electron and (b) muon untagged category and in (c) the tagged category. All scale factors have been applied and systematic uncertainty estimates are included in the errors. The bands show systematic uncertainties from MC statistics.

mistag rates. The uncertainty on $E_{\mathrm{T}}^{\mathrm{miss}}$ is estimated by propagating the uncertainties on individual objects into the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, as discussed in Section 4.2.4. The above detector-related uncertainties are applied to all MC processes. The dominant uncertainty in the tagged sample comes from the b-tagging efficiency. In the untagged sample, the uncertainty on the jet energy scale would normally be expected to have the maximum effect. However, due to the kinematic fit on the dijet system, its effect is found to be marginal.

4.5.2 Signal and background modelling systematics

The theoretical uncertainties on the Higgs boson production cross sections [62,87] are 15–20% for the gluon fusion process ($gg \to H$) and 3–9% for the vector-boson fusion process ($qq \to qqH$), for the Higgs boson mass interval covered by this analysis.

The normalization uncertainties of the DY+jets and top quark backgrounds are estimated with the data driven methods discussed in Section 4.4. The DY+jets uncertainty is estimated from the m_{jj} SBs and is approximatelly 3-4% for the untagged

Source of uncertainty	Treatment in analysis
Jet energy scale (JES)	2–7%, as a function of p_{T} and η
Jet pile-up uncertainty	3–7%, as a function of p_{T} and η
b-quark energy scale	2.5–1% as a function of p_{T}
Jet energy resolution	1–4%
Electron selection efficiency	0.7–3%, as a function of p_{T}; 0.4–6%, as a function of η
Electron Trigger Efficiency	$0.4-1\%$ as a function of η
Electron reconstruction efficiency	0.7–1.8%, as a function of η
Electron energy scale	0.1–6%, as a function of η, pile-up, material effects, etc.
Electron energy resolution	Sampling term 20%; a small constant term has a large variation with η
Muon selection efficiency	0.2–3%, as a function of p_{T}
Muon reconstruction efficiency	0.2–3%, as a function of p_{T}
Muon Trigger efficiency	$< 1\%$
Muon momentum scale	2–16%, as a function of η
Muon momentum resolution	p_{T} and η dependent resolution smearing functions, systematic $\leq 1\%$
b-tagging efficiency	5–15%, as a function of p_{T}
b-tagging mistag rate	10-22%, as a function of p_{T} and η
Missing transverse energy	Add/subtract object uncertainties in $E_{\mathrm{T}}^{\mathrm{miss}}$ + uncertainty on "SoftJet" and "CellOut" terms

Table 4.12: Systematic uncertainties related to object reconstruction and identification.

and 17% for the tagged sample. The top background uncertainty is estimated to be 10% and 7% for the untagged and tagged samples respectively. The diboson cross sections have a combined 5% QCD scale and PDF uncertainty [87]; adding an additional 10% uncertainty, corresponding to the maximum difference seen between MC@NLO and k-factor scaled PYTHIA results, yields an overall uncertainty of 11% on the diboson background normalisation.

As shown in Figure 4.13, the shape of the DY+jets background, after normalization corrections, is reasonably described by the MC simulation. The shape uncertainty is estimated by parameterizing the residual difference of the $m_{\ell\ell jj}$ distribution in the m_{jj} SBs, as a function of $m_{\ell\ell jj}$. The uncertainty measured with the untagged sample is also assigned to the tagged sample due to limited statistics.

The normalization uncertainty for the multijet background is calculated as described in Section 4.4.4, from the difference between the estimations of the two data-driven procedures used to derive the background. The shape uncertainty is estimated by comparison of the template with the one obtained from the m_{jj} sidebands.

4.6 Results

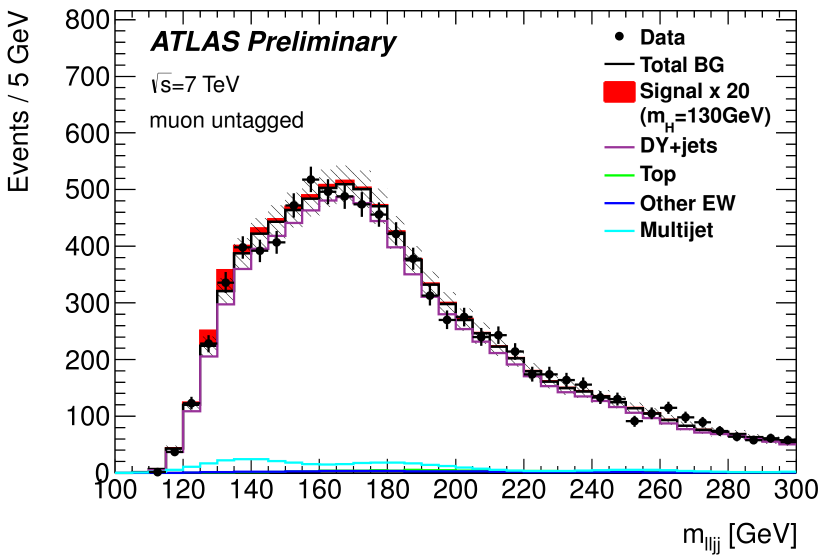
The final results of this search are drawn by comparing the $m_{\ell\ell jj}$ distribution in the data to the background prediction. With the kinematic fit used to rescale the jets momenta, as described in Section 4.3.2, the resolution of the core signal distribution is expected to be ≈ 3 GeV with a long tail also present above the mass peak. In the background the distribution is instead broad, peaking at ≈ 170 GeV with a width of ≈ 40 GeV.

The final $m_{\ell\ell jj}$ distributions are shown in Figures 4.14 and 4.15. The expected signal peak for the hypothesis of a Higgs boson with mass 130 GeV is displayed on top

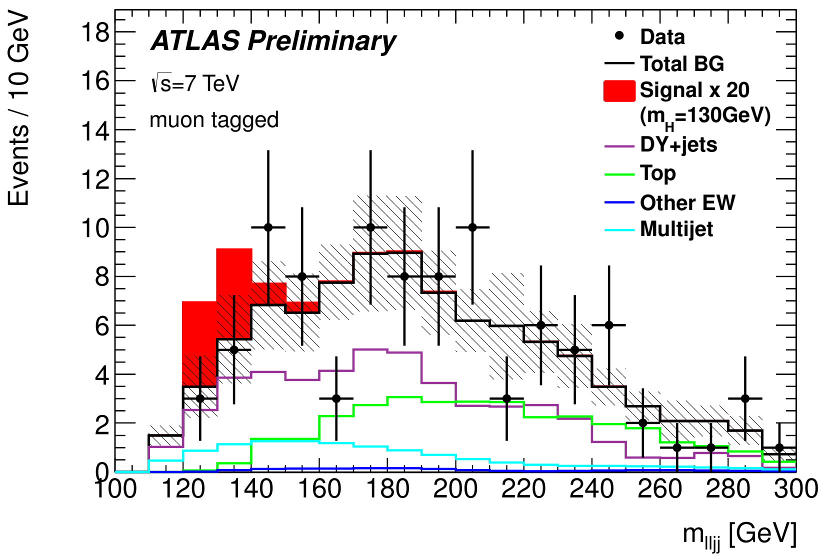
Untagged Channel										
Source	Muons N_{evt}		σ_{stat}		σ_{sys}	Electrons N_{evt}		σ_{stat}		σ_{sys}
DY+jets	9635	±	101	±	409	4654	±	42	±	161
Top quark	99.0	±	1.8	±	9.8	69.0	±	1.5	±	7.8
Multijet	388	±	26	±	50	36.3	±	1.0	±	5.3
Diboson	60.9	±	1.2	±	9.1	1.7	±	0.2	±	0.3
W+jet	10.9	±	2.5	±	1.6	30.2	±	10.7	±	4.3
Total background	10161	±	105	±	461	5291	±	63	±	231
Data	9714					5197				
Signal $m_H = 120$ GeV	2.07	±	0.10	±	0.12	0.90	±	0.06	±	0.07
Signal $m_H = 130$ GeV	7.26	±	0.28	±	0.38	3.19	±	0.18	±	0.24
Signal $m_H = 150$ GeV	21.1	±	0.60	±	0.73	9.70	±	0.40	±	0.55
Signal $m_H = 180$ GeV	4.86	±	0.20	±	0.17	2.85	±	0.15	±	0.14
Tagged Channel										
Source	Muons N_{evt}		σ_{stat}		σ_{sys}	Electrons N_{evt}		σ_{stat}		σ_{sys}
DY+jets	53.0	±	5.3	±	7.7	30.5	±	3.5	±	4.4
Top quark	33.8	±	1.0	±	2.2	22.2	±	0.8	±	1.6
Multijet	11.0	±	4.4	±	1.6	7.0	±	3.2	±	2.5
Diboson	1.7	±	0.2	±	0.3	1.1	±	0.2	±	0.2
Total background	99.5	±	7.0	±	8.2	60.8	±	4.8	±	5.3
Data	105					51				
Signal $m_H = 120$ GeV	0.080	±	0.018	±	0.010	0.042	±	0.016	±	0.006
Signal $m_H = 130$ GeV	0.431	±	0.067	±	0.051	0.288	±	0.055	±	0.037
Signal $m_H = 150$ GeV	1.561	±	0.167	±	0.178	0.594	±	0.100	±	0.072
Signal $m_H = 180$ GeV	0.299	±	0.049	±	0.034	0.177	±	0.039	±	0.021

Table 4.13: Number of observed events, along with background estimation for the various processes, and signal expectation for various m_H hypotheses in the interval 120-180GeV. Both statistical and systematic uncertainties are shown.

of the background. A factor of 20 (5) multiplies the signal in the untagged (tagged) sample to make the peak visible. The final number of data events measured along with the background prediction and signal expectation are listed in Table 4.13.

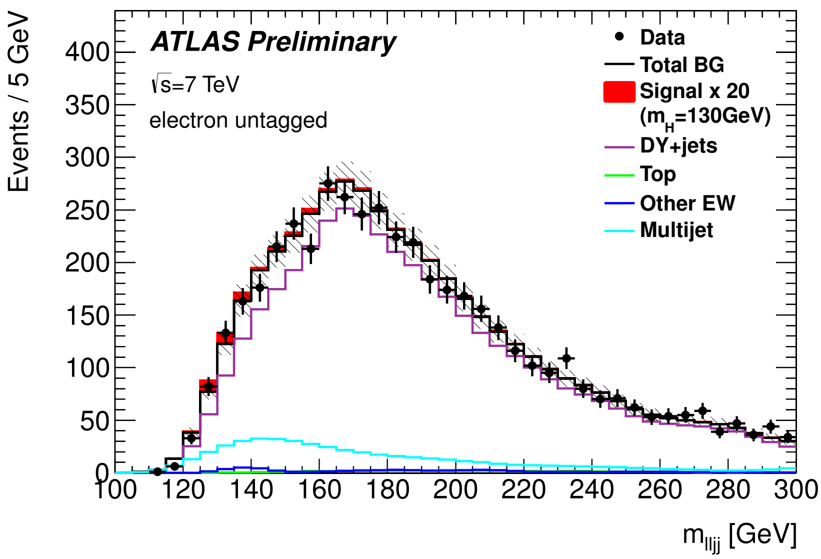


(a)

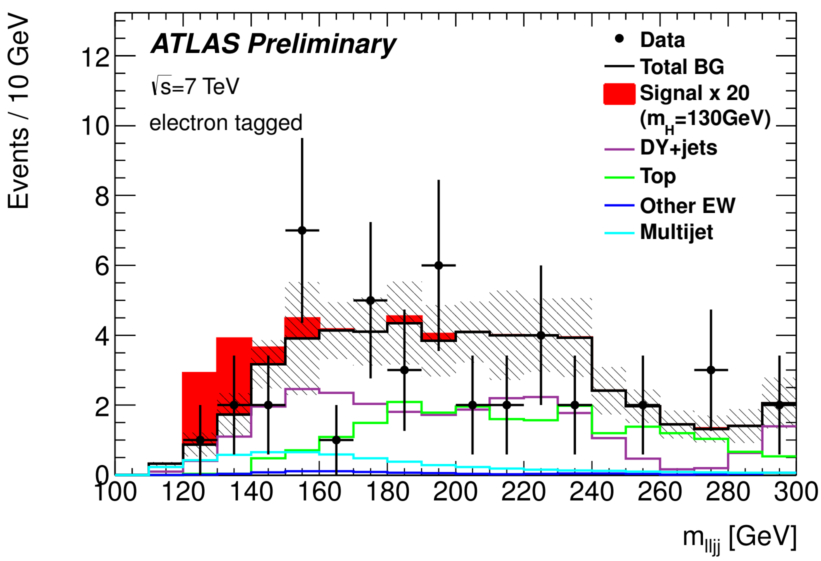


(b)

Figure 4.14: Distribution of $m_{\ell\ell jj}$ in the muon channel compared to the background prediction and including the signal for a Higgs boson mass of 130 GeV in the (a) untagged and (b) tagged samples. A factor of 20 (5) multiplies the signal in the untagged (tagged) case to make the peak visible.



(a)



(b)

Figure 4.15: Distribution of $m_{\ell\ell jj}$ in the electron channel compared to the background prediction and including the signal for a Higgs boson mass of 130 GeV in the (a) untagged and (b) tagged samples. A factor of 20 (5) multiplies the signal in the untagged (tagged) case to make the peak visible.

4.7 Limit extraction

No significant excess of observed events over the background is found, therefore limits are set on the Higgs boson cross section as a function of its mass m_H. The limits are calculated using the CL_s modified frequentist formalism with the profile likelihood test statistic [88, 89].

4.7.1 Results with MCLimit

Preliminary estimations of the upper limits on the Higgs boson cross section as a function of m_H are derived using MCLimit [90]. The inputs are the $m_{\ell\ell jj}$ distributions in the 4 independent analysis channels (shown in Figures 4.14 and 4.15) with the systematic uncertainties described in Section 4.5. The observed 95% CL_s limits, along with the median expected limits (in the absence of signal) and their $\pm1\sigma$ and $\pm2\sigma$ fluctuations, are listed in Table 4.14 for the various m_H hypotheses.

m_H	Observed	Expected μ/μ_{SM}				
(GeV)	μ/μ_{SM}	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$
120	50.4	7.7	28.1	51.7	88.0	113
125	32.2	3.1	11.3	21.3	38.3	64.0
130	12.9	1.2	4.0	9.5	16.0	22.8
135	7.3	1.0	3.2	5.4	8.7	13.4
140	5.8	1.0	3.2	4.8	7.4	10.5
145	3.7	1.0	2.0	3.6	6.1	8.0
150	5.7	1.0	1.9	4.1	6.3	8.1
155	8.0	0.4	1.9	3.7	5.8	8.0
160	6.6	2.0	4.0	7.8	11.9	15.4
165	4.7	0.8	6.2	13.8	19.0	30.3
170	4.0	1.0	6.8	13.6	23.3	34.8
175	11.5	0.9	8.1	16.0	25.7	32.0
180	19.5	1.0	12.5	19.3	31.2	46.5

Table 4.14: 95% CL_s upper limits in the $H \to ZZ \to \ell^+\ell^- q\bar{q}$ channel, as a multiple of the SM rate, for an integrated luminosity of 4.71 fb^{-1}. The limits were obtained with MCLimit.

4.7.2 Results with RooStats

The final limit calculation is performed with the HistFactory tool [91], member of the RooStats statistical analysis framework [92], using the asymptotic approximation [88]. The most notable benefit HistFactory offers over MCLimit is the option to include the bin-by-bin statistical uncertainties of background templates in the limit calculation (see also Section 6.4).

Figure 4.16 presents the final results of this search, showing the expected and observed 95% CLs upper limits as multiples of the SM cross section and considering all systematic uncertainties. The limits are also listed in Table 4.15 and appear to be consistent with the results obtained using MCLimit. Finally, limits are also shown separately for each of the four individual analysis channels, in Figure 4.17. As observed, all channels contribute with approximately equal sensitivity.

m_H	Observed	Expected $\mu/\mu_{\rm SM}$				
(GeV)	$\mu/\mu_{\rm SM}$	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$
120	52.38	34.78	46.69	64.79	94.22	139.71
125	22.72	14.07	18.88	26.21	38.05	56.04
130	16.43	6.11	8.21	11.39	16.42	23.77
135	8.80	3.31	4.45	6.17	8.93	12.96
140	5.79	3.08	4.13	5.74	8.23	11.80
145	3.45	2.20	2.95	4.10	5.91	8.53
150	4.45	2.22	2.98	4.14	5.94	8.52
155	5.39	2.37	3.18	4.41	6.37	9.19
160	6.16	4.63	6.21	8.62	12.40	17.79
165	6.54	7.83	10.51	14.58	21.02	30.43
170	6.98	8.74	11.73	16.28	23.47	33.77
175	10.61	10.52	14.12	19.60	28.15	40.24
180	13.28	11.13	14.94	20.74	29.78	42.63

Table 4.15: 95% CL_s upper limits in the $H \to ZZ \to \ell^+\ell^- q\bar{q}$ channel, as a multiple of the SM rate, for an integrated luminosity of 4.71 fb^{-1}.

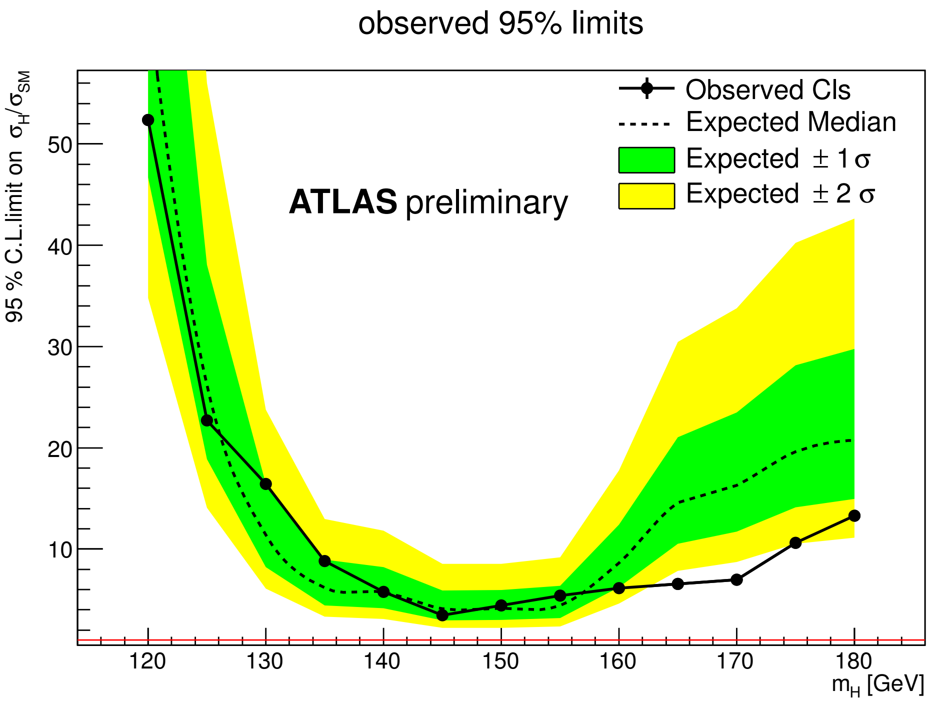
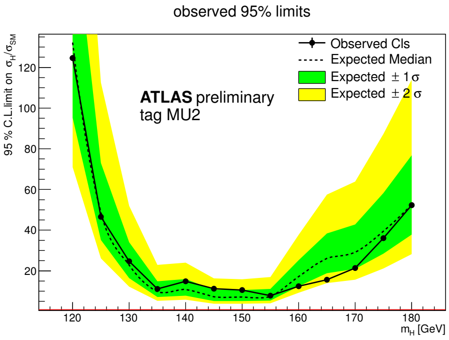
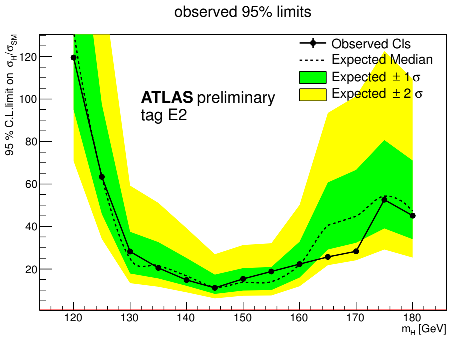


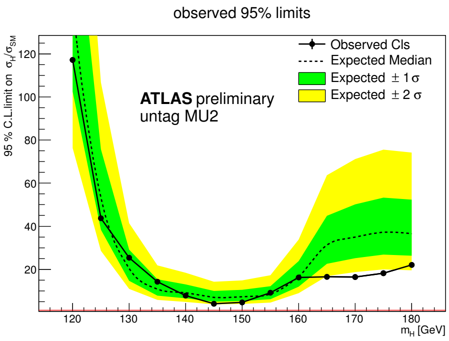
Figure 4.16: The expected (dashed line) and observed (solid line) upper limits on the total cross section as multiples of the expected SM The expected (dashed line) and observed (solid line) upper limits on the total cross section divided by the expected SM Higgs boson cross section, calculated using CL_s at 95%. The inner and outer bands indicate the $\pm 1\sigma$ and $\pm 2\sigma$ ranges in which the limit is expected to lie in the absence of signal. The horizontal dashed line shows the SM value of 1.



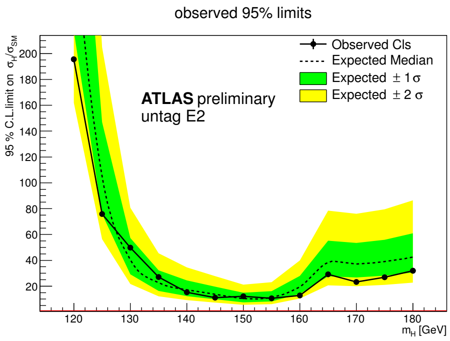
(a)



(b)



(c)



(d)

Figure 4.17: The expected (dashed line) and observed (solid line) upper limits on the total cross section as multiples of the expected SM The expected (dashed line) and observed (solid line) upper limits on the total cross section divided by the expected SM Higgs boson cross section, calculated using CL_s at 95%. The inner and outer bands indicate the $\pm 1\sigma$ and $\pm 2\sigma$ ranges in which the limit is expected to lie in the absence of signal. The horizontal dashed line shows the SM value of 1. (a) tagged muon channel, (b) tagged electron channel, (c) untagged muon channel (d) untagged electron channel.

Chapter 5

Beyond the Standard Model

With the discovery of the Higgs boson, with a mass of $\sim$ 125 GeV, the ATLAS and CMS collaborations added the missing keystone to the SM foundation. One of the remaining questions though is if the newly discovered boson is part of an extended scalar sector as postulated by various extensions to the SM. As discussed in Section 1.3, these models predict additional Higgs bosons and thereby motivate the continuation of the Higgs hunting at higher mass regions.

The study presented in the following sections probes the scenario of a new, heavy Higgs boson with a narrow width, as well as Type-I and Type-II 2HDMs (see Section 1.3), using data taken by ATLAS during the 2012 LHC run, at a center-of-mass energy of 8 TeV. A Higgs boson is searched for, in the mass range $200 < m_H < 1000$ GeV, with the decay mode $H \to ZZ \to \ell^+\ell^- q\bar{q}$ ($\ell = e, \mu$). The event selection is therefore tuned to reconstruct Higgs boson candidates, from a pair of leptons and a pair of high-p_{T} jets, and employs a number of criteria in order to maximize the signal significance against the background.

There are several background processes that give rise to final states with signatures similar to the signal. The dominant background process is Z+jets production, with the Z boson decaying to a pair of leptons. Top quark production also becomes important when two b-tagged jets are required in the final state. These backgrounds are taken from MC simulations but are subsequently corrected to better describe the data, as described in Section 5.4. Secondary background contributions arise from diboson ($WW/WZ/ZZ$) and W+jets production; these are modelled by MC simulation, while the small QCD multijet background is derived from the data, as discussed in Section 5.4.3.

The integrated luminosity achieved in the 2012 LHC run, allows optimization of the analysis by distinguishing subchannels for the dominant Higgs production mechanisms, gluon fusion (ggF) and vector-boson fusion (VBF), and also by classifying events according to the multiplicity of b-tagged jets. For high Higgs boson masses ($m_H > 700$ GeV), an additional category is adopted searching for $Z \to q\bar{q}$ decays in which the Z boson decay products are merged into a single, high-p_{T} jet. The conclusions of this search are finally drawn by looking for a peak above the background expectation in the invariant mass of the $\ell\ell jj$ system, $m_{\ell\ell jj}$.

5.1 Data and Monte Carlo samples

After the implementation of the Good Runs List (see Section 4.1), the filtered data set corresponds to an integrated luminosity of 20.3 fb^{-1}. Trigger requirements employ both single-lepton and dilepton non-prescaled triggers. The main single-lepton triggers have a minimum p_{T} (E_{T} for electrons) threshold of 24 GeV and require that the leptons are isolated. These are augmented with triggers with higher thresholds

(60 GeV for electrons and 36 GeV for muons) but with no isolation requirement in order to recover acceptance at high p_{T}. The dilepton triggers require two same-flavour leptons with a threshold of 12 GeV for electrons and 13 GeV for muons.

MC samples simulating the signal and background processes are generated at 8 TeV center-of-mass energy and passed through the simulation of the ATLAS detector [39]. To account for any differences in the distribution of pile-up interactions between MC simulation and data, the MC samples are reweighted accordingly (see Section 4.1.3).

5.1.1 Signal samples

Signal samples, simulating $H \to ZZ \to \ell^+\ell^- q\bar{q}$, with the Higgs boson line shape described by a fixed 1 GeV wide Breit-Wigner distribution, are generated separately for the ggF and VBF modes. The simulation uses POWHEG [43, 44] which calculates separately the gluon and vector-boson fusion (ggF and VBF) Higgs boson production mechanisms up to next-to-leading order (NLO) in α_{S}. The generated signal events are hadronized with PYTHIA 8 using the NLO CT10 [93] parton distribution function (PDF) and the AU2 set of tunable parameters for the underlying event [94, 95].

In this narrow width approximation (NWA), the interference effect between signal and continuum background is negligible over the full mass range [96, 97]. The optimization of event selection criteria is performed assumming the SM cross-section for the signal, as benchmark. The estimation of the final limits is of course independent of this choice, since the limits are set on the cross-section times branching ratio itself. Details on the calculation of SM cross-sections can be found in Section 4.1.1.

NWA signal samples are also used for the 2HDM interpretation, since the natural width of the heavy Higgs boson is expected to be small over most of the 2HDM parameter space (see Ref. [27]), specifically in low tanβ and up to moderate m_H. To account for variations of the Higgs width across the parameter space, Gaussian smearing is applied to the reconstructed m_{lljj} distribution, to have it reflect the natural width at a given model point, which is significant given the good m_{lljj} mass resolution, as discussed in Appendix C). The production cross sections for both the ggF and VBF processes in the 2HDM are calculated using SUSHI [49, 52, 98–101] and the branching ratios are calculated with 2HDMC [102]. For the calculation of branching ratios, it is assumed that $m_{\mathrm{A}} = m_H = m_H^{\pm}$, $m_{\mathrm{h}} = 125$ GeV, and $m_{12}^2 = m_{\mathrm{A}}^2 \tan\beta/(1+\tan\beta^2)$.

5.1.2 Background samples

The background processes contributing to the analysis are summarized in Table 5.1.

5.1.2.1 Vector boson (V)+jets production

W/Z+jets events are produced with SHERPA [103], interfaced with CT10 PDFs [93], using filters to distinguish events containing $\geq 1b$, $\geq 1c$ (and $0b$) and those containing only light-flavoured hadrons. The samples are further divided in different intervals of the vector boson transverse momentum p_{T}^V, namely: $40 < p_{\mathrm{T}}^V < 70$, $70 < p_{\mathrm{T}}^V < 140$, $140 < p_{\mathrm{T}}^V < 280$, $280 < p_{\mathrm{T}}^V < 500$ and $p_{\mathrm{T}}^V > 500$ GeV. Due to this classification,

Process	Generator	$\sigma \times BR$	N_{events}
Vector boson + jets			
$W \to \ell\nu$	SHERPA 1.4.1	12.07 nb	390M
$Z/\gamma* \to \ell\ell$	SHERPA 1.4.1		66M
$m_{\ell\ell} > 40$ GeV		1.24 nb	
$Z \to \ell\ell$	ALPGEN	1.16 nb	
Top quark			
$t\bar{t}$	POWHEG-BOX	252.89 pb	100M
$t-$channel	ACERMC	87.76 pb	9M
$s-$channel	POWHEG-BOX	5.61 pb	6M
$Wt-$channel	POWHEG-BOX	22.37 pb	20M
Diboson (moving to POWHEG-BOX)			
WW	POWHEG-BOX	52.44 pb	10M
WZ	POWHEG-BOX		20M
$m_{\ell\ell} > 20$ GeV + 1 boson decaying hadronically		9.241 pb	
ZZ	POWHEG-BOX		7.5M
$m_{\ell\ell} > 20$ GeV + 1 boson decaying hadronically		3.171 pb	

Table 5.1: Monte Carlo programs used for the modelling of the various background processes and the cross-section times branching ratio (BR) values used to normalize the different processes at $\sqrt{s} = 8$ TeV. Branching ratios correspond to the decays shown.

SHERPA offers abundant statistics up to high p_{T}^{V} (and hence high $m_{\ell\ell jj}$) regions for all flavours of the associating jets; for this reason it is prefered over ALPGEN for the ggF analysis.

Although SHERPA fits the requirements of the ggF analysis, the VBF analysis uses ALPGEN, interfaced to PYTHIA [45] for hadronization, since it better models the higher jet multiplicities expected in VBF events. Samples are generated with a fixed number of hard partons in the final state (0,1,2,3,4,$\geq$5) and with the partons matched to final-state jets. Dedicated samples for $Z(\to \ell\ell)+b\bar{b}$ and $Z(\to \ell\ell)+c\bar{c}$ processes are also available with 0,1,2,$\geq$3 partons. Any possible overlap of events between these samples is removed as described in Section 4.1.2.1.

5.1.2.2 Top quark pair and single top quark production

Samples of top quark pairs are generated with POWHEG-BOX [40–44], interfaced to PYTHIA, using a filter which requires that at least one W boson (from the top quark decay) decays into a charged lepton (e, μ or τ). Parton showering and hadronization is generated according to the Perugia2011C tune [104] using the LO CTEQ6L1 PDF set [105]. For single-top processes, s-channel and Wt-channel, samples are generated with POWHEG-BOX +PYTHIA, while the t-channel sample is generated with ACERMC+PYTHIA. All single-top channels use the CTEQ6L1 PDF set and the Perugia2011C tune.

5.1.2.3 Diboson production

The diboson background includes processes with two vector bosons in the final state (WW, WZ, ZZ). The background from ZZ production, where one Z boson decays leptonically and the other decays hadronically, is largely irreducible since it gives rise to the same final state as the signal process. Some contribution is also expected from WZ production where the Z boson decays leptonically and the W boson decays hadronically. The contribution of WW production is minimal. For the simulation of diboson production, the POWHEG generator provides a next-to-leading-order (NLO) estimate, relying on the CT10nlo PDF and the PYTHIA8 parton shower and hadronization model.

5.1.2.4 SM $Zh, h \to bb$ production

The SM $Zh(h \to bb)$ production constitutes a background in the 2HDM scenario only and it may be non-negligible in the ggF-category when two b-tagged jets are required in the final state (see also Section 5.4.6). These events are modelled with PYTHIA 8, for the simulation of $qq \to Zh$, and POWHEG for $gg \to Zh$.

5.2 Definition of physics objects

The first step in the selection procedure is to identify the physics objects that will form the building blocks of the analysis. Reconstructed muon, electron and jet candidates are classified as either "loose" or "tight". The loose class is defined by basic kinematic and reconstruction quality criteria; leptons and jets failing the loose criteria are removed from the analysis. Tight criteria are imposed only on loose candidates which have been selected to reconstruct the Higgs boson final state.

5.2.1 Muons

The reconstruction of muon objects is outlined in Section 4.2.1. In addition to combined (CB) and segment-tagged (ST) muons, this search also employs calorimeter-tagged (CaloTag) and standalone (SA) muons. CaloTag muons are formed whenever an inner detector track can be associated to energy deposits in the calorimeter which are consistent with a minimum-ionizing particle. Muons of this category are useful for recovering acceptance around $|\eta| = 0$ where there is a gap in the muon spectrometer coverage. Standalone muons use unmatched tracks, measured only in the muon spectrometer. The direction of flight and the impact parameter at the interaction point are determined by back-extrapolating the track to the point of closest approach to the beam line, taking into account energy losses in the calorimeters. Standalone muons are included in the analysis in order to increase the acceptance in pseudorapidity regions beyond the inner detector acceptance.

The definitions of "loose" and "tight" muons are given in Table 5.2. The muon selection is summarized in Table 5.3. Clarifications on the selection criteria can be found in Section 4.2.1.

Muon momenta in MC simulated samples are smeared according to measurements of $Z \to \mu\mu$ decays [106], and appropriate weights are applied to account for the

Class	Family	Kinematics
Loose	CB+ST	$p_{\mathrm{T}} > 7$ GeV, $\lvert\eta\rvert < 2.7$
	CaloTag	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 0.1$ ($\Delta R > 0.1$ from CB/ST/SA muons)
	Standalone	$p_{\mathrm{T}} > 7$ GeV, $2.5 < \lvert\eta\rvert < 2.7$
Tight	CB+ST	$p_{\mathrm{T}} > 25$ GeV, $\lvert\eta\rvert < 2.5$

Table 5.2: Definition of loose and tight muons in terms of the muon category and kinematic requirements.

difference in offline trigger efficiencies between the simulation and the data. Muon isolation efficiency is also corrected. Figure 5.1 shows the p_{T} and η distributions of muons which have been selected to reconstruct $Z \to \mu\mu$ decays[1].

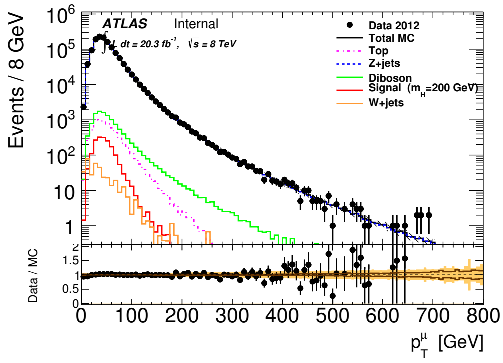
Identification	Loose or tight (see Table 5.2)
Inner detector	$N_{\mathrm{hits}}^{\mathrm{pixel}} + N_{\mathrm{dead}}^{\mathrm{pixel}} > 0$ $N_{\mathrm{hits}}^{\mathrm{SCT}} + N_{\mathrm{dead}}^{\mathrm{SCT}} > 4$ $N_{\mathrm{holes}}^{\mathrm{pixel}} + N_{\mathrm{holes}}^{\mathrm{SCT}} < 3$ $\lvert\eta\rvert < 1.9$: $N_{\mathrm{tot}}^{\mathrm{TRT}} > 5$ and $0.1 < N_{\mathrm{outliers}}^{\mathrm{TRT}} < 0.9 \times N_{\mathrm{tot}}^{\mathrm{TRT}}$ where $N_{\mathrm{tot}}^{\mathrm{TRT}} = N_{\mathrm{hits}}^{\mathrm{TRT}} + N_{\mathrm{outliers}}^{\mathrm{TRT}}$.
Cosmic rejection	$\lvert d_0\rvert < 1$ mm (not for SA muons) $\lvert z_0\rvert < 10$ mm (not for SA muons)
Track isolation	$\sum_{\mathrm{tracks}} p_{\mathrm{T}}(\Delta R < 0.2)/p_{\mathrm{T}}^{\mu} < 0.1$ (not for SA muons)

Table 5.3: Summary of muon selection criteria. N_{hits} (N_{holes}) represents the number of hits (missing hits) in a particular subdetector of the inner tracker, while N_{dead} refers to the number of dead sensors crossed by the muon in a particular subdetector.

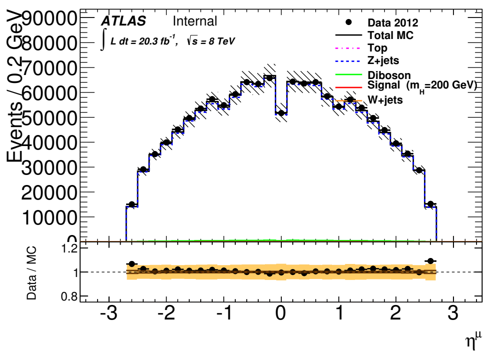
5.2.2 Electrons

The selection criteria for electrons are listed in Table 5.4. All electron candidates are required to satisfy the "VeryLoose" standard ATLAS identification quality criteria [107]. The selection then continues as in Section 4.2.2, with an increased E_{T} threshold imposed to tight candidates, $E_{\mathrm{T}} > 25$ GeV. Electron momenta in MC simulation are corrected for both energy scale and resolution, based on measurements of $Z \to ee$ decays [108]. Differences in reconstruction and identification efficiencies are

[1]In this and all subsequent plots, the MC prediction is normalized to the luminosity of the data, unless otherwise specified, and the background normalisation is defined from the final fit. The shaded (orange) band in the main (ratio) plot shows the quadratic sum of the MC statistical uncertainty (indicated by the brown histogram on the ratio) and the shape-dependent systematic uncertainty on the total background; Normalisation-only systematic uncertainties (i.e. luminosity, QCD/diboson normalisation and uncertainties on the fitted MC background normalisation) are not included.



(a)



(b)

Figure 5.1: The p_{T} and η distributions of muons used to reconstruct Z boson candidates, after the dilepton mass selection $83 < m_{\mu\mu} < 99$ GeV.

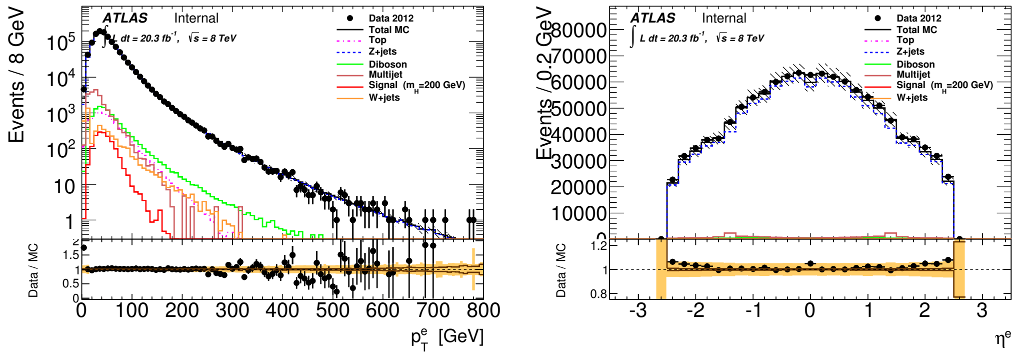
also corrected to match those in the data. The p_{T} and η distributions of electrons which have been identified as the products of $Z \to ee$ decays, are shown in Figure 5.2.

Identification	Author: Electron IsEM: VeryLooseLH
Kinematic cuts	$E_T > 7$ GeV (loose) $E_T > 25$ GeV (tight) $\lvert\eta_{\mathrm{cluster}}\rvert < 2.47$
Track isolation	$\sum_{\mathrm{tracks}} p_{\mathrm{T}}(\Delta R < 0.2)/p_{\mathrm{T}}^{e} < 0.1$

Table 5.4: Summary of electron selection criteria.

5.2.3 Jets

Jet candidates are reconstructed as described in Section 4.2.3. In this analysis however, global sequential calibration (GSC) [109] is applied on top of the calibration to EM+JES, since it has been found to improve the dijet invariant mass resolution by ∼8%. The jet selection is summarized in Table 5.5 and is similar to the selection followed in the SM Higgs search (see Section 4.2.3). The forward detector region $2.5 \leq |\eta| < 4.5$ is included for the identification of VBF signatures in which two hard jets are expected, with large separation in η, in addition to the jets from the Higgs boson decay. An increased p_{T} threshold in that region filters out low-p_{T} jets originating from pile-up interactions. In the central detector region, pile-up jets are removed using a cut on the jet vertex fraction [79]. Finally, due to the pile-up effect, the jet energy scale is expected to be shifted therefore an offset correction, derived from MC simulation, is applied. The p_{T} and η distributions of loose jets are shown in Figure 5.3 for events containing at least 2 tight jets.



(a) (b)

Figure 5.2: The p_{T} and η distributions of the two electrons forming the leptonic Z boson candidate, after the Z boson mass selection $83 < m_{ee} < 99$ GeV.

Identification	Anti-k_T $R = 0.4$ topological jets
Kinematic cuts	$p_{\mathrm{T}} > 20$ GeV $\quad |\eta| < 2.5$ (tight)
	$p_{\mathrm{T}} > 30$ GeV $\quad 2.5 \leq |\eta| < 4.5$
Pile-up	$|\mathrm{JVF}| > 0.5$ (if $|\eta| < 2.4$ and $p_{\mathrm{T}} < 50$ GeV)

Table 5.5: Summary of jet selection criteria.

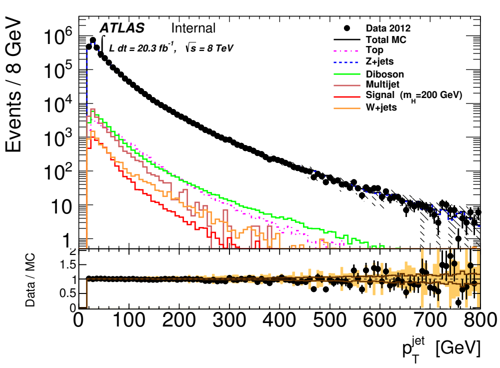
5.2.3.1 Jet flavour labelling

The true flavour of reconstructed jets in MC samples is determined from the hadrons within a cone $\Delta R = 0.4$ of the jet axis. If there is a B-hadron within this cone, the jet is labeled as b otherwise, if there is a D-hadron, it is labelled as c. If neither of these conditions are met, the jet is labelled as a light jet.

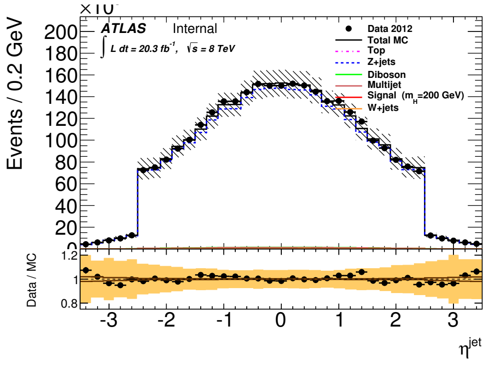
5.2.3.2 Identification of b jets

For the discrimination of jets that originate from fragmentation of b quarks, the b-tagging algorithm `MV1c` is used. `MV1c` is the successor of `MV1` and has been trained against a mixture of c and light jets to improve c-jet rejection (`MV1` has been trained only against light-jets). Figure 5.4 shows the charm rejection rates as a function of the b-tagging efficiency for jets stemming from simulated $t\bar{t}$ events, produced according to the SM predictions. The selection is applied at the operating point that gives an efficiency of 70%, on average, for identifying true b-jets, while the efficiencies for accepting c jets or light quark jets are 1/5 and 1/140 respectively [81,82,110,111].

The b-tagging selection has been calibrated so that the efficiency in MC simulation matches that in the data. Usually, calibration is performed for specific operating points of the tagging algorithm but, in this case, data-to-MC scale factors are available over the full `MV1c` spectrum. This is referred to as “continuous” or “pseudo-continuous” tagging. As discussed in Section 5.4.1.1, this allows using the



(a)



(b)

Figure 5.3: p_{T} and η distributions of "loose" jets, after the $Z \to ll$ boson mass selection, at least two "tight" jets and $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_T}$ cut. The MC is normalized to the luminosity of the data and the shaded band shows the systematic uncertainty on the total MC background.

`MV1c` output itself as a natural variable in order to correct the flavour composition of the dominant Z+jets background. The data-to-MC scale factors used to correct the b-tagging efficiency, were derived using PYTHIA6 for b jets and PYTHIA8+EVTGEN for c-jets. Since the b-tagging efficiency has been found to be generator dependent, MC-to-MC correction factors are additionally applied to take this dependence into account.

5.2.3.3 Truth tagging

In the ggF channel (which discriminates events according to the number of b-tagged jets) the flavour composition of the Z+jets MC simulated samples is corrected to match that in the data, by means of a "flavour fit". As discussed in Section 5.4.1.1, this fit makes use of three control regions, corresponding to the number of b-tagged jets per event (0, 1 and 2). Due to the powerful discrimination of the `MV1c` algorithm against non-b jets, Z+jets events with only light or c jets are very few in the 1 and 2b-tag control regions used by the flavour fit. In order to increase the statistical power of mis-tagged Z+jets events in the 2b-tag control region, "truth tagging" is applied. If neither of the two leading (in p_{T}) jets is a true b jet, each of them is assigned a random `MV1c` weight, above the operating point. This weight is obtained by sampling the tagger's cumulative efficiency distribution that corresponds to the particular process type as well as the flavour, transverse momentum and pseudorapidity of the jet. Finally, the event is weighted by the actual efficiency of each jet to pass the operating point.

Truth tagging is only applied to Z+jets MC simulated events in which neither of the two leading jets is a true b jet. In all other cases, the `MV1c` output values are used directly ("direct tagging").

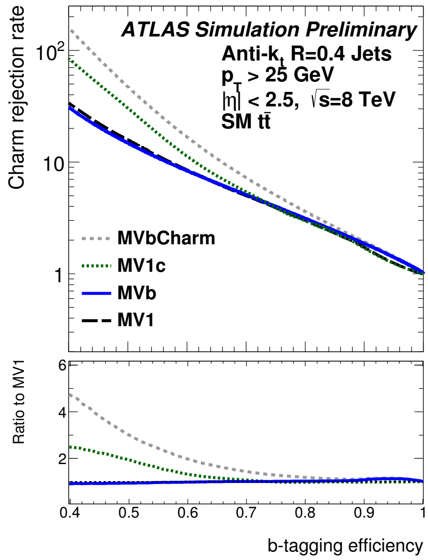


Figure 5.4: Charm rejection rates as a function of the b-tagging efficiency for jets stemming from simulated $t\bar{t}$ events produced according to the SM predictions [112]. The `MV1` and `MV1c` algorithms are shown along with the newer `MVb` and `MVbcharm`.

5.2.3.4 Energy corrections to the selected b jets

Two corrections are applied to b jets, selected to reconstruct $Z \to b\bar{b}$ decays [113]. The first correction ("muon-in-jet") accounts for energy lost due to semi-leptonic decays of b hadrons to muons. If a reconstructed muon with $p_{\mathrm{T}} > 4$ GeV is identified within $\Delta R < 0.4$ of a b-tagged jet and satisfies the inner detector hit requirements for muons (see Section 5.2.1), then its four-momentum is added to that of the jet, after having subtracted the energy that the muon has deposited in the calorimeter. The second correction (resolution correction) depends on the reconstructed p_{T} and accounts for discrepancies ($\sim 5\%$), measured in MC simulation, between the reconstructed and the true p_{T} of b jets.

5.2.4 Overlap removal

Since it may occur that the same tracks or calorimeter energy deposits are used in multiple reconstructed objects, a series of tests are applied in order to resolve cases of overlapping candidates.

1. First, jet candidates within $\Delta R < 0.4$ of an electron candidate are removed.

2. Next, fake jets caused by mis-reconstructed muon energy deposits in the calorimeter are removed. Specifically, for jets within $\Delta R < 0.4$ of a muon, the JVF variable is recalculated excluding the muon track, and the jet is removed if either the "corrected" JVF is less than 0.5 or it appears to have a single associated track with $p_{\mathrm{T}} > 1$ GeV.

3. Next, muons with $p_{\mathrm{T}} < 20$ GeV and within $\Delta R < 0.4$ of a remaining jet are removed. This procedure removes muons from semi-leptonic decays of heavy flavour hadrons (which are likely to have low p_{T}) while keeping cases where a muon deposits significant energy in the calorimeter (faking a jet) and also cases where a muon from the decay of a Z boson happens to overlap with a jet.

4. As a final step, any electron within $\Delta R < 0.2$ of a non-CaloTag muon is removed. For CaloTag muons, the muon is removed instead.

5.3 Event selection

All triggered events are required to contain a reconstructed primary vertex formed by at least three tracks with $p_{\mathrm{T}} > 150$ MeV. The $E_{\mathrm{T}}^{\mathrm{miss}}$-cleaning procedure, described in Chapter 4, is then followed in order remove events containing mismeasured jets which give rise to false $E_{\mathrm{T}}^{\mathrm{miss}}$ estimations.

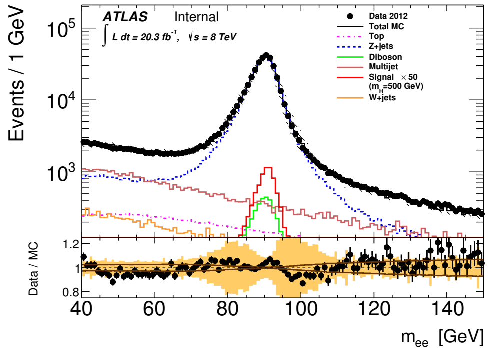
Starting with the reconstruction of the $Z \to \ell\ell$ decay, two main analysis channels are distinguished, one of which is tuned to select signal events produced by gluon fusion (ggF) while the other selects events produced by vector boson fusion (VBF). In both channels the $Z \to q\bar{q}$ decay is reconstructed from the two leading (in p_{T}) jets. The former is called the "resolved-ggF" channel. In the high mass region ($m_H > 700$ GeV), the "merged-ggF" channel is also added, in order to recover case where the $Z \to q\bar{q}$ decay is reconstructed as a single jet with a large mass. The classification of events into the three channels is described in Section 5.3.2.

5.3.1 $H \to ZZ \to \ell^+\ell^- q\bar{q}$ selection

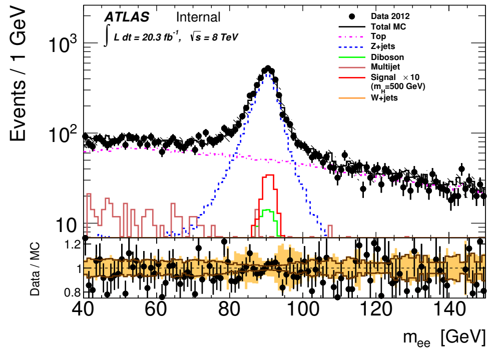
The $Z \to \ell\ell$ decay is reconstructed from two same-flavour loose leptons, one of which should pass the tight criteria, as defined in Sections 5.2.1 and 5.2.2 for muons and electrons respectively. For muons, opposite charge sign is also required. Although this requirement would also be expected for electrons, it is not applied due to the higher rate of charge misidentification[2]. The event is rejected if any additional loose leptons are found.

The invariant mass, $m_{\ell\ell}$, of the selected lepton pair is required to be close to the nominal Z boson mass, in the range $83 < m_{\ell\ell} < 99$ GeV. This requirement suppresses backgrounds without a resonant lepton pair, namely top quark and multijet production. Since no high-p_{T} neutrinos are expected in the final state, the "$E_{\mathrm{T}}^{\mathrm{miss}}$-significance", defined as $(E_{\mathrm{T}}^{\mathrm{miss}}/\ \mathrm{GeV})/\sqrt{(H_T/\ \mathrm{GeV})}$ (all quantities in GeV), where H_{T} is the vector-sum of momenta of all loose leptons and jets, is required to be less than 6. For cases of the resolved-ggF channel with 2 b-tagged jets, the cut value is tightened to 3.5, due to the larger top quark background. Using $E_{\mathrm{T}}^{\mathrm{miss}}$-significance is prefered over a cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ itself, since the former has been found to provide a roughly constant efficiency versus m_H, while with $E_{\mathrm{T}}^{\mathrm{miss}}$ the efficiency

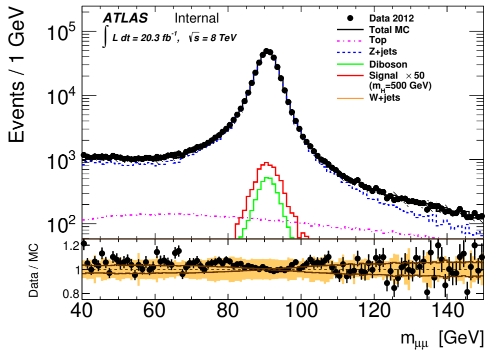
[2]The charge-identification efficiency for electrons has been measured with 2011 ATLAS data [114]. It is $> 99\%$ for reconstructed electrons everywhere in barrel region but decreases to $\sim 93\%$ at the boundaries of the tracking acceptance.



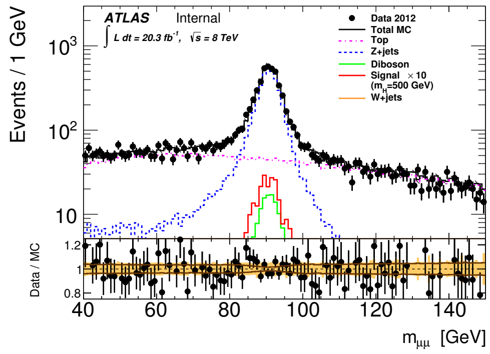
(a) ee, < 2 b-tagged jets.



(b) ee, 2 b-tagged jets.



(c) $\mu\mu$, < 2 b-tagged jets.



(d) $\mu\mu$, 2 b-tagged jets.

Figure 5.5: $m_{\ell\ell}$ spectrum for events with $\geq$ 2 jets after the $E_{\mathrm{T}}^{\mathrm{miss}}$-significance requirement in the (a) untagged electron, (b) tagged electron, (c) untagged muon, and (d) tagged muon cases of the resolved ggF channel.

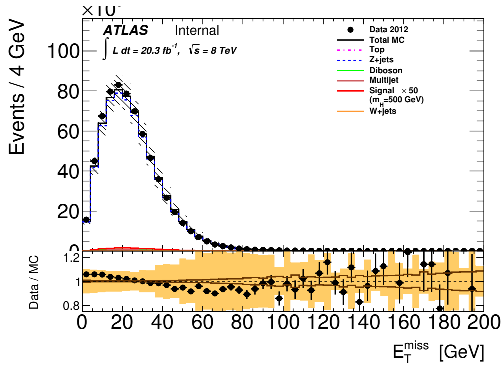
decreases as m_H increases. The dilepton mass spectrum is shown in Figure 5.5; the $E_{\mathrm{T}}^{\mathrm{miss}}$-significance and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions are shown in Figure 5.6.

The analysis next requires at least one jet with $p_{\mathrm{T}} > 45$ GeV which has been found to reduce the Z+jets background. Finally, for the reconstruction of the $Z \to q\bar{q}$ decay, the analysis is split into three channels; the resolved-ggF channel, the merged-ggF channel and the VBF channel which are discussed in the sections that follow.

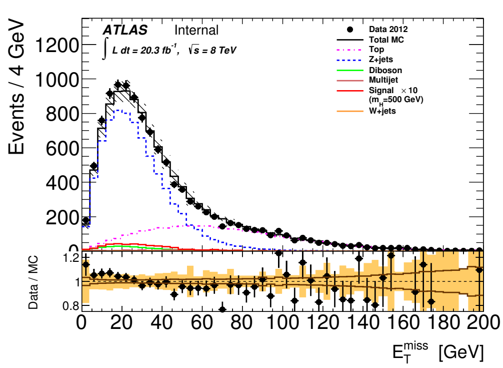
5.3.1.1 Resolved ggF channel

Over most of the mass range under study ($m_H \lesssim 700$ GeV), the $Z \to q\bar{q}$ decay produces two well separated jets that can be individually resolved. Hence, events are required to contain at least two tight jets, in addition to the selected pair of leptons. Figure 5.7 shows the distribution of jet multiplicity after the dilepton selection.

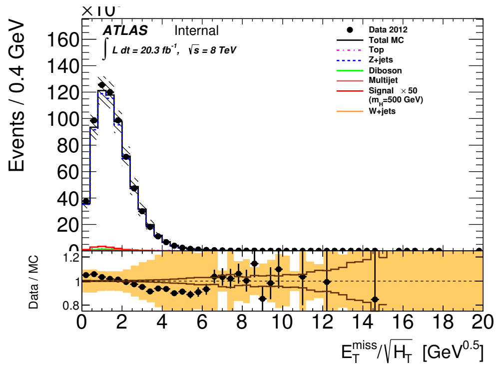
Next, the analysis classifies events into "tagged", with two b-tagged jets, and "untagged", with less than two b-tagged jets. This classification intends to exploit the fact that a significant fraction ($\sim$22%) of signal events are expected to contain b jets, from $Z \to b\bar{b}$ decays [72], which are rare in the dominant $Z(\to \ell\ell) +$jets background.



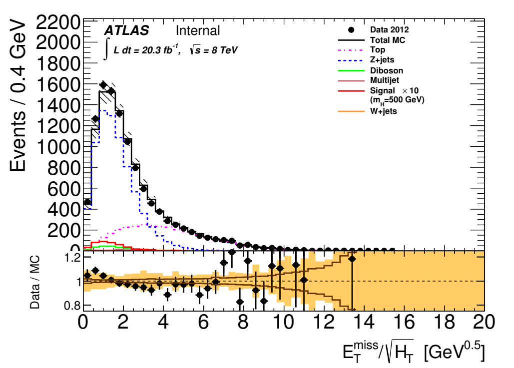
(a) < 2 b-tagged jets.



(b) 2 b-tagged jets.



(c) < 2 b-tagged jets.



(d) 2 b-tagged jets.

Figure 5.6: $E_{\mathrm{T}}^{\mathrm{miss}}$distribution for events with ≥ 2 jets after the $m_{\ell\ell}$ selection in the (a) untagged and (b) tagged (resolved) ggF channels. The $E_{\mathrm{T}}^{\mathrm{miss}}$-significance in the same regions is shown in (c) and (d).

Events with more than two b-tagged jets are discarded. The distribution of the `MV1c` discriminant and the resulting multiplicity of b-tagged jets are shown in Figure 5.8.

The $Z \to q\bar{q}$ decay is reconstructed as follows. In the tagged event category, the two b-tagged jets are selected, while for events with no b-tagged jets, the two leading in p_{T} jets are chosen. In case there is exactly one b-tagged jet then that jet is selected along with the leading in p_{T} non b-tagged jet, as it is likely that one of the b jets from a $Z \to b\bar{b}$ decay fails to be identified. Once the jet candidates have been selected, it is required that the invariant mass of the dijet system, m_{jj}, is close to the nominal Z boson mass, within the range $70 < m_{jj} < 105$ GeV. The m_{jj} window is wider than the one for $m_{\ell\ell}$ due to the larger energy resolution for jets. The distribution of m_{jj} in the two event categories is shown in Figure 5.9.

The mass of the $H \to ZZ$ candidate is estimated from the invariant mass of the $\ell\ell jj$ system, $m_{\ell\ell jj}$. The mass resolution is improved by imposing a Z-mass constraint on the two jets, since the reconstructed $Z \to jj$ mass resolution is worse than the intrinsic Z resonance width. A simple approach is chosen; the four-momentum of each jet is scaled by m_Z/m_{jj}. More complicated approaches, involving kinematic fits, were not found to provide significant improvement in sensitivity.

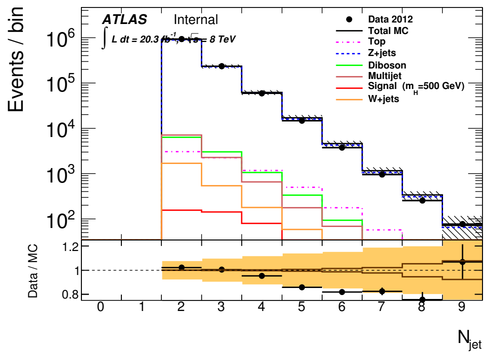
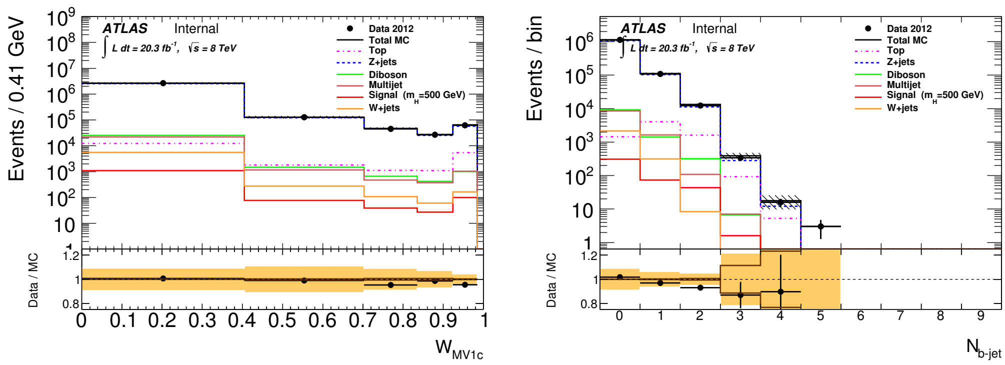


Figure 5.7: Number of tight jets after the $m_{\ell\ell}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_T}$ cuts.



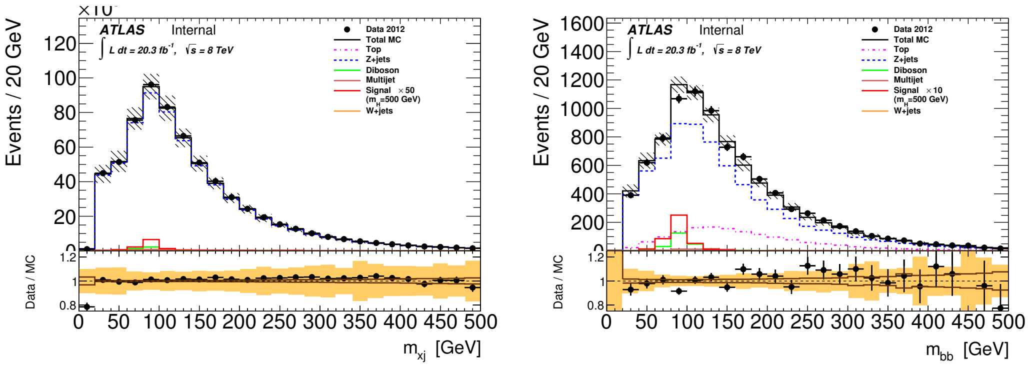
(a) (b)

Figure 5.8: Distribution of (a) the `MV1c` b-tagging discriminant for all tight jets and (b) the number of b-tagged jets after the $m_{\ell\ell}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_T}$ cuts.

Once the $\ell\ell jj$ candidate has been found, an optimized kinematic selection is applied for further separation between signal and background. Several variables have been investigated; the most sensitive ones are the transverse momenta p_{T}^{j} of the selected jets, the transverse momentum $p_{\mathrm{T}}^{\ell\ell}$ of the dilepton system and the azimuthal angle $\Delta\phi_{\ell\ell}$ between the two leptons. The selection has been optimized as a function of the Higgs candidate mass and the resulting criteria are the following:

$$\begin{aligned}
p_{\mathrm{T}}^{j} &> 0.1 \times m_{\ell\ell jj} \text{ for untagged} \qquad (5.1)\\
p_{\mathrm{T}}^{\ell\ell} &> \min[-54.04 + 0.455 \times m_{\ell\ell jj}, 275] \text{ for untagged}\\
&> \min[-79.18 + 0.439 \times m_{\ell\ell jj}, 275] \text{ tagged,}\\
\Delta\phi_{\ell\ell} &< 3.22 \times 10^{8}/{m_{\ell\ell jj}}^{3.50} + 1 \text{ for untagged,}
\end{aligned}$$

where all kinematic quantities are expressed in GeV. The criteria for p_{T}^{j} and $\Delta\phi_{\ell\ell}$ are not applied on the tagged sample since they provide no further improvement to the significance after the $p_{\mathrm{T}}^{\ell\ell}$ requirement. The cut values are defined as functions of the reconstructed mass $m_{\ell\ell jj}$ rather than the nominal Higgs mass m_H since the



(a) < 2 b-tagged jets (b) 2 b-tagged jets

Figure 5.9: Dijet mass distribution for the (left) untagged and (right) tagged samples in the ggF channel. x indicates either a tagged or untagged jet and xj represents the two jets in the untagged sample, which combines 0 and 1 b-tag events.

latter would force the dominant Z+jets background to peak under the Higgs signal; this is undesirable since any uncertainties in the size and shape of the background distribution will have a large effect. In addition, using $m_{\ell\ell jj}$ allows having a single background shape for all m_H hypotheses. Further details on the optimization of the above selection can be found in Ref. [115] and also in Appendix D where the same criteria are examined for the case of the VBF channel.

5.3.1.2 Merged (boosted) ggF channel

In the high mass range ($m_H \geq 700$ GeV), the Z bosons produced from a Higgs boson decay are highly boosted, hence the opening angle between the decay products is expected to be small in the lab frame. Since reconstructed jets have a finite size (see Section 5.2.3), a boosted, hadronically decaying Z boson is likely to give rise to jets that overlap with each other. Such cases would normally be discarded by an event selection requiring two resolved jets with m_{jj} close to the Z boson mass. The "merged" selection attempts to recover this efficiency loss by searching for $Z \to q\bar{q}$ decays which are reconstructed as a single jets.

The merged selection is applied on events that fail the selection for the resolved-ggF channel. The first case of such events are those containing only one high-p_{T} tight jet; such events are named "monojet" events. The second case are events where two or more tight jets exist, but with the selected jets having invariant mass outside the range $50-150$ GeV, thereby failing both the Z boson mass window and its sidebands, which are used to control the Z+jets background (see Section 5.4). Furthermore, the $Z \to \ell\ell$ candidate is required to have a large boost, with $p_{\mathrm{T}}^{\ell\ell} > 280$ GeV. Then, the leading in p_{T} jet is selected as the $Z \to q\bar{q}$ decay candidate and it is required to satisfy the kinematic requirements:

$$\begin{aligned} p_{\mathrm{T}}^{j} &> 200\ \mathrm{GeV} \\ m_j/p_{\mathrm{T}}^{j} &> 0.05 \end{aligned} \tag{5.2}$$

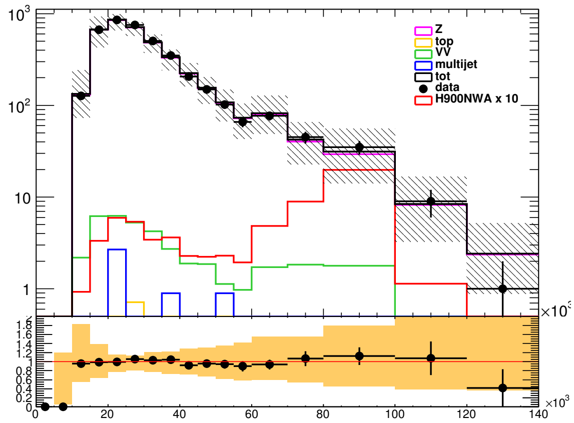


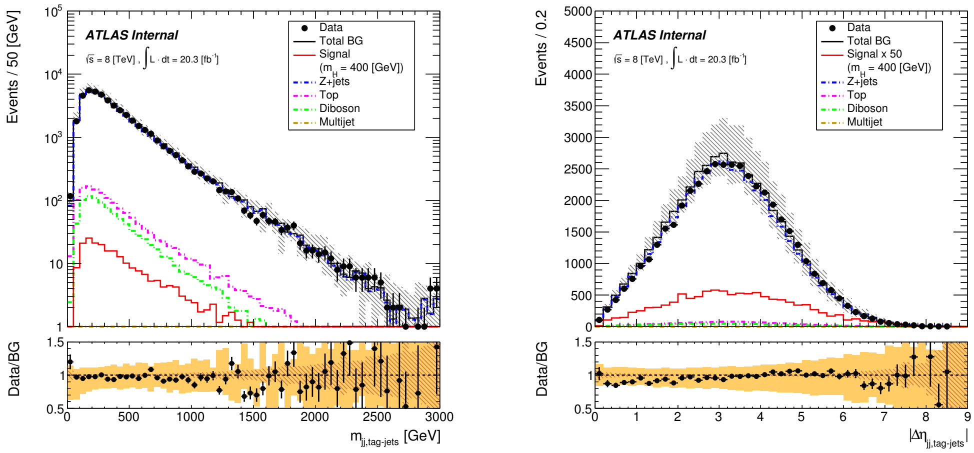
Figure 5.10: Invariant mass of the leading jet in events selected by the merged analysis, after the mass calibration and the kinematic selection.

where m_j is the single jet invariant mass. The distribution of m_j is shown in Figure 5.10 for the selected jets. It is observed that a mass window requirement, namely $70 < m_j < 105$ GeV, separates the signal from most of the non-resonant background. Calibration of the jet mass has been carried out by studying its response (m_j^{reco}/m_j^{truth}) as a function of several variables, both in QCD multijet and signal samples. Unlike the resolved-ggF channel selection, the merged selection does not distinguish events according to the number of b-tagged jets since the b-tagging efficiency is poor for merged jets. The final discriminating variable in this channel is the mass $m_{\ell\ell j}$ of the $\ell\ell j$ system.

5.3.1.3 VBF channel

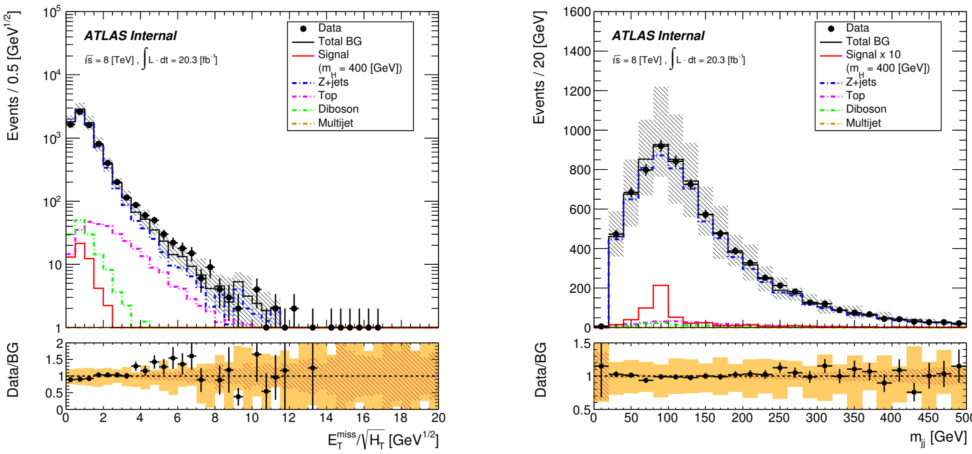
In the VBF production mechanism, two jets (called "tag-jets") are expected close to the beam pipe, in opposite hemispheres, in addition to the jets from the $Z \to q\bar{q}$ decay. On that basis, the selection begins by looking for two non b-tagged jets, with $\eta_{\mathrm{j1}} \times \eta_{\mathrm{j2}} < 0$. If more than one such pair is found, the one with the highest invariant mass is kept. The distribution of the invariant mass $m_{jj,\mathrm{tag}}$ and the pseudorapidity gap $\Delta\eta_{jj,\mathrm{tag}}$ between the tag-jets are shown in Figure 5.11. In order to reduce the Z+jets background, the tag jets are required to satisfy $m_{jj,\mathrm{tag}} > 500$ GeV and $\Delta\eta_{jj,\mathrm{tag}} > 4$. These cuts have been optimized for maximum significance, as discussed in Appendix D.

Once the tag-jets have been identified, the $Z \to q\bar{q}$ decay is reconstructed as in the resolved-ggF channel (Section 5.3.1.1), using the remaining jets. Due to limited statistics, categorization of events according to the number of b-tagged jets does not improve the significance, hence inclusive treatment is followed. The reconstructed dijet invariant mass m_{jj} is shown in Figure 5.12. Finally, optimization of the kinematic selection, as outlined in Section 5.3.1.1 for the resolved-ggF channel, has also been performed for the VBF case. The same selection has been found to be close to optimal (within limited statistics), therefore the $m_{\ell\ell jj}$-dependent criteria of Equation (5.1) are applied. Studies for the optimization of the criteria for the VBF channel



(a) (b)

Figure 5.11: Distribution of the invariant mass (left) and pseudorapidity gap (right) for the VBF tag-jet pair. In the VBF channel, ALPGEN is used for the $Z+$jets background.



(a) (b)

Figure 5.12: Distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (left) and the invariant mass of the dijet pair forming the $Z \to q\bar{q}$ candidate in the VBF channel (right).

are presented in Appendix D.

5.3.2 Summary of event selection criteria

As discussed in the previous sections, events are classified into three categories (resolved-ggF, merged-ggF, VBF) according to the topology of the reconstructed jets. The criteria that determine this classification are presented in Table 5.6 which

Criteria	Description
Pre-selection	Triggered event.
	Primary vertex formed by ≥ 3 tracks with $p_{\mathrm{T}} > 150$ MeV.
Leptons	Exactly 2 leptons (VeryLooseLH electrons/CB+ST muons).
	Opposite charge (muons only).
	Trigger matching.
	$83 < m_{\ell\ell} < 99$ GeV.
VBF candidate	Among pairs of:
	- non-b-tagged jets
	- with $\eta_{\mathrm{jet1}} \times \eta_{\mathrm{jet2}} < 0$,
	obtain the one with the highest invariant mass.
	Require $m_{jj,\mathrm{tag}} > 500$ GeV and $\lvert\Delta\eta_{jj,\mathrm{tag}}\rvert > 4$.
$Z \to qq$	≥ 2 tight jets (selection based on the number of b-tagged jets).
Resolved-ggF candidate	If the event is not a VBF candidate, then require:
$Z \to qq$	≥ 2 tight jets (selection based on the number of b-tagged jets)
untagged-ggF	< 2 b-tagged jets.
tagged-ggF	2 b-tagged jets.
	$50 < m_{jj} < 150$ GeV (signal region and control region).
Merged-ggF candidate	($m_H \geq 700$ GeV) if the event fails the above criteria, then require:
	$p_{\mathrm{T}}^{\ell\ell} > 280$ GeV
$Z \to qq$	≥ 1 tight jet (selection of the leading, in p_{T}, jet).
	Events with > 2 b-tagged jets are rejected.
	$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_T} < 6.5$ GeV$^{1/2}$ (3.5 GeV$^{1/2}$ for tagged-ggF).
	$p_{\mathrm{T}}^{j\ leading} > 45$ GeV.
	$70 < m_{jj} < 105$ GeV (m_j for merged-ggF).
Optimized selection	$p_{\mathrm{T}}^{j} > 0.1 \times m_{\ell\ell jj}$ (untagged-ggF/VBF).
	$p_{\mathrm{T}}^{j} > 0.1 \times m_{\ell\ell jj}$ (untagged-ggF/VBF).
	$p_{\mathrm{T}}^{\ell\ell} > \min[-54.04 + 0.455 \times m_{\ell\ell jj}, 275]$ (untagged-ggF/VBF).
	$p_{\mathrm{T}}^{\ell\ell} > \min[-79.18 + 0.439 \times m_{\ell\ell jj}, 275]$ (tagged-ggF).
	$\Delta\phi_{\ell\ell} < 3.22 \times 10^{8}/m_{\ell\ell jj}{}^{3.50} + 1$ (untagged-ggF/VBF).
	$p_{\mathrm{T}}^{j} > 200$ GeV (merged-ggF).
	$m_j/p_{\mathrm{T}}^{j} > 0.05$ (merged-ggF).

Table 5.6: Summary of event selection criteria. The definitions of object candidates are given in Tables 5.2- 5.5.

summarizes the event selection, as described in Sections 5.3.1.1, 5.3.1.2 and 5.3.1.3. The efficiencies of the applied selection criteria are listed in Appendix E.

5.4 Background Modelling

The MC simulations that model the various backgrounds to this analysis may not describe the data accurately. Such backgrounds are studied by defining control regions (CRs) in the data, dominated by the background under study and mostly free of signal. Corrections to the normalization and/or shape of the background is then possible by direct comparison of the prediction with the data in that region.

In summary, the shape and the normalization of the Z+jets background is corrected from data while the top background shape is taken from MC but normalized to the data in the CRs; the multijet background is taken from a purely data-driven method while all other minor backgrounds are taken entirely from MC simulation.

5.4.1 Z+jets background

The modelling of the Z+jets production is studied using the sidebands (SBs) of the m_{jj} distribution, which provide an almost pure sample of Z+jets events. The control region is defined by following the nominal selection, but replacing the m_{jj} criterion with $50 < m_{jj} < 70$ GeV (low-mass SB) or $105 < m_{jj} < 150$ GeV (high-mass SB). The SBs are chosen to be close to the signal region (SR), so that kinematics are kept similar, but the ranges are wide enough to provide sufficient statistics for the correction of the background. The CR is defined as the union of the two sidebands, although each sideband is also tested separately in order to appraise systematic uncertainties.

5.4.1.1 Z+jets in the resolved ggF channel

For the resolved-ggF channel, which distinguishes events according to the number of b-tagged jets, the proportion of Z+light-, c- and b-jet events are corrected from the data by fit. This "flavour fit" employs separate CRs for events with 0, 1 and 2 b-tagged jets, and exploits the `MV1c` discriminant, which is designed to distinguish the various jet flavours and in particular b-jets from both light- and c-jets. The simulated Z+jets events, entering the fit, are classified into subsamples ($Z + bb, bc, bl, cc, cl, ll$), according to the true flavour of the selected jets (see Section 5.2.3.1). The input distribution to the flavour fit is the sum of the `MV1c` weights of the selected jets. Specifically, for each jet, the central value of the bin, in which the `MV1c` weight lies, is acquired, following the binning of the `MV1c` distribution that was used to calibrate the tagger. Therefore, discrete values are assigned to the various jet combinations. In order to correctly take into account correlations between systematic uncertainties, the flavour fit is performed as part of the final profile likelihood fit, described in Chapter 6. Finally, the fit also handles the overall normalization of the Z+jets MC simulation to the data.

Sherpa does not model well the azimuthal separation $\Delta\phi_{jj}$ of the jets from the hadronic Z decay; discrepancy is observed in both the 0 and 1 b-tag event categories, at $p_{\mathrm{T}}^{\ell\ell} < 120$ GeV. This suggests that, at low $p_{\mathrm{T}}^{\ell\ell}$, a correction is needed for the Z+light-jet Sherpa sample which is dominant in the 0 b-tag and non-negligible in the 1 b-tag category. The correction is derived using the $\Delta\phi_{jj}$ distribution in the CR, for 0 b-tag events, with $p_{\mathrm{T}}^{\ell\ell} < 120$ GeV, and it is used to reweight the whole Z+light-jets MC simulation. Discrepancy is further observed in the description of the $p_{\mathrm{T}}^{\ell\ell}$ distribution, in the 1 and 2 b-tag cases, indicating that a $p_{\mathrm{T}}^{\ell\ell}$ correction is needed for the $Z + c/b$ Sherpa samples. Due to limited statistics in the 2 b-tag category, the correction is derived using the 1 and 2 b-tag distributions combined. A systematic uncertainty of half of the correction is assigned to the entire Z+jets simulation, but decorrelated between Z+light jets and $Z + b/c$.

Mismodelling of the m_{jj} distribution does not directly affect the $m_{\ell\ell jj}$ discriminant, since the dijet invariant mass is constrained to the Z boson mass when reconstructing $m_{\ell\ell jj}$. However, the modelling of this variable is still important since it affects the extrapolation between the Z+jets CR and the SR of both the normalization scale factor and the flavour fit. Consequently, no correction is applied, but a systematic uncertainty is associated to the shape description, to account for any residual data/MC disagreement in the CR, following the procedure described for the VBF case, in Section 5.4.1.3.

Figure 5.13 shows the final $m_{\ell\ell jj}$ discriminant in the Z+jets CR after the application of all the modelling corrections described in this section, including the relative flavour composition from the final combined fit. The distributions are shown both before and after the optimized selection of Equation (5.1). The data are well described by the MC simulation within the assigned uncertainties. Consequently, no further correction or uncertainty is considered.

5.4.1.2 Z+jets in the merged ggF channel

The Z+jets process is also the dominant background in the merged channel. Similarly to the resolved ggF channel, the control region is obtained from the sidebands of the m_j distribution, defined in the interval $30 < m_j < 70$ GeV. The distribution of the three-body mass $m_{\ell\ell j}$ in the sidebands is shown in Figure 5.14. It appears that the MC simulation reproduces correctly the shape of the data distribution, however there is disagreement in the normalization. Therefore, the CR is included in the final profile likelihood so that the Z+jets normalization can be determined from the fit.

5.4.1.3 Z+jets in the VBF channel

In the VBF channel, the Z+jets background normalization is estimated from the data, using the m_{jj} SBs. Due to limited statistics, the CR is defined at an earlier stage of the event selection, right after the acquisition of the $Z \to q\bar{q}$ decay candidate and before the optimized selection. The normalization of the final $m_{\ell\ell jj}$ distribution is derived as part of the final profile likelihood fit, as described in Chapter 6. The systematic uncertainty associated to the extrapolation of the normalization between the CR and the signal region, is estimated from the m_{jj} distribution, as in the case of the resolved-ggF channel. The uncertainty is evaluated by reweighting the Z+jets simulation so that it covers any residual data/MC disagreement in the CR, both from above and from below, as shown in Figure 5.15. The reweight used is a linear parameterization of the form $a(m_{jj} \times 10^{-3} - b)$.

The Z+jets CR is also used to derive a correction to modelling of the $m_{\ell\ell jj}$ distribution. The correction is derived directly from the binned ratio of the data, after subtracting the small contributions from the other background processes, to the Z+jets distribution. In order to reduce the effect of statistical fluctuations, the bin widths are defined so that all bins have less than 10% statistical uncertainty. The derived shape correction is then propagated to the signal region and to the later stages of the selection. Figure 5.16 shows the agreement of the MC simulation to the data after the correction has been applied. The entirety of the shape variation caused by the correction is taken as systematic uncertainty.

The variables that were used to correct the modelling of the $Z+$jets background in the ggF channel, were also tested for the VBF channel. However, the latter uses the Alpgen MC generator, as discussed in Section 5.1.2.1, so the variables did not provide a better description of the background than the above approach.

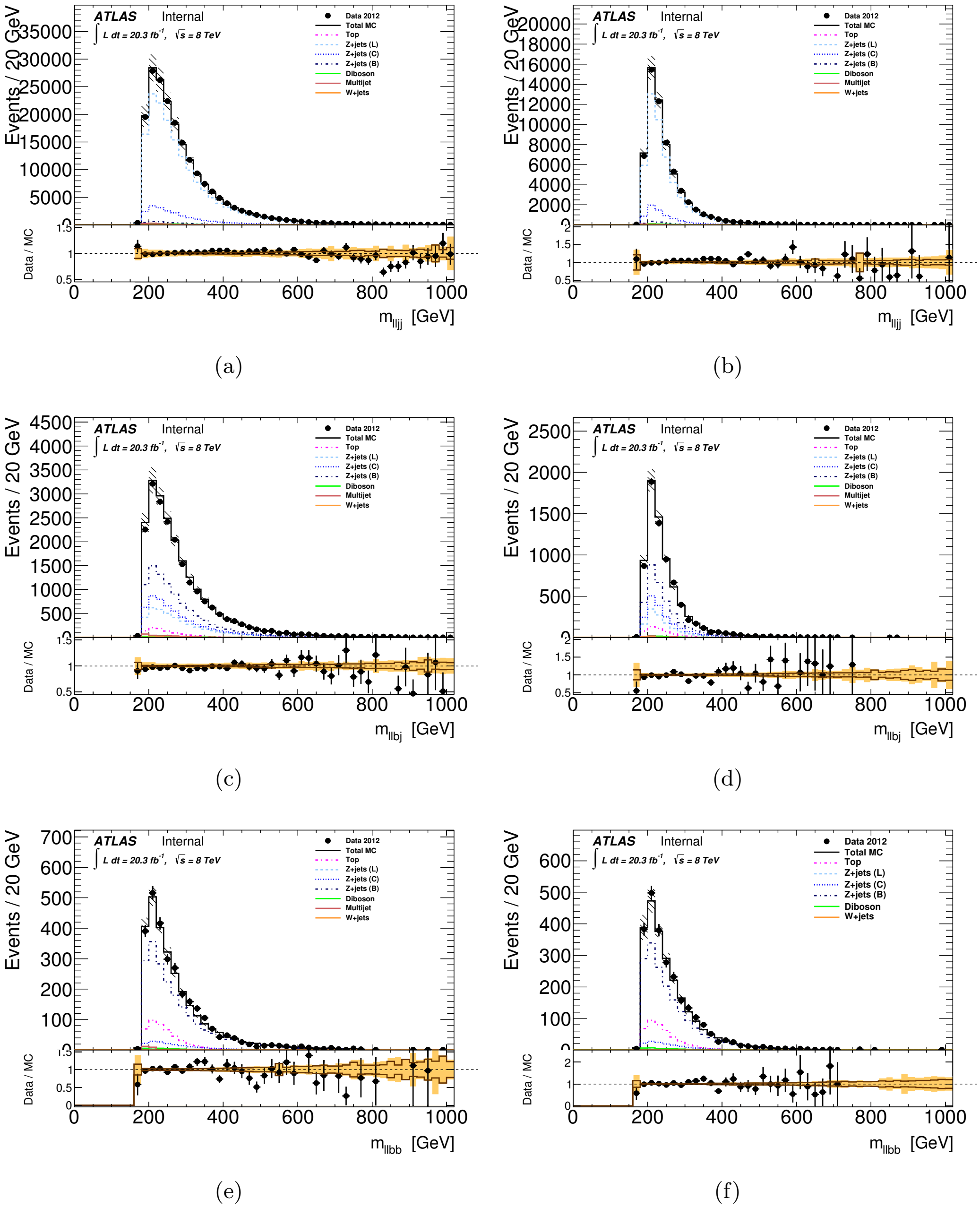
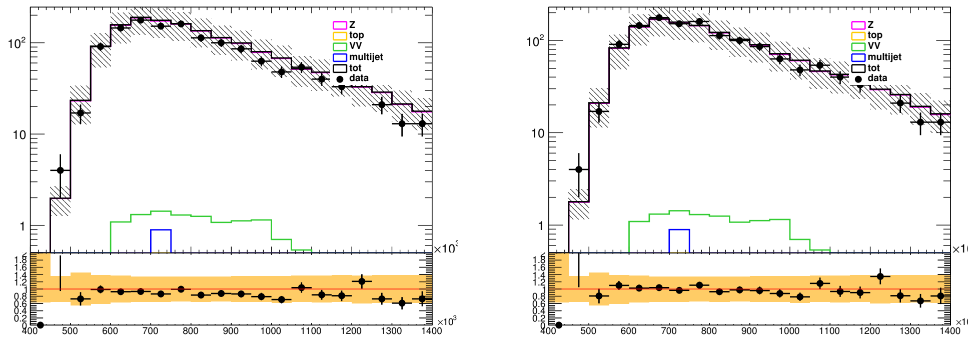


Figure 5.13: $m_{\ell\ell jj}$ distribution for (a,c) 0, (b,d) 1, and (e,f) 2 b-tag events, after the corrections in the text, before (left) and after (right) the optimized selection. The systematic uncertainty bands include the uncertainties associated with these corrections.



(a) $30 < m_j < 70$ GeV, before scale factor. (b) $30 < m_j < 70$ GeV, after scale factor.

Figure 5.14: $m_{\ell\ell j}$ distribution in the m_j sideband (a) before any correction and (b) after the correction.

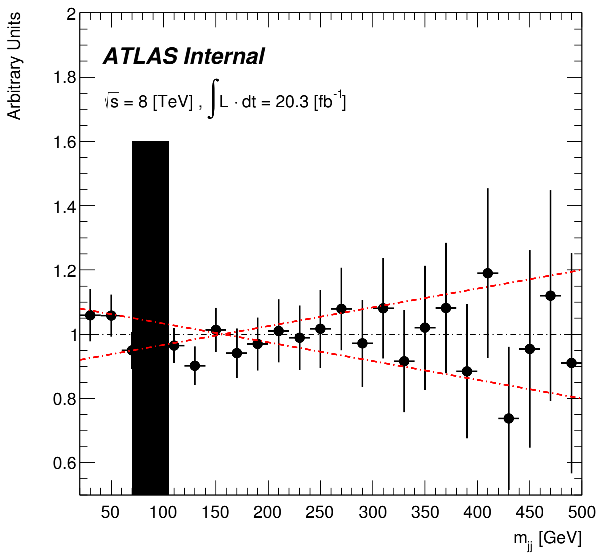
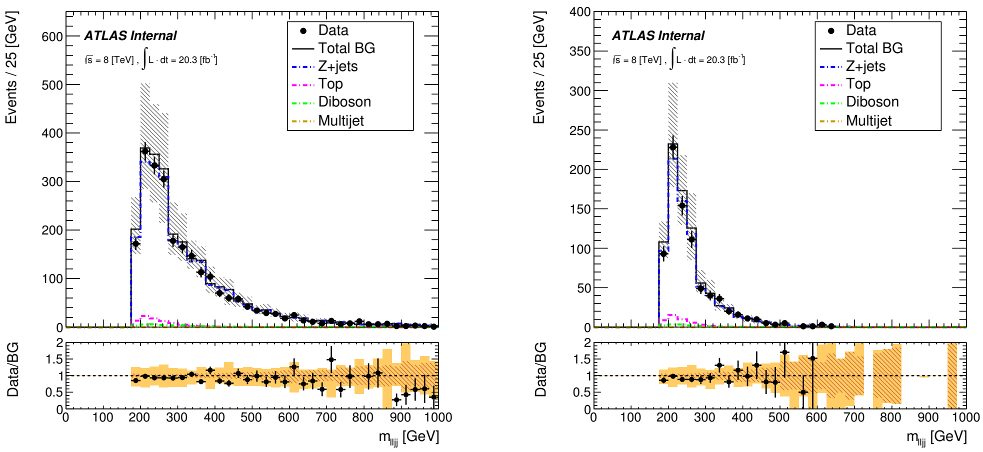


Figure 5.15: Ratio of the m_{jj} distribution in the data, after subtracting the non-Z simulated backgrounds, over the nominal Z+jets MC expectation. The signal region is blinded. The dashed lines show the reweighting of the Z+jets distribution in order to acquire each of the $\pm 1\sigma$ systematic variations associated to the normalization from the Z + jets CR.



(a) (b)

Figure 5.16: $m_{\ell\ell jj}$ distribution in the CR of the Z+jets background, in the VBF channel, before (a) and after (b) the optimized selection.

5.4.2 Top quark background

Top quark production is a significant background in the 2 b-tag ggF event category. This background is dominated by $t\bar{t}$ decays in which both W bosons decay into leptons, resulting in a final state with a pair of leptons and a pair of b-jets, the invariant masses of which happens to be close to that of a Z boson. The contribution from single top production, primarily Wt, is small. In the 2 b-tag sample, only 3.3% of the top quark background is from single top processes and 85% of that comes from Wt production. Another source of top quark background includes leptons that originate from decays of the b-jet daughters of top quarks.

A sample dominated by top quark processes is obtained by selecting events with opposite-flavour (i.e., $e\mu$), opposite-charge leptons. The remaining parts of the nominal selection are then applied, including the $m_{\ell\ell}$ requirement. Since top quark production is a small background in all other cases, the control region is defined primarily for 2 b-tag events. However, in order to cross-check the results with higher statistics, the 1 b-tag sample control region, is also studied. For the latter, a $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_T} > 3.5$ requirement is applied to obtain a sample dominated by $t\bar{t}$ decay. Figure 5.17 shows the m_{jj} and $m_{\ell\ell jj}$ distributions in the top quark CR for 1 and 2 b-tag events. The data are reasonably described by MC simulation, hence no corrections are applied. As regards the normalization of the top quark background, it is determined from the final fit, using the $m_{\ell\ell jj}$ distribution from the 2 b-tag control region, as described in Chapter 6.

Finally, the unfolded 7 TeV ATLAS $t\bar{t}$ measurement [116] shows that the top quark p_{T} distribution in MC simulation is harder than that observed in data. This difference between data and MC is used to correct the $t\bar{t}$ simulation and half or double of the correction applied is taken as a systematic uncertainty (see Section 5.5.2.2).

5.4.3 Multijet background

Multijet events form a background whenever two jets are mistakenly identified as leptons and their invariant mass happens to be compatible to the Z boson mass. Photon conversions also contribute in the case of electrons, while in-flight pion decays add to the muons channel. In addition to "fake" leptons, true leptons from the semi-leptonic decay of heavy flavour hadrons may also contribute to this background in both the electron and muon channels.

The multijet background in the electron channel is estimated from the data. The templates to describe the shape of the multijet background in the various distributions are obtained from a region dominated by multijet events, defined by reversing the track isolation criterion of the two electrons and applying the remaining analysis selections. The normalization of the templates to the nominal region is estimated by fitting the dielectron invariant mass distribution at an early stage of the event selection, right after the requirement of ≥ 2 jets. The fit is performed over the range $40 < m_{ee} < 150$ GeV using three components:

- The multijet template, derived from data using the loosened electron selection.
- The $Z \to ee$ background distribution from the Monte Carlo simulation, using the nominal electron selection.

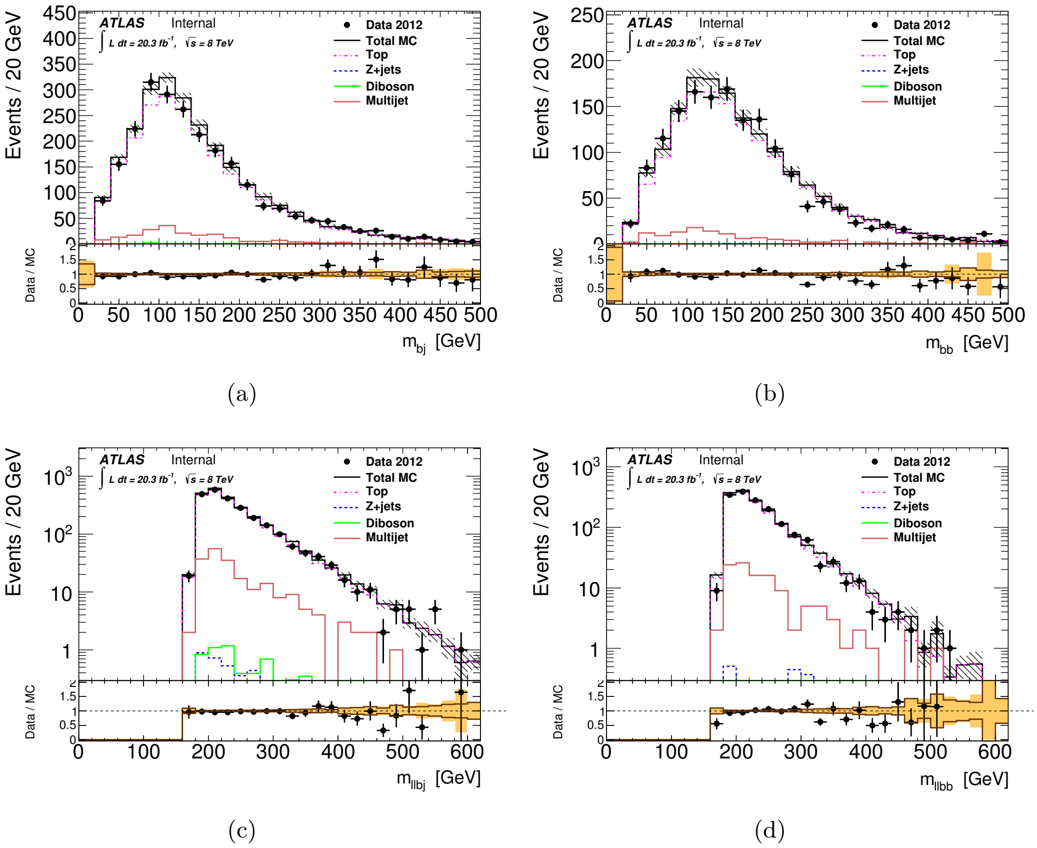


Figure 5.17: The (a,b) m_{jj} and (c,d) $m_{\ell\ell jj}$ distributions for the top quark CR in the samples with 1 (left) and 2 (right) b-tagged jets.

- The sum of all the other background distributions from the Monte Carlo simulation, using the nominal electron selection.

Only the normalization of the multijet template and $Z \to ee$ background are allowed to vary; the other backgrounds are fixed. The fit is performed separately for events with 0, 1, and 2 b-tagged jets, however, due to the low statistics and the large contribution from top quark background in the 2 b-tag category, the scale factor obtained for the 0 b-tag case is used for all b-tag event categories and a 50% uncertainty is assigned to account for the difference. In the $e\mu$ TopCR, the residual small multijet background is taken from the data, by selecting $e\mu$ pairs of same charge; the same sample also accounts for the small W+jets background in the particular region.

The multijet background in the muon channel was investigated by comparing the MC simulation with the data in the $m_{\mu\mu}$ sidebands (see Figure 5.5) and was found to be negligible. In the merged-ggF channel, the multijet background is estimated as described above; the estimate is carried out separately for the mono-jet and multiple-jet subsamples.

5.4.4 Diboson background

Since no suitable data CR can be defined for the diboson background, it is estimated directly from MC simulation, as discussed in Section 5.1.2.3.

5.4.5 W+jets background

The W+jets background is found to be negligible except for the top quark CR, where the small contribution is taken, together with the multijet background, by selecting $e\mu$ lepton pairs of same charge.

5.4.6 SM $Zh, h \to bb$ production

The SM $Zh(h \to bb)$ production is taken from MC simulation, as described in 5.1.2.4, and is found to be negligible, contributing to the background by $\sim 0.5\%$ in the Z+jets CR of 2 b-tag sample. Although, in the BSM senarios probed in this search, the coupling may not have the value predicted by the SM, measurements [113] have shown that it is $\mu_{bb} = 0.52 \pm 0.32$ (stat.) $\pm\, 0.24$ (syst.), therefore compatible with the SM prediction.

5.5 Systematic uncertainties

The following sections describe the systematic uncertainties on the measurement itself (Section 5.5.1) as well as those associated with the modelling of the signal and background processes (Section 5.5.2). Except where explicitly specified, systematic uncertainties are treated as fully correlated across all subchannels/categories (a given correlated uncertainty is modelled in the fit by using a nuisance parameter common to all categories/subchannels). The name of the nuisance parameter for each systematic uncertainty, as used in the fit model described in Chapter 6, is given in `typewriter text`.

5.5.1 Experimental systematic uncertainties

The uncertainties regarding efficiency corrections and calibrations of simulated objects are summarized in Table 5.7. The largest uncertainties of this type are those on jets, as can be seen in Section 6.5. Details are given in the sections that follow. The majority of these uncertainties are provided by the combined performance groups of ATLAS.

5.5.1.1 Luminosity and pile up

The uncertainty on the integrated luminosity is determined to be $\pm 2.8\%$ in a calibration following the methodology detailed in Ref. [38], using beam-separation scans. This uncertainty is applied to the signal and to those backgrounds which are determined from MC simulation (i.e. only the diboson background). There is also an uncertainty of 4% in the average number of interactions per bunch crossing, which leads to an uncertainty on distributions sensitive to pile-up.

Nuisance parameter	Description	NP count	Section
	Luminosity (2)		
`Lumi`	Total integrated luminosity	1	5.5.1.1
`MuScale`	$<\mu>$ (average number of interactions per bunch crossing) profile	1	
	Electrons (3)		
`ElecEffic`	Reconstruction and identification efficiencies	1	5.5.1.2
`ElecE`	Energy scale	1	
`ElecEResol`	Energy resolution	1	
	Muons (3)		
`MuonEffic`	Trigger, reconstruction, and identification efficiencies	1	5.5.1.2
`MuonEResolID`	Energy resolution from inner detector	1	
`MuonEResolMS`	Energy resolution from muon system	1	
	Jet energy scale (23)		
`JetNPX`	Eigenvector decomposition of in-situ calibration studies (X = 1-6rest)	6	5.5.1.3.1
`JetEtaModel`	η inter-calibration model	1	
`JetEtaStat`	Statistical uncertainty of η inter-calibration	1	
`JetNonClos`	Calibration non-closure	1	
`JetMu`	μ correction based on average number of pile-up interactions	1	
`JetNPV`	Uncertainty due to correction of the number of primary vertices	1	
`JetPilePt`	Pile-up in jet area correction	1	
`JetPileRho`	Pile-up in jet area correction	1	
`JetFlavB`†	b-jet energy scale	1	
`JetBE`†	b-jet scale for μ and ν energy uncertainties	1	
`JetFlavComp_X`‡	Knowledge of light quark vs gluon fraction	4	
`JetFlavResp_X`‡	Different response of light quarks vs gluon jets	4	
	Jet energy resolution (2)		
`JetEResol`	Resolution applied to all jets	1	5.5.1.3.2
`BJetReso`†	b-jet specific resolution	1	
	Jet quality (1)		
`JetJVF`	Jet vertex fraction efficiency	1	5.5.1.3.3
	$E_{\mathrm{T}}^{\mathrm{miss}}$(2)		
`METResoSoftTerms`	Resolution of soft component	1	5.5.1.4
`METScaleSoftTerms`	Scale of soft component	1	
	Flavour tagging (36)		
`BTagBNEffic`	b-jet uncertainty in 10 eigenvector ($N = 0-9$)	10	5.5.1.5
`BTagCNEffic`	c-jet uncertainty in 15 eigenvector ($N = 0-14$)	15	
`BTagLNEffic`	Light-jet uncertainty in 10 eigenvector ($N = 0-9$)	10	
`TruthTagDR`	Correction to ΔR(cc) bias from truth-tagging	1	

Table 5.7: Summary of experimental systematic uncertainties. The associated nuisance parameter names are also given. For the flavour composition and response systematics, `X`=`Wjets`, `Zjets`, `Top`, or `signal/VV`.
† Applied only to truth-matched b-jets.
‡ Applied only to non-truth-matched jets.

5.5.1.2 Leptons

Systematic uncertainties on the trigger, reconstruction and identification efficiencies for leptons [106, 108, 114] are relatively small ($\mathcal{O}(1\%)$). Since the trigger efficiency for electrons is very high, no uncertainty is applied for it. Each efficiency correction weight is shifted ($\pm 1\sigma$) coherently to evaluate one systematic variation for the combined effect on efficiency.

Uncertainties on the lepton energy and resolution are also taken into account. These uncertainties depend on the object's p_{T} and η as described in Refs [107, 114]. The respective systematic variations are evaluated separately by shifting the energy/resolution and re-running the event selection.

5.5.1.3 Jets

The experimental uncertainties related to jets are uncertainties on the energy scale, the resolution and the efficiency of the jet vertex fraction (JVF) criterion.

5.5.1.3.1 Jet energy scale

The uncertainty on the jet energy scale has several sources, including uncertainties in the in situ calibration analysis, corrections for pile-up, and the flavour composition of the sample [117, 118]. These uncertainties are decomposed into 56 independent components as follows:

- 47 from the various in-situ JES calibration analyses. These are reduced by eigenvector decomposition into 6 parameters.
- 2 from η inter-calibration, estimated by comparison between Pythia and Herwig and including the statistical component from the comparison itself. These uncertainties occur mainly due to difference in the modelling of the additional radiation which may affect the p_{T} and η of the dijet system.
- 1 from the propagation of uncertainties of single hadrons. This uncertainty affects highly energetic jets ($p_{\mathrm{T}} > 1\mathrm{TeV}$), and is ignored as negligible.
- 1 from MC non-closure.
- 4 related to pile-up corrections.
- 1 for the jet responce in the presence of close-by jets (not used).

For central jets, the total relative uncertainty on the jet energy scale ranges from about 3% for jets with a p_{T} of 20 GeV to about 1% for a p_{T} of 1 TeV. Four additional components are included for flavour and kinematic uncertainties:

- 1 for differences in the response of true b-jets, observed across different MC simulations.
- 1 related to μ and ν energies from b-hadron decays (true b-jets only).
- 1 for the sample's unknown mixture of light-quarks and gluons (non-true b-jets only).
- 1 for the difference in response between light-quarks and gluons assessed from MC comparisons (non-true b-jets only).

The b-jet energy scale uncertainty is $\sim$ 1–2%. The flavour composition and response are uncorrelated between different processes as the quark/gluon mixture can be different. The quark/gluon mixture (of non-b-jets) is assumed to be 50% with 100% uncertainty, thus the effect of this uncertainty is maximal. After including all sources of uncertainty, the total fractional systematic uncertainty corresponding to the jet energy scale ranges from $\approx 3\%$ at 20 GeV to $\approx 1\%$ for a 1 TeV jet.

5.5.1.3.2 Jet energy resolution

The relative jet energy resolution varies from $\approx 25\%$ at 20 GeV to $\approx 5\%$ near 1 TeV [119]. Two systematic uncertainties are defined by in-situ studies of the differences in resolution between data and MC simulation; an inclusive uncertainty, applied to all jets, and a second one, specifically on the resolution of true b-jets. The effect of these uncertainties in the analysis is obtained by smearing the p_{T} of each jet according to a Gaussian distribution centered at 1, with a width equal to the true jet resolution plus the the relative uncertainty which depends on the jet's p_{T} and η. The effect on the final discriminating variable is then symmetrized in order to obtain a two-sided uncertainty.

5.5.1.3.3 Jet vertex fraction

The uncertainty on the jet vertex fraction efficiency is estimated from the differences between data and MC simulation in Z+jets events. The respective systematic variation is obtained by varying the cut value of 50% between 47% and 53%, according to the measurements described in [79].

5.5.1.3.4 Jet mass scale

The uncertainty on the jet mass scale is of crucial importance to the merged analysis, where the jet mass is used as discriminant. A flat uncertainty of 10% was found to sufficiently contain the estimations from various comparisons between data and MC simulation. A second source of uncertainty occurs from the different topology characterizing jets produced from decays of boosted bosons against those that originate from parton hadronization. This effect is accounted for with an additional 10%. The total uncertainty for the jet mass scale is obtained by summing in quadrature the two estimates, leading to a total uncertainty of 14%.

5.5.1.4 Missing transverse energy

All systematic variations on object energies are propagated to the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$. Additional uncertainties on $E_{\mathrm{T}}^{\mathrm{miss}}$ originate from variations of the energy scale and resolution of calorimeter clusters which are not associated to reconstructed objects (the so-called "soft term").

5.5.1.5 Flavour tagging

The data-to-MC scale factors, used to correct the b-tagging efficiency in MC simulation according to the data (see Section 5.2.3.2), have been estimated for each jet flavour (see Section 5.2.3.1) as functions of the jet p_{T} and the `MV1c` output. The associated uncertainties receive contributions from experimental components (i.e. JES), theoretical components (i.e. the top quark p_{T} spectrum in $t\bar{t}$ events) as well as the statistical uncertainty of the data in each $p_{\mathrm{T}} \times$ MV1c ($\times\eta$ for light jets) bin. On top of those, an additional uncertainty is introduced to account for the MC-to-MC corrections applied to the scale factors. This uncertainty is defined as half of the applied correction for each MC generator.

As discussed in Section 5.2.3.3, truth-tagging is used in order to populate the $2b$-tag control regions with events that do not contain true b jets. However, a bias has been measured, as a function of ΔR_{jj}, in events with two c jets. Therefore, a correction has been derived for these cases and the respective systematic uncertainty is applied. This effect is not seen in light-light or c-light events.

5.5.2 Signal and background modelling systematics

This section describes the systematic uncertainties on the signal acceptance and interference and the modelling of the Z+jets, top quark, and multijet backgrounds.

5.5.2.1 Z+jets backgrounds

For the $\Delta\phi_{jj}$ correction applied to the Z+light-jet MC samples at low $p_{\mathrm{T}}^{\ell\ell}$ (see Section 5.4.1.1), a systematic uncertainty of half of the applied correction is assigned, while the full correction is considered as systematic uncertainty in the $Z + b/c$-jet simulation, where the correction is not applied. In the high-$p_{\mathrm{T}}^{\ell\ell}$ region, where there is no correction applied to any of the flavours, a linear fit is performed to the data/MC ratio in the 0 b-tag CR and the statistical uncertainty on the fitted slope is taken as a systematic uncertainty for all Z+jets flavours [`ZPtV`]. As regards the correction to the $p_{\mathrm{T}}^{\ell\ell}$ distribution of the $Z + c/b$-jets simulation, an uncertainty of half the correction is considered for the entire Z+jet simulation [`ZPtV`]. All the above uncertainties are treated as uncorrelated between Z+light-jets and $Z + b/c$-jets MC samples.

Finally, mismodelling of the m_{jj} distribution does not directly affect the $m_{\ell\ell jj}$ discriminant, since the dijet invariant mass is constrained to the Z boson mass when reconstructing $m_{\ell\ell jj}$. However, the modelling of this variable is still important since it affects the extrapolation between the Z+jets CR and the SR of both the normalization scale factor and the flavour fit. Consequently, no correction is applied, but a systematic uncertainty is associated to the shape description to account for any residual data/MC disagreement in the CR, as described in Section 5.4.1.1. This uncertainty is used for all Z+jets events, but is uncorrelated between the Z+light-jets and $Z + b/c$-jets samples [`ZMjj`].

In the VBF channel, in which a correction is applied to the $m_{\ell\ell jj}$ distribution (see Section 5.4.1.3), a conservative systematic uncertainty is considered, estimated by removing or doubling this correction [`ZMlljj`]. The uncertainty on the modelling of the m_{jj} distribution by the Z+jets MC simulation is estimated similarly to the resolved-ggF channel, as described in Section 5.4.1.1.

As mentioned in Section 5.4.1.1, the flavour composition of the Z+jets sample in the resolved-ggF channel is determined by the flavour fit. The estimated heavy flavour scale factors have some dependence on the MC model used to unfold them, therefore a truth-level comparison of SHERPA with ALPGEN+PYTHIA was performed showing an uncertainty of 12% on the ratio of $Z + bc$ to $Z + bb$ events [ZbcZbbRatio] and 12% on the ratio of $Z + cc$ to $Z + bb$ [ZccZbbRatio] events.

5.5.2.2 Top quark background

As discussed in Section 5.4.2, the top quark p_{T} distribution is corrected using ATLAS $t\bar{t}$ measurements [116]. A systematic uncertainty on this correction is defined by halving or doubling the applied correction [TopPt].

An uncertainty on the shape of the m_{jj} distribution is derived by comparing the default $t\bar{t}$ NLO simulation with POWHEG+PYTHIA, to other models that probe different sources of modelling uncertainty [TtbarMBBCont]. To investigate the uncertainty on the modelling of the parton showering, ACERMC samples, with either more or less parton showering, are used. POWHEG+HERWIG is used to investigate the effects of a different parton showering and hadronization model. The effects of modelling higher order perturbations are estimated by comparing with the LO MC generator ALPGEN, while another NLO MC generator, aMC@NLO, is used to estimate effects due to different matrix element calculations. The dependence of the cross-sections on the PDF set are investigated with POWHEG+PYTHIA using the HERA PDF.

A similar to the above procedure is used to derive uncertainties on the single-top background. Normalization uncertainties are derived by varying, in cross-section calculations, the renormalization and factorization scales, the value of α_{S} and the PDF eigenvectors. The resulting uncertainty is 7% for the single-top Wt channel, 4% for the t-channel and 4% for the s-channel. Since the Wt channel is by far the dominant component, the 7% uncertainty associated to it is applied to the full single top background [stopNorm]. Moreover, for the dominant Wt channel, shape uncertainties on the m_{jj} and leading jet distributions are derived from comparisons with HERWIG [WtChanPythiaHerwig] and ACERMC [WtChanAcerMC].

5.5.2.3 Diboson background

The uncertainty on the diboson background is determined by evaluating the perturbative QCD uncertainties of the fixed-order NLO calculation, using MCFM [96]. The uncertainties on the cross-section are derived by varying the renormalization and factorization scales [120, 121]; normalization and shape uncertainties are then derived as functions of $p_{\mathrm{T}}^{\ell\ell}$ for the exclusive two-jet [VVJetScalePtST2] and three-jet cross-sections [VVJetScalePtST1]. Variations of the PDF set and the values of α_{S} used in the calculations were found to have no dependence on $p_{\mathrm{T}}^{\ell\ell}$ and are therefore applied as normalization uncertainties [VVJetPDFAlphaPt]. These are 3% for ZZ/WW production and 4% for WZ. Additional shape uncertainties on the m_{jj} distribution are obtained by comparing the LO MC simulation HERWIG with the NLO POWHEG+PYTHIA [VVMbb].

5.5.2.4 Multijet background

As described in Section 5.4.3, the multijet background templates and normalization are determined from data. The normalization factor derived for the untagged event category is also used for the tagged category and a 50% systematic uncertainty is assigned [`MJ`]. A separate (uncorrelated) multijet normalisation parameter is applied in the top quark $e\mu$ control region [`MJ_regiontopemu`] since the multijet background there is determined using a different method (see Section 5.4.3).

5.5.2.5 SM $Zh, h \to bb$ background

A 50% uncertainty is assigned to the $Zh(h \to bb)$ background, as described in Section 5.4.6.

5.5.2.6 Signal

An uncertainty in the experimental acceptance, due to the modelling of Higgs production, is evaluated by varying the parameters of the POWHEG+PYTHIA generator and comparing the resulting samples, after applying the analysis selection at generator level. The following variations are considered:

- Renormalisation (μ_R) and factorisation (μ_F) scales are varied up and down both separately and coherently by a factor of two.
- The amount of initial state radiation (ISR) and final state radiation (FSR) are increased and decreased separately. This is done by changing the Pythia8 tunes used for the simulation of the signal processes. The variations of the ISR and FSR parameters induce a change in the overall energy and particle flow, which affects the underlying event activity.
- The nominal CT10 PDF is replaced by either the MSTW2008nlo68cl or the NNPDF23_nnlo_noLHC_as_0120 PDF.

For the μ_F and μ_R variations no change is observed in signal acceptance, within statistical uncertainties, therefore this variation is neglected. The PDF variations give rise to a small change in acceptance which is independent of the Higgs mass and amounts to a flat 2% in the resolved-ggF and VBF, and 3% in the merged-ggF channel. In the case of ISR/FSR variations, the dominant effect comes from FSR and depends on the Higgs mass. The acceptance change due to FSR increases for low and high values of the hypothesized m_H in the untagged-ggF and VBF channels; this is due to the criterion on p_{T}^{j}, which is not applied in the tagged case.

The ISR and FSR variations are added in quadrature and the overall change in acceptance is approximated by a quadratic function in m_H, which is symmetric about the nominal values. In the resolved-ggF channel, this variation amounts to $\approx 5\%$ at low m_H, decreasing to 1% for intermediate masses and then increasing to about 10(5)% for the untagged (tagged) sample at high m_H. In the VBF channel, it is $\approx$ 10% at low and high m_H, decreasing to 5% at intermediate m_H. This uncertainty is added in quadrature with the flat 2%, estimated from PDF variations, to obtain the overall signal acceptance uncertainty as a function of m_H. In the merged-ggF channel, the effect of ISR and FSR variations is found to be larger than in the resolved-ggF channel. The uncertainty is parametrised as a linear function and amounts to $\approx 30\%$ at $m_H = 800$ GeV, going down to around 10% at $m_H = 1$ TeV; this is added in quadrature to the flat 3% measured from PDF variations.

5.6 Results

This section presents the distribution of the final discriminant in the signal region, for the various analysis categories and for several Higgs boson mass m_H hypotheses, using NWA signal samples with the SM cross-section as a benchmark. The MC simulation has all the corrections listed in Section 5.4 applied, including background normalization and flavour composition.

5.6.1 Resolved ggF

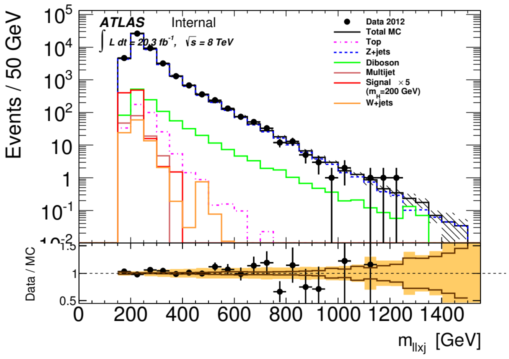
Figures 5.18 and 5.19 show the final $m_{\ell\ell jj}$ discriminant (defined in Section 5.3.1.1) for the untagged and tagged subchannels, after the optimized selection of Equation (5.1). The signal is shown for 200 GeV $\leq m_H \leq$ 1 TeV, in steps of 200 GeV.

5.6.2 VBF

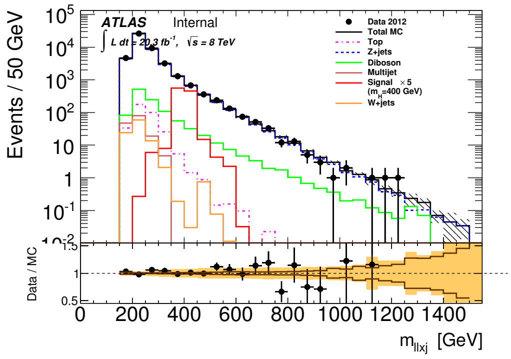
Figure 5.20 shows the final $m_{\ell\ell jj}$ discriminant (as described in Section 5.3.1.3) for the VBF category. The signal is shown for 200 GeV $\leq mH \leq$ 1 TeV, in 200 GeV steps.

5.6.3 Merged ggF

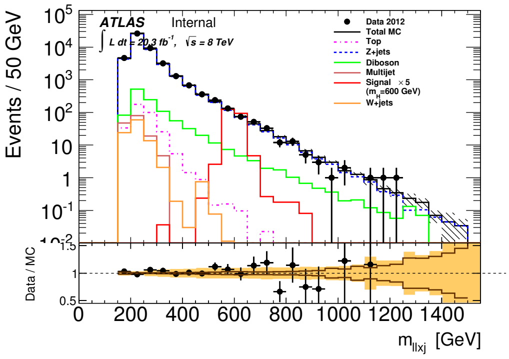
Figure 5.21 shows the final $m_{\ell\ell j}$ discriminant (defined in Section 5.3.1.2) for the merged ggF category. The signal includes the ggF production mode only and is shown for $m_H = 900$ GeV.



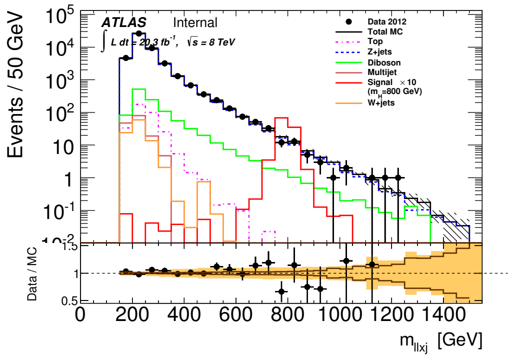
(a) $m_H = 200$ GeV.



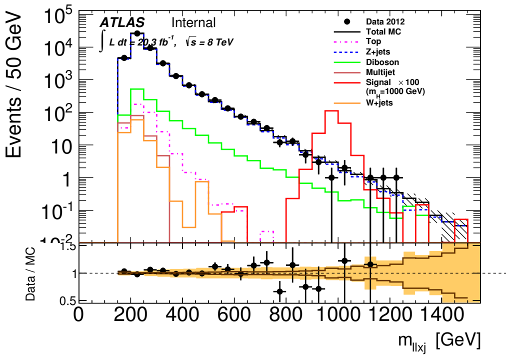
(b) $m_H = 400$ GeV.



(c) $m_H = 600$ GeV.

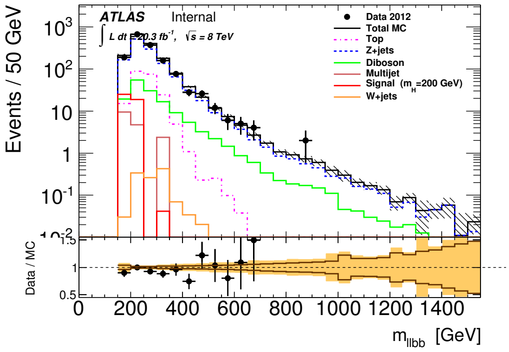


(d) $m_H = 800$ GeV.

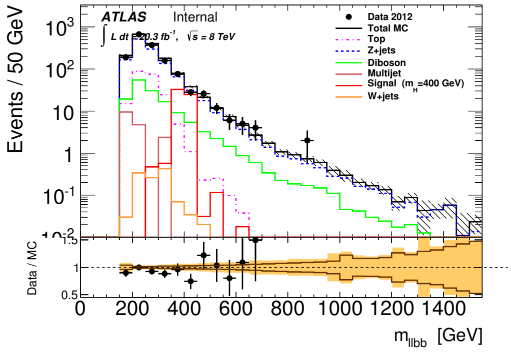


(e) $m_H = 1$ TeV.

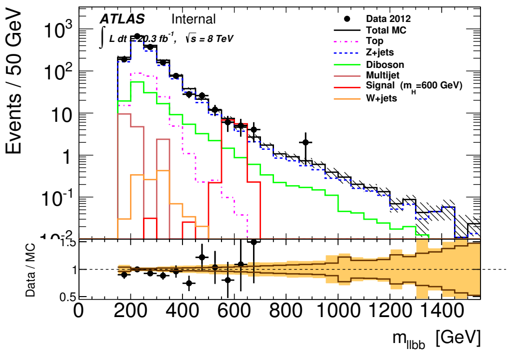
Figure 5.18: $m_{\ell\ell jj}$ distribution for different m_H hypotheses, in the untagged sub-channel of the (resolved) ggF category. The hashed band indicates the systematic uncertainty. Note that the signal is multiplied by various scale factors for clarity.



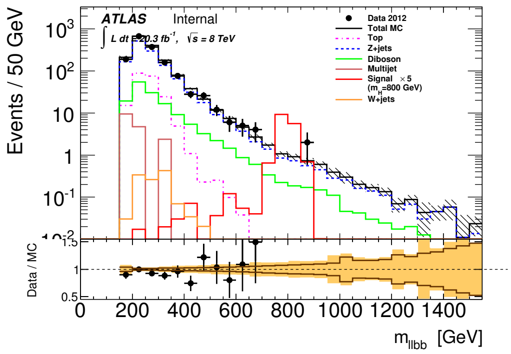
(a) $m_H = 200$ GeV.



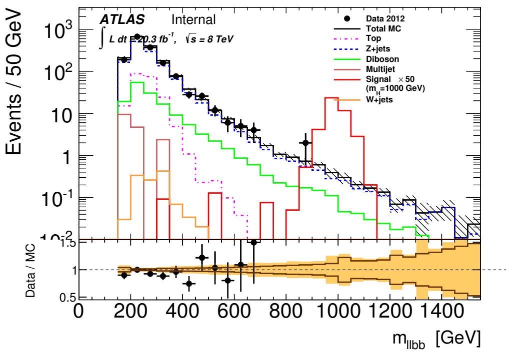
(b) $m_H = 400$ GeV.



(c) $m_H = 600$ GeV.

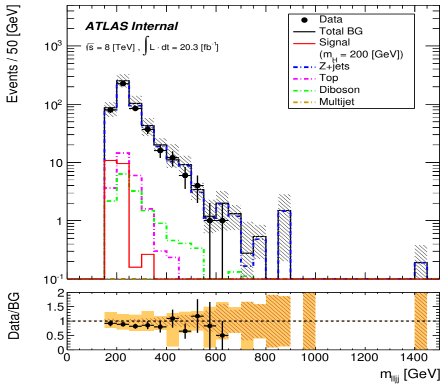


(d) $m_H = 800$ GeV.

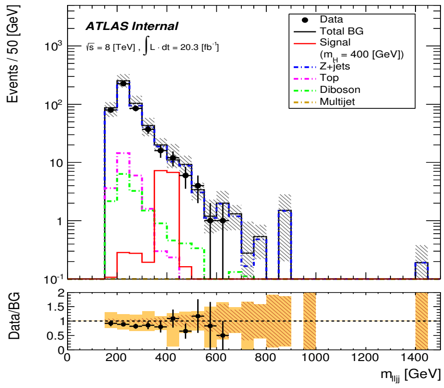


(e) $m_H = 1$ TeV.

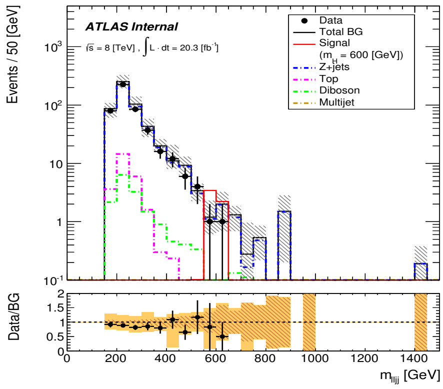
Figure 5.19: $m_{\ell\ell jj}$ distribution for different m_H hypotheses, in the tagged sub-channel of the (resolved) ggF category. The hashed band indicates the systematic uncertainty. Note that the signal is multiplied by various scale factors for clarity.



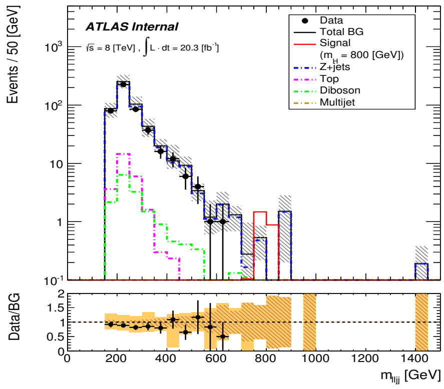
(a) $m_H = 200$ GeV.



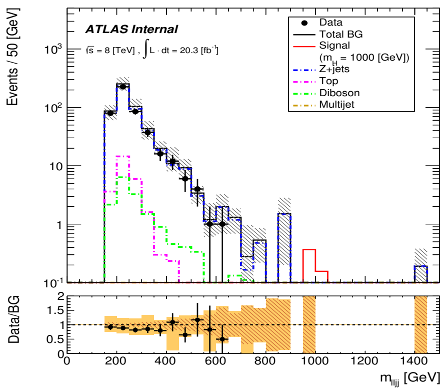
(b) $m_H = 400$ GeV.



(c) $m_H = 600$ GeV.



(d) $m_H = 800$ GeV.



(e) $m_H = 1$ TeV.

Figure 5.20: $m_{\ell\ell jj}$ distribution for different Higgs boson signal hypotheses in the VBF category.

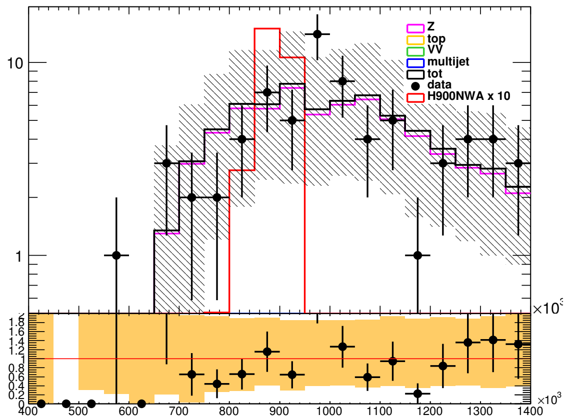


Figure 5.21: $m_{\ell\ell j}$ distribution for the merged ggF category, with $m_H = 900$ GeV. The dashed band shows the systematic uncertainty.

5.7 Limit extraction

As no significant excess is observed, exclusion limits are calculated with a modified frequentist method [89], also known as CL_s, using the $\tilde{q}_\mu$ test statistic in the asymptotic approximation [88]. The observed limits can be compared with expectations by generating "Asimov" data sets, which are representative event samples of the median expectation for an experimental result or its statistical variation in the asymptotic approximation. When producing the Asimov data set for the expected limits, the background-only hypothesis is assumed and the cross-sections for both ggF and VBF production of the heavy Higgs boson are set to zero. The remaining nuisance parameters are set to the value that maximizes the likelihood function for the observed data (profiled). When using the asymptotic procedure to calculate limits it is necessary to generate an Asimov data set both for the background-only hypothesis and for the signal hypothesis. When setting the observed limits, the cross-section for the production mode not under consideration is profiled to the data before generating the background-only Asimov data set.

5.7.1 Exclusion limits on narrow-width Higgs

Figure 5.22 shows the final exclusion limits on $\sigma \times \mathrm{BR}$, at 95% confidence level (CL). The limits are presented separately for the ggF and VBF channels, assuming narrow width approximation for the signal, as discussed in Section 5.1.1. The limit values are also given in Tables 5.8 and 5.9. In the case of the VBF limit, the wavy behaviour of the expected limit is attributed to uncertainties from MC statistics. Drifts of the observed limit into the 2σ band about the expected limit follow the deviations in the input distributions, which are illustrated in Figures 5.18-5.20.

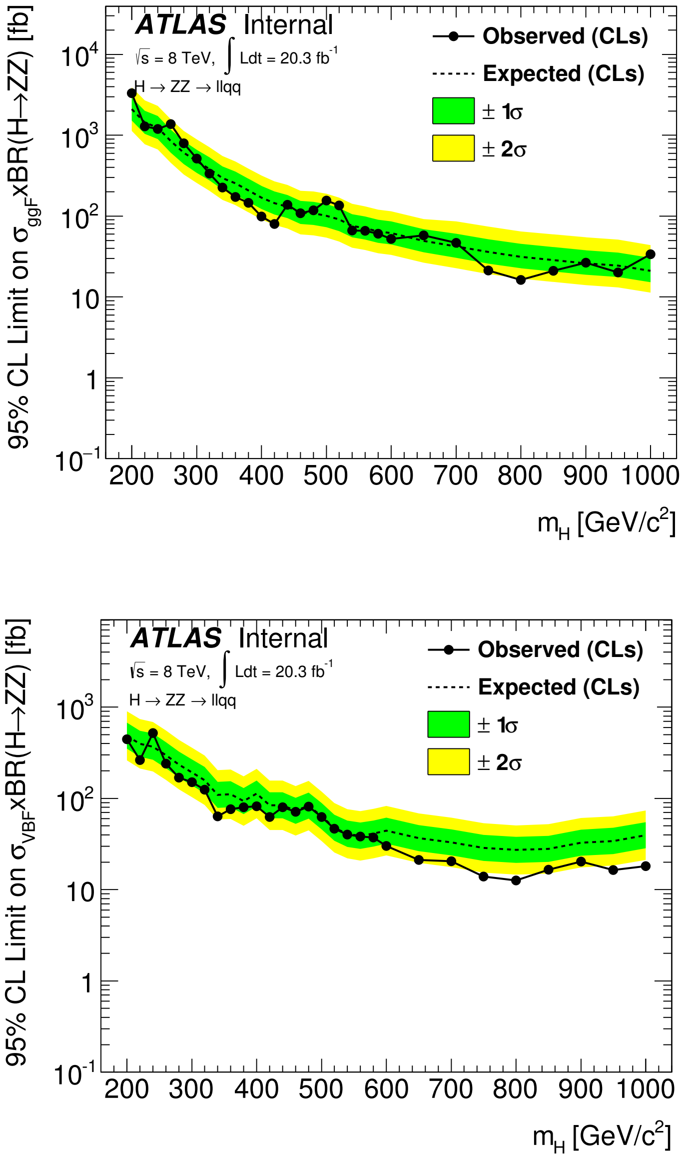


Figure 5.22: 95% CL upper limits on $\sigma \times$ BR as a function of m_H, for ggF (top) and VBF (bottom) production. The solid black line and points indicate the observed limit. The dashed black line indicates the expected limit and the bands the 1-σ and 2-σ uncertainty ranges about the expected limit.

mass	Obs.	-2σ	-1σ	Exp.	$+1\sigma$	$+2\sigma$
200	3329.10	1133.12	1521.21	2111.16	2938.13	3938.77
220	1291.63	775.59	1041.23	1445.04	2011.08	2695.99
240	1200.54	667.78	896.50	1244.18	1731.54	2321.25
260	1371.30	455.57	611.61	848.80	1181.29	1583.60
280	797.99	327.07	439.09	609.38	848.08	1136.91
300	514.11	256.12	343.85	477.20	664.12	890.30
320	335.58	206.01	276.57	383.83	534.19	716.11
340	226.87	159.11	213.61	296.45	412.57	553.07
360	172.57	136.99	183.91	255.23	355.21	476.18
380	146.81	111.49	149.68	207.73	289.09	387.55
400	99.18	91.30	122.56	170.10	236.73	317.35
420	79.95	77.81	104.46	144.97	201.75	270.46
440	138.69	70.16	94.19	130.72	181.93	243.89
460	108.31	59.49	79.87	110.85	154.27	206.80
480	117.91	58.11	78.01	108.26	150.67	201.98
500	155.38	54.27	72.86	101.11	140.72	188.64
520	135.38	48.70	65.37	90.73	126.27	169.27
540	66.65	40.77	54.73	75.95	105.71	141.71
560	65.72	38.07	51.11	70.93	98.72	132.34
580	60.79	34.71	46.60	64.67	90.01	120.66
600	52.38	33.13	44.48	61.73	85.91	115.17
650	57.97	26.56	35.66	49.49	68.88	92.34
700	46.79	22.65	30.41	42.20	60.60	85.84
750	21.18	19.28	25.88	35.92	51.78	74.01
800	16.27	16.84	22.61	31.38	45.30	64.95
850	21.13	15.42	20.70	28.73	41.59	59.76
900	26.66	14.04	18.85	26.16	37.92	54.67
950	20.09	13.17	17.68	24.53	35.53	51.12
1000	33.68	11.32	15.20	21.09	30.53	43.87

Table 5.8: 95% CL upper limits on $\sigma \times$ BR for ggF production. The observed and expected limits, along with $\pm 1\sigma$ and $\pm 2\sigma$ variations, are given in fb.

m_H(GeV)	Obs.	-2σ	-1σ	Exp.	$+1\sigma$	$+2\sigma$
200	446.04	260.30	349.46	484.99	674.96	904.83
220	264.24	212.99	285.94	396.83	552.28	740.37
240	521.00	198.18	266.05	369.23	513.87	688.88
260	241.07	159.51	214.15	297.20	413.61	554.48
280	169.05	125.52	168.51	233.85	325.46	436.30
300	150.91	103.26	138.62	192.38	267.74	358.93
320	124.51	85.12	114.27	158.59	220.71	295.88
340	63.36	58.61	78.68	109.20	151.97	203.73
360	76.14	59.47	79.85	110.81	154.22	206.74
380	80.06	50.26	67.48	93.65	130.33	174.72
400	81.98	60.73	81.53	113.15	157.48	211.11
420	62.78	45.04	60.47	83.93	116.80	156.58
440	80.24	44.99	60.40	83.82	116.65	156.38
460	71.66	39.15	52.55	72.94	101.51	136.08
480	81.27	44.69	60.00	83.27	115.89	155.35
500	62.59	34.23	45.95	63.77	88.76	118.98
520	46.49	25.55	34.31	47.61	66.26	88.83
540	40.13	21.99	29.52	40.96	57.01	76.42
560	38.27	20.91	28.07	38.96	54.22	72.69
580	37.25	22.04	29.60	41.07	57.16	76.63
600	30.25	23.73	31.85	44.21	61.52	82.47
650	21.14	19.65	26.38	36.61	50.95	68.30
700	20.48	17.65	23.69	32.88	45.75	61.34
750	13.90	15.37	20.63	28.63	39.85	53.42
800	12.63	14.60	19.60	27.21	37.86	50.76
850	16.60	15.02	20.16	27.98	38.94	52.20
900	20.33	17.51	23.50	32.62	45.40	60.86
950	16.37	18.35	24.64	34.19	47.58	63.79
1000	18.08	21.23	28.50	39.56	55.05	73.80

Table 5.9: 95% CL upper limits on $\sigma \times$ BR for VBF production. The observed and expected limits, along with $\pm 1\sigma$ and $\pm 2\sigma$ variations, are given in fb.

5.7.2 Exclusion limits on 2HDM

For the 2HDM limits it is necessary to take into account that the natural width of the heavy Higgs boson, as well as the ratio of the ggF to VBF production cross section, vary across the parameter space [27]. The non-zero width is taken into account by smearing each signal histogram to include a natural width up to 5% of the generated heavy Higgs boson mass m_H. For such widths the interference with the ZZ continuum background is negligible. The smearing is performed by looping over each bin of the input histogram and redistributing events according to a relativistic Breit-Wigner, centered at the bin center.

In order to avoid performing the limit fit at each point in the 2HDM parameter space, which is computationally intensive, the following approach is adopted. The limits are first extracted as a function of both the width/m_H and the $\sigma_{VBF}/(\sigma_{ggF}+\sigma_{VBF})$ production ratio in a 2D scan. The width is varied from 0% to 6% in 1% steps and for each width, the $\sigma_{VBF}/(\sigma_{ggF}+\sigma_{VBF})$ is varied from 0 to 1 in 0.1 steps. Once the limits are obtained as a function of the production ratio and the width, they are used to construct a 2D graph which allows to linearly interpolate between points. For each point in the 2HDM plane the predicted $\sigma_{VBF}/(\sigma_{ggF}+\sigma_{VBF})$ and width/m_H are used to look up the limit in the graph.

Figures 5.23 and 5.24 present the final exclusion limit on $\tan\beta$ vs $\cos(\beta-\alpha)$ and m_H, respectively, at 95% CL in the Type-II 2HDM. The results for Type-I are very similar (the various types of 2HDMs have been discussed in Section 1.3).

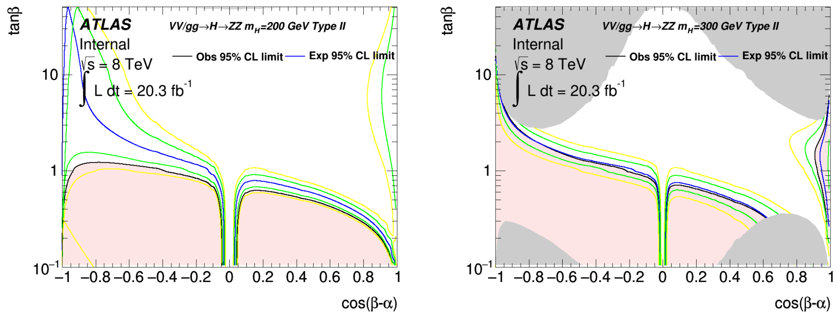


Figure 5.23: 95% CL exclusion contours in the Type-II 2HDM, for (a) $m_H = 200$ GeV and (b) $m_H = 300$ GeV, shown as a function of the parameters $\cos(\beta - \alpha)$ and $\tan\beta$. ggF signal production is included only. The green and yellow lines represent, respectively, the $\pm 1\sigma$ and $\pm 2\sigma$ variations of the expected limit. The grey band masks the region where the limits are not valid since $\Gamma_H/m_H > 5\%$. Results for Type-I 2HDM are very similar.

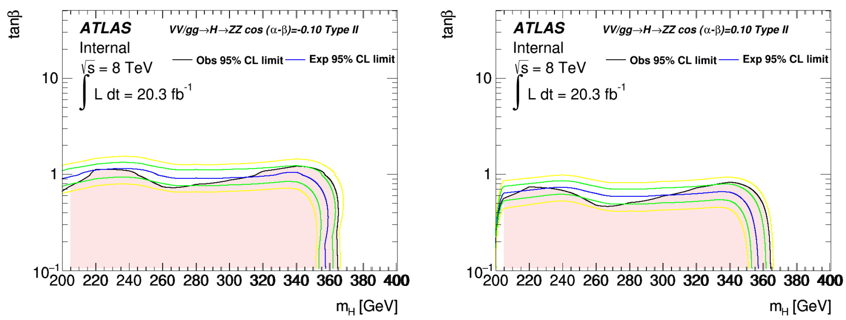


Figure 5.24: 95% CL exclusion contours, in the Type-II 2HDM, for (a) $\cos(\beta-\alpha) = -0.1$ and (b) $\cos(\beta - \alpha) = 0.1$, shown as a function of the heavy Higgs boson mass m_H and the parameter $\tan\beta$. Results for Type-I 2HDM are very similar. The green and yellow lines represent, respectively, the $\pm 1\sigma$ and $\pm 2\sigma$ variations of the expected limit. Only regions where $\Gamma_H/m_H < 5\%$ are shown.

Chapter 6

Statistical treatment

The culmination of the analysis described in Chapter 5 is a combined profile-likelihood in which our knowledge, and lack thereof, is parameterized and tested against the data.

6.1 Likelihood definition

The statistical treatment of the data uses a binned likelihood function, constructed as the product of Poisson probability terms:

$$\prod_{i \in \text{bins}} \text{Pois}\,(n_i|\mu s_i + b_i), \tag{6.1}$$

where the signal strength parameter μ multiplies the expected signal yield s_i in each histogram bin i, b_i represents the background content and n_i is the number of observed events in bin i.

The dependence of the signal and background predictions on systematic uncertainties is described by a set of nuisance parameters (NP) θ which have to be determined from the fit. Auxiliary measurements (such as calibration measurements, control regions etc.) are exploited in order to constrain the nuisance parameters. These measurements are represented by Gaussian constraint terms that multiply the likelihood; the likelihood increases as a nuisance parameter is shifted from the measurement. For normalization uncertainties, log-normal terms are prefered over Gaussian ones in order to maintain a positive likelihood. The fit result is finally obtained in terms of μ and its error σ_μ by maximising the likelihood function with respect to all parameters.

6.2 Fit inputs and variables

A simultaneous profile-likelihood fit is performed across the three analysis channels; resolved-ggF, merged-ggF and VBF. As already mentioned, the resolved-ggF channel discriminates between the untagged and tagged event categories, while the other channels are inclusive in b-tag multiplicity. For each of those cases, the input to the likelihood is the distribution of the reconstructed ZZ invariant mass; $m_{\ell\ell jj}$ for the ggF/VBF channel and $m_{\ell\ell j}$ for the merged channel. Control region distributions are also included for the normalization of the dominant Z+jets and top quark backgrounds, as discussed in Section 5.4. The rest of the backgrounds are taken from MC simulation (or data-driven techniques in the case of the multijet background). The control region distributions used are:

<table>
<tr><th rowspan="3">$N_{b-\text{tag}}$</th><th colspan="7">Channels</th></tr>
<tr><th colspan="3">ggF</th><th colspan="2">Merged</th><th colspan="2">VBF</th></tr>
<tr><th>m_{jj} SR</th><th>m_{jj} CR</th><th>$e\mu$ CR</th><th>m_j SR</th><th>m_j CR</th><th>m_{jj} SR</th><th>m_{jj} CR</th></tr>
<tr><td>0 b-tag</td><td rowspan="2">$m_{\ell\ell jj}$</td><td rowspan="2">MV1c</td><td>—</td><td rowspan="3">$m_{\ell\ell j}$</td><td rowspan="3">$m_{\ell\ell j}$</td><td rowspan="3">$m_{\ell\ell jj}$</td><td rowspan="3">$m_{\ell\ell jj}$</td></tr>
<tr><td>1 b-tag</td><td>—</td></tr>
<tr><td>2 b-tag</td><td>$m_{\ell\ell jj}$</td><td>MV1c</td><td>$m_{\ell\ell jj}$</td></tr>
</table>

Table 6.1: Summary of the regions entering the likelihood fit and the distributions used in each region. Vertically merged rows should be interpreted as regions treated with one distribution (i.e. there is no b-tag separation in merged/VBF channels and 0/1 b-tag regions are combined in the ggF channel). Rows with "—" mean that the region is not included in the fit. "SR" and "CR" stand for signal region and control region respectively.

Z+**jets:**

- **ggF:** The MV1c distribution in the combined m_{jj} sidebands for untagged and tagged events separately. As discussed in Section 5.4.1.1, this distribution is appropriate for the flavour fit.
- **VBF:** The $m_{\ell\ell jj}$ distribution in the combined m_{jj} sidebands (inclusive in b-tag multiplicity).
- **Merged:** The $m_{\ell\ell j}$ distribution in the combined m_j sidebands (inclusive in b-tag multiplicity).

Top: The $m_{\ell\ell jj}$ distribution in the 2 b-tag $e\mu$ control region for the ggF channel.

In total this amounts to 4 signal regions and 6 control regions which are summarized in Table 6.1.

6.3 Nuisance parameters: normalization and systematic uncertainties

Two different types of nuisance parameters are used in order to describe systematic uncertainties; freely-floating parameters and parameters with constraints. A floating normalization parameter is generally associated with the cross-section and acceptance in cases where there is absolute ignorance of the rate and therefore it is completely determined from the data. The fit contains nine normalization parameters which are determined from the signal and control regions:

Signal: Signal strengths for ggF [μ_{ggF}] and VBF [μ_{VBF}] production[1]. In the absence of a specific model the ratio of the two production mechanisms is unknown. For

[1] The signal strengths are defined as scale-factors on the SM cross section which is used as reference. This reference cross section is finally multiplied out in the final results to obtain limits on the cross-section times branching ratio.

this reason, fits for the ggF and VBF production modes are done separately and in each case the other process is "profiled" to the data by allowing it to float freely as an additional nuisance parameter.

Z+jets: The following scale factors are uncorrelated between channels:

Resolved-ggF: Normalization of flavour components Z+light-jet [`ZlNorm`], $Z+c$/light-jet [`ZclNorm`], $Z+b$/light-jet [`ZblNorm`], and Z+heavy-flavour [`ZhfNorm`]. The latter applies to $Z+bb/cc/bc$ events with the ratios $Z+cc/Z+bb$ and $Z+bc/Z+bb$ constrained, as discussed in Section 5.5.2.1.

Merged-ggF: Overall Z+jets production normalization [`ZMergedNorm`]. Since the merged-ggF subchannel selects a very different phase space, the normalization is separate from the case of resolved-ggF.

VBF: Overall Z+jets production normalization [`ZVBFNorm`]. This is separate from the ggF channels since the VBF analysis uses Alpgen rather than Sherpa to model the Z+jets processes.

Top: Overall top quark production normalization [`TopNorm`]. This is correlated across all channels/categories as the top quark background is small in the merged-ggF and VBF channels which are are inclusive in b-tag multiplicity.

The fit contains 72 nuisance parameters from experimental-related uncertainties (see Section 5.5.1) and 21 nuisance parameters from modelling uncertainties (see Section 5.5.2), in addition to the 7 normalization nuisance parameters described above.

6.4 Nuisance parameters: statistical uncertainties

In addition to the systematic uncertainties described above, one must take into account that the background MC samples do not have infinite statistics, hence the histograms are mere estimates of the underlying distributions with some statistical uncertainty. Statistical uncertainties are taken into account in the profile likelihood by using a light version of the Barlow-Beeston method [122]. This method adds extra nuisance parameters to account for the statistical uncertainty on the total MC background, in each bin. The parameters are completely uncorrelated across bins and are considered only for those bins in which the relative statistical uncertainty $> 1\%$ (the particular threshold has been selected by comparison of limit results for various threshold values).

6.4.1 Understanding the fit configuration

This section outlines the various tests preformed in order to verify the robustness of the fitting procedure and understand the results with the constructed model.

6.4.1.1 Nuisance parameter pulls and constraints

Nuisance parameter pulls (i.e. the pull of each parameter off its initial value) and constraints are shown in Figures 6.1 and 6.2 for a fit of the ggF-production mode

(VBF is profiled) and Figures 6.3 and 6.4 for a fit of the VBF mode (ggF is profiled). The hypothesis of a narrow width Higgs boson with $m_H = 400$ GeV is used and fits are performed on data as well as on two "Asimov" data sets, one generated with $\mu = 0$ (i.e. the background-only hypothesis) and the other with $\mu = 1$ (i.e. assumming SM-like signal).

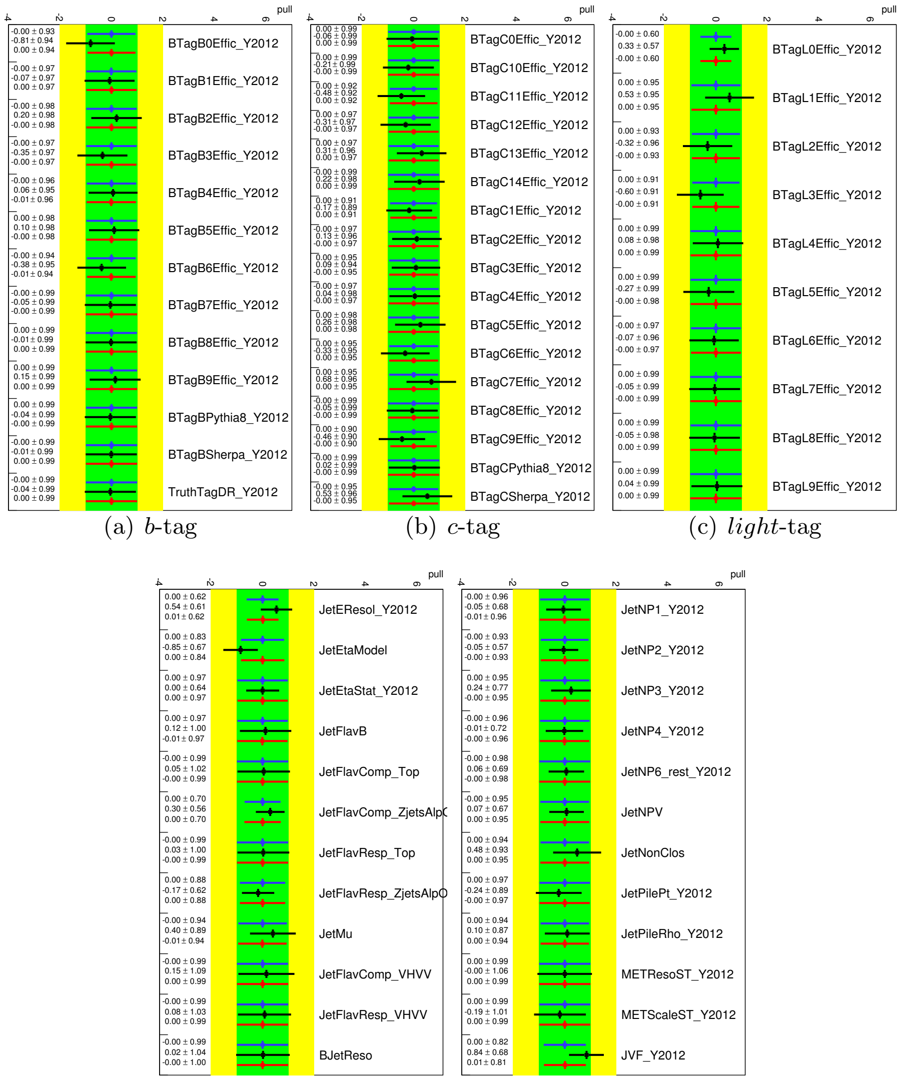
No peculiarities are observed in the results. The fitted normalization scale factors in the ggF fit are shown in Figure 6.2(e). As expected, the fitted values are close to 1 with the exception of the Zbb scale factor which acquires a value of 1.14 ± 0.06. This is consistent with measurements from different analyses [123]. The same plot also shows that the fit does not have enough statistical power to constrain the VBF normalization.

6.4.1.2 Nuisance parameter correlations

Figures 6.5 and 6.6 show an example of nuisance parameter correlations, for $m_H = 400$ GeV, with $\mu = 0$, in the ggF and VBF channel respectively. Only nuisance parameters that exhibit correlation greater than 2.5% with at least one other nuisance parameter are included. The same plot for the ggF channel is shown in Figure 6.7, with the $m_H = 900$ GeV hypothesis, when both the resolved and merged subchannels are considered.

6.5 Nuisance parameter ranking

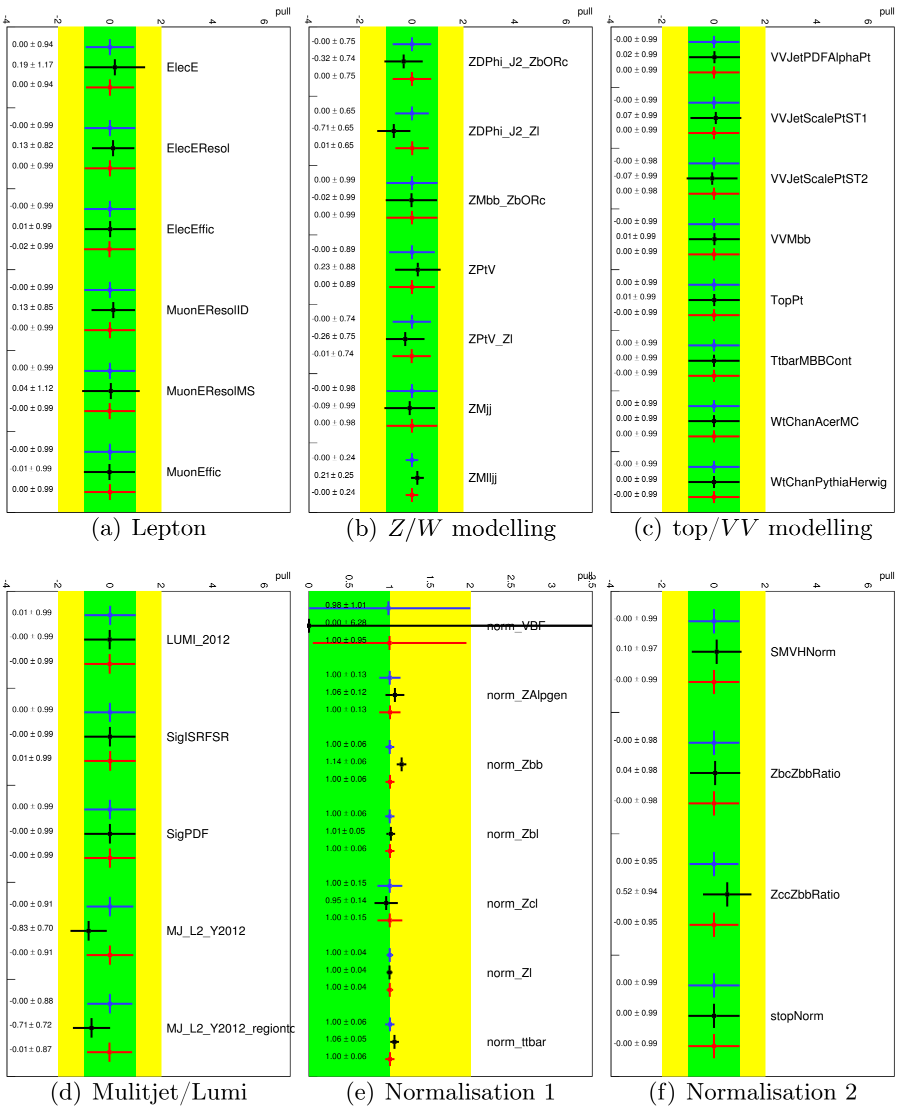
After the maximum log-likelihood value has been found, by fitting the data, each nuisance parameter is pulled at $\pm 1\sigma$ off its best-fit value and the likelihood is maximized again. The change in the fitted value of the signal strength provides the sensitivity to the particular nuisance parameter. Figures 6.8-6.11 show the nuisance parameter ranking for the top 15 ranked nuisance parameters in the ggF and VBF channels, for two signal hypotheses, $m_H = 200$ and 400GeV. Both pre-fit and post-fit impacts are shown, together with pulls from the fit to the data. Pre-fit impacts are estimated by fixing each nuisance parameter at $\pm 1\sigma$ off its initial value. As expected, the dominant systematics are found to be those related to jets, as well as those associated to the modelling of the dominant Z+jets background. Furthermore, it can be seen that the `JetEtaModelling` and `JVF` systematics are asymmetric. This comes from the fact that the input uncertainty is asymmetric. Finally, Figure 6.12 shows the nuisance parameter ranking in the ggF channel when both the merged and the resolved subchannels are considered.



(a) b-tag (b) c-tag (c) *light*-tag

(d) Jets 1 (e) Jets 2

Figure 6.1: The ggF nuisance parameter pulls for the jet and b-tagging related systematics for the Asimov fit with $\mu = 0/1$ (red/blue) and data fit (black). The pulls are for a narrow width Higgs, at $m_H = 400$ GeV.



(a) Lepton (b) Z/W modelling (c) top/VV modelling

(d) Mulitjet/Lumi (e) Normalisation 1 (f) Normalisation 2

Figure 6.2: The ggF nuisance parameter pulls for the lepton, background modelling and normalisation related systematics for the Asimov fit with $\mu = 0/1$ (red/blue) and data fit (black). The pulls are for a narrow width Higgs, at $m_H = 400$ GeV.

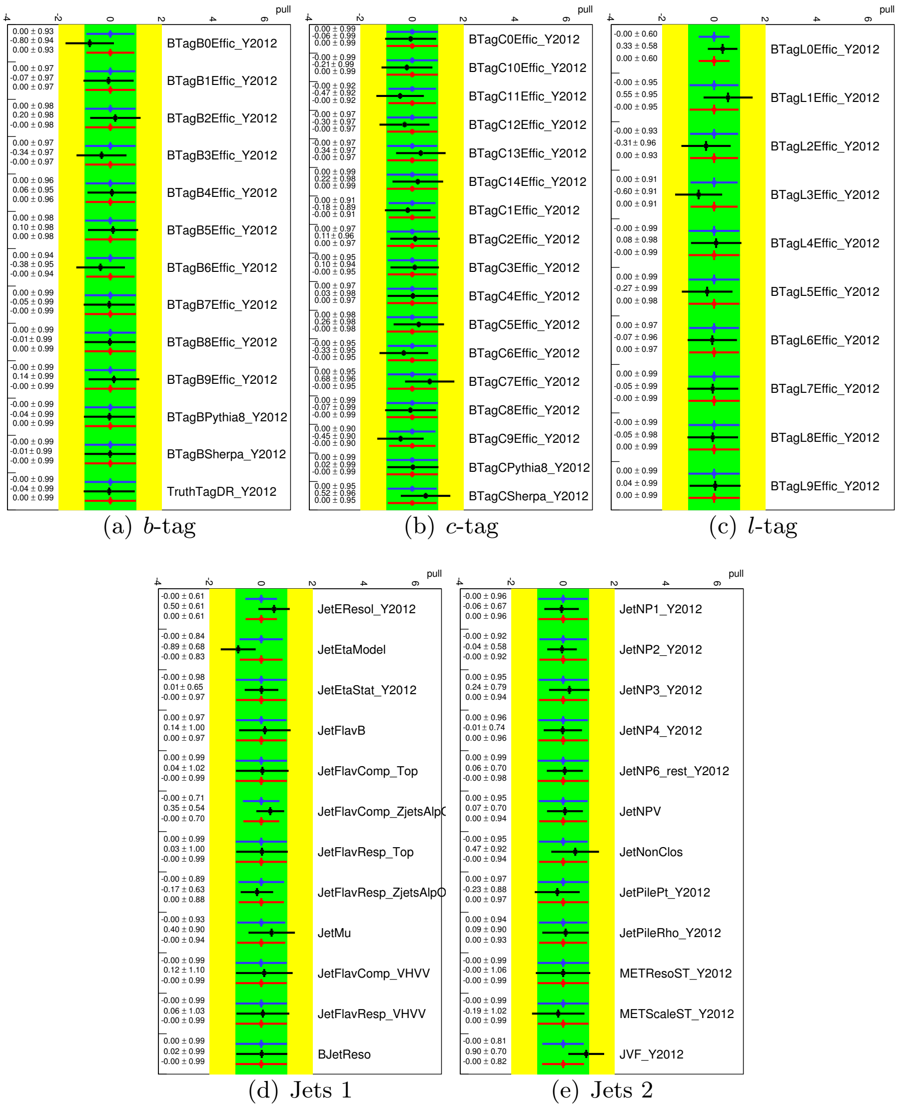
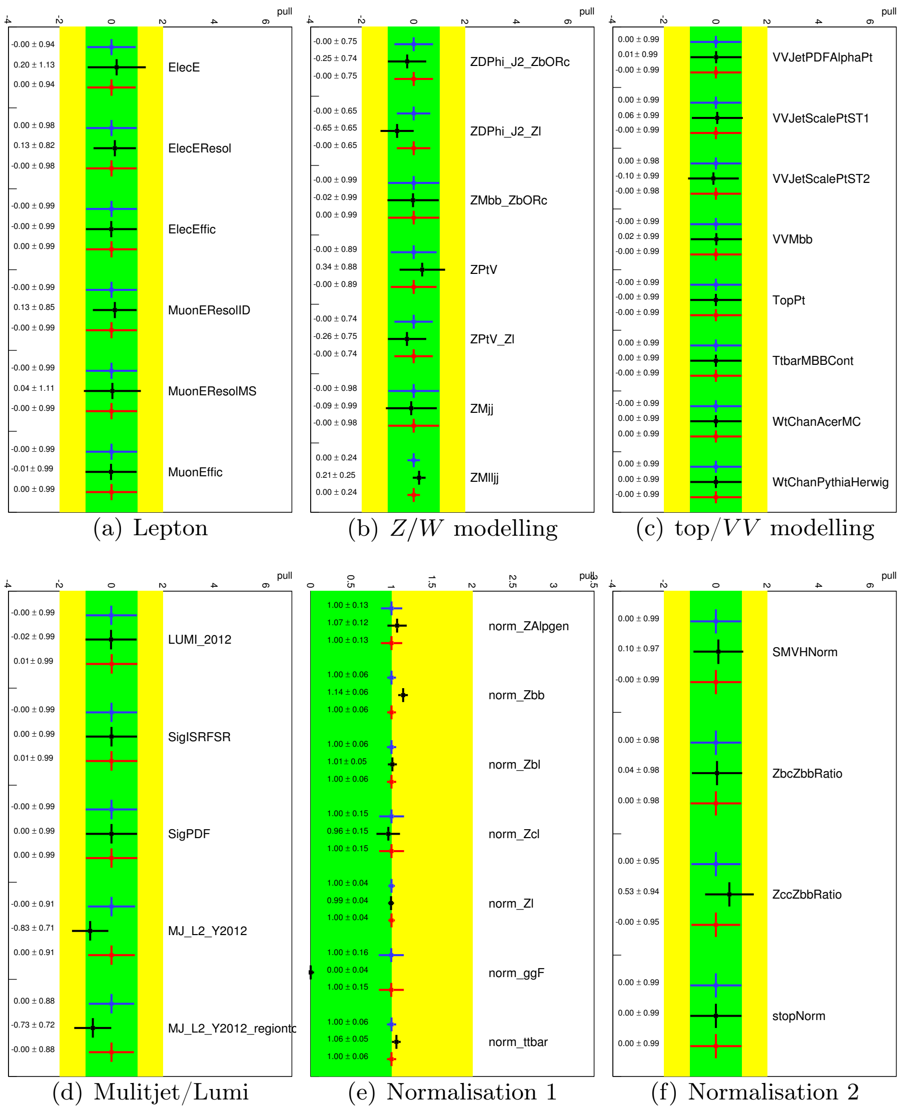


Figure 6.3: The VBF nuisance parameter pulls for the jet and b-tagging related systematics for the Asimov fit with $\mu = 0/1$ (red/blue) and data fit (black). The pulls are for a narrow width Higgs at $m_H = 400$ GeV.



(a) Lepton (b) Z/W modelling (c) top/VV modelling

(d) Mulitjet/Lumi (e) Normalisation 1 (f) Normalisation 2

Figure 6.4: The VBF nuisance parameter pulls for the lepton, background modelling and normalisation related systematics for the Asimov fit with $\mu = 0/1$ (red/blue) and data fit (black). The pulls are for a narrow width Higgs, at $m_H = 400$ GeV.

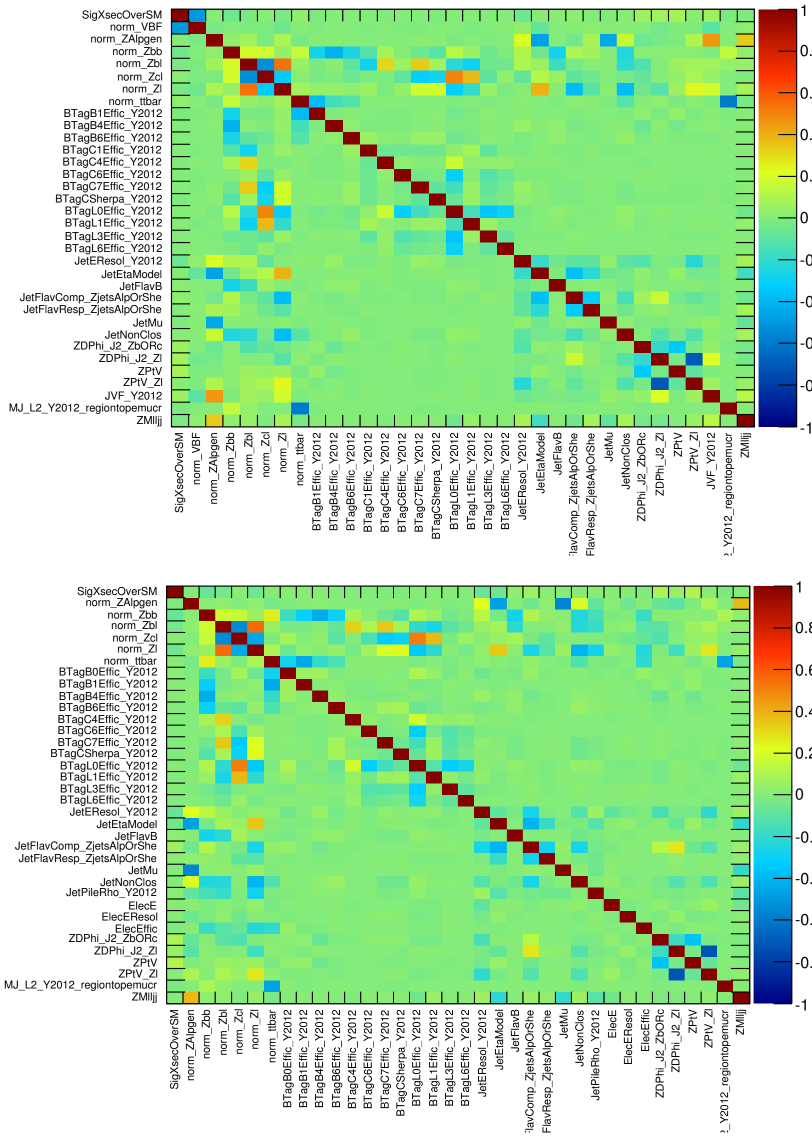
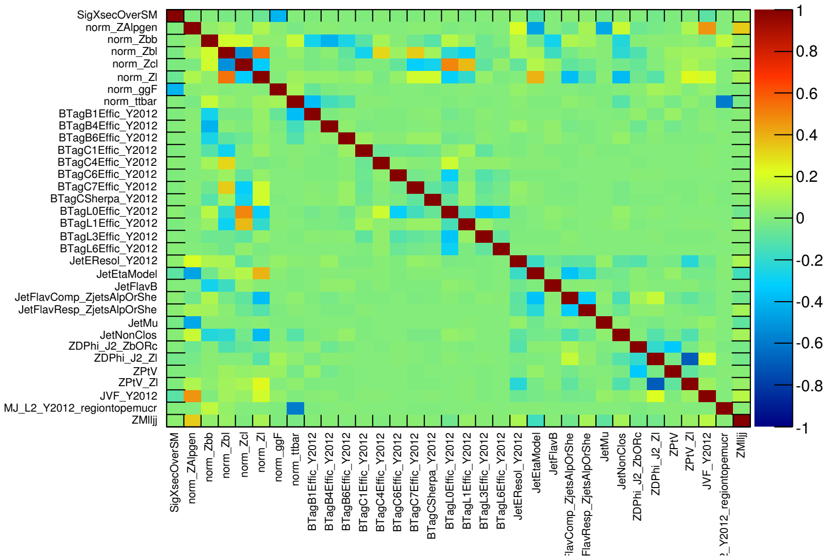


Figure 6.5: Correlation of nuisance parameter which have a correlation >2.5% with any other nuisance parameter for the $\mu = 0$ Asimov fit (top) and data fit (bottom) in the ggF channel.



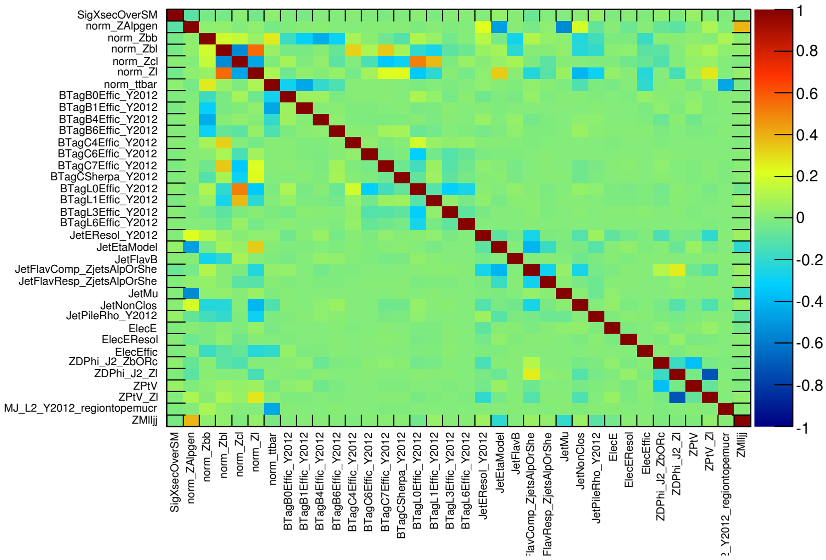
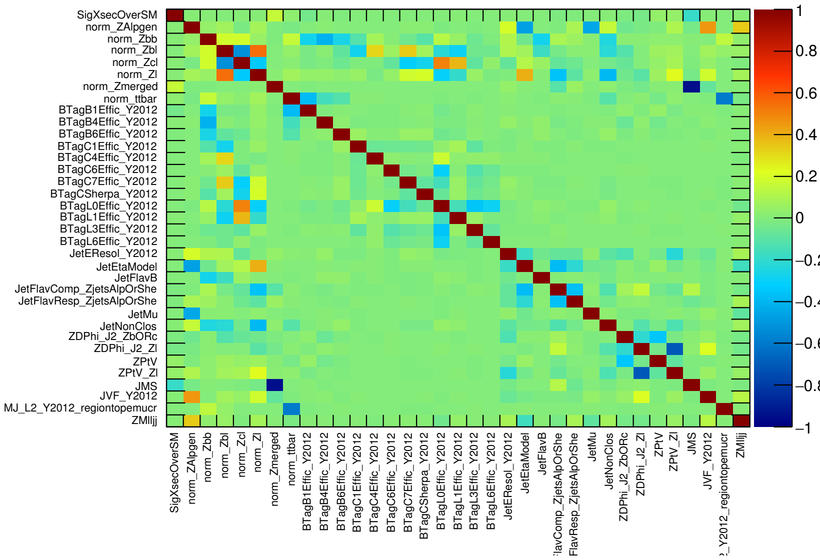


Figure 6.6: Correlation of nuisance parameter which have a correlation >2.5% with any other nuisance parameter for the $\mu = 0$ Asimov fit (top) and data fit (bottom) in the VBF channel.



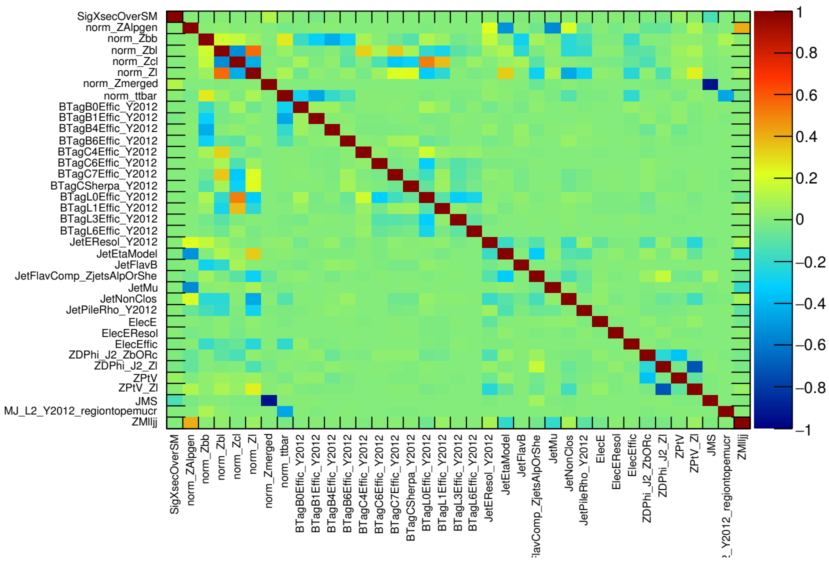


Figure 6.7: Correlation of NPs (that exhibit correlation >2.5% with at least one other NP) for the $\mu = 0$ Asimov fit (top) and data fit (bottom) in the ggF channel when both the merged and the resolved subchannels are considered.

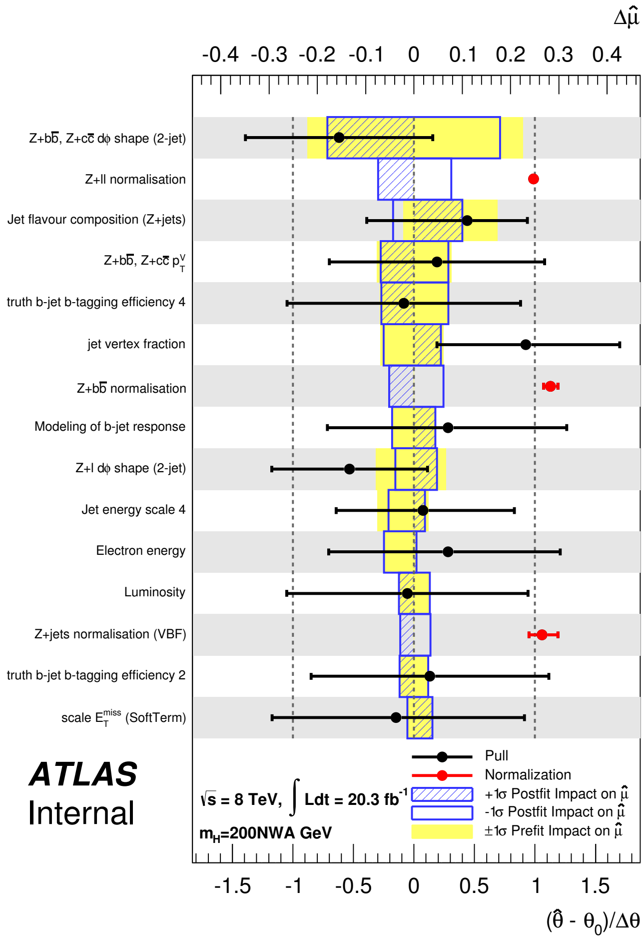


Figure 6.8: Ranking of the top 15 nuisance parameters in the ggF fit for m_H = 200 GeV

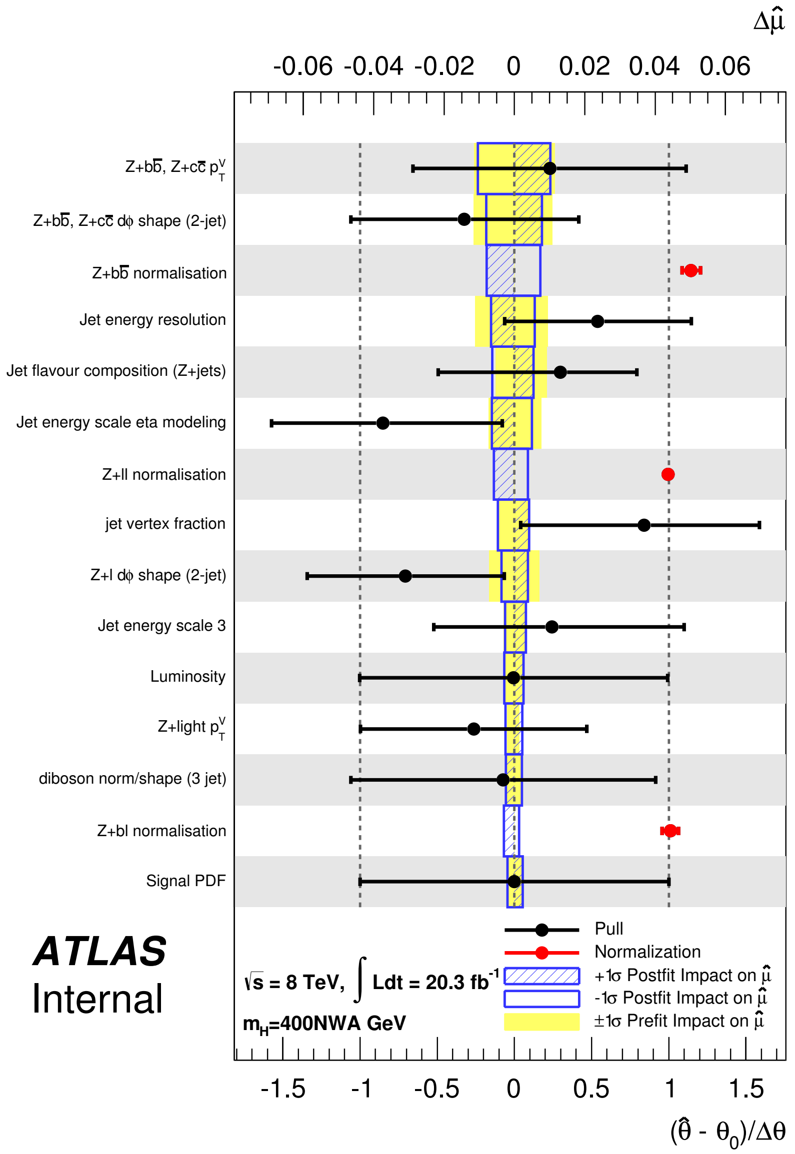


Figure 6.9: Ranking of the top 15 nuisance parameters in the ggF fit for m_H = 400 GeV

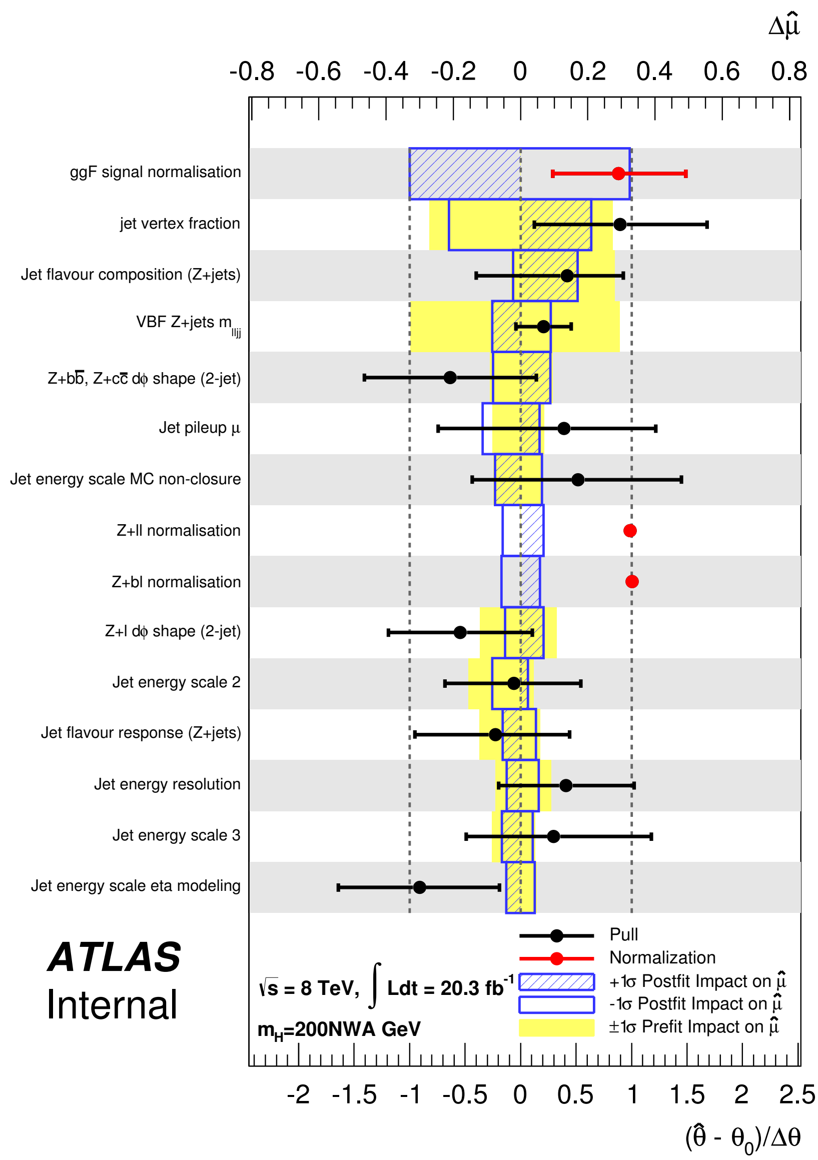


Figure 6.10: Ranking of the top 15 nuisance parameters in the VBF fit for $m_H =$ 200 GeV

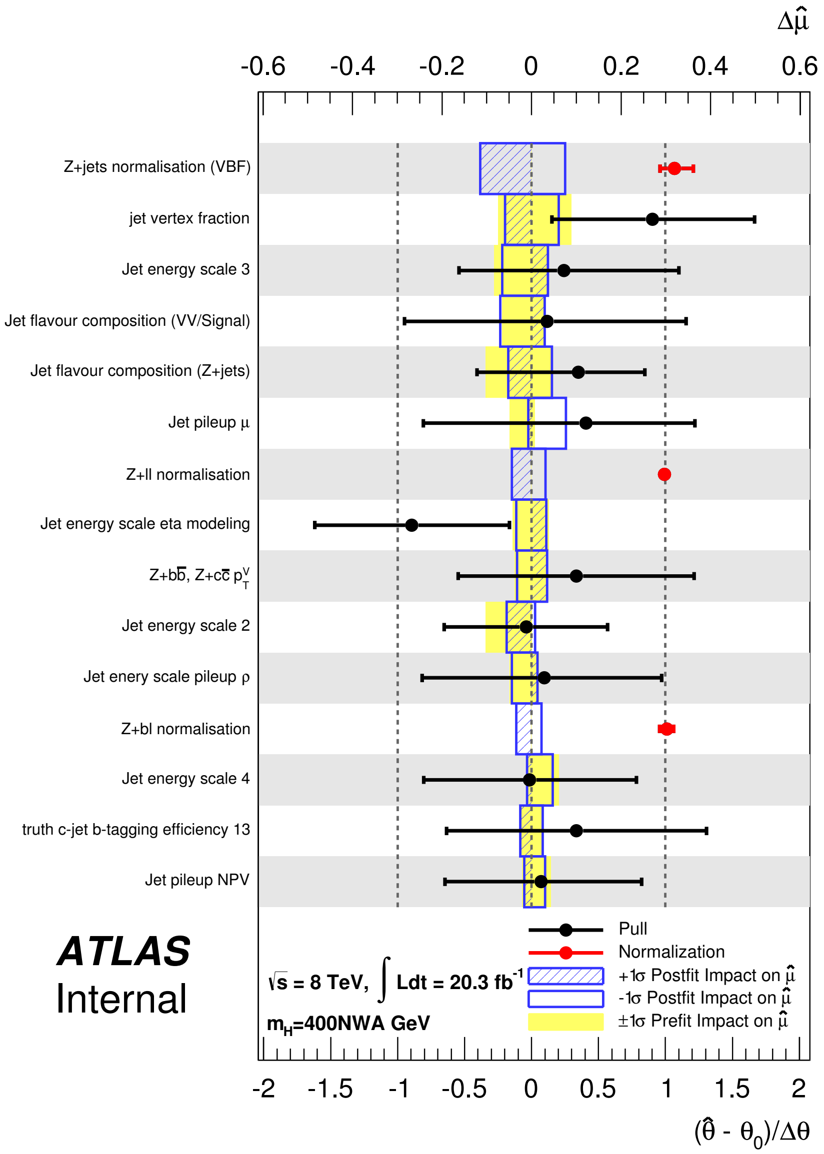


Figure 6.11: Ranking of the top 15 nuisance parameters in the VBF fit for m_H = 400 GeV

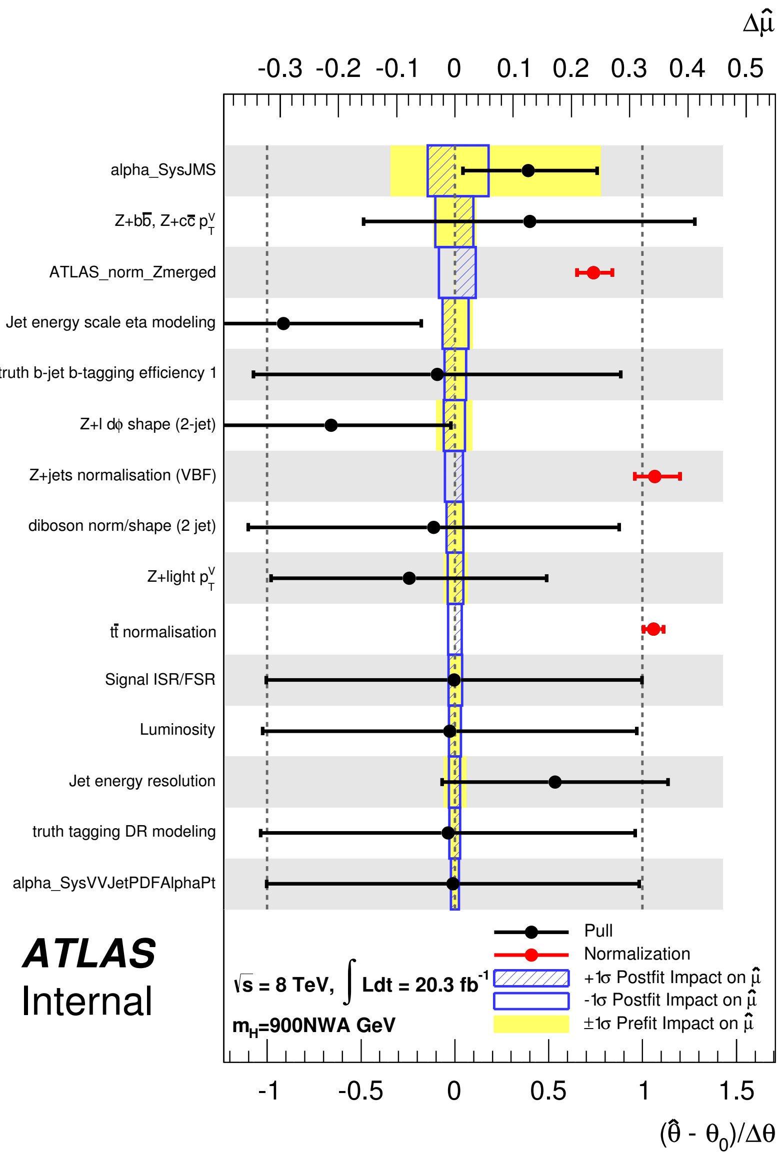


Figure 6.12: Ranking of the top 15 nuisance parameters in the ggF fit for m_H = 900 GeV considering both the resolved and merged-ggF subchannels

Chapter 7

Combined interpretation

In order to provide an answer to the question of whether the recently observed Higgs boson state is the only one or if there exists an extended scalar sector, as described in Section 1.3, several ATLAS groups, studying the $H \to ZZ$ decay mode to different final states, combined their results toward a common statistical interpretation [110]. Specifically, the final states $ZZ \to \ell\ell\ell\ell$, $ZZ \to \ell^+\ell^-\nu\bar{\nu}$, $ZZ \to \ell^+\ell^- q\bar{q}$ and $ZZ \to \nu\bar{\nu}q\bar{q}$, where "$\ell$" stands for either an electron or a muon, were studied in the scenario of a new Higgs boson with narrow width, as well as Type-I and Type-II 2HDMs. These decay modes are referred to, respectively, as $\ell\ell\ell\ell$, $\ell\ell\nu\nu$, $\ell\ell qq$, and $\nu\nu qq$.

It is assumed that additional Higgs bosons would be produced predominantly via the gluon fusion (ggF) and vector-boson fusion (VBF) processes but that the ratio of the two production mechanisms is unknown in the absence of a specific model. For this reason, results are interpreted separately for the ggF and VBF production modes. For Higgs boson masses (m_H) below 200 GeV, associated production (VH, where V stands for either a W or a Z boson) is important as well. In this mass range, only the $\ell\ell\ell\ell$ decay mode is considered.

Due to its excellent mass resolution and high signal-to-background ratio, the $\ell\ell\ell\ell$ decay mode is well-suited for the search of a narrow resonance in the range $140 < m_H < 500$ GeV; hence, the $\ell\ell\ell\ell$ analysis covers the m_H range down to 140 GeV and includes channels sensitive to VH production as well as to VBF and ggF. The $\ell\ell qq$ and $\ell\ell\nu\nu$ searches scan the m_H spectrum down to 200 and 240 GeV respectively and use dedicated ggF and VBF channels. The $\nu\nu qq$ study focuses above 400 GeV and does not distinguish between Higgs production mechanisms. Due to their higher branching ratios, the $\ell\ell qq$, $\ell\ell\nu\nu$ and $\nu\nu qq$ searches dominate at higher masses and determine the overall sensitivity of the combined result. The m_H range for all four searches extends up to 1000 GeV.

For each search channel, a discriminating variable, sensitive to m_H, is identified and used in a likelihood fit. The $\ell\ell\ell\ell$ and $\ell\ell qq$ searches use the invariant mass of the four-fermion system as the final discriminant, while the $\ell\ell\nu\nu$ and $\nu\nu qq$ searches use the transverse mass. Distributions of these discriminants, for each channel, are combined in a simultaneous likelihood fit which estimates the rate of heavy Higgs boson production and, simultaneously, the nuisance parameters corresponding to systematic uncertainties. Additional distributions from background-dominated control regions also enter the fit in order to constrain nuisance parameters.

7.1 $H \to ZZ \to \ell^+\ell^-\ell^+\ell^-$ selection

In the $\ell\ell\ell\ell$ search channel, Higgs candidates are reconstructed from two same-flavour, opposite-charge lepton pairs. Muons must satisfy $p_{\mathrm{T}} > 6$ GeV and $|\eta| < 2.7$, while electrons must satisfy $p_{\mathrm{T}} > 7$ GeV. The three leading in p_{T} leptons must satisfy, in order, $p_{\mathrm{T}} > 20$, 15, and 10 GeV. To ensure well-measured leptons and reduce backgrounds containing electrons from bremsstrahlung, same-flavour leptons must be separated from each other by $\Delta R > 0.1$ and different-flavour leptons by $\Delta R > 0.2$. Jets that are $\Delta R < 0.2$ from electrons are removed from the analysis.

Final states are classified according to the flavour of the leptons as 4μ, $2e2\mu$, $2\mu 2e$ and $4e$. Each event is tested for each of the four categories, in the given order that is according to highest expected signal acceptance. The lepton pair with invariant mass closest to that of a Z boson (the "leading" pair) is first selected; its mass is required to lie in the range $50 < m_{12} < 106$ GeV. The "subleading" pair is the combination of remaining leptons with invariant mass closest to that of a Z boson. For the subleading pair, the mass requirement is $m_{\min} < m_{34} < 115$ GeV, where $m_{\min}$ depends upon the reconstructed 4-lepton mass $m_{\ell\ell\ell\ell}$; starting at 12 GeVfor $m_{\ell\ell\ell\ell} = 140$ GeV, it raises linearly up to 50 GeVfor $m_{\ell\ell\ell\ell} = 190$ GeV and remains constant for higher masses. For 4μ and $4e$ events, if any opposite-charge, same-flavour combination is found with $m_{\ell\ell}$ below 5 GeV, the event is vetoed in order to reject J/ψ decays. To improve the mass resolution, the four-momentum of any reconstructed photon, consistent with having been radiated from one of the leptons in the leading pair, is added to the final state, while the four-momenta of the leading pair leptons are adjusted by means of a Z-mass-constrained kinematic fit [124].

Signal events can be produced via gluon fusion (ggF), vector-boson fusion (VBF) or associated production (VH, where V stands for either a W or a Z boson). Events containing at least two jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$ or $p_{\mathrm{T}} > 30$ GeV and $2.5 < |\eta| < 4.5$, and with the leading two such jets satisfying $m_{jj} > 130$ GeV, are classified as VBF events. Otherwise, if a pair of jets is present, satisfying the same p_{T} and η requirements but with $40 < m_{jj} < 130$ GeV, the event may be classified as VH, provided that it passes a multivariate selection that utilizes jet parameters in order to distinguish VH from ggF events. To account for leptonic decays of V, events failing this selection may still be classified as VH if an additional lepton with $p_{\mathrm{T}} > 8$ GeV is present. All remaining events are classified as ggF.

The dominant background in this channel is continuum ZZ production. Other background components are small and consist of $t\bar{t}$ and $Z + b\bar{b}$ events where, in the latter, muons arise mostly from heavy-flavour semileptonic decays and to a lesser extent, from π/K in-flight decays. The contribution from single-top production is negligible.

7.2 $H \to ZZ \to \ell^+\ell^-\nu\bar{\nu}$ selection

The event selection begins with the reconstruction of a $Z \to e^+e^-$ or $Z \to \mu^+\mu^-$ decay candidate from two same-flavour, opposite-charge leptons. Evidence of neutrino presence in the final state is required by imposing $E_{\mathrm{T}}^{\mathrm{miss}} > 70$ GeV. The selected leptons should satisfy $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ and their invariant mass should be in the range $76 < m_{\ell\ell} < 106$ GeV. Events containing a third lepton or muon with $p_{\mathrm{T}} > 7$ GeV are vetoed. The final discriminating variable is the transverse mass m_{T}^{ZZ}, reconstructed from the momentum of the dilepton system and the missing transverse momentum and defined as:

$$(m_{\mathrm{T}}^{ZZ})^2 \equiv \left(\sqrt{m_Z^2 + |p_{\mathrm{T}}^{\ell\ell}|^2} + \sqrt{m_Z^2 + |E_{\mathrm{T}}^{\mathrm{miss}}|^2} \right)^2 - \left| \vec{p}_{\mathrm{T}}^{\,\ell\ell} + \vec{E}_{\mathrm{T}}^{\mathrm{miss}} \right|^2. \tag{7.1}$$

Two subchannels are distinguished in order to measure event rates separately for the ggF and VBF production mechanisms, taking advantage of the different final state topologies. A VBF event candidate has at least two jets with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 4.5$, satisfying $m_{jj} > 550$ GeV and $\Delta\eta_{jj} > 4.4$. If the event fails the VBF selection, it is classified as ggF candidate if it contains no more than one jet with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.5$, otherwise the event is rejected.

A series of additional requirements are finally applied in order to maximize the signal significance. To suppress the Z+jets background, the azimuthal angle between the dilepton system and the missing transverse momentum $\Delta\phi(p_{\mathrm{T}}^{\ell\ell}, E_{\mathrm{T}}^{\mathrm{miss}})$ must be greater than 2.8 (2.7) for the ggF (VBF) channel and the fractional p_{T} difference, defined as $|p_{\mathrm{T}}^{\mathrm{miss,jet}} - p_{\mathrm{T}}^{\ell\ell}|/p_{\mathrm{T}}^{\ell\ell}$, must be less than 20%, where $p_{\mathrm{T}}^{\mathrm{miss,jet}} = \left|\vec{E}_{\mathrm{T}}^{\mathrm{miss}} + \sum_{\mathrm{jet}} \vec{p}_{\mathrm{T}}{}^{\mathrm{jet}}\right|$. Also, since Z bosons, originating from the decay of a high-mass state, are boosted, the azimuthal angle between the two leptons $\Delta\phi_{\ell\ell}$ is expected to be less than 1.4. Finally, events containing a b-tagged jet with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ are rejected, in order to reduce the top-background and all jets in the event must have an azimuthal angle greater than 0.3 relative to the missing transverse momentum.

The dominant background is ZZ production, followed by WZ production. Other important backgrounds to this search include the WW, $t\bar{t}$ Wt, and $Z \to \tau^+\tau^-$ processes, and also the Z+jets process with poorly reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ but these processes tend to yield final states with low M_T. Backgrounds from W+jets, $t\bar{t}$ single-top quark and multijet processes (with at least one jet misidentified as an electron or muon) are very small.

7.3 $H \to ZZ \to \nu\bar{\nu}q\bar{q}$ selection

Candidate $\nu\nu qq$ events contain neither electrons nor muons but large missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV, indicating the presence of high-p_{T} neutrinos in the final state. The $Z \to q\bar{q}$ decay is then reconstructed using the two leading jets satisfying $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$. One of the jets is required to satisfy $p_{\mathrm{T}} > 45$ GeV, while the dijet invariant mass should be in the range $70 < m_{jj} < 105$ GeV.

The discriminating variable in this study is the transverse mass of the $\nu\nu qq$ system, defined as in Equation (7.1) but with p_{T}^{jj} replacing $p_{\mathrm{T}}^{\ell\ell}$. In order to improve the resolution, the momenta of the selected jets are scaled by a multiplicative factor that sets the dijet invariant mass m_{jj} to the nominal Z boson mass, in the same manner as in the $\ell\ell qq$ search (see Section 5.3.1.1). Finally, the "tagged" (exactly two b-tagged jets) and "untagged" (fewer than two b-tagged jets) subchannels are distinguished, in order to exploit the low production rate of b-jets in the Z+jets and W+jets background processes. Events with more than two b-tags are rejected.

The dominant backgrounds in this search are Z+jets, W+jets and $t\bar{t}$ production. The contribution from multijet background is sufficiently suppressed with additional criteria that utilize a track-based missing transverse momentum $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, defined as the negative vectorial sum of the transverse momenta of all the good-quality inner detector tracks. It is required that the magnitude of $p_{\mathrm{T}}^{\mathrm{miss}}$ satisfies $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV, that the azimuthal angle between $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is $\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}) < \pi/2$ and that the azimuthal angle between $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the nearest jet satisfies $\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, j) > 0.6$.

The sensitivity of this search is further improved with an additional requirement on the jet transverse momenta. As in the $\ell\ell qq$ search (see Section 5.3.1.1), the optimal threshold appears to depend on m_H, however in this case it is not possible to define a single criterion as a function of the reconstructed diboson mass m_{ZZ}, due to the neutrinos in the final state. Instead, different criteria are implemented for the different hypothesized m_H values and the background contribution has to be estimated separately for each one of them. The subleading jet must satisfy $p_{\mathrm{T}}^{j2} > 0.1 \times m_H^{\mathrm{bin}}$ in events with no b-tagged jets and $p_{\mathrm{T}}^{j2} > 0.1 \times m_H^{\mathrm{bin}} - 10$ GeV in events with at least one b-tagged jet, where m_H^{bin} varies from 400 GeV to 1000 GeV in steps of 100 GeV. For a given generated m_H the requirement with the closest m_H^{bin} is used.

7.4 Combination and statistical interpretation

The statistical treatment of the data is similar to that described in Refs. [125, 126], using a simultaneous profile-likelihood-ratio fit (see also Chapter 6). The parameter of interest is the cross-section times branching ratio for heavy Higgs boson production which is treated as correlated between the searches. It is assumed that an additional Higgs boson would be produced predominantly via the ggF and VBF processes which are fitted separately in the absence of a specific model. The VH production mechanism is included in the fit for the $\ell\ell\ell\ell$ search and is assumed to scale with the VBF signal since both the VH and VBF production mechanisms depend on the coupling of the Higgs boson to vector bosons.

The simultaneous fit proceeds as follows. For each channel of each search, there is a distribution of the data with respect to some discriminating variable; these distributions are fitted with a sum of signal and backgrounds. The particular variables used are summarized in Table 7.1. The distributions for the $\ell\ell\ell\ell$ search are unbinned, since the resolution of $m_{\ell\ell\ell\ell}$ is very good, while other searches have binned distributions. For the VBF channels of the $\ell\ell\nu\nu$ search, only the overall event counts are used, rather than distributions, as the sample sizes are very small. The $\ell\ell qq$ and $\nu\nu qq$ searches include additional distributions in control regions in order to constrain the background, using either distributions of the mass variable or of the `MV1c` b-tagging category (see Sections 5.2.3.2 and 5.4.1.1).

A description of the systematic uncertainties contributing to each analysis can be found in Ref. [110]. Detector related uncertainties are common for all of the search channels and are outlined in Section 5.5.1. In the combined fit, systematic uncertainties on the signal acceptance, as well as many of the background theoretical and experimental uncertainties, are treated as fully correlated between the searches. The mass hypothesis for the heavy Higgs boson strongly affects which sources of systematic uncertainty have the greatest effect on the result. At lower masses, the ZZ background theory uncertainties, the Z+jets modelling uncertainties, and the uncertainties on the jet energy scale have the highest impact. At higher masses,

Search	Channel		SR	Z CR	W CR	Top CR
$\ell\ell\ell\ell$	ggF		m_{eeee}, $m_{\mu\mu\mu\mu}$, $m_{ee\mu\mu}$, $m_{\mu\mu ee}$			
	VBF		$m_{\ell\ell\ell\ell}$			
	VH		$m_{\ell\ell\ell\ell}$			
$\ell\ell\nu\nu$	ggF		m_{T}^{ee}, $m_{\mathrm{T}}^{\mu\mu}$			
	VBF		N_{evt}^{ee}, $N_{\mathrm{evt}}^{\mu\mu}$			
$\ell\ell qq$	ggF	untagged	$m_{\ell\ell jj}$	`MV1c`		$m_{\ell\ell jj}$
		tagged	$m_{\ell\ell jj}$	`MV1c`		
		merged-jet	$m_{\ell\ell j}$	$m_{\ell\ell j}$		
	VBF		$m_{\ell\ell jj}$	$m_{\ell\ell jj}$		
$\nu\nu qq$	ggF	untagged	m_{T}		`MV1c` (0 b-tags) `MV1c` (1 b-tag)	
		tagged	m_{T}			

Table 7.1: Summary of the distributions entering the likelihood fit for each channel of each search, both in the signal region (SR) and the various control regions (CR) used to constrain the background. Each entry represents one distribution; some channels have several distributions for different lepton flavours. The distributions are unbinned for the $\ell\ell\ell\ell$ search and binned elsewhere. The VBF channels of the $\ell\ell\nu\nu$ search use only the overall event counts.

uncertainties in the $\ell\ell\nu\nu$ non-ZZ background, the jet mass scale, and the Z+jets background in the merged-jet regime dominate.

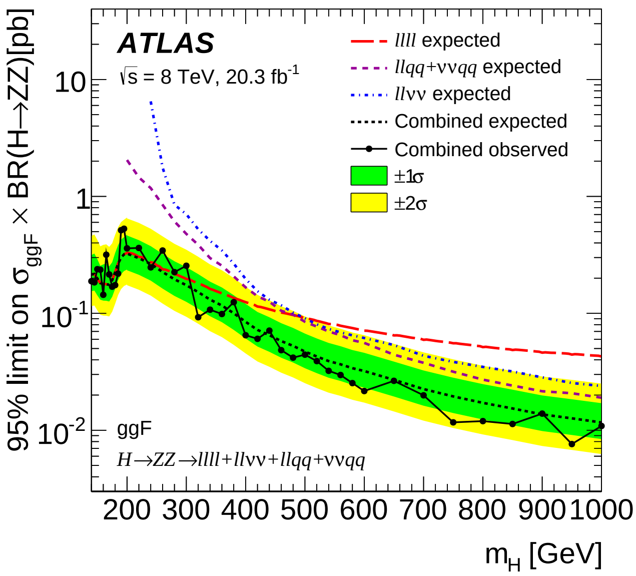
7.5 Limit extraction

As no significant excess is observed, exclusion limits are finally calculated with the CL_s method [89], as described in Section 5.7. Limits on the cross-section times branching ratio, from the combination of all searches, are shown in Figure 7.1. Also shown are expected limits from the $\ell\ell\ell\ell$, $\ell\ell\nu\nu$ and the combined $\ell\ell qq$+$\nu\nu qq$ searches (the latter two searches are only shown in combination as they share control regions). At low mass, the $\ell\ell\ell\ell$ search has the best sensitivity while at higher masses the sensitivity of the combined $\ell\ell qq$+$\nu\nu qq$ search is the greatest, with the sensitivity of the $\ell\ell\nu\nu$ channel only slightly inferior. In the mass range considered for this search, the 95% confidence level (CL) upper limits on the cross-section times branching ratio for heavy Higgs boson production vary between 0.53 pb at $m_H = 195$ GeV and 0.008 pb at $m_H = 950$ GeV in the ggF channel and between 0.31 pb at $m_H = 195$ GeV and 0.009 pb at $m_H = 950$ GeV in the VBF channel. Drifts of the observed limit into the 2σ band about the expected limit originate from local deviations in the input distributions.

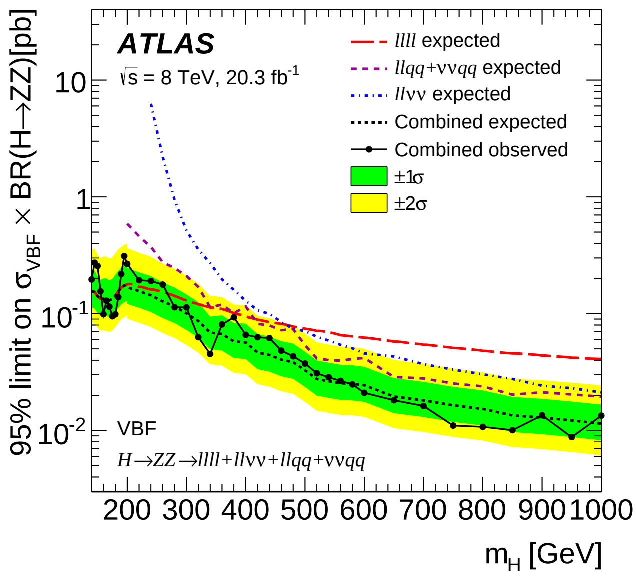
Figure 7.2 shows the exclusion limits in the $\cos(\beta - \alpha)$ versus $\tan\beta$ plane for

Type-I and Type-II 2HDMs, for a heavy Higgs boson with mass $m_H = 200$ GeV[1]. The range of $\cos(\beta - \alpha)$ and $\tan\beta$ explored is limited within the region where the assumption of a narrow-width heavy Higgs boson with negligible interference is valid. When calculating the limits at a given choice of $\cos(\beta - \alpha)$ and $\tan\beta$, the relative rate of ggF and VBF production in the fit is set according to the prediction of the 2HDM for that parameter choice. Figure 7.3 shows exclusion limits as a function of the heavy Higgs boson mass m_H and the parameter $\tan\beta$, for $\cos(\beta - \alpha) = -0.1$. The white areas in the plots indicate regions of the parameter space not excluded by the present analysis; in these regions the cross-section predicted by the 2HDM is below the experimental sensitivity.

[1]This m_H value has been chosen so the assumption of a narrow-width Higgs boson is valid over most of the parameter space (see Ref [27]) and the experimental sensitivity is at a maximum.

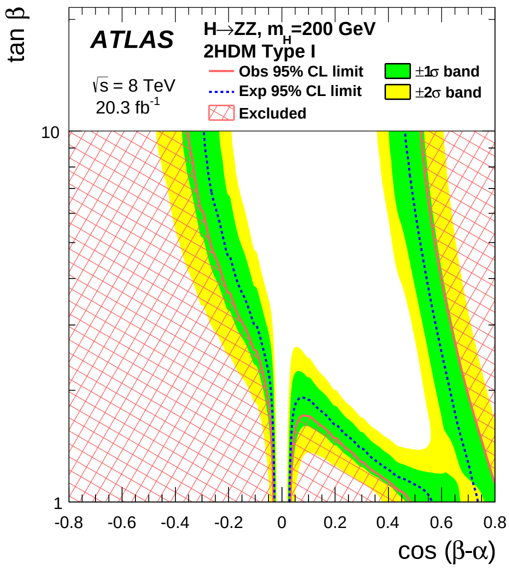


(a) ggF

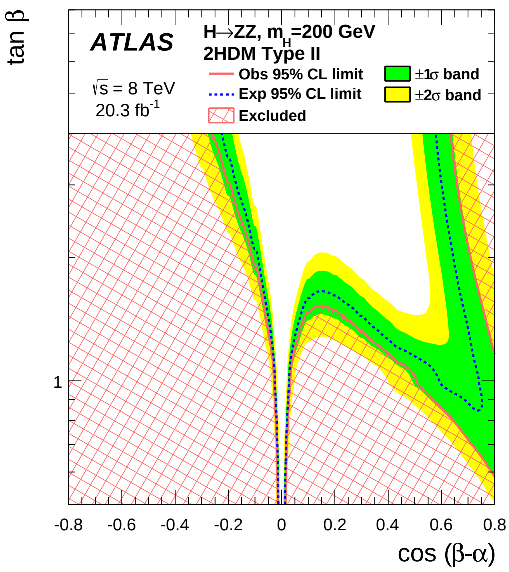


(b) VBF

Figure 7.1: 95% CL upper limits on $\sigma \times \mathrm{BR}(H \to ZZ)$ as a function of m_H, for (a) ggF and (b) VBF production, resulting from the combination of all of the searches. The solid black line and points indicate the observed limit. The dashed black line indicates the expected limit and the bands the 1-σ and 2-σ uncertainty ranges about the expected limit. The dashed coloured lines indicate the expected limits obtained from the individual searches; for the $\ell\ell qq$ and $\nu\nu qq$ searches, only the combination of the two is shown as they share control regions.

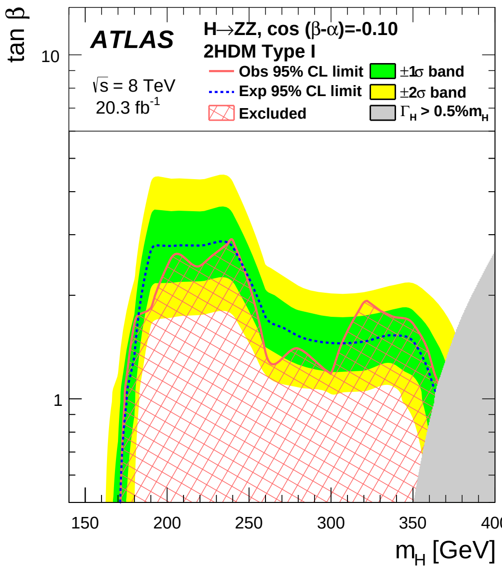


(a) Type-I

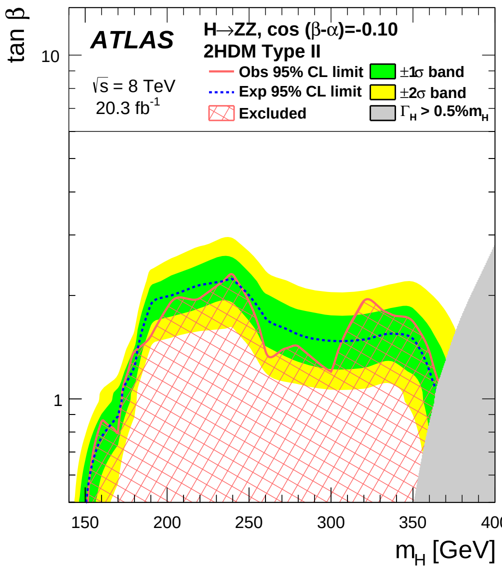


(b) Type-II

Figure 7.2: 95% CL exclusion contours in the 2HDM (a) Type-I and (b) Type-II models for $m_H = 200$ GeV, shown as a function of the parameters $\cos(\beta - \alpha)$ and $\tan\beta$. The red hashed area shows the observed exclusion, with the solid red line denoting the edge of the excluded region. The dashed blue line represents the expected exclusion contour and the shaded bands the 1-σ and 2-σ uncertainties on the expectation. The vertical axis range is set such that regions where the light Higgs couplings are enhanced by more than a factor of three from their SM values are avoided.



(a) Type-I



(b) Type-II

Figure 7.3: 95% CL exclusion contours in the 2HDM (a) Type-I and (b) Type-II models for $\cos(\beta - \alpha) = -0.1$, shown as a function of the heavy Higgs boson mass m_H and the parameter $\tan\beta$. The shaded area shows the observed exclusion, with the black line denoting the edge of the excluded region. The blue line represents the expected exclusion contour and the shaded bands the 1-σ and 2-σ uncertainties on the expectation. The grey area masks regions where the width of the boson is greater than 0.5% of m_H. For the choice of $\cos(\beta - \alpha) = -0.1$ the light Higgs couplings are not altered from their SM values by more than a factor of two.

Bibliography

[1] S. Weinberg, *The Making of the standard model*, Eur. Phys. J. **C34** (2004) 5–13, arXiv:hep-ph/0401010 [hep-ph], [,99(2005)].

[2] C. Quigg, *Gauge theories of the strong, weak and electromagnetic interactions*, Front. Phys. **56** (1983) 1–334.

[3] E. Noether, *Invariant Variation Problems*, Gott. Nachr. **1918** (1918) 235–257, arXiv:physics/0503066 [physics], [Transp. Theory Statist. Phys.1,186(1971)].

[4] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** no. 4, (1961) 579 – 588, http://www.sciencedirect.com/science/article/pii/0029558261904692.

[5] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. (1967) 1264–1266, http://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[6] A. Salam, *Weak and Electromagnetic Interactions*, Nobel Symposium 8, Aspenasgarden, Sweden, July 19-25, 1968. Elementary Particle Theory: Relativistic Groups and Analyticity, 367-377. Almqvist and Wiksell, Stockholm, 1968.

[7] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323, http://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[8] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509, http://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585–587, http://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[10] F. Halzen and A. D. Martin, *Quarks and Leptons: An introductory course in modern particle physics*. Wiley, New York, USA, 1984.

[11] G. 't Hooft and M. J. G. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44** no. 1-3, (1972) 189–213, https://cds.cern.ch/record/419895.

[12] Gargamelle Neutrino Collaboration, F. J. Hasert et al., *Observation of Neutrino Like Interactions without Muon or Electron in the Gargamelle Neutrino Experiment*, Nucl. Phys. **B73** (1974) 1–22.

[13] R. Cashmore, L. Maiani, and J. Revol, *Prestigious Discoveries At CERN*. Springer, 2013.

[14] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s}$ = 540-GeV*, Phys. Lett. **B122** (1983) 103–116.

[15] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c^2 at the CERN SPS Collider*, Phys. Lett. **B126** (1983) 398–410.

[16] UA1 Collaboration, C. Albajar et al., *Studies of Intermediate Vector Boson Production and Decay in UA1 at the CERN Proton - Antiproton Collider*, Z. Phys. **C44** (1989) 15–61.

[17] UA2 Collaboration, M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. **B122** (1983) 476–485.

[18] UA2 Collaboration, P. Bagnaia et al., *Evidence for Z0 —> e+ e- at the CERN anti-p p Collider*, Phys. Lett. **B129** (1983) 130–140.

[19] LEP Electroweak Working Group, DELPHI, OPAL, ALEPH, L3 Collaboration, S. Schael et al., *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP*, Phys. Rept. **532** (2013) 119–244, `arXiv:1302.3415 [hep-ex]`.

[20] Electroweak Working Group, DELPHI, OPAL, ALEPH, L3, SLD Electroweak Group, SLD, SLD Heavy Flavour Group Collaboration, S. Schael et al., *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257–454, `arXiv:hep-ex/0509008 [hep-ex]`.

[21] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[22] ATLAS Collaboration,, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** no. 1, (2012) 1 – 29, `http://www.sciencedirect.com/science/article/pii/S037026931200857X`.

[23] CMS Collaboration,, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** no. 1, (2012) 30 – 61, `http://www.sciencedirect.com/science/article/pii/S0370269312008581`.

[24] ATLAS Collaboration, G. Aad et al., *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. **B726** (2013) 120–144, `arXiv:1307.1432 [hep-ex]`.

[25] ATLAS Collaboration, G. Aad et al., *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, Phys. Lett. **B726** (2013) 88–119, `arXiv:1307.1427 [hep-ex]`, [Erratum: Phys. Lett.B734,406(2014)].

[26] A. Hill and J. van der Bij, *Strongly interacting singlet-doublet Higgs model*, Phys. Rev. D **36** (1987) 3463–3473.

[27] G. Branco, P. Ferreira, L. Lavoura, M. Rebelo, M. Sher, et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rep. **516** (2012) 1–102, arXiv:1106.0034 [hep-ph].

[28] ATLAS Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics*, tech. rep., Geneva, 2009. https://cds.cern.ch/record/1125884.

[29] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, J. Instrum. **3** (2008) S08003. 437 p, https://cds.cern.ch/record/1129811.

[30] ATLAS Collaboration,, *ATLAS detector and physics performance: Technical Design Report, 1*, tech. rep., Geneva, 1999. https://cds.cern.ch/record/391176.

[31] ATLAS Collaboration,, *ATLAS inner detector: Technical Design Report, 2*, tech. rep., Geneva, 1997. https://cds.cern.ch/record/331064.

[32] G. Aad et al., *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007.

[33] M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013, ATLAS-TDR-19, CERN, Geneva, Sep, 2010. https://cds.cern.ch/record/1291633.

[34] ATLAS Collaboration,, *ATLAS liquid argon calorimeter: Technical design report*,.

[35] ATLAS Collaboration,, *ATLAS muon spectrometer: Technical design report*, atlas note, 1997.

[36] S. van der Meer, *Calibration of the effective beam height in the ISR*, Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968. http://cds.cern.ch/record/296752.

[37] ATLAS Collaboration,, *Study of the channel $H \to Z^*Z \to \ell^+\ell^- q\bar{q}$ in the mass range 120-180 GeV with the ATLAS Detector at $\sqrt{s}$=7 TeV*, Tech. Rep. ATLAS-CONF-2012-163, CERN, Geneva, Nov, 2012. https://cds.cern.ch/record/1494191.

[38] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **73** no. arXiv:1302.4393. CERN-PH-EP-2013-026, (2013) 2518. 40 p, https://cds.cern.ch/record/1517411.

[39] ATLAS Collaboration,, *The ATLAS Simulation Infrastructure*, Eur. Phys. J **C70** (2010) 823, arXiv:1005.4568 [physics.ins-det].

[40] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv:hep-ph/0409146.

[41] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[42] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **1006** (2010) 043, arXiv:1002.2581 [hep-ph].

[43] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **04** (2009) 002, arXiv:0812.0578 [hep-ph].

[44] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **02** (2010) 037, arXiv:0911.5299 [hep-ph].

[45] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[46] A. Djouadi, M. Spira, and P. M. Zerwas, *Production of Higgs bosons in proton colliders: QCD corrections*, Phys. Lett. B **264** (1991) 440.

[47] S. Dawson, *Radiative corrections to Higgs boson production*, Nucl. Phys. **B359** (1991) 283.

[48] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas, *Higgs boson production at the LHC*, Nucl. Phys. **B453** (1995) 17, arXiv:hep-ph/9504378.

[49] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, Phys. Rev. Lett. **88** (2002) 201801, arXiv:hep-ph/0201206 [hep-ph].

[50] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, **B646** (2002) 220, arXiv:hep-ph/0207004 [hep-ph].

[51] V. Ravindran, J. Smith, W. L. van Neerven, *NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions*, **B665** (2003) 325, arXiv:hep-ph/0302135 [hep-ph].

[52] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini, *Two loop light fermion contribution to Higgs production and decays*, Phys. Lett. **B595** (2004) 432, arXiv:hep-ph/0404071 [hep-ph].

[53] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders*, Phys. Lett. **B670** (2008) 12, arXiv:0809.1301 [hep-ph].

[54] S. Catani, D. de Florian, M. Grazzini, and P. Nason, *Soft-gluon re-summation for Higgs boson production at hadron colliders*, JHEP **0307** (2003) 028, arXiv:hep-ph/0306211.

[55] C. Anastasiou, R. Boughezal, and F. Petriello, *Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion*, JHEP **0904** (2009) 003, arXiv:0811.3458 [hep-ph].

[56] D. de Florian, M. Grazzini, *Higgs production through gluon fusion: updated cross sections at the Tevatron and the LHC*, Phys. Lett. **B674** (2009) 291, arXiv:0901.2427 [hep-ph].

[57] J. Baglio and A. Djouadi, *Higgs production at the lHC*, JHEP **1103** (2011) 055, arXiv:1012.0530 [hep-ph].

[58] M. L. Ciccolini, A. Denner, and S. Dittmaier Phys. Rev. Lett. **99** (2007) 161803, arXiv:0707.0381 [hep-ph].

[59] M. L. Ciccolini, A. Denner, and S. Dittmaier Phys. Rev. **D77** (2008) 013002, arXiv:0710.4749 [hep-ph].

[60] K. Arnold, M. Bahr, G. Bozzi, F. Campanario, C. Englert, et al., *VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons*, Comput.Phys.Commun. **180** (2009) 1661–1670, arXiv:0811.4559 [hep-ph].

[61] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, *Higgs production via vector-boson fusion at NNLO in QCD*, Phys. Rev. Lett. **105** (2010) 011801, arXiv:arXiv:1003.4451 [hep-ph].

[62] LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier et al., *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*. CERN, Geneva, 2011. https://cds.cern.ch/record/1318996.

[63] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:hep-ph/0206293.

[64] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through O(alpha(s)**2)*, Phys. Rev. **D74** (2006) 114017, arXiv:hep-ph/0609070 [hep-ph].

[65] ATLAS Collaboration, G. Aad et al., *Measurement of the cross-section for b-jets produced in association with a Z boson at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Phys. Lett. **B706** (2012) 295–313, arXiv:1109.1403 [hep-ex].

[66] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, arXiv:0204244.

[67] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C72** (1996) 637–646, arXiv:hep-ph/9601371.

[68] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_s^4)$*, Phys. Rev. Lett. **110** (2013) 252004, arXiv:1303.6524 [hep-ph].

[69] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) 091503, arXiv:1103.2792 [hep-ph].

[70] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028, arXiv:1001.5034 [hep-ph].

[71] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys. Rev. D **82** (2010) 054018, arXiv:1005.4451 [hep-ph].

[72] Particle Data Group Collaboration, C. Amsler et al. Phys. Lett. **B667** (2008) 1, (and 2009 partial update for the 2010 edition).

[73] J. M. Campbell and R. K. Ellis, *An Update on vector boson pair production at hadron colliders*, Phys. Rev. **D60** (1999) 113006, arXiv:hep-ph/9905386 [hep-ph].

[74] S. et al., *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages*, Nucl. Instr. and Meth. **A 572** (2007) 77–99.

[75] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATLAS note ATL-LARG-PUB-2008-002, CERN, 2008.

[76] M. Cacciari, G. P. Salam, and G. Soyez, *Anti-KT jet clustering Algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189.

[77] ATLAS Collaboration,, *Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-053 (2010).

[78] ATLAS Collaboration,, *ATLAS jet energy scale uncertainties using tracks in proton proton collisions at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2011-067 (2011).

[79] ATLAS Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, ATLAS-CONF-2013-083, 2013, http://cdsweb.cern.ch/record/1570994.

[80] ATLAS Collaboration Collaboration,, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011. https://cds.cern.ch/record/1369219.

[81] ATLAS Collaboration Collaboration,, *Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-043, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1435197.

[82] ATLAS Collaboration Collaboration,, *Measurement of the Mistag Rate with 5 fb^{-1} of Data Collected by the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-040, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1435194.

[83] ATLAS Collaboration, G. Aad et al., *Search for a heavy Standard Model Higgs boson in the channel $H \to ZZ \to llqq$ using the ATLAS detector*, Phys. Lett. B **707** (2012) 27–45, arXiv:1108.5064 [hep-ex].

[84] ATLAS Statistics Forum, *ABCD method in searches*, 2012. https://twiki.cern.ch/twiki/pub/AtlasProtected/ATLASStatisticsFAQ/ABCD.pdf.

[85] ATLAS Collaboration,, *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector in 2011*, Tech. Rep. ATLAS-CONF-2011-116, CERN, Geneva, Aug, 2011. https://cds.cern.ch/record/1376384.

[86] ATLAS Collaboration, G. Aad et al., *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector at the LHC*, Eur. Phys. J. C **71** no. arXiv:1101.2185. CERN-PH-EP-2010-069, (2011) 1630. 61 p, https://cds.cern.ch/record/1323586.

[87] LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier et al., *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, Tech. Rep. arXiv:1201.3084. CERN-2012-002, Geneva, 2012. https://cds.cern.ch/record/1416519.

[88] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv:1007.1727 [physics.data-an], [Erratum: Eur. Phys. J.C73,2501(2013)].

[89] A. L. Read, *Presentation of search results: The CL(s) technique*, J.Phys.G **G28** (2002) 2693–2704.

[90] T. Junk, *Confidence Level Computation for Combining Searches with Small Statistics*, Nucl. Instrum. Meth. **A434** (1999) 435, arXiv:hep-ex/9902006, https://plone4.fnal.gov:4430/P0/phystat/packages/0711001.

[91] ROOT Collaboration Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*,.

[92] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv:1009.1003 [physics.data-an].

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

[94] *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*, Tech. Rep. ATL-PHYS-PUB-2011-009, CERN, Geneva, Jul, 2011. https://cds.cern.ch/record/1363300.

[95] *New ATLAS event generator tunes to 2010 data*, Tech. Rep. ATL-PHYS-PUB-2011-008, CERN, Geneva, Apr, 2011. https://cds.cern.ch/record/1345343.

[96] J. Campbell, K. Ellis, and C. Williams, *Vector boson pair production at the LHC*, JHEP **07** (2011) 018, arXiv:1105.0020 [hep-ph].

[97] J. Campbell, K. Ellis, and C. Williams, *Gluon-Gluon Contributions to W^+W^- Production and Higgs Interference Effects*, JHEP **1110** (2011) 005, arXiv:1107.5569 [hep-ph].

[98] R. V. Harlander, S. Liebler, and H. Mantler, *SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM*, Comput. Phys. Commun. **184** (2013) 1605–1617, arXiv:1212.3249 [hep-ph].

[99] R. V. Harlander and W. B. Kilgore, *Higgs boson production in bottom quark fusion at next-to-next-to leading order*, Phys. Rev. **D 68** (2003) 013001, arXiv:hep-ph/0304035 [hep-ph].

[100] R. Bonciani, G. Degrassi, and A. Vicini, *On the Generalized Harmonic Polylogarithms of One Complex Variable*, Comput. Phys. Commun. **182** (2011) 1253–1264, arXiv:1007.1891 [hep-ph].

[101] R. Harlander and P. Kant, *Higgs production and decay: Analytic results at next-to-leading order QCD*, JHEP **0512** (2005) 015, arXiv:hep-ph/0509189 [hep-ph].

[102] D. Eriksson, J. Rathsman, and O. Stal, *2HDMC: Two-Higgs-Doublet Model Calculator Physics and Manual*, Comput. Phys. Commun. **181** (2010) 189–205, arXiv:0902.0851 [hep-ph].

[103] T. Gleisberg et al., *Event generation with Sherpa 1.1*, JHEP **02** (2009) 007, 0811.4622 [hep-ph].

[104] B. Cooper, J. Katzy, M. Mangano, A. Messina, L. Mijovic, et al., *Importance of a consistent choice of alpha(s) in the matching of AlpGen and Pythia*, Eur. Phys. J C **72** (2012) 2078, arXiv:1109.5295 [hep-ph].

[105] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[106] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, Eur. Phys. J. C **74** (2014) 3130, arXiv:1407.3935 [hep-ex].

[107] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014. https://cds.cern.ch/record/1706245.

[108] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv:1407.5063 [hep-ex].

[109] ATLAS Collaboration, *Jet global sequential corrections with the ATLAS detector in proton–proton collisions at $\sqrt{s} = 8$ TeV*, ATLAS-CONF-2015-002, 2015, http://cdsweb.cern.ch/record/2001682.

[110] ATLAS Collaboration, G. Aad et al., *Search for an additional, heavy Higgs boson in the $H \to ZZ$ decay channel at $\sqrt{s} = 8$ TeV in pp collision data with the ATLAS detector*, arXiv:1507.05930 [hep-ex].

[111] ATLAS Collaboration Collaboration,, *b-jet tagging calibration on c-jets containing D^{*+} mesons*, Tech. Rep. ATLAS-CONF-2012-039, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1435193.

[112] *b-tagging in dense environments*, Tech. Rep. ATL-PHYS-PUB-2014-014, CERN, Geneva, Aug, 2014. https://cds.cern.ch/record/1750682.

[113] ATLAS Collaboration Collaboration,, *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector*, J. High Energy Phys. **01** no. arXiv:1409.6212. CERN-PH-EP-2014-214, (2014) 069. 69 p, https://cds.cern.ch/record/1757231.

[114] ATLAS Collaboration, G. Aad et al., *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur. Phys. J. **C74** no. 7, (2014) 2941, arXiv:1404.2240 [hep-ex].

[115] ATLAS Collaboration,, *Search for a heavy Higgs boson in the mass range 200-1000 GeV in the $H \to ZZ \to \ell\ell qq$ channel*, Tech. Rep. ATL-COM-PHYS-2014-269, CERN, Geneva, Apr, 2014. https://cds.cern.ch/record/1693159.

[116] ATLAS Collaboration, G. Aad et al., *Measurements of normalized differential cross sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. **D90** no. 7, (2014) 072004, arXiv:1407.0371 [hep-ex].

[117] ATLAS Collaboration, G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 17, arXiv:1406.0076 [hep-ex].

[118] ATLAS Collaboration, G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C73** no. 3, (2013) 2304, arXiv:1112.6426 [hep-ex].

[119] ATLAS Collaboration, G. Aad et al., *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73** no. 3, (2013) 2306, arXiv:1210.6210 [hep-ex].

[120] I. W. Stewart and F. J. Tackmann, *Theory uncertainties for Higgs mass and other searches using jet bins*, Phys. Rev. D **85** (2012) 034011, http://link.aps.org/doi/10.1103/PhysRevD.85.034011.

[121] S. Gangal and F. J. Tackmann, *Next-to-leading-order uncertainties in Higgs + jets from gluon fusion*, Phys. Rev. D **87** (2013) 093008, http://link.aps.org/doi/10.1103/PhysRevD.87.093008.

[122] J. S. Conway, *Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra*, in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*. arXiv:1103.0354 [physics.data-an].

[123] ATLAS Collaboration, G. Aad et al., *Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector*, JHEP **10** (2014) 141, arXiv:1407.3643 [hep-ex].

[124] ATLAS Collaboration, G. Aad et al., *Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. **D91** no. 1, (2015) 012006, arXiv:1408.5191 [hep-ex].

[125] ATLAS Collaboration, G. Aad et al., *Combined search for the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. **D86** (2012) 032003, arXiv:1207.0319 [hep-ex].

[126] ATLAS Collaboration,, *Procedure for the LHC Higgs boson search combination in summer 2011*, Tech. Rep. ATL-PHYS-PUB-2011-011, CERN, Geneva, Aug, 2011. https://cds.cern.ch/record/1375842.

[127] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81–86.

Appendices

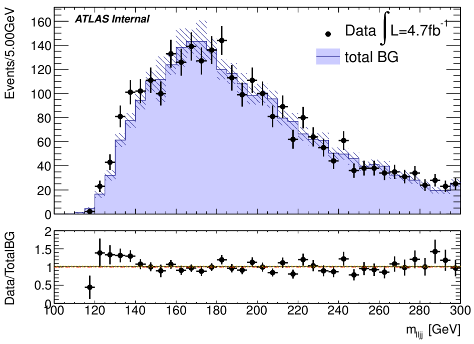
Appendix A

Study of the DY+jets control region

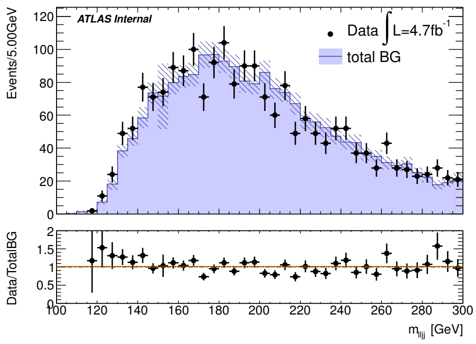
As discussed in Section 4.4.1, both the overall normalization and the shape of the DY+jets background is estimated using control regions in the side bands of the m_{jj} distribution. The control region, defined as the union of $40 < m_{jj} < 60$ GeV and $115 < m_{jj} < 160$ GeV is considered as the nominal one. In order to estimate possible systematic uncertainties induced be the side band description, the behavior of the low and high side bands is studied separately and furthermore the window defining the control region are modified. Examples of these variations are shown in Figures A.1 and A.2 for the electron and muon channels of untagged case respectively and in Figure A.3 for the tagged case (electron and muon channel together). The observed differences in the corresponding DY +jets scale factors are presented in Table A.1 and are used in the estimation of the systenmatic uncertainty.

Category	Side band region		
	$30 < m_{jj} < 50$, $125 < m_{jj} < 170$ GeV	$40 < m_{jj} < 60$ GeV	$115 < m_{jj} < 160$ GeV
Untagged electron:	0.99 ± 0.01	1.02 ± 0.03	0.99 ± 0.02
Untagged muon:	1.01 ± 0.01	0.97 ± 0.02	1.04 ± 0.02
Tagged:	0.94 ± 0.07	0.98 ± 0.21	1.00 ± 0.09

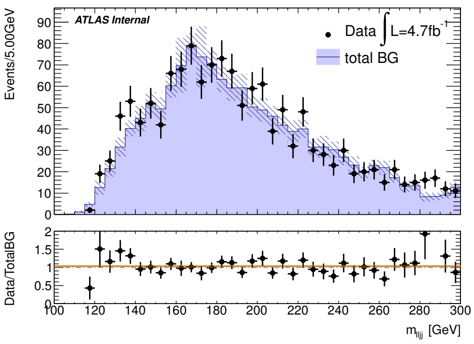
Table A.1: $DY+$jets scale factors obtained from modified SBs w.r.t the nominal one ($40 < Mjj < 60$, $115 < Mjj < 160$ GeV).



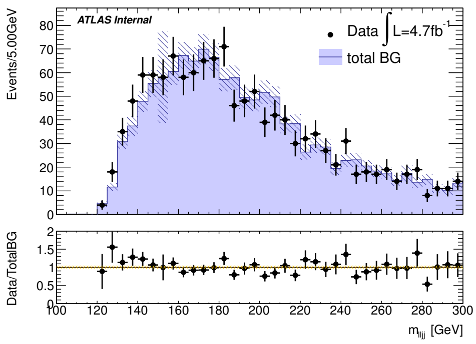
(a) $40 < m_{jj} < 60$ GeV OR $115 < m_{jj} < 160$ GeV



(b) $30 < m_{jj} < 50$ GeV OR $125 < m_{jj} < 170$ GeV

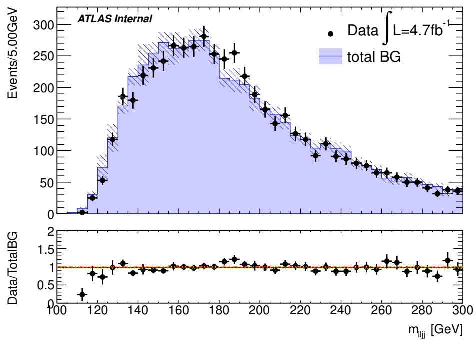


(c) $40 < m_{jj} < 60$ GeV

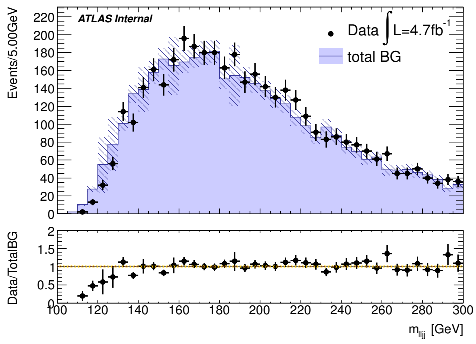


(d) $115 < m_{jj} < 160$ GeV

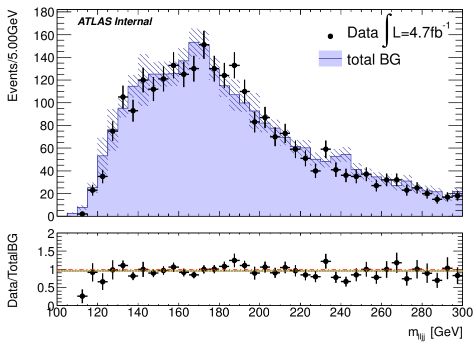
Figure A.1: m_{lljj} distribution in the electron untagged category for various definitions of the m_{jj} SBs, which are used to control the DY+jets background. The bands show systematic uncertainties from MC statistics.



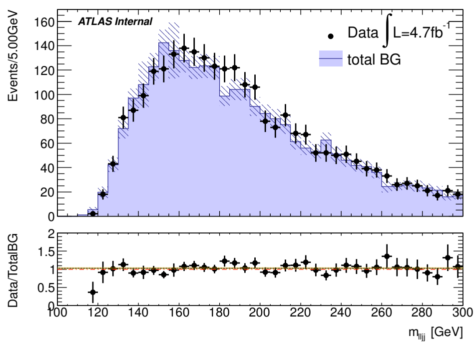
(a) $40 < m_{jj} < 60$ GeV OR $115 < m_{jj} < 160$ GeV



(b) $30 < m_{jj} < 50$ GeV OR $125 < m_{jj} < 170$ GeV

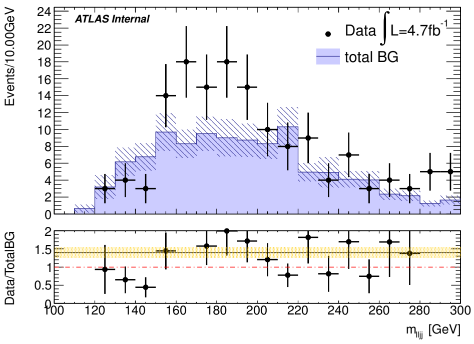


(c) $40 < m_{jj} < 60$ GeV

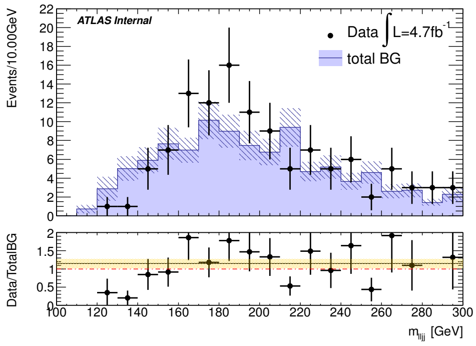


(d) $115 < m_{jj} < 160$ GeV

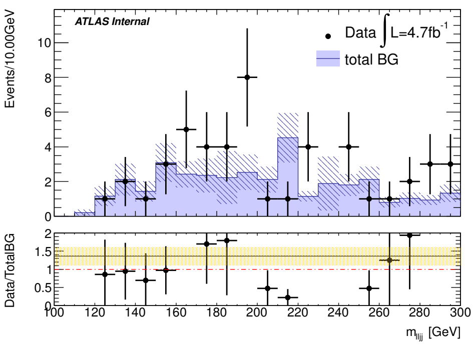
Figure A.2: m_{lljj} distribution in the muon untagged category for various definitions of the m_{jj} SBs, which are used to control the DY+jets background. The bands show systematic uncertainties from MC statistics.



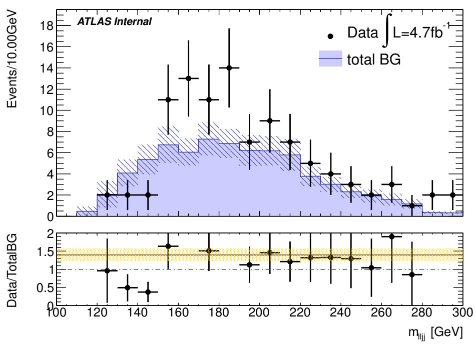
(a) $40 < m_{jj} < 60$ GeV OR $115 < m_{jj} < 160$ GeV



(b) $30 < m_{jj} < 50$ GeV OR $125 < m_{jj} < 170$ GeV



(c) $40 < m_{jj} < 60$ GeV



(d) $115 < m_{jj} < 160$ GeV

Figure A.3: m_{lljj} distribution in the tagged category for various definitions of the m_{jj} SBs, which are used to control the DY+jets background. The bands show systematic uncertainties from MC statistics.

Appendix B

Multijet background (ABCD method)

B.1 Shape studies

B.1.1 Multijet background shape in the muon channel

In the search for the SM Higgs boson (Section 4.4.4), the candidate control regions (CRs) for the acquisition of the multijet background shape in the case of muon events are defined by:

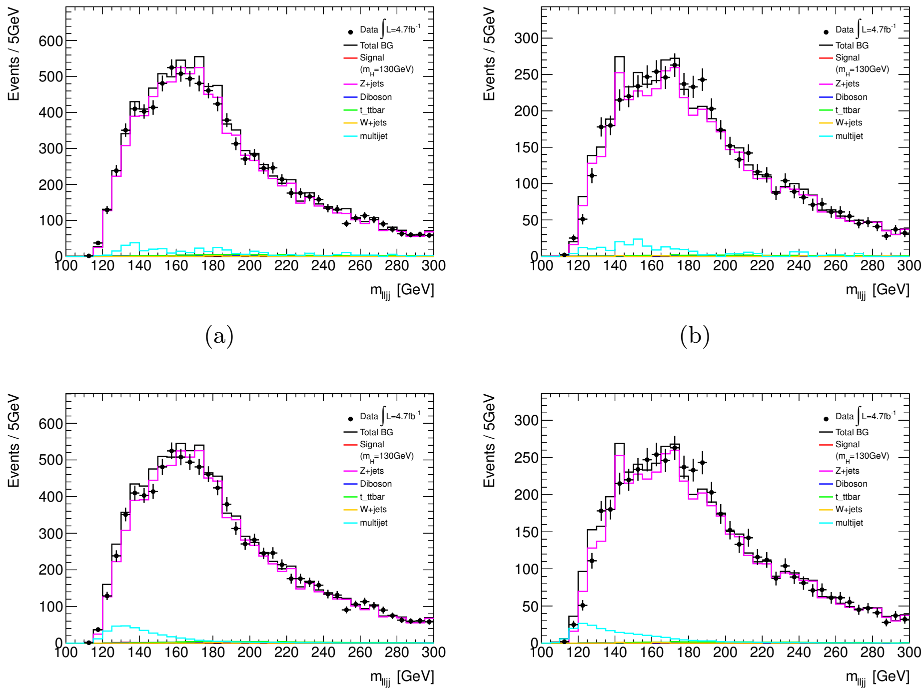
- events with a pair of same charge, isolated muons,
- events with muons of opposite charge, one of which should fail the track-isolation requirement.

In the second definition, anti-isolation is conservatively required from only one of the two muons in order to keep the CR as close to the original selection (i.e. of opposite charge, isolated muons) as possible. Figure B.1 shows the distribution of $m_{\mu\mu jj}$, using the two CRs for the description of the multijet background.

The evaluation of the provided shapes is carried out using Kolmogorov-Smirnov (K-S) tests [127] to measure the level of agreement between the total background and the data. The tests are performed both in the m_{jj} side bands (SBs) as well as the chargeal region (SR) and include comparisons using different variables. The results are summarized in Table B.1. Apparently, the best description is provided by the sample of same charge, isolated muons. The third column demonstrates the results using a modified anti-isolation requirement; both muons are required to exhibit relative track-isolation between 20% and 100%. The description is still worse than the one provided by the sample of isolated, same charge muons.

Distribution and selection	Kolmogorov Probability		
	same charge isolated	opposite charge non-isolated	opposite charge modified non-isolated
m_{jj} dilepton +2 jets	0.17	0.01	0.05
$m_{\ell\ell jj}$ final untagged SR	0.96	0.22	0.26
$m_{\ell\ell jj}$ final untagged SB	0.23	0.02	0.04
$m_{\ell\ell jj}$ final tagged SR	0.92	0.55	0.56
$m_{\ell\ell jj}$ final tagged SB	0.99	0.71	0.83

Table B.1: Comparisons between data and total background, in the muon channel, using different CRs to obtain the multijet background shape. Results are shown for different variables.



(a) (b) (c) (d)

Figure B.1: $m_{\mu\mu jj}$ distributions in the SR (left) and the m_{jj} SBs (right). The shape of the multijet background is obtained from the sample of same charge, isolated muons (top) or opposite charge muons with reversed track-isolation requirement for one of them (bottom).

B.1.2 Multijet background shape in the electron channel

In the case of electrons, three CRs are tested as candidates for the acquisition of the multijet background shape:

- events with same charge electrons,
- opposite charge electrons, one of which should fail the track-isolation requirement,
- opposite charge, isolated electrons, one of which passes the *loose* but fails the *medium* identification quality criteria.

Figure B.2 shows the distribution of m_{eejj}, using different shapes to describe the multijet background. The K-S probabilities derived using for the above CR definitions, are summarized in Table B.2. The conclusion is not as straightforward as in the case of muons; a different shape appears to be optimal for each case. Among them, the CR defined by anti-isolation of one of the electrons is slightly favoured since it exhibits very good agreement in the SR of the untagged event category.

Distribution and selection	Kolmogorov Probability same charge isolated, tight	opposite charge non-isolated, tight	opposite charge isolated, loose non medium
m_{jj} dilepton +2 jets	0.04	0.04	0.01
$m_{\ell\ell jj}$ final untagged SR	0.75	0.96	0.12
$m_{\ell\ell jj}$ final untagged SB	0.92	0.35	0.06
$m_{\ell\ell jj}$ final tagged SR	0.86	0.86	0.76
$m_{\ell\ell jj}$ final tagged SB	0.22	0.33	0.52

Table B.2: Comparisons between data and total background, in the electron channel, using different CRs to obtain the multijet background shape. Results are shown for different variables.

B.2 Normalization studies

The normalization of the multijet background is obtained with the ABCD method, as described in Section 4.4.4.2. In order to evaluate the robustness of the method, tests are performed by changing the criteria used to define CRs B and C and also by obtaining the normalization, as a percentage of the total background, at an earlier stage of the event selection.

In the case of muons, CR B is substituted with the requirement that both muons exhibit relative track-isolation between 20% and 100%. The results are summarized in Table B.3. The observed effect from the modification of the anti-isolation criterion is $\sim 9\%$ in both the untagged and tagged categories. The stage of the event selection at which the estimation is obtained appears to affect the normalization by $\sim 4\%$ in the untagged category and as much as 33% in the tagged category. All of the yields shown in the Table B.3 have been obtained in the mass interval $100 < m_{\ell\ell jj} < 300$ GeV.

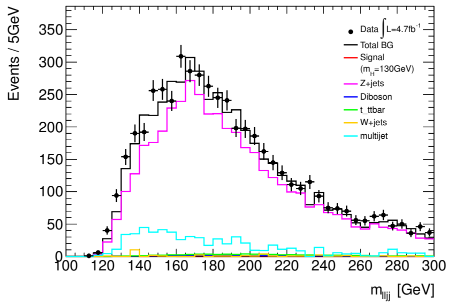
In the case of electrons, an alternative definition of CR B is obtained by exploiting the electron identification quality. Namely, the anti-isolation requirement is substituted by the requirement that one of the electrons passes the *loose*++but fails the *medium*++ criteria. The results are presented in Table B.4. The modified definition of CR B leads to a ~31%(15%) variation of the result, while the stage of the event selection at which the estimation is made affects the normalization by ~4%(10%) in the untagged (tagged) category. It must be stated that the systematic uncertainty estimated using the above measurements is very conservative since the modified CR fails to descibe the data and thus gives rise to such a big variation. All yields are obtained within the mass interval $100 < m_{\ell\ell jj} < 300$ GeV.

Selection	QCD event yeld ABCD	 ABCD modified isolation
Dilepton +2 jets	832±43 (0.69%)	746±38 (0.59%)
$m_{\ell\ell}$ +2 jets	698±37 (3.67%)	627±35 (3.31%)
Final untagged SR	356±27 (3.78%)	325±24 (3.46%)
Final untagged SB	214±21 (4.32%)	189±19 (3.83%)
Dilepton +2 *b*-tagged jets	24± 7 (11.2%)	20± 6 (9.66%)
Final tagged SR	14± 5 (16.7%)	12± 5 (15.3%)
Final tagged SB	3± 4 (5.45%)	2± 3 (3.85%)

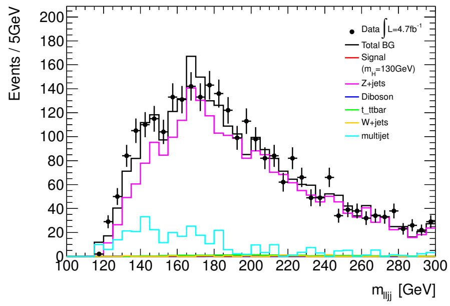
Table B.3: Multijet background yield in the muon channel, at different stages of the event selection using different criteria for the application of the ABCD method. The numbers in parentheses show the ratio of the multijet to the total background.

Selection	QCD event yeld charge - Isolation	 charge - Identification
Dilepton +2 jets	1375±91 (1.8%)	915±58 (1.2%)
$m_{\ell\ell}$ +2 jets	1207±62 (11.1%)	801±40 (8.0%)
Final untagged SR	579±42 (11.6%)	375±26 (7.9%)
Final untagged SB	322±31 (11.8%)	226±19 (8.8%)
Dilepton +2 b jets	30±10 (20.4%)	25± 8 (17.3%)
Final tagged SR	10± 4 (18.6%)	8± 4 (15.5%)
Final tagged SB	7± 6 (16.9%)	4± 3 (9.8%)

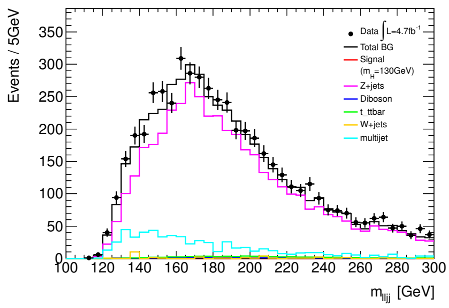
Table B.4: Multijet background yield in the electron channel, at different stages of the event selection using different criteria for the application of the ABCD method. The numbers in parentheses show the ratio of the multijet to the total background.



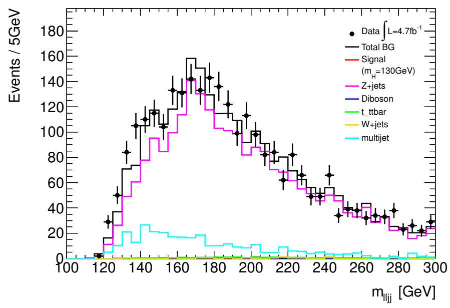
(a)



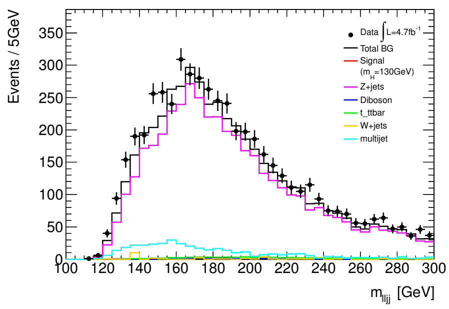
(b)



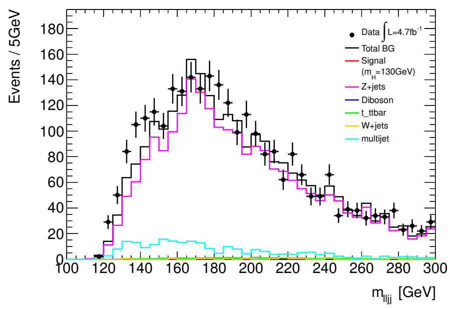
(c)



(d)



(e)



(f)

Figure B.2: m_{eejj} distributions in the SR (left) and the m_{jj} SBs (right). The shape of the multijet background is obtained from the sample of same charge electrons (top), opposite charge, non isolated electrons (middle) and the sample of opposite charge, isolated candidates where one is loose/non-medium (bottom).

Appendix C

Resolution and binning of m_{lljj}

C.1 Resolution

The resolution of the signal m_{lljj} distribution has important consequences for the applicability of the NWA signal samples in the 2HDM model; the NWA is only valid in regions of the 2HDM parameter space where the experimental resolution is less than the signal width predicted by the 2HDM [27]. To investigate this, the experimental resolution is extracted by iteratively fitting a Gaussian lineshape to the reconstructed m_{lljj} distribution of the NWA signal, over a $\pm 2\sigma$ window, until no change above 1% is observed. The results of the resolution versus m_H are shown in Figure C.1, where it can be seen that the width increases from ~ 5 GeV at $m_H = 200$ GeV to 25 GeV at $m_H = 1$ TeV. Figure C.2 shows the fractional resolution relative to m_H, which is relatively flat in the range $2-3\%$. This is significantly narrower than the natural width in some regions of the 2HDM parameter space, and therefore must be taken into account (see Section 5.7.2).

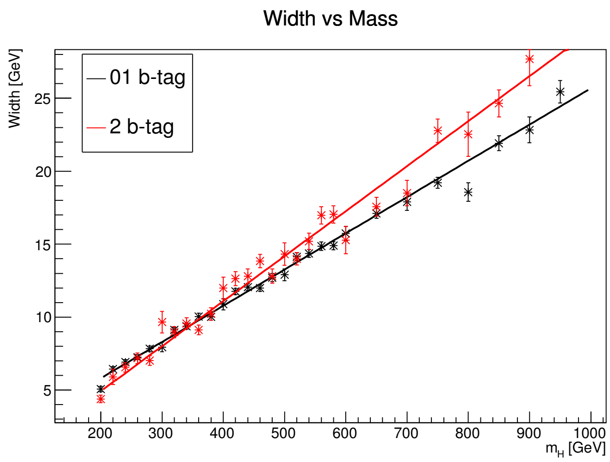


Figure C.1: Resolution of the m_{lljj} distribution as a function of m_H for both untagged and tagged ggF channels combined. The errors are statistical only.

C.2 Binning

Variable sized binning has been chosen for the m_{lljj} distribution since the background falls sharply with m_{lljj}, while the resolution increases, as seen in figure C.1. The binning scheme ensures a reasonable number of background events in each bin and that the bin size is finer than the signal resolution. The binning is defined as follows:

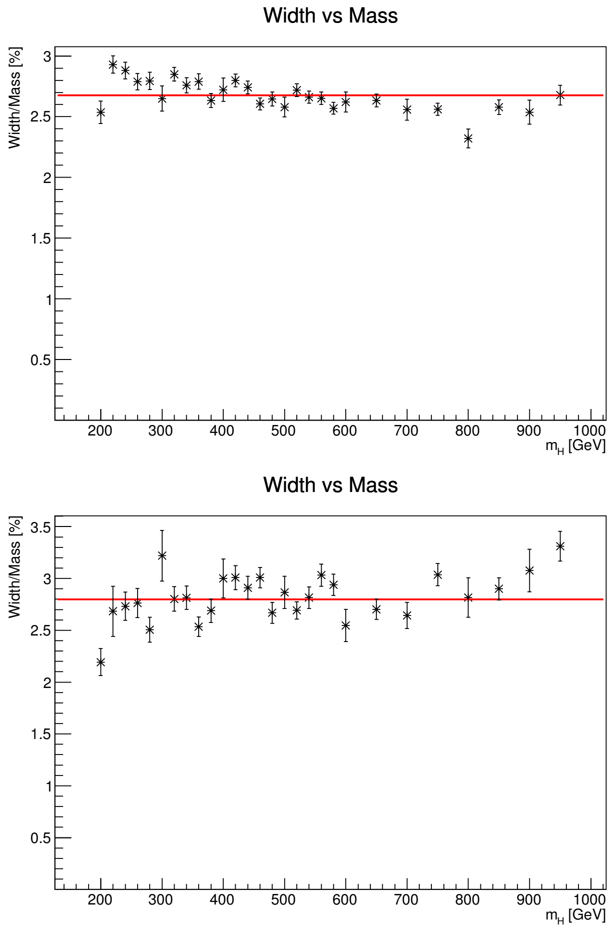


Figure C.2: Fractional resolution of the m_{lljj} distribution as a function of m_H for both untagged (up) and tagged (down) ggF channels. The errors are statistical only.

- Using the total MC background distribution and starting with a minimum bin width of 8 GeV,
- the bin width either remains the same or increases with increasing m_{lljj}.
- For the ggF channel, in the region $300 < m_{lljj} < 900$ GeV ($m_{lljj} > 900$ GeV) successive bins are merged so that no bin exhibits statistical error greater than 5% (15%).
- For the VBF channel and $300 < m_{lljj} < 600$ GeV ($m_{lljj} > 600$ GeV) the maximum statistical bin error is 15% (25%).

The maximum MC statistical error is increased at higher masses to prevent getting a very wide bin, wider than the signal resolution. The cuts on statistical errors for the VBF channel are larger, due to the lower background efficiency and worse MC statistics.

Appendix D

Optimization of the VBF selection

VBF induced final states have specific topology, with the two tag jets lying in regions of large $|\eta|$, pointing in opposite directions in z and exhibiting large invariant mass. On the other hand, tight jets are central and resemble the products of a Z boson decay. Starting from events with at least four jets, the first step is to decide which jets make up which pair. Two approaches have been studied.

First the tag-jets are selected and then the same procedure as in the resolved-ggF channel is followed, using the remaining jets. Two variables have been studied for the selection of the tag jets; the invariant mass $m_{jj,\mathrm{tag}}$ and the separation in pseudorapidity $\Delta\eta_{jj,\mathrm{tag}}$. Both variables lead to similar results in terms of signal acceptance, background rejection and significance, in slight favour of the invariant mass. This indicates that there is a small number of cases where one of the tag-jets lies between the tight jets in η.

Alternatively, one can reconstruct the $Z \to q\bar{q}$ candidate first and then choose, as tag-jets, the pair with the largest invariant mass among the remaining jets. For the former, a possible approach is to pick the two leading in p_{T} jets, mimicing the resolved-ggF selection; that has been found to be inadequate, since the jets from the Z boson decay do not necessarily have higher transverse momenta than the tag-jets. Another option would be to consider the jet pair with invariant mass closest to the Z boson mass, either by using the reconstructed mass directly or by performing a kinematic fit. This approach is also rejected since it often occurs that at least one of the VBF inducing jets lies within $|\eta| < 2.5$ and so gives rise to fake combinations that bias the resulting m_{jj} distribution.

From the above studies, the following procedure is finally decided:

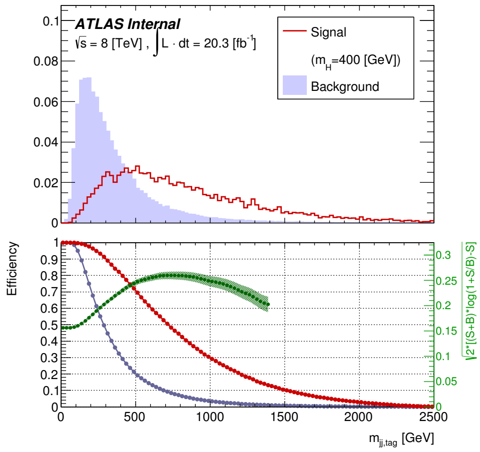
1. From the possible pairs of non b-tagged, loose jets, lying in opposite hemispheres ($\eta_1 \times \eta_2 < 0$), the pair exhibiting the highest invariant mass $m_{jj,\mathrm{tag}}$ is kept.

2. $m_{jj,\mathrm{tag}} > 500$ GeV and $|\Delta\eta_{jj,\mathrm{tag}}| > 4$ is required.

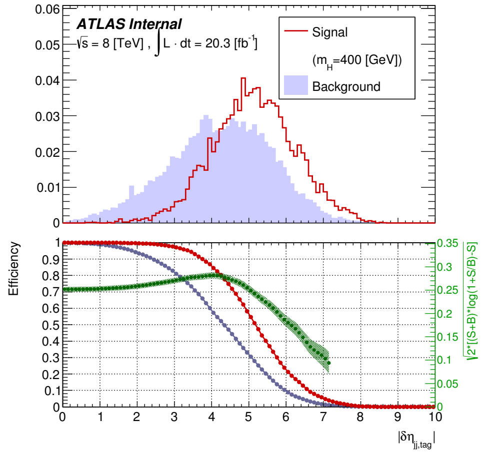
3. At least two tight jets should remain for further analysis.

The optimization of the criteria on the tag-jet invariant mass and pseudorapidity gap is done by measuring the significance from the expected number of signal and background events, in a window around the reconstructed Higgs boson mass peak, containing 90% of the MC signal. Results are shown in Figure D.1. The optimization is repeated iteratively for each of the tested variables until the results stabilize.

For the final part of the event selection, the procedure used for the resolved-ggF case has been found to be optimal for the VBF channel as well (see Figure D.2), within statistical uncertainty. Several additional variables have been tested in attempt to improve the significance. The best performing of those is the vector sum of



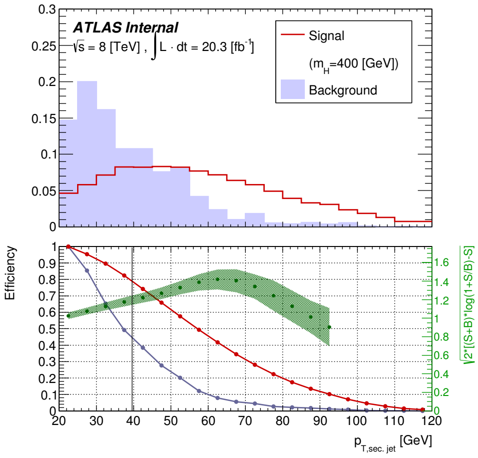
(a)



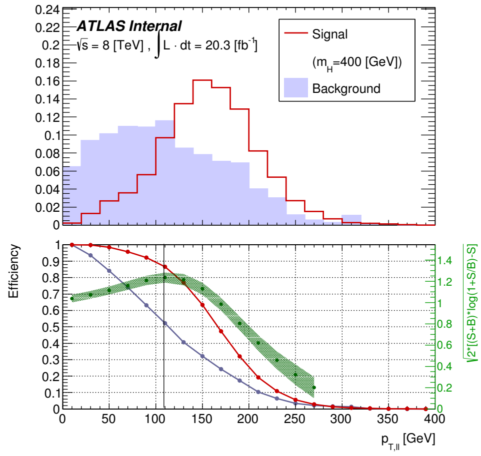
(b)

Figure D.1: Optimization of the cut on m_{jj} (left) and $\Delta\eta_{jj}$ (right). The VBF signal component and background distributions on the top are both normalized to unity. The curves on the bottom plots show the signal (red) and background (blue) efficiencies versus cut value, while the green curve shows the significance versus cut value.

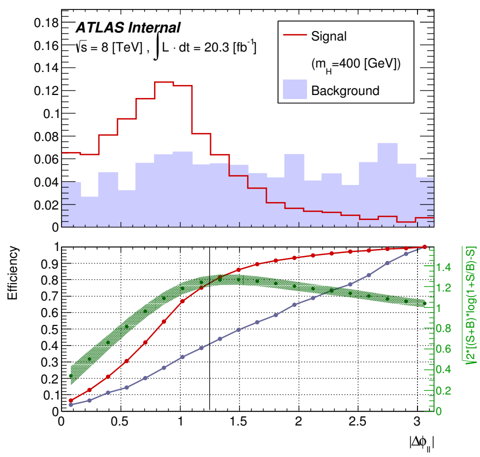
momenta $p_{\mathrm{T,sum}}$ of all final state objects, i.e. the two leptons and the four jets. As can be seen in Figure D.3, imposing a cut on this quantity provides some discrimination between signal and background but also between the ggF and VBF signal modes. Nevertheless, it has been decided not to use this variable since the available MC statistics is insufficient for a robust model of the background. The distribution of $p_{\mathrm{T,sum}}$ is shown in Figure D.4.



(a)

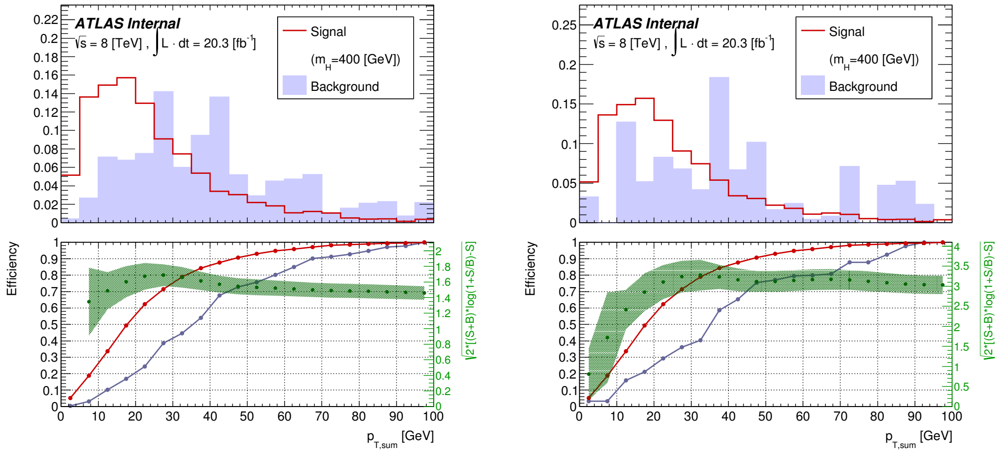


(b)



(c)

Figure D.2: Optimization (VBF-signal versus background) of the cutting values on (a) the jet p_{T} threshold, (b) $p_{\mathrm{T}}^{\ell\ell}$ and (c) $\Delta\phi_{\ell\ell}$, for a single Higgs boson mass point. The distributions on the top sides of the plots are normalized to unity. The curves on the bottom sides show the signal (red) and background (blue) efficiencies versus cut value, while the green curves show the significance versus cut value.



(a) (b)

Figure D.3: Optimization of the cutting value on the vector sum of transverse momenta of the final state objects. The VBF signal component is plotted against (a) the total background and (b) the ggF signal component. The distributions on the top are normalized to unity. The curves on the bottom plots show the signal (red) and background (blue) efficiencies versus cut value, while the green curve shows the significance versus cut value.

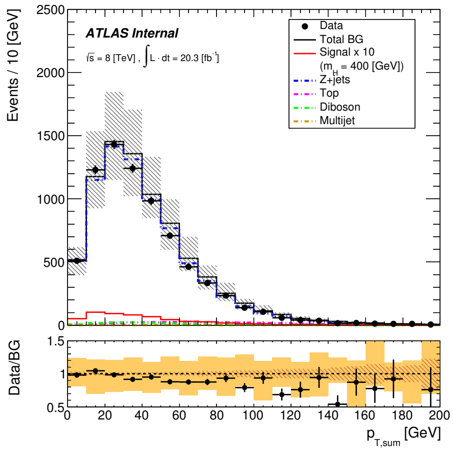


Figure D.4: Distribution of $p_{\mathrm{T,sum}}$ in the VBF channel.

Appendix E

Efficiency of event selection criteria

$\mu\mu jj$ channel	ggH130 MC	ggH130 MC
No cuts	30000	30000
HFOR weight	30000	30000
GoodRunsList	30000	30000
LAr error	30000	30000
Trigger	7641	7641
vertex	7641	7641
$E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning	7641	7641
LAr Hole	7618	7618
Exactly 2 leptons	4590	4591
Opposite charge	4587	4588
Extra kinematics	4297	4301
$E_{\mathrm{T}}^{\mathrm{miss}}$	3684	3689
≥ 2 jets	1574	1577
< 2 b-tagged jets	1521	1524
$m_{\ell\ell}$	742	739
m_{jj}	540	541
2 b-tagged jets	52	52
$m_{\ell\ell}$	34	34
m_{jj}	30	30
> 2 b-tagged jets	1	1
$m_{\ell\ell}$	1	1
m_{jj}	1	1

Table E.1: Efficiencies of the event selection criteria (unweighted) described in Section 4.3.1. The simulated samples contain $H \to ZZ^* \to \ell\ell qq$ decays, where $\ell = e, \mu, \tau$. The numbers of events in the left column are measured when no corrections have been applied to the reconstructed objects momenta.

$eejj$ channel	ggH130 MC	ggH130 MC
All	30000	30000
HFOR weight	30000	30000
GoodRunsList	30000	30000
LAr error	30000	30000
Trigger	6779	6779
Vertex	6779	6779
$E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning	6779	6779
LAr Hole	6714	6714
Exactly 2 leptons	2546	2546
Opposite charge	2536	2536
Extra kinematics	2500	2499
$E_{\mathrm{T}}^{\mathrm{miss}}$	2130	2129
2 jets	850	849
< 2 b-tagged jets	823	822
$m_{\ell\ell}$	373	372
m_{jj}	252	251
2 b-tagged jets	27	27
$m_{\ell\ell}$	23	23
m_{jj}	23	23
> 2 b-tagged jets	0	0
$m_{\ell\ell}$	0	0
m_{jj}	0	0

Table E.2: Efficiencies of the event selection criteria (unweighted) described in Section 4.3.1. The simulated samples contain $H \to ZZ^* \to \ell\ell qq$ decays, where $\ell = e, \mu, \tau$. The numbers of events in the left column are measured when no corrections have been applied to the reconstructed objects momenta.

$eejj$ channel	H400NWA MC	H900NWA MC
No cuts	30000	30000
MC weight	30000 / 30000.00	30000 / 30000.00
HFOR weight	30000 / 30000.00	30000 / 30000.00
POWHEG weight	30000 / 30043.56	30000 / 32932.46
GoodRunsList	30000 / 30043.56	30000 / 32932.46
Evt error	30000 / 30043.56	30000 / 32932.46
Trigger	11009 / 11005.15	12200 / 13382.00
PileUp reweight	11009 / 11142.89	12200 / 13370.48
Vertex	11007 / 11140.45	12195 / 13368.57
$E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning	10989 / 11131.20	12175 / 13351.35
LAr hole jet veto	10989 / 11131.20	12175 / 13351.35
Exactly 2 leptons	8098 / 8250.20	8945 / 9852.22
Trigger matching/SF	8095 / 8235.68	8942 / 9840.38
Opposite charge / SF	8095 / 8222.53	8942 / 9820.58
$\geq$ 2 tight jets	7316 / 7444.75	8372 / 9187.90
$m_{\ell\ell}$	6102 / 6192.07	7013 / 7745.32
$E_{\mathrm{T}}^{\mathrm{miss}}$	5781 / 5938.73	6080 / 6830.82
No binning (highest-pt pair only)		
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	5690 / 5837.87	6070 / 6826.29
m_{jj}	2915 / 2990.14	3044 / 3527.82
$\Delta\phi_{jj}$ corr.	2915 / 2831.53	3044 / 3231.27
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	2345 / 2258.25	1783 / 1884.01
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	2154 / 2071.78	1291 / 1405.58
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	1842 / 1774.79	1285 / 1395.14
0 b-tagged jets	4275 / 4413.67	4462 / 5068.78
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	4205 / 4336.99	4452 / 5064.02
m_{jj}	2168 / 2232.58	2246 / 2610.11
$\Delta\phi_{jj}$ corr.	2168 / 2114.67	2246 / 2390.36
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	1740 / 1680.44	1335 / 1415.43
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	1598 / 1548.10	956 / 1048.90
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	1361 / 1324.85	952 / 1039.39
1 b-tagged jet	973 / 1005.06	1052 / 1234.91
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	955 / 980.08	1052 / 1234.91
m_{jj}	454 / 451.01	485 / 604.64
$\Delta\phi_{jj}$ corr.	454 / 427.14	485 / 554.73
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	333 / 299.37	229 / 260.04
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	302 / 269.12	165 / 188.34
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	264 / 236.73	164 / 186.36
2 b-tagged jets	519 / 521.96	545 / 573.30
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	513 / 520.61	537 / 567.39
m_{jj}	410 / 408.43	414 / 443.32
$\Delta\phi_{jj}$ corr.	410 / 385.61	414 / 407.23
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	318 / 297.63	199 / 188.83
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	292 / 268.70	154 / 147.40
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	260 / 240.41	153 / 146.55
>= 3 b-tagged jets	14 / 11.08	21 / 19.60

Table E.3: Efficiencies of the event selection criteria described in Section 5.3.1. The simulated samples contain $H \to ZZ \to \ell^+\ell^- q\bar{q}$ decays, where $\ell = e, \mu, \tau$. Both unweighted and weighted numbers of events are presented.

$\mu\mu jj$ channel	H400NWA MC	H900NWA MC
No cuts	30000	30000
MC weight	30000 / 30000.00	30000 / 30000.00
HFOR weight	30000 / 30000.00	30000 / 30000.00
POWHEG weight	30000 / 30043.56	30000 / 32932.46
GoodRunsList	30000 / 30043.56	30000 / 32932.46
Evt error	30000 / 30043.56	30000 / 32932.46
Trigger	10033 / 10065.73	11041 / 12175.66
PileUp reweight	10033 / 9988.43	11041 / 12142.56
Vertex	10032 / 9986.64	11039 / 12140.71
$E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning	10011 / 9970.79	11012 / 12119.75
LAr hole jet veto	10011 / 9970.79	11012 / 12119.75
Exactly 2 leptons	7800 / 7837.75	8303 / 9187.04
Trigger matching/SF	7790 / 7819.92	8281 / 9150.91
Opposite charge / SF	7786 / 7814.26	8278 / 9146.01
$\geq$ 2 tight jets	7037 / 7111.20	7784 / 8589.31
$m_{\ell\ell}$	5831 / 5928.52	6023 / 6639.02
$E_{\mathrm{T}}^{\mathrm{miss}}$	5364 / 5570.76	4905 / 5536.34
No binning (highest-pt pair only)		
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	5263 / 5469.99	4893 / 5524.59
m_{jj}	2707 / 2795.07	2439 / 2779.05
$\Delta\phi_{jj}$ corr.	2707 / 2648.84	2439 / 2547.46
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	2147 / 2070.15	1427 / 1520.09
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	1944 / 1871.69	954 / 997.73
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	1673 / 1625.28	950 / 993.37
0 b-tagged jets	3896 / 4127.34	3689 / 4180.88
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	3823 / 4058.21	3679 / 4171.31
m_{jj}	1996 / 2093.72	1822 / 2086.91
$\Delta\phi_{jj}$ corr.	1996 / 1985.49	1822 / 1912.97
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	1562 / 1527.51	1075 / 1150.83
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	1404 / 1366.16	714 / 742.37
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	1208 / 1189.14	711 / 738.90
1 b-tagged jet	948 / 953.86	789 / 877.03
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	926 / 929.17	788 / 877.03
m_{jj}	475 / 454.26	365 / 395.39
$\Delta\phi_{jj}$ corr.	475 / 429.44	365 / 362.61
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	352 / 328.40	187 / 184.94
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	322 / 305.93	122 / 120.59
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	269 / 258.16	122 / 120.59
2 b-tagged jets	502 / 504.88	409 / 489.06
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	486 / 477.05	405 / 482.61
m_{jj}	371 / 356.55	325 / 395.61
$\Delta\phi_{jj}$ corr.	371 / 337.29	325 / 363.30
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	292 / 254.61	153 / 175.84
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	253 / 228.07	106 / 119.63
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	225 / 206.07	105 / 118.97
>= 3 b-tagged jets	18 / 17.36	18 / 24.32

Table E.4: Efficiencies of the event selection criteria described in Section 5.3.1. The simulated samples contain $H \to ZZ \to \ell^+\ell^- q\bar{q}$ decays, where $\ell = e, \mu, \tau$. Both unweighted and weighted numbers of events are presented.

$eejj$ channel (VBF)	H400NWA VBF MC	H900NWA VBF MC
No cuts	50000	50000
MC weight	50000 / 50000.00	50000 / 50000.00
HFOR weight	50000 / 50000.00	50000 / 50000.00
POWHEG weight	50000 / 49735.48	50000 / 49695.38
GoodRunsList	50000 / 49735.48	50000 / 49695.38
Evt error	50000 / 49735.48	50000 / 49695.38
Trigger	18476 / 18459.74	20243 / 20180.61
PileUp reweight	18476 / 18605.22	20243 / 20104.96
Vertex	18475 / 18603.60	20235 / 20097.86
$E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning	18369 / 18534.25	20121 / 20014.60
LAr hole jet veto	18369 / 18534.25	20121 / 20014.60
Exactly 2 leptons	13779 / 13968.82	14914 / 14979.09
Trigger matching/SF	13776 / 13940.75	14911 / 14962.51
VBF jet pair & ≥ 2 tight jets	7077 / 7201.91	6213 / 6283.81
VBF $m_{jj,tag}$	4963 / 5104.02	5426 / 5483.92
VBF $\Delta\eta_{jj,tag}$	4253 / 4353.67	4421 / 4464.33
Opposite charge / SF	4253 / 4346.97	4421 / 4455.97
$m_{\ell\ell}$	3560 / 3610.48	3720 / 3753.16
$E_{\mathrm{T}}^{\mathrm{miss}}$	3383 / 3469.60	3320 / 3399.82
No binning (highest-pt pair only)		
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	3300 / 3399.49	3291 / 3364.47
m_{jj}	2244 / 2295.30	2513 / 2584.15
$\Delta\phi_{jj}$ corr.	2244 / 2176.18	2513 / 2373.58
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	1700 / 1611.40	1295 / 1255.30
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	1429 / 1369.09	948 / 917.27
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	1229 / 1186.25	945 / 915.96
0 b-tagged jets	2493 / 2555.38	2452 / 2582.36
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	2429 / 2499.11	2427 / 2554.11
m_{jj}	1712 / 1756.53	1862 / 1955.35
$\Delta\phi_{jj}$ corr.	1712 / 1665.83	1862 / 1796.54
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	1290 / 1236.18	968 / 969.45
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	1088 / 1048.69	701 / 696.75
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	941 / 908.86	700 / 695.58
1 b-tagged jet	539 / 542.51	505 / 489.28
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	523 / 528.13	502 / 482.97
m_{jj}	313 / 319.32	354 / 359.34
$\Delta\phi_{jj}$ corr.	313 / 302.27	354 / 329.87
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	234 / 207.37	163 / 157.38
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	191 / 171.59	117 / 111.76
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	159 / 144.23	116 / 111.76
2 b-tagged jets	349 / 369.27	360 / 345.64
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	343 / 366.61	358 / 344.06
m_{jj}	282 / 297.06	310 / 301.79
$\Delta\phi_{jj}$ corr.	282 / 281.07	310 / 276.51
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	200 / 191.25	164 / 141.73
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	169 / 165.57	132 / 115.52
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	148 / 147.72	131 / 115.20
> 2 b-tagged jets	2 / 4.62	3 / 4.83

Table E.5: Efficiencies of the event selection criteria described in Section 5.3.1. The simulated samples contain $H \to ZZ \to \ell^+\ell^- q\bar{q}$ decays, where $\ell = e, \mu, \tau$. Both unweighted and weighted numbers of events are presented.

$\mu\mu jj$ channel (VBF)	H400NWA VBF MC	H900NWA VBF MC
No cuts	50000	50000
MC weight	50000 / 50000.00	50000 / 50000.00
HFOR weight	50000 / 50000.00	50000 / 50000.00
POWHEG weight	50000 / 49735.48	50000 / 49695.38
GoodRunsList	50000 / 49735.48	50000 / 49695.38
Evt error	50000 / 49735.48	50000 / 49695.38
Trigger	16955 / 16843.86	18447 / 18317.47
PileUp reweight	16955 / 16749.30	18447 / 18373.53
Vertex	16953 / 16747.12	18435 / 18362.03
$E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning	16856 / 16684.80	18325 / 18279.55
LAr hole jet veto	16856 / 16684.80	18325 / 18279.55
Exactly 2 leptons	13258 / 13119.31	13881 / 13941.47
Trigger matching/SF	13240 / 13110.85	13826 / 13886.88
VBF jet pair & ≥ 2 tight jets	6719 / 6676.64	5879 / 6006.57
VBF $m_{jj,tag}$	4711 / 4613.36	5087 / 5210.24
VBF $\Delta\eta_{jj,tag}$	3967 / 3832.78	4115 / 4205.24
Opposite charge / SF	3967 / 3831.66	4112 / 4201.90
$m_{\ell\ell}$	3257 / 3147.93	3192 / 3272.44
$E_{\mathrm{T}}^{\mathrm{miss}}$	2990 / 2951.37	2658 / 2759.06
No binning (highest-pt pair only)		
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	2881 / 2838.26	2632 / 2729.37
m_{jj}	1906 / 1889.24	2004 / 2106.41
$\Delta\phi_{jj}$ corr.	1906 / 1787.72	2004 / 1936.41
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	1426 / 1332.19	1000 / 942.78
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	1167 / 1104.16	664 / 609.42
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	1021 / 963.00	662 / 607.83
0 b-tagged jets	2201 / 2208.14	1977 / 2086.37
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	2118 / 2112.27	1951 / 2056.48
m_{jj}	1417 / 1413.18	1497 / 1602.94
$\Delta\phi_{jj}$ corr.	1417 / 1336.89	1497 / 1474.17
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	1050 / 1000.76	766 / 739.28
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	847 / 817.36	514 / 476.74
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	738 / 710.76	513 / 476.32
1 b-tagged jet	497 / 464.88	397 / 416.47
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	475 / 451.50	397 / 416.47
m_{jj}	303 / 287.75	273 / 294.73
$\Delta\phi_{jj}$ corr.	303 / 272.92	273 / 271.39
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	214 / 181.19	119 / 106.01
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	183 / 155.27	75 / 62.60
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	166 / 138.28	74 / 60.90
2 b-tagged jets	291 / 282.64	283 / 264.70
$p_{\mathrm{T}}^{j\ lead} > 45GeV$	284 / 274.37	283 / 264.70
m_{jj}	227 / 227.87	244 / 233.97
$\Delta\phi_{jj}$ corr.	227 / 215.40	244 / 214.34
$p_{\mathrm{T}}^{j}(m_{\ell\ell jj})$	163 / 153.31	113 / 93.89
$p_{\mathrm{T}}^{\ell\ell}(m_{\ell\ell jj})$	139 / 132.39	74 / 68.14
$\Delta\phi_{\ell\ell}(m_{\ell\ell jj})$	119 / 112.06	74 / 68.14
> 2 b-tagged jets	1 / 0.00	1 / 2.11

Table E.6: Efficiencies of the event selection criteria described in Section 5.3.1. The simulated samples contain $H \to ZZ \to \ell^+\ell^- q\bar{q}$ decays, where $\ell = e, \mu, \tau$. Both unweighted and weighted numbers of events are presented.

List of Figures

List of Tables