SEARCH FOR HIGGS BOSON PAIR PRODUCTION IN THE $b\bar{b}WW^*$ CHANNEL IN $\sqrt{S} = 13$ TEV PROTON-PROTON COLLISIONS AT THE LARGE HADRON COLLIDER USING THE ATLAS DETECTOR

by

JOHNCHRISTOPHER STEPHEN MYERS

A DISSERTATION

Presented to the Department of Physics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2019

DISSERTATION APPROVAL PAGE

Student: JohnChristopher Stephen Myers

Title: Search for Higgs Boson Pair Production in the $b\bar{b}WW^*$ Channel in $\sqrt{s} =$ 13 TeV Proton-Proton Collisions at the Large Hadron Collider Using the ATLAS Detector

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Physics by:

Stephanie Majewski	Chair
Eric Torrence	Advisor
Tim Cohen	Member
Hank Childs	Institutional Representative

and

Janet Woodruff-Borden	Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2019

DISSERTATION ABSTRACT

JohnChristopher Stephen Myers

Doctor of Philosophy

Department of Physics

June 2019

Title: Search for Higgs Boson Pair Production in the $b\bar{b}WW^*$ Channel in $\sqrt{s} =$ 13 TeV Proton-Proton Collisions at the Large Hadron Collider Using the ATLAS Detector

This dissertation presents a search for double Higgs production in the $b\bar{b}WW^*$ final state in proton-proton collisions at the ATLAS detector at the Large Hadron Collider. Double Higgs production is predicted in the Standard Model with a cross section of $\sigma_{HH} = 33.53\text{fb}\ ^{+4.3\%}_{-6.0\%} \pm 5.9\%$. Many Beyond the Standard Model theories predict enhancements to the production cross section through resonant production.

The 2015-2016 ATLAS dataset has an integrated luminosity of 36.1 fb^{-1} with a center of mass energy of $\sqrt{s} = 13$ TeV. Candidate events are broken into two kinematic regions: a resolved selection containing one lepton (either an electron or muon), 2 b-tagged calorimeter jets, 2 light-flavor jets, and missing transverse energy; and a boosted analysis containing one lepton (electron or muon), two large radius jets, one with two ghost associated, b-tagged track-jets and missing transverse energy. No significant deviation from background was observed a cross section upper limit

was set for the SM double Higgs production of 10 pb and for resonant production as a function of HH invariant mass from 500 GeV to 3000 GeV.

This dissertation includes previously published coauthored material.

CURRICULUM VITAE

NAME OF AUTHOR: JohnChristopher Stephen Myers

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
The Ohio State University, Columbus, Ohio

DEGREES AWARDED:

Doctor of Philosophy, Physics, 2018 University of Oregon
Bachelor of Science, Physics, 2013, The Ohio State University

AREAS OF SPECIAL INTEREST:

Boosted object reconstruction
Trigger and detector operations

PROFESSIONAL EXPERIENCE:

Graduate Research Assistant, University Of Oregon: ATLAS Collaboration, 2015-Present

Graduate Teaching Assistant, University Of Oregon: Department of Physics, 2013-2015

PUBLICATIONS:

List of publications with significant contributions:

The ATLAS collaboration. "Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2019.4 (Apr. 2019), p. 92. ISSN: 1029-8479. DOI: 10.1007/JHEP04(2019)092. URL: https://doi.org/10.1007/JHEP04(2019)092

Co-authored on 124 publications with minor contributions can be found at Inspire (search parameters: exactauthor:John.Myers.1)

ACKNOWLEDGEMENTS

I would like to thank my advisor Eric Torrence for giving me this opportunity and guiding me throughout my time at Oregon. Thank you to the remaining members of my committee; especially Stephanie Majewski for your mentorship over the years. Thank you to David Strom for your help during my time at CERN. And to Anne McGinley for all of your help. Thank you to all of the post-docs who watched out for me at CERN, both with physics and life in general: Aparajita Dattagupta, Walter Hopkins, Jochen Heinrich, and Kate Whalen. Thank you to all my colleges who suffered with me: Ben Allen, Jason Barkeloo, Johan Bonilla, Liza Brost, Tyler Burch, Elliot Parrish, Ian Snyder, Bri Stamas, Amanda Steinhebel , and Chaowaroj 'Max' Wanotayaroj.

Thank you to my family. You have always believed in me and your encouragement got me through many tough times. And finally and most importantly, thank you to my wife, Alison Jelden, and her family. Alison, without your support and love, I would not be where I am today.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

LIST OF TABLES

Table Page

Table Page

Table Page

CHAPTER I

INTRODUCTION

The Standard Model (SM) is the culmination of more than a century of work. The first piece added to the puzzle was the electron, discovered in 1891. Since then, 24 other particles have been discovered, with the final piece, the Higgs Boson, being added in 2012. Since it was theorized, the SM has held up to rigorous experimentation. Even though the SM is widely successful, it fails to explain all observed phenomena. Dark matter, baryogenesis, the hierarchy problem, along with other observations, all lack explanation within the SM. The remaining task is to probe the extremes of the SM to either more precisely measure the parameters or to find evidence for new physics effects. In this dissertation I will test the SM by searching for anomalous production of di-Higgs events that is not predicted by the SM. This anomalous production would be a clear signature of new phenomena and would expand the current knowledge of the physical world by measuring new, previously undiscovered particles.

This chapter gives an overview of the Standard Model with a focus on Electro-Weak Symmetry Breaking and the Higgs mechanism. Chapter II includes a discussion of di-Higgs production, both in the SM and beyond the SM. Chapter III describes the experimental apparatus used in this dissertation. Chapter IV discusses the generation of background samples and reconstruction of data for the search. Chapter V and Chapter VI gives a detailed description of the analysis along with an additional analysis technique. Chapter VII concludes the dissertation. Chapter V includes material coauthored with the ATLAS Collaboration.

1.1. The Standard Model

The Standard Model defines the basic building blocks of matter and force carriers and the interactions between them. Everyday matter is made of protons, neutrons, and electrons. Electrons are a fundamental particle, called a lepton, meaning they are not made of smaller constituents. However, protons and neutrons are not fundamental particles. They are a composite built of up and down quarks, along with sea quarks and gluons, which are more fundamental particles. The protons are "made of" two ups and a down and the neutrons are "made of" two downs and an up. The up and down quarks, along with the electron, are types of fermions.

Fermions are spin-$\frac{1}{2}$ particles that make up all matter in the SM. The fermions can be broken down into three "generations". A generation contains two quarks, one with electric charge $+\frac{2}{3}$ and one with electric charge $-\frac{1}{3}$, one electrically charged lepton, charge -1, and one electrically neutral lepton, along with their antiparticles. The quarks have an additional color charge, of which there are three charges. This is an additional quantum number associated with the strong force. In all, this gives 12 fermions.

Gauge bosons are spin-1 particles responsible for communicating the fundamental forces in the SM. There are four physical gauge bosons in the electroweak sector. The photon γ is a massless, charge neutral force carrier for the electromagnetic force. The nuclear forces are carried by two massive gauge bosons: an electrically neutral Z boson and a charged W boson, which has electron charge $= \pm 1$. Together, these two bosons control the electroweak interactions in the SM. The remaining gauge boson is the gluon, which comes in eight color-anticolor combinations. The gluon is the force carrier for the strong nuclear interaction. A

	Particle	Spin	Charge	Mass
Quarks				
u type	u	$\frac{1}{2}$	$\frac{2}{3}$	$2.4^{+0.6}_{-0.4}$ MeV
	c			1.28 ± 0.03 GeV
	t			173.1 ± 0.6 GeV
d type	d	$\frac{1}{2}$	$-\frac{1}{3}$	$4.7^{+0.5}_{-0.4}$ MeV
	s			96^{+8}_{-4} MeV
	b			$4.18^{+0.04}_{-0.03}$ GeV
Leptons				
e doublet	e	$\frac{1}{2}$	-1	$0.5109989461 \pm 0.000000003$ MeV
	ν_e		0	< 2 eV
μ doublet	μ	$\frac{1}{2}$	-1	$105.6583745 \pm 0.0000024$ MeV
	ν_μ		0	< 2 eV
τ doublet	τ	$\frac{1}{2}$	-1	1776.86 ± 0.12 GeV
	ν_τ		0	$<$ 2eV
Bosons				
Vector	γ	1	0	$< 10^{-18}$ eV
	g	1	0	0
	W	1	$\pm$	80.385 ± 0.0015 GeV
	Z	1	0	91.1876 ± 0.0021 GeV
Scalar	H	0	0	$125.09 \pm 0.21 \pm 0.11$ GeV

TABLE 1.1. Particles of the Standard Model [2]

gluon is a massless, electrically neutral particle that has two color charges.

The remaining piece of the SM is the Higgs Boson. The Higgs boson is a massive scalar, spin-0, chargeless boson. The Higgs boson is responsible for giving mass to the massive fundamental particles. The full list of SM particles and their properties are in Table 1.1.

Along with the previously outlined particles, there are also anti-particles. Every particle in the SM has a partner of identical mass but opposite charge. For example, the anti-particle to the electron is the positron; a particles with electric charge of +1. For the W boson, the positive and negative Ws are particle-antiparticle pairs. Some particles are their own anti particle, for example, the anti-particle of a photon is a photon, the same is true for the Z boson.

1.1.1. Interactions

The SM explains interactions between fermions by the exchange of bosons (integer spin). For leptons, the overarching theory is the electroweak theory. This can be broken down further into the electromagnetic interaction (Quantum Electrodynamics (QED)) and the weak interaction. The electromagnetic interaction defines the interaction of electrically charged particles with photons. An example electromagnetic interaction diagram is electron-positron annihilation Figure 1.1, where an electron and a positron collide and produce two photons. This can also be reversed, two photons interact and produce an electron-positron pair. The strength of this interaction is the electrical charge e.

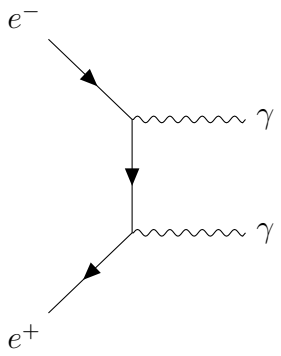

FIGURE 1.1. Electron-Positron Annihilation

The weak interaction defines the interaction of particles under the weak isospin quantum number and hypercharge. In the SM, every fermion is a mix of a left and a right-handed chirality. Particles with a right-handed chirality have a weak isospin T = 0. Left-handed particles have a weak isospin T = $\frac{1}{2}$. These particles are part of a doublets as illustrated in Table 1.2. For these particles, the third component of the weak isospin T_3, $+\frac{1}{2}$ for up-type quarks and charged leptons and $-\frac{1}{2}$ for down-type

quarks and neutral leptons.

Left Handed Fermions, $T = \frac{1}{2}, T_3 = \pm\frac{1}{2}$	Right Handed Fermions, $T = 0, T_3 = 0$
$\binom{u}{d}$, $\binom{c}{s}$, $\binom{t}{b}$, $\binom{e}{\nu_e}$, $\binom{\mu}{\nu_\mu}$, $\binom{\tau}{\nu_\tau}$	u, d, c, s, t, b, e, ν_e, μ, ν_μ, τ, ν_τ

TABLE 1.2. Fermion doublets and singlets [3]

The remaining pieces of the weak interaction are the W and Z bosons. The W has an isospin of T = 1. This gives three options for the third component of isospin, $T_3 = +1, 0, -1$ which give the W^+, the W^0, and the W^-. The W^0 will be discussed more in section 1.1.2. The $W^\pm$ either raises or lowers the T_3 of the fermions. The Z boson has a weak isospin of 0 meaning it does not change the isospin of the fermions. Instead, the Z boson transfers momentum, energy, and spin in interactions that do not change electric charge or weak isospin. Additionally, the Z boson couples to both the left and right handed fermions. Figure 1.2 is an example of a weak interaction.

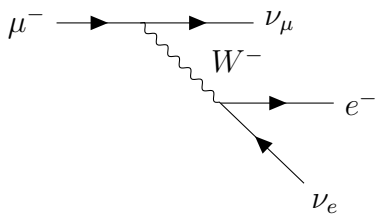

FIGURE 1.2. Muon emitting a muon neutrino and a W Boson

The last fundamental forces in the SM is quantum chromodynamics (QCD). Unlike the electroweak force, as the distance between a pair of colored particles increases, the force between them increases. A consequence of this is known as color confinement. As the distance between the particles becomes greater, the energy

stored grows until a new quark-antiquark pair is spontaneously is generated from the vacuum between them to neutralize the color of the original quarks. This generation of color neutral hadrons is known as hadronization and occurs whenever a quark gains enough momentum to be ejected from an object. The same thing happens to gluons, as they too cannot exist by themselves. After hadronization, the resultant hadrons produce showers of particles when they interact with dense matter. These showers are recontructed in detectors as objects called jets.

Gravity is outside the scope of the SM and is not explained here.

1.1.2. The Higgs Mechanism and Higgs Boson

The Electroweak portion of the Standard Model has four massless bosons: the W_1, W_2, W_3, and B. In order to generate observed massive bosons, the W and Z, and the photon, Electroweak symmetry must be broken. This process of Spontaneous Electroweak Symmetry Breaking is known as the Higgs Mechanism.

In Electroweak Theory, the four fundamental gauge bosons couple to a complex scalar doublet, $\Phi \equiv \binom{\phi^+}{\phi^0}$. This doublet has a scalar potential of the form

$$V(\Phi) = \mu^2|\Phi^\dagger\Phi| + \lambda(|\Phi^\dagger\Phi|)^2. \tag{1.1}$$

When $\mu^2 < 0$, the potential looks like the wine bottle shaped potential seen in Figure 1.3, with a minimum energy at

$$\langle\phi\rangle = \sqrt{-\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}} \tag{1.2}$$

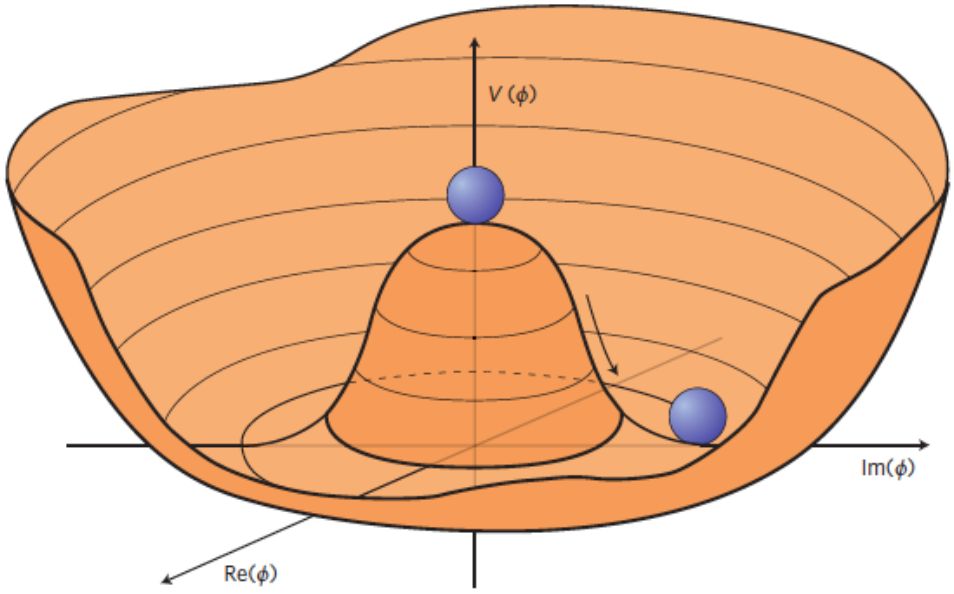

FIGURE 1.3. The Higgs potential.[4]

called the vacuum expectation value (VEV) of ϕ. The choice of the direction of fluctuation is arbitrary. By convention, we choose

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix}. \tag{1.3}$$

After the direction is chosen, three Goldstone bosons are eaten and the only remaining degree of freedom is the real scalar field h(x), giving

$$\phi(x) = \phi_0 + h(x). \tag{1.4}$$

The doublet can now be described by

$$\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v+h(x)\end{pmatrix}. \tag{1.5}$$

This field couples to the gauge bosons as

$$(\frac{g}{2}\overrightarrow{\tau}\cdot\overrightarrow{W}+\frac{g'}{2}B)\phi_0. \tag{1.6}$$

Where $\overrightarrow{\tau}$ are the Pauli matrices, $\overrightarrow{W}$ are $W_{1,2,3}$ and g, g' are the coupling constants for weak and hypercharge respectively. The result of the coupling is the four eigenstates, which are the observed bosons,

$$\begin{aligned} W^{\pm} &= \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \\ Z^\mu &= \frac{-g'B_\mu + gW^3_\mu}{\sqrt{g^2+g'^2}} \\ A^\mu &= \frac{gB_\mu + g'W^3_\mu}{\sqrt{g^2+g'^2}} \end{aligned} \tag{1.7}$$

Three of these eigenstates are massive, the $W^{\mu^{\pm}}$ and the Z^μ. A^μ, the photon, remains massless.

$$\begin{aligned} M_W^2 &= \frac{1}{4}g^2v^2 \\ M_Z^2 &= \frac{1}{4}(g^2+g'^2)v^2 \\ M_A &= 0 \end{aligned} \tag{1.8}$$

As we just saw, an additional scalar field, the Higgs Field, is required.

The Higgs Boson is an excitation of the scalar Higgs field theorized by Peter Higgs [5] [6]; François Englert and Robert Brout[7]; and Gerald Guralnik, Carl Hagen and Tim Kibble [8]. In 2012, a Higgs like scalar boson was discovered at the LHC by the ATLAS and CMS experiments with a mass of 125 GeV [9].

Since the discovery, many measurements have been made of this Higgs Boson to compare it to the SM Higgs Boson. So far, the Higgs Boson has held up to these tests.

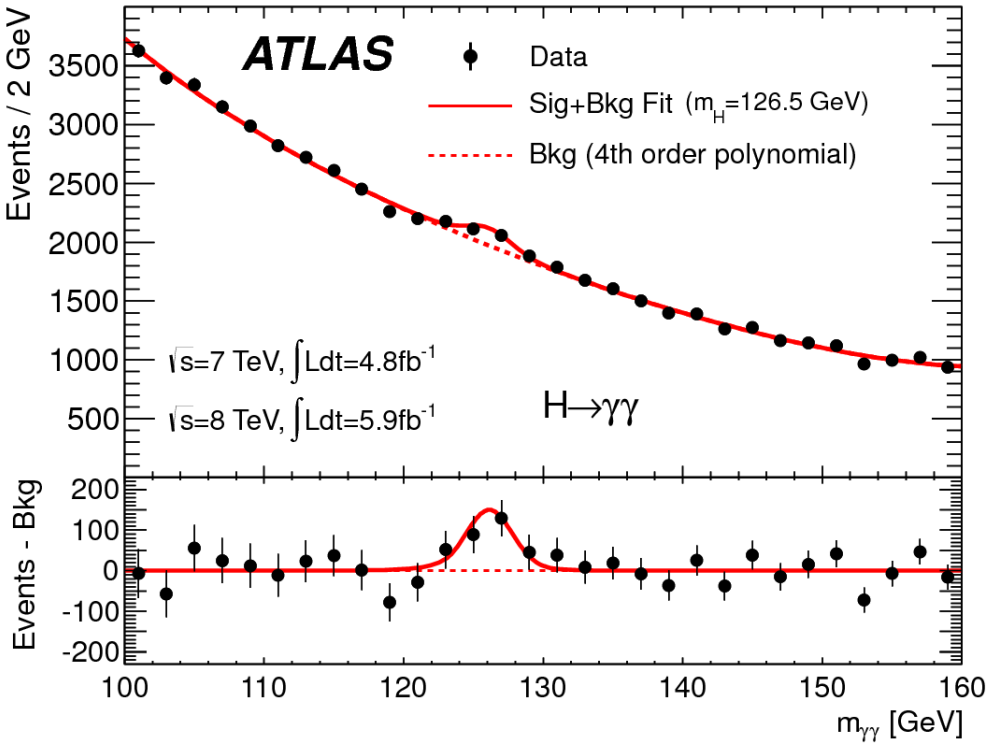

FIGURE 1.4. Invariant mass distribution of di-photon candidates for the combined root $\sqrt{s}$= 7 TeV and $\sqrt{s} = 8$ TeV data samples. The result of a fit to the data of the sum of a signal component fixed to $m_H = 126.5$ GeV and a background component described by a fourth-order Bernstein polynomial is superimposed. The bottom inset displays the residuals of the data with respect to the fitted background component.

The Higgs Boson has spin-parity $J^P = 0^+$ [10], and its decay branching fractions to bb [11], $\gamma\gamma, \tau\tau$ [12], WW and ZZ have been measured with appropriate signal strengths, and no significant deviations with respect to the SM have been observed in any Run 2 analyses. However, there are still many parameters of the Higgs Boson that still need measured. One of which is the triple Higgs coupling.

CHAPTER II

DI-HIGGS PRODUCTION

Measuring di-Higgs production is necessary to further confirm the SM or find evidence for beyond-the-SM physics. The small SM production rate makes the channel an important place to look for new physics. In particular, the di-Higgs production rate gives a handle to more accurately measure the Higgs potential. This dissertation looks at both the measurement of the SM di-Higgs production rate and to search for new physics through resonant di-Higgs production.

2.1. Standard Model

The Higgs self coupling potential in the SM is

$$V_{\mathrm{self-coupling}} = \lambda(|\Phi^{\dagger}\Phi|)^2 \tag{2.1}$$

When Φ is expanded around the vacuum expectation value (VEV), v, the self coupling term can be written as

$$V_{\mathrm{self-coupling}} \supset \lambda v\Phi^3 + \frac{\lambda}{4}\Phi^4 \tag{2.2}$$

where the first term, $\lambda v\Phi^3$ is the coupling between three Higgs bosons with strength $\lambda_{HHH} \equiv \lambda v$[13]. The tri-linear Higgs coupling can be probed at the LHC by measuring the cross section of events with two Higgs Bosons in the final state.

Currently at the LHC, there are two dominant ways to produce di-Higgs events, the tri-linear Higgs coupling gluon-gluon fusion (ggF) diagram and a box diagram,

Figure 2.1. These diagrams interfere destructively, resulting in a SM prediction for the cross section,

$$\sigma_{\mathrm{HH}} = 33.53\mathrm{fb}^{+4.3\%}_{-6.0\%}(\text{QCD unc.}) \pm 5.9\%(\text{other unc.}) \tag{2.3}$$

in pp collisions at 13 TeV [14]. With the current data sample, this small cross section is challenging to measure. Additionally, since the cross section is so small, it is an promising place to look for deviations from the SM, since any enhancement over the SM predicted cross section would be indicative of new physics.

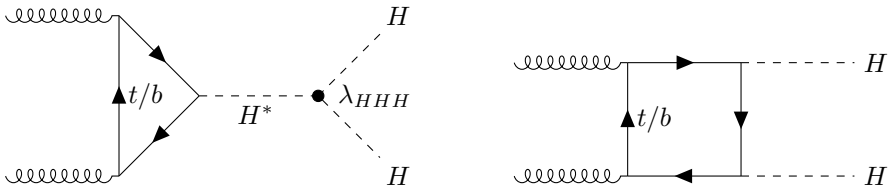

FIGURE 2.1. The dominate production method for di-Higgs events at the LHC with $\sqrt{s} = 13$ TeV, with the tri-linear Higgs coupling on the left

The SM di-Higgs production is a continuum production, with events produced with an invariant mass above 2 m_H. There can be events below this (off-shell H) but the rate will go down quickly as you get further off-shell. Figure 2.2 shows the continuum distribution, as expected, it is a falling power law distribution peaked around 400 GeV.

2.2. Resonant Production

There are several BSM models that may enhance the rate of di-Higgs production at the LHC. This section will give an overview of a few of these models

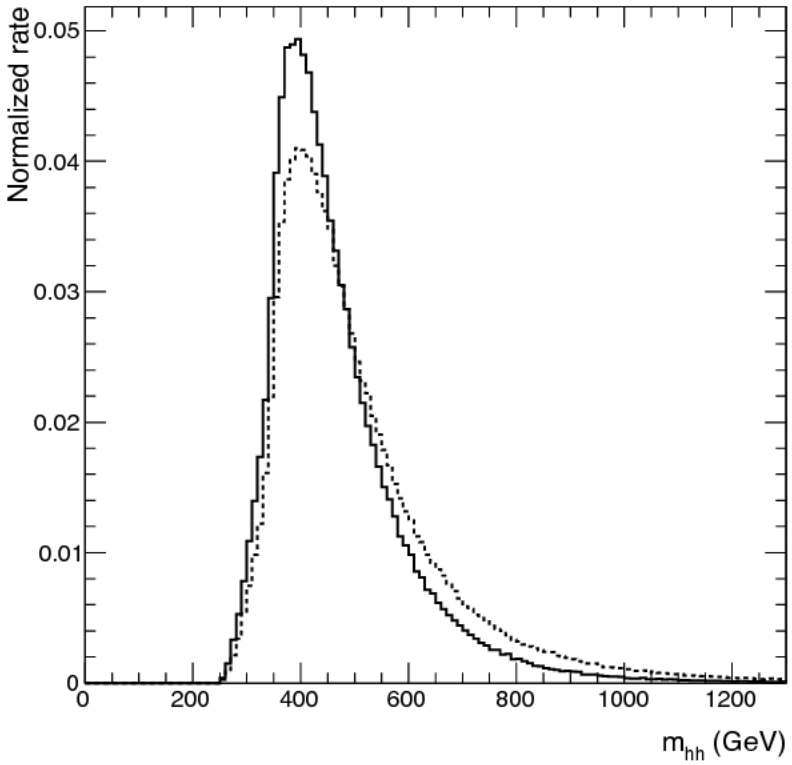

FIGURE 2.2. Normalized differential cross section for pp $\rightarrow$ hh in the SM as a function of the invariant mass of the two Higgs bosons. The solid and dotted lines correspond respectively to $\sqrt{s} = 14$ and 100 TeV .[15]

2.2.1. Complex Higgs Singlet

The addition of a complex scalar singlet to the SM results in three neutral scalar particles after spontaneous symmetry breaking, which mix to give mass eigenstates, including the observed 125 GeV scalar [16].

The normalizable scalar potential is

$$\begin{aligned} V(\Phi, S_c) = \frac{\mu^2}{2}\Phi^\dagger\Phi + \frac{\lambda}{4}(\Phi^\dagger\Phi)^2 \\ +(\frac{1}{4}\delta_1\Phi^\dagger\Phi S_c + \frac{1}{4}\delta_3\Phi^\dagger\Phi S_c^2 \\ +a_1 S_c + \frac{1}{4}b_1 S_c^2 + \frac{1}{6}e_1 S_c^3 + \frac{1}{6}e_2 S_c|S_c|^2 \\ +\frac{1}{8}d_1 S^4 + \frac{1}{8}d_3 S_c^2|S_c|^2 + \text{H.C.}) \\ +\frac{1}{4}d_2(|S_2|^2)^2 + \frac{\delta_2}{2}\Phi^\dagger\Phi|S_c|^2 + \frac{1}{2}b_2|S_c|^2 \end{aligned} \tag{2.4}$$

After spontaneous symmetry breaking, the fields are defined as:

$$\Phi = \begin{pmatrix} 0 \\ \frac{h+v}{\sqrt{2}} \end{pmatrix}; S_c = \frac{1}{\sqrt{2}}(S + v_s + i(A + v_A)). \tag{2.5}$$

v_A is set to 0 to conserve CP symmetry.

The mass eigenstate fields are given by:

$$\begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \begin{pmatrix} c_1 & -s_1 & 0 \\ s_1c_2 & c_1c_2 & s_2 \\ s_1s_2 & c_1s_2 & -c_2 \end{pmatrix} \begin{pmatrix} h \\ S \\ A \end{pmatrix} \tag{2.6}$$

where $c_i = \cos\theta_i$, h is the SU(2) doublet field, and S and A are the real and imaginary components of the complex scalar S_c. Assigning the SM-like Higgs boson as h_1 with $m_1 = 125$ GeV and $v = 246$ GeV, h_2 and h_3 are physical heavy Higgs bosons with $m_2, m_3 > m_1$. The coupling of h_1 to SM particles is dominant, suppressed by a factor of c_1 from the SM rate, with the h_2 couplings suppressed by s_1c_2 and h_3 couplings suppressed by s_1s_2. ATLAS has set limits on $c_1 > 0.94$ at 95% CL in RUN I. As the h_1 couplings become more SM-like ($\theta_1 \to 0$), the allowed h_2 couplings become

suppressed.

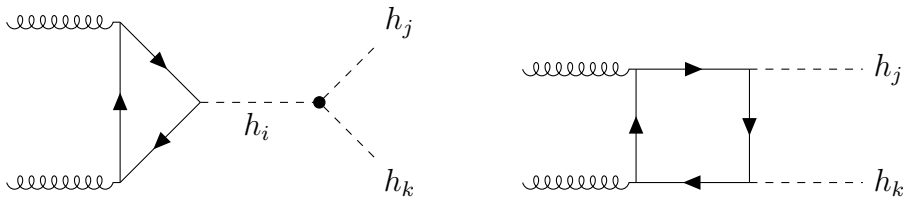

FIGURE 2.3. Feynman diagrams for the production of $h_j h_k$, $i, j, k = 1, 2, 3$.

In the limit of $\theta_2 \to 0$, which is in agreement with the single Higgs rates, h_3 does not directly couple to SM fermions or vector bosons. The only way to produce h_3 is through h_1 or h_2, with the largest production rate from $gg \to h_2 \to h_1 h_3$, Figure 2.3. For a range of masses m_2 and m_3 the rate of production of $h_1 h_3 \gg h_1 h_1$, Figure 2.4.

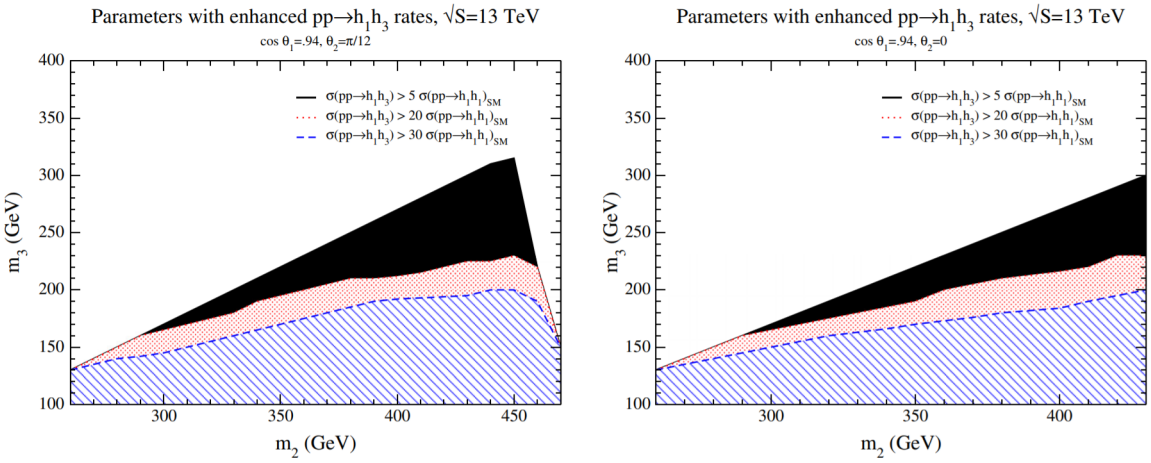

FIGURE 2.4. Regions of parameter space allowed by limits on oblique parameters S T and U from Ref [17], perturbative unitarity of $2 \to 2$ scattering process [18], and the minimization of the potential where the rate for $h_1 h_3$ production is significantly larger than the SM $h_1 h_1$ rate at $\sqrt{S} = 13$ TeV.

The enhancement can be see in the potential where

$$V \to \frac{1}{2}\lambda_{211}h_1^2h_2 + \frac{1}{2}\lambda_{311}h_1^2h_3 + \frac{1}{2}\lambda_{331}h_1h_3^2 + \frac{1}{2}\lambda_{321}h_1h_2h_3 + ... \tag{2.7}$$

So while the SM tri-linear Higgs coupling is determined by m_h, with this extension, the coupling is much less constrained. This leads to enhanced values seen in Figure 2.5. So while this model would definitely show up in SM di-Higgs production, through the first two terms in equation 2.7, a search for one SM Higgs and a heavy Higgs would be more sensitive. This is a promising search moving forward but is not the focus of this dissertation.

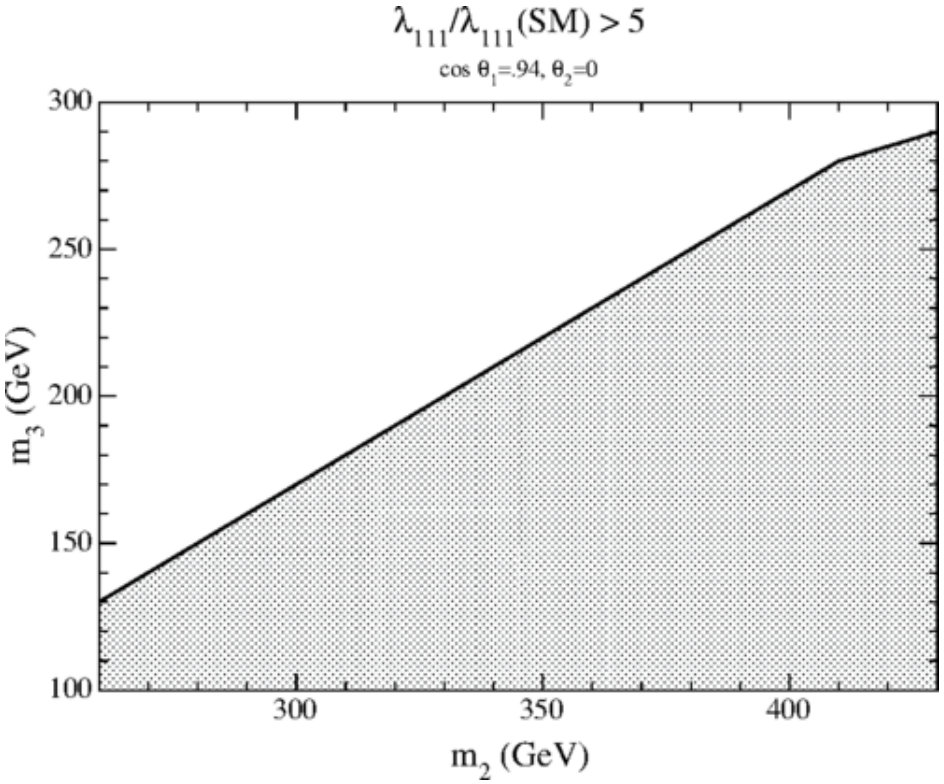

FIGURE 2.5. Region of parameter space allowed by limits on oblique parameters, perturbative unitarity and the minimization of the potential where the $h_1h_1h_1$ tri-linear coupling is greater than five times the SM value.

2.2.2. Real Higgs Singlet Extension

One simple explanation of an enhanced di-Higgs production rate at the LHC is the addition of a real scalar Higgs singlet, S [19]. In this model, S can only interact with the SM through the Higgs field. In the case where there is no Z_2 symmetry, where $S \to -S$ the scalar field S mixes with the SM Higgs boson. If the mass is large enough, it is possible for S to decay to two on-shell SM Higgs Bosons, significantly enhancing the di-Higgs production rate.

The most general potential that can be added is

$$V(\Phi, S) = -\mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 + \frac{a_1}{2}\Phi^\dagger\Phi S + \frac{a_2}{2}\Phi^\dagger\Phi S + b_1 S + \frac{b_2}{2}S^2 + \frac{b_3}{3}S^3 + \frac{b_4}{4}S^4. \quad (2.8)$$

Where Φ is $\phi_0 = \frac{(h+v)}{\sqrt{2}}$ and $< \phi_0 >= \frac{v}{2}$, while $S = s + x$ where x is the VEV of S. By shifting the field, it is possible the set $x = 0$. After electroweak symmetry breaking the fields mix to give the two mass eigenstates

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} h \\ s \end{pmatrix} \quad (2.9)$$

With $m_1 = 125GeV$, the free parameters are m_2, θ, a_2, b_2 and b_4. For di-Higgs production, in the case of $m_2 > 2m_1$, the important piece of the potential is

$$V(h_1 h_2) \supset \frac{\lambda_{111}}{3!}h_1^3 + \frac{\lambda_{211}}{3!}h_2 h_1^2 \quad (2.10)$$

This gives an additional resonant double Higgs production diagram , Figure 2.6, for 250 GeV $\leq m_2$.

Varying the values of b_4 and $\sin^2\theta$, it is found that the maximum branching ratio (BR) for $h_2 \to h_1 h_1$ if obtained with $b_4 = 4.2, \sin^2\theta = 0.12$. Figure 2.7, shows

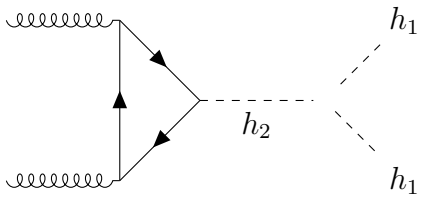

FIGURE 2.6. Feynman diagram for $h_2 \to h_1 h_1$.

the minimum and maximum BR as a function of m_2. The largest BR is when $m \approx 280$ GeV at $BR(h_2 \to h_1 h_1) = 0.76$. This corresponds to an enhancement in di-Higgs production of approximately 30 times the SM cross section.

2.3. Summary

The SM di-Higgs production rate is an important and achievable measurement for the LHC and HL-LHC. It gives insight to the shape of the Higgs potential through measurement of the tri-linear Higgs coupling. It is also a valuable discovery channel for BSM physics, especially for models with an extended Higgs sector, through resonant di-Higgs production. This dissertation will present results for both SM and resonant di-Higgs production.

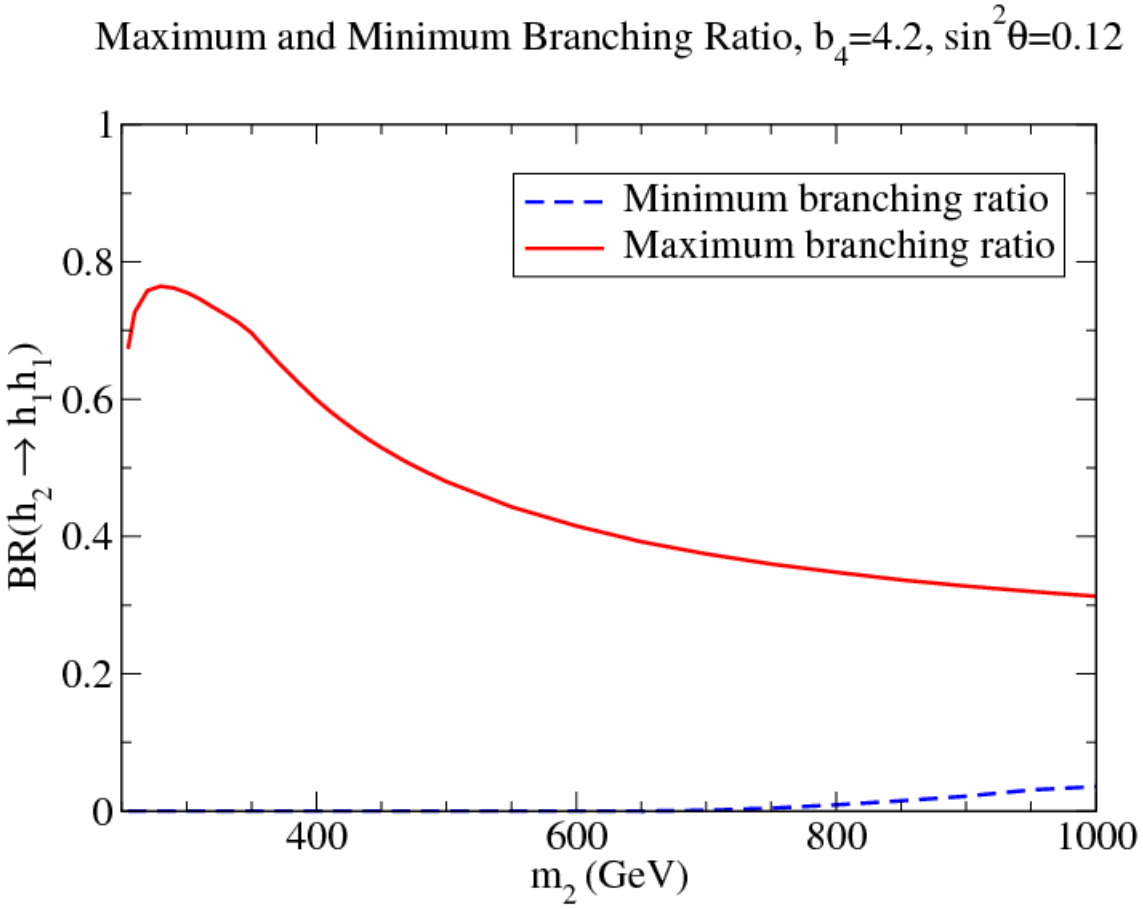

FIGURE 2.7. Maximum and minimum allowed $BR(h_2 \to h_1 h_1)$ as a function of m_2 for $b_4 = 4.2$ and $\sin^2\theta = 0.12$.

CHAPTER III

EXPERIMENTAL SETUP

3.1. Hadron Colliders

Hadron colliders use two beams of hadrons, particles made of three quarks, typically proton-proton or proton-antiproton. The large mass of the hadrons results in smaller synchrotron radiation in circular accelerators, when compared to lepton colliders with equivalent energies, as the radiated power scales as $\frac{1}{m^4}$. This allows hadron colliders to reach much larger center of mass energies with the same size circular ring.

While hadron colliders typically have larger collision energies, they also have significantly "messier" collisions. In lepton colliders, the final state particles only come from the colliding particles. In hadron colliders, not all of the constituent partons of the hadrons interact in the hard collision. This means it is impossible to know the exact longitudinal momentum.

3.2. The Large Hadron Collider

The Large Hadron Collider (LHC) is a 27 kilometer ring underneath the Franco-Swiss border. The LHC accelerates beams of protons (or ions) to a center of mass energy of up to 13 TeV(5 TeV) in two countercirculating beams around the ring. The particles are then collided at four primary interaction points each of which has a dedicated detector: ATLAS, CMS, ALICE, and LHCb.

In order to get the protons up to the collision energy, the LHC uses a series of smaller accelerators in the injection chain. The start of the chain, and the source of protons for the LHC, is the Linear Accelerator 2 (Linac 2) [20]. Hydrogen gas

is taken from a bottle and the electrons are stripped off by an electric field. When only the protons remain, they pass through radiofrequency (RF) cavities. These RF cavities are shaped such that the electromagnetic waves are resonant inside of the cavities and build up energy. When charged particles pass through the cavity, they feel the force of the electric field and are pushed forward. The field in the RF cavities oscillates at a frequency specific to the distance from the previous cavity, giving a specific energy to a passing particle depending on the momentum. When a particle arrives early to the cavity, the field gives a smaller impulse to the particle, when it arrives late it gets a boost from the cavity to get it back to the target energy. When a proton reaches the end of Linac 2, it has an energy of 50 MeV.

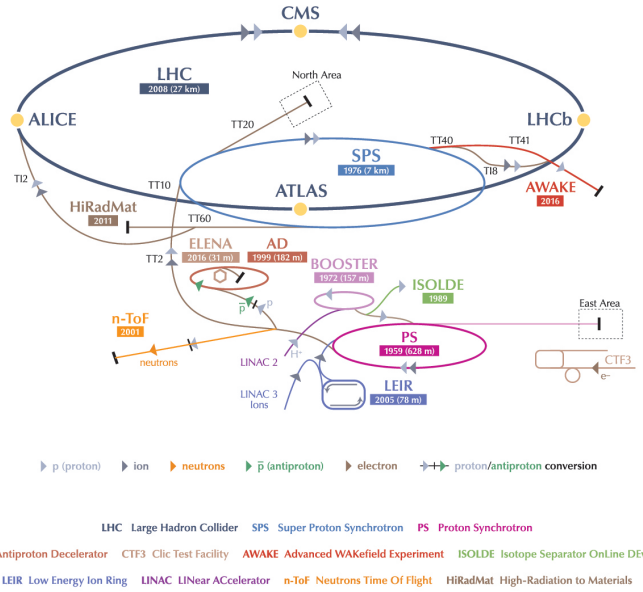

FIGURE 3.1. CERN accelerator complex [21]

After the protons leave the Linac 2, they enter the Proton Synchrotron Booster (PSB)[22]. The PSB accelerates the protons to an energy of 1.4 GeV and are passed to the Proton Synchrotron (PS) in two batches, 1.2 seconds apart. The PS accelerates protons to 26 GeV and delivers them to the next step in the injection chain, the Super Proton Synchrotron (SPS) in a series of four batches, 3.6 seconds apart. The SPS is the second largest accelerator at CERN. As protons pass through, they as boosted to an energy of 450 GeV [23]. Once the particles reach this energy, they are split and injected into the LHC in two opposite directions. Once in the LHC, the protons are brought up to their target energy of 6.5 TeV per beam.

3.2.1. LHC Magnets

Accelerators depend on powerful magnets to bend and focus the colliding particles. The LHC is the most powerful particle accelerator that has ever existed. In order to make this possible, the LHC was constructed with the most advanced magnet technology that could be produced on an industrial scale. These magnets are cooled to below 2 K and produce an extremely uniform field of over 8 T.

The LHC uses 1104 superconducting, dipole magnets to bend the beam of particles around the ring, and another 128 in the beam dump. Each dipole is 15 m long and weighs 35 tonnes. A current of 11,000 Amps pass through an octant in series to produce the magnetic field. Each octant is powered independently. A cross section of a dipole can be seen in Figure 3.2. Inside of these dipoles, there are sextupole, octopole and decapole magnets to correct for small imperfections in the magnetic field at the outside of the dipole [24]. A key feature of the dipole magnets is the 2-in-1 configuration. Where each dipole generates a magnetic field in the opposite direction in the two pipes, Figure 3.3. Allowing a single dipole to bend the two

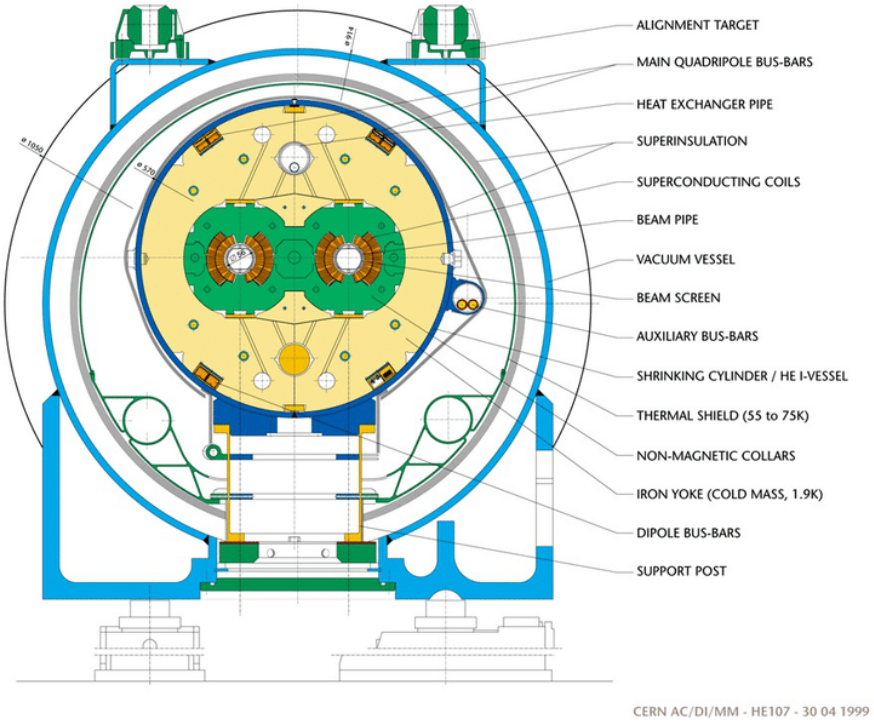

FIGURE 3.2. Cross-section of cryodipole (lengths in mm)

beams in opposite directions.

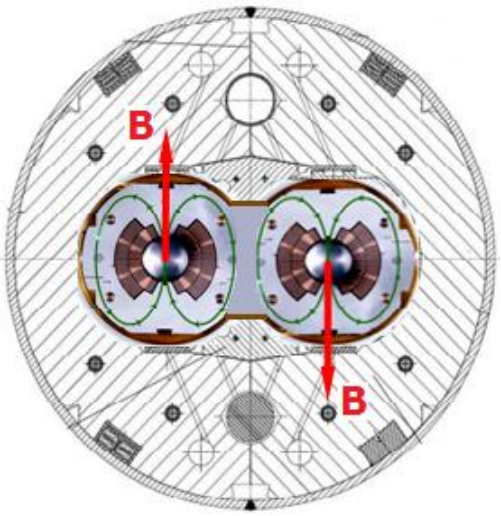

FIGURE 3.3. Magnetic field configuration of the dipole magnets [25]

It is also important to focus the beam to a small area at each interaction point, in order to maximize the collision rate. This is achieved by pair of quadrupole magnets. One magnet focuses in the horizontal plane, defocusing in the vertical, and the other focuses in the vertical plane, defocusing in the horizontal. These magnets work together to focus the beam in both planes.

3.2.2. Luminosity and Pileup

In a collider, a key statistic is the rate of events produced. This value is called luminosity ($\mathcal{L}$) and is defined by equation 3.1 [26], where σ is the cross section, or probability of collision.

$$\mathcal{L} = \frac{1}{\sigma}\frac{dN}{dt} \tag{3.1}$$

In order to maximize this number, the LHC collides bunches of protons. These bunches are made up of 1.15×10^{11} protons. In the LHC during Run II, the LHC beams had around 2500 bunches organized into a train of bunches, separated by small gaps with a larger gap at the end of the train. The luminosity per proton pair crossing, equation 3.2

$$\mathcal{L} = \frac{1}{4\pi\sigma_x\sigma_y} \tag{3.2}$$

is then multiplied by the number of protons per bunch for the two beams, N_1 and N_2, and by the number of bunches N_b and the frequency f. The use of bunches changes the luminosity calculation to equation 3.3)

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y} \tag{3.3}$$

By integrating the luminosity over the running time of the LHC, you obtain the total delivered luminosity, Fig 3.4. By multiplying the integrated luminosity by the

probably of a particular final state, or the cross section, it is possible to obtain the number of times a final state is produced.

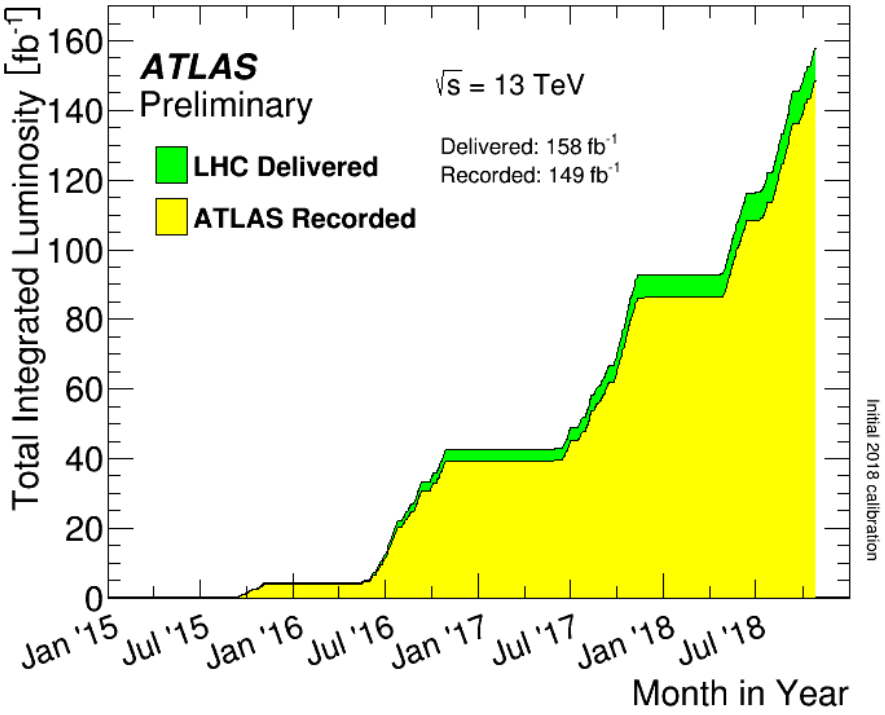

FIGURE 3.4. Cumulative luminosity versus time delivered to ATLAS (green) and recorded by ATLAS (yellow) during stable beams for pp collisions at 13 TeV center-of-mass energy in LHC Run II. (figure from the ATLAS Collaboration)

3.2.3. Pileup

By interacting bunches of protons at a time, it is possible for more that one pair of protons to undergo an inelastic collision. In fact, during Run II, the average number of interactions per bunch crossing, or pileup, was roughly 32. The profile of the pileup for all of Run II can be seen in Figure 3.5.

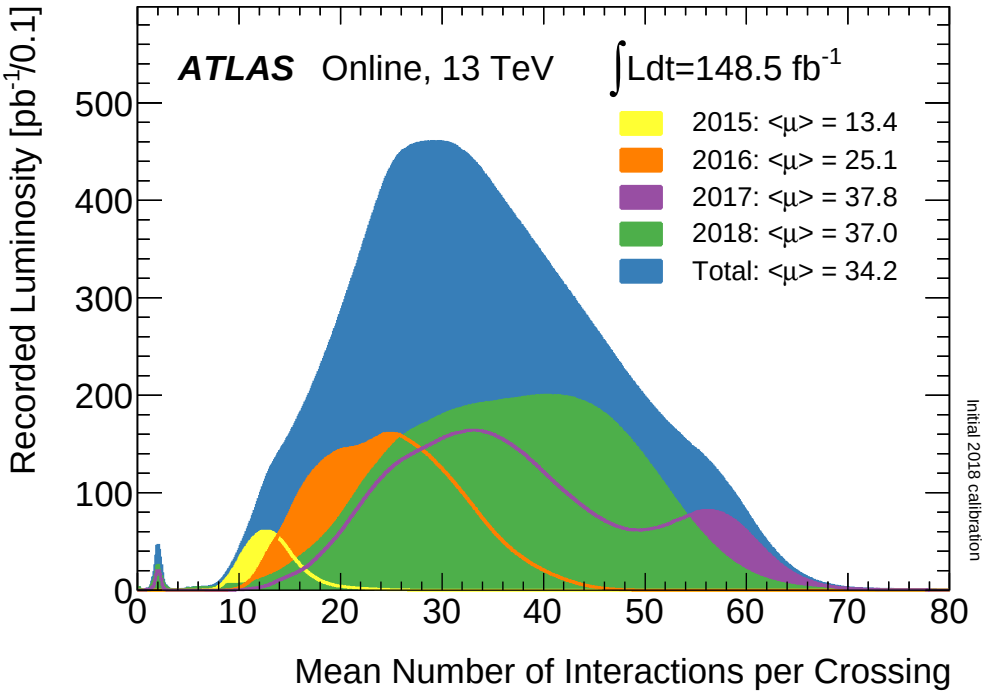

FIGURE 3.5. Shown is the luminosity-weighted distribution of the mean number of interactions per crossing for the 2015-2018 pp collision data at 13 TeV centre-of-mass energy. All data recorded by ATLAS during stable beams is shown, and the integrated luminosity and the mean mu value are given in the figure. (figure from the ATLAS Collaboration)

3.2.4. Detector Coordinates

Within the ATLAS detector, the interaction point defines the origin of the coordinate system. The z-axis, the longitudinal axis, runs along the beam line, the positive x-axis points toward the center of the LHC ring, and the positive y-axis points toward the surface. The detector is also described in r, η, ϕ coordinates. With the transverse plane, the plane perpendicular to the beam line, being described by r and ϕ. The radial coordinate, r, describes the distance from the beam line. The azimuthal angle, ϕ, is the angle from the x-axis around the beam line. The final coordinate, η, is referred to pseudorapidity and is defined as $\eta = -\ln(\tan(\frac{\theta}{2}))$.

With θ being the angle from the y-axis. Differences in η are Lorentz invariant under longitudinal boosts. This means the difference in the rest frame of colliding particles are not important for massless particles. In ATLAS, the large particle boost allows for pseudo-rapidity to be a good estimate for the true rapidity of the particle. Additionally, massless particle are produced uniformly in η. For these reasons, η is preferred over θ. A pictoral representation can be seen in Figure 3.6. The variable $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is often used to describe the distance between detector objects.

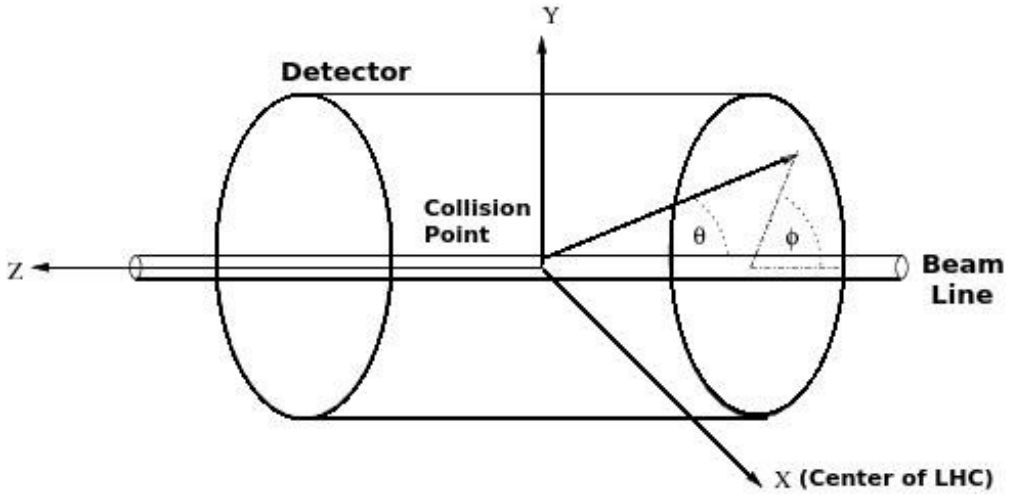

FIGURE 3.6. Detector coordinate system [27]

3.3. Detector Overview

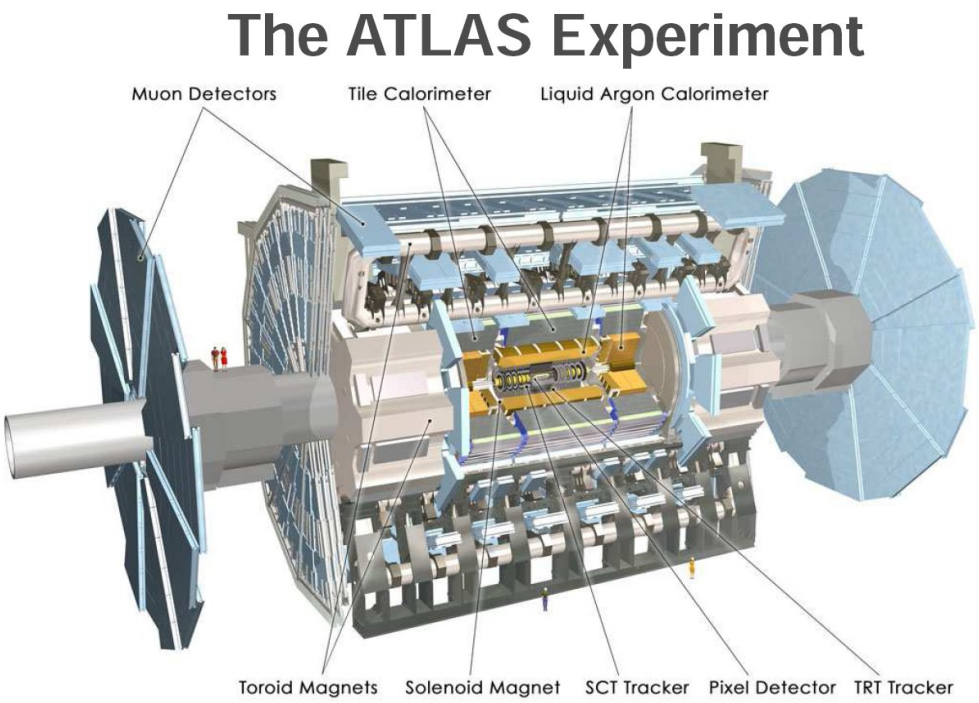

FIGURE 3.7. The ATLAS detector [28]

The ATLAS detector, Figure 3.7, is a general purpose detector and the largest at the LHC. It is made up of concentric subsystems, each with a specialized task: the inner detector, which is responsible for measuring the charge and momentum of charged particles; the calorimeters, which are responsible for measuring the energy of different electromagnetic and hadronic particles; the muon spectrometer, which measures the momentum of minimum ionizing particles (MIP), like muons; and the magnet system, which is responsible for bending the charged particles in the detector, allowing their charge and momentum to be measured. The subdetectors feed into a vast Trigger and Data Acquisition (TDAQ) system that is responsible for selecting collision events with interesting characteristics and reading out detector elements.

3.3.1. The Inner Detector

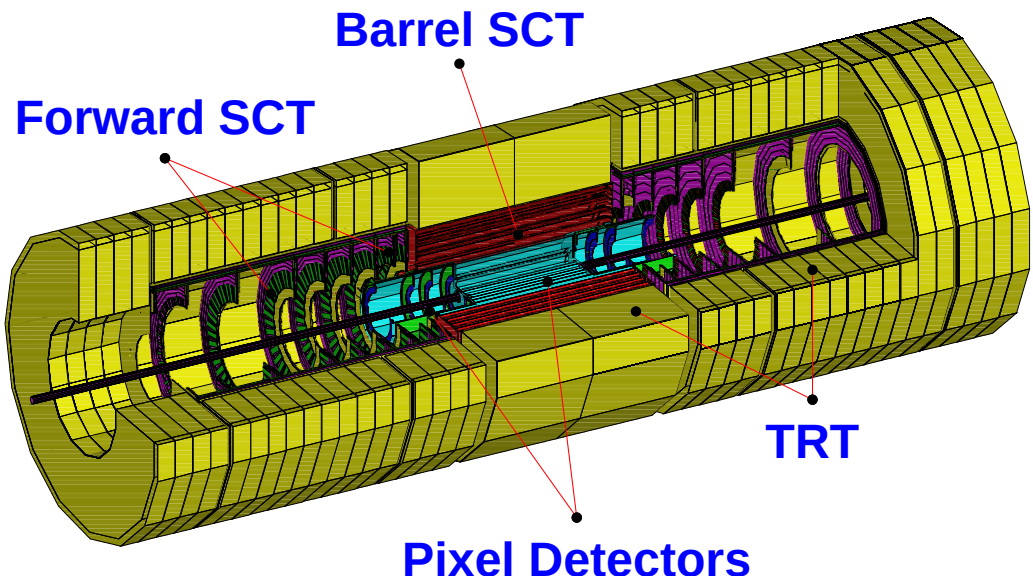

FIGURE 3.8. Cross section of the Inner Detector.

The inner detector (ID), Figure 3.8, is the closest system to the beam pipe and is made of four separate pieces. In order of distance from the beam pipe: the Insertable

B-Layer (IBL) [29], the Pixel Detectors, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT) [30]. These subsystems work together to give charged particle tracking information within the pseudorapidity range of $|\eta| < 2.5$. When a charged particle passes through the silicon semiconductor detector in the IBL, Pixel and SCT, an electron-hole pair is created in the silicon. These electron-hole pairs drift toward the charged readout electrodes on the surface of the detector. This gives a "hit" in the detector. The inner detector is immersed by a 2T solenoid magnet, Section 3.3.4. The magnetic field causes charged particles to curve as they pass through the ID. The hits they leave along the way are connected together into "tracks" that show the trajectory of the particle as it passes through. The radius of curvature and direction of the track gives the sign of the charge, positive or negative, along with a momentum measurement of the particle. The other task of the ID is vertexing, or determining if the transient particle came from the interaction point or a slightly displaced point. This is used to identify long lived particles, like bottom or charm quarks, which travel a small distance before decaying. This is discussed further in section 4.2.4.

The IBL is the newest addition to the ID, being installed during the 2016 shutdown. It is placed directly outside the beam pipe in order to maintain good vertexing and b tagging in increased pileup environments. In order to facilitate the insertion of the IBL, the beam pipe inner radius was decreased by 4 mm between Run I and Run II(from 29 mm to 25 mm). The IBL utilizes planar sensors, similar to the Pixel Detector, and 3D sensors, allowing electrons to interact with the bulk of the sensor as opposed to just the surface, and functions as a fourth pixel layer of the Pixel Detector. The addition of the IBL significantly improves the vertexing performance of the ID. Figure 3.9 show the improvements the vertexing with the IBL

included. This improved vertexing allows for an increase in b-tagging performance, described in Section 4.2.4.

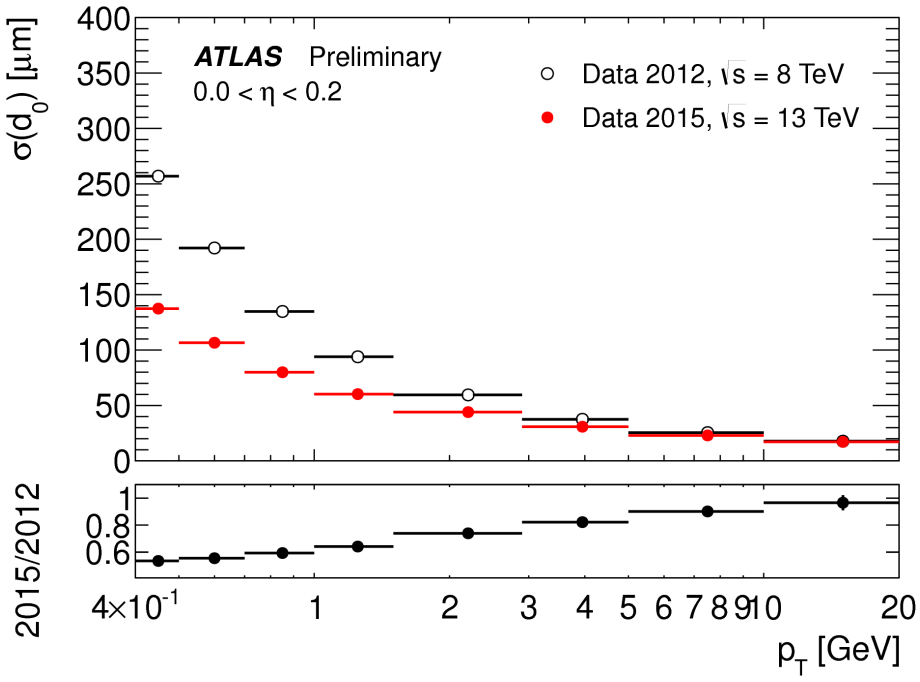

FIGURE 3.9. Unfolded transverse impact parameter resolution measured with and without the IBL as a function of pT,.[31]

The Pixel Detector is a network of high granularity, silicon pixels which measure the 2D position of passing charged particles. The silicon pixels are n-doped silicon wafers. A high voltage is applied to the wafer and when a charged particle passes through the silicon many electron hole pairs are created. The electron drifts to the electrode and creates a signal that is read out by the electronics. The Pixel detector central barrel is divided into three cylindrical layers, the innermost layer is the B-layer, followed by Layer 1 and Layer 2. Each is covered in $50\mu m$ x $400\mu m$ silicon pixels. In order to ensure complete coverage, end cap modules are placed on each side of the barrel. The end-caps consist of four wheels, each with an inner and outer ring of trapezoid shaped silicon detectors.

The SCT is made of a barrel detector and two end-cap detectors. The barrel SCT has four cylindrical layers made up square modules covered with silicon microstrip detectors. The end-cap SCT tracker is made of rings of SCT modules with either silicon or galium arsenide. These rings are arranged into nine wheels on each side of the barrel. In total, the SCT contains $61m^2$ of silicon detectors with 6.2 million readout channels.

Outside of the silicon detectors lies the TRT. The TRT is a straw detector comprised of 50,000, 4mm diameter straws in the barrel and 320,000 radial straws in the end-caps. There are 420,000 electronic channels, which give a spatial resolution of $170\mu m$ per straw. The straws are filled with various mixtures of xenon argon, carbon dioxide, tetrafluoromethane and nitrogen gas. When a charged particle passes through the radiator between the straws, made of polypropylene, they emit transition radiation photons. These photons ionize the gas in the straws and the free electrons are attracted to the positively charged wire and produce a signal that is later amplified and read out. The xenon in the gas mixture allows for accurate particle identification from the transition radiation photon detection. Transition radiation is emitted when a particle moves between two materials with different dielectric constants and is proportional to the Lorentz factor of the particle [32]. This gives a good discrimination between electrons and charged pions. This entire system is enclosed within a solenoid magnet to bend the charged particles inside the ID.

3.3.2. Calorimeters

Outside of the solenoid magnet lies the calorimetry system. The calorimeters are responsible for measuring the energy of both charged and neutral particles, with the

exception of MIPs and non-interacting particles such as neutrinos. The calorimeters can be broken into two distinct pieces, the liquid Argon calorimeter (LAr)[33] and the tile calorimeter (TileCal)[34].

The Liquid Argon (LAr) calorimeter is a sampling calorimeter that is used for electromagnetic calorimetry for the entire range of acceptance ($|\eta| < 4.8$). It is also used for hadronic calorimetry for higher pseudorapidity $1.4 < |\eta| < 4.8$. The central barrel of the calorimeter ($|\eta| < 1.4$) is made up of 1024 lead-stainless-steel converters with copper-polyimde multilayer readout boards. The plates and readouts are arranged in an "accordion-shaped" geometry. This allows for complete azimuthal coverage with no gaps, giving an electromagnetic energy resolution that is uniform in azimuth. In between the accordion layers, liquid argon is used as the active medium. The system is enclosed in a cryostat to maintain the temperature of the detector. The LAr barrel is divided radially into four sampling layers. The granularity of the layers can be found in Figure 3.10. The layer closest to the beamline is the Presampler. This layer sits inside of the cryostat and is responsible for correcting for the energy loss in front of the calorimeter (the same is done in the endcap). Inside the cryostat, there are three additional layers. The thickness of the layers is often described in terms of radiation lengths χ_0. A radiation length is the distance a electron travels before it loses approximately 1/2 of it's energy to photon emission. The front layer has a thickness of $4.3\chi_0$, followed by the middle layer with a thickness of $16\chi_0$ and the back layer of thickness $2\chi_0$. The shower maximum is contained in the second layer of the calorimeter, resulting in the bulk of the energy being absorbed in that layer. The design of the calorimeter allows for an energy resolution for electrons of $\sigma_E/E \sim 10\%/\sqrt{\frac{E}{GeV}} \bigoplus 0.7\%$ [35].

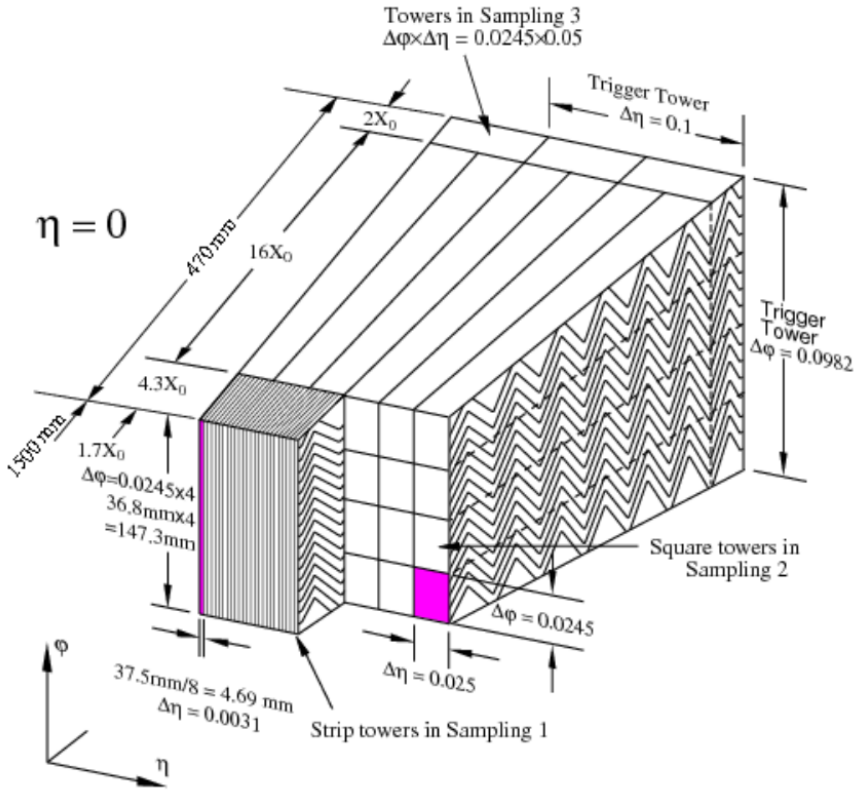

FIGURE 3.10. Sketch of the accordion structure of the EM calorimeter

Forward from the barrel, there are two electromagnetic endcap (EMEC) wheels with a similar accordion structure to the barrel. One covering $1.4 < |\eta| < 2.5$ and one from $2.5 < |\eta| < 3.2$. Outside of the EMEC is the Hadronic endcap (HEC). This is also a copper-LAr sampling calorimeter. It has a simpler parallel plate design. Finishing out the LAr calorimeter is the Forward Calorimeter (FCal), which is contained in the endcap cryostat. This calorimeter is in the very forward region of the detector. In this region, the particle flux is very high, so a dense calorimeter is necessary to avoid energy leaking into other pieces of the detector. There are three layers in the FCAL, the first is made of copper and the other two are made of tungsten. They are matrices of metal with concentric tubes filled with Argon, see

Figure 3.11

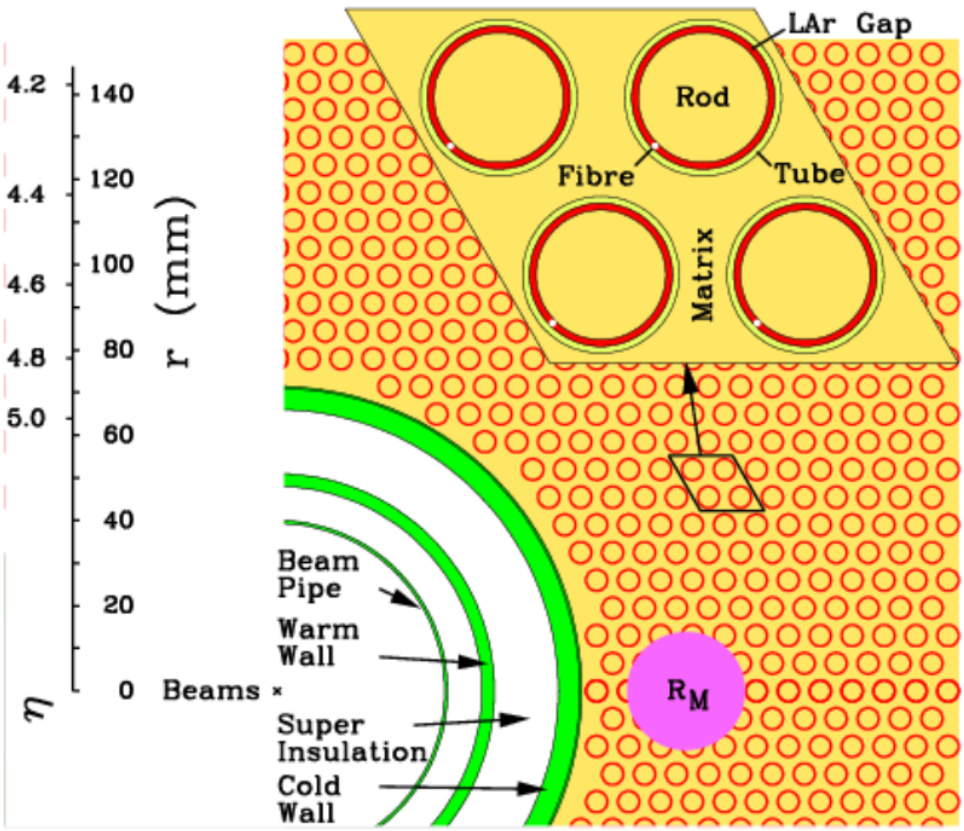

FIGURE 3.11. Sketch of the matrix and rods in the forward calorimeter

In the central region $|\eta| < 1.7$, the tile calorimeter (TileCal) is responsible for the hadronic calorimetry. The TileCal is a sampling calorimeter with alternating iron plate absorbers and plastic scintillating tiles; the orientation can be seen in Figure 3.12. The scintillating tiles are placed perpendicular to the beamline and are read out by wave-length shifting fibers on both ends of the module. The light is passed to photomultiplier tubes (PMTs) on the outside of the system, and then passed to the front-end electronics. It has a fixed central barrel and two extended barrel sections that can be moved. The TileCal has a depth of 7.4λ, where λ is the

nuclear interaction length, the mean distance a hadronic particle travels before it undergoes an inelastic interaction. The readout has a granularity of $0.1 \times 0.1(\eta \times \phi)$

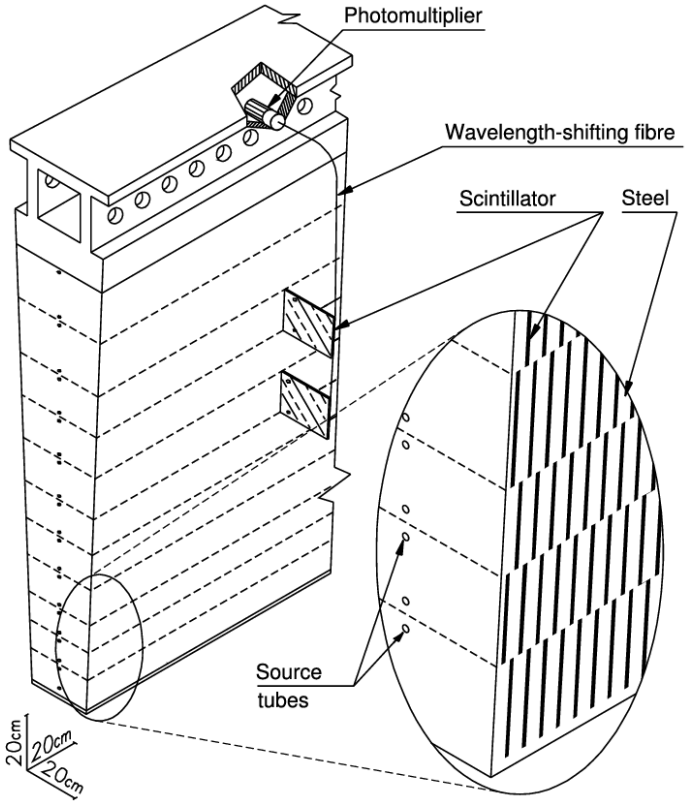

FIGURE 3.12. Schematic showing the mechanical assembly and the optical readout of the Tile Calorimeter, corresponding to a ϕ wedge. The various components of the optical readout, namely the tiles, the fibers and the photomultipliers, are shown. The trapezoidal scintillating tiles are oriented perpendicular to the colliding beam axis and are read out by fibers coupled to their non-parallel sides [36]

3.3.3. Muon Detectors

To detect muons, ATLAS uses four different technologies to measure muons with $|\eta| < 2.7$ [37]. For precision energy and position measurements, monitored drift tubes (MDT) and cathode strip chambers (CSC) are used. The CSCs are used in regions of high flux ($2.0 < |\eta| < 2.7$, where the MDTs are not suitable. For the muon

trigger system, a fast detector is needed to keep up with the timing requirements and high collision rate of the LHC. In the central region, resistive plate chambers (RPC) are used, while in the forward region, where flux is higher, thin gap chambers (TGC) are used. The muon system, much like the ID utilize a magnetic field to determine the charge and momentum of passing particles. The magnet system is further discussed in Section 3.3.4.

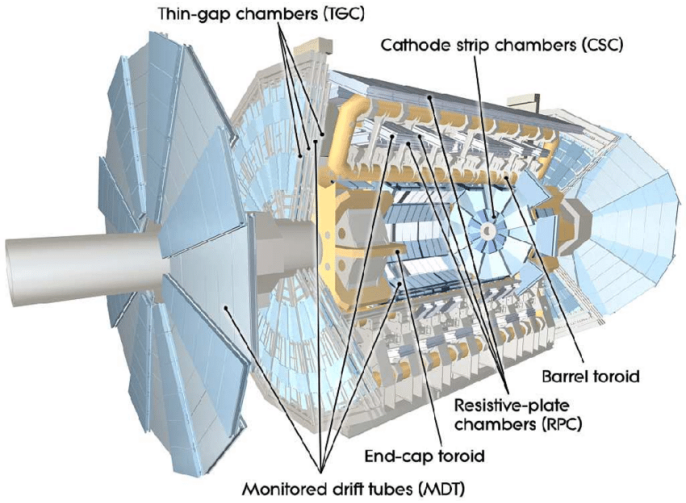

FIGURE 3.13. Cut-away view of the ATLAS muon system [38]

The MDTs are made up of six parallel layers of cylindrical aluminum drift tubes with a tungsten-rhenium wires. The drift tubes are filled with a mixture of argon, nitrogen and methane. The tubes are assembled on a support spacer and are monitored for deformation by a built-in optical system, hence the **monitored** drift tubes. The monitoring ensures a high accuracy in the position of the measurement points. Allowing the MDTs to achieve a sagitta precision of $50\mu m$ and thus a momentum precision at 1 TeV of $\frac{\Delta p_T}{p_T} = 10\%$

While the MDTs are very good at precision measurements. However, they

are not appropriate in areas with high rate counts ($> 200Hz/cm^2$) due to their large diameter and high operating pressure. This is the case for the first layer of muon measurement with pseudorapidities of $|\eta| > 2.0$. For this region, CSCs are the detector of choice. CSCs are multiwire proportional chambers with a cathode strip readout. They gives good single and two track resolution in this high rate region.

In the barrel region, the muon trigger system employs RPCs, a low occupancy chamber with fast response. RPCs are gaseous parallel-plate detectors of Bakelite with a coating of linseed oil based paint. The system can operate in two modes, avalanche and streamer. In streamer mode, a large potential across the plates generates a discharge around the ionizing particle. For avalanche mode, a smaller potential difference and large signal amplification in the electronics allows for increased rate capability.

Finally, in the end-cap of ATLAS, TGCs provide two important components. For the trigger system, TGCs have good timing resolution compared to the MDTs and can deal with a rate of up to 100 kHz/cm^2. For measurement, TGCs provide the azimuthal coordinate to compliment the bending coordinate from the MDTs. The TGCs are made up of anode wires and graphite cathodes in between layers of fiberglass laminate.

3.3.4. Magnet System

The signature piece of the ATLAS detector is the Large Toroid magnet system. The toroid has eight coils in the barrel and two endcaps, with eight coils each (Figure 3.14) [39]. The Toroid system provides a magentic field of 3.9 T(4.1 T) in the barrel (end-cap) to the muon system. The coils of the three toroids are assembled radially and symmetriclly around the beam axis. A toroid has two advantages over a solenoid. The first is the field at the edges of the detector remains perpendicular

to the outgoing particles, allowing for a better measurement at high pseudorapidity. The other advantage is cost. It takes much less material to build a large toroid than an equivalently sized solenoid. This allows ATLAS to have a very large volume for particle bending in the muon system.

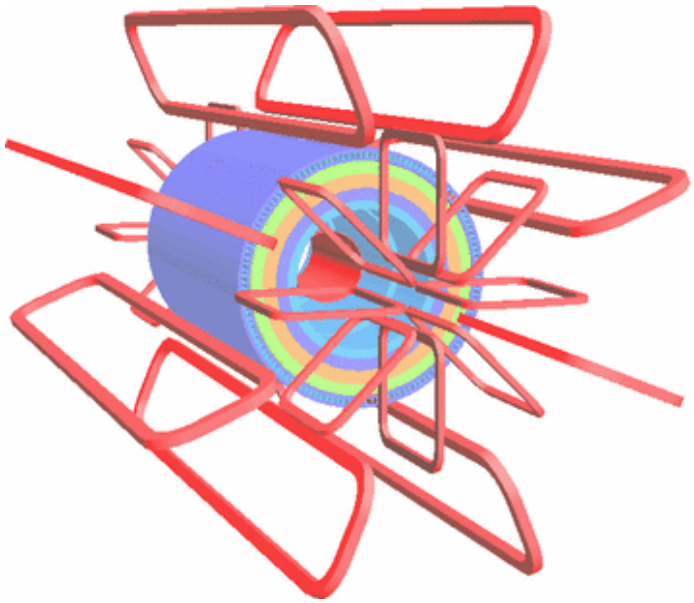

FIGURE 3.14. Geometry of magnet windings and tile calorimeter steel [38]

Along with the toroid, ATLAS has a solenoid magnet inside of the calorimeter. This solenoid provides a 2 T magnetic field to the ID for bending of charged particles. The solenoid is a single layer coil in a suporting cylinder. It is supported by the LAr cryostat. It is very important the solenoid is thin, in order to minimize the amount of material in front of the calorimeters. To achieve this, the vacuum of the solenoid and LAr are combined into one and the coil is designed to be as thin as possible.

3.3.5. Trigger System

The LHC delivers collisions at a rate of 40MHz. With each raw event being about 1.6MB [40], this would give an output rate of 64TB/s. This rate of output is beyond what can be handled by the computing resources available. In order to reduce this rate to a manageable level, ATLAS employs a two level trigger system. The first level trigger is a hardware based trigger, referred to as the L1 trigger. This trigger lowers the rate to between 75 kHz and 100 kHz. This is sent to the second level, software based, trigger, the High Level Trigger (HLT), where the rate is further reduced to below 2kHz for full event readout. When combined with the partial event readout, the total bandwidth around 3GB/s Fig 3.15 illustrates the ATLAS trigger system data flow.

The L1 trigger begins with signals from either the calorimeters or the muon detectors. A signal from the calorimeter is sent to the Level-1 Calo (L1 Calo) system. The L1 Calo system uses low granularity calorimeter information to identify Regions of Interest (RoIs) for electrons, photons, taus, jets, as well as high total energy and missing transverse energy ($\not{E}_T$). L1 Calo received an upgrade in Run II in the form of a new Multi-Chip Module (nMCM). This module allows for L1 Calo to suppress the effects of pile-up on the system.

Signals from the muon system are fed into the L1 Muon system. L1 Muon uses information from the RPC and TGC in the barrel and end-caps. In Run II, the muon end-cap triggers required a coincident hit in the innermost muon chamber to reduce the fake muon rate in the forward regions.

The signals from L1 Calo and L1 Muon are passed to the L1 Topological trigger processor (L1 Topo) and the Central Trigger Processor (CTP) simultaneously. In the

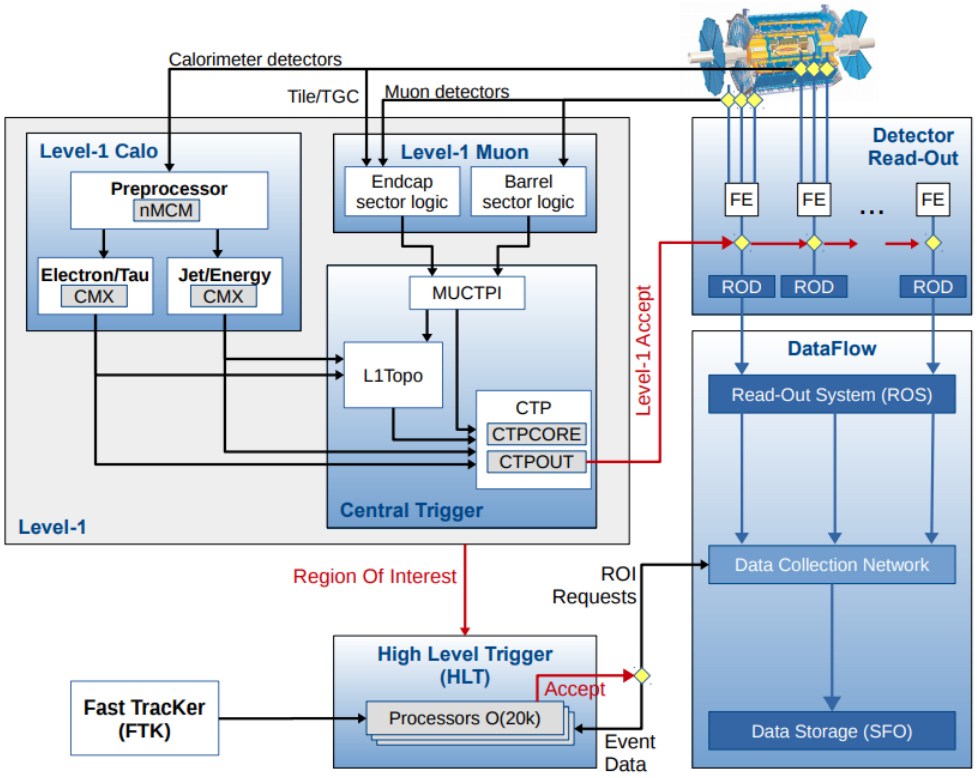

FIGURE 3.15. Schematic layout of the ATLAS trigger and data acquisition system in Run-2.[41]

case of L1 Muon, the signals from the barrel and the end-cap are merged in the muon Central Trigger Processor Interface (MuCTPi). In L1 Topo, a new system for Run II, kinematic information from L1 Calo and L1 Muon are used to make topological selections, like angular separation, invariant mass requirements, and total $\not{E}_T$, at level 1.

Information from L1 Topo, and directly from L1 Calo and L1 Muon, are sent to the CTP where the L1 trigger Accept and the LHC timing information are provided to the sub-detector readout system. The CTP also sends the RoIs to the HLT to use as inputs for higher level algorithms.

The HLT is made up of 40,000 processing cores which run around 2,500 independent trigger chains. A chain is an offline-like algorithm run over the RoIs

from L1. Partial event reconstruction, and even full event reconstruction can be done within the HLT depending on the event stream. There are four types of data streams depending on the purpose of the data, physics analysis, trigger level analysis, monitoring, or detector calibration. For the physics analysis stream, the full event is written out. For the other streams, only partial event information is written out, saving bandwidth to allow for higher trigger rates.

The configuration of the trigger system is defined by the trigger menu. The trigger menu is made from the primary physics analysis triggers along with support triggers for efficiency measurements and monitoring, alternative triggers, backup triggers and calibration triggers [42]. These triggers, L1 items and HLT triggers, are regulated by prescale factors. A prescale factor is used to reduce the rate of selected events. For a prescale of N, only 1/N events that meet the trigger requirement are accepted. Prescales can be set to individual L1 items and to specific HLT chains independently and as a function of luminosity. This allows triggers to maintain a specific rate regardless of the luminosity. Typically, the L1 primary physics triggers are unprescaled, this means, the other triggers are prescaled to a low rate to give priority to the physics triggers. As a result, most of the bandwidth is occupied by the physics chain.

CHAPTER IV

SIMULATION AND EVENT RECONSTRUCTION

4.1. Simulation

In order to draw conclusion from ATLAS data, it is necessary to compare to theoretical predictions. For particle collisions, it is not practical to create exact predictions, especially including detector effects such as resolution. To get the best estimate of these effects, ATLAS uses the Monte Carlo (MC) method to simulate data and detector response to the incident particles. This is done in multiple steps as illustrated by figure 4.1. These steps are the simulation of the hard process, where the deep inelastic collision simulated using the initial state (Parton Distribution Functions) and interaction amplitudes; the parton shower; the hadronization; the detector simulation; and finally the reconstruction. These steps together form the complete MC simulation of ATLAS data.

4.1.1. Parton Distribution Functions

At the energies at the LHC, collisions usually do not involve entire protons. Instead, they involve constituents known as partons. Protons, while often described as two up quarks and a down quark, also contain a sea of gluons. This sea of gluons creates many virtual quark-antiquark pairs known as sea quarks. The valance quarks for the proton are up and down quarks. These valance quarks are the primary role players in lower energy inelastic interactions. At the LHC, the collision energies are sufficient for deep inelastic scattering, where the affects of the sea quarks and gluons are non-trivial. This internal structure of the proton is described by a Parton

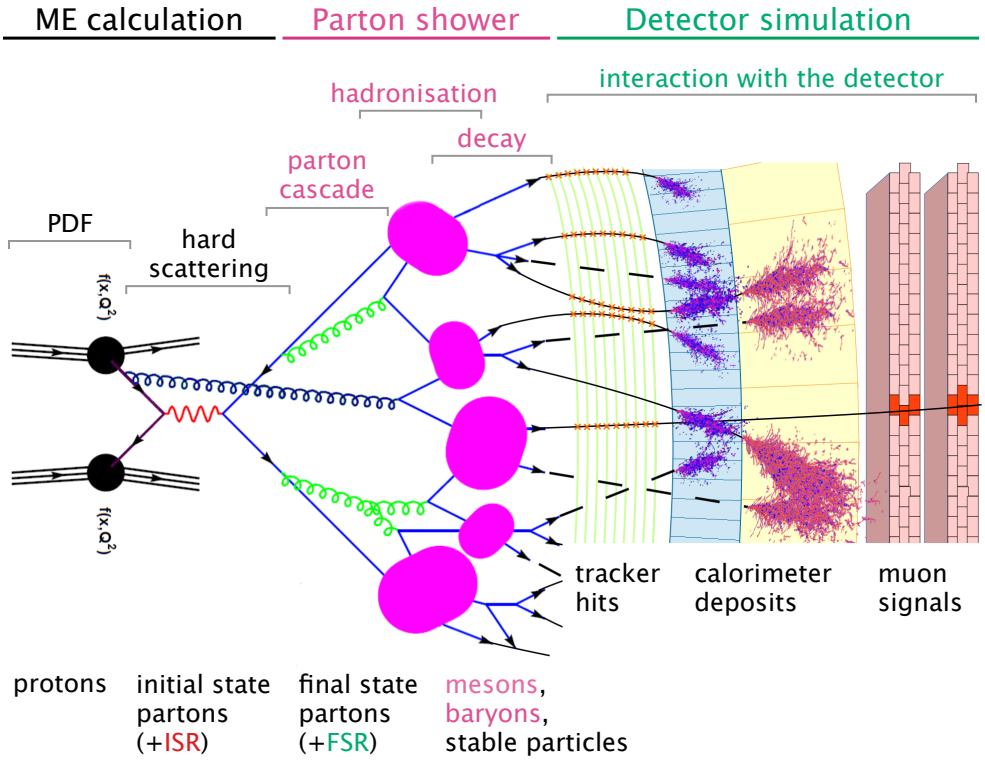

FIGURE 4.1. Pictorial representation of how an event is generated [43]

Distribution Function (PDF), figure 4.2. A PDF shows the probability density of finding a parton carrying a momentum fraction x at a squared energy scale.

4.1.2. Hard Scattering

The hard scattering process can be described using Feynman diagrams. These diagrams are a pictorial representation of amplitudes. These amplitudes go into calculating the matrix elements (ME) of various interactions. This ME describes the probability of a certain interaction occurring. In the event generation, these MEs are calculated to a specified order in perturbation theory. Common examples are

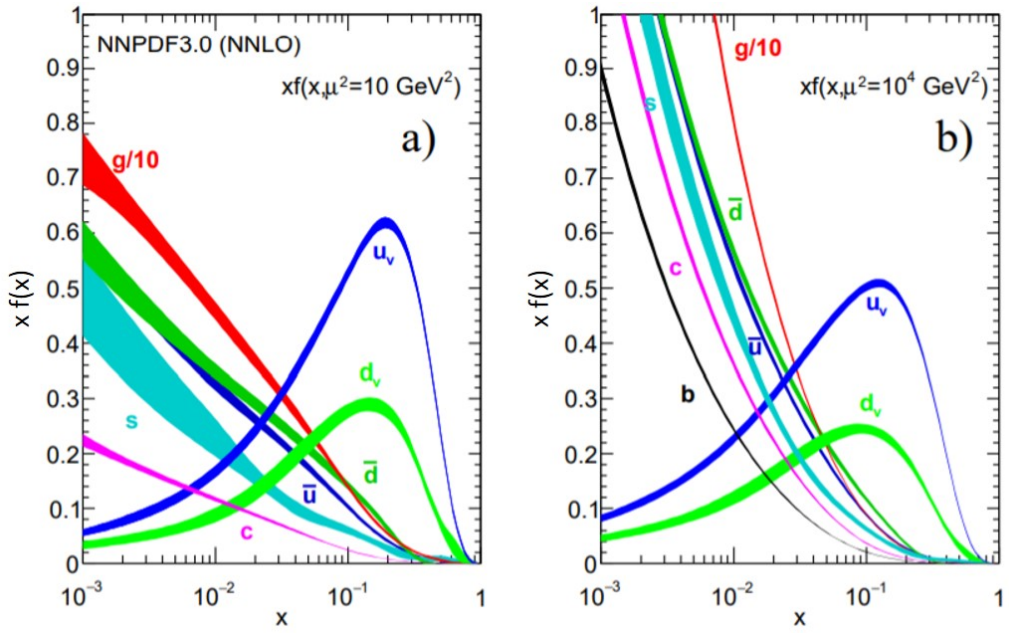

FIGURE 4.2. The bands are x times the unpolarized parton distributions $f(x)$ (where $f = u_v, d_v, \bar{u}, \bar{d}, s \simeq \bar{s}, c = \bar{c}, b = \bar{b}, g$) obtained in NNLO NNPDF3.0 global analysis at scales $\mu^2 = 10$ GeV^2 (a) and $\mu^2 = 100$ GeV^2 (b), with $\alpha_s(M_Z^2) = 0.118$.

leading order (LO), next-to-leading order (NLO), and so on. The higher the order of the calculation, the more accurate the predictions. However, higher orders can be extremely hard to theoretically calculate, often restricting the level of the event generator.

4.1.3. Parton Shower Calculation

After the ME generator, the hard partons are used as the inputs to the Parton Shower (PS) calculation. A parton shower is the evolution from the quarks and gluons produced in the hard interaction to the final state hadrons and other particles seen in the detector through QCD processes. The PS calculation models this showering process.

In an interaction, colored particles can spontaneously emit gluons. These gluons,

in turn, create either more gluons, or quark-antiquark pairs. This can happen either before (ISR) or after (FSR) the hard scattering process. The PS generator can also describes the hadronization and subsequent decay of the hadrons into the final state particles.

The precision of the PS generators are described similarly to the ME, with their contributions coming in as leading log (LL), next-to-leading-log (NLL), etc. for the parton showering process.

4.1.4. Detector Simulation

The MC simulation up to this point can be done with generators that are written outside of the ATLAS collaboration. These generators are used by ATLAS to simulate the underlying processes which are fed into the detector simulation software. ATLAS uses GEANT4 to handle this propagation[44] through the detetor. GEANT4 uses a detailed geometric description of the ATLAS detector to simulate particle interactions with the detector material. This includes simulation of energy deposition and the readout process.

The final result of the MC event generation is a set of simulated data that resembles actual data from the p-p collisions in the ATLAS detector.

4.1.5. Reconstruction

Once the data has been simulated, it is necessary to transform it into meaningful objects through reconstruction. There are two main types of reconstruction in ATLAS: turning patterns of hits in tracking detectors (Inner Detector and Muon Spectrometer) into tracks with direction and momentum information, and turning energy deposits in the calorimeter into calibrated energy deposits. These objects are

then used to build a picture of the physics event through particle identification and event reconstruction.

4.2. Particle Identification

For all events, either MC or actual collision data, it is important to be able to identify and reconstruct the underlying physics event. In particle collisions, the energy from the final state particles is deposited in the various subdetectors within ATLAS. These energy deposits must be translated to physically meaningful objects. This is the task of the event reconstruction, to use the ATLAS detector to recreate the final state particles for any given interaction. For this analysis, the final state particles present in the signal events are a lepton, either an electron or a muon; a neutrino, in the from of missing transverse energy; two light flavor quarks; and two b quarks. Each of these particles has a particular signal in each of the subdetectors, figure 4.3.

4.2.1. Electrons

Electrons are reconstructed by fitting a track using the Inner Detector and matching this track to an energy cluster in the EM calorimeter[46]. As an electron passes through the EM calorimeter, it produces Bremsstahlung radiation photons. These photons then convert back to electron-positron pairs and the process repeats. This shower of electrons, positrons, and photons give the signature energy cluster in the calorimeter. Particles with the required Inner Detector track and matching EM energy cluster are selected as electron candidates.

Electron identification algorithms are applied to these electron candidates. These algorithms separate prompt, isolated electron candidates from backgrounds such as converted photons and misidentified jets. The electron identification

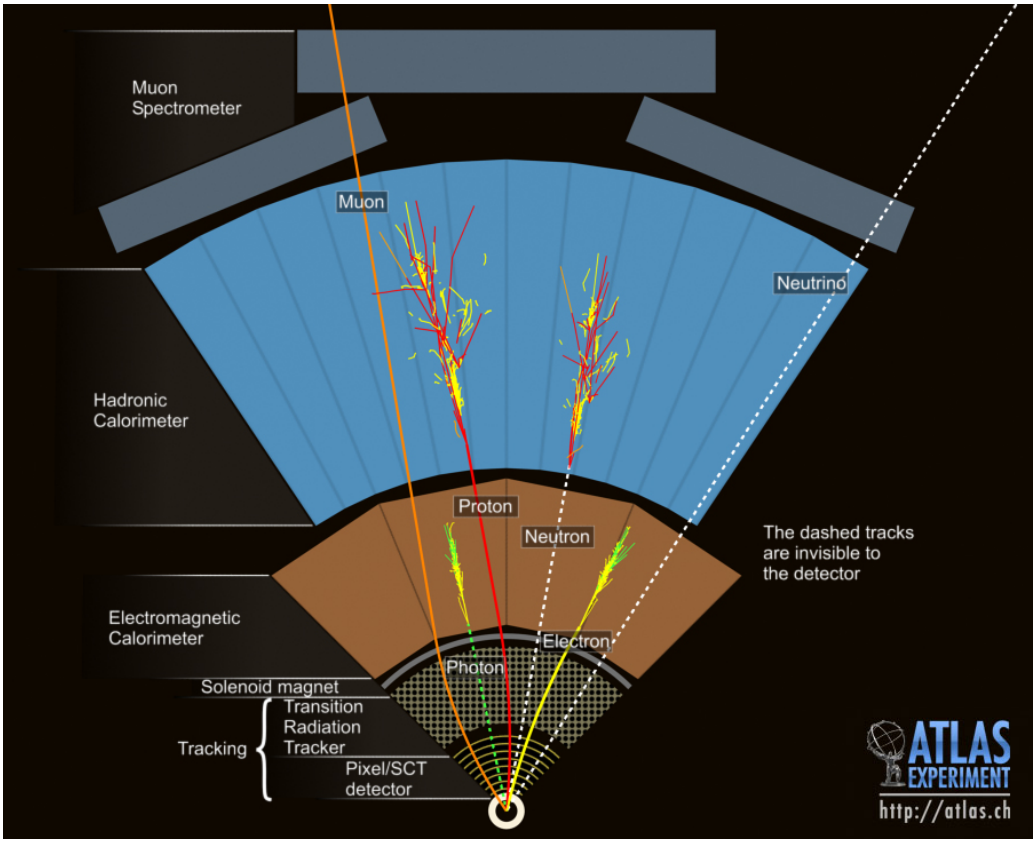

FIGURE 4.3. Event Cross Section in a computer generated image of the ATLAS detector [45]

algorithm uses the energy-momentum ratio, shower shape, track and track-to-cluster matching to identify electron candidates, with E/p being the most important discriminant. There are three identification working points for electron identification: Loose, Medium, and Tight. The operating points with higher background rejection are a subset of electron candidates with lower background rejection with a tighter background rejection giving a lower electron efficiency.

The isolation variables quantify the energy around the electron candidate and allow us to disentangle prompt electrons from other, non-isolated electron candidates such as electrons originating from converted photons produced in hadron decays, electrons from heavy flavor hadron decays, and light hadrons mis-identified as

electrons. The isolation variable we use for reconstructed electrons is *Track*-based isolation, $p_T^{\text{varcone0.2}}$, defined as the sum of transverse momenta of all tracks, satisfying quality requirements, within a cone around the electron candidate of $\Delta R = 0.2$ or of 10 GeV/E_T for high energy electrons, where E_T is the transverse energy of the electron candidate.

A more detailed discussion on the electron likelihood identification and isolation variables and their performance with Run 2 data can be found in Ref. [47]. The electron energy scale is calibrated such that it is uniform throughout the detector and the residual differences between data and simulation are corrected. The calibration strategy is based on the same strategy developed in Run 1 [48] and updates to the calibration strategy for Run 2 is documented in Ref. [49].

4.2.2. Muons

The Muon Spectrometer (MS) specializes in muon detection and precision momentum measurements. Unsurprisingly, this makes the MS a vital part of muon identification, but it is not the only subdetector used. The Inner Detector is also has an important part in reconstructing muons. In ATLAS, muon reconstruction is performed independently in the Inner Detector and the MS. The information is then combined to form muon tracks. In the Inner Detector, the muons are reconstructed similarly to any other charged particle.

In the MS, the reconstruction looks for a hit pattern within each chamber to form segments [50]. The MDT segments are combined using a straight-line fit within a single layer. Segments in the CSCs are combined using a combinatorial search in the η and ϕ planes.

Muon candidates are built by fitting together hits from segments in different layers. A combinatorial search, using segments in the middle layer as seeds, is

performed. The inner and outer layers are then used as seeds as the search is extended. A minimum of two segments are required to build a track. It is possible for a segment to be included in multiple tracks, an overlap removal algorithm selects the best assigned track or can allow for a segment to be shared between two tracks. A global χ^2 fit is performed on the hits of each track. If the χ^2 of the fit passes a selection criteria, the track is accepted.

The information from the Inner Detector and the MS are then combined to give a muon signature. The combination method depends on the information available. The main method used is the Combined Muon reconstruction, where track reconstruction is performed in the Inner Detector and MS independently. Most of these muons are reconstructed using an "outside-in" reconstruction. This means tracks in the MS are extrapolated inward and matched to an Inner Detector track.

The muon isolation variables are similar to the electron isolation variables above which is the *track*-based isolation, $p_T^{\text{varcone0.3}}$, defined as the sum of transverse momenta of all tracks, satisfying quality requirements, within a cone of $\Delta R = \min(0.3, 10 \text{ GeV}/p_T)$ around the candidate muon.

The performance of the muon identification and isolation variables are documented in Ref. [50].

Corrections to the muon momentum scale and resolution are applied to MC simulation using the MuonCalibrationAndSmearingTool[1] to correct for data/MC differences. The correction factors were derived from data/MC simulation comparisons with $Z \to \mu\mu$ and $J/\Psi \to \mu\mu$ events (the calibration procedure to derive the factors is documented in Ref. [50]).

[1] as prescribed in MCPAnalysisGuidelinesMC15 twiki

4.2.3. Jets

Due to the strong force, quarks very quickly undergo showering, a process of emitting gluons which then further produce quark-antiquark pairs and the process repeats. Once the average energy of a quark (or gluon) reaches 1 GeV, the particles hadronize, producing quark bound states such as pions and mesons that ultimately deposit energy in the ATLAS detector. The top quark is the only quark that decays before hadronizing. If we could measure every hadron and correctly assign them to the underlying quarks, energy and momentum conservation would allow the exact momentum and energy of the quark could be determined. Jet finding is an algorithm that tries to do this. Collections of hadrons deposit energy in the ATLAS detector. These collections of energy are called jets and can be made from various detector object. In this analysis in particular, two different types of jets are used: calo-jets, jets constructed from energy deposited in the calorimeters; and track-jets, jets constructed from tracks in the Inner Detector. To form calo-jets, ATLAS used topological clusters(topo-clusters) [51]. Topo-clusters are three dimensional collections of topologically connected energy deposits. These topo-clusters are formed by a growing-volume algorithm starting from a high significance seed signal. Topo-clustering suppresses noise by removing cells with insignificant signals, making it superior to basic energy summing.

Since a jet is not a physical object, rather a collection of energy-momentum 4-vectors, there are many ways to define a jet. A jet algorithm takes a set of input 4-vectors and combines them into one or more jet objects based upon some criteria for separating and grouping inputs. Jets can be made with either energy deposits (calo-jets) or tracks (track-jets). The process of creating calo-jets is described here

but the method for making track-jets is similar. Two important characteristics of any jet algorithm are Infrared (IR) Safety and Collinear (CL) Safety. For a jet algorithm to be IR Safe, the addition or subtraction of small energy deposits will not change the jet collection. A jet algorithm is CL Safe if splitting or merging high transverse momentum particles does not change the jet collection. Figure 4.4 illustrates both IR and CL Safety.

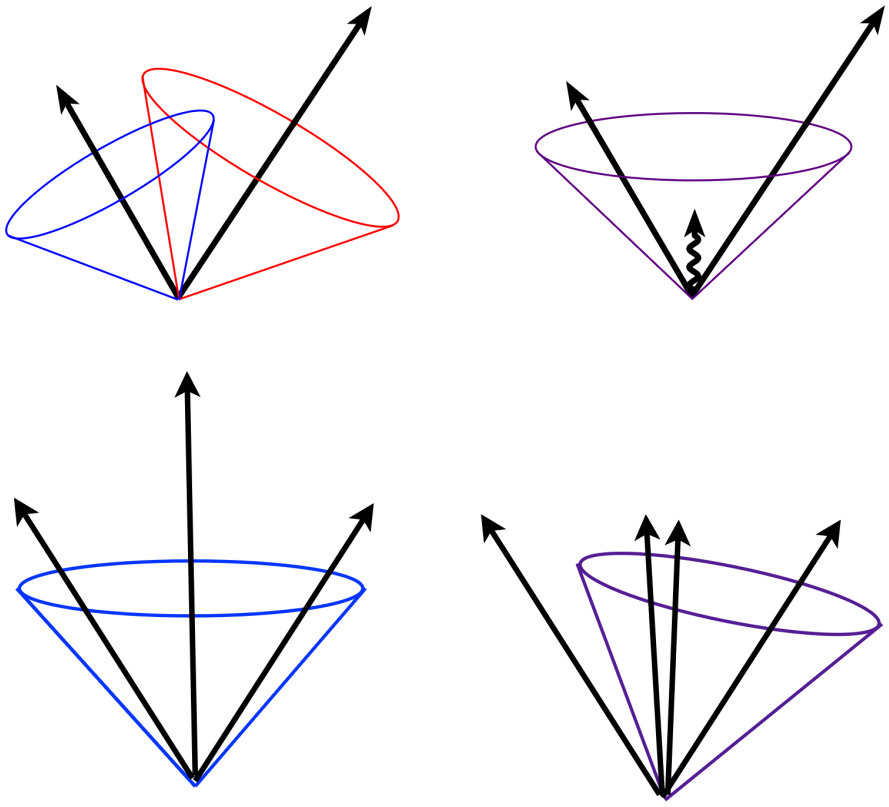

FIGURE 4.4. Illustration of the infrared sensitivity of a cursory designed jet algorithm (top). Illustration of the product of a collinear unsafe jet algorithm. A collinear splitting changes the number of jets (bottom). [52].

Some examples of jet algorithms are visualized in figure 4.5. For this analysis, the anti-k_t algorithm is selected. In addition to being IR and CL safe, the anti-k_t algorithm gives roughly circular jets. This makes calculating the energy density

much easier than non-circular jets and thus easier to calibrate the jets. The anti-k_t algorithm calculates the distance between objects i and j (d_{ij}) and i and the beam B (d_{iB}). If d_{ij} is smaller than d_{iB}, the objects are combined. If d_{iB} is smaller the object is removed and the algorithm is rerun. An important distinction between anti-k_t and other jet algorithms is the definition of the distances d_{ij} and d_{iB}

$$\begin{aligned} d_{ij} = \ \min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2}, \\ d_{iB} = k_{ti}^{2p} \end{aligned} \tag{4.1}$$

where k_{ti} is the transverse momentum, Δ is the distance between objects, and $p = -1$. The anti-k_t algorithm has a radius parameter R. R acts as a cutoff radius for energy clustering and is not strictly a radius, as objects with a $\Delta > R$ can still be clustered together. The track-jets used in the analysis have R = 0.2, while the R = 0.4 (small-R) and R = 1.0 (large-R) calo-jets are used.

4.2.3.1. Large-R jets

For decays with a high momentum to rest-mass ratio, such as the $W \rightarrow qq$ decay, it is impossible to separate energies cleanly into jets with R = 0.4. Instead, to measure the energy/momentum of the W it is advantageous to use a larger radius paramenter. The large-R jets are clustered using the anti$-k_t$ jet algorithm [53] with topological calorimeter clusters as inputs. The clusters are calibrated to the "local hadronic cell weighting"(LCW) scale [54]. In order to minimize the effects from pileup on the large-R jet kinematics, the large-R jet is then groomed using the trimming algorithm. The trimming algorithm removes subjets if the ratio of the subjet p_T over the large-R jet p_T is below some threshold[55]. This removes energy from pileup that is contained in the jet. The large-R jet energy and mass is then

calibrated to the particle-level scale. The calibration factors were derived from MC simulation of multijet events [56].

4.2.3.2. Track jets

Track jets are built by clustering Inner Detector tracks using the anti-k_t algorithm with a radius parameter R = 0.2. The selected tracks are required to have p_T greater than 400 MeV and pass a loose set of cuts, as listed in reference [57]. The smaller R parameter coupled with the fact that tracks have better angular resolution than calorimeter clusters, mean that the decay products of highly boosted heavy objects can still be resolved. The selected track jets are then associated to the large-R calorimeter jets via ghost association [58] method. A b-tagging algorithm is used to identify track jets which are likely to contain b-hadrons which consist of the b-quarks from the Higgs boson decay. The MV2c10 algorithm exploit the relatively long lifetime of B-hadrons with respect to lighter hadrons, as well as the kinematics of the charged particle tracks.

4.2.3.3. Small-R jets

Small-R jets are reconstructed from three-dimensional topological calorimeter clusters [54] using the anti-k_t jet algorithm [53] with a radius parameter of 0.4. This is the standard jet used in most ATLAS analyses. Jet energies are corrected [59] for detector inhomogeneities, the non-compensating nature of the calorimeter, and the impact of multiple overlapping pp interactions. Correction factors are derived using test beam, cosmic ray, pp collision data, and a detailed GEANT4 detector simulation.

Jet cleaning is applied to remove events with jets built from noisy calorimeter cells or non-collision backgrounds, requiring that jets are not of "bad" quality.[2]

To avoid selecting jets originating from pile-up interactions a "jet vertex tagger" (JVT) criterion [60] is applied for jets with $p_T < 60$ GeV and $|\eta| < 2.5$ requiring a JVT > 0.59 cut. This cut corresponds to the `Default` working point, as described on the JVTCalibration twiki.

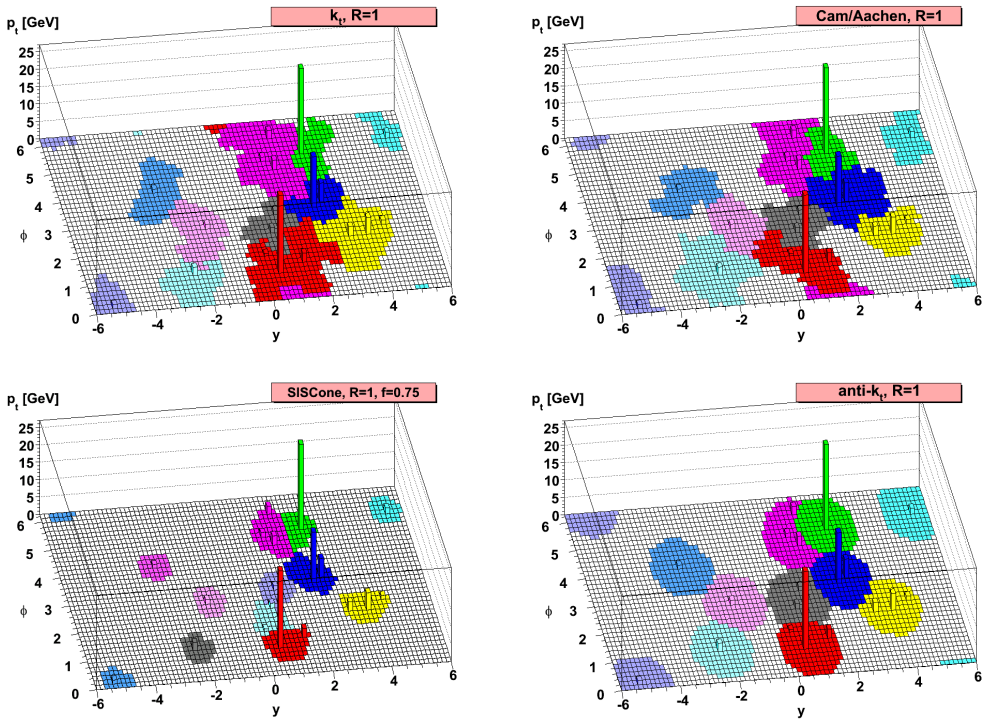

FIGURE 4.5. A sample parton-level event, together with many random soft "ghosts", clustered with four different jets algorithms, illustrating the "active" catchment areas of the resulting hard jets[61].

[2] *LooseBad* jets, defined on the HowToCleanJets2016 twiki, are removed.

4.2.4. b Tagging

Jets that originate from b-quarks have unique properties that allow them to be distinguished from other jets. When a b-quark is produced, it hadronizes into B-Hadrons. These hadrons have a relatively long lifetime compared to many other hadrons that are produced. This long lifetime, combined with the relativistic speeds the hadron is traveling allow the particle to travel a measurable distance before it decays. Figure 4.6 illustrates a b-hadron decay. To tag a jet as a b-jet, a jet that came from a b-quark, ATLAS relies on the tracks from the inner detector. By reconstructing the tracks in an event, it is possible to find a place, outside of the interaction point, where multiple tracks originate. This is referred to as the secondary vertex. Additionally, it is possible to backtrack the tracks in the displaced vertex to measure their impact parameter, or the minimum distance between the track and the interaction point. Lastly, a decay chain MVA attempts to fully reconstruct the decay chain of the jet. These three methods; secondary vertex identification, impact parameter measurements, and decay reconstruction; are all used to identify jets coming from b-quarks and reject those coming from light flavor quarks[62].

In this analysis, the MV2c10 is used to tag b-jets [62]. MV2 is a multivariate discriminant that combines the b-tagging algorithms described above. The c10 signifies a 10% c-jet fraction in the background training sample. The three algorithms that are used as inputs to the MV2 discriminant are: an impact parameter-based algorithm, an inclusive secondary vertex reconstruction algorithm, and a decay chain multi-vertex reconstruction algorithm. For this analysis, the 85% efficiency fixed-cut working point is used for b-jet identification. This corresponds to a c-jet rejection rate of 3.1 and a light-jet rejection rate of 33, Figure 4.7.

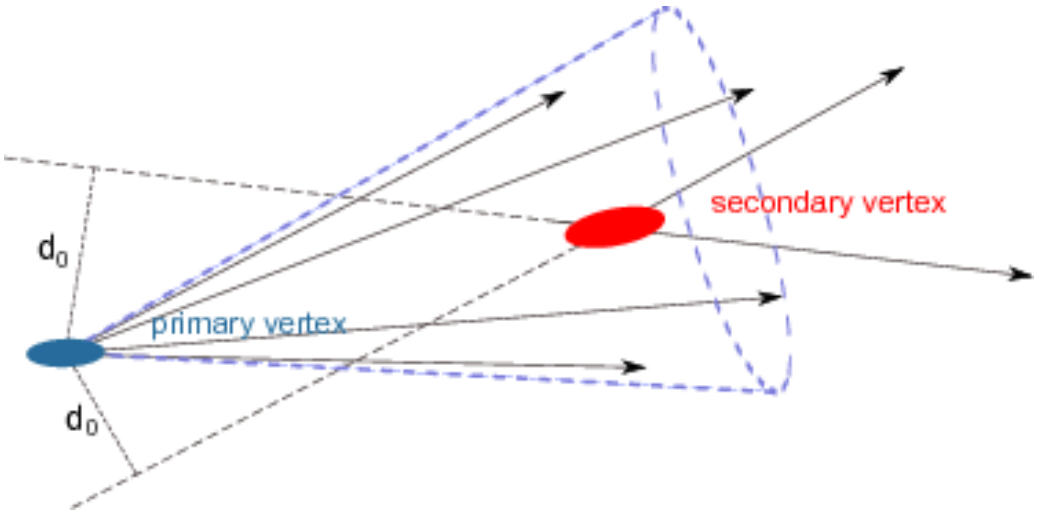

FIGURE 4.6. Schmatic view of the tracks in a b-jet [63].

The difference in the efficiency of b-tagging between data and simulation is taken into account by applying scale factors provided by the Flavour Tagging CP group, as prescribed on the BTagCalib2015 twiki. The uncertainties associated with b-tagging are considered for b-, c- and light-flavor-induced jets, separately.

4.2.5. Missing Transverse Momentum

Neutrinos do not interact with the detector as they pass through. This means they cannot be measured like the other particles. In order to measure neutrinos, ATLAS relies on the conservation of momentum. As previously mentioned, the exact collision energy is unknown, as each partons does not carry a consistent fraction of the proton energy. However, in the transverse plane, the plane perpendicular to the beam line, the total momentum is known to be very small. Before the collision, there is very little momentum in the transverse plane, on the order of 1 GeV. After the collision, this must also be true. This implies the vector summation of all objects should have approximately zero momentum in the transverse plane. Any imbalance

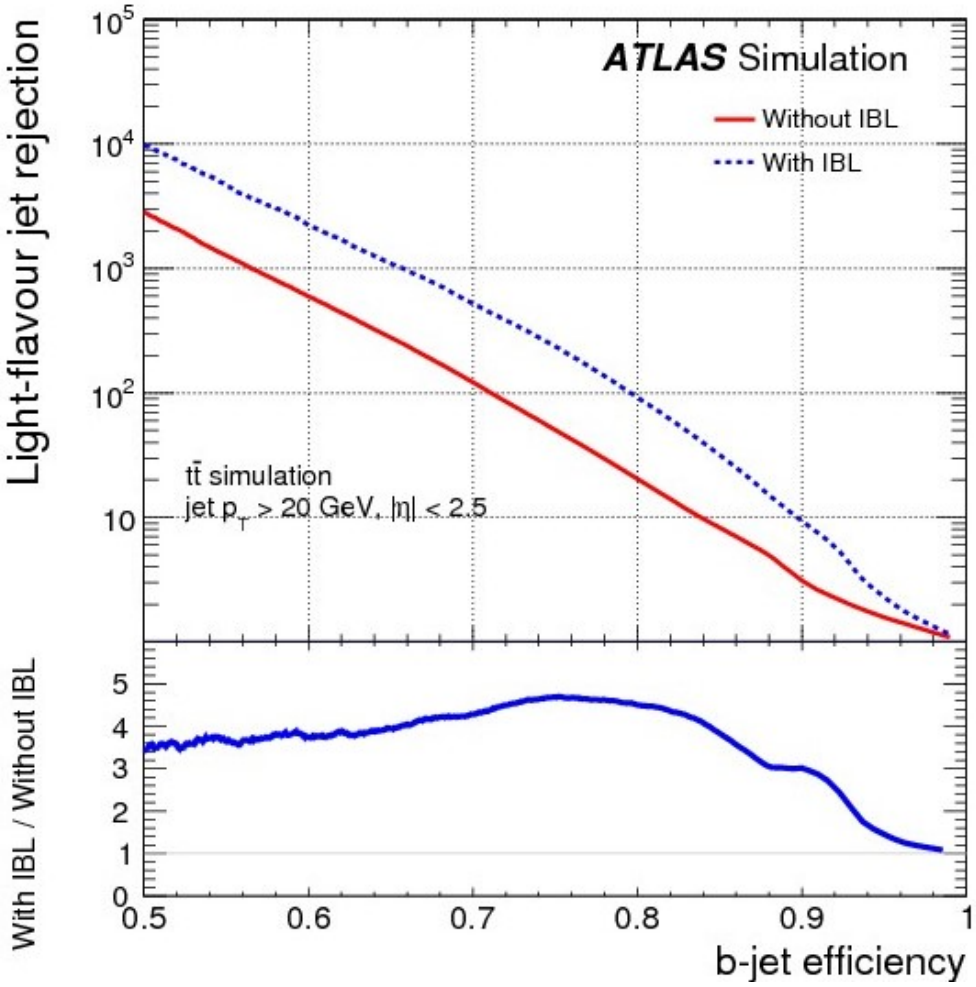

FIGURE 4.7. Comparison of light-jet rejection as a function of b-jet tagging efficiency for the Run 1 and Run 2 detector layouts.[64]

in this momentum is referred to as Missing Transverse Momentum ($\not{E}_T$). The energy symbol is used, however we really mean the magnitude of the vector sum. The $\not{E}_T$ is constructed as the negative vector sum of all reconstructed objects with an additional soft term reconstructed from detector signal objects not associated with any object[65].

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss,e}} + E_{x(y)}^{\text{miss},\gamma} + E_{x(y)}^{\text{miss},\tau} + E_{x(y)}^{\text{miss,jets}} + E_{x(y)}^{\text{miss},\mu} + E_{x(y)}^{\text{miss,soft}} \tag{4.2}$$

From the x and y components of E^{miss}, the magnitude and azimuthal angle are calculated.

$$\begin{aligned} \not{E}_T &= \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}, \\ \phi^{\text{miss}} &= \arctan(E_y^{\text{miss}}/E_x^{\text{miss}}) \end{aligned} \tag{4.3}$$

In this analysis, the $\not{E}_T$ is reconstructed using VHLooseElectrons, VHLooseMuons the analysis jets, and the track-based soft term.

The $\not{E}_T$ vector is a vector in the transverse plane, meaning it does not directly correspond to a neutrino. Additional information is needed to reconstruct a neutrino. In this analysis, a Higgs mass constraint is used to supply the direction of the signal neutrino.

CHAPTER V

ANALYSIS

This chapter will present the results of a search for Higgs boson pair production published in $JHEP$ [66]. This chapter contains material coauthored with the ATLAS Collaboration. I developed the framework for the analysis, optimized the signal regions and developed the method for estimating QCD background. I was the primary contributor to the kinematic figures. Other members of the analysis group, members of the ATLAS Collaboration, estimated the other backgrounds that were used to produce the final result presented in this chapter.

In this analysis, one Higgs boson decays via $H \to b\bar{b}$ and the other via $H \to WW^*$. The WW^* system decays into $l\nu qq$ (where l is either an electron or a muon). There is a contamination from the leptonic τ decays but it is small and not explicitly vetoed in the analysis. The Higgs boson decay modes chosen for this analysis are a compromise between signal efficiency and background rejection. The $H \to WW^*$ branching ratio of approximately 25% is the second largest after $H \to b\bar{b}$ (approximately 58%), 5.1.

The final state contains two b-quarks consistent with coming from one H, two light jets, an identified electron or muon plus $\not{E}_T$, consistent with a WW decay. The 1-lepton final state gives a strong discriminator against multijet background. The dominant backgrounds are $t\bar{t}$ production, which has the same final state but with different kinematic properties; W bosons produced in association with jets (W+jets), where two of the associated jets come from b-quarks; and multijet events where a jet is misidentified as a lepton. There are smaller background contributions from single top-quark production, Z bosons produced in association with jets (Z+jets),

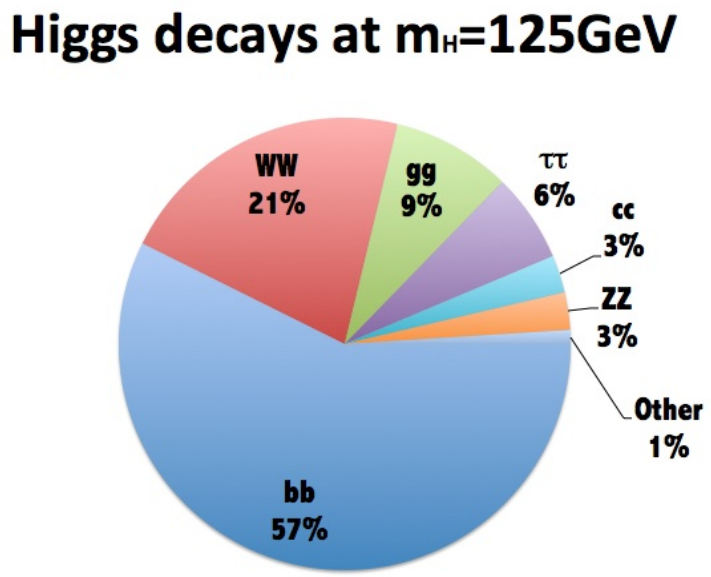

FIGURE 5.1. Illustration of Higgs boson branching ratios[67]

and diboson production.

This analysis sets limits on both SM Higgs boson pair production and on resonant production. Both production methods are discussed in detail in chapter II. Figure 5.2 shows a Feynman diagram of resonant production of the Higgs boson pairs with the subsequent decays $H \to WW^*$ and $H \to b\bar{b}$.

5.1. Analysis Overview

Two complementary techniques are used to reconstruct the Higgs boson candidates that decays into two b-quarks. Both techniques use the anti-k_t jet algorithm but with different radius parameters. The first technique uses jets with a radius parameter R = 0.4 and it is used when each b-quark from the $H \to b\bar{b}$ decay can be reconstructed as a distinct b-jet. This is referred to as the "resolved analysis"[68]. The second technique uses jets with a radius parameter R = 1.0, also

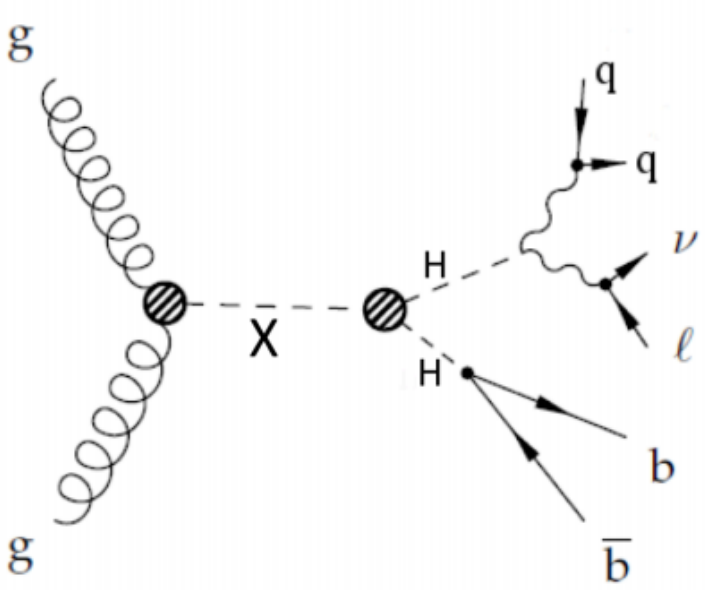

FIGURE 5.2. Schematic diagram of resonant Higgs boson pair production with the subsequent Higgs and W boson decays.

know as large-R jets, and is used when the b-quarks cannot be reconstructed as two distinct b-jets. Instead the $H \to b\bar{b}$ candidate is identified as the single large-R jet. This technique is referred to as the "boosted analysis"[69]. In both analyses, the jets from the hadronically decaying W boson are reconstructed as anti-k_t jets with radius parameter R = 0.4. The non-resonant, SM production, search uses the resolved analysis exclusively, while the resonant analysis is performed using resolved analysis for resonant masses below 1300 GeV and the boosted analysis for resonant masses above 1300 GeV. The cross-over point was chosen to maximize the sensitivity of the search.

5.2. Data and Monte Carlo Samples

5.2.1. Data

The analysis presented uses the full proton-proton collision dataset collected in 2015 and 2016 as the center-of-mass energy of 13 TeV passing data quality checks requiring good conditions of all sub-detectors. The data that are currently used correspond to an integrated luminosity of 36.1 fb^{-1} (3.2 fb^{-1} from 2015 plus 32.8 fb^{-1} from 2016)[1].

.

5.2.2. Monte Carlo Samples

With the exception of the QCD multijet background described in 5.4.8, MC simulated events are used to estimate SM backgrounds and the signal acceptances. Table 5.1 summarizes the MC samples used for background estimation.

The $t\bar{t}$ and single top-quark samples are generated with Powheg-Box v2 [70] using CT10 parton distribution functions (PDF) interfaced to Pythia 6.428 [71] for parton shower, using the Perugia2012 [72] tune with CTEQ6L1 [73] PDF for the underlying event descriptions. EvtGen v1.2.0 [74] is used for properties of the bottom and charm hadron decays and charmed hadron decays. The mass of the top quark is set to $m_t = 172.5$ GeV. At least one top quark in the $t\bar{t}$ event is required to decay to a final state with leptons, excluding taus. The cross section of $t\bar{t}$ is known to NNLO in QCD including re-summation of next-to-next-to-leading logarithmic

[1] The following GoodRunLists (GRL) are used:
data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml
and
data16_13TeV.periodAllYear_DetStatus-v88-pro20-21_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml.
The GRLs were retrieved from the GoodRunListsForAnalysisRun2 twiki

Process	Generator	$\sigma \times$ BR [pb]
$t\bar{t} \to WWbb \to l\nu bb + X$	POWHEG+PYTHIA6	451.65
Wt incl.	POWHEG+PYTHIA6	71.7
single t, s-channel, $\to l\nu + X$	POWHEG+PYTHIA6	3.31
single t, t-channel, $\to l\nu + X$	POWHEG+PYTHIA6	69.5
W+jets, $W \to l\nu$	SHERPA	61510
Z+jets, $Z \to ll$	SHERPA	6425
Dibosons incl.	SHERPA	47.3
ggh *incl.*	POWHEG+PYTHIA8	48.5
tth, $\to l\nu + X$	aMC@NLO + HERWIG++	0.223

TABLE 5.1. SM MC samples used for background estimation.

(NNLL) soft gluon terms, and the reference value used in ATLAS is calculated using TOP++ 2.0 [75]. The parameter HDAMP, used to regulate the high-p_T radiation in POWHEG, is set to m_t for good data/MC agreement in the high p_T region [76]. Each process of single top-quark (t-channel, s-channel and Wt-channel) is generated separately. The cross section of single-top is calculated with the prescriptions in Ref. [77, 78].

SHERPA v2.2.1 [79] with the NNPDF 3.0 [80] PDF set is used as the baseline generator for the $(W \to \ell\nu)/(Z \to \ell\ell)$+jets background. The diboson processes (WW, WZ and ZZ) are generated with SHERPA with the CT10 PDF set.

The ggH and VBF inclusive samples are generated with POWHEG using the CT10 PDF set interfaced to PYTHIA8 for parton shower, while ttH is a semi-leptonic sample generated with MADGRAPH5_aMC@NLO interfaced to HERWIG++. The ggF cross section is normalized by using computations including up to three QCD loops (N3LO) [81]. VBF, Wh and Zh samples, with inclusive h, W and Z decays are also generated using PYTHIA8.

Signal samples are generated with MADGRAPH5_aMC@NLO [82] interfaced to HERWIG++ according to the procedure defined in Ref. [83]. Events are generated

with an effective Lagrangian in the infinite top-quark mass approximation, and re-weighting the generated events with form factors that take into account the finite mass of the top quark. This procedure partially accounts for the finite top-quark mass effects [84]. After the full analysis chain was developed, there were also developments in the theoretical front, which took full NLO calculation and top mass into account [85, 86]. This led to a slight difference in m_{HH} shape. A re-weighting scheme was then developed to correct m_{HH} shape as described in these slides. [2] The overall effect in the sensitivity is a loss of signal efficiency by about 30%, which is also seen by other analysis such as $HH \to bbbb$.

Table 5.2 shows the list of HH signals. They use a heavy Higgs scalar model as the signal hypothesis. The masses of the heavy Higgs range from 260 GeV to 3000 GeV while the Higgs width is set to 10 MeV, therefore the model is valid in the Narrow Width Approximation (NWA). The non-resonant signal is normalized to $\sigma(\text{pp} \to \text{HH}) \times \text{Br}(\text{HH} \to \text{WWbb}) = 0.590$ pb (the expected SM cross section) , the resonant ones are normalized to 0.044 pb for $m_H < 2000$ GeV and to 0.041 for $m_H \geq 2000$ GeV (the Run I limits).

Process	Generator
HH SM	MADGRAPH5_AMC@NLO + Herwig++ including Form Factor
$S \to HH$ ($m_S = 260 - 3000$) GeV	MADGRAPH5_AMC@NLO + Herwig++including Form Factor

TABLE 5.2. Di-Higgs signal samples used in the analysis.

Additional pp collisions generated with PYTHIA 8.186 are overlaid to model the effects of the pileup for all simulated events. All simulated events are processed with the same reconstruction algorithm used for data. All background samples are processed through the full ATLAS detector simulation [87] based on GEANT4 [88] while signal samples use the Atlas Fast simulation.

[2] https://indico.cern.ch/event/652372/

5.3. Object Reconstruction

The observable particles of this analysis includes electrons, muons, neutrinos and jets, including b-jets. The identification criteria and the selection applied to the reconstructed objects are defined in the present section.

5.3.1. Electrons

5.3.1.1. Electron reconstruction

For this analysis, two set of electron selections are defined. They are denoted as VHLooseElectron and SignalElectron. The selections are defined as the following:

VHLooseElectron: The electron p_T is required to be greater than 7 GeV. The electron cluster should be in the range of $|\eta| < 2.47$. Loose likelihood identification is applied in this criteria. Impact parameter significance ($|d_0^{\text{sig}}| = d_0/\sigma_{d_0}$) less than 10 standard deviations. and $|\Delta z_0^{\text{IBL}} \sin\theta| < 0.5$ mm are also required, where IBL refers to the ATLAS Insertable B-Layer.

SignalElectron: The electron is required to pass the VHLooseElectron selection with its p_T required to be greater than 27 GeV. The electron cluster should be in the range of $|\eta| < 2.47$ but excluded from the crack region ($1.37 < |\eta| < 1.52$). Tight likelihood identification is applied in SignalElectron criteria with the impact parameter significance required to be less than two. In addition, the electron is required to be isolated by passing the `FixedCutTightTrackOnly` isolation working point which corresponds to a cut on the ratio of $p_T^{\text{varcone0.2}}$ to electron p_T of 0.06 (i.e $p_T^{\text{varcone0.2}}/p_T < 0.06$).

A summary of the electron selections is shown in Table 5.3.

Electron Selection	p_T	$\lvert\eta\rvert$	ID	$\lvert d_0^{\text{sig}}\rvert$	$\lvert\Delta z_0^{\text{IBL}} \sin\theta\rvert$	Isolation
VHLoooseElectron	>7 GeV	< 2.47	LH Loose	< 10	< 0.5 mm	-
SignalElectron	>27 GeV	< 2.47 and $\notin [1.37, 1.52]$	LH Tight	< 2	< 0.5 mm	`FixedCutTightTrackOnly`

TABLE 5.3. Electron selection requirements.

5.3.2. Muons

5.3.2.1. Muon reconstruction

For this analysis, two sets of muon selections are defined. They are denoted as VHLooseMuon and SignalMuon. The selections are defined as the following:

VHLooseMuon: The muon p_T is required to be greater than 7 GeV. The muon cluster should be in the range of $|\eta| < 2.7$. Loose identification is applied in this criteria. Impact parameter significance ($|d_0^{sig}|$) less than 6 standard deviations. and $|\Delta z_0^{\text{IBL}} \sin\theta| < 0.5$ mm are also required.

SignalMuon: The muon is required to pass the VHLooseMuon selection with its p_T required to be greater than 27 GeV and should be in the range of $|\eta| < 2.4$. Medium identification is applied in SignalMuon criteria with the impact parameter significance required to be less than 2. In addition, the muon is required to be isolated by passing the `FixedCutTightTrackOnly` isolation working point which corresponds to a cut on the ratio of $p_T^{\text{varcone0.3}}$ to muon p_T of 0.06 (i.e $p_T^{\text{varcone0.3}}/p_T < 0.06$).

A summary of the muon selections is shown in Table 5.4.

Muon Selection	p_T	$\lvert\eta\rvert$	ID	$\lvert d_0^{\text{sig}}\rvert$	$\lvert\Delta z_0^{\text{IBL}} \sin\theta\rvert$	Isolation
VHLoooseMuon	>7 GeV	< 2.7	Loose quality	< 6	< 0.5 mm	-
SignalMuon	>27 GeV	< 2.4	Medium quality	< 2	< 0.5 mm	`FixedCutTightTrackOnly`

TABLE 5.4. Muon selection requirements.

5.3.3. Jets

5.3.3.1. Large-R jets

For signal processes with a large resonant mass the b-jets produced by the Higgs may be too close together to be resolved by the R=0.4 calorimeter-based jets (calo-jets). This effect is expected to be noticeable when $p_T^H > 500$ GeV[3]. Our approach to reconstructing the $H \to b\bar{b}$ system in this "boosted" regime is to use a large radius (large-R) jet with radius parameter R =1.0. The large-R jets are required to have $p_T > 250$ GeV and $|\eta| < 2.0$.

5.3.3.2. Track jets

To identify a large-R jet that is consistent, the decay of $H \to b\bar{b}$, a method developed by ATLAS is to reconstruct subjets within the large-R jet and identify the subjets whether it is a b-jet or not by using a b-tagging algorithm. The baseline method is to use subjets built from tracks (track jets). For the boosted analysis, track jets are required to have $p_T > 10$GeV and $|\eta| < 2.5$ for them to be within the inner detector acceptance. They are also required to have at least two track constituents. The MV2c10 working point for track jets is the 77% Fixed Cut efficiency.

5.3.3.3. Small-R jets

Signal jets are defined as jets which passes the jet cleaning and JVT criteria, described in the previous section. They are further required to have $p_T > 20$ GeV and $|\eta| < 2.5$.

Signal jets are labeled **b-jets** if they pass the `MV2c10` 85% WP cut and labeled as **light-jets** if they fail the cut.

[3] Using the rule of thumb $\Delta R = 2m/p_T$, where $m = m_H$ and $\Delta R = 0.4$

Table 5.5 summarizes the jets selection.

	Signal Jets
Algorithm	anti$-k_t$
p_T	20 GeV
$\|\eta\|$	< 2.5
Quality	not "bad" jet
Pile-up jet removal	JVT > 0.59 when $\|\eta\| < 2.5$ *and* $p_T < 60$ GeV
b-tagging	`MV2c10`, 85% fixed-cut WP, labelled as b-jets pass cut, light-jets if fail cut

TABLE 5.5. Selection for jets with distance parameter $R = 0.4$.

5.3.4. Missing transverse momentum ($\not{E}_T$)

The missing transverse momentum (MET, or $\not{E}_T$) [65] used in this analysis is computed by using electrons that pass the VHLooseElectron selection, muons passing the VHLooseMuon selection and jets of the analysis.[4] The track-based soft term[5] (TST) is the recommended soft term component for the MET calculation. Photons and hadronically decaying taus are included in the $\not{E}_T$ calculation as jets since they are not used explicitly in the event reconstruction.

5.3.5. Overlap removal

Each object identification algorithm in ATLAS runs independently. This means the same physical object can be identified as multiple things. One example is electrons and jets, both of which can have tracks and EM calorimeter signatures. In order to uniquely identify objects, overlapping objects are removed according to the overlap removal procedure defined in this section. Electrons and muons that pass the VHLooseElectron and VHLooseMuon selections (as defined in Sec. 5.3.1.1

[4] From MET_Core_AntiKt4EMTopo with the MissingETAssociationMap using the METMaker tool. All calibrated jets are passed to the METMaker tool as prescribed on the EtMiss subgroup `twiki`

[5] Defined on this twiki.

and 5.3.2.1) are considered for overlap removal. Calorimeter jets which pass the JVT requirement are also considered for overlap removal. The procedure is defined as follows.

If an electron and a muon shares a track, the muon is removed if it is *calo-tagged*. Otherwise, the electron is removed. Calorimeter jets are then removed if they are within $\Delta R(\text{calo-jet}, \text{electron}) < 0.2$ of surviving electrons. Electrons that satisfy $\Delta R(\text{electron}, \text{calo-jet}) < \min(0.4,\ 0.04 + 10\ \text{GeV}/E_{\text{T}}^{\text{electron}})$ are removed. The surviving calorimeter jets are removed if they are within $\Delta R(\text{calo-jet}, \text{muon}) < 0.2$ and do **not** pass any of the following criteria:

- The number of tracks in the jet are more than two.
- $p_T^{\text{muon}}/p_T^{\text{calo-jet}} < 0.5$ AND $p_T^{\text{muon}}/p_T^{\text{tracks in calo-jet}} < 0.7$.

Muons that satisfy $\Delta R(\text{muon}, \text{calo-jet}) < \min(0.4,\ 0.04 + 10\ \text{GeV}/p_T^{\text{muon}})$ are removed. The overlap removal procedure is implemented using ASG's `AssociationUtils` package and summarized in Table 5.6.

Overlapping Objects	Removal Procedure
Electron - Muon	If share track, remove muon if calo-tagged. Otherwise remove electron.
Electron - Calo-jet	If $\Delta R(\text{calo-jet}, \text{electron})$ ¡ 0.2, remove calo-jet. If $\Delta R(\text{electron}, \text{calo-jet})$ ¡ $\min(0.4,\ 0.04 + 10\ \text{GeV}/E_{\text{T}}^{\text{electron}})$, remove electron.
Muon - Calo-jet	If $\Delta R(\text{calo-jet}, \text{muon})$ ¡ 0.2, remove calo-jet if: a) Number of tracks in calo-jet ≤ 2, OR b) $p_T^{\text{muon}}/p_T^{\text{calo-jet}} > 0.5$ AND $p_T^{\text{muon}}/p_T^{\text{tracks in calo-jet}} > 0.7$. If $\Delta R(\text{muon}, \text{calo-jet})$ ¡ $\min(0.4,\ 0.04 + 10\ \text{GeV}/p_T^{\text{muon}})$, remove muon.

TABLE 5.6. A summary of the overlap removal procedure.

5.4. Resolved Analysis

5.4.1. Event Selection

The final state of interest consists of one charged lepton, one neutrino, and four quarks, two of which are b-quarks. Hence the detector signature consists of

one charged lepton (e/μ), large $\not{E}_T$, and four or more anti-k_t jets of which two are b jets from the h decay while the other two are light jets from the hadronic decay of the W boson. One challenge in the event reconstruction is to correctly identify the pair of light jets from the W boson decay. This information is also used to solve the z component of the neutrino momentum. For the HH signal there is an additional complication due to the fact that one of the W bosons is off-shell, and thus for this W there is no W mass constraint. This section details the stages of the event reconstruction and the progression towards the final selection which defines the signal region. In addition, signal depleted control regions are defined in the next section which are used to check the consistency of the SM background predictions with the data in the control regions. The search has been kept "blinded" until the comparison between data and simulation of backgrounds are well understood in the signal depleted control regions.

5.4.2. Trigger requirement

Events are selected using the unprescaled single lepton triggers. The list of triggers used in this analysis is shown in Table 5.7. Events are selected with a logical OR between the triggers listed in Table 5.7.

5.4.2.1. Pre-selection

The following selection criteria are applied at the pre-selection level to the recorded events:

– Good detector conditions are required based on data quality assessment. See Appendix B.0.1 for more details.

Dataset	Trigger items
2015	mu20_iloose_L1MU15 mu50 e24_lhmedium_L1EM18VH (MC) e24_lhmedium_L1EM20VH (data) e60_lhmedium e120_lhloose
2016 - Period A	mu24_iloose_L1MU15 (MC) mu24_iloose (data) mu40 e26_lhtight_nod0_ivarloose e60_lhmedium_nod0 e60_medium e140_lhloose_nod0 e300_etcut
2016 - Period B-D3	mu24_ivarmedium mu50 e26_lhtight_nod0_ivarloose e60_lhmedium_nod0 e60_lhmedium e140_lhloose_nod0 e300_etcut
2016 - Period D4-E3	mu26_ivarmedium mu50 e26_lhtight_nod0_ivarloose e60_lhmedium_nod0 e60_lhmedium e140_lhloose_nod0 e300_etcut
2016 - Period $\geq$ F	mu26_ivarmedium mu50 e26_lhtight_nod0_ivarloose e60_lhmedium_nod0 e60_lhmedium e140_lhloose_nod0 e300_etcut

TABLE 5.7. Summary of trigger items used for 2015 and 2016 data. For 2016 data, different triggers were used for different data run periods. All triggers are unprescaled.

– The presence of a primary vertex with at least two tracks. Among all primary vertices, that with the highest $\sum p^2_{\mathrm{T,trk}}$, where $p_{\mathrm{T,trk}}$ is the transverse momentum of tracks associated with the vertex, is retained as the primary interaction vertex;

– at least one SignalElectron (e) or SignalMuon (μ), as defined in Section 5.3.1.1 and Section 5.3.2.1, and it must be trigger matched to the corresponding HLT object which fires the trigger;

– at least four jets, of which two and only two are b-tagged.

5.4.3. Event Reconstruction

Events are reconstructed by first requiring exactly two 2 b-tag jets and at least two light jets and at most three light jets. In events with three light jets, the pair with the lowest ΔR between them are selected as W jet candidates. This procedure yields the correct jet assignment in 70% of the cases for signal events where the hadronic daughters of the W boson can be correctly matched to reconstructed jets.

The event kinematics of the $H \rightarrow WW^* \rightarrow l\nu qq$ topology can be fully reconstructed. In fact, among all four-momenta of the final state particle, only the component of the neutrino momentum along the beam axis, p_z in the following, is unknown while its transverse momentum is the measured $\not\!\!E_T$. Imposing the relation:

$$m_h^2 = (p^l + p^\nu + p^{j1} + p^{j2})^2 \tag{5.1}$$

where p^i is the four-momenta of particle i, the neutrino p_z can be reconstructed using the relations:

$$p^\nu_E = E^\nu = \sqrt{P_T^2 + p_z^2} \quad p^\nu_x = P_T cos(\phi) \quad p^\nu_y = P_T sin(\phi)$$

where ϕ is the azimuthal angle of the $\not\!\!E_T$, E^ν the neutrino energy, p_x and p_y the two transverse spatial components of the neutrino momentum. Equation 5.1 is a quadratic expression in p_z. It can have two real, one real or two complex solutions. In the last case only the real part of the complex solution is taken into account, therefore a single value of p_z is obtained. In the first case the solution with the neutrino direction closest to the charged lepton is retained. It has been shown that this algorithm selects the correct solution in approximately 60% of the cases (see Appendix F).

5.4.4. bb$\tau\tau$ analysis overlap removal

In order to remove overlap with the $bb\tau\tau$ analysis we reject any event containing at least one hadronic τ candidate that could be identified by the $bb\tau\tau$ analysis, that fullfill the following requirements:

- $p_T > 20$ GeV and $|\eta| < 2.5$;
- one or three prongs;
- unit charge;
- pass the medium τ ID BDT working point.

The rejection of such events causes a signal efficiency drop of about 3%.

5.4.5. Kinematic selection

Kinematic selection is used to suppress mainly $t\bar{t}$ background while keeping high signal efficiency. A schematic view of the $HH \to WW*b\bar{b}$ and the $t\bar{t} \to WWbb$ event topology is shown in Figure 5.3.

The $t\bar{t}$ events are typically characterized by two b-jets and two W bosons such that the ΔR separation between the two b-jets and between the W bosons is large. On the contrary, in particular when the invariant mass of the m_{HH} is high, the signal is characterized by two b-jets which are close together in ΔR and by two W bosons which are also relatively closer than in the $t\bar{t}$ case. Moreover, while for the signal the two b-jets have an invariant mass equal to m_h, this is not necessarily the case for the $t\bar{t}$ background. Following these considerations, the typical separation variables are:

- the p_T of the $b\bar{b}$ pair (p_T^{bb});
- the ΔR of the $b\bar{b}$ pair (ΔR^{bb});

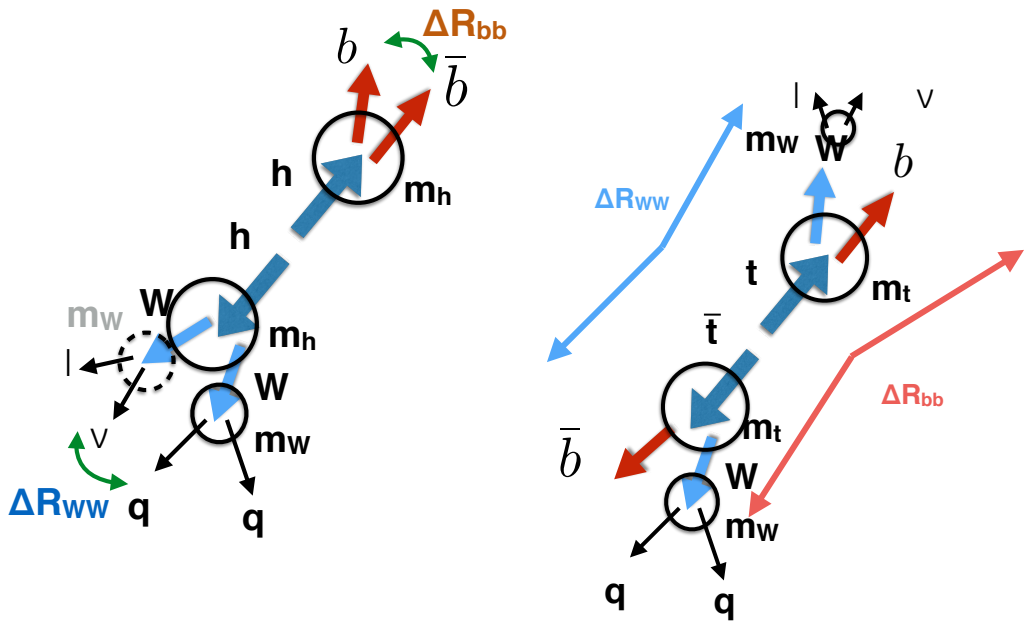

FIGURE 5.3. Schematic view of a $HH \to WWbb$ event compared to a $t\bar{t} \to WWbb$ event.

– the p_T of the WW pair (p_T^{WW});

– the ΔR of the WW pair (ΔR^{WW});

– the mass of the WW system computed using the calculated neutrino longitudinal momentum (m_{WW}). This value is exactly equal to m_h if a real solution is found, it is larger if no real solution is found;

– the invariant mass of the di-Higgs boson candidate system (m_{HH}).

– the invariant mass of the 2 b-jets boson system (m_{bb}).

5.4.6. Signal region definitions

The signal selection criteria have been optimized by maximizing the Poisson significance at the end of the selection based on MC simulation [6]. The Poisson significance formula depends on the absolute yield of expected signal and background events. For the optimization formula the $t\bar{t}$ background was normalized to data with $m_{bb} < 100$ GeV or $m_{bb} > 140$ GeV, where the majority of the signal is rejected. Four signal hypotheses have been used in the optimization:

- a heavy Higgs with $m_S = 500$ GeV and $m_S = 700$ GeV (defined as low-mass analysis),
- a heavy Higgs with $m_S = 2000$ GeV (defined as high-mass analysis) and
- a non-resonant di-Higgs production (defined as non-resonant analysis).

An additional mass point with $m_S = 1400$ GeV was also checked. The resulting selection and the corresponding sensitivity are very similar to the selection for $m_S = 2000$ GeV, and hence that selection is dropped.[7]

The signal regions for the reference signal hypotheses are summarized in Table 5.8.

The *non-res* and *m500* selections are exclusively used for non-resonant signal and resonant signal with mass 500 GeV respectively. The *low-mass* selection is used for signal masses from 600 to 1300 GeV, while the *high-mass* selection is used for signals with masses between 1400 and 3000 GeV. In addition, requirements are

[6] a two step procedure has been implemented. In the first step each selection criteria is optimized, in the second step, all selections are set to their optimal value and selections are varied one by one to look for a different optimization point. Correlation among variables could in fact spoil the results obtained at the first step.

[7] See `https://indico.cern.ch/event/641988/contributions/2604588/attachments/1465307/2265002/bbWW_Weekly_Optimization_Revisited_24May2018.pdf`

variable	*Non-Res*	*m500*	*low-mass*	*high-mass*
$\not\!\!E_T$ (GeV)	> 25	> 25	> 25	> 25
m_{WW} (GeV)	< 130	< 130	< 130	no-cut
p_{T}^{bb} (GeV)	> 300	> 210	> 210	> 350
p_{T}^{WW} (GeV)	> 250	> 150	> 250	> 250
ΔR_{WW}	no-cut	no-cut	no-cut	< 1.5
m_{bb} (GeV)	105-135	105-135	105-135	105-135

TABLE 5.8. Criteria for non-resonant, *m500*, *low-mass* and *high-mass* selection. The m_{HH} window is not applied for non-resonant signal, and for resonant signals m_{HH} depends on the mass.

m_S (GeV)	500	600	700	750	800
m_{HH} (GeV)	480-530	560-640	625-775	660-840	695 - 905
m_S (GeV)	900	1000	1100	1200	1300
m_{HH} (GeV)	760-970	840-1160	925-1275	1010-1390	1095-1505
m_S (GeV)	1400	1500	1600	1800	2000
m_{HH} (GeV)	1250-1550	1340-1660	1430-1770	1750-2020	1910-2170
m_S (GeV)	2250	2500	2750	3000	
m_{HH} (GeV)	2040-2460	2330-2740	2570-2950	2760-3210	

TABLE 5.9. Window selection on m_{HH} as a function of the resonance mass m_S.

placed on the reconstructed di-Higgs invariant mass m_{HH} as a function of the signal resonance mass m_S, as shown in Table 5.9. The resolution of the reconstructed m_{HH} ranges from 6% at 500 GeV to 10% at 3000 GeV.

5.4.7. Background Determination

In the present analysis we expect that at the end of the event selection the sample will be largely dominated by $t\bar{t}$ and multi-jet background, therefore the $t\bar{t}$ background normalization is derived from data while, as described in Sec. 5.4.8, the multi-jet background is derived using a data-driven ABCD method. For all the other backgrounds, e.g. di-boson, Higgs, W+jets, the MC is used appropriately

Selection	non-res	m500	m700	m2000
Buffer/SR	1.85	1.95	1.90	1.63
Sidebands/SR	20.5	12.6	13.4	5.6

TABLE 5.10. The ratios of S/B in the buffer zone and sidebands compared to the S/B in the final SR.

normalized by using the expected cross sections and the integrated luminosity that has been collected.

5.4.7.1. Top normalization and control region

The $t\bar{t}$ background is normalized and validated using dedicated control regions (CR). Three CR's are defined, one for the SR's of the *non-res* (CR1), one for the *low-mass* analysis (CR2), and one for the *high-mass* analysis (CR3). The CRs are defined in Table 5.11.

Table 5.12 through 5.15 show the number of observed events and expected background events in the top CRs, and also in the sideband across selections that serve as validation regions. The final signal region is defined by m_{bb} of 105 GeV $< m_{bb} <$ 135 GeV based on optimization. The sidebands are orthogonal to the SR by virtue of having the m_{bb} reversed. $m_{bb} <$ 100 GeV or $m_{bb} >$ 140 GeV defines the sidebands in which the control regions are defined. The 5 GeV buffer region is kept on both sides so as to be less affected by systematic effects at the edge. Fig. 5.4 shows m_{bb} for various signal mass points. A comparative study of signal over background in these three regions shows that S/B in the final SR is 5 (20) times higher than in the sidebands for m2000 (non-resonance) while S/B in the final SR is approximate twice as high as in the buffer zones. Table 5.10 shows the ratios of S/B in the buffer zones and sidebands compared to the S/B in the final SR.

variable	CR1	CR2	CR3
m_{bb} (GeV)	$m_{bb} < 100$ or $m_{bb} > 140$	$m_{bb} < 100$ or $m_{bb} > 140$	$m_{bb} < 100$ or $m_{bb} > 140$
m_{WW} (GeV)	< 130	< 130	no-cut
p_{T}^{bb} (GeV)	> 300	> 210	> 350

TABLE 5.11. Definition of the kinematic regions used to normalize the Top background. $m_{bb} < 100$ GeV or $m_{bb} > 140$ GeV defines the sidebands in which the control regions are defined. Expected SM backgrounds are then checked against data at each subsequent selection.

CR1: $m_{b\bar{b}}$ Sideband				
Sample	mww	bbpt210	bbpt300	wwpt250
$t\bar{t}$	23776.6 ± 87.2	531.7 ± 13.1	109.9 ± 5.9	63.9 ± 4.6
QCD	13310.5 ± 500.3	250.2 ± 30.6	33.7 ± 4.1	21.4 ± 2.6
W+jets	3938.9 ± 31.1	124.7 ± 3.5	29.3 ± 1.4	17.1 ± 1.1
SingleTop	1605.4 ± 18.0	76.0 ± 3.8	20.1 ± 2.0	13.5 ± 1.7
Dibosons	109.9 ± 2.7	8.3 ± 0.8	2.2 ± 0.4	1.5 ± 0.4
Z+jets	1107.6 ± 8.4	27.1 ± 0.8	6.7 ± 0.4	2.4 ± 0.2
Background Sum	43849.0± 509.2	1017.9± 33.7	201.9± 7.6	119.8± 5.7
XhhSM	44.6 ± 2.2	9.1 ± 0.7	1.5 ± 0.2	1.1 ± 0.1
Data	43902.0	1069.0	206.0	138.0

TABLE 5.12. The number of observed events and expected background events in the m_{bb} side-bands for the *non-res* selection. The top CR1 is defined at the bbpt300 selection. No NF has been applied to the background yields to show the level of data/expectation agreement before normalizing ttbar. Only statistical uncertainties are shown.

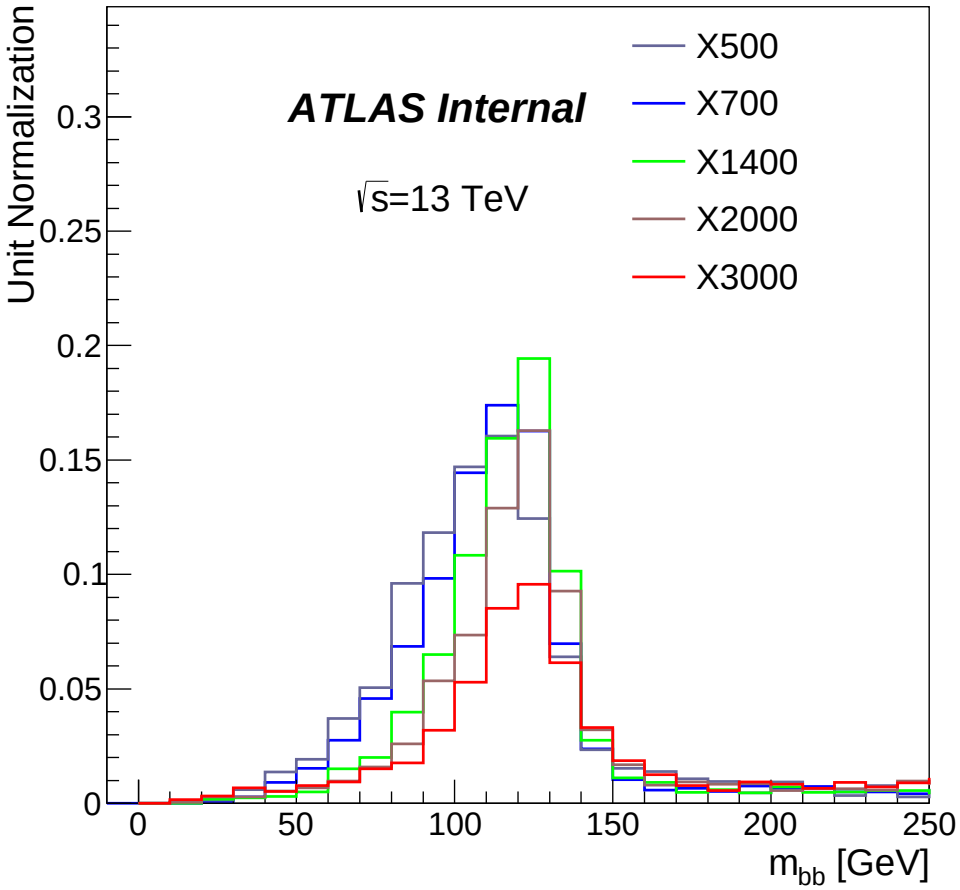

FIGURE 5.4. m_{bb} resolution for signal samples.

Table 5.12 through 5.15 show the number of observed events and expected background events in the top CRs before the normalization factors have been applied to the top background sample.

CR2: $m_{b\bar{b}}$ Sideband				
Sample	mww	bbpt210	wwpt150	hh500
$t\bar{t}$	23776.6 ± 87.2	531.7 ± 13.1	432.7 ± 11.8	35.5 ± 3.2
QCD	13310.5 ± 500.3	250.2 ± 30.6	206.3 ± 25.3	16.9 ± 2.1
W+jets	3938.9 ± 31.1	124.7 ± 3.5	105.9 ± 3.3	4.9 ± 0.6
SingleTop	1605.4 ± 18.0	76.0 ± 3.8	64.9 ± 3.5	2.8 ± 0.6
Dibosons	109.9 ± 2.7	8.3 ± 0.8	6.7 ± 0.8	0.9 ± 0.2
Z+jets	1107.6 ± 8.4	27.1 ± 0.8	19.0 ± 0.7	1.5 ± 0.2
Background Sum	43849.0± 509.2	1017.9± 33.7	835.5± 28.3	62.5± 3.9
Xhh500	3.2 ± 0.1	0.6 ± 0.1	0.6 ± 0.1	0.2 ± 0.1
Data	43902.0	1069.0	898.0	73.0

TABLE 5.13. The number of observed events and expected background events in the m_{bb} side-bands for the low-mass selection, m500. The top CR2 is defined at the bbpt210 selection. To show how well the prediction matches data, no NF has been applied to any background. Only statistical uncertainties are shown.

CR2: $m_{b\bar{b}}$ Sideband				
Sample	mww	bbpt210	wwpt250	hh700
$t\bar{t}$	23776.6 ± 87.2	531.7 ± 13.1	175.6 ± 7.5	49.9 ± 3.9
QCD	13310.5 ± 500.3	250.2 ± 30.6	72.4 ± 8.9	28.4 ± 3.5
W+jets	3938.9 ± 31.1	124.7 ± 3.5	45.7 ± 2.1	13.7 ± 1.4
SingleTop	1605.4 ± 18.0	76.0 ± 3.8	28.4 ± 2.4	6.9 ± 1.1
Diboson	109.9 ± 2.7	8.3 ± 0.8	2.8 ± 0.5	0.7 ± 0.2
Z+jets	1107.6 ± 8.4	27.1 ± 0.8	5.8 ± 0.4	2.0 ± 0.3
Background Sum	43849.0± 509.2	1017.9± 33.7	330.7± 12.1	101.5± 5.5
Xhh700	4.2 ± 0.2	2.2 ± 0.1	1.5 ± 0.1	1.0 ± 0.1
Data	43902.0	1069.0	367.0	124.0

TABLE 5.14. The number of observed events and expected background events in the m_{bb} side-bands for the low-mass selection, m700. The top CR2 is defined at the bbpt210 selection. To show how well the prediction matches data, no NF has been applied to any background. Only statistical uncertainties are shown.

CR3: $m_{b\bar{b}}$ Sideband				
Sample	bbpt350	wwpt250	drww15	hh2000
$t\bar{t}$	8568.7 ± 52.1	7095.6 ± 47.5	1940.5 ± 25.1	122.3 ± 6.5
QCD	1538.7 ± 252.7	1359.5 ± 75.9	392.7 ± 21.9	20.7 ± 1.2
W+jets	2259.5 ± 7.9	1952.1 ± 7.4	696.6 ± 4.6	55.5 ± 1.1
SingleTop	1778.1 ± 19.4	1601.6 ± 18.4	405.4 ± 9.2	29.6 ± 2.6
Dibosons	170.6 ± 3.9	147.1 ± 3.7	46.8 ± 2.1	3.4 ± 0.6
Z+jets	403.6 ± 2.1	307.6 ± 1.8	95.6 ± 1.1	7.5 ± 0.3
Background Sum	14719.1± 258.9	12463.5± 91.8	3577.5± 35.0	238.9± 7.2
Xhh2000	25.7 ± 0.4	24.0 ± 0.4	9.6 ± 0.3	2.9 ± 0.1
Data	14862.0	12450.0	3761.0	250.0

TABLE 5.15. The number of observed events and expected background events in the m_{bb} side-bands for the high-mass selection. The top CR3 is defined at the bbpt350 selection. No NF has been applied to the background yields. Only statistical uncertainties are shown.

Top background normalization factors in the two CRs.

region	NF	$\sigma_{stat.}$	$\sigma_{syst.}$
non-res	1.04	±0.20	±0.43
low-mass	1.14	±0.10	±0.35
high-mass	1.02	±0.02	±0.07

TABLE 5.16. Normalization factors for the two CRs, the statistical error includes only data statistics, the systematic error is obtained subtracting in quadrature the statistical error from the total error.

The top normalization factors are determined by a simultaneous fit of signal and control regions, which include both Top CR and QCD CR 5.4.8. It also depends slightly on the m_{HH} window due to the presence of top background in the signal region, and it is furthermore different for the *non-res*, *low-mass* and *high-mass* analyses. The normalization factors of the three top control regions are shown in Table 5.16.

Fig. 5.5 shows the m_{bb} and m_{HH} distributions in the two CRs.

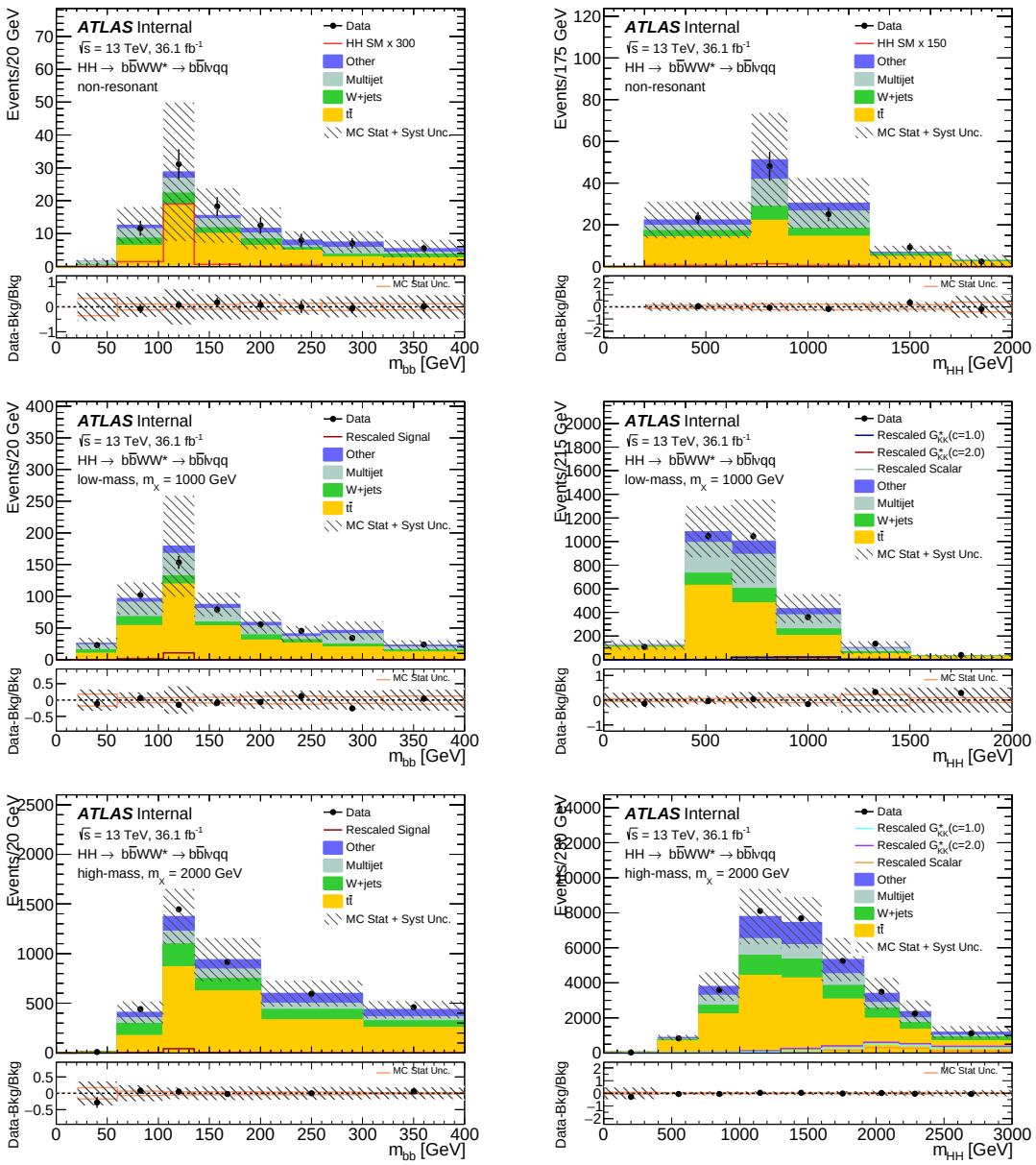

FIGURE 5.5. m_{bb} and m_{HH} in CR1, CR2 and CR3. $t\bar{t}$ NFs as described in 5.4.7.1 have been applied. The uncertainties shown include the statistical and systematic uncertainties described in 5.4.9. Data are blinded in the region $100 < m_{bb} < 140$.

5.4.8. Multi-jet background

Multi-jet backgrounds can enter in the event selection if a jet from heavy flavor decays is mis-identified as an electron or a muon and used as a lepton in the

analysis. Such phenomena are not accurately reproduced by MC simulation, due to large uncertainties in the jet shower shape simulation and uncertainties in muon fragmentation functions and kinematics. In order to estimate the contributions of multi-jet processes, a data-driven ABCD method is used to estimate this background in the present analysis.

The ABCD method uses three control regions (the B, C, and D regions) to estimate the contribution of a given background in the signal (A) region. Selections on two ideally orthogonal variables are used to create the signal and various control regions, e.g. the A region passes both selections, the B and C regions each pass one selection and fail the other, while the D region fails both selections. The absolute value of the significance of the lepton impact parameter and the missing transverse energy (MET) are used as the two variables used to define the regions in the ABCD method for this analysis. The regions are thus defined:

- A region: MET > 25 GeV, $|\sigma_{d_0}| < 2.0$
- B region: MET < 25 GeV, $|\sigma_{d_0}| < 2.0$
- C region: MET > 25 GeV, $|\sigma_{d_0}| > 2.0$
- D region: MET < 25 GeV, $|\sigma_{d_0}| > 2.0$

Figure 5.6 shows a pictoral representation of the four regions.

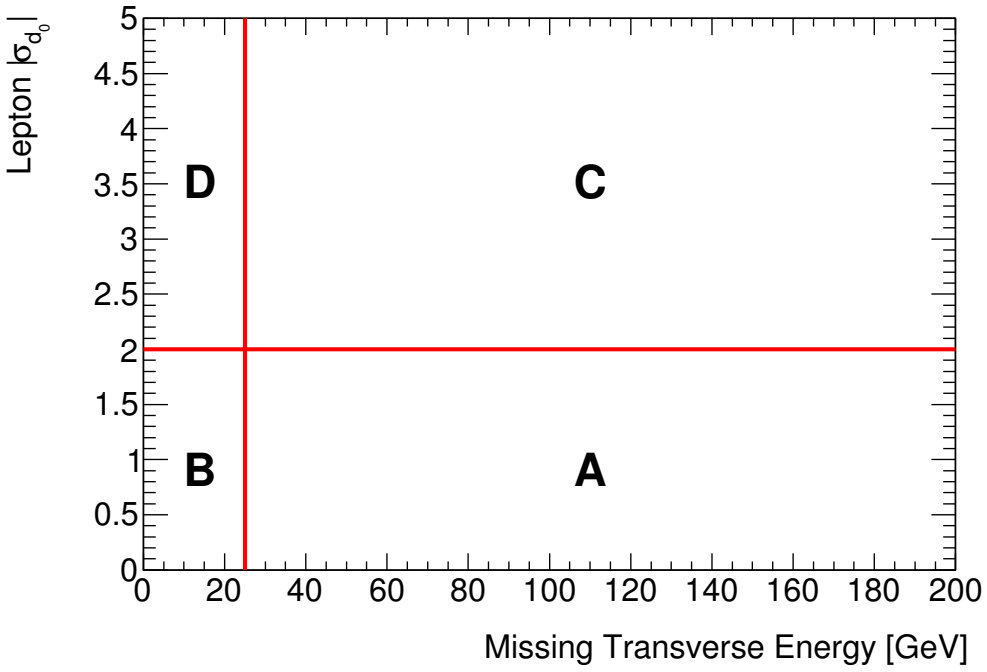

FIGURE 5.6. A pictoral representation of the four regions used in the ABCD calculation.

Assuming that the two variables chosen to define the ABCD regions are completely uncorrelated, the yield of the process being modeled (QCD multi-jets in this case) in the A region is given by

$$N_A = N_C \frac{N_B}{N_D} \tag{5.2}$$

where the yields N_i are yields calculated from data - all Monte Carlo backgrounds ($t\bar{t}$, W/Z+jets, single top, diboson processes) in region i ($N_i = N_i^{\text{data}} - N_i^{\text{MC Bkgs}}$). The assumption underlying Equation 5.2 is that the relationship between the yields in the B and D regions is the same as the relationship between the A and C regions, i.e.

$$\frac{N_A}{N_C} = \frac{N_B}{N_D} \tag{5.3}$$

Using equation 5.3, the quantity $R = \frac{N_C N_B}{N_A N_D}$ can be defined. In the case of two completely uncorrelated variables, $R = 1$ and the ABCD estimation reduces to Equation 5.2. If the two variables are not completely uncorrelated, the R factor enters as a correction to Equation 5.2 for the multi-jet estimation in the A region, and the expression can be rewritten as

$$N_A = R\frac{N_C N_B}{N_D} \tag{5.4}$$

The R factor is calculated for each selection (*non-res*, *low-mass*, and *high-mass*) individually, and the results at each step of the selection is provided in Table 5.17.

QCD R Values, Non-resonant Selection			
mww	bbpt210	bbpt300	wwpt250
0.74 ± 0.04	0.79 ± 0.23	1.07 ± 1.18	—
QCD R Values, Low Mass Selection (m500)			
mww	bbpt210	wwpt150	hh500
0.74 ± 0.04	0.79 ± 0.23	—	—
QCD R Values, Low Mass Selection (m700)			
mww	bbpt210	wwpt250	hh700
0.74 ± 0.04	0.79 ± 0.23	0.09 ± 0.14	—
QCD R Values, High Mass Selection			
bbpt350	wwpt250	drww15	hh2000
0.48 ± 0.09	0.43 ± 0.08	0.50 ± 0.16	4.28 ± 5.30

TABLE 5.17. Values calculated for R at each stage in the *non-res*, *low-mass*, and *high-mass* selections. The estimate of multi-jet contribution in the A region uses the R value calculated after the first criteria of each selection.

In order to minimize the statistical error on R, the R value calculated after the first criteria of each selection (0.74 and 0.48) is used in Equation 5.4 to estimate the multi-jet background after each subsequent selection. This can be done given the compatibility of the R value at the end of the cutflow with that at the point where R is evaluated. In order to check such compatibility with higher statistics, the R

value has been calculated applying each selection just after R is evaluated, in order to check that R is not correlated with each of the selections. The result is shown in Table 5.18.

QCD R Values, Non-resonant Selection			
mww	mww + bbpt210	mww + bbpt300	mww + wwpt250
0.74 ± 0.04	0.79 ± 0.23	1.12 ± 1.22	0.25 ± 0.20
QCD R Values, Low Mass (m500) Selection			
mww	mww + bbpt210	mww + wwpt150	mww + hh500
0.74 ± 0.04	0.79 ± 0.23	0.50 ± 0.08	0.52 ± 0.09
QCD R Values, Low Mass (m700) Selection			
mww	mww + bbpt210	mww + wwpt250	mww + hh700
0.74 ± 0.04	0.79 ± 0.23	0.25 ± 0.20	0.63 ± 0.13
QCD R Values, High Mass Selection			
bbpt350	bbpt350 + wwpt250	bbpt350 + drww15	bbpt350 + hh2000
0.47 ± 0.06	0.44 ± 0.08	0.52 ± 0.17	1.07 ± 0.67

TABLE 5.18. Values of R obtained applying a single selection after R is nominally evaluated.

Once the normalization of the multi-jet background in the A region is calculated using Equation 5.4, the shape of the multi-jet template is taken from the data minus Monte Carlo distribution in the C region since the two are kinematically identical except for $|\sigma_{d_0}|$.

The uncertainty due to the limited statistics in the B and D regions is the main source of the multi-jet estimation method systematics. In order to minimise such error, the yields from the B and D regions used in the ABCD calculation are frozen at a level of the cutflow to minimise statistical fluctuations. The B and D region yields are frozen after the $p_T^{bb} > 210$ GeV for the *non-res* and *low-mass* selection, and after the $p_T(WW) > 250$ GeV for the *high-mass* selection. Appendix G details the study carried out to select the stage at which to freeze the B and D regions.

To further reduce the error coming from the C region, the shape of the data minus Monte Carlo, i.e. non-prompt, $m_{b\bar{b}}$ distribution was studied as a function of each individual criteria for each selection. If the $m_{b\bar{b}}$ shape is unchanged by the additions of further kinematic selections, the C region shape can be taken from an earlier stage in the cutflow, reducing the shape uncertainty and overall statistical error on the QCD yield. To determine the stability of the $m_{b\bar{b}}$ shape, the ratio of events in the $m_{b\bar{b}}$ signal region ([100, 140] GeV) over the numbers of events in the full $m_{b\bar{b}}$ spectrum was computed for each individual criteria using the C region of the ABCD method. This ratio was found to be stable across each selection, and the results of this calculation are provided in Table 5.19.

reOptNonRes: $m_{b\bar{b}}$ SR/Total Ratios For Individual Selections			
mww	bbpt210	bbpt300	wwpt250
0.17 ± 0.02	0.15 ± 0.03	0.13 ± 0.05	0.16 ± 0.05

reOpt500: $m_{b\bar{b}}$ SR/CR Ratios For Individual Selections			
mww	bbpt210	wwpt150	hh500
0.17 ± 0.02	0.15 ± 0.03	0.18 ± 0.02	0.22 ± 0.03

reOpt700: $m_{b\bar{b}}$ SR/Total Ratios For Individual Selections			
mww	bbpt210	wwpt250	hh700
0.17 ± 0.02	0.15 ± 0.03	0.16 ± 0.05	0.16 ± 0.02

reOpt2000: $m_{b\bar{b}}$ SR/Total Ratios For Individual Selections			
bbpt350	wwpt250	drww15	hh2000
0.12 ± 0.08	0.16 ± 0.05	0.19 ± 0.02	0.11 ± 0.14

TABLE 5.19. The ratio of events in the $m_{b\bar{b}}$ signal region ([100, 140] GeV) over the numbers of events in the full $m_{b\bar{b}}$ spectrum for each individual selection. Both event yields are calculated in the C region of the ABCD method. The ratios are found to be stable across each selection.

The normalized shapes of the full $m_{b\bar{b}}$ distributions after each selection are shown in Figure 5.7. The earliest selection with a consistent shape was chosen as the shape for each cutflow and this corresponds to the shape obtained after the $p_T^{bb} > 210$ GeV

for the non-resonant the m500 and m700 selections, and after the $p_T^{WW} > 250$ GeV for the m2000 selection.

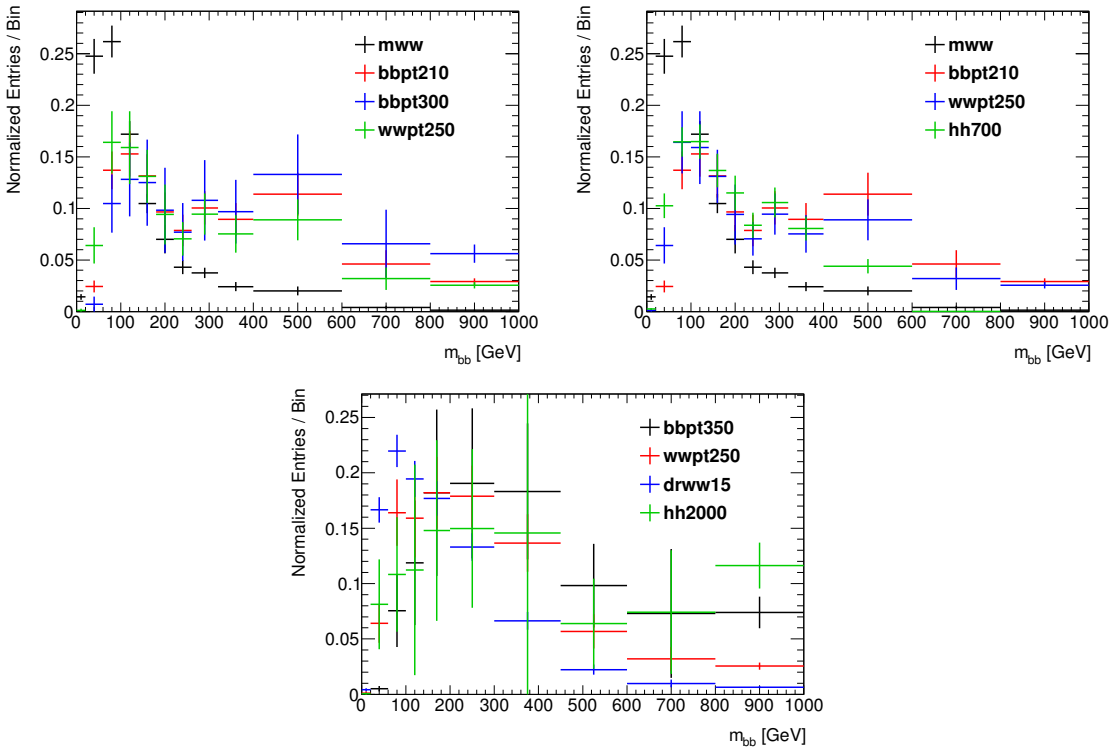

FIGURE 5.7. The normalized shapes of the full $m_{b\bar{b}}$ distributions after each in each selection. Shapes from the non-resonant cutflow are shown at the top left, shapes from the *low-mass* selection are shown at the top left, and shapes from the *high-mass* selection are shown at the bottom. The earliest selection with a consistent shape was chosen as the shape for each cutflow and this corresponds to the shape obtained after the m_{WW} and $p_T^{bb} > 210$ GeV for the *non-res* and *low-mass* selections and after the $p_T^{bb} > 350$ GeV and $p_T^{WW} > 250$ GeV for the *high-mass* mass selection.

Since $t\bar{t}$ and multi-jet contaminate the control regions used for their estimation, additional studies were performed via an iterative procedure to ensure that the $t\bar{t}$ and QCD yields converge to stable values and that the estimation technique is able to disentangle between the two backgrounds. Appendix H details the results of the study, and the yields were found to converge (stable within $< 5\%$) after a few iterations for each cutflow.

To evaluate the systematic error of the estimation, a Sherpa multi-jet bb sample is used to compare the ABCD prediction to the Monte Carlo expectation using events with exactly one lepton and four jets. No b-tagging requirements or other event selection criteria are applied. Pseudo-data is produced using the multi-jet Monte Carlo and events from the nominal $t\bar{t}$ Monte Carlo. The R factor is calculated using the inclusive b-tag (0, 1, or 2 b-tags) and used to estimate the QCD contribution in the two b-tag exclusive region. The percent difference between the number of events from the ABCD estimation and the number of multi-jet events from Monte Carlo used to produce the pseudo-data is taken as the systematic uncertainty. This uncertainty is calculated to be 36%. In addition, the bb MC is used to calculate the R factor for each selection (*non-res*, *low-mass*, and *high-mass*) and each lepton channel, and the results at each kinematic criteria in each selection are provided in the Table 5.20.

5.4.9. Background Shape and Cutflow

The modeling of the background was checked at all selection stages and, in general, shows good agreement with data. Figure 5.8 shows the m_T distribution of the leptonic W boson candidate in the three top control regions. The m_T variable is defined as:

$$m_T = \sqrt{2p_T^{\ell} \not{E}_T \cdot (1 - \cos\Delta\phi)} \ ,$$

where $\Delta\phi$ is the azimuthal angle between p_T^{ℓ} and $\not{E}_T$. The multijet background populates the low values of the m_T distribution, so any mis-modeling of the multijet background would be clearly visible in the m_T distribution.

Figures 5.9 and 5.10 show the $m_{b\bar{b}}$ distributions at the selection stage where all requirements, including m_{HH}, are applied except the one on $m_{b\bar{b}}$ itself. The expected

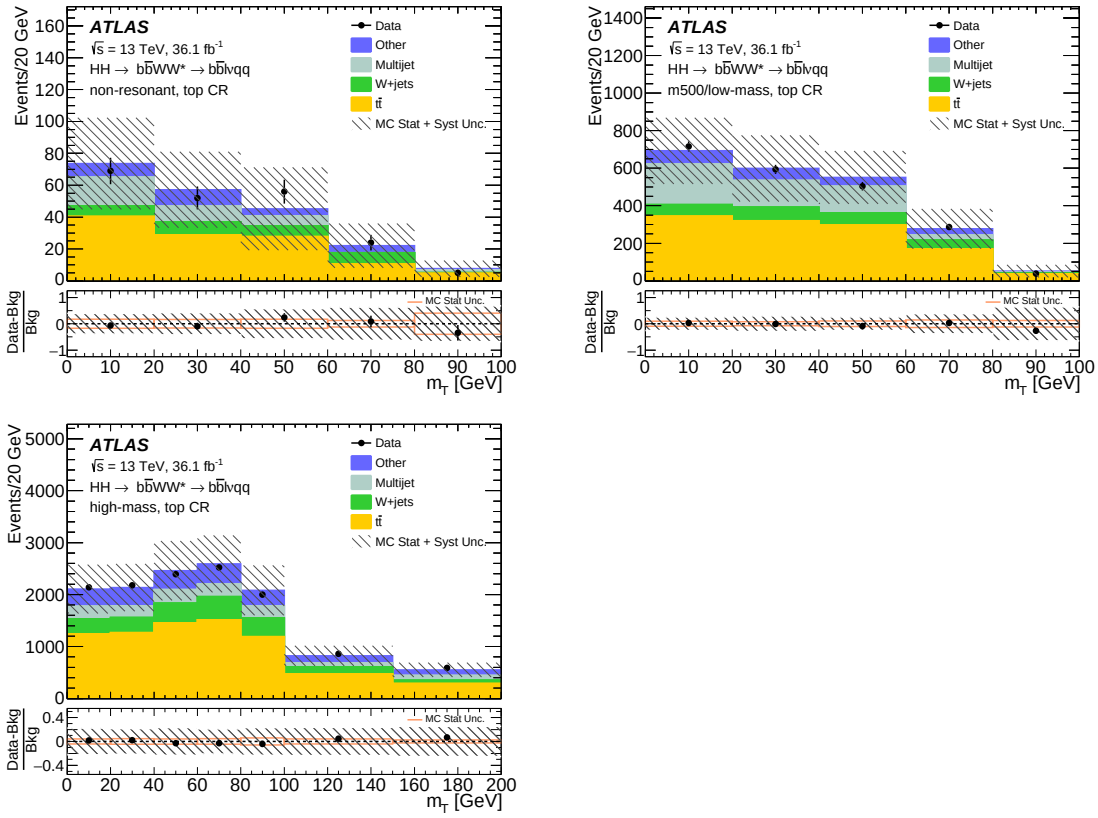

FIGURE 5.8. The m_T distribution in the three top-background control regions for the *non-res*, *low-mass*, and the *high-mass* selections of the resolved analyses. The signal contamination is negligible, and hence not shown. The lower panel shows the fractional difference between the data and the total expected background with the corresponding statistical and total uncertainty.

Non-resonant: QCD $R_{non-prompt}$ Values				
	mww	bbpt210	bbpt300	wwpt250
Combined	0.74 ± 0.04	0.80 ± 0.05	0.64 ± 0.08	0.62 ± 0.05
Electrons	0.72 ± 0.05	0.75 ± 0.06	0.68 ± 0.10	0.64 ± 0.06
Muons	0.81 ± 0.06	0.92 ± 0.11	0.55 ± 0.15	0.59 ± 0.10
Low-mass (m500): QCD $R_{non-prompt}$ Values				
	mww	bbpt210	wwpt150	hh500
Combined	0.74 ± 0.04	0.80 ± 0.05	0.68 ± 0.03	0.50 ± 0.03
Electrons	0.72 ± 0.05	0.75 ± 0.06	0.65 ± 0.03	0.44 ± 0.04
Muons	0.81 ± 0.06	0.92 ± 0.11	0.75 ± 0.06	0.58 ± 0.07
Low-mass (m700): QCD $R_{non-prompt}$ Values				
	mww	bbpt210	wwpt250	hh700
Combined	0.74 ± 0.04	0.80 ± 0.05	0.62 ± 0.25	0.60 ± 0.03
Electrons	0.72 ± 0.05	0.75 ± 0.06	0.64 ± 0.06	0.52 ± 0.04
Muons	0.81 ± 0.06	0.92 ± 0.11	0.59 ± 0.10	0.74 ± 0.08
High-mass: QCD $R_{non-prompt}$ Values				
	bbpt350	wwpt250	drww15	hh2000
Combined	0.47 ± 0.09	0.62 ± 0.05	0.68 ± 0.03	0.58 ± 0.13
Electrons	0.48 ± 0.10	0.64 ± 0.06	0.60 ± 0.04	0.57 ± 0.14
Muons	0.44 ± 0.20	0.59 ± 0.10	0.82 ± 0.07	0.65 ± 0.39

TABLE 5.20. $R_{non-prompt}$ calculated for the QCD Monte Carlo sample at using each kinematic criteria individually from all selections. The lepton channels are shown separated and combined.

background is in agreement with the data over the entire distribution, and close to the signal region in particular. All simulated backgrounds are normalized according to their theoretical cross-sections, except $t\bar{t}$, which is normalized in the top CRs.

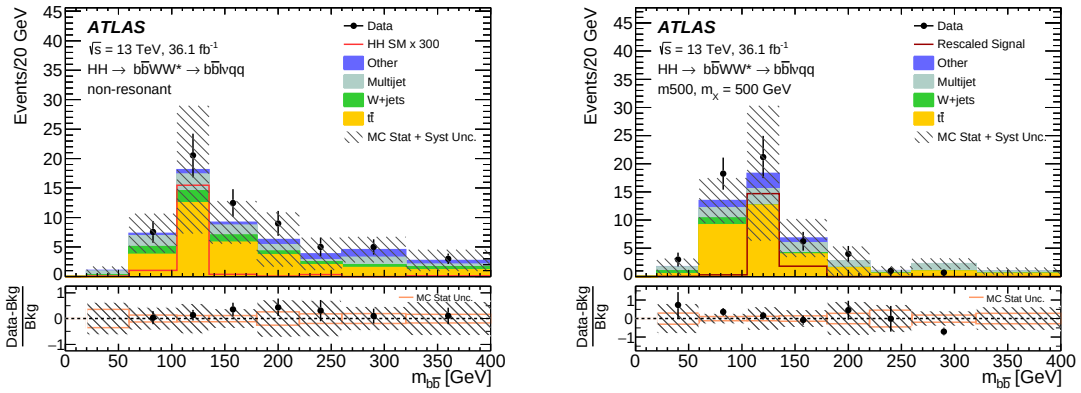

FIGURE 5.9. The $m_{b\bar{b}}$ distribution in the resolved analysis for the *non-res* and *m500* selections at the end of the selection sequence, before applying the $m_{b\bar{b}}$ requirement. The signals shown are from SM non-resonant HH production scaled up by a factor of 300 (left) and from a scalar resonance with mass 500 GeV scaled to the expected upper-limit cross section reported in Section 5.6.1 (right). The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty.

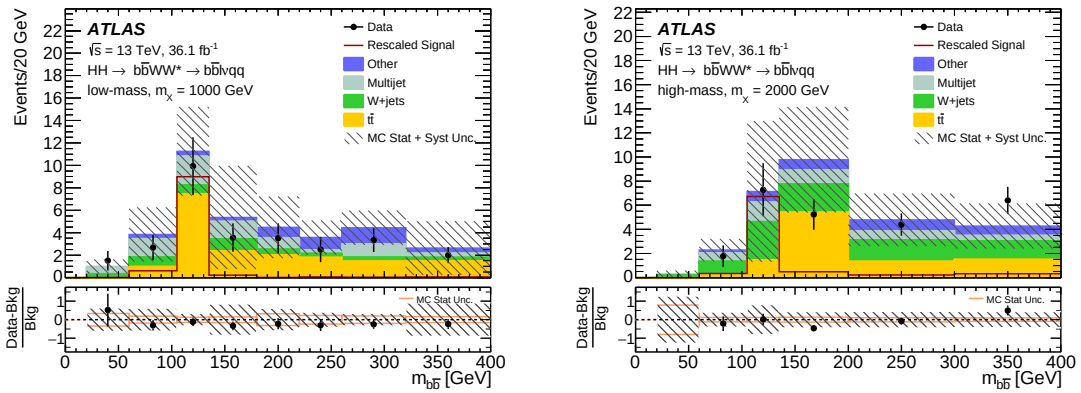

FIGURE 5.10. The $m_{b\bar{b}}$ distribution in the resolved analysis for the *low-mass* and *high-mass* selections at the end of the selection sequence, before applying the $m_{b\bar{b}}$ requirement. The signals shown are from scalar resonances with mass 1000 GeV (left) and 2000 GeV (right) scaled to the expected upper-limit cross section reported in Section 5.6.1. The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty.

SR: $100 < m_{b\bar{b}} < 140$ GeV					
Sample	mww	bbpt210	bbpt300	wwpt250	mbb
$t\bar{t}$	7461.0 ± 48.6	162.9 ± 7.3	27.9 ± 2.9	18.4 ± 2.4	15.4 ± 2.2
QCD	2756.2 ± 210.5	48.7 ± 14.2	6.6 ± 1.9	4.2 ± 1.2	3.6 ± 1.6
Wv221	640.8 ± 12.7	19.1 ± 1.4	5.0 ± 0.6	3.1 ± 0.5	2.3 ± 0.4
SingleTop	452.2 ± 9.6	14.3 ± 1.7	1.7 ± 0.5	1.0 ± 0.4	0.6 ± 0.3
Dibosonsv221	21.6 ± 1.3	0.6 ± 0.2	0.4 ± 0.2	0.0 ± 0.0	0.0 ± 0.0
Zv221	262.8 ± 4.4	3.1 ± 0.3	1.0 ± 0.2	0.2 ± 0.1	0.2 ± 0.1
Background Sum	11594.7± 216.7	248.6± 16.1	42.6± 3.6	27.0± 2.8	22.1± 2.8
XhhSM	68.3 ± 2.4	20.7 ± 0.9	6.7 ± 0.4	5.5 ± 0.3	4.8 ± 0.3
Data	11450.0	232.0	47.0	31.0	22.0

TABLE 5.21. The number of expected background and signal events in the m_{bb} SR for the non-resonant selection. Only statistical uncertainties are shown. No NF has been applied.

SR: $100 < m_{b\bar{b}} < 140$ GeV					
Sample	mww	bbpt210	wwpt150	hh500	mbb
$t\bar{t}$	7461.0 ± 48.6	162.9 ± 7.3	141.7 ± 6.8	17.3 ± 2.2	12.6 ± 1.9
QCD	2756.2 ± 210.5	48.7 ± 14.2	40.2 ± 11.7	3.3 ± 1.0	2.9 ± 1.3
Wv221	640.8 ± 12.7	19.1 ± 1.4	15.3 ± 1.3	0.1 ± 0.0	-0.2 ± 0.1
SingleTop	452.2 ± 9.6	14.3 ± 1.7	12.2 ± 1.6	3.6 ± 0.8	2.8 ± 0.7
Dibosonsv221	21.6 ± 1.3	0.6 ± 0.2	0.5 ± 0.2	0.1 ± 0.0	0.1 ± 0.0
Zv221	262.8 ± 4.4	3.1 ± 0.3	1.9 ± 0.2	0.5 ± 0.1	0.4 ± 0.1
Background Sum	11594.7± 216.7	248.6± 16.1	211.8± 13.7	24.9± 2.5	18.6± 2.4
Xhh500	6.6 ± 0.2	1.9 ± 0.1	1.7 ± 0.1	0.9 ± 0.1	0.8 ± 0.1
Data	11450.0	232.0	194.0	32.0	26.0

TABLE 5.22. The number of expected background and signal events in the m_{bb} SR for the low-mass selection, m500. Only statistical uncertainties are shown. No NF has been applied.

SR: $100 < m_{b\bar{b}} < 140$ GeV					
Sample	mww	bbpt210	wwpt250	hh700	mbb
$t\bar{t}$	7461.0 ± 48.6	162.9 ± 7.3	61.5 ± 4.7	21.9 ± 2.7	15.3 ± 2.2
QCD	2756.2 ± 210.5	48.7 ± 14.2	14.1 ± 4.1	5.5 ± 1.6	4.8 ± 2.2
Wv221	640.8 ± 12.7	19.1 ± 1.4	9.7 ± 1.1	4.1 ± 0.8	2.6 ± 0.6
SingleTop	452.2 ± 9.6	14.3 ± 1.7	2.6 ± 0.7	0.5 ± 0.2	0.3 ± 0.2
Dibosonsv221	21.6 ± 1.3	0.6 ± 0.2	0.2 ± 0.1	0.2 ± 0.1	0.2 ± 0.1
Zv221	262.8 ± 4.4	3.1 ± 0.3	0.6 ± 0.1	0.1 ± 0.0	0.1 ± 0.0
Background Sum	11594.7± 216.7	248.6± 16.1	88.7± 6.4	32.3± 3.2	23.3± 3.1
Xhh700	9.2 ± 0.3	7.8 ± 0.2	5.9 ± 0.2	5.0 ± 0.2	4.4 ± 0.2
Data	11450.0	232.0	75.0	25.0	22.0

TABLE 5.23. The number of expected background and signal events in the m_{bb} SR for the low-mass selection, m700. Only statistical uncertainties are shown. No NF has been applied.

SR: $100 < m_{b\bar{b}} < 140$ GeV					
Sample	bbpt350	wwpt250	drww15	hh2000	mbb
$t\bar{t}$	1307.8 ± 20.2	1024.9 ± 17.7	287.5 ± 9.4	2.2 ± 0.8	1.4 ± 0.6
QCD	207.2 ± 99.5	191.2 ± 29.0	55.2 ± 8.4	2.9 ± 0.4	2.2 ± 0.5
Wv221	341.3 ± 3.4	291.5 ± 3.2	110.7 ± 2.1	4.8 ± 0.3	3.4 ± 0.3
SingleTop	144.1 ± 5.6	126.6 ± 5.3	29.2 ± 2.6	0.5 ± 0.3	0.5 ± 0.3
Dibosonsv221	25.9 ± 1.5	21.8 ± 1.3	6.6 ± 0.7	0.0 ± 0.0	0.0 ± 0.0
Zv221	53.8 ± 0.8	40.4 ± 0.7	13.2 ± 0.4	0.8 ± 0.1	0.7 ± 0.1
Background Sum	2080.1± 101.8	1696.5± 34.6	502.5± 13.1	11.2± 1.0	8.2± 0.8
Xhh2000	21.0 ± 0.4	19.3 ± 0.4	8.4 ± 0.2	3.4 ± 0.1	2.9 ± 0.1
Data	2182.0	1830.0	587.0	11.0	9.0

TABLE 5.24. The number of expected background and signal events in the m_{bb} SR for the high-mass selection, m2000. Only statistical uncertainties are shown. No NF has been applied.

5.4.10. Systematic Uncertainties

This section describes the sources of systematic uncertainties considered in the analysis. These uncertainties are divided into four categories: experimental uncertainties, uncertainties on the data driven background estimation, uncertainties on the modeling of background processes estimated from simulation, and theoretical uncertainties on the signal processes. In the statistical analysis each systematic uncertainty is treated as a nuisance parameter the names of which are defined below. These systematic variations are estimated on the final expected yield in the signal regions.

5.4.10.1. Experimental uncertainties

Each reconstructed object has several sources of uncertainties, each of which are evaluated separately. Wherever possible, we follow the latest available recommendations from the ATLAS combined performance (CP) groups. The leading instrumental uncertainties are the uncertainty on the b-tagging efficiency and the

jet energy scale (JES). The summary of experimental uncertainties is presented in Table 5.25.

Source	Description	Analysis Name
Muons	p_T resolution MS	MUON_MS
Muons	p_T resolution ID	MUON_ID
Muons	p_T scale	MUON_SCALE
Muons	Isolation efficiency SF	MUON_ISO_SYS
Muons	Isolation efficiency SF	MUON_ISO_STAT
Muons	Reconstruction efficiency SF	MUON_EFF_SYS
Muons	Reconstruction efficiency SF	MUON_EFF_STAT
Muons	Trigger efficiency SF	MUON_EFF_TrigStatUncertainty
Muons	Trigger efficiency SF	MUON_EFF_TrigSystUncertainty
Electrons	p_T resolution	EG_RESOLUTION_ALL
Electrons	p_T scale	EG_SCALE_ALL
Electrons	Isolation efficiency SF	EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR
Electrons	Reconstruction efficiency SF	EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR
Electrons	Trigger efficiency SF	EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR
Electrons	Identification efficiency SF	EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR
Tau	Energy scale model	TAUS_TRUEHADTAU_SME_TES_MODEL
Tau	Energy scale detector	TAUS_TRUEHADTAU_SME_TES_DETECTOR
Tau	In-situ energy calibration	TAUS_TRUEHADTAU_SME_TES_INSITU
MET	Soft term	MET_SoftTrk_ResoPerp
MET	Soft term	MET_SoftTrk_ResoPara
MET	Soft term	MET_SoftTrk_Scale
Small-R Jets	JES strongly reduced	JET_SR1_JET_GroupedNP_1
Small-R Jets	JES strongly reduced	JET_SR1_JET_GroupedNP_2
Small-R Jets	JES strongly reduced	JET_SR1_JET_GroupedNP_3
Small-R Jets	JES strongly reduced	JET_SR1_JET_EtaIntercalibration_NonClosure
Small-R Jets	Energy resolution	JET_JER_SINGLE_NP
Small-R Jets	JVT efficiency SF	JET_JvtEfficiency
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_0
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_1
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_2
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_3
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_4
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_5
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_6
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_7
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_8
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_9
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_10
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_11
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_12
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_13
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_0
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_1
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_2
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_3
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_4
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_0
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_1
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_2
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_3
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_4
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_extrapolation
b-tagging	Flavor tagging scale factors	FT_EFF_Eigen_extrapolation_from_charm

TABLE 5.25. Qualitative summary of the object systematic uncertainties included in this analysis.

5.4.10.1.1. Luminosity The uncertainty in the combined 2015+2016 integrated luminosity is 2.1%[8]. It is derived, following a methodology similar to that detailed in [26], from a preliminary calibration of the luminosity scale using x-y beam-separation scans performed in August 2015 and May 2016. The luminosity uncertainty is applied to those backgrounds estimated from simulation and the signal samples.

5.4.10.1.2. Trigger Systematic uncertainties on the efficiency of electron and muon triggers are evaluated as recommended by the corresponding combined performance groups as documented here.[9]

5.4.10.1.3. Muons The following systematic uncertainties are applied to muons in estimations based on the simulation:

- Identification efficiency: The efficiencies are measured with the tag and probe method using the Z mass peak.
- Energy and Momentum scales: These are also measured with Z mass line shape, and provided by the CP groups.

5.4.10.1.4. Electrons The following systematic uncertainties are applied to electron in estimations based on the simulation:

- Identification efficiency: The efficiencies are measured with the tag and probe method using the Z mass peak. They include contributions from reconstruction, identification and isolation;

[8] `https://twiki.cern.ch/twiki/bin/view/Atlas/LuminosityForPhysics`

[9] el:`https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronEfficiencyRun2`
mu:`https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC15`

– Energy and Momentum scales: These are also measured with Z mass line shape, and provided by the CP groups.

5.4.10.1.5. Jet uncertainties The jet energy uncertainties are derived based on in situ measurements performed during Run 2 conditions [89]. The jet energy resolution uncertainty is evaluated by smearing jet energies according to the systematic uncertainties of the resolution measurement [90]. The uncertainty in the b-tagging efficiency is evaluated by propagating the systematic uncertainty in the measured tagging efficiency for b-jets [91]. The "Loose" reduction scheme is used.

5.4.10.1.6. Missing transverse energy The systematic uncertainties related to the missing transverse energy are obtained by the propagation of the systematic uncertainty on the objects that build the MET, in particular the muon, electron and jets energy resolution and scale. The resolution and scale of the MET soft-term is broken down into its components: METScale, METResoPara, METResoPerp, and full uncertainties from each component is taken into account in the final fit.

5.4.10.1.7. d_0^{sig} uncertainties The uncertainty due to the d_0^{sig} criteria has been evaluated by making the ratio between the efficiency of the criteria for data and the efficiency of the criteria for the MC background samples. We selected di-electron or di-muon event, requiring an invariant di-leptons mass within 80-100 GeV Z Mass window. To be as similar to our signal region as possible but to keep high statistics, loose pre-selection criteria are applied in selecting the events. The leading lepton is required to have $p_T > 27$ GeV and $\not{E}_T > 25$ GeV. At least four resolved jets are required of which exactly two are b-jets. The d_0^{sig} distributions for data and MC samples for each lepton channel are shown in the Figure 5.11, in Figure 5.12 the relative ratio of the total distributions is shown. The ratio of the efficiency of the

d_0^{sig} criteria for data over MC samples si about 96%, this is equivalent if the ratio is estimated by using only muons or only electrons. The difference of this ratio from one is the fractional uncertainty due to the d_0^{sig} criteria efficiency.

This results in about 4% for the d_0^{sig} uncertainty independent from the lepton flavor.

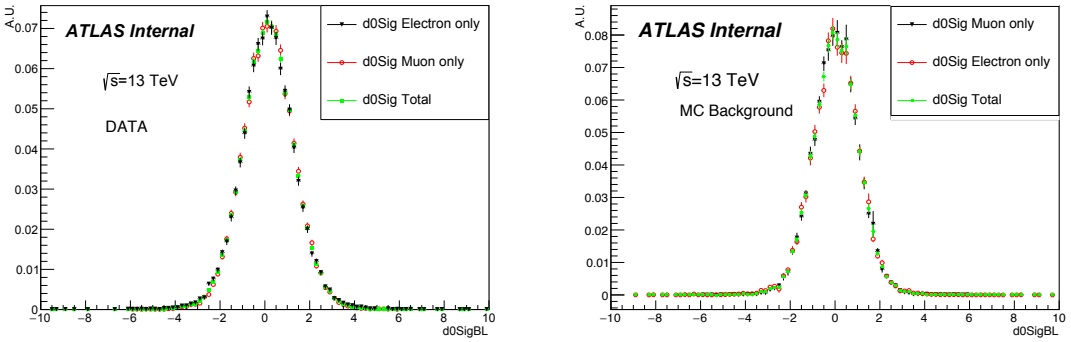

FIGURE 5.11. d_0^{sig} distributions for data and background MC samples, identifying the lepton channel.

5.4.10.2. Background modeling uncertainties

Several systematics have been evaluated to take into account the uncertainties on the modeling of backgrounds.

5.4.10.2.1. Uncertainties from the modeling of $t\bar{t}$ The dominant background $t\bar{t}$ is normalized in dedicated CRs. MC is used to extrapolate the shapes from the control regions to the signal region, so theoretical uncertainties are related to such extrapolation. PDF and scale uncertainties are evaluated by applying event selection at truth level. The resulting uncertainties are approximately 5% to 6% and are included in the final fit.

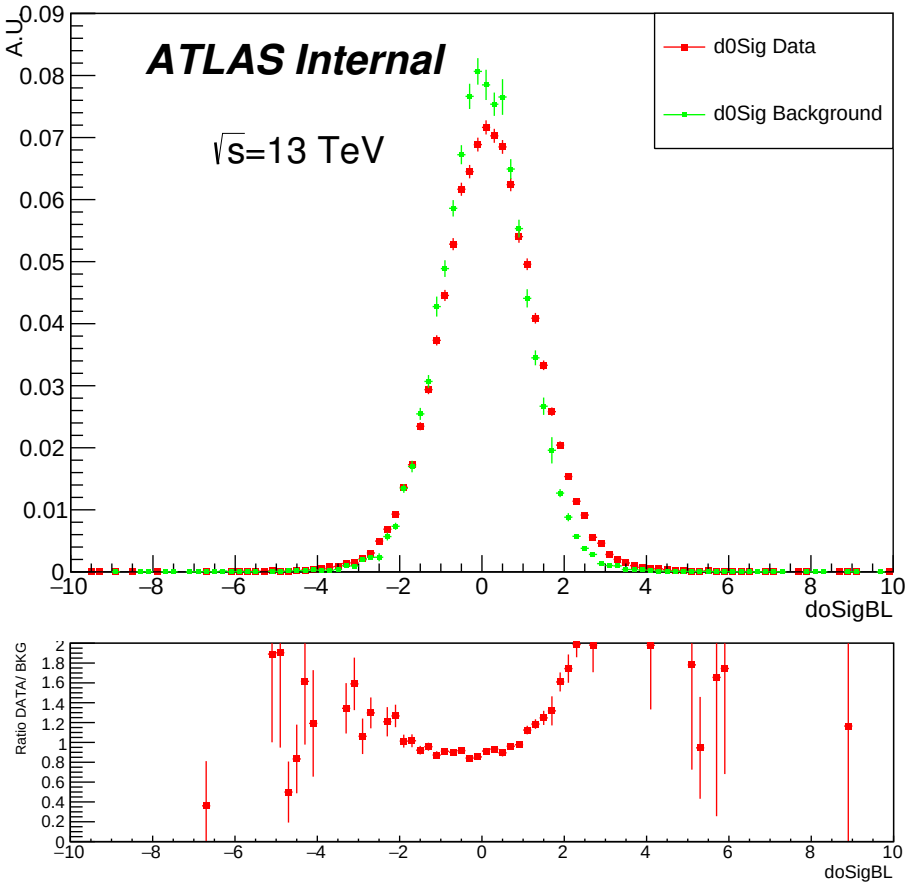

FIGURE 5.12. d_0^{sig} distributions and the relative ratio for data and background MC samples.

Additional uncertainties in $t\bar{t}$ modeling stems from the difference in the matrix element (ME) implementation across generators, hadronization and fragmentation modeling (called parton shower, PS), and the amount of initial and final state radiation (ISR/FSR). The ME uncertainty is computed by comparing the events generated by aMC@NLO with the events generated by Powheg-Box v2, both interfaced to Herwig++ for parton shower. The difference computed close to the signal region with enough statistics is used. The PS uncertainty is computed

Source	Non-resonant	Low-Mass	High-Mass
Matrix Element	7.2	0.5	4.1
Parton Shower	3.7	16.4	9.5
ISR/FSR	14.7	4.9	8.2
PDF	5.2	3.5	6.1
Scale	3.3	2.2	3.7

TABLE 5.26. Extrapolation uncertainties in percentage from the CR to the SR for all selections, provided to the fit for the $t\bar{t}$ modeling systematics.

by comparing the the nominal POWHEG+PYTHIA6 sample with the PS variation POWHEG+HERWIG++ sample in a region close to the SR but with enough statistics. For ISR/FSR, the dedicated RADHI and RADLO samples with modified HDAMP parameter are compared. The sample with the higher impact on the fit is kept as the uncertainty due to ISR/FSR. Table 5.26 shows the numbers provided to the fit for the various $t\bar{t}$ modeling systematics for the *low-* and *high-mass* selections. Since the uncertainties are computed based on extrapolation from the CRs, these uncertainties are input to the statistical fit as rate uncertainties in the SR only.

5.4.10.2.2. Single top uncertainty Theoretical cross section uncertainties of 5.3% is assigned to the associated Wt production, 3.9% to the s-channel and 4.2% to t-channel single top production. The dominant production for this analysis is the Wt channel. The single top modeling systematic uncertainties have been calculated employing the difference between the nominal (DR scheme) and the (available) systematic variation sample (DS scheme) in SR for the Wt channel. The uncertainties are computed to be 48%, 48%, and 84% for non-resonant, low-mass and high-mass analyses respectively. Single top is a very small background in the analysis. However, the large difference between the two schemes warranted getting feedback from the Top Group. We narrowed the huge difference down to very tight $p_T(bb)$. See Figure 5.13. The Top Group is working on a better prescription but in the meantime this is

the recommended method and the huge difference is what we keep. Additional uncertainties on single top have been calculated employing the difference between the nominal and the (available) systematic variation samples. The recommendation is taken from the Top Twiki.[10] The comparison is done between the nominal sample and the Powheg+Herwig for fragmentation, MC@NLO for Matrix Element, and RadHi/RadLo for ISR/FSR uncertainties. The comparison is done at the selection where MC statistical uncertainties are small. This leads to $p_T(bb) < 120$ for non-resonance and low-mass selections, and $\Delta R(WW) < 1.5$ for high-mass selection. The uncertainties vary across selections being 3.5% at the smallest to 23% at the largest.

5.4.10.2.3. W/Z+jets modeling uncertainty Uncertainties on the modeling of W+jets background were computed in each SR and top CR. Three sources of uncertainties were considered: scale variation uncertainties, PDF uncertainties and PS/modeling uncertainties. Scale uncertainties were computed by varying by a factor of two the nominal renormalization and factorization scales, PDF uncertainties were computed according to the NNPDF recipe, that is computing the standard deviation of the 100 PDF eigenset, while modeling uncertainties were computed comparing Sherpa with Alpgen+Pythia6. The values obtained in each region are summarized in Table 5.27. Same uncertainties are used for the subleading Z+jets background.

5.4.10.2.4. QCD uncertainty An overall 36% uncertainty is assigned to the QCD multijet background. The systematic uncertainty was calculated following the steps described in Section 5.4.8.

[10] `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopSystematics2015`

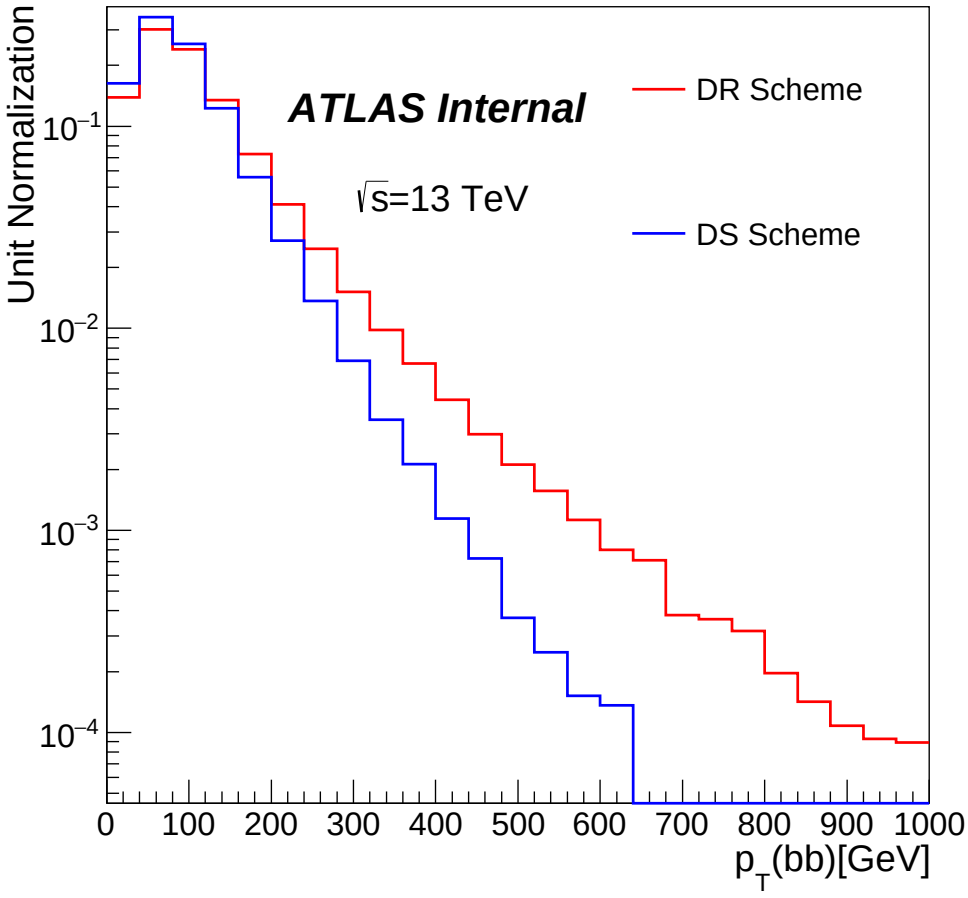

FIGURE 5.13. $p_T(bb)$ for DR and DS schemes for single top modeling.

5.4.10.3. Model uncertainties on the signal

Systematics on signal acceptance are computed generating multiple-weight samples that include weights corresponding to variation of the normalization and factorization scales by a factor $(\xi_r, \xi_f) = 2$. The envelope is built excluding the cases where $\xi_r/\xi_f = 4$ and $\xi_r/\xi_f = 1/4$. The fractional uncertainty is obtained dividing 1/2 of the envelop by the central value. PDF uncertainties are compute using PD4LHC_mc_30 pdf sets, that include the envelope of three PDF sets, namely

Source	Non-resonant		Low-Mass		High-Mass	
	SR	CR	SR	CR	SR	CR
modeling/PS	37	37	37	37	18	18
PDF	7	30	10	36	17	31
Scale	25	35	17	31	29	28

TABLE 5.27. Theoretical uncertainties in percentage on W/Z+jets event yield computed in the CR and the SR of all selections, provided to the fit for the W/Z jets modeling systematics.

Signal model	Non-resonant	$m_H 500 - 1300$ GeV	$m_H > 1300$ GeV
scale	$\pm 1.1\%$	± 0.8 %	± 0.7 %
PDF	$\pm$ 1.3%	± 1.3 %	$\pm$ 1.3 %

TABLE 5.28. Theoretical uncertainties in percentage on the signal acceptance.

the CT14, MMHT'14, NNPDF3. The CTEQ error formula is used to compute the uncertainties. Results are summarized in Table 5.28 for each signal hypothesis.

5.5. Boosted Analysis

5.5.1. Event Reconstruction

For the boosted analysis, events are reconstructed by requiring at least one reconstructed lepton. This lepton will be referred to as the selected lepton. To reconstruct the $H \to b\bar{b}$ candidate, there should be at least one large-R jet with $\Delta R > 1.0$ from the selected lepton. The highest p_T large-R jet is chosen as the $H \to b\bar{b}$ candidate. The large-R jet is then required to have two at least two track jets associated to it. Events with the large-R jet mass to be in the range of 30 GeV $< m_{\text{large-R jet}} <$ 300 GeV are retained for further analysis.

In order to reconstruct $W \to q\bar{q}$ candidate, events are required have at least two signal small-R jets with $\Delta R > 1.4$ of the $H \to b\bar{b}$ candidate. The ΔR requirement ensures that the small-R jets used for $W \to q\bar{q}$ reconstruction do not overlap with

the $H \to b\bar{b}$ candidate large-R jet. If there are exactly two signal small-R jets, they are used to reconstruct the $W \to q\bar{q}$ candidate. If there are at least three signal small-R jets, the $W \to q\bar{q}$ candidate is reconstructed from the combination of pairs of the three highest p_T small-R jets with the smallest ΔR between the small-R jets.

The full $h \to WW^* \to l\nu qq$ is reconstructed identically to the resolved analysis as described in Section 5.4.3. With the $H \to b\bar{b}$ candidate identified and $h \to WW^* \to l\nu qq$ system fully reconstructed, the di-Higgs (HH) system is reconstructed by the sum of four-momenta of the $H \to b\bar{b}$ candidate large-R jet and the reconstructed $h \to WW^* \to l\nu qq$ system. Figure 5.14 shows a diagram of the event topology after the event reconstruction.

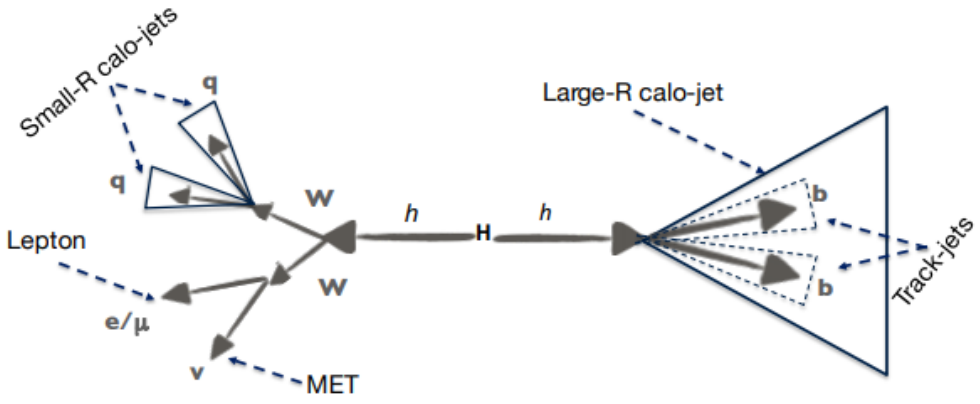

FIGURE 5.14. Diagram of the reconstructed event topology

5.5.2. Event Selection

After the event is reconstructed, a b-jet veto is applied on the event by requiring all signal small-R jets do not pass b-tagging requirement to reject $t\bar{t}$ events. The two highest p_T track-jets in the large-R jet are required to be b-tagged and events that

passes this requirement is considered to be in the "2-tag" region. In addition, the $\not{E}_T$ is required to be more than 50 GeV to reject events from QCD multijet background.

5.5.3. Kinematic Selection

5.5.3.1. Signal Region

In order to enhance the sensitivity to a resonant HH signal, similarly to the resolved analysis, it is required that the $H \to b\bar{b}$ candidate large-R jet has a mass consistent with the Standard Model Higgs boson mass. Events which have the $H \to b\bar{b}$ candidate large-R jet mass in a window of 90 GeV $< m_{\text{Large-Rjet}} <$ 140 GeV is considered to be in the signal region (SR). The signal region is blinded and the blinding strategy is implemented by removing any events in data that passes the signal region requirement on the large-R jet mass.

5.5.3.2. mBB Control Region

In order to asses the modeling of the backgrounds, a control region is defined to be any events which fails the large-R jet signal region mass window requirement. Any event which has a large-R jet mass 90 GeV $< m_{\text{Large-Rjet}}$ or $m_{\text{Large-Rjet}} >$ 140 GeV falls in the mBB control region (mBBcr). By construction, this region is orthogonal to the signal region.

5.5.4. Multijet Background

As with the resolved analysis, the QCD multijet background is estimated using the same data-driven method with slight modifications.

The ABCD method uses three control regions (the B, C, and D regions) to estimate the contribution of a given background in the signal (A) region. Cuts on

two ideally orthogonal variables are used to create the signal and various control regions, e.g. the A region passes both cuts, the B and C regions each pass one cut and fail the other, while the D region fails both cuts.

For the boosted analysis, the ABCD regions are defined by using the same variables as in the resolved analysis, which are the significance of the lepton impact parameter (d_0^{sig}) and the missing transverse momentum ($\not{E}_T$). A small difference would be the higher cut value for the $\not{E}_T$ with respect to the resolved analysis. The regions are defined as follow:

- Region A: $\not{E}_T >$ 50 GeV, $|d_0^{\text{sig}}| < 2.0$
- Region B: $\not{E}_T <$ 50 GeV, $|d_0^{\text{sig}}| < 2.0$
- Region C: $\not{E}_T >$ 50 GeV, $|d_0^{\text{sig}}| > 2.0$
- Region D: $\not{E}_T <$ 50 GeV, $|d_0^{\text{sig}}| > 2.0$

Figure 5.15 shows the four regions represented on on the lepton d_0 significance vs $\not{E}_T$ plane. Assuming that the two variables chosen to define the ABCD regions are completely uncorrelated, the QCD multijet yield in region A can be predicted. The The correlation between $|d_0^{\text{sig}}|$ vs $\not{E}_T$ is estimated in multiple MC background samples and also in data, and they are found to negligible.

The ABCD method is done separately between the muon and electron channel as it is expected that the $\frac{N_B^{\text{QCD}}}{N_D^{\text{QCD}}}$ ratio and QCD multijet contribution to the total predicted background will be significantly different between the channels.

5.5.4.1. Yield prediction

Table 5.29 lists the MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region B and D before the large-R jet mass is

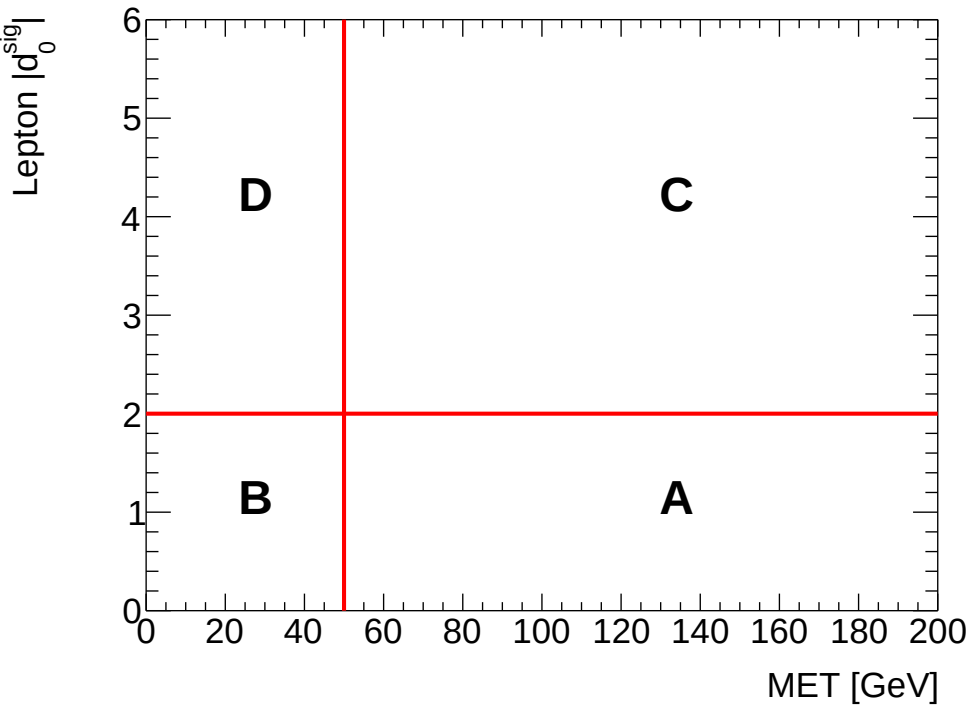

FIGURE 5.15. Regions defined in the ABCD method based on the lepton d_0 significance vs $\not{E}_T$ plane. Region A is the signal enriched region which we want to estimate the multijet background. Region C is where the shape template is derived from and used as shape prediction of the multijet background in region A. The ratio of the multijet yields in region B to region D is used to scale the multijet yield in region C to predict the multijet background yield in region A.

applied and Table 5.30 shows the yields in Region C mBB control region and signal region.

The $\frac{N_B^{\text{QCD}}}{N_D^{\text{QCD}}}$ ratio are calculated inclusively in the large-R jet mass distribution. In other words, the ratio is calculated without the jet mass window selection, which defines the SR and mBBcr, applied. The ratio is then used to scale the QCD multijet yield in the SR and mBBcr of region C to predict the QCD multijet yield in region A.

Table 5.31 shows the ratio in the electron channel and muon channel. As expected, the electron channel ratio is larger than the muon channel. The muon

	Region B		Region D	
Samples	Electron	Muon	Electron	Muon
$t\bar{t}$	307.7 ± 11.5	279.8 ± 10.6	21.3 ± 2.7	18.3 ± 2.6
W+Jets	173.2 ± 5.2	179.3 ± 5.6	11.6 ± 1.4	10.6 ± 1.1
Single-top	42.9 ± 3.4	33.5 ± 3.6	3.0 ± 0.9	0.8 ± 0.5
Z+Jets	78.5 ± 1.9	72.5 ± 1.7	6.4 ± 0.6	5.5 ± 0.5
Dibosons	19.1 ± 1.5	17.7 ± 1.5	1.6 ± 0.4	2.2 ± 0.8
Total Prompt	621.3 ± 13.2	582.7 ± 12.7	44.0 ± 3.3	37.4 ± 3.0
Data	1003 ± 31.7	711 ± 26.7	144 ± 12.0	98 ± 9.9
QCD	381.7 ± 34.3	128.3 ± 29.5	100.0 ± 12.4	60.6 ± 10.4

TABLE 5.29. MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region B and D. The multijet yield is calculated by subtracting the estimated total prompt lepton backgrounds from the observed data. The statistical uncertainty on the yields is shown.

	mBBcr		SR	
Samples	Electron	Muon	Electron	Muon
$t\bar{t}$	38.7 ± 4.2	46.8 ± 7.9	28.5 ± 3.1	22.0 ± 2.7
W+Jets	22.3 ± 2.0	20.0 ± 1.7	9.6 ± 1.3	10.0 ± 1.8
Single-top	7.6 ± 2.1	6.5 ± 1.3	7.1 ± 1.5	2.7 ± 0.8
Z+Jets	4.6 ± 0.8	3.8 ± 0.5	1.6 ± 0.3	1.9 ± 0.6
Dibosons	2.2 ± 0.6	1.2 ± 0.4	0.8 ± 0.3	1.7 ± 0.4
Total Prompt	75.4 ± 5.2	78.4 ± 8.2	47.5 ± 3.7	38.3 ± 3.4
Data	148 ± 12.2	126 ± 11.2	91 ± 9.5	71 ± 8.4
QCD	72.6 ± 13.2	47.6 ± 13.9	43.5 ± 10.2	32.7 ± 9.1

TABLE 5.30. MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region C mBBcr and SR. The multijet yield is calculated by subtracting the estimated total prompt lepton backgrounds from the observed data. The statistical uncertainty on the yields is shown.

Multijet yield in region	Electron	Muon
N_B^{QCD}	381.7 ± 34.3	128.3 ± 29.5
N_D^{QCD}	100.0 ± 12.4	60.6 ± 10.4
$N_B^{\mathrm{QCD}}/N_D^{\mathrm{QCD}}$	3.8 ± 0.6 (15.3%)	2.1 ± 0.6 (28.7%)

TABLE 5.31. Multijet yields in region B and region D and also the ratio of the yields for each lepton channel. The error on the $\frac{N_B^{\mathrm{QCD}}}{N_D^{\mathrm{QCD}}}$ ratio is propagated from the statistical uncertainties on the multijet yields in each region.

channel ratio has a larger uncertainty due to the more limited statistics in region B and region D compared to the electron channel. The predicted yields of the QCD multijet background in the mBB control region and signal region are presented in Table 5.32. The QCD multijet background is estimated to be 19% of the total background in the signal region (Table 5.52).

Multijet yield in region	Electron	Muon
SR		
N_C^{QCD}	43.4 ± 10.2	32.7 ± 9.1
N_A^{QCD}	165.9 ± 46.6 (28.1%)	69.3 ± 27.7 (39.9%)
mBBcr		
N_C^{QCD}	72.6 ± 13.2	47.6 ± 13.9
N_A^{QCD}	277.1 ± 66.0 (23.8%)	100.8 ± 41.3 (41.0%)

TABLE 5.32. Multijet yield in region C and predicted yield in region A in the SR. The error on N_A^{QCD} are propagated from the error on the $N_B^{\mathrm{QCD}}/N_D^{\mathrm{QCD}}$ ratio and statistical uncertainty on N_C^{QCD} yield. The numbers in brackets are the relative uncertainty in percentage.

5.5.4.2. Shape prediction

In order to predict the shape of the HH mass distribution (and also other kinematic distribution) of the QCD multijet background, the shape template of all kinematic distributions are obtained by subtracting all the MC backgrounds from data in region C.

It was found that distributions in region C suffer from lack of statistics due the low number of data events which results in shape templates with severe statistical fluctuations. To overcome this, the shape templates are derived from a sample of 1 b-tagged (1-tag) events in region C. This sample requires that one of the two leading track-jet is b-tagged but not both at the same time.

5.5.4.3. Multijet yield uncertainties

5.5.4.3.1. Statistical The uncertainty on the predicted yield of the multijet background is determined by propagating the statistical uncertainty of the $\frac{N_B^{\mathrm{QCD}}}{N_D^{\mathrm{QCD}}}$ ratio, as shown in Table 5.31, and the statisical uncertainty on the multijet yield in region C (N_C^{QCD}), as in Table 5.32.

5.5.4.3.2. 1-tag/2-tag jet mass acceptance Another source of uncertainty on the multijet yield is the the difference of acceptance of the large-R jet mass cut between 1-tag and 2-tag. This uncertainty is included since the template for or the multijet shape prediction uses the multijet shape from the 1-tag region C. Table 6.6 shows the acceptance of the large-R jet mass signal and mBB control region selection in the multijet 1-tag region C and 2-tag region C yields. The relative difference between the acceptance in 1-tag region C and in 2-tag region C is considered as an uncertainty on the normalization of the QCD multijet prediction.

5.5.4.3.3. mBB control region fit A likelihood fit of the large-R jet mass in the mBB control region is performed with the normalization of the multijet background to be unconstrained in the fit. In the muon channel, the post-fit normalization factor is consistent with unity and in the electron channel, the normalization factor is 0.676 $\pm$ 0.130. Due to this significant deviation from unity for the electron channel, we

Region	Electron	Muon
SR		
1-tag $\frac{N_{\mathrm{SR}}}{N_{\mathrm{Inc}}}$	31.6 ± 2.7 %	27.9 ± 2.7 %
2-tag $\frac{N_{\mathrm{SR}}}{N_{\mathrm{Inc}}}$	37.5 ± 9.7 %	40.7 ± 9.7 %
Rel. difference between 1-tag and 2-tag	15.6 %	31.5 %
mBBcr		
1-tag $\frac{N_{\mathrm{mBBcr}}}{N_{\mathrm{Inc}}}$	68.4 ± 2.7 %	72.1 ± 2.7 %
2-tag $\frac{N_{\mathrm{mBBcr}}}{N_{\mathrm{Inc}}}$	62.5 ± 7.0 %	59.3 ± 7.0 %
Rel. difference between 1-tag and 2-tag	9.4 %	21.6 %

TABLE 5.33. The acceptance of the large-R jet mass signal region selection on the multijet 1-tag and 2-tag region C. $N_{\mathrm{SR}}(N_{\mathrm{Inc}})$ is the multijet yield with (without) the signal region large-R jet mass selection.

assign a normalization uncertainty of 32.4% for the multijet background in both the mBB control region and signal region.

5.5.4.3.4. $t\bar{t}$ and W+jets MC modeling The uncertainties on the MC modeling of the $t\bar{t}$ and W+jets, the two largest prompt background predicted by MC in region B,D and C are taken as a systematic on the predicted multijet background in region A since the multijet background is calculated by subtracting the prompt background from observed data. The uncertainties on the normalization of $t\bar{t}$ and W+jets in each region are calculated by comparing the yields between the nominal $t\bar{t}$ and W+jets samples with their alternative samples (see Section 5.5.10.2.2 and 5.5.10.2.3).

The uncertainty on the multijet yield prediction in region A is then calculated by recalculating the multijet yield in each region with the $t\bar{t}$ and W+jets yields be varied up and down, simultaneously in all regions, by the uncertainty due to the MC modeling of the background. The resulting multijet yield prediction in region A for each background uncertainty is then compared to the nominal prediction in region A and the difference is then taken as the uncertaintyon the multijet yield prediction in

region A. Table 5.34 shows the uncertainty on the multijet yield prediction in region A signal and mBB control region due to the the uncertainty on the $t\bar{t}$ and W+jets MC modeling.

	Electron	Muon
SR		
$t\bar{t}$	26.5 %	60.1 %
W+jets	24.7 %	70.4 %
mBBcr		
$t\bar{t}$	37.4 %	101.0%
W+jets	29.5 %	77.6%

TABLE 5.34. The uncertainty on the multijet yield prediction in region A due to the normalization uncertainty of the $t\bar{t}$ and W+jets backgrounds in region C.

5.5.4.3.5. Detector modeling of prompt backgrounds The detector modeling systematic uncertainties on the prompt background in regions B, D and C are propagated through the ABCD method to estimate the uncertainty on the multijet yield prediction region A. Table 5.35 shows the uncertainties on the predicted multijet yield in both lepton channels.

	Electron	Muon
SR		
Total Uncertainty	46.0%	105.6%
mBBcr		
Total Uncertainty	45.5%	127.3 %

TABLE 5.35. The total uncertainty on the multijet yield prediction in region A due to the detector modeling uncertainties of the prompt backgrounds in region B, C and D.

5.5.4.3.6. $|\mathbf{d_0^{sig}}|$ cut efficiency modeling The $|d_0^{\text{sig}}|$ cut efficiency modeling uncertainty for the prompt MC backgrounds are also taken into account in the

ABCD method due to the tigther cut on the leptons' $|d_0^{\text{sig}}|$, than the recommended value. The determination of the $|d_0^{\text{sig}}|$ cut efficiency modeling uncertainty is discussed in Section 5.5.10.1.1. This uncertainty is propagated through the ABCD method by varying the normalization of the prompt backgrounds in regions B, D and C simultaneously to estimate the uncertainty on the multijet yield prediction region A and is treated as anti-correlated between regions B and regions D,C. Table 5.36 shows the uncertainties on the predicted multijet yield due $|d_0^{\text{sig}}|$ cut efficiency modeling uncertainty on prompt MC backgrounds in both lepton channels.

	Electron	Muon
SR		
Total Uncertainty	46.4%	50.9%
mBBcr		
Total Uncertainty	42.0%	110.6%

TABLE 5.36. The total uncertainty on the multijet yield prediction in region A due to the $|d_0^{\text{sig}}|$ cut efficiency modeling uncertainties of the prompt backgrounds in region B, D and C.

5.5.4.3.7. Total uncertainty on the yield The total uncertainty on the multijet prediction is calculated as the sum in quadrature of the uncertainties from the two sources explained above. Table 5.37 summarizes the systematic uncertainties on the predicted multijet yield in the electron channel and the muon channel.

5.5.4.4. Multijet shape uncertainties

The uncertainty on the m_{HH} distribution prediction for the QCD multijet background is estimated by comparing the m_{HH} distribution shape in the 2-tag region C and 1-tag region C (Figure 5.16).

Source of uncertainty	Electron	Muon
SR		
Statistical	28.1 %	39.9 %
2-tag/1-tag jet mass acceptance	15.6 %	31.5 %
mBB control region fit	32.3 %	-
$t\bar{t}$ MC modeling	26.5 %	60.1%
W+jets modeling	24.7 %	70.4%
Detector modeling of prompt backgrounds	46.0 %	105.6%
$\lvert d_0^{\mathrm{sig}}\rvert$ cut efficiency	46.4 %	50.9%
Total	87.5 %	157.8 %
mBBcr		
Statistical	23.8 %	41.0 %
2-tag/1-tag jet mass acceptance	9.4 %	21.6 %
mBB control region fit	32.3 %	-
$t\bar{t}$ MC modeling	37.4 %	101.0%
W+jets MC modeling	29.5 %	77.6%
Detector modeling of prompt backgrounds	45.5 %	127.3%
$\lvert d_0^{\mathrm{sig}}\rvert$ cut efficiency	42.0 %	110.6%
Total	88.3 %	216.4 %

TABLE 5.37. Summary of systematic uncertainties on the QCD multijet yield in the signal region and mBB control region for each lepton channel. The total uncertainty calculated by adding in quadrature the uncertainties from all sources.

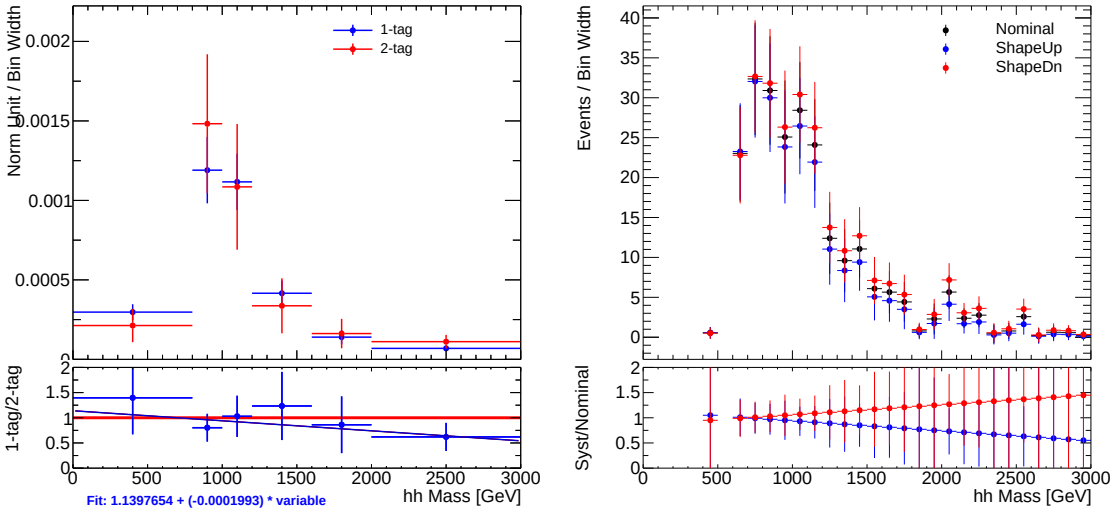

FIGURE 5.16. (a) Shape comparison of the m_{HH} distribution in region C of the 1-tag and 2-tag region. The linear fit to the ratio is used as the shape systematic of QCD background prediction. (b) The up and down QCD shape systematic variation for the predicted m_{HH} distribution.

5.5.4.5. Final prediction and validation

The predicted multijet yield and its uncertainty in the signal region and mBB control region are shown in Table 5.38.

Region	Electron	Muon	Combined
SR	165.9 ± 145.2(87.5%)	69.3 ± 109.3(157.8%)	235.2 ± 181.8 (77.3%)
mBBcr	277.1 ± 244.8(88.3%)	100.8 ± 218.1(216.4%)	377.9 ± 327.8 (86.8%)

TABLE 5.38. Predicted multijet yield with uncertainties in the signal region (SR) and mBB control region (mBBcr) for each lepton channel.

The ABCD method is validated by assessing the agreement of data and total background prediction with the QCD multijet background included in the mBB control region. This is discussed in Section 5.5.9.

5.5.5. $t\bar{t}$

MC simulation is used to model the shape of the $t\bar{t}$ background. The background is predicted to be the largest ($\sim 52\%$) of the total background. A top-enriched control region is used to validate the modeling of the $t\bar{t}$ background.

5.5.6. V+jets

MC simulation is used to model the shape and predict the yield of the W+jets and Z+jets background. The W+jets background is predicted to be the third largest background ($\sim 18\%$) of the total background while the Z+jets background is expected to be $\sim 2\%$ in the signal region (Table 5.52).

5.5.7. Single Top

MC simulation is used to model the shape and predict the yield of background from single-top processes. This background is predicted to be $\sim 9\%$ of the total background (Table 5.52). The single-top background is predominantly consist of Wt production process ($\sim 90\%$), followed by the t-channel production ($\sim 9\%$) and the s-channel production ($\sim 1\%$).

5.5.8. Diboson

MC simulation is used to model the shape and predict the yield of the diboson background. This background is predicted to be 2% of the total background in the signal region (Table 5.52).

5.5.9. Data/Prediction comparisons in control regions

To asses the modeling of the background, the total predicted background is asses with data in the mBB control region. The electron channel and the muon channel are combined into a single channel.

Table 5.39 shows the predicted yield of each background. The data, collected in 2015+2016, corresponding to the integrated luminosity of 36.1 fb^{-1} are used. The MC backgrounds $t\bar{t}$, W+jets, Single-top, Z+jets and Dibosons are normalized to luminosity. The QCD multijet prediction is estimated from the ABCD method. Statistical errors are shown for the individual backgrounds and the total predicted background while the error due to systematic uncertainties are shown only for the total predicted background. Detector modeling uncertainties, MC background modeling uncertainties and uncertainties from the ABCD method for QCD multijet background are considered. The predicted total background yield has an error of about 27% due to systematic uncertainties and the observed yield in data is in good agreement with the total background yield.

Figure 5.17, 5.18, 5.19, 5.20, 5.21, 5.22, 5.23 and 5.24 show the distributions of kinematic variables for events which fall into the mBB control regions. The observed data (black circle) corresponds to an integrated luminosity of 36.1 fb^{-1}. The MC backgrounds $t\bar{t}$ (orange), W+jets (blue), Single-top (red), Z+jets (green) and Dibosons (yellow) are normalized to cross-section prediction scaled to luminosity of 36.1 fb^{-1}. The QCD multijet background (grey) is predicted from the ABCD method. The hashed grey band is the statistical uncertainty on the predicted background and the red box band is the statistical+systematics uncertainty on the predicted background. The systematics uncertainty on the predicted background consists of the

Sample	Yield	Stats Unc	Systs Unc
$t\bar{t}$	1005.6	± 20.6	$^{+283.6(+28.2\%)}_{-288.8(-28.7\%)}$
W+Jets	565.6	± 10.3	$^{+277.9(+49.1\%)}_{-270.0(-47.7\%)}$
QCD	377.9	± 19.6	$^{+328.0(+86.8\%)}_{-328.0(-86.8\%)}$
Single-top	161.3	± 7.2	$^{+114.4(+70.9\%)}_{-114.4(-70.9\%)}$
Z+Jets	55.9	± 1.6	$^{+27.7(+49.5\%)}_{-27.2(-48.6\%)}$
Dibosons	39.7	± 2.6	$^{+23.4(+58.9\%)}_{-23.3(-58.7\%)}$
Prediction	2206.0	± 31.2	$^{+593.7(+26.9\%)}_{-586.1(-26.6\%)}$
Data	2179	-	-
Data/Pred	0.99	-	-

TABLE 5.39. Predicted and observed yields in the mBB control region. Detector modeling uncertainties, MC background modeling uncertainties and QCD background modeling uncertainties from ABCD method are considered for the systematic uncertainties.

detector modeling systematic uncertainties, MC background modeling uncertainties and uncertainties from the ABCD method for QCD multijet background.

Figure 5.17 is the invariant mass of the reconstructed di-Higgs (HH)system distribution, the $\not\!\!E_T$, and $W \to l\nu$ system transverse mass distributions and as it can be seen that it is reasonably modelled. The good modeling observed of the distributions gives confidence to the QCD multijet prediction as the events from the background tend to have low values of $\not\!\!E_T$ and transverse mass.

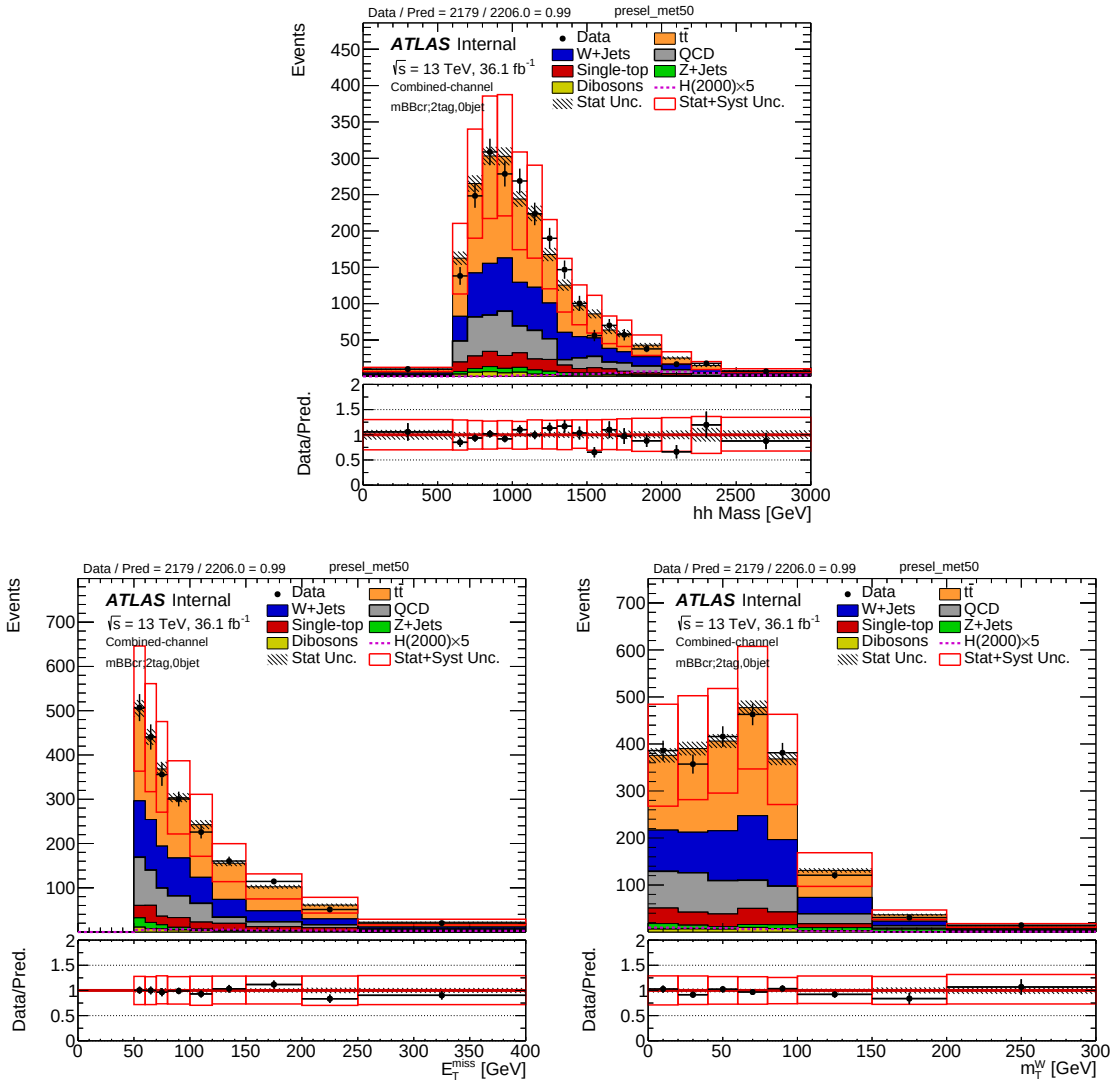

FIGURE 5.17. The invariant mass of the reconstructed di-Higgs (HH) system, $\not{E}_T$and transverse mass of the $W \rightarrow l\nu$ system distributions of events in the mBB control region (mBBcr).

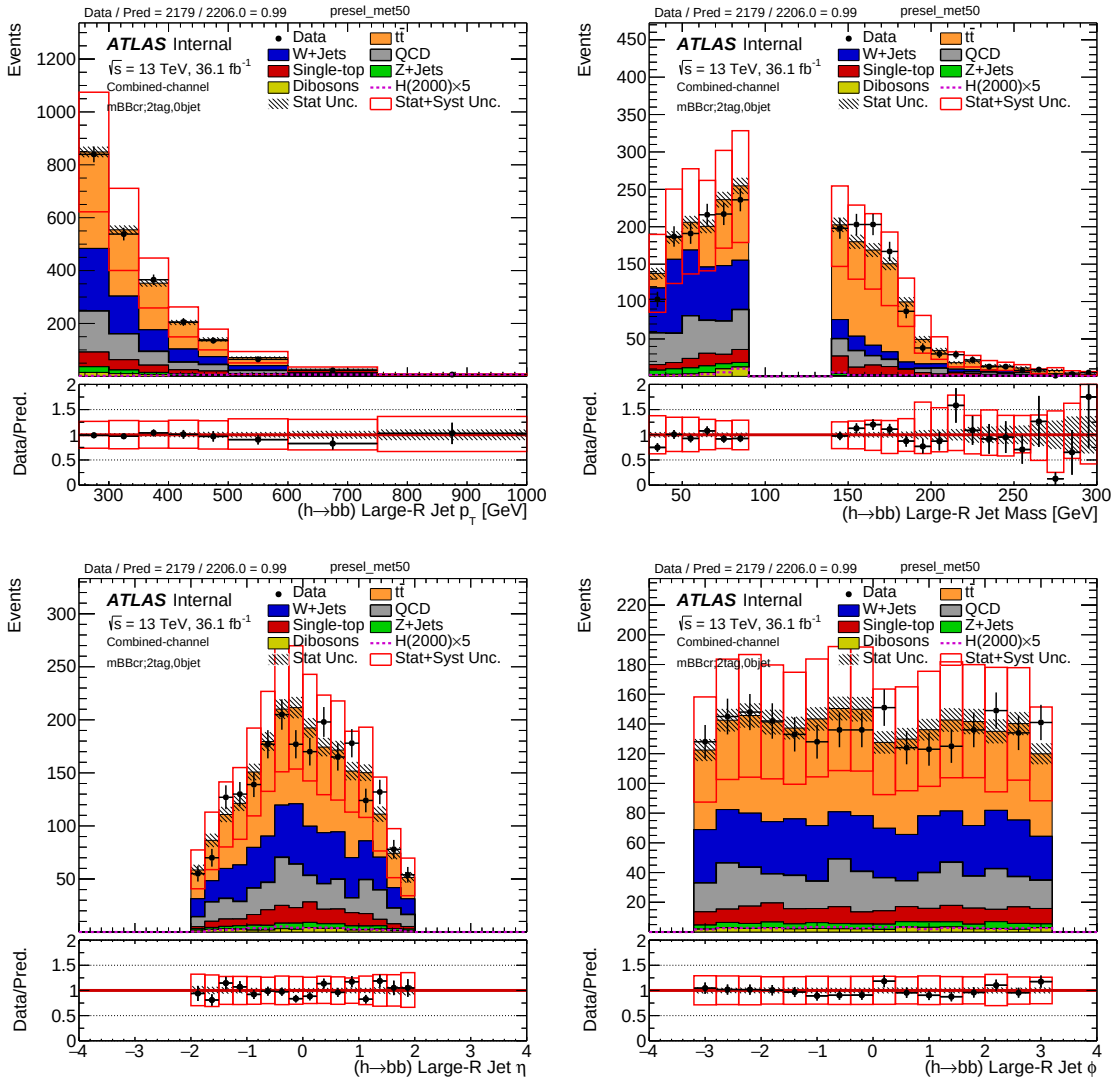

FIGURE 5.18. Kinematic distributions of the reconstructed large-R jet in the mBB control region (mBBcr).

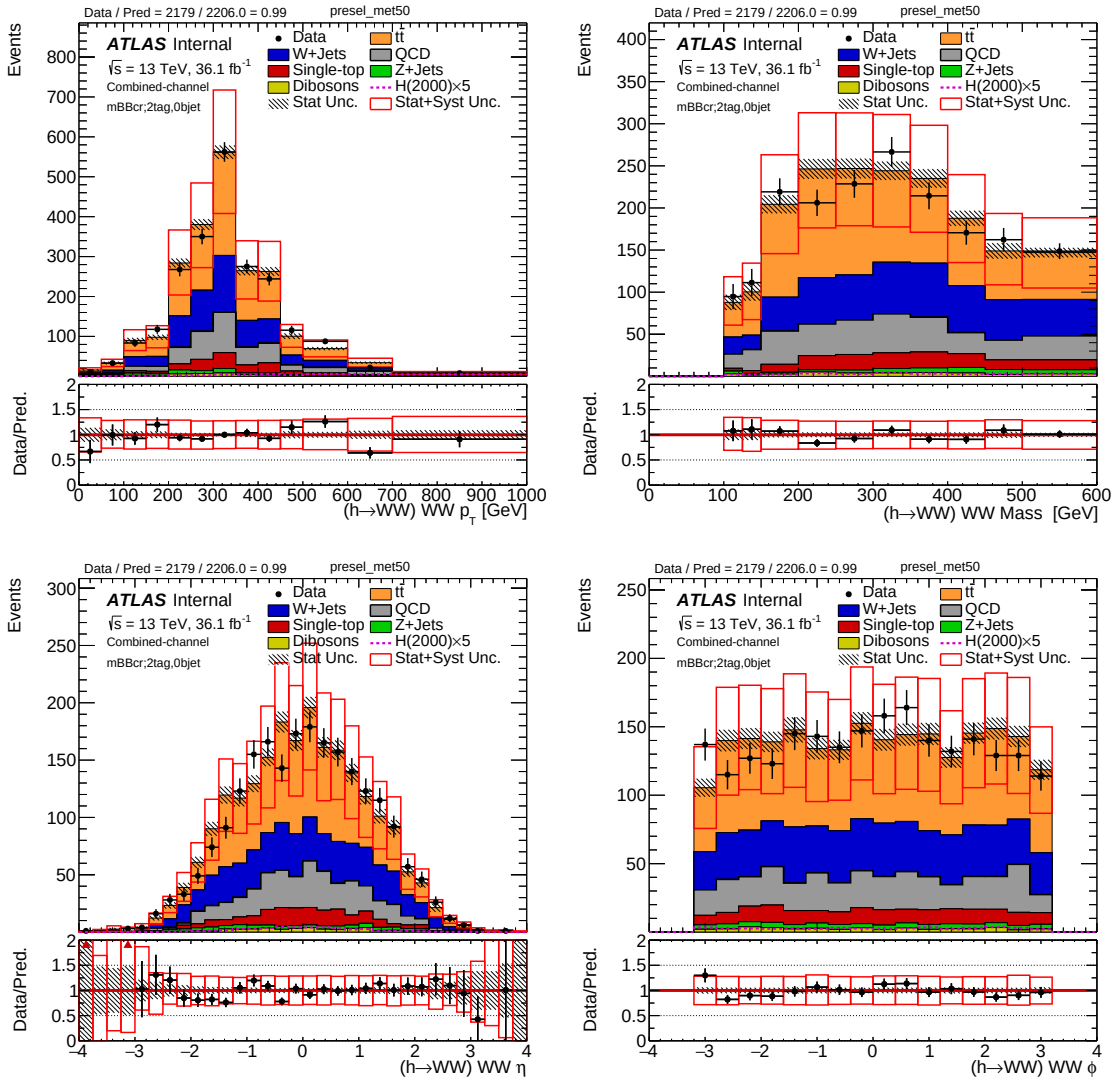

FIGURE 5.19. Kinematic distributions of the reconstructed $h \to WW$ system in the mBB control region (mBBcr).

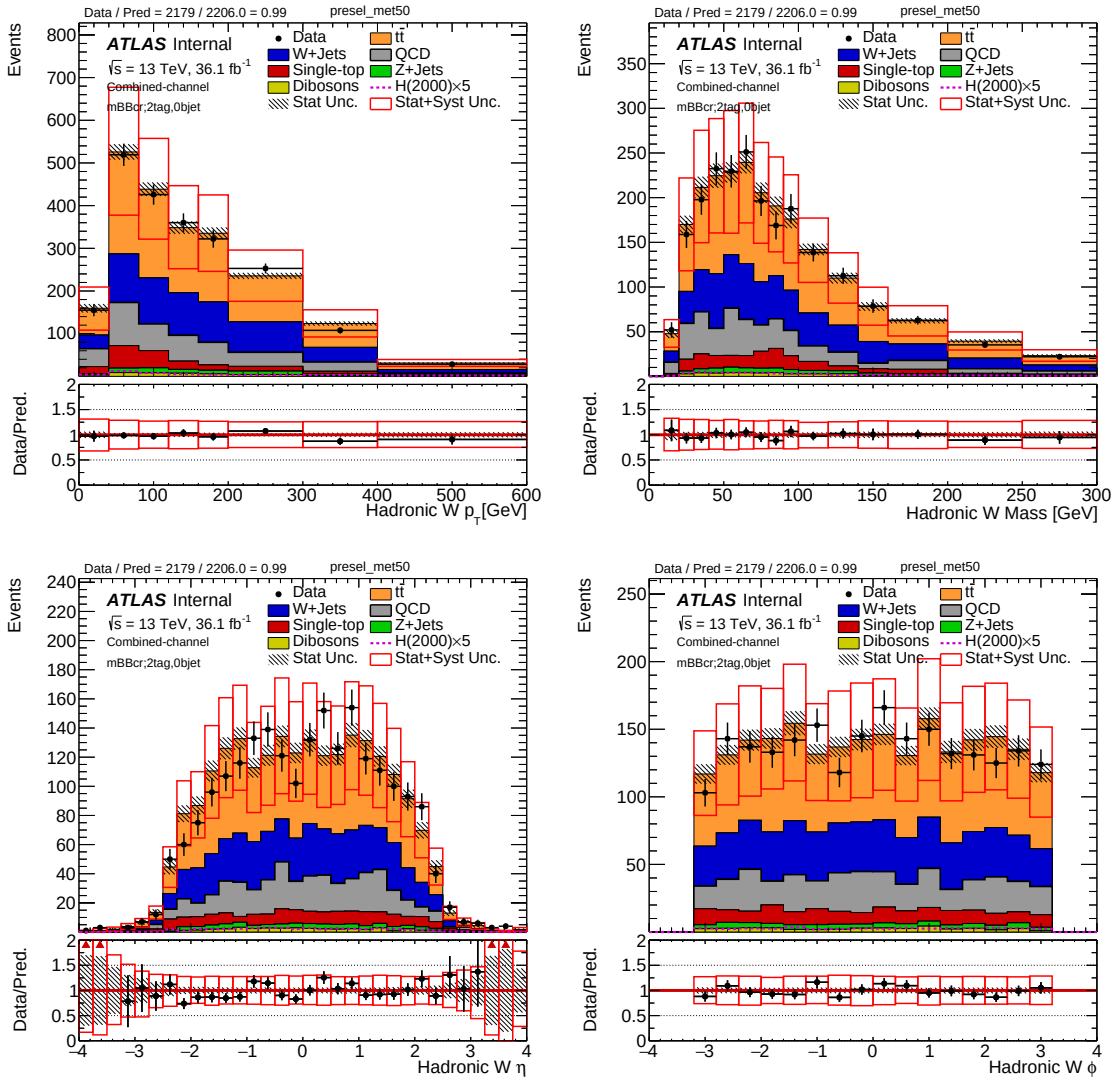

FIGURE 5.20. Kinematic distributions of the reconstructed $W \to q\bar{q}$ system in the mBB control region (mBBcr).

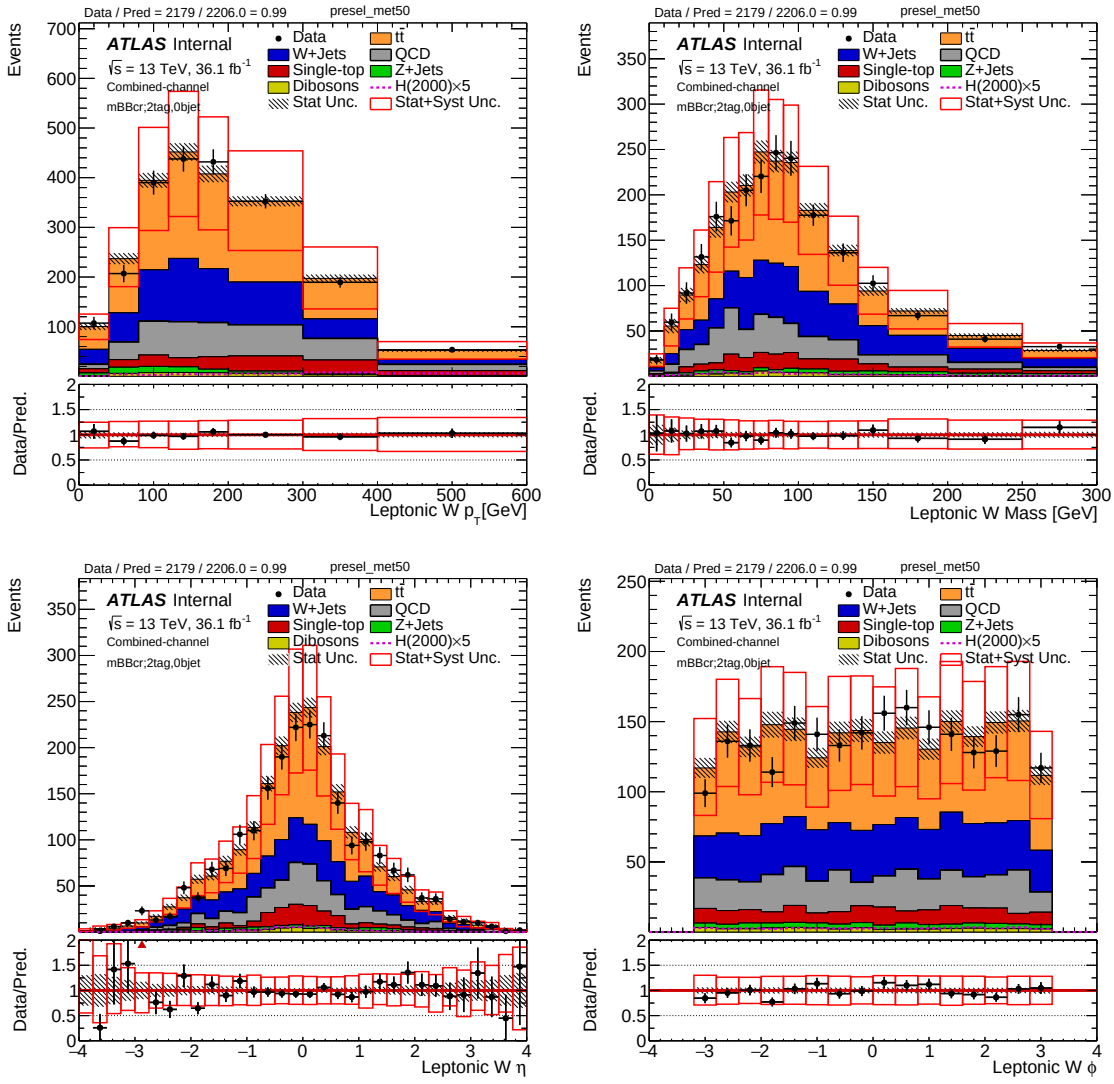

FIGURE 5.21. Kinematic distributions of the reconstructed $W \to l\nu$ system in the mBB control region (mBBcr).

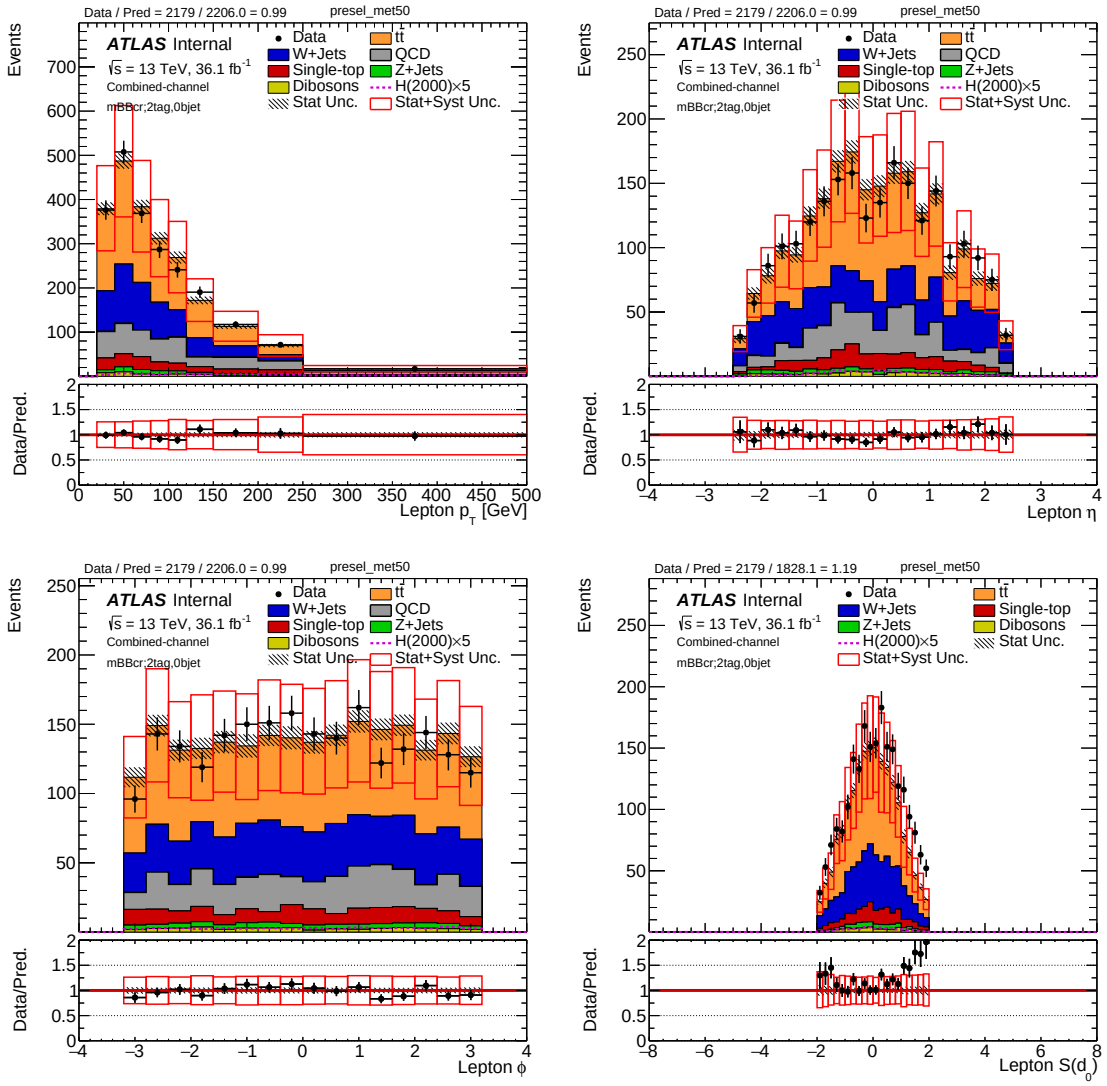

FIGURE 5.22. Kinematic distributions of the selected lepton in the mBB control region (mBBcr).

5.5.10. Systematics

5.5.10.1. Dectector modeling uncertainties

The experimental uncertainties considered in the analysis are listed in Table 5.40. The uncertainties are applicable to signal and background processes that are modelled using MC simulation.

5.5.10.1.1. $|d_0^{\mathrm{sig}}|$ cut efficiency modeling In this analysis, the $|d_0^{\mathrm{sig}}|$ cut value used for electrons and muons are not the recommended value by CP groups. The recommended value for electrons is 5 while for muons it is 3. The modeling of the $|d_0^{\mathrm{sig}}|$ significance is assesed in a top-enriched control region. The event reconstruction and selection criteria for the top-enriched control region are exactly the same as outlined in Section 5.5.1 to 5.5.3.1 with the exception that the b-tagging requirement on the event is different. For this control region, each event is required to have either the leading or sub-leading track-jet to be b-tagged but not both track-jets to b-tagged. The event is also required to have at least one b-tagged signal small-R jets, which is in other words the b-jet veto is reversed. For the purposes of studying the modeling the of the $|d_0^{\mathrm{sig}}|$ distribution, no $|d_0^{\mathrm{sig}}|$ requirement is applied on the reconstructed leptons.

Figure 5.25 shows the $|d_0^{\mathrm{sig}}|$ distribution for the electron and the muon in the top-enriched control region. A clear bias in Data can be observed and this is consistent with other studies throught ATLAS. In order to take into account the effect of the mismodeling of the d_0^{sig} bias on the tight d_0^{sig} cut used in this analysis, the efficiency of the $|d_0^{\mathrm{sig}}|$ cut is evaluated and compared between Data and MC. Table 5.41 shows the efficiency of the $|d_0^{\mathrm{sig}}|$ cut in Data and MC and the ratio of effieciency between Data and MC in each lepton channel. The relative difference between Data/MC

Systematic uncertainty	Short description
Event	
ATLAS_LUMI_2015_2016	uncertainty on total integrated luminosity
PRW_DATASF	pile-up reweighting uncertainty
Electrons	
EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	ID efficiency uncertainty
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty
EG_SCALE_ALL	energy scale uncertainty
EG_RESOLUTION_ALL	energy resolution uncertainty
Muons	
MUON_EFF_TrigStatUncertainty MUON_EFF_TrigSystUncertainty	trigger efficiency uncertainty
MUON_EFF_STAT MUON_EFF_SYS	reconstruction and ID efficiency uncertainty for muons
MUON_ISO_STAT MUON_ISO_SYS	isolation efficiency uncertainty
MUONS_SCALE	energy scale uncertainty
MUONS_ID	energy resolution uncertainty from inner detector
MUONS_MS	energy resolution uncertainty from muon system
Small-R jets	
JET_SR1_JET_GroupedNP_1 JET_SR1_JET_GroupedNP_2 JET_SR1_JET_GroupedNP_3 JET_SR1_JET_EtaIntercalibration_NonClosure	energy scale uncertainties strongly-reduced to 4 components.
JET_JER_SINGLE_NP	energy resolution uncertainty
JET_JvtEfficiency	JVT efficiency uncertainty
Large-R jets	
FATJET_Medium_JET_Comb_Baseline_Kin FATJET_Medium_JET_Comb_modeling_Kin FATJET_Medium_JET_Comb_TotalStat_Kin FATJET_Medium_JET_Comb_Tracking_Kin	energy scale uncertainties (p_T and mass scales are fully correlated)
FATJET_JER	energy resolution uncertainty
FATJET_JMR	mass resolution uncertainty
Track-jets and Small-R jets	
FT_EFF_Eigen_B FT_EFF_Eigen_C FT_EFF_Eigen_L	b-tagging efficiency uncertainties ("BTAG_MEDIUM"): 3 components for b jets, 4 for c jets and 5 for light jets
FT_EFF_Eigen_extrapolation	b-tagging efficiency uncertainty on the extrapolation to high p_T jets
FT_EFF_Eigen_extrapolation_from_charm	b-tagging efficiency uncertainty on tau jets
$\not{E}_T$	
MET_SoftTrk_ResoPara	track-based soft term related longitudinal resolution uncertainty
MET_SoftTrk_ResoPerp	track-based soft term related transverse resolution uncertainty
MET_SoftTrk_Scale	track-based soft term related longitudinal scale uncertainty

TABLE 5.40. Summary of the (nuisance parameter) names and meanings of the detector modeling systematic uncertainties.

ratio and unity is taken as the uncertainty on the $|d_0^{\text{sig}}|$ significance cut efficiency and it is assigned for all processes with prompt leptons modelled by MC simulation. The nuisance parameter name associated to this uncertainty is `LEP_d0_CutEff`

	Electron	Muon
$\|d_0^{\text{sig}}\| < 2.0$		
Data	0.884	0.897
MC	0.926	0.937
Data/MC	0.955	0.957
$\|d_0^{\text{sig}}\| > 2.0$		
Data	0.115	0.104
MC	0.074	0.063
Data/MC	1.572	1.640

TABLE 5.41. Efficiency of the $|d_0^{\text{sig}}|$ cut for electrons and muons in Data and MC. The Data/MC ratio is also calculated and the difference between the ratio and unity is taken as the systematic uncertainty on the $|d_0^{\text{sig}}|$ efficiency modeling for events with leptons that pass (< 2.0) or fail (> 2.0) the $|d_0^{\text{sig}}|$ cut.

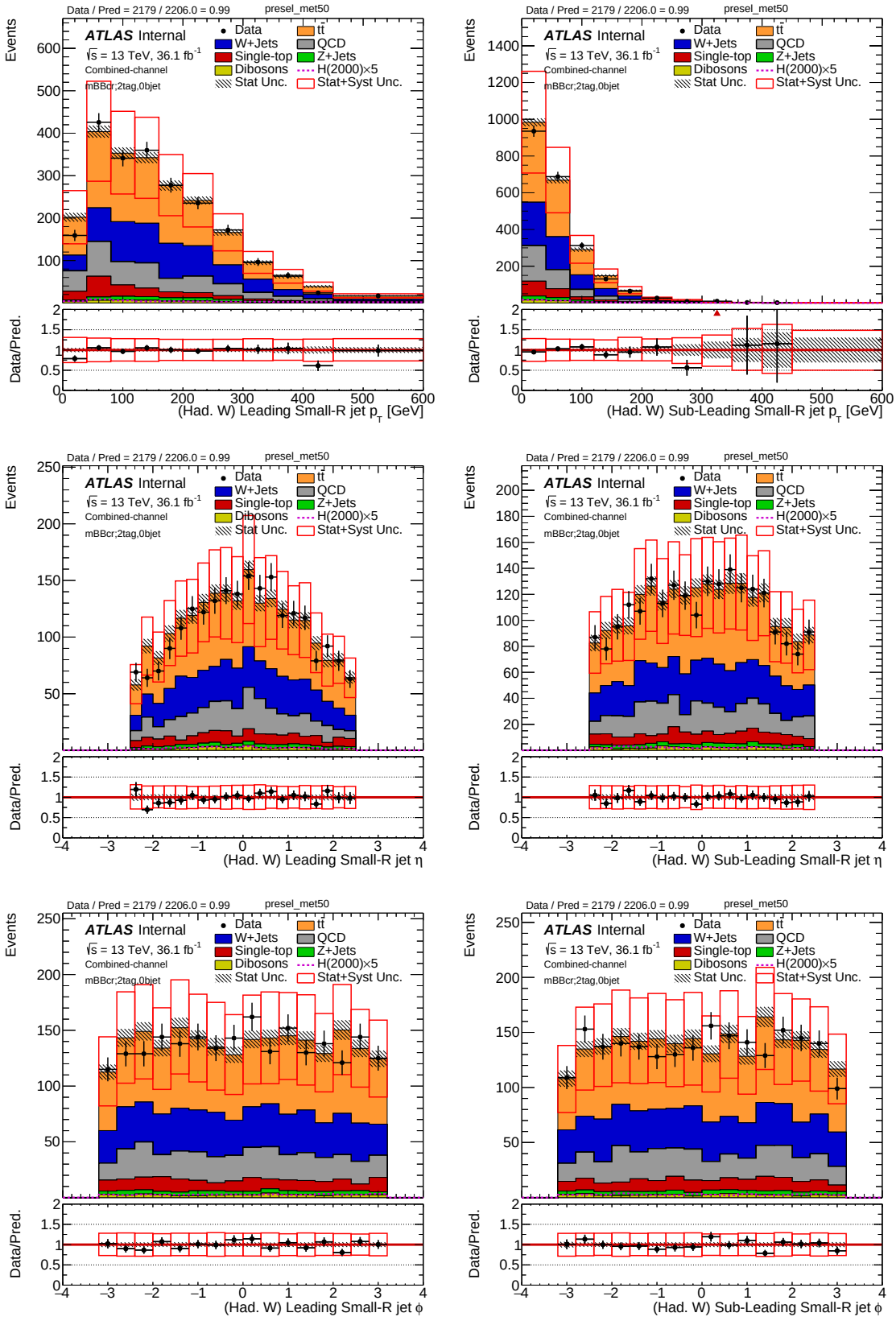

FIGURE 5.23. Kinematic distributions of the leading and sub-leading small-R jets (of the reconstructed hadronic W) in the mBB control region (mBBcr).

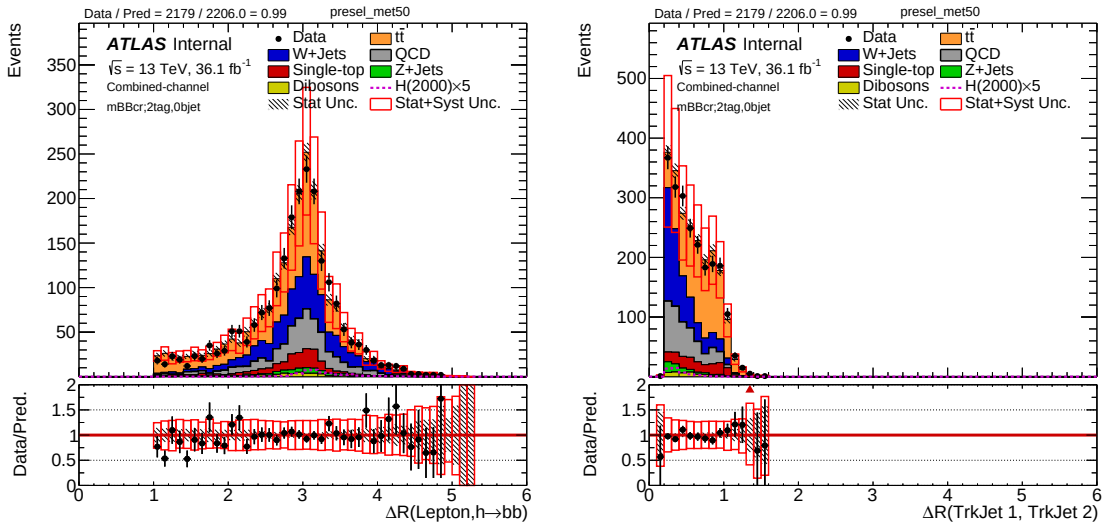

FIGURE 5.24. ΔR distribution between the selected lepton and the large-R jet and ΔR distribution between the track-jets inside the large-R jet in the mBB control region (mBBcr).

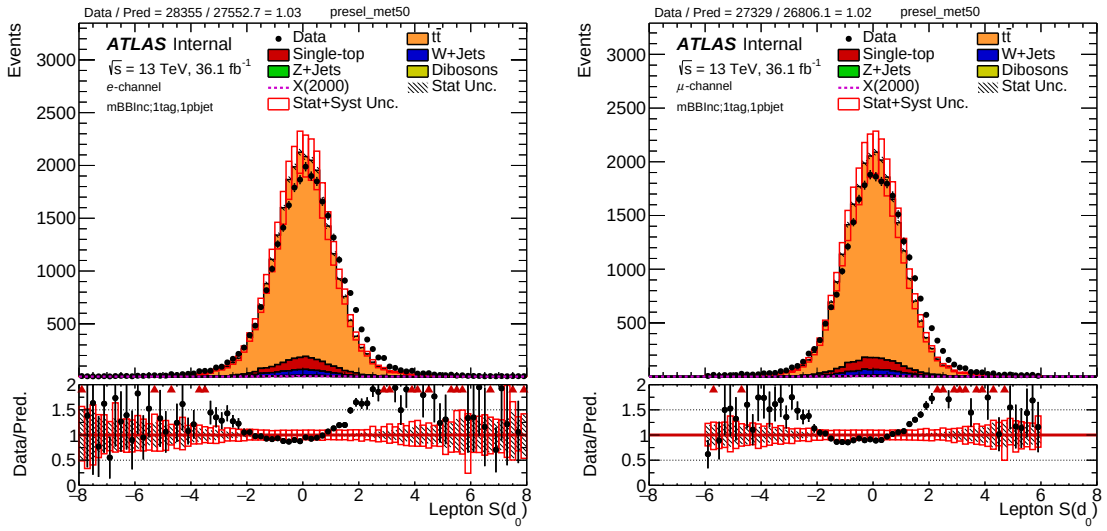

FIGURE 5.25. Electron and muon d_0^{sig} distributions in the top control region without the large-R jet mass cut.

5.5.10.2. Background and Signal modeling Systematics

5.5.10.2.1. Methodology Uncertainties in the m_{HH} distributions are assigned to the dominant backgrounds, $t\bar{t}$ and V+Jets and single top by comparing the nominal MC samples to a number of alternative MC generators at the reconstruction level. The comparisons are performed in with the same event selection in Section 5.5. Each uncertainty contains two components, a shape systematic and a normalization systematic due to acceptance.

The shape systematic corresponds to a reweighting function derived by fitting a 1^{st} order polynomial to the ratio of m_{HH} distribution of the variation sample over the nominal sample. The m_{HH} distribution for the the variation sample is normalized to the same number of events of the norminal sample.

5.5.10.2.2. Top-quark processes ($t\bar{t}$ & single top) Four alternative MC $t\bar{t}$ samples are used to assess 3 aspects of the MC modeling, whilst five alternative MC samples are used to assess 4 aspects of the MC modeling for single top quark pair production. The alternative samples considered are:

- **Powheg +Herwig++:** The ME Powheg generator uses the same setup as that used for the nominal Powheg +Pythia 6 configuration, but the parton shower (PS) generator is swapped out for Herwig++ version 2.7.1 using the UE-EE-5 tune and CTEQ6L1 PDF set. The purpose therefore of this comparison is to test the PS, hadronisation, underlying event (UE) and Multiple Parton Interation (MPI) models whilst maintaining the same hard scattering model given by Powheg.
- **aMC@NLO+Herwig++:** The ME generator is swapped out for aMC@NLO using the CT10 PDF set, interfaced with Herwig++ using the CTEQ6L1-UE-

EE-5 tune and CTEQ6LI PDF set. This sample is compared to the previous POWHEG +Herwig++ sample. This fixes the PS generator component, but alters the hard scattering generator, making this variation sensitive to the hard scatter model.

– **Powheg +Pythia 6 Radhi/RadLo:** Using the same setup as that used for the nominal POWHEG +PYTHIA 6 sample, the RadHi and RadLo samples correspond to either the enhancement (high) or reduction (low) of initial/final state radiation (IFSR). The two samples are compared to the nominal sample setup, and so are sensitive to variations of IFSR models.

  * RadHi: The renormalization (μ_R) and factorisation scale (μ_F) scales are decreased by a factor of 0.5, the POWHEG hdamp parameter is doubled ($2 \times m_{top}$), and the high radiation PERUGIA2012 tune is used.

  * RadLo: The renormalization (μ_R) and factorisation scale (μ_F) scales are increased by a factor of two, the POWHEG hdamp parameter is kept at m_{top}, and the low radiation PERUGIA2012 tune is used.

– **Powheg +Pythia 6 Diagram Subtraction:** For the production of a single top quark in association with a W-boson (Wt) the interference with the $t\bar{t}$ production process at NLO in QCD is removed by subtracting the cross-section associated with the $t\bar{t}$ double resonance amplitude terms, rather than subtracting the same terms from the amplitude prior to the calculation (Diagram Removal).

Table 5.42 shows the estimated uncertainty on the normalization of the $t\bar{t}$ background in the signal region from the comparison of the nominal $t\bar{t}$ sample to the alternative samples. The largest uncertainty comes the RadLo variation, which

is about ∼8.4% with similar level of uncertainties from alternative ME generator choice and alternative PS generator choice. The normalization of $t\bar{t}$ background is assinged with a single nuisance parameter with the total uncertainty set as the prior uncertainty.

Shape comparisons of the m_{HH} distribution between the nominal $t\bar{t}$ sample to the alternative samples were made and they are shown in Figure 5.26 and Figure 5.27.

Variation	Uncertainty (%)
RadHi	1.4
RadLo	8.4
aMC@NLO	7.1
Herwig++	7.8
PDF	1.9
Scale	5.0
Total	13.5

TABLE 5.42. The normalization uncertainty for the $t\bar{t}$ background in the signal region from different sources. The total uncertainty is calculated as the sum of quadrature from all the sources.

For the single-top background, only the uncertainties on the modeling of the Wt production process are considered since it is the dominant single-top process in the signal region. Table 5.43 shows the estimated uncertainty on the normalization of the single-top background. The biggest uncertainties comes from the PS generator choice and comparisons to the DS sample. The total uncertainty is abnormally large, which is larger than 100%.

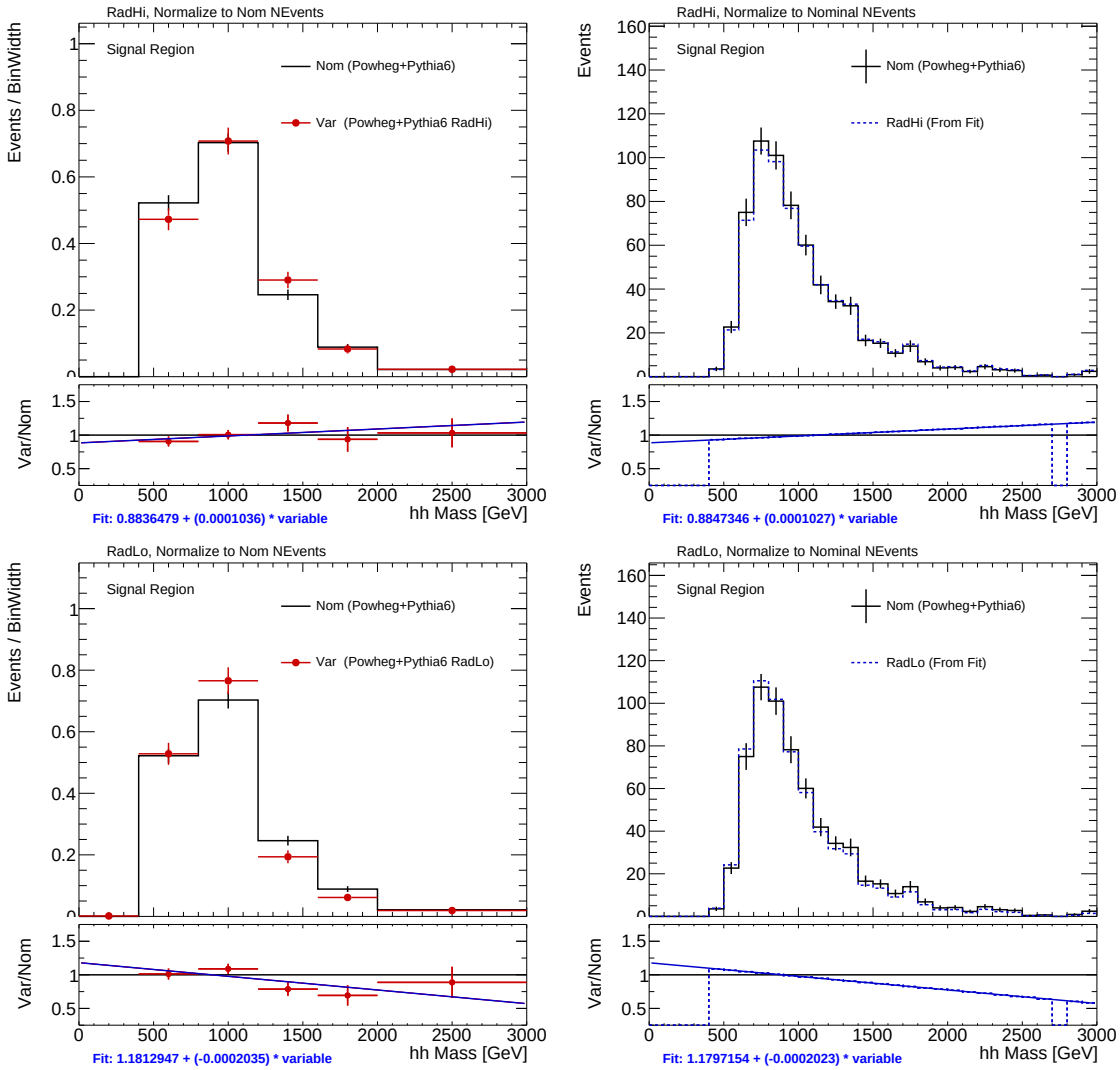

FIGURE 5.26. m_{HH} distribution shape comparison between nominal $t\bar{t}$ sample and alternative samples. Plot on the left is a direct comparison between the nominal and alternative sample while on the right, the variation comes from the reweighted function applied to the nominal $t\bar{t}$ sample. The linear fit in the ratio of the left plot is used as the reweighted function.

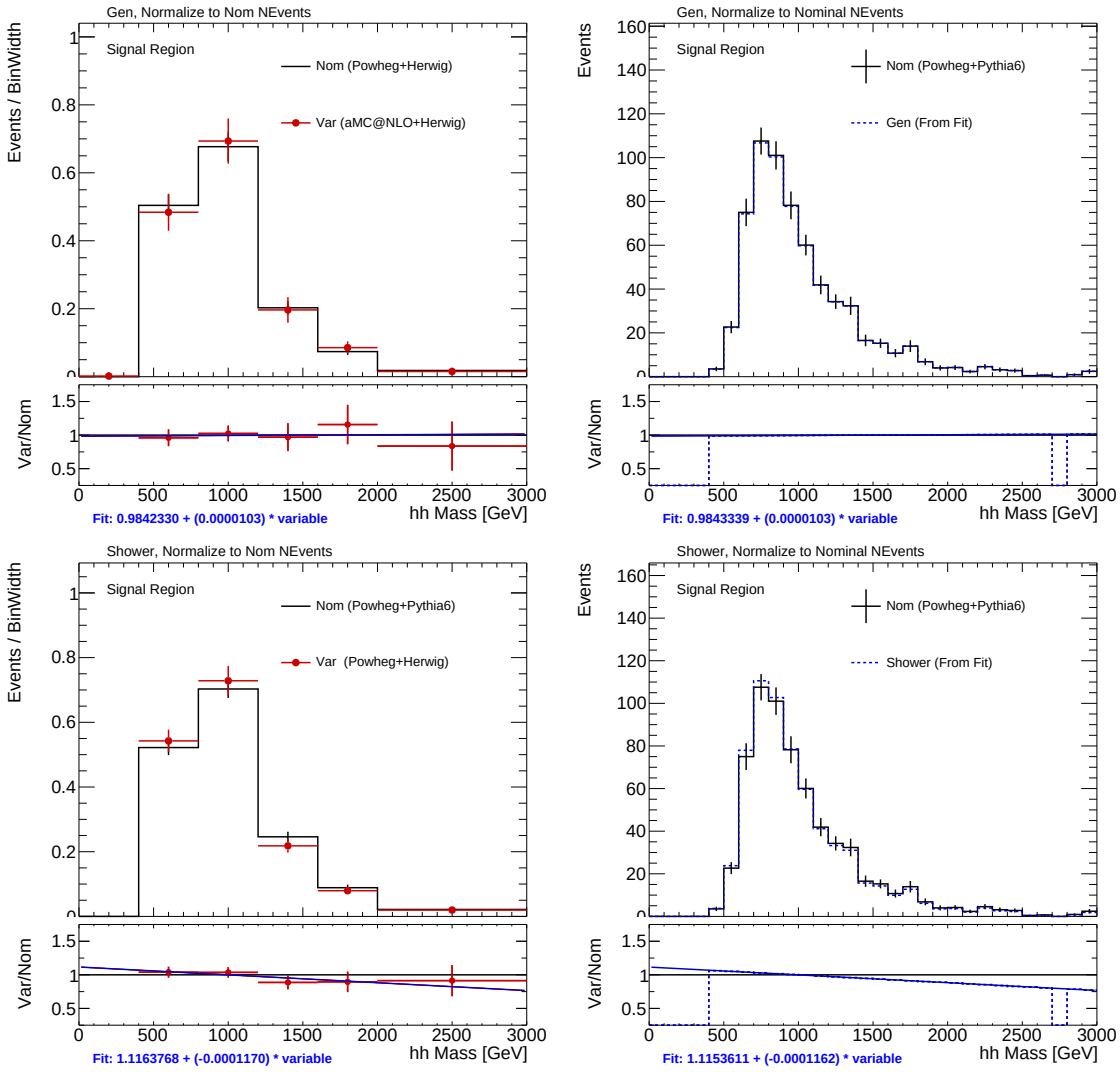

FIGURE 5.27. m_{HH} distribution shape comparison between nominal $t\bar{t}$ sample and alternative samples. Plot on the left is a direct comparison between the nominal and alternative sample while on the right, the variation comes from the reweighted function applied to the nominal $t\bar{t}$ sample. The linear fit in the ratio of the left plot is used as the reweighted function.

Variation	Uncertainty (%)
RadHi	15.1
RadLo	19.0
Herwig++	33.5
DR	72.5
aMC@NLO	25.4
Total	85.9

TABLE 5.43. The normalization uncertainty for the $t\bar{t}$ background in the signal region from different sources. The total uncertainty is calculated as the sum of quadrature from all the sources.

5.5.10.2.3. V+Jets processes The nominal V+Jets prediction, uses the ME+PS generator Sherpa 2.2.1 interfaced with the NNPDF 3.0 NNLO PDF set. This default configuration provides a prediction for vector boson production plus associated jets at NLO accuracy at the ME level for up to 2 extra partons, and LO accuracy for 3 and 4 extra partons in QCD. The merging of additional parton multiplicities arising from the internal Sherpa PS, is regulated by the MEPS@NLO merging technique.

The alternative samples used to assess the modeling uncertainties are:

– **MadGraph5+Pythia 8.186 :** The LO ME generator MadGraph5 using the NNPDF3.0(2.3) NLO(LO) PDF set interfaced with Pythia 8 version 8.186 using the A14 tune, offers a LO+NLL accurate prediction for vector boson production in association with jets for up to four extra partons from the ME and 4+ parton from Pythia 8 at NLL accuracy. The comparison between the nominal Sherpa 2.2.1 sample with this sample convolves the ME and PS model variation. Due to the unavailability of this sample at reconstruction level, the comparison is made at (particle) truth level.

– **Sherpa 2.2.1 scale variations:** Configured in the same manner as the nominal V+Jets sample, the renormalization μ_R and resummation μ_F scales are varied up/down by a factor of two.

– **Sherpa 2.2.1 PDF variations:** Configured in the same manner as the nominal V+Jets sample. The 100 NNPDF3.0NNLO replicas variations are available. The central values of two alternative PDF sets, MMHT2014NNLO 68% CL and CT14NNLO are also available.

– **Sherpa 2.2.1 $\alpha_s(PDF)$ variations:** Configured in the same manner as the nominal V+Jets sample, the α_s value used by the nominal NNPDF 3.0 NNLO PDF is varied up and down according to a variation of the μ_R scale by a factor of two.

Table 5.44 shows the estimated uncertainty on the normalization of the W+jets background in the signal region from the comparison of the nominal W+jets sample to the alternative samples. The largest uncertainty comes from the renormalization and resummation scale, which is about ∼42% and dominates the total uncertainty on W+jets background. The normalization of W+jets background is assinged with a single nuisance parameter with the total uncertainty set as the prior uncertainty.

Shape comparisons of the m_{HH} distribution between the nominal W+jets sample to the alternative samples were made and only one variation was found to have noticeable difference: the scale variation of μ_R=0.5, μ_F=0.5 as in Figure 5.28.

Table 5.45 shows the estimated uncertainty on the normalization of the Z+jets background in the signal region from the comparison of the nominal Z+jets sample to the alternative samples. The largest uncertainty comes from the renormalization and resummation scale, which is about ∼48% and dominates the total uncertainty on Z+jets background. The normalization of Z+jets background is assinged with

Variation	Uncertainty (%)
Scale	41.9
α_S(PDF)	8.4
PDF alternative set	1.6
NNPDF replicas	5.6
Madgraph+Pythia8	11.0
Total	44.5

TABLE 5.44. The normalization uncertainty for the W+jets background in the signal region from different sources. The total uncertainty is calculated as the sum of quadrature from all the sources.

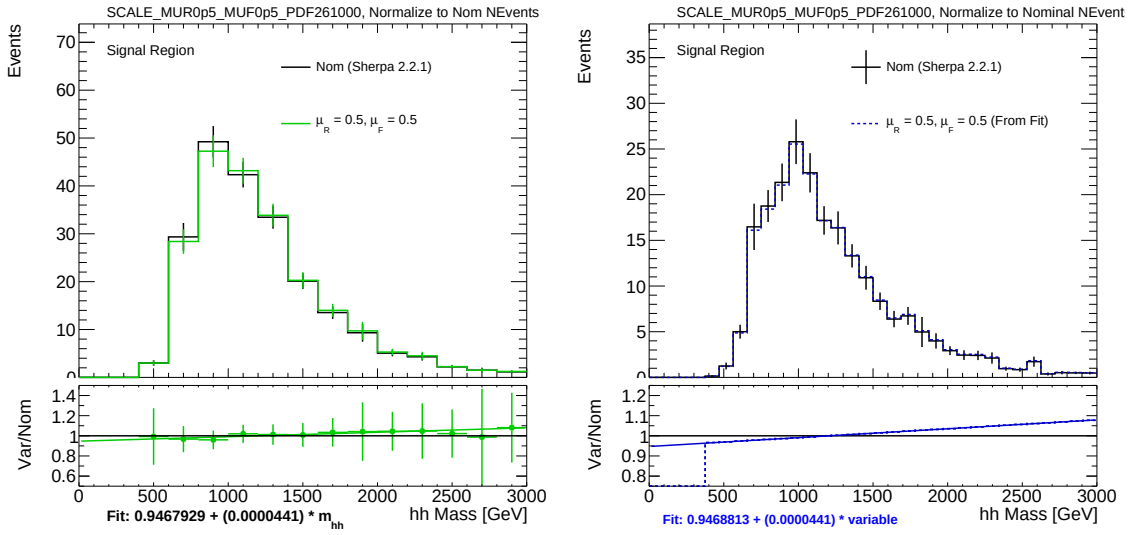

FIGURE 5.28. m_{HH} distribution shape comparison between nominal W+jets sample and scale variation (μ_R=0.5, μ_F=0.5) sample. Plot on the left is a direct comparison between the nominal and variation sample while on the right, the variation comes from the reweighted function applied to the nominal W+jets sample. The linear fit in the ratio of the left plot is used as the reweighted function.

a single nuisance parameter with the total uncertainty set as the prior uncertainty. The uncertainty on the m_{HH} shape is found to be negligible and therefore ignored.

Variation	Uncertainty (%)
Scale	48.3
α_S(PDF)	1.6
PDF alternative set	2.7
NNPDF replicas	1.4
Total	48.4

TABLE 5.45. The normalization uncertainty for the Z+jets background in the signal region from different sources. The total uncertainty is calculated as the sum of quadrature from all the sources.

5.5.10.2.4. Diboson processes The systematic uncertainty on the normalization of the Diboson background is assigned to be 40%. This uncertainty is taken from the resolved analysis. As the background is small, the uncertainty is considered to be conservative.

5.5.10.2.5. Production Systematic uncertainties on the acceptance of signal processes are computed by generating alternative variation signal samples and then compare their acceptance with respect to the nominal signal samples. The difference sources of uncertainty considered are:

- **Scale variations:** Configured in the same manner as the nominal signal samples but the renormalization and resummation scales are varied up/down by a factor of two.
- **Parton shower choice:** Configured in the same manner as the nominal signal samples but with Pythia 8 chosen as the shower generator instead of Herwig++.

Table 5.46 lists the systematic uncertainties for four different scalar signal sample mass points.

Variation	Xhh1000	Xhh1500	Xhh2000	Xhh2500	Xhh3000
Scale	0.2	0.2	0.4	0.4	0.4
PDF	0.4	0.2	0.4	0.2	0.1
Shower	0.4	0.8	1.6	3.4	4.1

TABLE 5.46. Theoretical uncertainties (in percentage) on the acceptance of several signal mass points.

5.5.10.3. QCD multijet modeling

Systematic uncertainties related to the modeling of the multijet background were discussed in Section 6.3.4.3 for the predicted yield and in Section 6.3.4.3 for the predicted m_{HH} distribution.

5.6. Results

5.6.1. Resolved analysis results

The resolved analysis is described in detail in Section 5.4. The event selection is described in Section 5.4.1 and summarized in Table 5.8. For each selected event, the invariant mass of the HH system (m_{HH}) is reconstructed and its distribution is shown in Figure 5.29 for the *non-res* and the *m500* analyses, and in Figure 5.30 for the *low-mass* and the *high-mass* analyses. Data are generally in good agreement with the expected background predictions within the total uncertainty. The signal m_{HH} distribution is shown in the figure for the non-resonant and the scalar resonance. Because the scalar-resonance samples are simulated in the narrow-width approximation, the reconstructed resonance width is exclusively due to the detector resolution.

The m_{HH} distribution is sampled with resonance-mass-dependent m_{HH} requirements as reported in Table 5.9. The numbers of events in the signal and

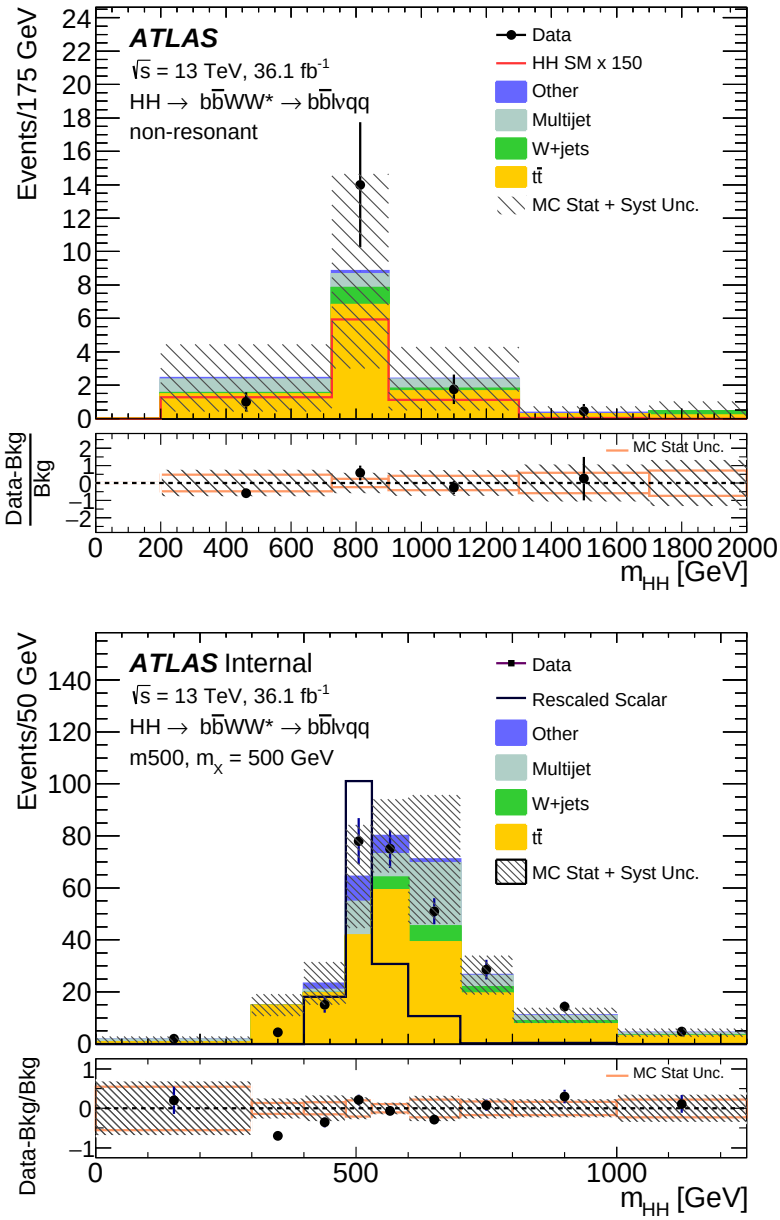

FIGURE 5.29. m_{HH} distributions for non-resonant and *m500* selections in the resolved analysis. For each selection the corresponding signal hypothesis, non-resonant and scalar resonance, is shown. For the scalar signal, resonances with mass 500 GeV are shown. The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty. The non-resonant signal is multiplied by a factor of 150 with respect to the expected SM cross section. The scalar signal is multiplied by a factor of five with respect to the expected upper-limit cross section.)

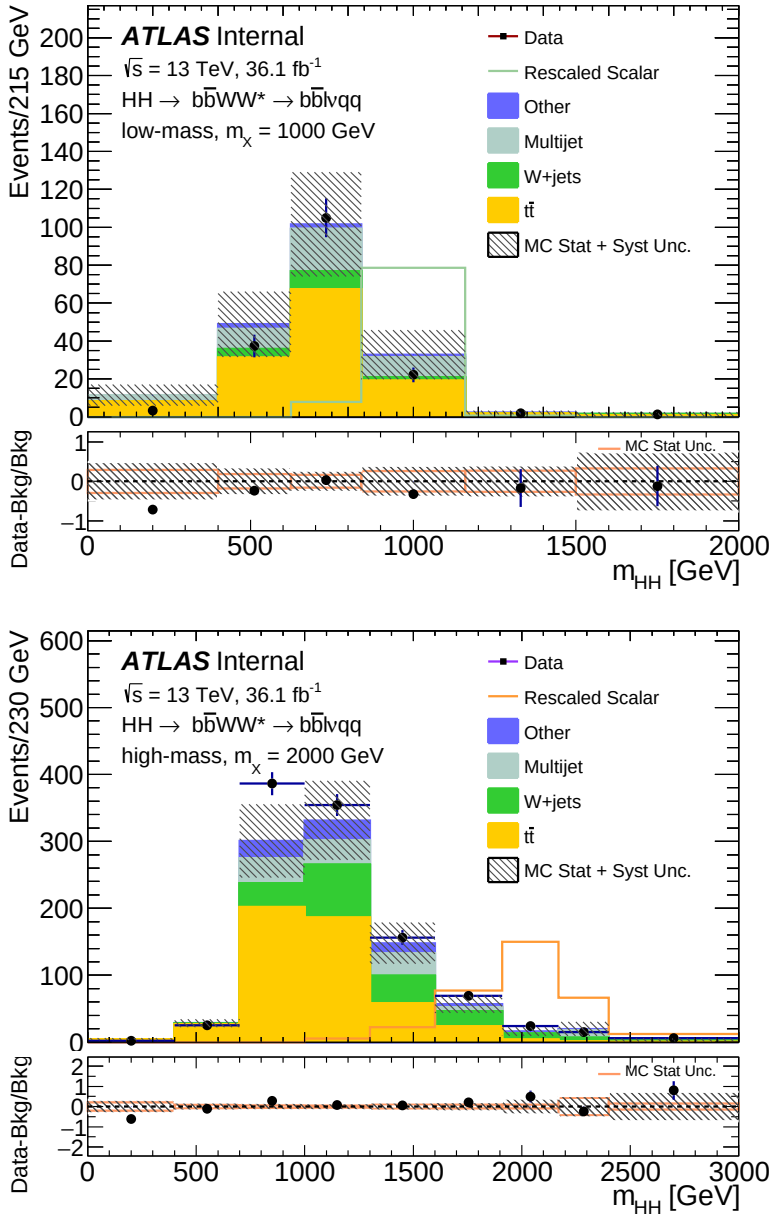

FIGURE 5.30. m_{HH} distributions in the resolved analysis selections. For each selection the corresponding signal hypothesis and mass 1000 (2000) GeV for the *low-mass* (*high-mass*) analysis, are shown. The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty. In the plot the scalar signal is multiplied by a factor of eight with respect to the expected upper-limit cross section;for the plot on the right the multiplying factor is 20 for the scalar signal.

control regions (the $t\bar{t}$ control region and the C region of the multijet estimation procedure) are simultaneously fit using a maximum-likelihood approach. The fit includes six contributions: signal, W+jets, Z+jets, $t\bar{t}$, single-top-quark production, diboson and multijet. The $t\bar{t}$ and multijet normalisations are free to float, the C region of the ABCD method being directly used in the fit, while the diboson, W+jets and Z+jets backgrounds are constrained to the expected SM cross sections within their uncertainties.

The fit is performed after combining the electron and muon channel distributions. Statistical uncertainties due to the limited sample sizes of the simulated background processes are taken into account in the fit by means of nuisance parameters, which are parameterised by Poisson priors. Systematic uncertainties are taken into account as nuisance parameters with Gaussian constraints. For each source of systematic uncertainty, the correlations across bins and between different kinematic regions, as well as those between signal and background, are taken into account. Table 5.47 shows the post-fit number of predicted backgrounds, observed data, and the signal events normalized to the expected upper limit cross sections. Expected event yields vary across mass because of varying selections. For instance, the requirement on $p_{\mathrm{T}}^{b\bar{b}}$ is higher in *non-res* selection than in *low-mass* selection. Similarly, even within *low-mass* or *high-mass* selection, the requirement on m_{HH} vary across mass.

No significant excess over the expectation is observed and the results are used to evaluate an upper limit at the 95% confidence level (CL) on the production cross section times the branching fraction for the signal hypotheses under consideration. The exclusion limits are calculated with a modified frequentist method [92], also known as $\mathrm{CL_s}$, and the profile-likelihood test statistic [93]. None of the considered systematic uncertainties is significantly constrained or pulled in the likelihood fit.

Resonant analysis			
m_X [GeV]	S	Total Bkg.	Data
500	18 ± 5	19 ± 6	26
600	13 ± 2	17 ± 6	16
700	16 ± 2	25 ± 8	22
750	20 ± 2	22 ± 9	27
800	18.4 ± 1.5	20 ± 8	28
900	16.3 ± 1.6	20 ± 7	23
1000	12.0 ± 1.3	14 ± 5	11
1100	9.6 ± 1.2	8 ± 3	8
1200	8.1 ± 0.9	6 ± 3	5
1300	5.1 ± 0.7	3.5 ± 1.8	1
1400	4.3 ± 0.3	1.1 ± 0.2	0
1500	3.5 ± 0.3	1.1 ± 0.2	0
1600	3.1 ± 0.3	0.4 ± 0.3	1
1800	14.1 ± 1.8	17 ± 5	21
2000	8.7 ± 1.0	8 ± 3	9
2250	7.9 ± 1.1	6 ± 2	7
2500	5.5 ± 0.8	3.3 ± 1.4	3
2750	5.7 ± 1.0	3.1 ± 1.3	3
3000	4.3 ± 0.7	2.1 ± 1.0	1
Non-resonant analysis			
Rescaled SM signal		Total Bkg.	Data
17 ± 2		21 ± 8	22

TABLE 5.47. Data event yields, and post-fit signal and background event yields in the final signal region for the non-resonant analysis and the resonant analysis in the 500–3000 GeV mass range. The errors shown are the MC statistical and systematic uncertainties described in Section 5.4.9. The yields are shown for a scalar (S) signal model. Signal event yields are normalized to the expected upper-limit cross section.

In the SM signal hypothesis the observed (expected) upper limit on the $\sigma(pp \to HH) \times \mathcal{B}(HH \to b\bar{b}WW^*)$ at 95% CL is:

$$\sigma(pp \to HH) \cdot \mathcal{B}(HH \to b\bar{b}WW^*) < 2.5 \left(2.5^{+1.0}_{-0.7}\right) \text{ pb.}$$

The branching fraction $\mathcal{B}(HH \to b\bar{b}WW^*) = 2 \times \mathcal{B}(H \to b\bar{b}) \times \mathcal{B}(H \to WW^*) = 0.248$ is used to obtain the following observed (expected) limit on the HH production cross section at 95% CL:

$$\sigma(pp \to HH) < 10 \left(10^{+4}_{-3}\right) \text{ pb,}$$

which corresponds to 300 (300^{+100}_{-80}) times the SM predicted cross section. Including only the statistical uncertainty, the expected upper limit for the non-resonant production is 190 times the SM prediction. This result, when compared with other HH decay channels, is not competitive. This is mainly due to the similarity of the reconstructed m_{HH} spectrum between the non-resonant SM signal and the $t\bar{t}$ background that makes the separation between the two processes difficult.

Figure 5.31 shows the expected and observed limit curves for the production cross section of a scalar S particle. Different selections are used in different resonance mass ranges without attempting to statistically combine them. The switch from one selection to another is performed based on the best expected limit for that resonance mass. The outcome of this procedure is that the *m500* selection is used to set limits on resonances of mass of 500 GeV, the *low-mass* selection is used up to masses of 1600 GeV, while the *high-mass* selection is used in the mass range 1600-3000 GeV.

Overall, the resolved analysis is most sensitive for a mass value of 1300 GeV with an expected upper limit of 0.35 pb on $\sigma(pp \to HH)$. At this mass the observed exclusion limit is 0.2 pb. In both the non-resonant and resonant cases, the impact of

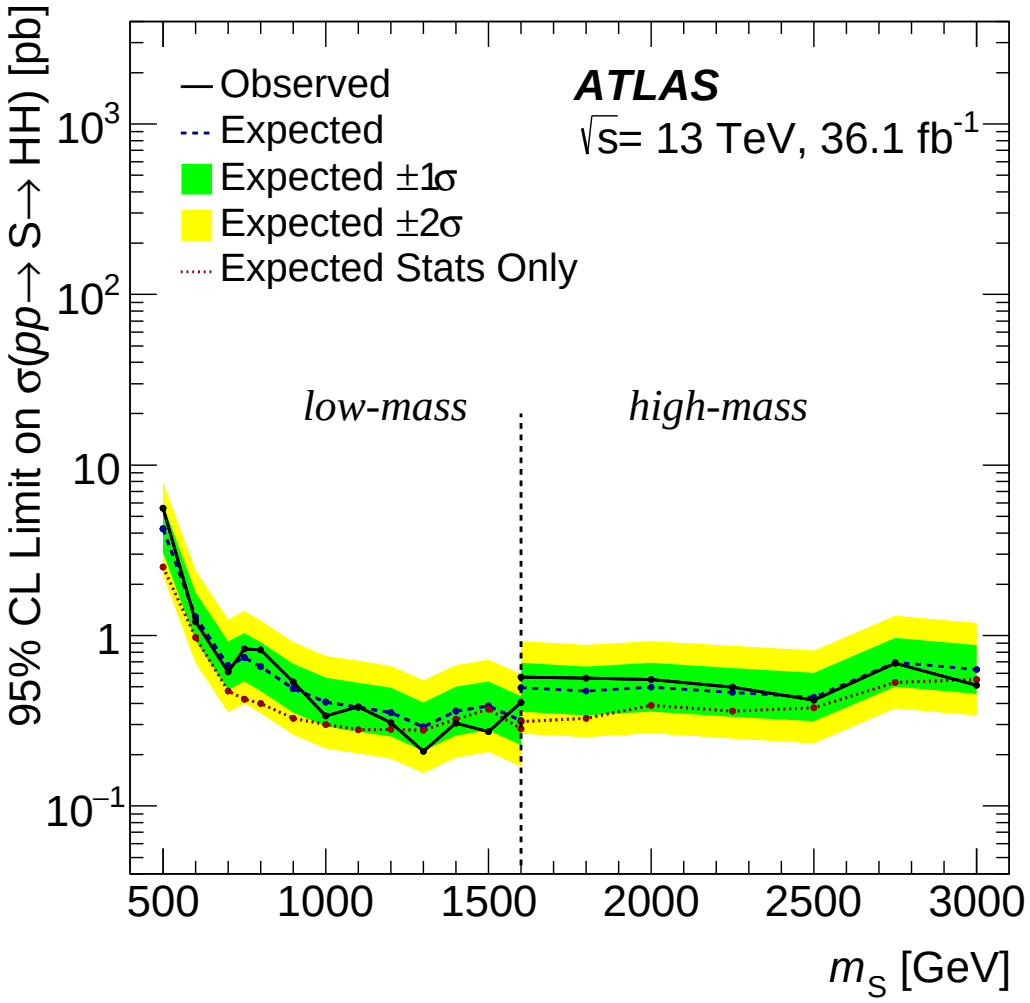

FIGURE 5.31. Expected and observed upper limit at 95% CL on the cross section of resonant pair production for the resolved analysis in the heavy scalar boson S model. The plot also shows the expected limit without including the systematic errors in order to show their impact.

the systematic uncertainties is observed to be large. In order to quantify the impact of the systematic uncertainties, a fit is performed where the estimated signal yield, normalized to an arbitrary cross-section value, is multiplied by a scaling factor α_{sig}, which is treated as the parameter of interest in the fit. The fit is performed using pseudo-data and the contribution to the uncertainty in α_{sig} from several sources is

Statistical source	Resolved analysis			
	Non-Res (%)	500 GeV (%)	1000 GeV (%)	2000 GeV (%)
Signal region	+60/–40	+60/–60	+70/–60	+80/–70
Top control region	+40/–30	+28/–30	+20/–12	+13/–13
Multijet control region	+40/–30	+24/–26	+30/–30	+30/–30
Total statistical	+80/–60	+70/–70	+80/–70	+90/–80

TABLE 5.48. Statistical contribution (in percentage) to the total error in the scaling factor α_{sig} for the non-resonant signal and three scalar-signal mass hypotheses, 500 GeV, 1000 GeV and 2000 GeV, in the resolved analysis. The values are extracted by calculating the difference in quadrature between the total statistical error and the error obtained after setting constant the normalisation factor of the background that dominates the region of interest.

determined. The contribution of the statistical uncertainty to the total uncertainty in α_{sig}, shown in Table 5.48, is decomposed into signal region statistics, top CR statistics and multijet CR statistics. The contribution of the systematic uncertainties to the total uncertainty is decomposed into the dominant components and shown in Table 5.49. The dominant systematic uncertainties vary across the mass range, but some of the most relevant ones are due to $t\bar{t}$ modelling, b-tagging systematic uncertainties, and those related to jet measurements.

Systematic source	Resolved analysis			
	Non-Res (%)	500 GeV (%)	1000 GeV (%)	2000 GeV (%)
$t\bar{t}$ modelling ISR/FSR	+30/–20	+10/–5	+7 / –4	+2/–2
Multijet uncertainty	+10/–10	+20/–10	+20 / –20	+30/–30
$t\bar{t}$ Matrix Element	+10/–10	—	—	—
W+jets modelling PDF	+4/–7	+10/–10	+2 / –6	+7/–5
W+jets modelling scale	+9/–10	+9/–4	+9 / –2	+20/–10
W+jets modelling gen.	+10/–8	+10/–10	+9 / –1	+9/–9
$t\bar{t}$ modelling PS	+3/–2	+30/–20	+20 / –20	+2/–2
b tagging	+30/–20	+11/–5	+7 / –6	+30/–30
JES/JER	+13/–20	+20/–20	+50 / –50	+10/–6
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft term res.	+20/–20	+8/–1	+9 / –7	+7/–7
Pile-up reweighting	+3/–10	+5/–3	+9 / –10	+6/–6
Total systematic	+60/–80	+70/–70	+60/–70	+40/–60

TABLE 5.49. Systematic contributions (in percentage) to the total error in the scaling factor α_{sig} for the non-resonant signal and three scalar-signal mass hypotheses, 500 GeV, 1000 GeV and 2000 GeV, in the resolved analysis. The first column quotes the source of the systematic uncertainty. The "$-$" symbol indicates that the specified source is negligible. The contribution is obtained by calculating the difference in quadrature between the total error in α_{sig} and that obtained by setting constant the nuisance parameter(s) relative to the contribution(s) under study.

m_X [GeV]	S	G^*_{KK} $(c = 1.0)$	G^*_{KK} $(c = 2.0)$	Total Bkg.	Data
2000	28 ± 0.5	36.4 ± 0.8	43.0 ± 0.7	1255 ± 27	1107

TABLE 5.50. Data event yields, and post-fit signal and background event yields in the final signal region for the boosted analysis and the scalar S particle hypothesis. The errors shown are the MC statistical and systematic uncertainties described in Section 5.5.10. For illustration a signal mass point of 2000 GeV is reported in the table. The signal samples are normalized to the expected upper limit cross sections.

5.6.2. Boosted analysis results

The boosted analysis applies the selection criteria described in Section 5.5.2. After applying the large-R jet mass requirement $90 < m_{\text{Large-}R\text{ jet}} < 140$ GeV, the m_{HH} distribution is reconstructed and its shape is fit to data using MC signal and background templates. The distribution is fit using 17 bins, with almost uniform width except at low and high m_{HH}, where the bin width is modified in order to have a MC statistical uncertainty smaller than 20%. All backgrounds, except multijet, are simulated using MC generators and normalized using the cross section of the simulated process. The multijet background is estimated using the ABCD method, and its normalisation obtained from this method is kept fixed in the fit. The bias due to possible signal contamination in the ABCD regions was studied and found to have negligible effect on the result. The integral of the m_{HH} distribution for the boosted analysis is shown in Table 5.50.

Systematic uncertainties affecting the m_{HH} shape are parameterised as linear functions of m_{HH}, and the function parameters are treated as nuisance parameters in the fit. Statistical uncertainties due to the limited sample sizes of the simulated background processes are taken into account in the fit by means of further nuisance parameters, which are parameterised by Poisson priors.

Uncertainty source	Boosted analysis			
	1500 GeV [%]	2000 GeV [%]	2500 GeV [%]	3000 GeV [%]
Data statistics	+50/–52	+59/–61	+64/–66	+70/–72
Total systematic	+87/–85	+81/–79	+76/–75	+71/–69
MC statistics	+42/–48	+42/–50	+39/–48	+39/–49
$t\bar{t}$ modelling	+29/–31	+36/–38	+40/–45	+32/–39
Multijet uncertainty	+11/–14	+19/–23	+16/–20	+11/–16
W+jets modelling	+27/–30	+8/–12	+11/–10	+11/–10
Single-top modelling	+22/–26	+5/–6	+4/–5	+5/–5
b tagging	+31/–19	+36/–22	+36/–17	+34/–14
JES/JER	+14/–14	+6/–6	+14/–11	+7/–9
Large-R jet	+29/–10	+27/–8	+27/–7	+29/–8

TABLE 5.51. Statistical and systematic contributions (in percentage) to the total error in the scaling factor α_{sig} in the boosted analysis for four mass hypotheses: 1500 GeV, 2000 GeV, 2500 GeV and 3000 GeV. The first column quotes the source of the uncertainty. The contribution is obtained by calculating the difference in quadrature between the total error in α_{sig} and that obtained by setting constant the nuisance parameter(s) relative to the contribution(s) under study.

The systematic uncertainties included in the fit are described in Section 5.5.10. The contribution of the systematic uncertainties to the total uncertainty is decomposed into the dominant components and summarized in Table 5.51. The most relevant systematic uncertainties are due to the limited size of the MC samples, the $t\bar{t}$ modelling and the b-tagging systematic uncertainties.

Sample	Yield	Stats Err	Systs Err
$t\bar{t}$	648.7	± 16.4	$^{+177.3(+27.3\%)}_{-169.2(-26.1\%)}$
W+Jets	217.0	± 6.5	$^{+104.3(+48.1\%)}_{-100.9(-46.5\%)}$
QCD	235.2	± 18.9	$^{+181.8(+77.3\%)}_{-181.8(-77.3\%)}$
Single-top	109.2	± 6.0	$^{+86.0(+78.8\%)}_{-85.8(-78.6\%)}$
Z+Jets	20.5	± 1.1	$^{+11.2(+54.6\%)}_{-10.9(-52.9\%)}$
Dibosons	24.4	± 1.9	$^{+15.3(+62.6\%)}_{-14.7(-60.1\%)}$
Prediction	1255.0	± 26.7	$^{+324.3(+25.8\%)}_{-311.3(-24.8\%)}$
Data	1107	-	-
Data/Pred	0.88	-	-

TABLE 5.52. Predicted and observed yields in the signal region. Detector modeling uncertainties, MC background modeling uncertainties are considered for the systematic uncertainties. The expected background yields are predicted from MC and no normalization factors are applied.

Figure 5.32 shows the m_{HH} distribution for data and the background components for the boosted analysis. Data are generally in good agreement with the background expectations within the quoted systematic errors. The signal m_{HH} distribution is shown in the figure for the scalar resonance. Figure 5.33 shows the observed and the expected upper limit on the production cross section of the scalar S particle.

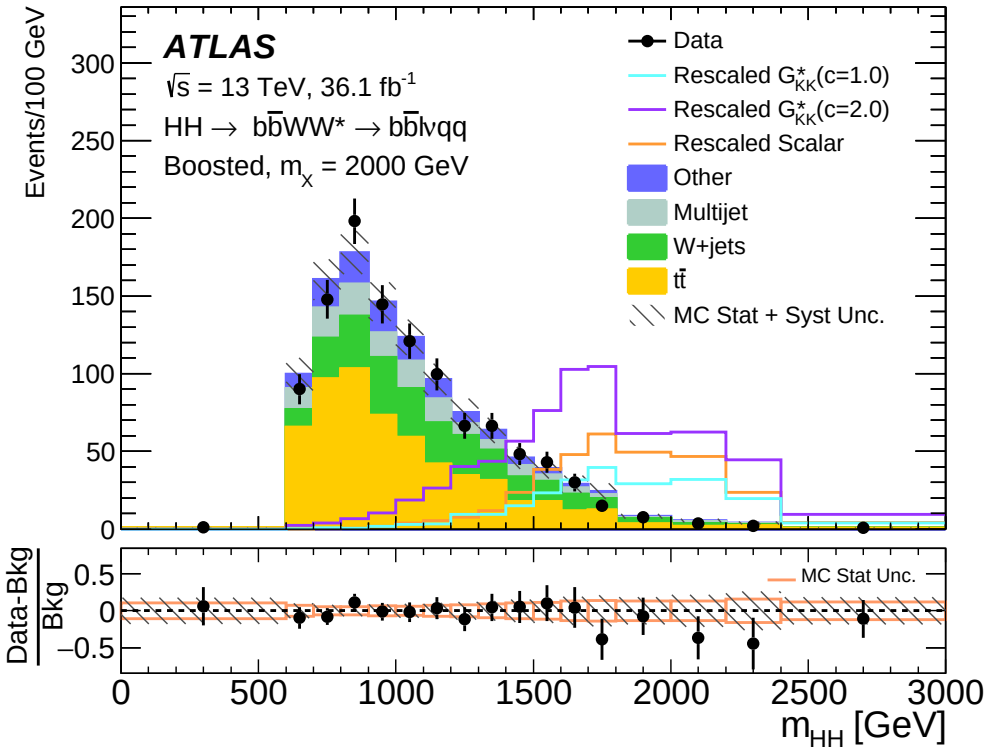

FIGURE 5.32. m_{HH} distributions after the global likelihood fit for the boosted analysis. The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty. The signals shown correspond to resonances of mass 2000 GeV. The scalar signal is multiplied by a factor of four with respect to the expected upper-limit cross section.

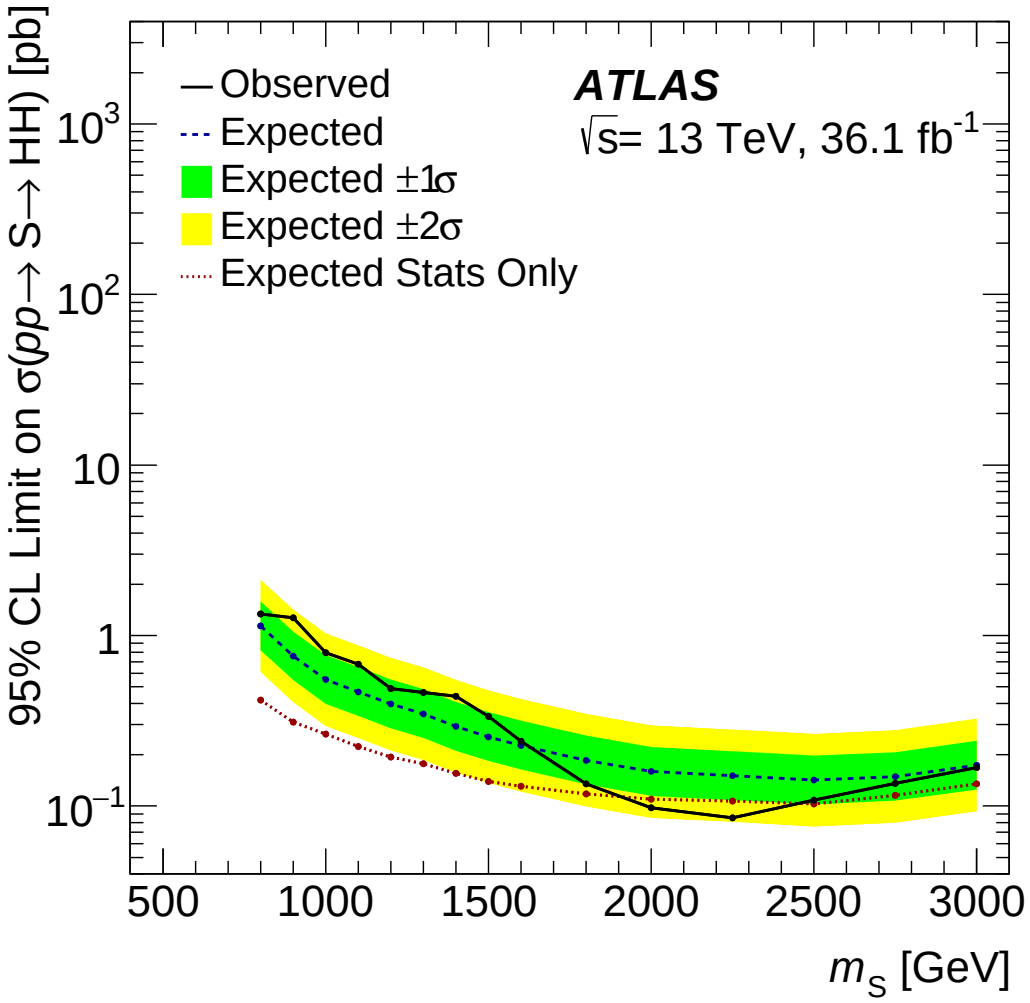

FIGURE 5.33. Expected and observed upper limits at 95% CL on the cross section of resonant pair production for the heavy scalar boson S model in the boosted analysis. The plot also shows the expected limits without including the systematic errors in order to show their impact.

5.6.3. Combined results

Results of the two analyses are summarized in Figure 5.34 for the scalar interpretation. The sensitivity of the boosted analysis is higher than the resolved analysis (the expected limit being lower) for masses larger than 1300 GeV in the

scalar interpretation. For masses lower than these values, the limits of the resolved analysis are presented in the figure, otherwise the boosted-analysis limits are shown. In addition, the expected limits of both analyses are shown near the mass values where the switch between the two limit curves occurs.

Finally, the observed upper limits on the production cross sections range from 5.6 pb for $m_X = 500$ GeV to 0.51 pb for $m_X = 3000$ GeV for the scalar signal model. No boosted analysis was performed for the non-resonant SM signal model.

For the non-resonant signal hypothesis the observed (expected) upper limit on the $\sigma(pp \to HH) \times \mathcal{B}(HH \to b\bar{b}WW^*)$ at 95% CL is:

$$\sigma(pp \to HH) \cdot \mathcal{B}(HH \to b\bar{b}WW^*) < 2.5 \left(2.5^{+1.0}_{-0.7}\right) \text{ pb},$$

which corresponds to 300 (300^{+100}_{-80}) times the SM predicted cross section.

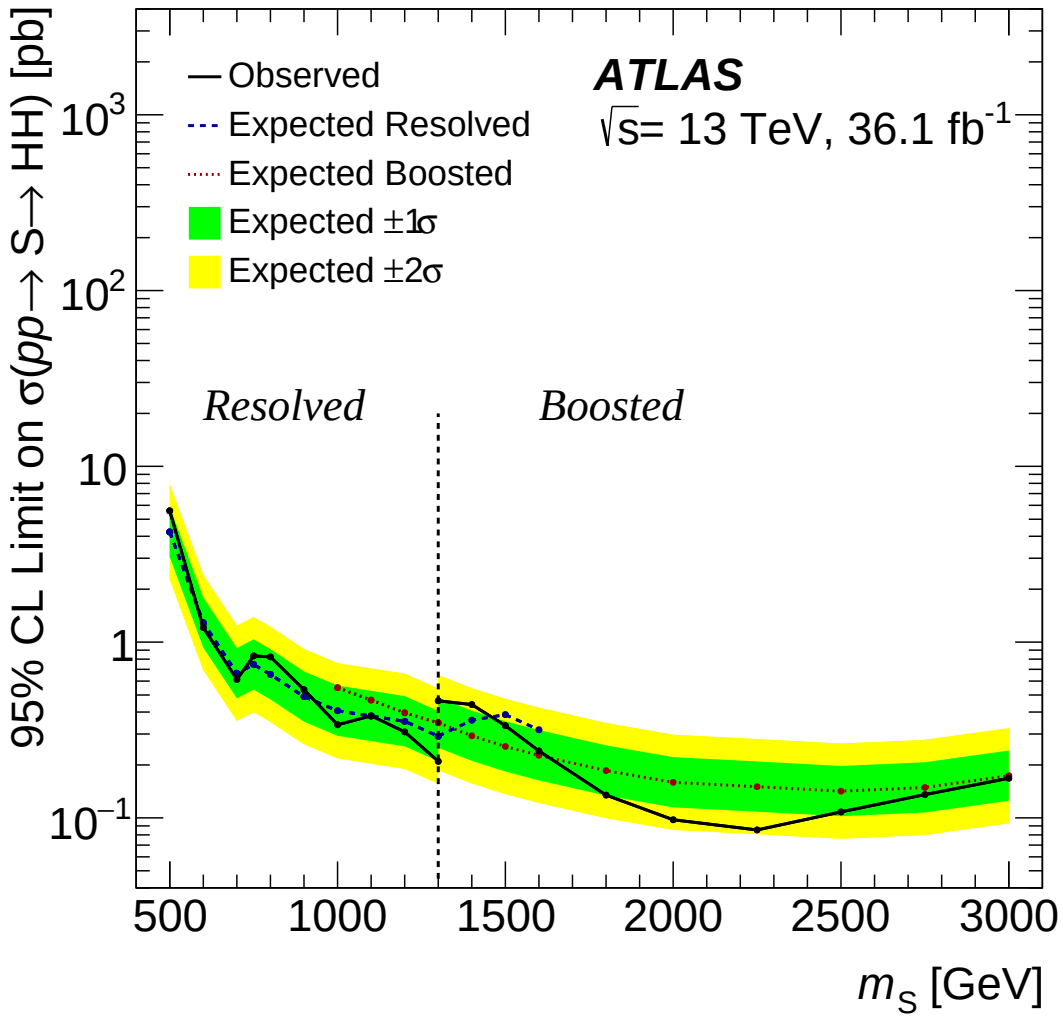

FIGURE 5.34. Expected and observed upper limits at 95% CL on the cross-section of the resonant scalar signal model hypotheses. The observed limits of the scalar signal models are switched at a mass of 1300 GeV. The expected limits of both analyses are shown in a region around the switching points. The switching point is chosen at the mass value where the boosted analysis becomes more sensitive than the resolved analysis.

5.7. Conclusion

A search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}WW^*$ decay mode is performed in the $b\bar{b}\ell\nu qq$ final state using pp collision data

corresponding to an integrated luminosity of 36.1 fb^{-1}, collected at $\sqrt{s} = 13$ TeV by the ATLAS detector at the Large Hadron Collider. No evidence of a significant excess of events over the background expectation is found. Limits are set on resonant production as a function of the resonance mass for a scalar resonance and for spin-2 gravitons in the mass range 500 to 3000 GeV. Any excesses seen are local excesses and need to be evaluated with the look elsewhere effect before anything can be said about a global significance. An upper limit is set on the cross section of non-resonant pair production $\sigma(pp \to HH) \cdot \mathcal{B}(HH \to b\bar{b}WW^*) < 2.5$ pb at 95% CL corresponding to 300 times the predicted SM cross section. Given the result of this work, in order to bring relevant sensitivity improvement to the HH non-resonant SM searches in this channel at the LHC and at future colliders, more advanced analysis techniques, development of new methods for the normalisation of the $t\bar{t}$ background, and a more refined estimation of the multijet background, need to be deployed.

CHAPTER VI

IMPROVEMENTS TO THE BOOSTED ANALYSIS

The boosted analysis described in Chapter V is optimized for the case where the $H \to b\bar{b}$ system is boosted and the $H \to WW^*$ system is resolved. However, at high resonant masses, one would expect both the $H \to b\bar{b}$ and the $H \to WW^*$ to be boosted. However, the semileptonic decay of the W boson pair adds an additional complication of having a lepton within the radius of a large-R jet. This chapter will describe a method for reconstructing this complex topology and the improvements it offers to the boosted, semi-leptonic $HH \to b\bar{b}WW^*$ analysis.

6.1. Motivation

The resonant analysis covers a large range of mass hypotheses, from 400 - 3000 GeV. As the resonant mass increases, the two Higgs systems become more and more collimated. Figure 6.1 shows the average distance between final state partons as a function of the resonant mass for simulated HH events. The WW bosons become very collimated (within $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.5$) around 1000 GeV.

On the $H \to b\bar{b}$ side, this is accounted for through the boosted analysis selection described in Secion 5.5.1. Moving toward the full Run II analysis, it is worthwhile to look at the potential gain from including a boosted $H \to WW^*$ selection. This "fully-boosted" analysis can be used in conjunction with the current boosted and resolved analysis to increase sensitivity and reach.

The "fully-boosted" analysis piggybacks off of the analysis presented in Chapter V. This means the data and Monte Carlo Samples , object reconstruction, and trigger

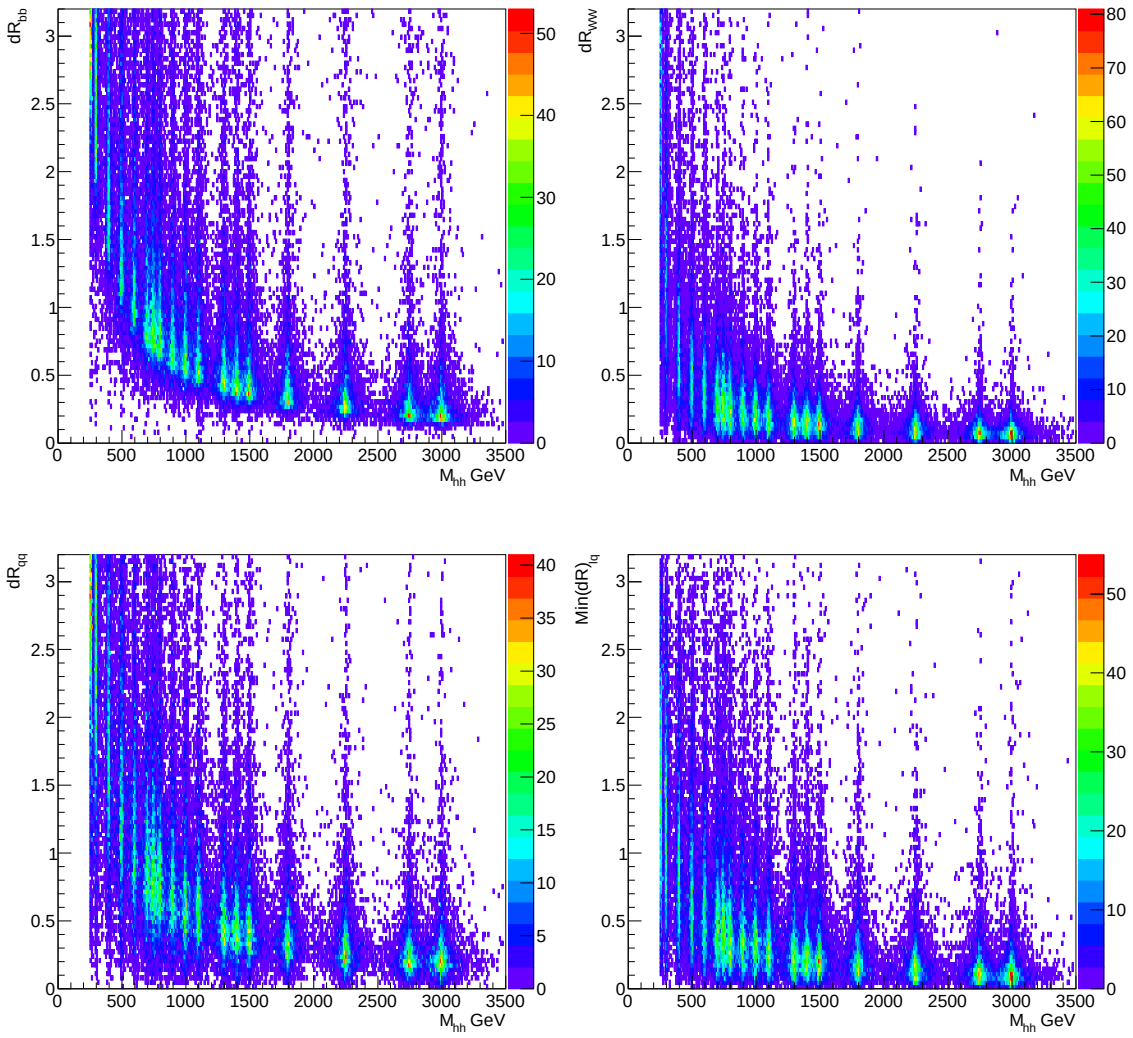

FIGURE 6.1. Distance between the two b partons (top left); the two W bosons (top right); the two light quarks (bottom left); and the lepton and the closest light quark (bottom right) for resonant HH production as a function of resonant mass.

requirements are the same as the previously presented analysis.

6.2. Event Reconstruction

Identically to the boosted analysis, Section 5.5, events are reconstructed by requiring at least one reconstructed lepton. To reconstruct the $H \to b\bar{b}$ candidate, there should be at least one large-R jet with $\Delta R > 1.0$ from the selected lepton. The highest of these large-R jets is selected as the $H \to b\bar{b}$ candidate. This large-R jet is then required to have at least two track jets associated to it. Events with a $H \to b\bar{b}$ in the range 30 GeV $< m_{\mathrm{bb}} <$ 300GeV are retained for further analysis.

To reconstruct the $H \to WW^*$ candidate, there should be at least one large-R jet with $\Delta R < 1.0$ from the selected lepton. This large-R jet is selected as the $H \to WW^*$ jet candidate. Once the $H \to b\bar{b}$ jet has been selected, they are split into either electron or muon channel for the full reconstruction.

Calorimeter jets are clusters of energy that are grouped together into an object based on distances. If an electron, which deposits the majority of its energy into the calorimeter, were to fall within the radius of a calorimeter jet, its energy should be measured as part of the jet energy. Using this, it is possible to use a single large-R to measure the energy of the $W \to qq$ system and the electron. With the large-R jet and the $\not{E}_T$, it is possible to fully reconstruct the $H \to WW^*$ system. The neutrino is reconstructed using a similar method as in Section 5.4.3. Imposing the relation:

$$m_h^2 = (p^\nu + p^{\mathrm{large-Rjet}})^2 \tag{6.1}$$

the neutrino p_z can be reconstructed using the relations:

$$p_E^\nu = E^\nu = \sqrt{P_T^2 + p_z^2} \quad p_x^\nu = P_T cos(\phi) \quad p_y^\nu = P_T sin(\phi)$$

where ϕ is the azimuthal angle of the $\not{E}_T$, E^{ν} the neutrino energy, p_x and p_y the two transverse spatial components of the neutrino momentum.

Muons do not deposit a significant amount of energy in the calorimeters, this means we cannot use the same reconstruction as the electrons. Instead, the muons are treated in a more traditional fashion. In the muon channel, the large-R jet contains the energy of the $W \to qq$ system. The muon is reconstructed using the MS and ID information and the neutrino is reconstructed identically to Section 5.4.3, with the hadronic W as a single object.

Figure 6.2 shows a diagram of the event topology after the event reconstruction.

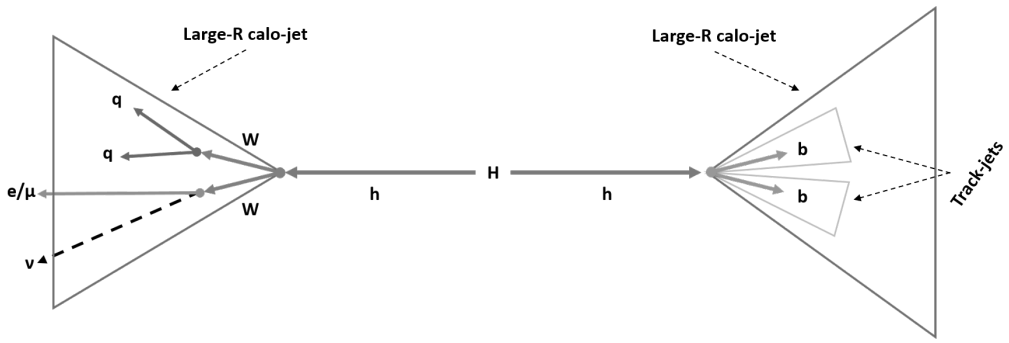

FIGURE 6.2. Diagram of the fully-boosted event topology

6.3. Event selection

After the event is reconstructed, a b-tag requirement is applied to the two track-jets in the $H \to b\bar{b}$ candidate. The $\not{E}_T$ is required to be more than 50 GeV to reject events from QCD background. Finally a p_T requirement is placed on the $H \to WW^*$ candidate. It is important to cut on the same physics objects. To accomplish this, an "adjusted p_T " ($p_T^{'}$) cut of 250 GeV is applied. Table 6.1 defines the $p_T^{'}$ for both channels.

Lepton Channel	Alternative p_T definition
Muon Channel	$p_T^{'} = p_T^{\text{Large-Rjet}}$
Electron channel	$p_T^{'} = \sqrt{(p_x^{\text{Large-R jet}} - p_x^{\text{electron}})^2 + (p_y^{\text{Large-R jet}} - p_y^{\text{electron}})^2}$

TABLE 6.1. Alternative p_T definition for the electron and muon channels. A cut of $p_T^{'} > 250$ GeV is applied to the selected $H \to WW^*$ large-R jet.

6.3.1. Comparison of reconstruction objects

Figure 6.3 and 6.4 show a comparison of the old boosted analysis reconstruction and the new fully-boosted analysis for the electron and muon selection respectively. The m_{WW^*} distribution shows a much shorter tail in the fully-boosted analysis than in the old boosted analysis. Along with this, the m_{HH} distribution has a sharper peak around the resonant mass. A sharper peak allows for tighter cuts around the signal mass, leading to a better signal/background ratio.

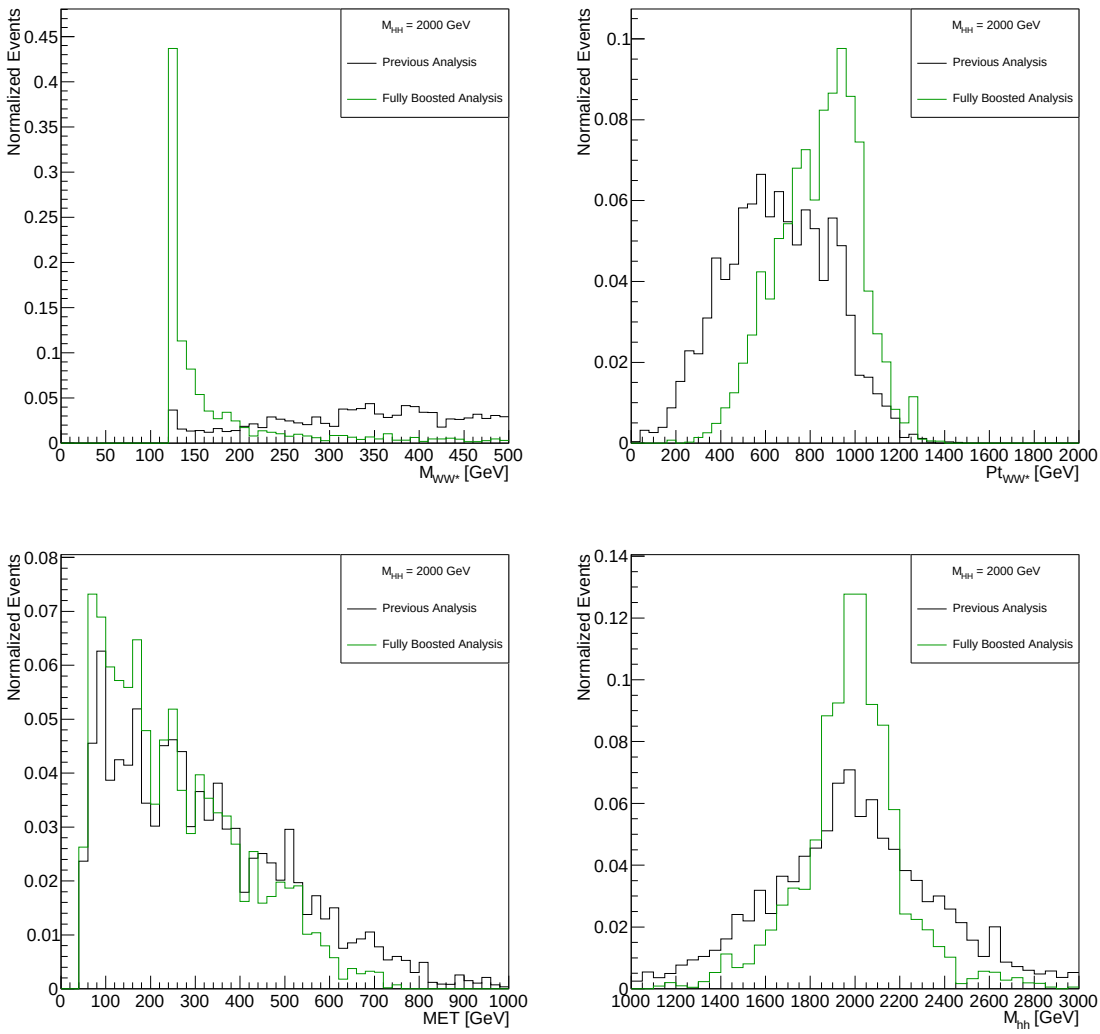

FIGURE 6.3. Comparison of reconstructed $H \to WW^*$ mass (top left), $H \to WW^*$ p_T (top right), $\not{E}_T$(bottom left), and HH mass for the previous boosted analysis reconstruction and new fully boosted selection for a resonant signal with a mass of 2000 GeV in the electron channel.

6.3.2. Signal Region Definition

As with the boosted analysis in Section 5.5, the $h \to b\bar{b}$ candidate must have a jet mass in the window 90 GeV $< m_{bb} <$ 140 GeV to be considered in the signal region (SR). The previous boosted analysis included a b-jet veto for all jets outside of

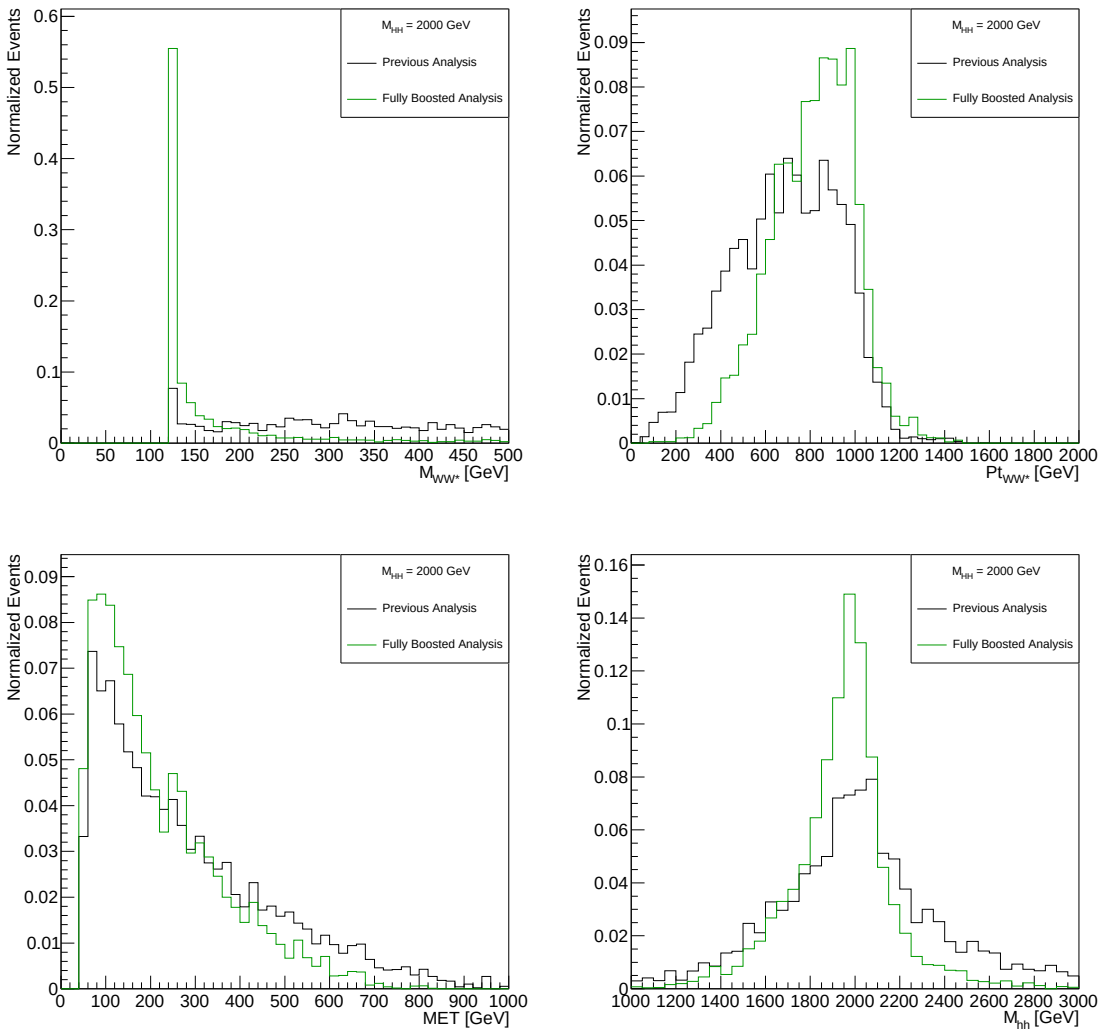

FIGURE 6.4. Comparison of $H \to WW^*$ mass (top left), $H \to WW^*$ p_T (top right), $\not{E}_T$(bottom left), and HH mass for the previous boosted analysis reconstruction and new fully boosted selection for a resonant signal with a mass of 2000 GeV in the muon channel.

the $H \to b\bar{b}$ candidate. In order to increase statistics for the QCD multijet estimate, this requirement was removed from the fully boosted analysis.

6.3.3. mBB Control Region

To check the modeling of the background, a control region is created with an inverted m_{bb} cut. Section 6.3.5 shows various kinematic distributions in the mBB control region to check the background shape with respect to data.

6.3.4. Multijet Background

The QCD multijet background is estimated using the same data-driven background as the boosted analysis. The ABCD method with the regions defined as:

- Region A: $\not{E}_T >$ 50 GeV, $|d_0^{\text{sig}}| < 2.0$
- Region B: $\not{E}_T <$ 50 GeV, $|d_0^{\text{sig}}| < 2.0$
- Region C: $\not{E}_T >$ 50 GeV, $|d_0^{\text{sig}}| > 2.0$
- Region D: $\not{E}_T <$ 50 GeV, $|d_0^{\text{sig}}| > 2.0$

6.3.4.1. Yield Prediction

Table 6.2 lists the MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region B and D before the $H \to b\bar{b}$ mass cut is applied and Table 6.3 shows the yields in Region C mBB control region and signal region.

Table 6.4 shows the ratio in the electron channel and muon channel. The predicted yields of the QCD multijet background in the mBB control region and signal region are presented in table 6.5. The QCD multijet background is estimated to be 13% of the total background in the signal region (Table 6.8).

	Region B		Region D	
Samples	Electron	Muon	Electron	Muon
$t\bar{t}$	138.7 ± 7.3	146.1 ± 8.3	14.4 ± 2.7	6.3 ± 1.5
W+Jets	27.3 ± 1.7	27.9 ± 1.9	1.1 ± 0.3	2.0 ± 0.4
Single-top	7.1 ± 1.4	5.2 ± 1.4	0.2 ± 0.2	0.4 ± 0.3
Z+Jets	18.3 ± 0.8	9.3 ± 0.5	1.9 ± 0.3	0.9 ± 0.2
Dibosons	3.1 ± 0.5	1.3 ± 0.3	0.2 ± 0.1	0.2 ± 0.1
Total Prompt	194.4 ± 7.7	189.9 ± 8.7	17.8 ± 2.7	9.8 ± 1.6
Data	274.0 ± 16.6	218.0 ± 14.8	34.0 ± 5.8	30.0 ± 5.5
QCD	79.6 ± 18.3	28.1 ± 17.1	16.2 ± 6.4	20.2 ± 5.7

TABLE 6.2. MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region B and D. The multijet yield is calculated by subtracting the estimated total prompt lepton backgrounds from the observed data. The statistical uncertainty on the yields is shown.

	mBBcr		SR	
Samples	Electron	Muon	Electron	Muon
$t\bar{t}$	21.7 ± 4.1	12.8 ± 2.1	6.4 ± 1.4	5.4 ± 1.2
W+Jets	3.4 ± 0.6	2.3 ± 0.3	1.4 ± 0.3	1.2 ± 0.3
Single-top	0.7 ± 0.5	0.3 ± 0.2	0.4 ± 0.3	0.2 ± 0.2
Z+Jets	1.0 ± 0.2	0.4 ± 0.1	0.4 ± 0.1	0.2 ± 0.1
Dibosons	0.2 ± 0.1	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.0
Total Prompt	26.9 ± 4.1	15.9 ± 2.2	8.7 ± 1.5	7.0 ± 1.3
Data	53.0 ± 7.3	33.0 ± 5.7	12.0 ± 3.5	20.0 ± 4.5
QCD	26.1 ± 8.4	17.1 ± 6.1	3.3 ± 3.8	13.0 ± 4.7

TABLE 6.3. MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region C mBBcr and SR. The multijet yield is calculated by subtracting the estimated total prompt lepton backgrounds from the observed data. The statistical uncertainty on the yields is shown.

Multijet yield in region	Electron	Muon
N_B^{QCD}	79.6 ± 18.3	28.1 ± 17.1
N_D^{QCD}	16.2 ± 6.4	20.2 ± 5.7
$N_B^{\mathrm{QCD}}/N_D^{\mathrm{QCD}}$	4.9 ± 2.62 (46.0%)	1.4 ± 0.94 (67.2%)

TABLE 6.4. Multijet yields in region B and region D and also the ratio of the yields for each lepton channel. The error on the $\frac{N_B^{\mathrm{QCD}}}{N_D^{\mathrm{QCD}}}$ ratio is propagated from the statistical uncertainties on the multijet yields in each region.

Multijet yield in region	Electron	Muon
SR		
N_C^{QCD}	3.3 ± 3.8	13.0 ± 4.7
N_A^{QCD}	16.4 ± 20.5 (126.8%)	18.1 ± 13.9 (76.8%)
mBBcr		
N_C^{QCD}	26.1 ± 8.4	17.1 ± 6.1
N_A^{QCD}	128.3 ± 79.4 (61.9%)	23.9 ± 18.1 (75.7%)

TABLE 6.5. Multijet yield in region C and predicted yield in region A in the SR. The error on N_A^{QCD} are propagated from the error on the $N_B^{\text{QCD}}/N_D^{\text{QCD}}$ ratio and statistical uncertainty on N_C^{QCD} yield. The numbers in brackets are the relative uncertainty in percentage.

6.3.4.2. Shape prediction

The shape prediction follows the same procedure as described in Section 5.5.4.2.

6.3.4.3. Multijet yield uncertainties

6.3.4.3.1. Statistical The uncertainty on the predicted yield of the multijet background is determined by propagating the statistical uncertainty of the $\frac{N_B^{\text{QCD}}}{N_D^{\text{QCD}}}$ ratio, as shown in Table 6.4, and the statisical uncertainty on the multijet yield in region C (N_C^{QCD}), as in Table 6.5.

6.3.4.3.2. 1-tag/2-tag jet mass acceptance Another source of uncertainty on the multijet yield is the the difference of acceptance of the large-R jet mass cut between the single b-tag (1-tag) and two b-tag (2-tag) selections. This uncertainty is included since the template for or the multijet shape prediction uses the multijet shape from the 1-tag region C. Table 6.6 shows the acceptance of the large-R jet mass signal and mBB control region selection in the multijet 1-tag region C and 2-tag region C yields. The relative difference between the acceptance in 1-tag region

C and in 2-tag region C is considered as an uncertainty on the normalization of the QCD multijet prediction.

Region	Electron	Muon
SR		
1-tag $\frac{N_{\mathrm{SR}}}{N_{\mathrm{Inc}}}$	32%	23 %
2-tag $\frac{N_{\mathrm{SR}}}{N_{\mathrm{Inc}}}$	11%	43 %
Rel. difference between 1-tag and 2-tag	26.4 %	46 %
mBBcr		
1-tag $\frac{N_{\mathrm{mBBcr}}}{N_{\mathrm{Inc}}}$	68%	77 %
2-tag $\frac{N_{\mathrm{mBBcr}}}{N_{\mathrm{Inc}}}$	89%	56 %
Rel. difference between 1-tag and 2-tag	24 %	37 %

TABLE 6.6. The acceptance of the large-R jet mass signal region selection on the multijet 1-tag and 2-tag region C. $N_{\mathrm{SR}}(N_{\mathrm{Inc}})$ is the multijet yield with (without) the signal region large-R jet mass selection.

6.3.5. mBB Control Region Plots

Figures 6.5 and s6.6 hows the m_{HH} distribution in the mBB control region for the electron and muon channels. The figure shows a reasonable agreement between data and background in the mBB control region but with large uncertainties. Appendix J has a more complete set of kinematic plots.

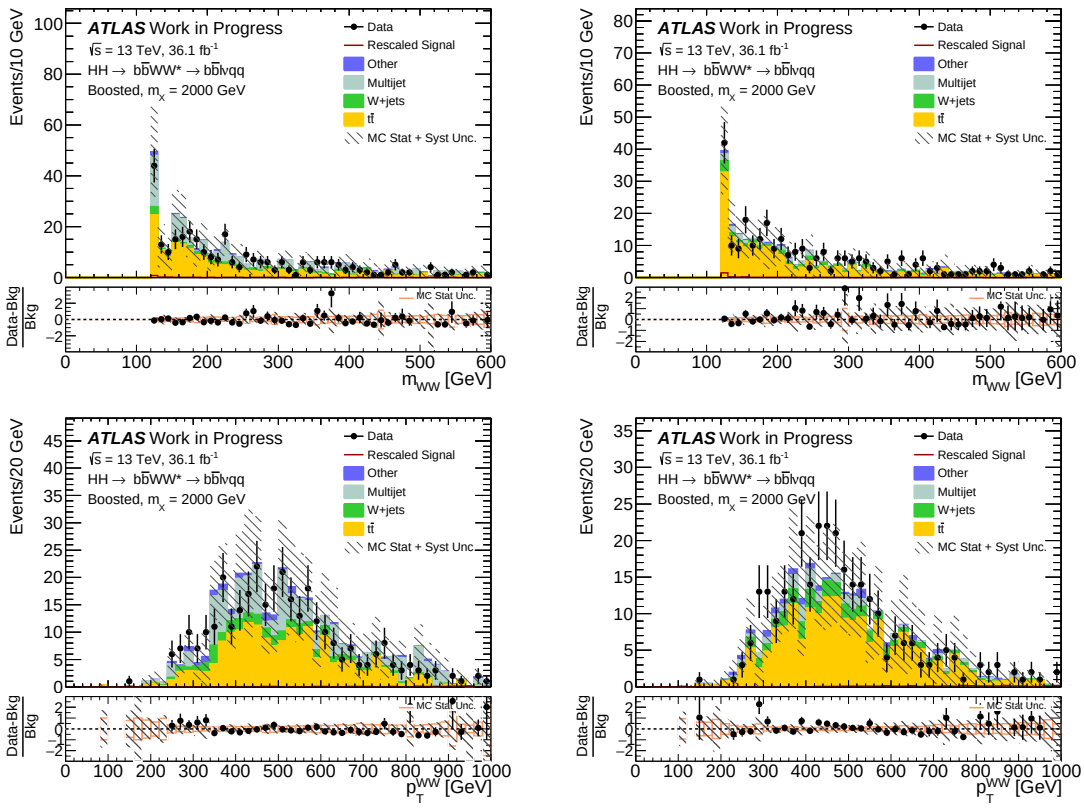

FIGURE 6.5. m_{WW} (top) and $p_T^W W$ (bottom) distribution in the mBB control region for the electron (left) and muon (right) channels.

m_X [GeV]	S	Total Bkg.	Data
2000	26.8 ± 0.5	271.0 ± 10.7	268

TABLE 6.7. Data event yields, and signal and background event yields in the final signal region for the boosted analysis and the scalar S particle hypothesis. The errors shown are the MC statistical uncertainties. For illustration a signal mass point of 2000 GeV is reported in the table. The signal samples are normalized to the expected upper limit cross sections.

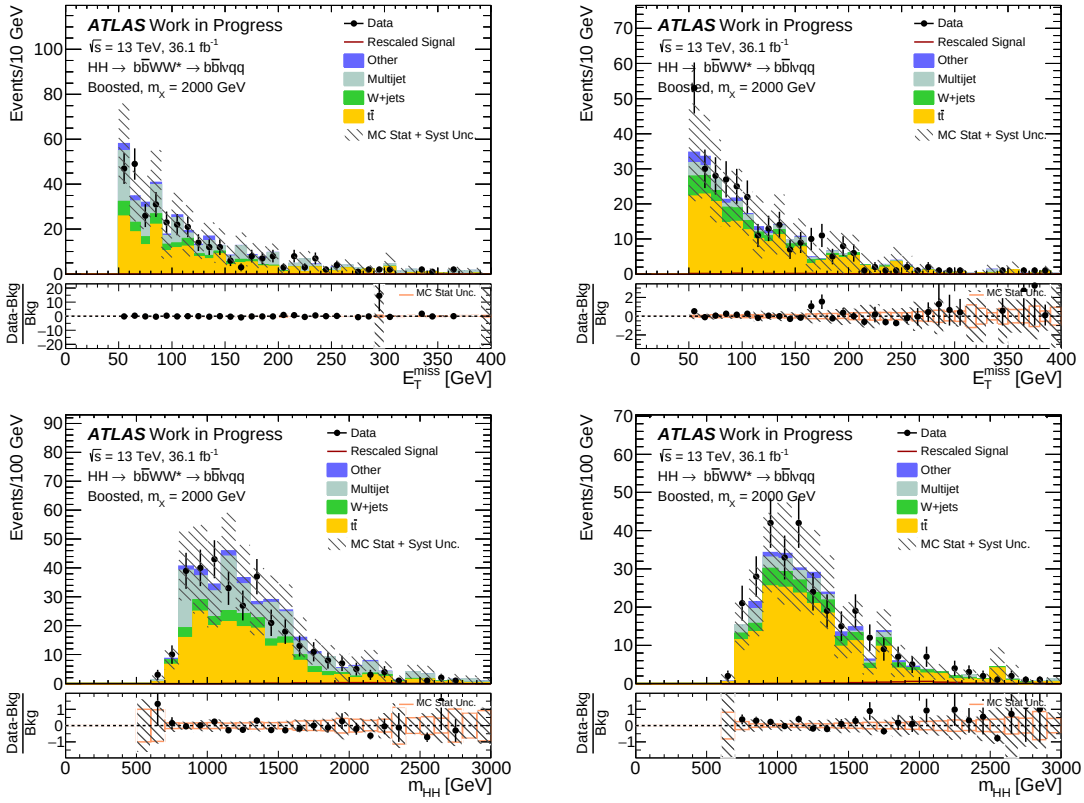

FIGURE 6.6. $\not{E}_T$(top) and m_{HH} (bottom) distribution in the mBB control region for the electron (left) and muon (right) channels.

6.4. Results

The fully-boosted analysis reconstructs the m_{HH} distribution and the shape is fit to data using MC signal and background templates. The distribution is fit using

Sample	Yield	Stats Unc
$t\bar{t}$	187.7	± 8.8
W+Jets	33.7	± 1.9
QCD	34.5	± 5.5
Single-top	7.0	± 1.3
Z+Jets	4.7	± 0.4
Dibosons	3.3	± 0.6
Prediction	271.0	± 10.7

TABLE 6.8. Predicted and observed yields in the Signal region. Detector modeling uncertainties, MC background modeling uncertainties and QCD background modeling uncertainties from ABCD method are considered for the systematic uncertainties.

17 bins, with almost uniform width except at low and high m_{HH}. All backgrounds, except multijet, are simulated using MC generators and normalised using the cross section of the simulated process. The multijet background is estimated using the ABCD method, and its normalisation obtained from this method is kept fixed in the fit. The bias due to possible signal contamination in the ABCD regions was studied and found to have negligible effect on the result. The integral of the m_{HH} distribution for the boosted analysis is shown in Table 6.7.

The dominant components of the systematic uncertainties are summarized in Table 6.9.

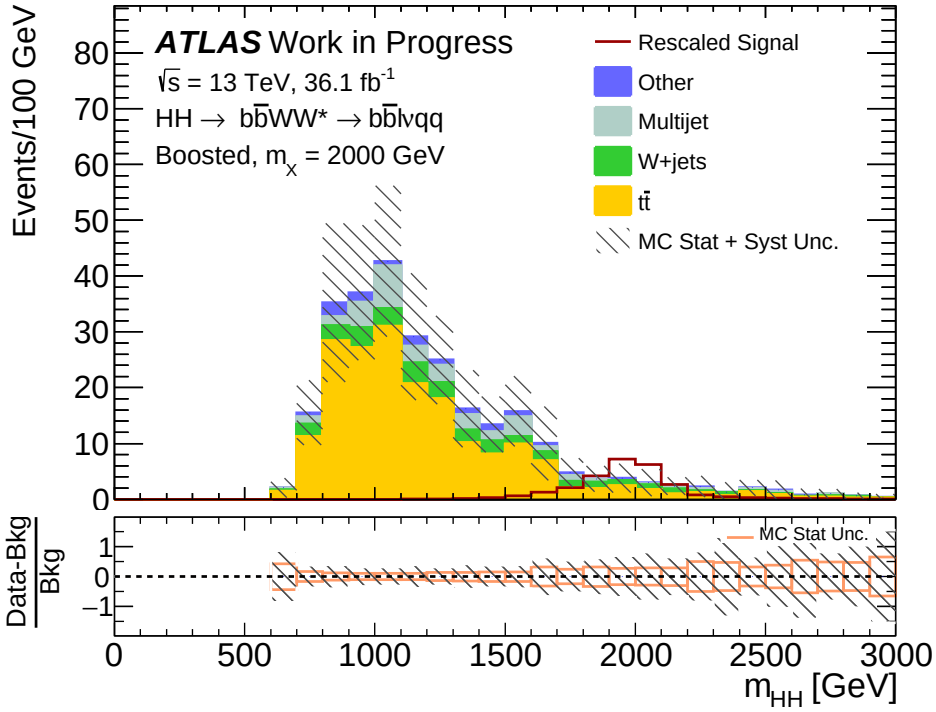

FIGURE 6.7. The m_{HH} distribution for data and background in the final signal region.

Figure 6.7 shows the m_{HH} distribution for data and the background components for the boosted analysis. Appendix K shows a more complete set of kinematic plots. Data are generally in good agreement with the background expectations within the quoted systematic errors. The signal m_{HH} distribution is shown in the figure for the scalar resonance. Figure 6.8 shows the observed and the expected upper limit on the production cross section of the scalar S particle.

Uncertainty	Up/Down
SysFT_EFF_Eigen_Light_0_AntiKt2PV0TrackJets__1down	-12.9/12.5
SysFT_EFF_Eigen_C_0_AntiKt2PV0TrackJets__1down	-12.6/12.1
SysFT_EFF_Eigen_C_0_AntiKt2PV0TrackJets__1up	11.3/-11.9
SysFT_EFF_Eigen_Light_0_AntiKt2PV0TrackJets__1up	11.3/-11.9
SysFATJET_Medium_JET_Comb_Baseline_Kin__1up	-6.47/5.95
SysFATJET_Medium_JET_Comb_Baseline_Kin__1down	5.83/-6.43
SysFT_EFF_Eigen_B_0_AntiKt2PV0TrackJets__1down	-3.49/2.97
SysFT_EFF_Eigen_B_1_AntiKt2PV0TrackJets__1down	-3.32/2.8
SysFT_EFF_Eigen_C_1_AntiKt2PV0TrackJets__1down	-2.97/2.45
SysFT_EFF_Eigen_B_0_AntiKt2PV0TrackJets__1up	2.39/-2.94
SysFT_EFF_Eigen_B_1_AntiKt2PV0TrackJets__1up	2.23/-2.78
SysFT_EFF_Eigen_C_1_AntiKt2PV0TrackJets__1up	1.9/-2.45
SysFATJET_Medium_JET_Comb_Tracking_Kin__1down	1.77/-2.38
SysFATJET_Medium_JET_Comb_Tracking_Kin__1up	-2.2/1.68
SysFT_EFF_extrapolation_AntiKt2PV0TrackJets__1up	-2.07/1.59
SysFATJET_JMR__1up	1.41/-1.97
SysFT_EFF_extrapolation_from_charm_AntiKt2PV0TrackJets__1up	-1.62/1.1
SysFT_EFF_extrapolation_AntiKt2PV0TrackJets__1down	0.963/-1.54
SysPRW_DATASF__1down	-1.47/0.94
SysJET_SR1_JET_GroupedNP_1__1down	0.764/-1.33
SysJET_SR1_JET_GroupedNP_1__1up	-1.2/0.678
SysFATJET_JER__1up	0.574/-1.16
SysFT_EFF_Eigen_Light_1_AntiKt2PV0TrackJets__1up	-1.12/0.576
SysFT_EFF_extrapolation_from_charm_AntiKt2PV0TrackJets__1down	0.553/-1.1
SysFATJET_Medium_JET_Comb_TotalStat_Kin__1down	-1.03/0.498
Total Up	27.2
Total Do	27.9

TABLE 6.9. List of dominant systematic uncertainties for the fully boosted analysis. The full list of systematic uncertainties is listed in Appendix H

6.5. Conclusion

The fully-boosted $HH \to b\bar{b}WW^*$ semi-leptonic analysis can strengthen the limit on the cross section of the resonant pair production by approximately a factor of two over the limit shown in Figure 5.34 for a large range of resonant masses.

Several improvements can be made to the fully-boosted analysis moving toward the full Run II analysis. The isolated lepton trigger that is used for these studies is not suitable for highly collimated topologies and likely leads to a loss of signal efficiency at high resonant masses. A large-R jet trigger or a non-isolated lepton trigger would allow for leptons much closer to other objects and could increase the number of events, especially in the high resonant mass region. Additionally, the QCD multijet estimation suffers from a lack of statistics in the C and D regions. This gives a large error on the QCD background estimation. A different method of QCD estimation, such as the matrix method, may be more suitable.

The fully-boosted analysis can be combined with the previous resolved and boosted analysis for the full Run-2 search. By using all three analysis strategies, it is possible to maximize the reach of the analysis across the entire mass range, even extending the range higher. The main findings of the fully boosted analysis are as follows:

- A Large-R jet offers a better resolution for the $H \to WW^*$ reconstruction than the resolved reconstruction for $m_S > 1000$ GeV.
- The lepton-inside-of-jets reconstruction accurately accounts for the lepton energy for boosted topologies.
- An increase in resolution translates to a more sensitive search and should be prioritized for the $HH \to b\bar{b}WW^*$

    * This can be done through improving the neutrino reconstruction.

– Additionally, the ABCD data driven QCD estimate is limited by statistics and needs to be improved.

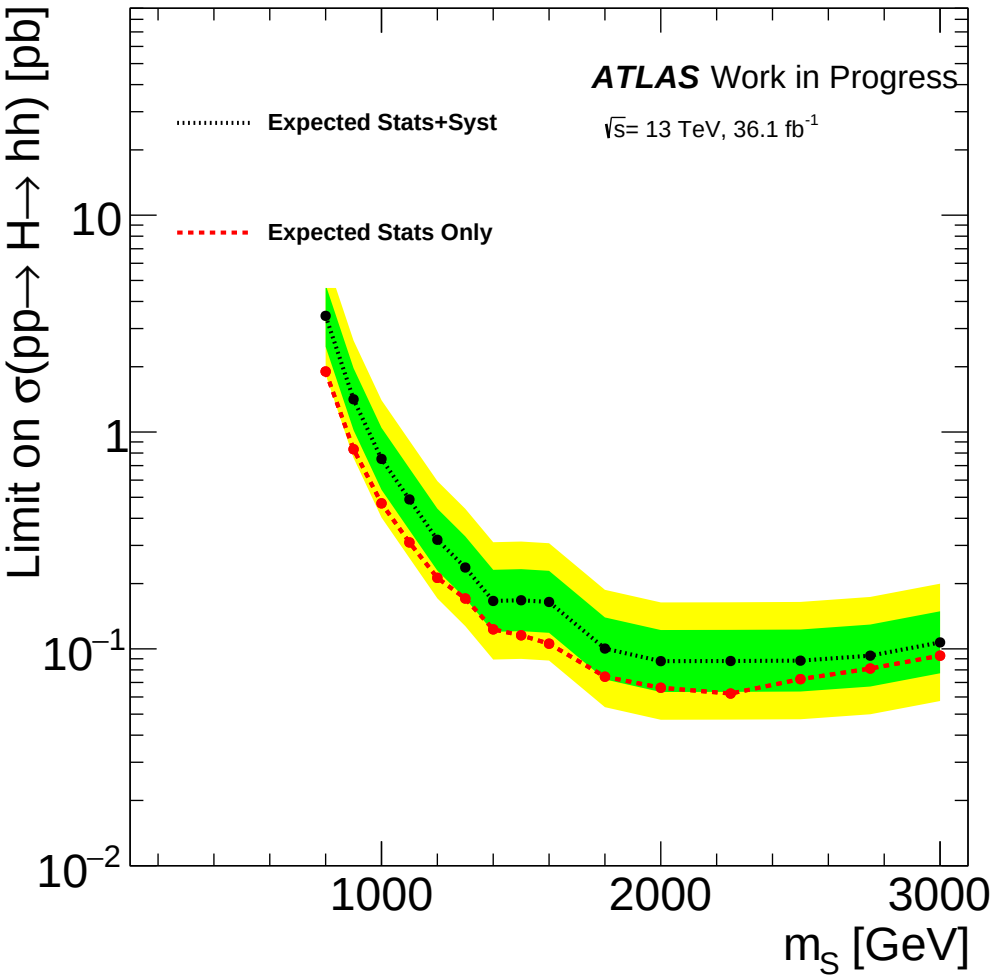

FIGURE 6.8. Expected and observed upper limits at 95% CL on the cross section of the resonant pair production for heavy scalar boson S model. The expected limits without including the systematic errors is also shown to show their impact.

CHAPTER VII

CONCLUSION

This dissertation presented a search for anomalous production of di-Higgs events not predicted in the SM. The SM prediction for di-Higgs production was observed and no evidence for new physics was found.

A search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}WW^*$ decay mode is done in the $b\bar{b}l\nu qq$ final state using pp collision data with an integrated luminosity of 36.1 fb^{-1} collected at $\sqrt{s} = 13$ TeV by the ATLAS detector at the LHC. No excess of events over the background only expectation is found. Limits are set on resonant and non-resonant production.

In addition to the complete analyses presented in this dissertation, an complimentary event reconstruction is presented. This new fully boosted analysis offers roughly a factor of two increase in sensitivity for the same dataset and is a promising addition to the $HH \to b\bar{b}WW^* \to b\bar{b}l\nu qq$ analysis for the full Run I analysis.

APPENDIX A

DERIVATION (HIGG5D2)

Pre-selection is applied to both data and MC samples using the derivation framework in order to reduce the xAOD sample size. We use the HIGG5D2 derivation for our DxAOD sample production. More information on the derivation framework can be found in the Higgs group's Twiki.[1]

[1] https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HSG2xAODMigration

APPENDIX B

COMPLETE LIST OF MC SAMPLES

The following MC samples have been used to simulate the signal and SM backgrounds at the center-of-mass energy of 13 TeV. The detailed information including cross section, k-factor (where applicable), and the corresponding job options files are listed in the Twiki page[1]

B.0.1. Signal Samples

mc15_13TeV.342053.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_hh_WWbb.merge.DAOD_HIGG5D2.e4392_a766_a821_r7676_p2949

mc15_13TeV.343764.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m260_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343766.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m300_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343769.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m400_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343771.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m500_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343772.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m600_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343773.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m700_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343774.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m750_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343775.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m800_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343776.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m900_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343777.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1000_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343778.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1100_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343779.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1200_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343780.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1300_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343781.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1400_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343782.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1500_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343783.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1600_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343784.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1800_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343785.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m2000_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343786.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m2250_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343787.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m2500_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343788.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m2750_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

mc15_13TeV.343789.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m3000_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949

B.0.2. Background Samples

B.0.2.1. $t\bar{t}$

mc15_13TeV.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.DAOD_HIGG5D2.e3698_s2608_s2183_r7725_r7676_p2949

B.0.2.2. Sherpa W+jets

mc15_13TeV.364156.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

[1] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/CentralMC15ProductionList

mc15_13TeV.364157.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364158.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364159.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364160.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364161.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364162.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364163.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364164.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364165.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364166.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364167.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364168.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV500_1000.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364169.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV1000_E_CMS.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364170.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364171.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364172.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364173.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364174.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364175.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364176.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364177.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364178.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364179.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364180.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364181.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364182.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV500_1000.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364183.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV1000_E_CMS.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364184.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364185.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364186.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364187.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364188.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364189.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364190.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364191.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364192.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364193.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CVetoBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364194.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CFilterBVeto.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364195.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_BFilter.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364196.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV500_1000.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364197.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV1000_E_CMS.merge.DAOD_HIGG5D2.e5340_s2726_r7772_r7676_p2949

B.0.2.3. Z+jets

mc15_13TeV.364100.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CVetoBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364101.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CFilterBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364102.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_BFilter.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364103.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CVetoBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364104.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CFilterBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364105.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_BFilter.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364106.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CVetoBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364107.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CFilterBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364108.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_BFilter.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364109.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CVetoBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364110.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CFilterBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364111.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_BFilter.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364112.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV500_1000.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364113.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV1000_E_CMS.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364114.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CVetoBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364115.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CFilterBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364116.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_BFilter.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364117.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CVetoBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364118.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CFilterBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364119.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_BFilter.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364120.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364121.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CFilterBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364122.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_BFilter.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364123.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CVetoBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364124.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CFilterBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364125.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_BFilter.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364126.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364127.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364128.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CVetoBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364129.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CFilterBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364130.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_BFilter.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364131.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CVetoBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364132.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CFilterBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364133.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_BFilter.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364134.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CVetoBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364135.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CFilterBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364136.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_BFilter.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364137.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CVetoBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364138.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CFilterBVeto.merge.DAOD_HIGG5D2.e5313_s2726_r7772_r7676_p2949

mc15_13TeV.364139.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_BFilter.merge.DAOD_HIGG5D2.e5313_s2726_r7772_r7676_p2949

mc15_13TeV.364140.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364141.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

B.0.2.4. Dibosons

mc15_13TeV.361091.Sherpa_CT10_WplvWmqq_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949

mc15_13TeV.361092.Sherpa_CT10_WpqqWmlv_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949

mc15_13TeV.361093.Sherpa_CT10_WlvZqq_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949

mc15_13TeV.361094.Sherpa_CT10_WqqZll_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949

mc15_13TeV.361095.Sherpa_CT10_WqqZvv_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949

mc15_13TeV.361096.Sherpa_CT10_ZqqZll_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949

mc15_13TeV.361097.Sherpa_CT10_ZqqZvv_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949

B.0.2.5. Single Top

mc15_13TeV.410011.PowhegPythiaEvtGen_P2012_singletop_tchan_lept_top.merge.DAOD_HIGG5D2.e3824_s2608_s2183_r7725_r7676_p2949

mc15_13TeV.410012.PowhegPythiaEvtGen_P2012_singletop_tchan_lept_antitop.merge.DAOD_HIGG5D2.e3824_s2608_s2183_r7725_r7676_p2949

mc15_13TeV.410013.PowhegPythiaEvtGen_P2012_Wt_inclusive_top.merge.DAOD_HIGG5D2.e3753_s2608_s2183_r7725_r7676_p2949

mc15_13TeV.410014.PowhegPythiaEvtGen_P2012_Wt_inclusive_antitop.merge.DAOD_HIGG5D2.e3753_s2608_s2183_r7725_r7676_p2949

mc15_13TeV.410025.PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_top.merge.DAOD_HIGG5D2.e3998_s2608_s2183_r7725_r7676_p2949

mc15_13TeV.410026.PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_antitop.merge.DAOD_HIGG5D2.e3998_s2608_s2183_r7725_r7676_p2949

B.0.2.6. Single Higgs

mc15_13TeV.342282.PowhegPythia8EvtGen_CT10_AZNLOCTEQ6L1_ggH125_inc.merge.DAOD_HIGG5D2.e4850_a766_a821_r7676_p2949

mc15_13TeV.342285.Pythia8EvtGen_A14NNPDF23LO_ZH125_inc.merge.DAOD_HIGG5D2.e4246_s2608_s2183_r7772_r7676_p2949

B.0.2.7. QCD Multijet

mc15_13TeV.344715.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt30_50.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949

mc15_13TeV.344716.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt50_80.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949

mc15_13TeV.344717.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt80_130.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949

mc15_13TeV.344718.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt130_200.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949

mc15_13TeV.344719.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt200_E_CMS.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949

APPENDIX C

DATA SAMPLES

In order to assure good data quality, events with bad detector conditions, namely where large part of the detectors were missing from data acquisition due to problems during a run, or when the performance of the detectors were affected by large noise, have been rejected from the data analysis. A GRL selection taken from data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml and data16_13TeV.periodAllYear_DetStatus-v88-pro20-21_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml is applied. Moreover, incomplete events or events with bad detector information are rejected.

APPENDIX D

LEPTON SELECTION OPTIMISATION

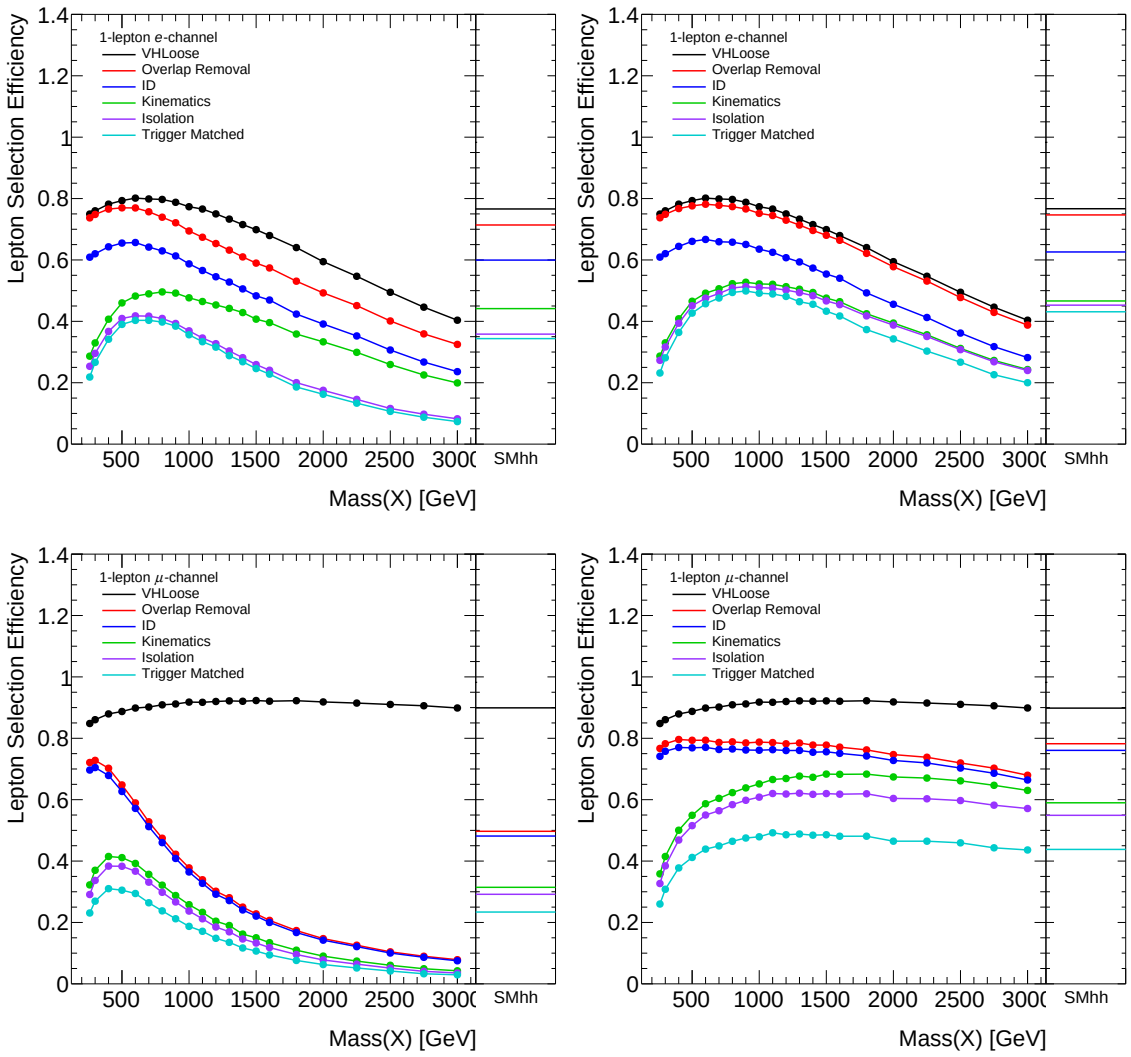

FIGURE D.1. Lepton selection effciency as a function of signal Scalar resonance mass for the electron-channel (top) and the muon-channel (bottom). The efficiency for the SM di-Higgs signal sample (SMhh) is also shown. The plots on the left are efficiences for the lepton selections as used in the ICHEP 2016 analysis while the plot on the right are the optimised baseline lepton selections.

APPENDIX E

LEPTON AND B-JET OVERLAP REMOVAL: IMPACT ON SIGNAL EFFICIENCY

The overlap removal procedure, as presented in Sec. 5.3.5 requires calorimeter-jets to be removed from the event if it is within. Due to the presence of b-jets from the $h \to b\bar{b}$ decay, it is possible for a reconstructed muon or electron to be in close proximity of the b-jets, due to semi-leptonic decays of the b-hadron, and the b-jets fail the overlap removal criteria. Figure E.1 shows the signal efficiency to find the leading and sub-leading b-jet in the event. In this study, the b-jets are identified by requiring at least one b-hadron within the b-jets.

The efficiency loss due to electrons, which pass the VHLooseElectron selection, to be within ΔR ¡ 0.2 of the b-jets are negligible as can be seen in the top figures. For muons that pass VHLooseMuon selection, the impact is slightly more pronounced but still small ($\sim$3%). Note that in the jet-muon overlap removal procedure, additional requirements are imposed on the muon and the overlapping jet before deciding to remove the jet and this is expected to mitigate the lost of efficiency. The impact of the additional requirements are not studied here.

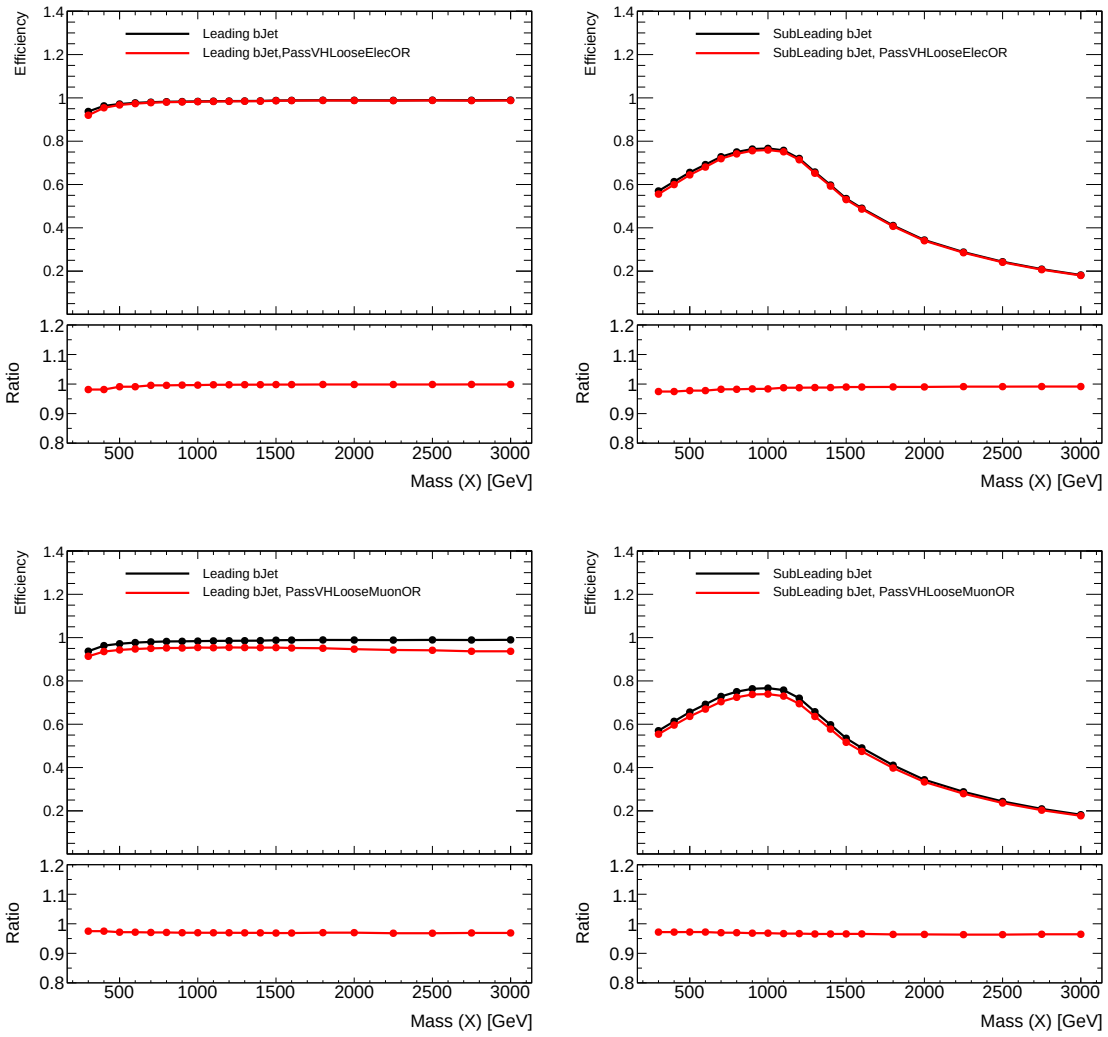

FIGURE E.1. The signal efficiencies to find the leading (left) and sub-leading (right) jets which has at least 1 b-hadron within the jets. The red curves correspond to the requirement that the jets do not have a VHLoose electrons (top) and muons (bottom) within ΔR ¡ 0.2 from the jet axis.

APPENDIX F

SOLVING FOR NEUTRINO LONGITUDINAL MOMENTUM

Neutrinos are not detected directly using the ATLAS detector. Instead, their transverse momentum is determined using conservation of momentum. The longitudinal momentum of neutrinos can be calculated in terms of the momentum and masses of other particles in the event, but it is a solution to a quadratic equation, and hence there are two possible solutions. Please see Section 5.4.3 for the equation. We investigated the efficiencies of several different methods of picking the solution.

Several methods for selecting the sign were tested on truth level using the signal Monte Carlo samples for resonant di-Higgs production with mass of 700 GeV, 2000 GeV and 5000 GeV. The choice of sign is said to be correct if the solution is within 10% of the truth value if the truth value is larger than 100 GeV or within 10 GeV if the truth value is smaller than 100 GeV. This criterion was chosen to take into account the fact that we do not need to know the exact momentum, just the momentum up to errors from the detector resolution.

F.0.1. W mass method

The mass of the W that decays leptonically can be calculated from the momentum of the lepton and neutrino since $P_W = P_l + P_\nu$. The mass of the leptonic W we calculate with the chosen solution depends on our choice of sign. As figure F.3 shows, the W mass distribution has peaks at 40 and 80 GeV. So for this method, we will use the solution that minimizes $|m_W - 80|$ or $|m_W - 40|$.

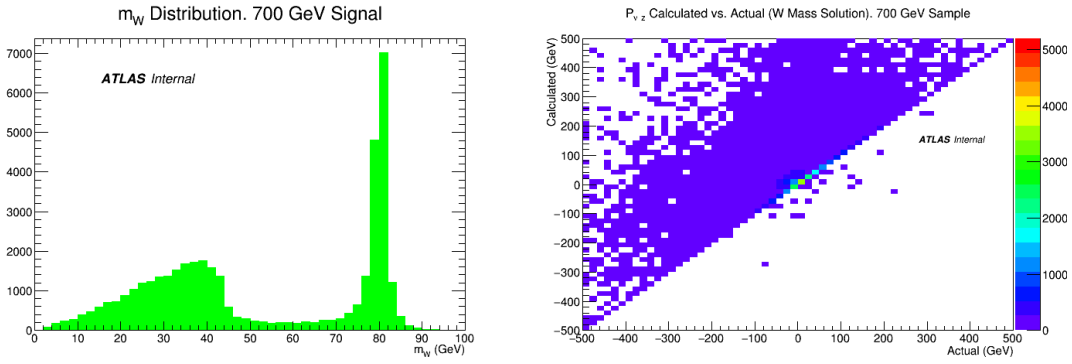

FIGURE F.1. m_W Distribution and the solution using W mass method.

F.0.2. η solution

The next method we tested was picking the sign in the solution that is the opposite of the sign of η of the lepton. This was motivated by studying the same problem for single W production and noticing that there is a correlation between the the correct sign choice and the sign of η. This correlation is shown in figure F.5. It is interesting to note that this method is mathematically equivalent to using the opposite sign of the longitudinal momentum of the lepton. This can be verified by looking at the definition of $\eta_l = \frac{1}{2}\log\frac{p_l+p_{l,z}}{p_l-p_{l,z}}$. If $p_{l,z} < 0$ then $\frac{p_l+p_{l,z}}{p_l-p_{l,z}} < 1$ so $\eta_l < 0$ and similarly if $p_{l,z} > 0$ then $\eta_l > 0$.

We also tested using the opposite sign of η of the $lq\bar{q}$ system. Using the additional information from the quark antiquark pair allows us to pick the correct sign more often than just using η_l, F.4

F.0.3. Δ R Method

We also investigated minimising $\Delta R(l,\nu)$. This is motivated by the idea that the neutrino and the charged lepton produced from the same W should be near each if the W boson is boosted, which is often the case for heavy resonances decaying into

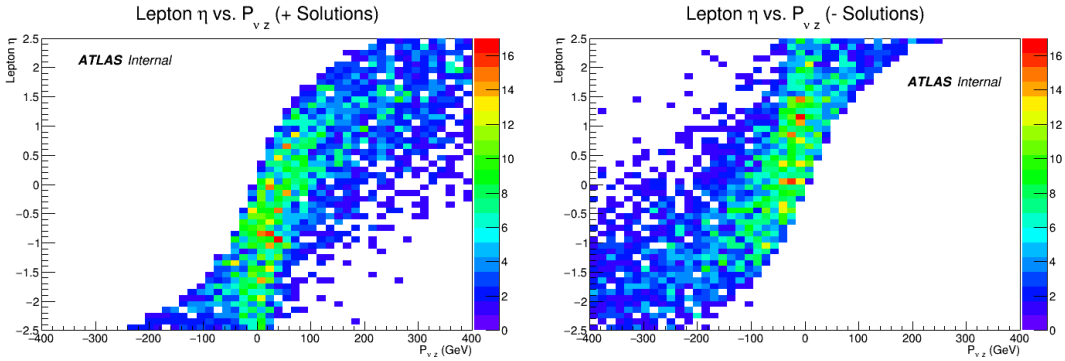

FIGURE F.2. Single W η_l vs Correct Sign Choice

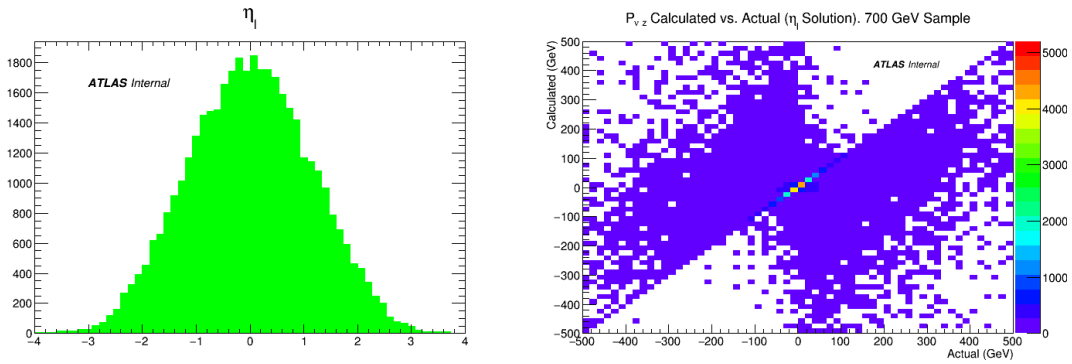

FIGURE F.3. η_l and the solution using η_l method in 700 GeV resonant sample.

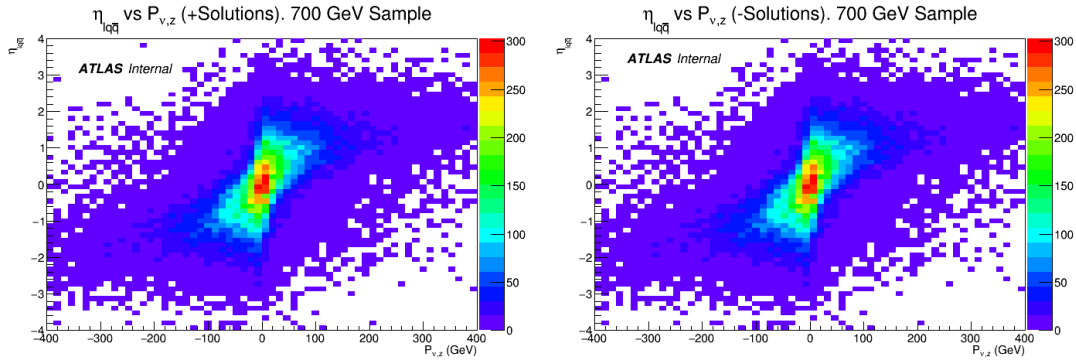

FIGURE F.4. Higgs pair $\eta_{lq\bar{q}}$ vs Correct Sign Choice

W boson. Fig. F.6 shows the $\Delta R(l,\nu)$ and the solution using $\Delta R(l,\nu)$ method in 700 GeV resonant sample.

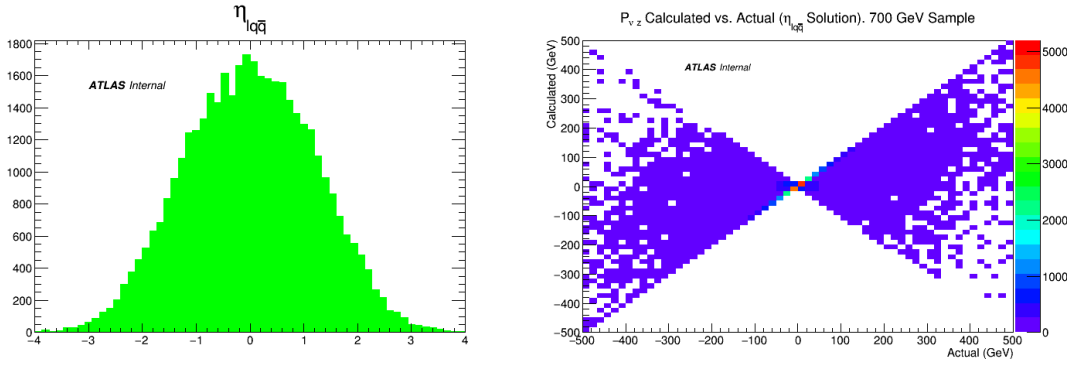

FIGURE F.5. η_l and the solution using η_l method in 700 GeV resonant sample.

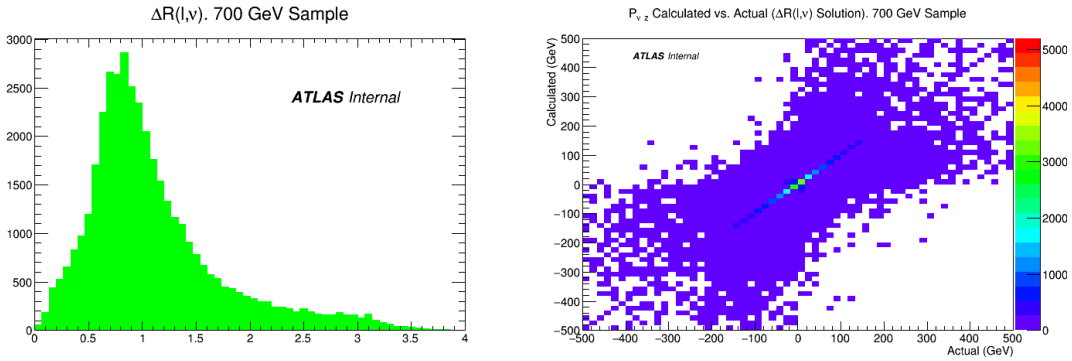

FIGURE F.6. $\Delta R(l,\nu)$ and the solution using $\Delta R(l,\nu)$ method in 700 GeV resonant sample.

F.0.4. Results

The fraction of the time each method picks the correct sign is given as a percentage in the table below.

Method	700 GeV Sample	2000 GeV Sample	5000 GeV Sample
m_W	55.7	56.3	58.3
η_l	52.9	47.6	47.4
$\eta_{l\nu q\bar{q}}$	59.9	60.5	63.3
$\Delta R(l,\nu)$	57	65	72.7

APPENDIX G

FREEZING B AND D REGIONS IN QCD ESTIMATE

This appendix summarizes the study undertaken to select regions to 'freeze' the cuts in the B and D regions used in the ABCD estimation for QCD background. For each selection, the B/D values diverge from the value calculated at the beginning of the selection when the statistics in the B and/or D region drops significantly. To avoid large statistical errors in the normalization calculated for the multi-jet contribution in the A region, the yields from the B and D regions used in the ABCD calculation are frozen, i.e. no further cuts applied after the earliest cut in the selection which has a B/D ratio consistent with the last statistically stable B/D ratio. Using this method, the multi-jet modelling is kept as close as possible to the phase space in the final signal region while taking advantage of higher statistics earlier in the cutflow. All numbers in this study were conducted with a $t\bar{t}$ normalization factor equal to 1.0, i.e. no data-driven normalization was applied to $t\bar{t}$.

QCD B/D Values, Non-resonant Selection			
mww	bbpt210	bbpt300	wwpt250
0.36 ± 0.01	0.35 ± 0.07	0.49 ± 0.30	0.14 ± 0.10
QCD B/D Values, Low Mass (m700) Selection			
mww	bbpt210	wwpt250	hh700
0.36 ± 0.01	0.35 ± 0.07	0.04 ± 0.02	0.11 ± 0.07
QCD B/D Values, High Mass Selection			
bbpt350	wwpt250	drww15	hh2000
0.29 ± 0.04	0.28 ± 0.04	0.22 ± 0.06	1.24 ± 0.86

TABLE G.1. Values calculated for B/D at each stage in the non-resonant, low mass, and high mass selections. The estimate of multi-jet contribution in the A region uses the R value calculated after the selection described in the text.

The QCD and total background yields obtained in the $m_{b\bar{b}}$ control region without freezing the B and D regions are shown in Tables G.2, G.3, and G.4.

Non-resonant Selection in $m_{b\bar{b}}$ Control Region, No B/D Freezing				
Sample	mww	bbpt210	bbpt300	wwpt250
QCD	13310.5 ± 500.3	250.2 ± 30.6	24.6 ± 3.0	54.8 ± 6.7
Background Sum	43849.0± 509.2	1017.9± 33.7	192.8± 7.1	153.2± 8.4
Data	43902.0	1069.0	206.0	138.0

TABLE G.2. QCD and total background yields for the non-resonant selection without freezing the selection cuts used in the B and D regions, i.e. the yields in the B and D region after each cut are used in the ABCD calculation up until the $m_{b\bar{b}}$ cut. Non-monotonic QCD yields are observed.

Low Mass (m700) Selection in $m_{b\bar{b}}$ Control Region, No B/D Freezing				
Sample	mww	bbpt210	wwpt250	hh700
QCD	13310.5 ± 500.3	250.2 ± 30.6	585.3 ± 71.7	54.8 ± 6.7
Background Sum	43849.0± 509.2	1017.9± 33.7	843.7± 72.1	104.7± 7.6
Data	43902.0	1069.0	367.0	89.0

TABLE G.3. QCD and total background yields for the low mass selection without freezing the selection cuts used in the B and D regions, i.e. the yields in the B and D region after each cut are used in the ABCD calculation up until the $m_{b\bar{b}}$ cut.

High Mass Selection in $m_{b\bar{b}}$ Control Region, No B/D Freezing				
Sample	bbpt350	wwpt250	drww15	hh2000
QCD	1538.7 ± 252.7	1359.5 ± 75.9	486.4 ± 27.1	4.6 ± 0.3
Background Sum	14719.1± 258.9	12463.5± 91.8	3671.3± 38.5	222.8± 7.1
Data	14862.0	12450.0	3761.0	250.0

TABLE G.4. QCD and total background yields for the high mass selection without freezing the selection cuts used in the B and D regions, i.e. the yields in the B and D region after each cut are used in the ABCD calculation up until the $m_{b\bar{b}}$ cut.

The B/D ratio at the last cut in each selection have large errors (near or larger than 100%) and are found to be unstable. The yields after freezing the B and D regions to their yields after the earliest selection cut with a B/D ratio consistent

with the last statistically stable B/D ratio (after $p_T^{bb} > 210$ GeV for the non-resonant and low-mass selections and after $p_T^{WW} > 250$ GeV for the high mass selection) are shown in Tables G.5, G.6, and G.7.

Non-resonant Selection in $m_{b\bar{b}}$ Control Region, B/D Frozen after $p_T^{bb} > 210$				
Sample	mww	bbpt210	bbpt300	wwpt250
QCD	13310.5 ± 500.3	250.2 ± 30.6	33.7 ± 4.1	21.4 ± 2.6
Background Sum	43849.0± 509.2	1017.9± 33.7	201.9± 7.6	119.8± 5.7
Data	43902.0	1069.0	206.0	138.0

TABLE G.5. QCD and total background yields for the non-resonant selection after freezing the selection cuts used in the B and D regions after requiring $p_T^{bb} > 210$ GeV. Monotonic QCD yields are now observed.

Low Mass (m700) Selection in $m_{b\bar{b}}$ Control Region, B/D Frozen after $p_T^{bb} > 210$				
Sample	mww	bbpt210	wwpt250	hh700
QCD	13310.5 ± 500.3	250.2 ± 30.6	72.4 ± 8.9	16.3 ± 2.0
Background Sum	43849.0± 509.2	1017.9± 33.7	330.7± 12.1	66.2± 4.1
Data	43902.0	1069.0	367.0	89.0

TABLE G.6. QCD and total background yields for the low mass (m700) selection after freezing the selection cuts used in the B and D regions after requiring $p_T^{bb} > 210$ GeV.

High Mass Selection in $m_{b\bar{b}}$ Control Region, B/D Frozen after $p_T^{WW} > 250$ GeV				
Sample	bbpt350	wwpt250	drww15	hh2000
QCD	1538.7 ± 252.7	1359.5 ± 75.9	392.7 ± 21.9	20.7 ± 1.2
Background Sum	14719.1± 258.9	12463.5± 91.8	3577.5± 35.0	238.9± 7.2
Data	14862.0	12450.0	3761.0	250.0

TABLE G.7. QCD and total background yields for the low mass selection after freezing the selection cuts used in the B and D regions after requring $p_T^{WW} > 250$ GeV.

After freezing, the yields in the non-resonant $m_{b\bar{b}}$ control region are monotonically decreasing as expected, and the absolute statistical error on the QCD

estimate in the final signal region is significantly reduced compared to the yields obtained without freezing the B and D regions.

G.1. QCD Lepton Flavour Composition after Preselection Criteria for σ_{d0} distribution

We estimate the lepton flavour composition in the QCD sample and SM signal sample just after the preselection criteria. By using the σ_{d0} distribution we identify the lepton flavour and the origin of each lepton. The majority of the muons results to be from bottom meson while most of the electrons come from photon conversion (about 70%). The impact of the σ_{d0} cut on the two population is mostly independent from the lepton flavour, namely the cut $|\sigma_{d0}| \leq 2.0$ removes for muon leptons (76 $\pm$ 100)% for QCD sample and (6.9 $\pm$ 1.1)% for signal sample, while for the electons these fractions are (27.9 $\pm$ 5.0)% for QCD sample and (5.7 $\pm$ 0.9)% for signal. The Table G.8 shows the complete lepton flavour composition.

Electrons	
Photon Conversion	65.3%
Not Defined	15.0%
Bottom Meson	11.9%
Dalitz Decay	4.6%
Others	$\leq$ 0.1%
Muons	
Bottom Meson	99%
Charm Barion	1%

TABLE G.8. Lepton flavor composition in QCD sample. The lepton origin is reported for each event passing the preselection criteria.

APPENDIX H

STABILITY OF $t\bar{t}$ NORMALIZATION IN QCD ESTIMATE

This appendix contains tables tables summarizing a study performed to test the stability of the normalization factor applied to $t\bar{t}$ events when calculated iteratively with the ABCD method for QCD estimation. Since both the ABCD method and the $t\bar{t}$ normalization are calculated using data-driven techniques, and they contaminate each other control region, the aim is to check that the $t\bar{t}$ normalization and the QCD estimate converge to stable values after a small number of iterations. The iterative procedure begins with the QCD estimate set to zero and continues as follows:

- Calulate $t\bar{t}$ NF at appropriate cut **before** QCD is calculated
- Start at beginning of cutflow, apply $t\bar{t}$ NF, and calculate QCD
- Move through cutflow to $t\bar{t}$ NF cut
- Calculate new $t\bar{t}$ NF
- Repeat steps 2 - 4

If the procedure converges, the estimate is said to be stable. The following tables show the iteration for the non-resonant, low mass, and high mass selections. Iteration 0 starts with zero QCD in the background estimate. Yields are provided for QCD and $t\bar{t}$ after the first selection cut (noted as sample_first) and in the final $m_{b\bar{b}}$ signal region (noted as sample_SR). The QCD yield at the $t\bar{t}$ normalization cut is also provided. The final column shows the calculated normalization factor for $t\bar{t}$. For all selections, convergence (change in $t\bar{t}$ normalization $<5\%$) is observed after just a few iterations.

Iteration	nQCD_first	n_QCD_nf	nQCD_SR	ttbar_first	ttbar_SR	ttbarNF
0	0.00	0.00	0.00	24637.64	65.92	1.43 ± 0.02
1	4226.57	273.95	3.30	35249.05	94.31	0.92 ± 0.02
2	15765.74	1381.06	19.04	22708.20	60.75	0.94 ± 0.03
3	15311.64	1335.16	18.33	23145.35	61.92	0.94 ± 0.03
4	15254.28	1329.37	18.24	23200.59	62.07	0.94 ± 0.03
5	15247.34	1328.67	18.23	23207.28	62.09	0.94 ± 0.03
6	15246.50	1328.58	18.23	23208.09	62.09	0.94 ± 0.03
7	15246.40	1328.57	18.23	23208.18	62.09	0.94 ± 0.03

TABLE H.1. Iteration of $t\bar{t}$ normalization and QCD estimate for the non-resonant selection. The '_first' region is defined after the $m_{WW} < 130$ GeV cut and the '_nf' region is defined after requiring $p_T^{bb} > 150$ GeV. Convergence is observed after two iterations.

Iteration	nQCD_first	n_QCD_nf	nQCD_SR	ttbar_first	ttbar_SR	ttbarNF
0	0.00	0.00	0.00	24637.64	11.28	1.43 ± 0.02
1	4226.57	273.95	1.13	35249.05	16.14	0.92 ± 0.02
2	15765.74	1381.06	8.14	22708.20	10.40	0.94 ± 0.03
3	15311.64	1335.16	7.81	23145.35	10.60	0.94 ± 0.03
4	15254.28	1329.37	7.76	23200.59	10.62	0.94 ± 0.03
5	15247.34	1328.67	7.76	23207.28	10.63	0.94 ± 0.03
6	15246.50	1328.58	7.76	23208.09	10.63	0.94 ± 0.03
7	15246.40	1328.57	7.76	23208.18	10.63	0.94 ± 0.03

TABLE H.2. Iteration of $t\bar{t}$ normalization and QCD estimate for the low mass selection. The '_first' region is defined after the $m_{WW} < 130$ GeV cut and the '_nf' region is defined after requiring $p_T^{bb} > 150$ GeV. Convergence is observed after two iterations.

Iteration	nQCD_first	n_QCD_nf	nQCD_SR	ttbar_first	ttbar_SR	ttbarNF
0	0.00	0.00	0.00	8810.24	6.70	1.26 ± 0.02
1	92.73	40.56	0.11	11133.25	8.47	0.99 ± 0.02
2	1595.54	541.28	1.40	8679.56	6.60	1.03 ± 0.03
3	1238.18	432.83	1.12	9063.72	6.90	1.03 ± 0.02
4	1196.43	419.79	1.09	9108.80	6.93	1.03 ± 0.02
5	1190.60	417.96	1.09	9115.10	6.94	1.03 ± 0.02
6	1189.77	417.70	1.09	9115.99	6.94	1.03 ± 0.02
7	1189.65	417.66	1.09	9116.12	6.94	1.03 ± 0.02

TABLE H.3. Iteration of $t\bar{t}$ normalization and QCD estimate for the low mass selection. The '_first' region is defined after the $p_T^{bb} > 350$ GeV cut and the '_nf' region is defined after requiring $p_T^{WW} > 360$ GeV. Convergence is observed after two iterations.

APPENDIX I

FULLY BOOSTED SYSTEMATIC UNCERTAINTIES

Uncertainty	Up/Down
SysFT_EFF_Eigen_Light_0_AntiKt2PV0TrackJets__1down	-12.9/12.5
SysFT_EFF_Eigen_C_0_AntiKt2PV0TrackJets__1down	-12.6/12.1
SysFT_EFF_Eigen_C_0_AntiKt2PV0TrackJets__1up	11.3/-11.9
SysFT_EFF_Eigen_Light_0_AntiKt2PV0TrackJets__1up	11.3/-11.9
SysFATJET_Medium_JET_Comb_Baseline_Kin__1up	-6.47/5.95
SysFATJET_Medium_JET_Comb_Baseline_Kin__1down	5.83/-6.43
SysFT_EFF_Eigen_B_0_AntiKt2PV0TrackJets__1down	-3.49/2.97
SysFT_EFF_Eigen_B_1_AntiKt2PV0TrackJets__1down	-3.32/2.8
SysFT_EFF_Eigen_C_1_AntiKt2PV0TrackJets__1down	-2.97/2.45
SysFT_EFF_Eigen_B_0_AntiKt2PV0TrackJets__1up	2.39/-2.94
SysFT_EFF_Eigen_B_1_AntiKt2PV0TrackJets__1up	2.23/-2.78
SysFT_EFF_Eigen_C_1_AntiKt2PV0TrackJets__1up	1.9/-2.45
SysFATJET_Medium_JET_Comb_Tracking_Kin__1down	1.77/-2.38
SysFATJET_Medium_JET_Comb_Tracking_Kin__1up	-2.2/1.68
SysFT_EFF_extrapolation_AntiKt2PV0TrackJets__1up	-2.07/1.59
SysFATJET_JMR__1up	1.41/-1.97
SysFT_EFF_extrapolation_from_charm_AntiKt2PV0TrackJets__1up	-1.62/1.1
SysFT_EFF_extrapolation_AntiKt2PV0TrackJets__1down	0.963/-1.54
SysPRW_DATASF__1down	-1.47/0.94
SysJET_SR1_JET_GroupedNP_1__1down	0.764/-1.33
SysJET_SR1_JET_GroupedNP_1__1up	-1.2/0.678
SysFATJET_JER__1up	0.574/-1.16
SysFT_EFF_Eigen_Light_1_AntiKt2PV0TrackJets__1up	-1.12/0.576
SysFT_EFF_extrapolation_from_charm_AntiKt2PV0TrackJets__1down	0.553/-1.1
SysFATJET_Medium_JET_Comb_TotalStat_Kin__1down	-1.03/0.498
SysFT_EFF_Eigen_C_2_AntiKt2PV0TrackJets__1down	-0.983/0.458

SysMET_SoftTrk_Scale__1down	-0.971/0.439
SysJET_JER_SINGLE_NP__1up	-0.936/0.375
SysEL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR__1up	-0.935/0.405
SysFATJET_Medium_JET_Comb_Modelling_Kin__1up	-0.903/0.381
SysMUON_EFF_SYS__1up	-0.88/0.353
SysJET_SR1_JET_GroupedNP_3__1up	-0.871/0.346
SysJET_SR1_JET_GroupedNP_2__1up	-0.826/0.289
SysFT_EFF_Eigen_Light_2_AntiKt2PV0TrackJets__1down	-0.825/0.301
SysJET_SR1_JET_EtaIntercalibration_NonClosure__1up	-0.792/0.261
SysEL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR__1up	-0.78/0.25
SysFT_EFF_Eigen_C_3_AntiKt2PV0TrackJets__1up	-0.767/0.233
SysFT_EFF_Eigen_Light_3_AntiKt2PV0TrackJets__1up	-0.762/0.229
SysEL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR__1up	-0.71/0.179
SysFT_EFF_Eigen_B_2_AntiKt2PV0TrackJets__1up	-0.705/0.173
SysMUON_ID__1down	-0.698/0.166
SysFT_EFF_Eigen_B_2_AntiKt4EMTopoJets__1down	-0.662/0.13
SysFT_EFF_extrapolation_AntiKt4EMTopoJets__1up	-0.651/0.119
SysFT_EFF_Eigen_B_0_AntiKt4EMTopoJets__1down	-0.637/0.107
SysEL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR__1up	-0.611/0.0791
SysMUON_EFF_STAT__1up	-0.604/0.0725
SysFT_EFF_Eigen_B_1_AntiKt4EMTopoJets__1down	-0.593/0.0613
SysMUON_ID__1up	-0.588/0.0559
SysFT_EFF_Eigen_Light_2_AntiKt4EMTopoJets__1down	-0.582/0.0499
SysMET_SoftTrk_Scale__1up	0.0485/-0.582
SysMUON_ISO_SYS__1up	-0.576/0.0453
SysFT_EFF_Eigen_Light_0_AntiKt4EMTopoJets__1down	-0.572/0.0358
SysFT_EFF_Eigen_Light_1_AntiKt2PV0TrackJets__1down	0.0513/-0.571
SysFT_EFF_Eigen_Light_1_AntiKt4EMTopoJets__1up	-0.567/0.0359
SysMUON_MS__1up	-0.558/0.0258
SysFT_EFF_Eigen_C_1_AntiKt4EMTopoJets__1up	-0.556/0.024
SysFT_EFF_Eigen_Light_3_AntiKt4EMTopoJets__1down	-0.555/0.0236

SysFT_EFF_Eigen_Light_4_AntiKt2PV0TrackJets__1down	-0.551/0.0151
SysFT_EFF_Eigen_C_0_AntiKt4EMTopoJets__1up	-0.546/0.0135
SysMUON_ISO_STAT__1up	-0.545/0.0136
SysFT_EFF_Eigen_C_3_AntiKt4EMTopoJets__1up	-0.543/0.0112
SysFT_EFF_Eigen_Light_4_AntiKt4EMTopoJets__1down	-0.54/0.00864
SysFT_EFF_Eigen_C_2_AntiKt4EMTopoJets__1up	-0.539/0.00807
SysFT_EFF_extrapolation_from_charm_AntiKt4EMTopoJets__1down	-0.537/0.00495
SysLEPTON_d0Eff_Acc__1down	-0.532/0
SysLEPTON_d0Eff_Acc__1up	-0.532/0
SysMODEL_DIBOSONS_Norm_Inc_Acc__1down	-0.532/0
SysMODEL_DIBOSONS_Norm_Inc_Acc__1up	-0.532/0
SysMODEL_DIBOSONS_Norm_Inc_XS__1down	-0.532/0
SysMODEL_DIBOSONS_Norm_Inc_XS__1up	-0.532/0
SysMODEL_QCD_Norm_Inc_1down	-0.532/0
SysMODEL_QCD_Norm_Inc_1up	-0.532/0
SysMODEL_STOP_Norm_Inc_Acc__1down	-0.532/0
SysMODEL_STOP_Norm_Inc_Acc__1up	-0.532/0
SysMODEL_STOP_Norm_Inc_XS__1down	-0.532/0
SysMODEL_STOP_Norm_Inc_XS__1up	-0.532/0
SysMODEL_TTBar_Norm_Inc_Acc__1down	-0.532/0
SysMODEL_TTBar_Norm_Inc_Acc__1up	-0.532/0
SysMODEL_TTBar_Norm_Inc_XS__1down	-0.532/0
SysMODEL_TTBar_Norm_Inc_XS__1up	-0.532/0
SysMODEL_WJETS_Norm_Inc_Acc__1down	-0.532/0
SysMODEL_WJETS_Norm_Inc_Acc__1up	-0.532/0
SysMODEL_WJETS_Norm_Inc_XS__1down	-0.532/0
SysMODEL_WJETS_Norm_Inc_XS__1up	-0.532/0
SysMODEL_ZJETS_Norm_Inc_Acc__1down	-0.532/0
SysMODEL_ZJETS_Norm_Inc_Acc__1up	-0.532/0
SysMODEL_ZJETS_Norm_Inc_XS__1down	-0.532/0
SysMODEL_ZJETS_Norm_Inc_XS__1up	-0.532/0

SysMUON_SCALE__1down	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_Shower__1up	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_Shower__1down	-0.532/0
SysMODEL_WJETS_Norm_mbbcr_XS__1up	-0.532/0
SysMODEL_WJETS_Norm_mbbcr_XS__1down	-0.532/0
SysMODEL_ZJETS_Norm_SR_XS__1up	-0.532/0
SysMODEL_ZJETS_Norm_SR_XS__1down	-0.532/0
SysMODEL_ZJETS_Norm_SR_Acc__1up	-0.532/0
SysMODEL_ZJETS_Norm_SR_Acc__1down	-0.532/0
SysMODEL_WJETS_Shape_hhMass_SR_Scale__1up	-0.532/0
SysMODEL_WJETS_Shape_hhMass_SR_Scale__1down	-0.532/0
SysMODEL_DIBOSONS_Norm_mbbcr_XS__1up	-0.532/0
SysMODEL_DIBOSONS_Norm_mbbcr_XS__1down	-0.532/0
SysMODEL_DIBOSONS_Norm_mbbcr_Acc__1up	-0.532/0
SysMODEL_DIBOSONS_Norm_mbbcr_Acc__1down	-0.532/0
SysMODEL_QCD_Norm_mbbcr__1up	-0.532/0
SysMODEL_QCD_Norm_mbbcr__1down	-0.532/0
SysMODEL_STOP_Norm_mbbcr_Acc__1up	-0.532/0
SysMODEL_STOP_Norm_mbbcr_Acc__1down	-0.532/0
SysMODEL_STOP_Norm_mbbcr_XS__1up	-0.532/0
SysMODEL_STOP_Norm_mbbcr_XS__1down	-0.532/0
SysMODEL_TTBar_Norm_mbbcr_Acc__1up	-0.532/0
SysMODEL_TTBar_Norm_mbbcr_Acc__1down	-0.532/0
SysMODEL_TTBar_Norm_mbbcr_XS__1up	-0.532/0
SysMODEL_TTBar_Norm_mbbcr_XS__1down	-0.532/0
SysMODEL_WJETS_Norm_mbbcr_Acc__1up	-0.532/0
SysMODEL_WJETS_Norm_mbbcr_Acc__1down	-0.532/0
SysMODEL_ZJETS_Norm_mbbcr_Acc__1up	-0.532/0
SysMODEL_ZJETS_Norm_mbbcr_Acc__1down	-0.532/0
ZJETS_Norm_mbbcr_XS__1up	-0.532/0
ZJETS_Norm_mbbcr_XS__1down	-0.532/0

SysMODEL_DIBOSONS_Norm_SR_Acc__1up	-0.532/0
SysMODEL_DIBOSONS_Norm_SR_Acc__1down	-0.532/0
SysMODEL_DIBOSONS_Norm_SR_XS__1up	-0.532/0
SysMODEL_DIBOSONS_Norm_SR_XS__1down	-0.532/0
SysMODEL_QCD_Norm_SR__1up	-0.532/0
SysMODEL_QCD_Norm_SR__1down	-0.532/0
SysMODEL_QCD_Shape_hhMass_SR__1up	-0.532/0
SysMODEL_QCD_Shape_hhMass_SR__1down	-0.532/0
SysMODEL_STOP_Norm_SR_Acc__1up	-0.532/0
SysMODEL_STOP_Norm_SR_Acc__1down	-0.532/0
SysMODEL_STOP_Norm_SR_XS__1up	-0.532/0
SysMODEL_STOP_Norm_SR_XS__1down	-0.532/0
SysMODEL_TTBar_Norm_SR_Acc__1up	-0.532/0
SysMODEL_TTBar_Norm_SR_Acc__1down	-0.532/0
SysMODEL_TTBar_Norm_SR_XS__1up	-0.532/0
SysMODEL_TTBar_Norm_SR_XS__1down	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_Gen__1up	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_Gen__1down	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_RadLo__1up	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_RadLo__1down	-0.532/0
SysMODEL_WJETS_Norm_SR_Acc__1up	-0.532/0
SysMODEL_WJETS_Norm_SR_Acc__1down	-0.532/0
SysMODEL_WJETS_Norm_SR_XS__1up	-0.532/0
SysMODEL_WJETS_Norm_SR_XS__1down	-0.532/0
SysFT_EFF_extrapolation_from_charm_AntiKt4EMTopoJets__1up	-0.527/- 0.00495
SysJET_JvtEfficiency__1up	-0.526/- 0.00613
SysJET_SR1_JET_GroupedNP_2__1down	-0.0414/-0.525
SysFT_EFF_Eigen_C_2_AntiKt4EMTopoJets__1down	-0.524/- 0.00814

SysFT_EFF_Eigen_Light_4_AntiKt4EMTopoJets__1up	-0.523/-0.00865
SysFT_EFF_Eigen_C_3_AntiKt4EMTopoJets__1down	-0.52/-0.0113
SysMUON_ISO_STAT__1down	-0.518/-0.0136
SysFT_EFF_Eigen_C_0_AntiKt4EMTopoJets__1down	-0.518/-0.0125
SysMUON_SAGITTA_RESBIAS__1up	-0.0125/-0.516
SysFT_EFF_Eigen_Light_4_AntiKt2PV0TrackJets__1up	-0.513/-0.015
SysFT_EFF_Eigen_Light_3_AntiKt4EMTopoJets__1up	-0.508/-0.0235
SysFT_EFF_Eigen_C_1_AntiKt4EMTopoJets__1down	-0.508/-0.0239
SysMET_SoftTrk_ResoPara__1up	-0.505/-0.0311
SysFT_EFF_Eigen_Light_0_AntiKt4EMTopoJets__1up	-0.498/-0.0297
SysFT_EFF_Eigen_Light_1_AntiKt4EMTopoJets__1down	-0.496/-0.0358
SysMUON_ISO_SYS__1down	-0.487/-0.0453
SysFATJET_Medium_JET_Comb_TotalStat_Kin__1up	-0.048/-0.484
SysFT_EFF_Eigen_Light_2_AntiKt4EMTopoJets__1up	-0.482/-0.0499
SysJET_JvtEfficiency__1down	-0.478/-0.0538
SysJET_SR1_JET_EtaIntercalibration_NonClosure__1down	-0.474/-0.0887
SysEG_SCALE_ALL__1down	-0.471/-0.0611
SysMUON_MS__1down	-0.47/-0.0619
SysFT_EFF_Eigen_B_1_AntiKt4EMTopoJets__1up	-0.468/-0.0644
SysMUON_SCALE__1up	-0.462/-0.0696
SysMUON_EFF_STAT__1down	-0.46/-0.0725
SysFT_EFF_Eigen_C_2_AntiKt2PV0TrackJets__1up	-0.0799/-0.459
SysEL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR__1down	-0.453/-0.0791
SysMUON_SAGITTA_RHO__1down	-0.429/-0.106
SysMUON_SAGITTA_RHO__1up	-0.429/-0.106
SysFT_EFF_Eigen_B_0_AntiKt4EMTopoJets__1up	-0.426/-0.107
SysFT_EFF_extrapolation_AntiKt4EMTopoJets__1down	-0.413/-0.119
SysEL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR__1down	-0.129/-0.405
SysMET_SoftTrk_ResoPerp__1up	-0.403/-0.13
SysFT_EFF_Eigen_B_2_AntiKt4EMTopoJets__1up	-0.402/-0.13

SysEG_SCALE_ALL__1up	-0.132/-0.4
SysFATJET_Medium_JET_Comb_Modelling_Kin__1down	-0.181/-0.384
SysJET_SR1_JET_GroupedNP_3__1down	-0.369/-0.198
SysFT_EFF_Eigen_B_2_AntiKt2PV0TrackJets__1down	-0.358/-0.173
SysEL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR__1down	-0.353/-0.179
SysMUON_EFF_SYS__1down	-0.185/-0.352
SysMUON_SAGITTA_RESBIAS__1down	-0.335/-0.2
SysPRW_DATASF__1up	-0.195/-0.322
SysFT_EFF_Eigen_Light_3_AntiKt2PV0TrackJets__1down	-0.301/-0.229
SysFT_EFF_Eigen_Light_2_AntiKt2PV0TrackJets__1up	-0.24/-0.3
SysFT_EFF_Eigen_C_3_AntiKt2PV0TrackJets__1down	-0.296/-0.233
SysEL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR__1down	-0.284/-0.25
Total Up	27.2
Total Do	27.9

TABLE I.1. List of systematic uncertainties for the fully boosted analysis

APPENDIX J

MBB CONTROL REGION DISTRIBUTIONS

Kinematic distributions for the fully-boosted analysis in the mBB control region. Where $100\text{GeV} > m_{bb}$ or $m_{bb} > 140\text{GeV}$.

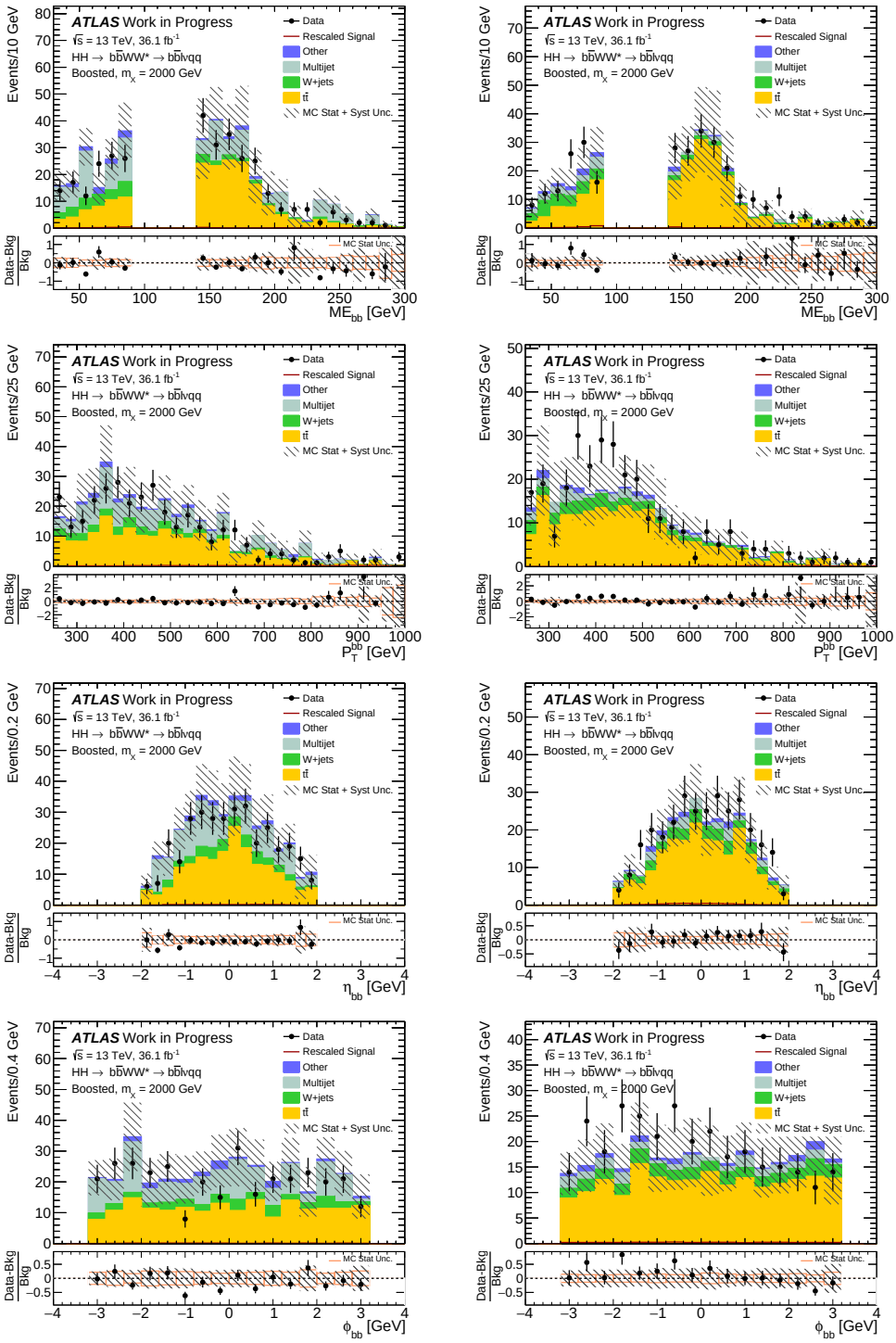

FIGURE J.1. m_{bb}, $p_T^b b$, η_{bb}, and ϕ_{bb} (from top to bottom) distribution in the mBB control region for the electron (left) and muon (right) channels.

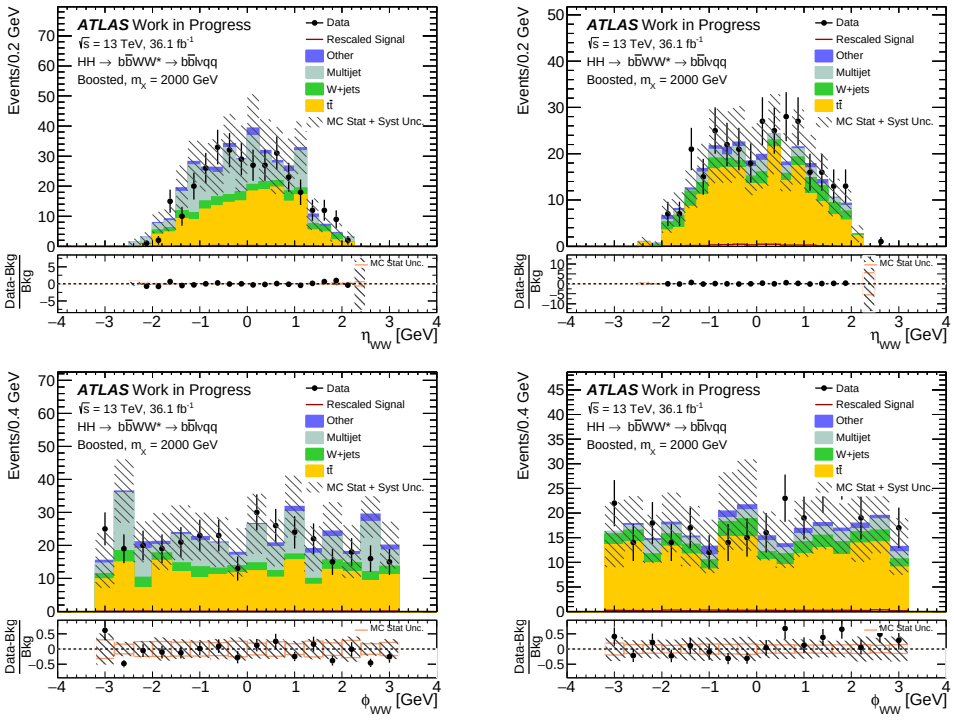

FIGURE J.2. η_{WW}, and ϕ_{WW} (from top to bottom) distribution in the mBB control region for the electron (left) and muon (right) channels.

APPENDIX K

SR DISTRIBUTIONS

Kinematic distributions for the fully-boosted analysis in the signal region. Where $100\text{GeV} < m_{bb} < 140\text{GeV}$.

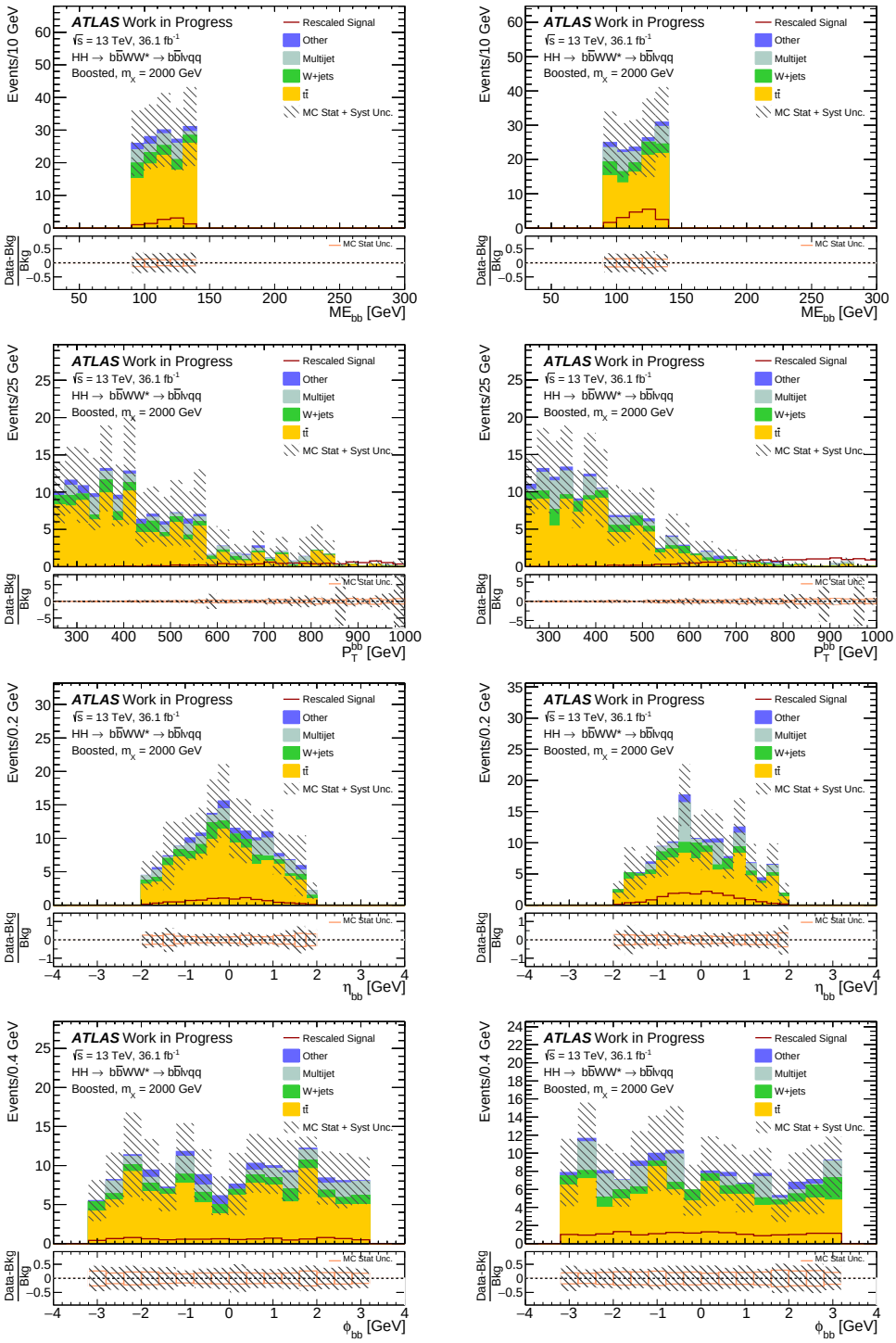

FIGURE K.1. m_{bb}, $p_T^b b$, η_{bb}, and ϕ_{bb} (from top to bottom) distribution in the signal region for the electron (left) and muon (right) channels.

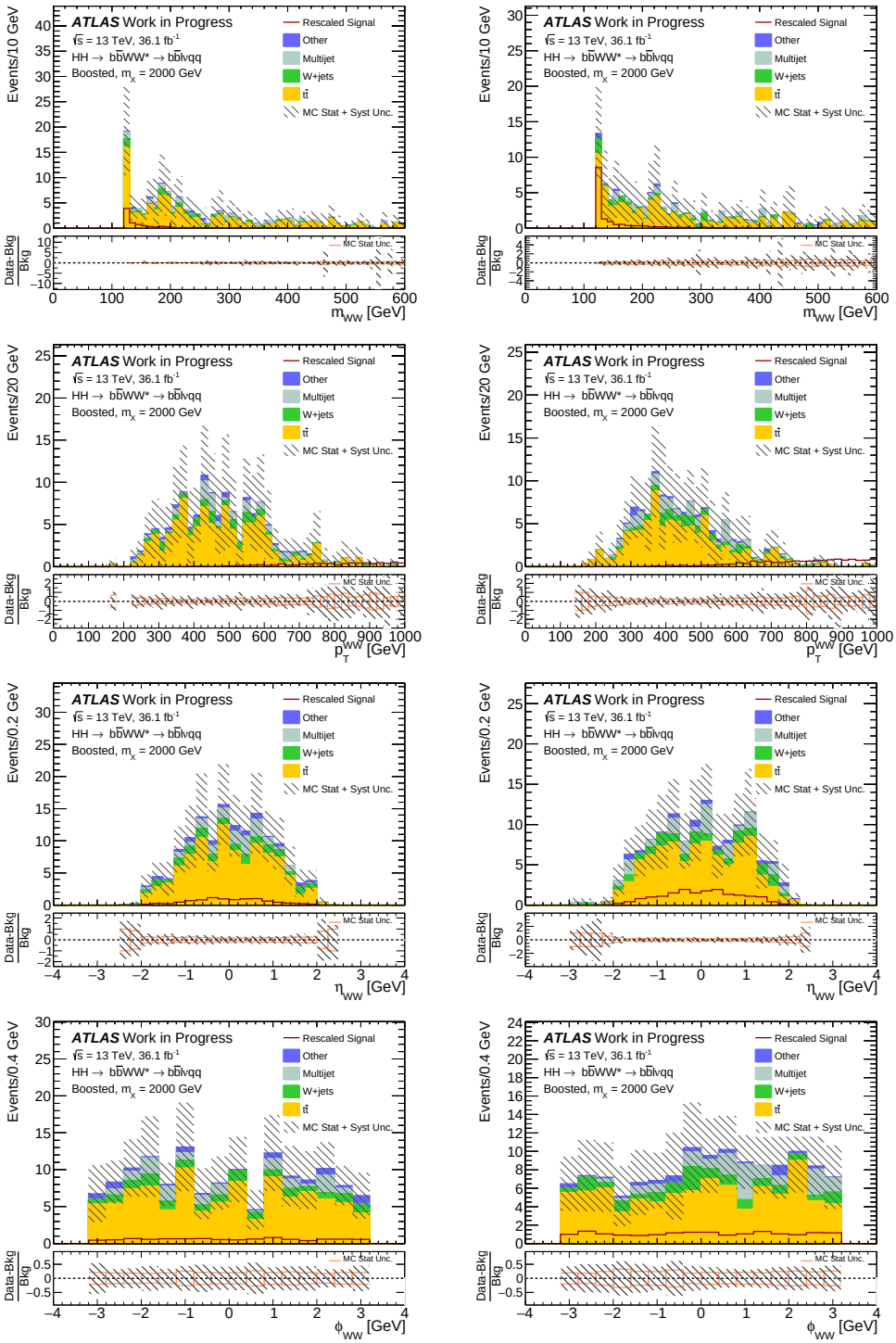

FIGURE K.2. m_{WW}, $p_T^W W$, η_{WW}, and ϕ_{WW} (from top to bottom) distribution in the signal region for the electron (left) and muon (right) channels.

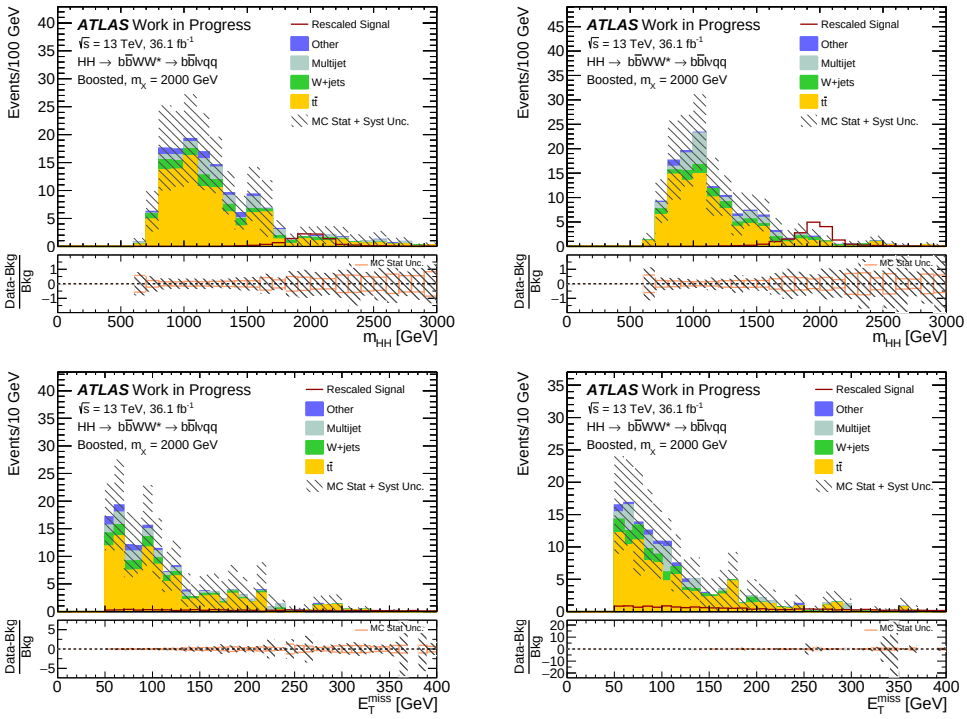

FIGURE K.3. m_{HH} (top) and $\not{E}_T$(bottom) distribution in the signal region for the electron (left) and muon (right) channels.

[1] The ATLAS collaboration. "Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2019.4 (Apr. 2019), p. 92. ISSN: 1029-8479. DOI: 10.1007/JHEP04(2019)092. URL: https://doi.org/10.1007/JHEP04(2019)092.

[2] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[3] Ian Snyder. "SEARCH FOR A SCALAR PARTNER OF THE TOP QUARK IN THE JETS+Missing ET T FINAL STATE IN PROTON-PROTON COLLISIONS AT Sqrt(s) =13 TEV WITH THE ATLAS DETECTOR". PhD thesis. University of Oregon, 2018.

[4] John Ellis. "Higgs Physics". In: arXiv:1312.5672. KCL-PH-TH-2013-49. LCTS-2013-36. CERN-PH-TH-2013-315 (Dec. 2013). 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report, 117–168. 52 p. URL: http://cds.cern.ch/record/1638469.

[5] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[6] P.W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Physics Letters* 12.2 (1964), pp. 132–133. ISSN: 0031-9163. DOI: https://doi.org/10.1016/0031-9163(64)91136-9. URL: http://www.sciencedirect.com/science/article/pii/0031916364911369.

[7] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[9] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[10] Georges Aad et al. "Evidence for the spin-0 nature of the Higgs boson using ATLAS data". In: *Phys. Lett.* B726 (2013), pp. 120–144. DOI: 10.1016/j.physletb.2013.08.026. arXiv: 1307.1432 [hep-ex].

[11] Morad Aaboud et al. "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector". In: *Phys. Lett.* B786 (2018), pp. 59–86. DOI: 10.1016/j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex].

[12] Morad Aaboud et al. "Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Submitted to: Phys. Rev.* (2018). arXiv: 1811.08856 [hep-ex].

[13] R. Belusevic and G. Jikia. "Higgs selfcoupling in gamma gamma collisions". In: *Phys. Rev.* D70 (2004), p. 073017. DOI: 10.1103/PhysRevD.70.073017. arXiv: hep-ph/0403303 [hep-ph].

[14] Albert M Sirunyan et al. "Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Submitted to: Phys. Rev. Lett.* (2018). arXiv: 1811.09689 [hep-ex].

[15] Aleksandr Azatov et al. "Effective field theory analysis of double Higgs production via gluon fusion". In: *Physical Review D* 92 (Feb. 2015). DOI: 10.1103/PhysRevD.92.035001.

[16] S. Dawson and M. Sullivan. "Enhanced di-Higgs boson production in the complex Higgs singlet model". In: *Phys. Rev. D* 97 (1 Jan. 2018), p. 015022. DOI: 10.1103/PhysRevD.97.015022. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.015022.

[17] J. de Blas et al. "Electroweak precision observables and Higgs-boson signal strengths in the Standard Model and beyond: present and future". In: *Journal of High Energy Physics* 2016.12 (Dec. 2016), p. 135. ISSN: 1029-8479. DOI: 10.1007/JHEP12(2016)135. URL: https://doi.org/10.1007/JHEP12(2016)135.

[18] Benjamin W. Lee, C. Quigg, and H. B. Thacker. "Weak interactions at very high energies: The role of the Higgs-boson mass". In: *Phys. Rev. D* 16 (5 Sept. 1977), pp. 1519–1531. DOI: 10.1103/PhysRevD.16.1519. URL: https://link.aps.org/doi/10.1103/PhysRevD.16.1519.

[19] Ian M. Lewis and Matthew Sullivan. "Benchmarks for Double Higgs Production in the Singlet Extended Standard Model at the LHC". In: *Phys. Rev.* D96.3 (2017), p. 035037. DOI: 10.1103/PhysRevD.96.035037. arXiv: 1701.08774 [hep-ph].

[20] "Linear accelerator 2". In: (Sept. 2012). URL: https://cds.cern.ch/record/1997427.

[21] Cinzia De Melis. "The CERN accelerator complex. Complexe des accélérateurs du CERN". In: (Jan. 2016). General Photo. URL: https://cds.cern.ch/record/2119882.

[22] K.H. Reich. "The CERN Proton Synchrotron Booster". In: (Jan. 1969). Round Table On Booster Injectors. URL: http://accelconf.web.cern.ch/Accelconf/p69/PDF/PAC1969_0959.PDF.

[23] Giorgio Brianti. "The CERN synchrotrons". In: (1997). URL: https://cds.cern.ch/record/340514.

[24] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001.

[25] Xabier Cid Vidal Ramon Cid Manzano. *Lorentz Force Taking a closer look at LHC*. URL: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.lorentz_force.

[26] Morad Aaboud et al. "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C76.12 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[27] Genessis Perez. "Unitarization Models For Vector Boson Scattering at the LHC". PhD thesis. Jan. 2018. DOI: 10.5445/IR/1000082199.

[28] The ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003.

[29] M Capeans et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. Sept. 2010. URL: https://cds.cern.ch/record/1291633.

[30] *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331063.

[31] The ATLAS Collaboration. "Approved Plots of the Tracking Combined Performance Group". URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-007/.

[32] Boris Dolgoshein. "Transition radiation detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 326.3 (1993), pp. 434–469. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(93)90846-A. URL: http://www.sciencedirect.com/science/article/pii/016890029390846A.

[33] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331061.

[34] *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331062.

[35] Georges Aad et al. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data". In: *Eur. Phys. J.* C74.10 (2014), p. 3071. DOI: 10.1140/epjc/s10052-014-3071-4. arXiv: 1407.5063 [hep-ex].

[36] Ana Maria Henriques Correia. *The ATLAS Tile Calorimeter*. Tech. rep. ATL-TILECAL-PROC-2015-002. Geneva: CERN, Mar. 2015. URL: https://cds.cern.ch/record/2004868.

[37] *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331068.

[38] Jeremy R Love. "A Search for Technicolor at The Large Hadron Collider". PhD thesis. Boston U., 2011-12-12.

[39] *ATLAS magnet system: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/338080.

[40] ATLAS Outreach. "ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC". 2010. URL: https://cds.cern.ch/record/1457044.

[41] Aranzazu Ruiz-Martinez and ATLAS Collaboration. *The Run-2 ATLAS Trigger System*. Tech. rep. ATL-DAQ-PROC-2016-003. Geneva: CERN, Feb. 2016. URL: https://cds.cern.ch/record/2133909.

[42] Martin zur Nedden. *The Run-2 ATLAS Trigger System: Design, Performance and Plan*. Tech. rep. ATL-DAQ-PROC-2016-039. Geneva: CERN, Dec. 2016. URL: https://cds.cern.ch/record/2238679.

[43] Chaowaroj Wanotayaroj and Jim Brau. "Search for a Scalar Partner of the Top Quark in the Jets+MET Final State with the ATLAS detector". Presented 25 Oct 2016. Nov. 2016. URL: https://cds.cern.ch/record/2242196.

[44] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* (2003).

[45] Joao Pequenao. "Event Cross Section in a computer generated image of the ATLAS detector." Mar. 2008. URL: https://cds.cern.ch/record/1096081.

[46] Grigore Tarna. *Electron identification with the ATLAS detector*. Tech. rep. ATL-PHYS-PROC-2017-173. Geneva: CERN, Sept. 2017. URL: https://cds.cern.ch/record/2286383.

[47] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*. ATLAS-CONF-2016-024. June 2016. URL: https://cds.cern.ch/record/2157687.

[48] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data". In: *Eur. Phys. J.* C74.10 (2014), p. 3071. DOI: 10.1140/epjc/s10052-014-3071-4. arXiv: 1407.5063.

[49] ATLAS Collaboration. *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at* $\sqrt{s} = 13$ TeV. Aug. 2016. URL: https://cds.cern.ch/record/2203514.

[50] Georges Aad et al. "Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur.Phys.J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y.

[51] Georges Aad et al. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J.* C77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[52] Bora Isildak. "Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s} = 7$ tev". PhD thesis. Bogazici U., 2011. arXiv: 1308.6064 [hep-ex].

[53] M. Cacciari, C. P. Salam and G. Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063.

[54] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: (2016). arXiv: 1603.02934 [hep-ex].

[55] David Krohn, Jesse Thaler, and Lian-Tao Wang. "Jet Trimming". In: *JHEP* 1002 (2010), p. 084. DOI: 10.1007/JHEP02(2010)084. arXiv: 0912.1342 [hep-ph].

[56] *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*. Tech. rep. ATLAS-CONF-2016-035. Geneva: CERN, July 2016. URL: https://cds.cern.ch/record/2200211.

[57] *Expected Performance of Boosted Higgs ($\rightarrow b\bar{b}$) Boson Identification with the ATLAS Detector at $\sqrt{s}$ = 13 TeV*. Tech. rep. ATL-PHYS-PUB-2015-035. Geneva: CERN, Aug. 2015. URL: https://cds.cern.ch/record/2042155.

[58] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The Catchment Area of Jets". In: *JHEP* 04 (2008), p. 005. DOI: 10.1088/1126-6708/2008/04/005. arXiv: 0802.1188 [hep-ph].

[59] ATLAS Collaboration. " Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at sqrt(s) = 13 TeV with the ATLAS detector ". In: (2017). arXiv: 1703.09665 [hep-ex].

[60] ATLAS Collaboration. "Performance of pile-up mitigation techniques for jets in *pp* collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector". In: *Eur. Phys. J.* C76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].

[61] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[62] *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*. Tech. rep. ATL-PHYS-PUB-2016-012. Geneva: CERN, June 2016. URL: https://cds.cern.ch/record/2160731.

[63] Per Ola Hansson Adrian. *The ATLAS b-Jet Trigger*. Tech. rep. arXiv:1111.4190. Comments: 4 pages, 6 figures, conference proceedings for PIC2011. Geneva: CERN, Nov. 2011. URL: https://cds.cern.ch/record/1397942.

[64] B. Abbott et al. "Production and Integration of the ATLAS Insertable B-Layer". In: *JINST* 13.05 (2018), T05008. DOI: 10.1088/1748-0221/13/05/T05008. arXiv: 1803.00844 [physics.ins-det].

[65] *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV*. Tech. rep. ATL-PHYS-PUB-2015-027. Geneva: CERN, July 2015. URL: https://cds.cern.ch/record/2037904.

[66] Morad Aaboud et al. "Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (2018). arXiv: 1811.04671 [hep-ex].

[67] AIDAN RANDLE-CONDE. "What next for the Higgs?" URL: https://www.quantumdiaries.org/2012/06/09/what-next-for-the-higgs/.

[68] Biagio Di Micco et al. *Search for Higgs boson pair production in the bbWW final state at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATL-COM-PHYS-2016-486. Geneva: CERN, May 2016. URL: https://cds.cern.ch/record/2151893.

[69] Cigdem Issever et al. *Boosted Analysis: Search for Higgs boson pair production in the bbWW final state at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATL-COM-PHYS-2017-1088. Geneva: CERN, July 2017. URL: https://cds.cern.ch/record/2276099.

[70] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[71] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 0605 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[72] Peter Zeiler Skands. "Tuning Monte Carlo Generators: The Perugia Tunes". In: *Phys.Rev.* D82 (2010), p. 074018. DOI: 10.1103/PhysRevD.82.074018. arXiv: 1005.3457 [hep-ph].

[73] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 0207 (2002), p. 012. arXiv: hep-ph/0201195 [hep-ph].

[74] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth.* A462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.

[75] Michal Czakon and Alexander Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: *Compt/ Phys. Commun.* (2014), p. 2930. arXiv: 1112.5675 [hep-ph].

[76] ATLAS Collaboration. *Comparison of Monte Carlo generator predictions for gap fraction and jet multiplicity observables in top-antitop events*. Tech. rep. ATL-PHYS-PUB-2014-005. Geneva: CERN, May 2014. URL: https://cds.cern.ch/record/1703034.

[77] Nikolaos Kidonakis. "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production". In: *Phys. Rev.* D83 (2011), p. 091503. DOI: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792 [hep-ph].

[78] Nikolaos Kidonakis. "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-". In: *Phys. Rev.* D82 (2010), p. 054018. DOI: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451 [hep-ph].

[79] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 0902 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[80] Hung-Liang Lai et al. "New parton distributions for collider physics". In: *Phys.Rev.* D82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

[81] C. Anastasiou et al. "High precision determination of the gluon fusion Higgs boson cross-section at the LHC". In: *JHEP* 05 (2016), p. 058. arXiv: 1602.00695 [hep-ph].

[82] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[83] R.Frederix et al. "Higgs pair production at the LHC with NLO and parton-shower effects". In: *Phys. Lett.* B732 (2014), p. 142. DOI: 10.1016/j.physletb.2014.03.026.

[84] G. Degrassi, P. P. Giardino, R. Groeber. "On the two-loop virtual QCD corrections to Higgs boson pair production in the Standard Model". In: *Eur.Phys.J.* C76 (2016), p. 411. DOI: 10.1140/epjc/s10052-016-4256-9.

[85] S. Borowka et al. "Full top quark mass dependence in Higgs boson pair production at NLO". In: *JHEP* 10 (2016), p. 107. DOI: 10.1007/JHEP10(2016)107. arXiv: 1608.04798 [hep-ph].

[86] S. Borowka et al. "Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence". In: *Phys. Rev. Lett.* 117.1 (2016). [Erratum: Phys. Rev. Lett.117,no.7,079901(2016)], p. 012001. DOI: 10.1103/PhysRevLett.117.079901, 10.1103/PhysRevLett.117.012001. arXiv: 1604.06447 [hep-ph].

[87] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur.Phys.J.* C70 (2010), pp. 823–874. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[88] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl.Instrum.Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[89] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*. Tech. rep. ATL-PHYS-PUB-2015-015. Geneva: CERN, July 2015. URL: https://cds.cern.ch/record/2037613.

[90] ATLAS Collaboration. "Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C75 (2015), p. 17. DOI: 10.1140/epjc/s10052-014-3190-y. arXiv: 1406.0076 [hep-ex].

[91] ATLAS Collaboration. *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*. Tech. rep. ATLAS-CONF-2014-004. Geneva: CERN, Feb. 2014. URL: http://cds.cern.ch/record/1664335.

[92] A.L. Read. "Presentation of search results: the CL_s technique". In: *J. Phys.* G28 (2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313.

[93] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1603.00385 [hep-ph].