THE UNIVERSITY *of* EDINBURGH

The search for Higgs boson pair production with the ATLAS detector at the Large Hadron Collider

Alan James Taylor

Doctor of Philosophy
The University of Edinburgh
August 2019

Lay summary

The Large Hadron Collider (LHC) at CERN (European Organisation for Nuclear Research) is the largest particle accelerator ever built. In a 27 km tunnel, the LHC collides protons together at high energies in order to recreate the conditions of the big bang. These collisions are picked apart in the hope it will be possible to gain new insights into the laws that govern fundamental particles.

In a high energy collision at the LHC, many of the particles produced do not make up the everyday matter that surrounds us. For example, the heaviest known particle is the top quark and only lives for approximately 10^{-25} seconds before decaying into a cascade of lighter, more stable particles. In order to understand the particles being produced, it is necessary to build large detectors that surround the collisions. The detectors work like large, 3-dimensional, 100 megapixel digital cameras taking up to 40 million images per second.

The Standard Model (SM) of particle physics represents our best understanding of nature on the subatomic scale. With only a few assumptions, the SM is able to describe almost all measurements performed in high energy physics. One of the biggest achievements at the LHC thus far has been the discovery of the Higgs boson by the ATLAS and CMS experiments which was the missing piece for many years. Despite the SM now being complete, there are still many gaps and mysteries which cannot be explained. For example there is no particle in the SM that is a candidate for the dark matter that exists in the universe.

Many measurements of the properties of the Higgs boson are now entering a precision era for the first time. For example, the mass of the Higgs boson has now been measured to a precision of approximately 0.2% and once the mass of the Higgs boson is known, all properties of the particle can be predicted. One prediction of the SM is that a single virtual Higgs boson can produce two Higgs bosons, a property known as the *Higgs self-coupling*. This property has not yet

been observed by the LHC experiments as the rate as which it occurs is very low.

The challenging task of observing Higgs boson pair production will require a data set significantly larger than the current data set. A number of upgrades are planned to both the LHC and the ATLAS detector in order to collect this data set. A simulation study was conducted to study the prospects for observing Higgs boson pair production in association with top quarks. In this study, realistic backgrounds were accounted for as well as the increased number of overlapping collisions that will be present during the upgraded LHC program. This method of studying Higgs boson pair production is found to be particularly challenging.

Although it is not possible to observe Higgs boson pair production as predicted in the SM with the current data set, it is important to study since new physics beyond the Standard Model may enhance the production rate. A search is performed with one Higgs boson decaying to a pair of bottom quarks and the other Higgs boson decaying to a pair of photons. The results obtained are in good agreement with the SM predictions and a constraint is placed on the Higgs self-coupling with respect to the value predicted in the SM.

Abstract

Since the observation of a new particle consistent with the Standard Model Higgs boson by the ATLAS and CMS collaborations, a number of important measurements have been made to understand this new particle. For example, the mass has been measured to a precision of approximately 0.2% and it has been confirmed to have spin-0 and positive parity as predicted in the SM. Recently, the Higgs boson has also been observed to couple to top and bottom quarks.

The Higgs mechanism also predicts that the Higgs field can interact with itself: a single virtual Higgs boson can produce two Higgs bosons. The central topic of this thesis is understanding this self-coupling property of the Higgs boson. Unfortunately the rate predicted for Higgs boson pair production is very low and the backgrounds are particularly challenging. This makes it one of the most difficult properties of the Higgs boson to measure at the Large Hadron Collider (LHC).

In this thesis three contributions are presented: performance studies for the upgraded ATLAS inner detector, the prospects for observing Higgs boson pair production in association with top quarks at the High Luminosity LHC (HL-LHC) and the search for Higgs boson pair production in the final state of two bottom quarks and two photons with the Run 2 data set.

In the near future, the LHC will be upgraded so it can deliver higher instantaneous luminosities. In order to maintain the current detector performance in these conditions, the ATLAS detector must also be upgraded. The silicon cluster sizes, channel occupancies and cluster densities are studied in two proposed layout designs for the future ATLAS inner detector, the ITk. The differences between the two layouts are compared and these studies were used as one of many inputs into the layout design decision which was taken at the end of 2017.

A number of phenomenological studies have shown that Higgs boson pair (HH)

production in association with top quarks is a promising channel to study the Higgs boson self-coupling. This channel is examined with a parameterised detector response and realistic backgrounds for the first time in this thesis. It is found that the sensitivity to this process is much smaller than that predicted in the phenomenological studies. The origin of this is due to a significant underestimation of the backgrounds in the phenomenological studies.

It is possible that new physics could enhance the rate of HH production, for example modifications to the Higgs boson self-coupling can increase the rate of HH production. A search is performed for HH production in the $b\bar{b}\gamma\gamma$ final state with the Run 2 data set collected by the ATLAS detector. A novel 2D fit is implemented and is found to improve the sensitivity to SM HH production by approximately 10% compared to the 1D fit which is typically used in $H \to \gamma\gamma$ physics analyses. No significant deviations from the SM are observed in this analysis and 95% CL upper limits are established. The observed (expected) limit for the SM HH cross section is 10.1 (7.0) times the SM value while the ratio of the Higgs boson self-coupling to its SM prediction, κ_λ is observed (expected) to be constrained at 95% CL to $-4.4 < \kappa_\lambda < 9.5$ ($-3.3 < \kappa_\lambda < 8.9$).

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Alan James Taylor, August 2019*)

Acknowledgements

I would like to thank many people for everything they have done for me over the last five years.

Firstly, I would like to thank Victoria Martin who took me on as a masters student on the CLIC experiment back in 2014. I think I have come along way from my first trip to CERN five years ago, where I could just about use a terminal and my coding skills and physics knowledge were minimal. All students have to make a slow start somewhere and with this a supervisor must believe in them. I am extremely grateful for that.

All members of the Edinburgh PPE Group. In particular: Tim Bristow, Sebastian Olivares and Yanyan Gao. Thanks to Tim and Sebastian, for the laughs in the early years, the squash battles and for the spare beds in Johannesburg and Santiago. Special thanks to Yanyan Gao for convincing me the $t\bar{t}HH$ analysis was a worthwhile study and for accompanying me on a trip to Durham for advice from Frank Krauss to resolve the $t\bar{t}HH$ MC troubles. Big thanks to Frank as well of course for this.

To my colleagues in the $H \to \gamma\gamma$ and $HH \to b\bar{b}\gamma\gamma$ groups, I am grateful for all the advice and for trusting me with many important responsibilities.

The CERN Visits service who have given me the opportunity to show people around many CERN sites. In particular, thanks to Marc for his infectious energy as Visits Traffic Controller. I hope we meet again.

To my family and high school friends who I have spent little time with the past years. I can only apologise for the time I have stolen and put into physics instead. To my Dad and Mum for the advice and support throughout the years and to my Gran for inspiring me - Thanks.

Finally, to Leo who has put up with me for almost two and a half years now. I am extremely happy we met and for all the happy memories we have together. Exploring many different places and for the support you have given me since we met and in particular the last few months. I look forward to a happy and bright future together.

Preface

The work presented in this thesis has been performed within the ATLAS collaboration. The work performed in the ATLAS collaboration is only possible due to the work of thousands of other people, both past and present. Every physics analysis relies on the calibrations, techniques and performance studies created by many others. The direct contributions presented by the author in this thesis are outlined below:

Chapter 3

All ITk performance plots shown in this chapter were created by the author unless otherwise stated.

Chapter 4

The author was the sole analyser for this analysis. The author was responsible for producing the Monte Carlo samples, writing the analysis framework, optimising the analysis and performing the statistical analysis.

Chapter 5

The author made important contributions to the $H \to \gamma\gamma$ framework. The author was responsible for the production and software development of $H \to \gamma\gamma$ and $H \to Z\gamma$ derivations. The author was also responsible for jets, flavour tagging and the $HH \to b\bar{b}\gamma\gamma$ selection tool within the framework.

The author was responsible for signal and background modelling, optimisation and performing the statistical analysis.

Contents

Chapter 1

The Standard Model of Particle Physics

The Standard Model (SM) of particle physics combines three of the four known fundamental interactions in nature into a comprehensive quantum field theory. The electromagnetic and weak interactions are unified into the electroweak force and combined with the theory of the strong interactions, Quantum Chromodynamics.

The Standard Model is built from a number of symmetry principles and from this, the SM is able to accomodate all experimental facts and precision measurements performed in high energy physics. However, for a long time, the SM remained incomplete. The mechanism for electroweak symmetry breaking was hypothesized, but the accompanying Higgs boson particle was still missing in experiments.

In 2012, the ATLAS and CMS collaborations announced the discovery of a new particle consistent with the Higgs boson and, hence, completed the Standard Model. A brief overview of the Standard Model will be presented in this chapter, with a particular emphasis on the Higgs mechanism and the Higgs boson. Phenomenology at proton-proton colliders is also presented in this chapter with details of the latest experimental results in Higgs physics.

1.1 Electromagnetic interactions

Matter is made up of spin-1/2 particles called fermions. The wave function associated to a fermion, $\psi(x)$ is a four component spinor where each component is a function of space-time x. The equation of motion for a free fermion is described by the Dirac equation:

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi(x) = 0 \tag{1.1}$$

where γ^{μ} are the Dirac matrices and m is the mass of the fermion particle. The corresponding Lagrangian for a free fermion is given by:

$$\mathcal{L} = i\bar{\psi}(x)\gamma^{\mu}\partial_{\mu}\psi(x) - m\,\bar{\psi}(x)\psi(x) \tag{1.2}$$

where $\bar{\psi}$ is the adjoint spinor, defined as $\bar{\psi} = \psi^{\dagger}\gamma^{0}$. Interactions between particles can be introduced by invoking the principle of gauge invariance. This principle states that under the the local phase transformations of the fields, the Lagrangian should not be modified by more than a total derivative and the equation of motion should remain unchanged. If a local phase transformation of $\psi \rightarrow \psi^{'} = e^{i\theta(x)}\psi(x)$ is considered, where $\theta(x)$ is an arbitrary function, the above condition is not fulfilled since:

$$\begin{aligned} i\bar{\psi}(x)\gamma^{\mu}\partial_{\mu}\psi(x) &\rightarrow i\bar{\psi}(x)e^{-i\theta(x)}\gamma^{\mu}\partial_{\mu}[e^{i\theta(x)}\psi(x)] \\ &= i\bar{\psi}(x)\gamma^{\mu}\partial_{\mu}\psi(x) - \bar{\psi}(x)\gamma^{\mu}\psi(x)[\partial_{\mu}\theta(x)] \end{aligned} \tag{1.3}$$

The $\partial_{\mu}\theta(x)$ term violates the invariance. This can be resolved by adding the spin-1 field A_{μ} that transforms like $A_{\mu} \rightarrow A_{\mu}^{'} = A_{\mu} - \frac{1}{e}\partial_{\mu}\theta(x)$ where e is a constant under the same U(1) symmetry. This transformation cancels the $\partial_{\mu}\theta(x)$ term which previously violated the invariance. The ordinary derivative ∂_{μ} can now be

replaced with the covariant derivative D_μ defined as

$$D_\mu(x)\psi(x) \equiv [\partial_\mu + ieA_\mu(x)]\psi(x) \tag{1.4}$$

The resulting Lagrangian is then given by:

$$\begin{aligned}\mathcal{L} &= i\bar{\psi}(x)\gamma^\mu D_\mu\psi(x) - m\,\bar{\psi}(x)\psi(x) \\ &= i\bar{\psi}(x)\gamma^\mu\partial_\mu\psi(x) - e\,\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu - m\,\bar{\psi}(x)\psi(x)\end{aligned} \tag{1.5}$$

The term $-e\,\bar{\psi}(x)\gamma^\mu\psi(x)\,A_\mu$ has emerged due to demanding gauge invariance. This can be interpreted as a fermion with charge e interacting with an electromagnetic potential A_μ. The principle of gauge invariance gives a recipe for generating interaction terms between forces and particles.

The dynamics of A_μ can be introduced by adding a kinetic term $-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ into the Lagrangian, where $F_{\mu\nu}$ is the electromagnetic field strength tensor given by $F_{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. The photon is massless and if a mass term $\frac{1}{2}m_\gamma^2 A_\mu A^\mu$ is added to the Lagrangian, gauge invariance would be broken again.

1.2 Strong interactions

The theory of the strong interactions, Quantum Chromodynamics predicts interactions between quarks and gluons. Quarks make up bound states like mesons ($q\bar{q}$) and baryons (qqq). Quarks carry electric charge so interact through the electromagnetic interaction while also carrying an additional quantum number called colour charge. The colours are conventionally taken as red, green and blue. The Lagrangian for a free quark $q(x)$ is:

$$\mathcal{L} = \sum_j \bar{q}_j(i\gamma^\mu\partial_\mu - m_j)q_j \tag{1.6}$$

where j is the sum over all quark flavours. Similar to before, quark interactions are generated by invoking the principle of gauge invariance, this time for an SU(3)

symmetry. The QCD Lagrangian is then given by:

$$\mathcal{L} = \sum_j \bar{q}_j (i\gamma^\mu \partial_\mu - m_j) q_j - g_s G^\mu_a \sum_j \bar{q}^\alpha_j \gamma_\mu \left(\frac{\lambda^a}{2}\right)_{\alpha\beta} q^\beta_j - \frac{1}{4} G^{\mu\nu}_a G^a_{\mu\nu} \tag{1.7}$$

where g_s is the strength of the interaction, universal for all quark flavours. The second term is interpreted as the interactions between the quarks and the gluon field G^a_μ where λ^a represents the eight Gell-Mann matrices and α and β label the quark colours. The third term is the kinetic term, with $G^{\mu\nu}_a = \partial^\mu G^\nu_a - \partial^\nu G^\mu_a - g_s f^{abc} G^\mu_b G^\nu_c$ where f^{abc} are the structure constants. There are eight gluons which also carry colour charge due to the non-abelian nature of the SU(3) group. This gives rise to self-interactions between the gluons and results in quarks only being observed in bound-states.

1.3 Electroweak interactions

Weak interactions explain a number of experimental observations such as parity violating beta decay, muon decay and the decays of charged pions. The weak and electromagnetic interactions are unified into a single theory described by an $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ group [1–3]. In electroweak theory, fermions have two additional quantum numbers, weak isospin T associated to $\mathrm{SU(2)_L}$ and hypercharge Y associated to $\mathrm{U(1)_Y}$. The hypercharge is related to the electric charge Q and the weak isospin by:

$$Q = T^3 + Y/2 \tag{1.8}$$

where T^3 is the 3rd component of weak isospin. All particles in the SM have non-zero hypercharge with the exception of gluons.

In the weak interaction, only left-handed chiral fermion states and right-handed chiral anti-fermion states participate. The left and right-handed chiral fermion states (ψ_L, ψ_R) are defined as:

$$\psi_L = \frac{1-\gamma^5}{2}\psi, \qquad \psi_R = \frac{1+\gamma^5}{2}\psi \tag{1.9}$$

where γ^5 is formed by the product of the Dirac matrices. In the case of massless particles, the chirality corresponds to the helicity. Right-handed (left-handed) helicity is defined for particles that have their spin pointing in the same (opposite) direction of their momentum. For anti-fermions this convention is reversed.

Electroweak interactions dictate how matter is organised in the Standard Model. All fermions are organised into three families that have identical quantum numbers but differ in mass. The left-handed chiral fermion states are combined into weak isospin doublets. Each doublet has weak isospin $T = \frac{1}{2}$ with the upper and lower members of each doublet having $T^3 = \pm\frac{1}{2}$ respectively. Table 1.1 shows the fundamental fermions in the Standard Model with their quantum numbers.

Leptons			s	T	T^3	Q	Y	
$\begin{pmatrix}\nu_L^e\\ e_L^-\end{pmatrix}$	$\begin{pmatrix}\nu_L^\mu\\ \mu_L^-\end{pmatrix}$	$\begin{pmatrix}\nu_L^\tau\\ \tau_L^-\end{pmatrix}$	$\frac{1}{2}$ $\frac{1}{2}$	$\frac{1}{2}$ $\frac{1}{2}$	$\frac{1}{2}$ $-\frac{1}{2}$	0 -1	-1 -1	
e_R^-	e_R^-	e_R^-	$\frac{1}{2}$	0	0	-1	-2	

Quarks			s	T	T^3	Q	Y	C
$\begin{pmatrix}u_L\\ d'_L\end{pmatrix}$	$\begin{pmatrix}c_L\\ s'_L\end{pmatrix}$	$\begin{pmatrix}t_L\\ b'_L\end{pmatrix}$	$\frac{1}{2}$ $\frac{1}{2}$	$\frac{1}{2}$ $\frac{1}{2}$	$\frac{1}{2}$ $-\frac{1}{2}$	0 -1	$\frac{1}{3}$ $\frac{1}{3}$	3 3
u_R	c_R	t_R	$\frac{1}{2}$	0	0	$\frac{2}{3}$	$\frac{4}{3}$	3
d_R	s_R	b_R	$\frac{1}{2}$	0	0	$-\frac{1}{3}$	$-\frac{2}{3}$	3

Table 1.1 *The fundamental fermions in the SM and their quantum numbers. The quantum numbers are spin s, weak isospin T, the 3^{rd} component of weak isospin T^3, electric charge Q, weak hypercharge Y. In addition, the quarks carry colour that can take the values red, green or blue.*

The electron, muon and tau particles each have a neutrino associated to them. The neutrinos carry no electric charge while the electron, muon and tau carry a charge of -1. The quarks carry fractional electric charge, $+2/3$ for up (u), charm (c) and top (t) and $-1/3$ for down (d), strange (s) and bottom (b). Each quark also has colour, conventionally taken as red, blue or green.

Particle interactions are introduced by invoking the principle of gauge invariance, as usual. Four vector fields, B_μ and W_μ^a (with a = 1,2,3), are introduced to restore the invariance under $\mathrm{U(1)_Y}$ and $\mathrm{SU(2)_L}$ respectively. For instance, if the first generation quark family is expressed as:

$$\psi_1(x) = \begin{pmatrix} u_L \\ d'_L \end{pmatrix} \qquad \psi_2(x) = u_R \qquad \psi_3(x) = d_R \tag{1.10}$$

then the associated electroweak Lagrangian is given by:

$$\mathcal{L} = \sum_{j=1}^{3} i\bar{\psi}_j(x)\gamma^\mu D_\mu \psi_j(x) - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \tag{1.11}$$

where j is the sum over the fermion generation as shown in equation 1.10 and D_μ is given by:

$$D_\mu \psi_j(x) = \left[\partial_\mu - ig\frac{\sigma_a}{2}W^a_\mu \delta_{1j} - ig'\frac{Y_j}{2}B_\mu\right]\psi_j(x) \tag{1.12}$$

where σ_a are the Pauli matrices. The second term is associated to $\mathrm{SU(2)_L}$ and acts only on ψ_L states with coupling g. The final term is associated to $\mathrm{U(1)_Y}$ which acts on both chiral states with coupling g'. The kinetic terms for each field are defined as:

$$B^a_{\mu\nu} = \partial_\mu B_\nu + \partial_\nu B_\mu, \tag{1.13}$$

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\epsilon^{abc}W^b_\mu W^c_\nu, \tag{1.14}$$

The charged vector bosons can then be written as a linear combination of the first two components of W^a_μ:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \tag{1.15}$$

The third component of W^a_μ mixes with B_μ via the weak mixing angle θ_W in order to form the neutral vector bosons A_μ and Z_μ.

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W^3_\mu \tag{1.16}$$

$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W^3_\mu \tag{1.17}$$

Similar to QCD, there are interactions between the bosons. There are three and

four point self-interactions between the $W^{\pm}$, Z and the photon with each of these interactions involving the presence of a $W^{\pm}$ pair.

1.4 The Higgs Mechanism

The Lagrangian in equation 1.11 describes the charged and weak interactions in the SM. However, there are no mass terms for the $W^{\pm}$ and Z bosons, or the fermions, which is in disagreement with experimental facts. Simply adding these terms to the Lagrangian would violate the gauge invariance. Mass has to be generated more subtly, a mechanism to do so was proposed by Peter Higgs [4], Robert Brout, Francois Englert [5] and others [6].

The Higgs mechanism introduces a doublet of complex scalar fields:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.18}$$

with the following Lagrangian:

$$\mathcal{L} = (D_\mu \phi)^\dagger (D_\mu \phi) - V(\phi) \tag{1.19}$$

where the potential $V(\phi)$ is postulated as:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \tag{1.20}$$

For the potential to have a minimum, the value of λ must be greater than 0. If $\mu^2 > 0$ is also chosen then the potential has a trivial minimum at $|\phi|^2 = 0$. However if $\mu^2 < 0$ is chosen, the potential no longer has a minimum at $|\phi|^2 = 0$ but a minima on the surface of fields ϕ that satisfy:

$$\frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2} \tag{1.21}$$

where v is the vacuum expectation value. An illustration of the potential for a complex scalar field with two degrees of freedom is shown in Figure 1.1.

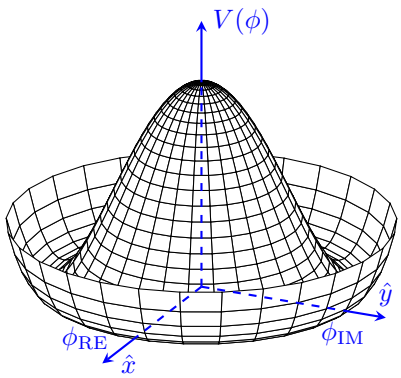

Figure 1.1 *Illustration of the Higgs potential*

The Lagrangian in equation 1.19 is invariant under $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$, however the choice of a particular minima breaks this symmetry. The theory now has to be developed around a minimum with the minimum usually chosen such that $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = v$. It is possible to write the expansion of the field ϕ around the minimum in the so called unitary gauge as:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ H(x) + v \end{pmatrix} \tag{1.22}$$

where $H(x)$ is the physical Higgs field. It is now possible to evaluate the Lagrangian in equation 1.19 in the unitary gauge, using equation 1.12. The Lagrangian obtained is:

$$\mathcal{L} = \frac{1}{4}(2\partial_\mu \partial^\mu + \lambda v^2 - 2\lambda v H - \lambda H^2)(v + H)^2 + \frac{g^2}{8}(2W_\mu^- W^{+\mu} + \frac{Z_\mu Z^\mu}{\cos^2\theta_W})(v + H)^2 \tag{1.23}$$

The masses of the $W^\pm$ and the Z bosons can now be identified by extracting the terms that contain the $W_\mu^\pm$ and Z_μ fields, but not the field H:

$$\frac{1}{2}\left(\frac{gv}{2}\right)^2 \left(2W_\mu^- W^{+\mu} + \frac{Z_\mu Z^\mu}{\cos^2\theta_W}\right) \tag{1.24}$$

The masses of the W and Z bosons are given by $M_W = \frac{1}{2}gv$ and $M_Z = \frac{gv}{2\cos\theta_W} = \frac{M_W}{\cos\theta_W}$. The doublet of complex scalar fields introduced four degrees of freedom, three of which have been absorbed by the $W^\pm$ and the Z bosons. The remaining degree of freedom is the Higgs boson, the mass of which can be identified by

finding the terms that are quadratic only in H^2, one finds $-\lambda v^2 H^2$ giving $M_H = \sqrt{2\lambda v^2}$.

The terms in equation 1.23 that contain the fields $W_\mu^\pm$ or Z_μ and the Higgs field describe interactions between the Higgs boson and the W or Z bosons with both VVH and $VVHH$ couplings existing. The strength of these interactions is proportional to M_W^2 and M_Z^2.

There are also terms that are proportional to H^3 and H^4 in the Lagrangian which describe the self-interactions of the Higgs boson. The Higgs self-couplings λ_{HHH} and λ_{HHHH} are given by:

$$\lambda_{HHH} = \frac{M_H^2}{2v} \qquad \lambda_{HHHH} = \frac{M_H^2}{8v^2} \tag{1.25}$$

The Higgs boson mass has been measured precisely in experiment and the vacuum expectation value can be determined from the ratio of the $W^\pm$ and Z boson masses which have also been measured in experiments with high precision. A measurement of λ_{HHH} can therefore test this relationship and confirm our understanding of electroweak symmetry breaking. The central topic of this thesis is the study of Higgs boson pair production which has a strong dependence on the value of λ_{HHH}.

1.4.1 Fermion masses

Fermion mass terms are not an intrinsic part of the Higgs mechanism but the terms can now be introduced without violating gauge invariance. For a fermion, the Lagrangian term postulated is:

$$\mathcal{L} = -\lambda_f[\bar{\psi}_L \, \phi \, \psi_R + \bar{\psi}_R \, \phi^\dagger \, \psi_L] \tag{1.26}$$

where λ_f is the Yukawa coupling for a fermion f. If one considers the electron and the scalar field in the unitary gauge, equation 1.26 becomes:

$$\begin{aligned}\mathcal{L} &= -\lambda_e \frac{1}{\sqrt{2}}[(\bar{\nu},\bar{e})_L \begin{pmatrix} 0 \\ v+H \end{pmatrix} e_R + \bar{e_R}(0, v+H) \begin{pmatrix} \nu \\ e \end{pmatrix}_L \\ &= \frac{\lambda_e \nu}{\sqrt{2}} \bar{e}e - \frac{\lambda_e}{\sqrt{2}} H\bar{e}e \end{aligned} \tag{1.27}$$

The mass of the electron can be identified as $m_e = \frac{\lambda_e \nu}{\sqrt{2}}$ and the coupling between the electron and the Higgs boson is given by $\lambda_e/\sqrt{2} = m_e/v$.

In the unitary gauge, the mass and interaction terms for a fermion with the Higgs boson can be described by the Lagrangian:

$$\mathcal{L} = -m_f f\bar{f}\left(1 + \frac{H}{v}\right) \tag{1.28}$$

1.5 Phenomenology

Lagrangians alone are not sufficient to make calculations and predictions for experimental measurements at the LHC. In proton-proton collisions at the LHC, a large number of perturbative and non-perturbative effects have to be taken into account to describe particle production from a collision. The non-perturbative nature of QCD in the low energy regime means that it is unfeasible to perform all the calculations necessary to make predictions at the LHC without using approximations and phenomenological models. Monte Carlo (MC) simulations are used to describe the initial collision to the interactions of the final state particles with the detector. A description of the necessary inputs to Monte Carlo simulations are given in this section.

1.5.1 Running of α_S

Experimental measurements of the QCD coupling constant, α_S ($=g_s^2/4\pi$) are shown as a function of momentum transfer Q in Figure 1.2. At low energy scales (Q ~ 1 GeV), α_s is approximately equal to 1 and perturbation theory cannot be used. However at higher energy scales (Q ~ 100 GeV) which is the typical energy scale for collisions of interest at the LHC, α_s decreases and can be of the order 0.1, a regime where perturbation theory can be used. This property of QCD is

known as asymptotic freedom, at high momentum transfer the quarks can be treated as free particles rather than bound to the protons. The evolution of α_S, only including Feynman diagrams with one loop in them is given by:

$$\alpha_S(Q^2) = \frac{\alpha_S(Q_0^2)}{1 + \frac{\beta_0}{4\pi}\alpha_s(Q_0^2)\ln\left(\frac{Q^2}{Q_0^2}\right)} \tag{1.29}$$

where Q_0 is some reference scale and β_0 is given by

$$\beta_0 = 11n - \frac{2}{3}f \tag{1.30}$$

where n is the number of colours (3) and f is the number of quark flavours (6). Since $\beta_0 > 0$, the QCD coupling constant decreases with increasing Q^2. Equation 1.29 can be simplified further in the regime where $\alpha_S(Q_0^2)$ diverges; the reciprocal of equation 1.29 can be written as:

$$\frac{1}{\alpha_S(Q^2)} = \frac{1}{\alpha_S(Q_0^2)} + \frac{\beta_0}{4\pi}\ln\left(\frac{Q^2}{Q_0^2}\right) \tag{1.31}$$

and if the limit is taken such that $\alpha_S(Q_0^2) \to \infty$, equation 1.31 can be written as:

$$\alpha_S(Q^2) = \frac{4\pi}{\beta_0 \ln\left(\frac{Q^2}{\Lambda_{QCD}^2}\right)} \tag{1.32}$$

where Λ_{QCD} is the QCD scale parameter, defined as the value of Q_0 such that $\alpha_S(Q_0^2) \to \infty$. Λ_{QCD} has a measured value of 210 $\pm$ 14 MeV [7].

1.5.2 Cross sections

The main interest at the LHC is the production of heavy particles like the Higgs boson. Heavy particles are usually produced in collisions where there is a large momentum transfer (Q^2) between the protons. In collisions with a large momentum transfer, the collision can be considered as occurring between two partons (a and b) which are constituents of the incoming protons (A and B). The cross section for the production of a particle c in a proton-proton collision is

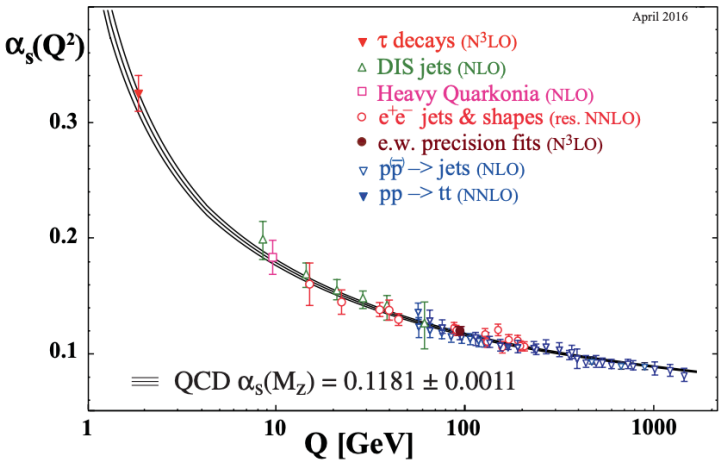

Figure 1.2 *Summary of experimental measurements and theoretical predictions of α_s as a function of the energy scale Q [7].*

given by

$$\sigma_{AB\to c} = \sum_{a,b} \int dx_a dx_b f_{a/A}\left(x_a, Q^2\right) f_{b/B}\left(x_b, Q^2\right) \hat{\sigma}_{ab\to c} \tag{1.33}$$

integrated over the fractions of proton momentum x_a and x_b carried by each interacting parton and where $\hat{\sigma}_{ab\to c}$ is the cross section for the elementary interaction between the partons producing particle c. The probability to have a parton within the proton carrying a fraction x of the total proton momentum ($f_{a/A}\left(x_a, Q^2\right)$) is called the parton distribution function (PDF). The product of the parton momentum fraction with the PDF as a function of parton momentum fraction is shown in Figure 1.3 for $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2.

In equation 1.33, Q^2 denotes the momentum transfer that characterises the production of particle c, but other interactions with lower Q^2 also exist in a proton-proton collision. These interactions can either be included as part of the perturbative calculation or they can be factorised into the PDFs. An unphysical scale parameter called the factorisation scale, μ_F is introduced to separate the high and low Q^2 interactions. Due to the running of α_S as discussed previously, a value of Q^2 must be chosen so that α_S can be evaluated. As a result of this, another unphysical scale parameter called the renormalisation scale, μ_R is introducd. Both scales are typically set to the mass(es) of the particle(s) being produced. It is conventional that the scale choices are varied up and down by a factor of two and the upper and lower cross section predictions are taken as an

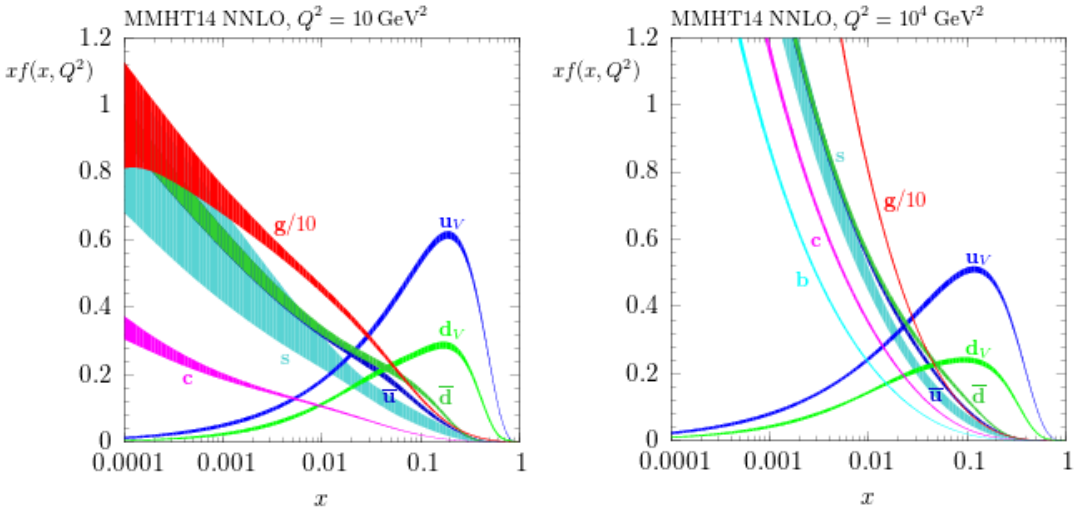

Figure 1.3 *The product of the parton momentum fraction x with the PDF computed for $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2 from the MMHT 2014 PDFs [8].*

uncertainty. Equation 1.33 becomes:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}\left(x_a, \mu_F^2\right) f_{b/B}\left(x_b, \mu_F^2\right) \times \left[\sigma_0 + \hat{\alpha}_S\left(\mu_R^2\right)\hat{\sigma}_1 + ...\right]_{ab\to X}, \quad (1.34)$$

where $\hat{\sigma}_0$ ($\hat{\sigma}_1$) is the leading order (next-to-leading order) cross section for the partonic process. The cross section dependency on the unphysical scales introduced decreases the more higher order corrections calculated and would vanish completely if it were possible to calculate to all orders in perturbation theory. This combined with the fact that α_s is approximately 0.1 for ($Q \sim$ 100 GeV) mean that it is essential for processes at the LHC to be calculated beyond leading order.

1.5.3 Event generation

The discussion in the previous section gave a recipe for calculating cross sections for a given process at the LHC. There are however a number of complex effects that still need to be accounted for.

The outgoing particle c referred to in the previous section can be unstable particles like electroweak bosons or it can be leptons, quarks or gluons. Some of these particles will decay before detection and, in the case of quarks or gluons, they will always hadronise, with the exception of the top quark. The hadronisation process

cannot be studied with perturbation theory since $\alpha_s \sim 1$ in this regime. Instead, these processes are simulated using phenomenological models that have had some experimental input. The proton remnants can also undergo further interactions which give rise to an underlying event. It is also possible that particles in the initial or final state emit radiation or additional particles. All these effects are all taken into account in Monte Carlo event generators. A pictorial representation of a $t\bar{t}H$ event is shown in Figure 1.4.

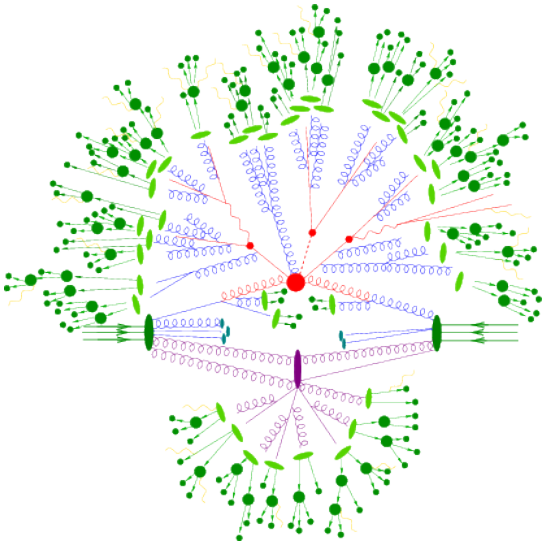

Figure 1.4 *Simulation of a $t\bar{t}H$ event in* SHERPA *[9].*

Matrix element

Event generators are programs capable of producing *events* for a given process. The first step of generating events is to calculate the matrix element for a given process as described previously. The total cross sections for all subprocesses that contribute to the overall process are evaluated and then an event in a given kinematic configuration is selected probabilistically.

Matrix element calculations are now automated at leading order and are usually possible for two incoming particles to up to six outgoing particles. As the number of outgoing particles increases so does the number of Feynman diagrams and this calculation becomes more computationally expensive.

Parton shower

After the matrix element calculation, the event has a fixed number of outgoing particles. Outgoing quarks or gluons can emit additional gluons or quark pairs which can in turn emit further particles, generating a shower of partons. These emissions can be included in the matrix element calculation, however their inclusion makes the calculation increasingly more complex and eventually unpractical. Instead, the emissions of quarks and gluons can be factorised from the matrix element and treated on a probabilistic basis in an algorithm referred to as the *parton shower*. The emissions start from the high energy partons in the matrix element and iteratively evolve towards a scale where non-perturbative physics sets in. The algorithm is then terminated and the hadronisation process is utilised.

Merging

The parton shower is based on approximations in the limit where the quarks and gluons emitted are almost collinear or with low p_{T} ("soft"). For high p_{T} ("hard") and well separated emissions, the parton shower modelling is likely to be inadequate and ideally these emissions should be included in the matrix element calculation. Combining the matrix element and the parton shower however can be non-trivial. For example, a Z boson plus an additional jet event can be obtained in two ways, the Z boson plus an additional parton can be calculated in the matrix element, or the Z boson alone can be calculated in the matrix element with the additional jet being produced from a hard well separated emission in the parton shower.

To resolve this issue, the event generators use merging procedures. The merging procedures decide on an event-by-event basis which of the two paths (matrix element or parton shower) should be followed. There are different approaches to this problem, but most of them follow a similar strategy.

One of the most common approaches is the MLM method [10] which is briefly summarised here. In the MLM method, matrix elements are calculated for the production of a particle, for example a Z boson plus up to an additional n partons where the number n has to be specified. Typically n can be no greater than 6 as

the calculation becomes increasingly more complex. All partons must have:

$$p_T^{part} > p_T^{min} \qquad |\eta_{part}| < \eta_{max} \qquad \Delta R_{ij} > R_{min} \tag{1.35}$$

where ΔR_{ij}[1] is the separation between two partons. The parameters p_T^{min}, η_{max} and R_{min} are all free parameters that must be specified. The parton shower algorithm is then executed and all particles are clustered to form jets using the k_T algorithm [11] which will be discussed in Section 2.3.6. All jets with p_{T} $> Q_{merge}$ are identified where Q_{merge} is another free parameter, usually taken between 20 - 30 GeV. If each parton that was calculated in the matrix element matches one of the jets (nearby in angle) and there are no additional jets above the scale Q_{merge}, the event is accepted, otherwise it is rejected. The rejection of events with additional jets is relaxed in the case where the matrix element has been calculated for $Z + n$ where n is the maximum number of additional partons specified. Additional jets from the parton shower are allowed here since there will be no overlap with any other subsample.

In practice

There are a large number of event generator programs now available. Some programs are capable of calculating the matrix element (e.g. MadGraph [12]) and then they are interfaced to another program which can handle the parton shower and hadronisation steps (e.g. Herwig [13]) but there are also others which have all the capabilities (e.g. Sherpa [14] and Pythia [15]).

Figure 1.5 shows comparisons between predictions from three different event generators against data from a 13 TeV Z + jets measurement [16] performed by the ATLAS collaboration. The jet multiplicity in Z + jets events appears to be well modelled up to four additional jets, then there is disagreement between the predictions and the data with Sherpa predicting too much jet activity and the others predicting too little.

The experimental uncertainties are shown with a grey band in the ratio panels. The theory uncertainties are estimated using Sherpa and are shown with an orange band in the ratio panels. This theory uncertainty band is a combination of the statistical uncertainty from the number of events generated in Sherpa, the

[1] ΔR is defined as $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$

PDF uncertainties, varying α_S and varying the factorisation and renormalisation scales.

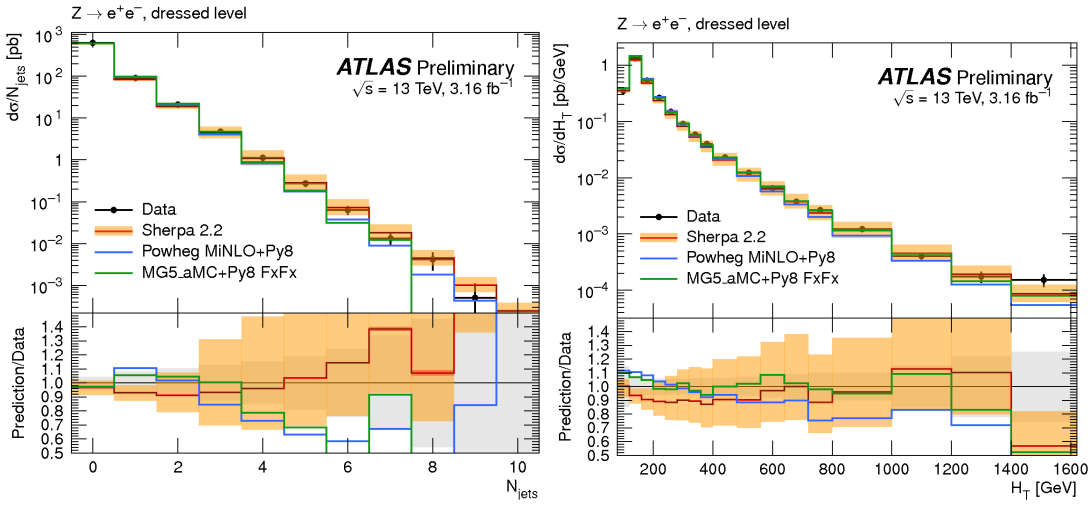

Figure 1.5 *Predictions for the differential cross sections as a function of the exclusive jet multiplicity (left) and the H_T (right) for different Monte Carlo event generators compared against data. The grey band shows the size of the experimental uncertainty while the orange band shows the theory uncertainty which was estimated using* Sherpa. *A full description of the different event generators can be found in Ref. [17].*

1.5.4 Summary

Event generation at the LHC is extremely complex due to the nature of QCD. Despite the relative success of the event generators in predicting data distributions, they are built on approximations with corresponding uncertainties in each step (PDF, α_s, choice of QCD scales etc.). It is difficult to imagine that the separation of the various steps from the hard scatter to the final state particles is as distinct in nature as it is in the current event generators.

Nevertheless, event generators are essential tools used in all physics analyses performed at the LHC. Every physics analysis relies upon them to model a given signal process and many heavily rely on them to model background processes as well. Given this, the event generators can have a huge impact on the precision of many measurements and further improvements are needed in order to progress.

1.6 Higgs physics at the LHC

The Higgs boson couplings to the fermions and bosons are proportional to the masses of the particles. The dominant Higgs boson production mechanisms therefore involve the heaviest particles in the Standard Model. Ideally, one would like to use all the production mechanisms and final states for measurements. However this isn't always possible due to the overwhelming QCD backgrounds and trigger requirements at the LHC. An efficient trigger requires charged leptons, photons, large $E_{\mathrm{T}}^{\mathrm{miss}}$ or high p_{T} jets at the LHC.

Higgs boson production modes at the LHC

There are four main production modes for the Higgs boson at the LHC which are shown in Figure 1.6 and described below.

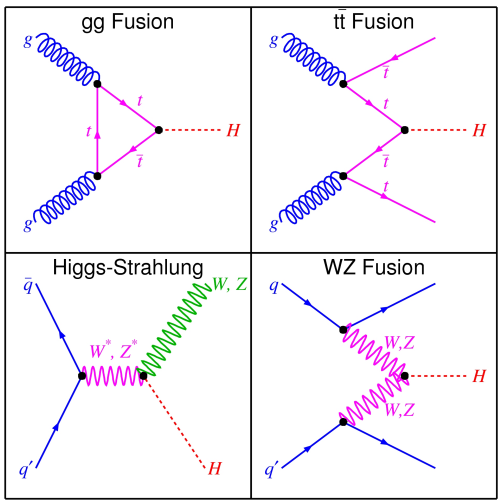

Figure 1.6 *Main Higgs boson production mechanisms [18].*

- **Gluon-gluon fusion (ggF)**. This process is mediated by a quark loop with the dominant contribution coming from the top quark. This production mode has the highest cross section at the LHC since gluons are the largest component of the protons PDF at high energy (see Figure 1.3). A huge effort has been made to calculate this process with great precision. It is now known at next-to-next-to-next-to-leading order ($\mathrm{N^3LO}$) with uncertainties on the cross section at the ~5% level [19]. Before the discovery of the Higgs

boson, uncertainties were as high as ~40% for this process. This production mechanism makes up approximately 88% of total Higgs production.

- **Vector boson fusion (VBF)**. Two incoming quarks exchange a W or Z boson which fuse together to make a Higgs boson. The outgoing quarks give rise to jets located in the forward region of the detector with a large pseudorapidity gap between them. At leading order, VBF is purely an electroweak process, and as a result of this, no colour is exchanged between the incoming quarks. This results in little additional jet activity in the pseudorapidity region between the two forward jets. This feature can be used to separate VBF production from the backgrounds, including ggF + 2 jets Higgs production. VBF contributes around 7% to Higgs production.

- **Associated production with a W or Z boson (VH)**. Higgs production in association with a W or Z boson is initiated by $q\bar{q}$. The decays of the vector bosons to charged leptons or neutrinos provide a useful trigger and help reduce QCD backgrounds. This process takes around 4% of Higgs production.

- **Associated production with top quarks ($t\bar{t}H$)**. This can be initiated by gluons or quarks with the Higgs being radiated from a top quark line. The $t\bar{t}H$ process is important for directly measuring the coupling between the top quark and the Higgs boson. This process only makes up approximately 1% of Higgs production.

The ATLAS collaboration has now observed ($\geq 5\,\sigma$) the four production modes discussed above [20–23]. There are additional small Higgs production modes that were not discussed above. Similar to $t\bar{t}H$, there is also associated production with bottom quarks ($b\bar{b}H$). This process has approximately the same cross section as $t\bar{t}H$, however the jets originating from the bottom quarks have low p_{T} and often fail the jet p_{T} thresholds that must be imposed in order to suppress spurious jets at the LHC. Finally there is also associated production with a single top quark (tHW and $tHjb$). The $tHjb$ and tHW processes have cross sections approximately a factor of 10 and 30 respectively smaller than the $t\bar{t}H$ process. They are also difficult to isolate from $t\bar{t}H$ production.

Higgs boson pair production at the LHC

The central topic of this thesis is not single Higgs boson production, but Higgs boson pair production which has a strong dependence on the Higgs self-coupling, λ_{HHH}. At $\sqrt{s} = 13$ TeV, the production cross section for HH is orders of magnitudes smaller than single Higgs production, for gluon-gluon fusion HH the cross section is 33 fb [24], in comparison to ggF single Higgs which has a cross section of 49000 fb [19]. The production mechanisms for HH are discussed below.

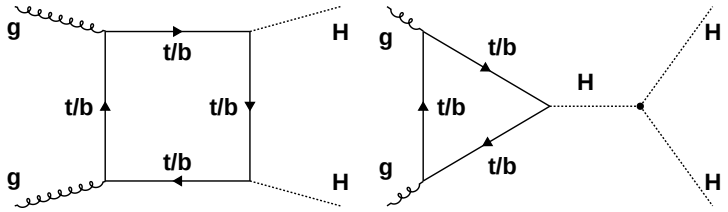

Figure 1.7 *Examples of two Feynman diagrams that contribute to $gg \rightarrow HH$ production.*

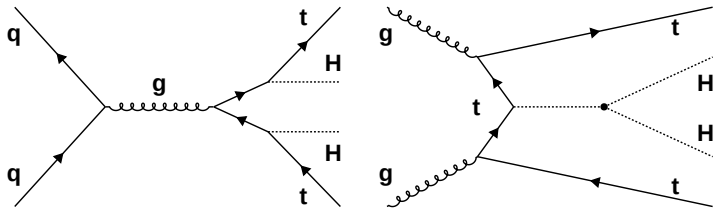

Figure 1.8 *Examples of two Feynman diagrams that contribute to $t\bar{t}HH$ production.*

- **Gluon-gluon fusion (ggHH)**. Two leading order Feynman diagrams are shown in Figure 1.7. There is interference between the *box* diagram (left) and *triangle* diagram (right). The impact of the interference is shown in Figure 1.9, the interference significantly reduces the cross section near the Standard Model value of λ_{HHH} with the maximum amount of destructive interference occurring at $\lambda_{HHH} \approx 2$ times the SM value.

- **Vector-boson fusion (qqHH)**. The cross section for this process is approximately 4% of the gluon-gluon fusion HH process. The characteristics are similar to single VBF Higgs production. However, as mentioned previously, one of the most powerful discriminants in separating VBF from

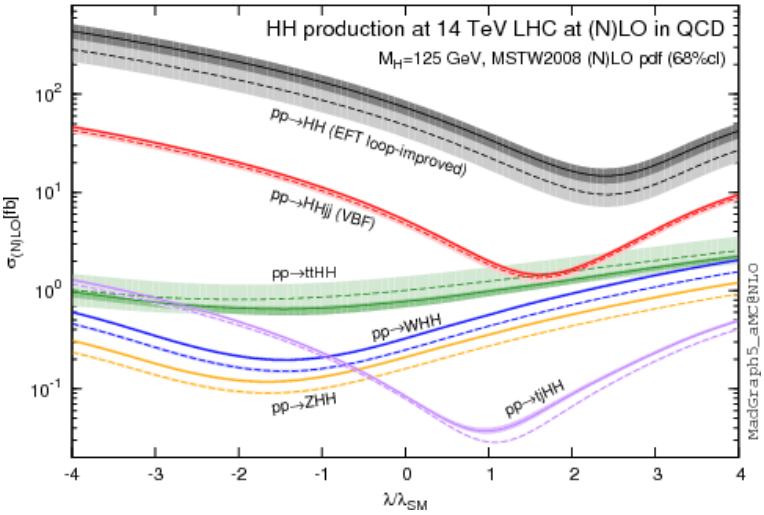

Figure 1.9 *The cross section for HH production mechanisms as a function of the Higgs self-coupling, λ_{HHH}. The dashed (solid) lines and light (dark) coloured bands correspond to the LO (NLO) results and to the scale and PDF uncertainties added together [25].*

the backgrounds is the lack of central jet activity. In HH searches usually one of the Higgs bosons is required to decay into $b\bar{b}$ due to its large branching ratio. This requirement therefore removes one of the main discriminants which is usually powerful in separating VBF production. This process is also sensitive to the $HHVV$ coupling.

- **Associated production with top quarks ($t\bar{t}HH$).** Two leading order Feynman diagrams for this process are shown in Figure 1.8. Since the top quarks almost always decay into W bosons and bottom quarks. The charged leptons from the decay of a W boson can be used for an efficient trigger. It can be seen in Figure 1.9, there is constructive interference between the diagrams. If $\lambda_{HHH} \geq \lambda_{SM}$, it could be possible that $t\bar{t}HH$ offers complementary information to ggHH production. The prospects for observing this process at the LHC with the upgraded ATLAS detector is the topic of Chapter 4.

There is also HH production in association with a W or Z boson, VHH and in association with a single top quark, $tjHH$. However the cross section for these processes are even smaller than those listed above and therefore they are unlikely to be of any interest at the LHC.

Decay channel	Branching ratio (%)
$H \to b\bar{b}$	57.5 ± 1.9
$H \to WW^*$	21.6 ± 0.9
$H \to gg$	8.56 ± 0.86
$H \to \tau^+\tau^-$	6.30 ± 0.36
$H \to c\bar{c}$	2.90 ± 0.35
$H \to ZZ^*$	2.67 ± 0.11
$H \to \gamma\gamma$	0.228 ± 0.011
$H \to Z\gamma$	0.155 ± 0.014
$H \to \mu^+\mu^-$	0.022 ± 0.001

Table 1.2 *The main Higgs boson branching ratios in descending order.*

Higgs boson decay modes

The branching ratios for the Standard Model Higgs boson are listed in Table 1.2. Despite the fact that $m_H < 2m_W$ and $m_H < 2m_Z$, the Higgs boson can still decay to W and Z bosons when one of the W or Z bosons is produced off-mass-shell. The Higgs boson can also decay to massless particles, like $H \to \gamma\gamma$, through loops of virtual particles. A description of the experimental search for each major decay mode is given below.

- $H \to b\bar{b}$ has the largest branching ratio at 58%. However its observation at the LHC suffers from large QCD backgrounds, orders of magnitude above the signal. This final state is usually paired with the VH production mechanism since the leptons from the decay of the vector boson offer an efficient trigger. The ATLAS collaboration now has a $5\,\sigma$ observation of the $VH(H \to b\bar{b})$ process [21]. The $b\bar{b}$ decay is also studied with $t\bar{t}H$ production, however the challenging $t\bar{t}$ + jets background make this difficult also. For a long time, $H \to b\bar{b}$ using the ggF production mechanism was assumed to be impractical due to the lack of an efficient trigger and the QCD backgrounds. However, the CMS collaboration [26] showed by selecting highly boosted Higgs candidates as a single large radius jet with p_{T} greater than 450 GeV, it was possible to have some sensitivity to this process. The CMS collaboration measured $1.5\,\sigma$ ($0.7\,\sigma$ expected) [27]. No other Higgs decay channel has the event statistics to have any sensitivity in the p_{T} regime greater than 450 GeV yet.

- $H \to WW^*$ is only observable at the LHC when both W bosons decay into a charged lepton and a neutrino. The presence of the charged leptons and

the missing energy provide an efficient trigger. However the $t\bar{t}$ and the Z + jets background make this analysis challenging. The analysis is separated into categories where the leptons have different flavours. The $e\nu\mu\nu$ final state is the most important as the different lepton flavours suppress the Z + jets background.

- $H \to ZZ^*$ with both of the Z bosons decaying into a pair of electrons or muons. Although the rate of $H \to ZZ^* \to 4l$ rate is extremely low, the backgrounds are even smaller. The signal to background ratio is greater than 2.

- $H \to \tau^+\tau^-$. Approximately 42% of the time, both taus decay hadronically, 46% of the time, one decays hadronically and the other leptonically and in 12% of cases, both taus decay leptonically. Again, this analysis is challenging because of the $Z \to \tau^+\tau^-$ background and for the hadronic tau decays, multijet background with jets faking taus.

- $H \to \gamma\gamma$. Despite the low branching ratio, it offers a distinct signature of two high energy isolated photons. The background consists of $\gamma\gamma$ production with small contributions from γ + jets and multijets. A signal and background fit can be performed to the $m_{\gamma\gamma}$ distribution as the Higgs boson appears as a narrow signal peak over a smooth monotonically falling background.

The ATLAS collaboration has observed ($\geq 5\,\sigma$) all decays discussed above [21, 28–31] and now explores all the main production modes for the above decays.

The other final states in Table 1.2 have not been observed. It will not be possible to observe $H \to gg$ or $H \to c\bar{c}$ at the LHC assuming their Standard Model rates. ATLAS has set an observed (expected) upper limit on $ZH(H \to c\bar{c})$ at the 95% confidence level (CL) of 110 (150) $\times$ the Standard Model value [32]. This result has been extrapolated to a data set of $3000\,\mathrm{fb}^{-1}$, approximately the data set size the ATLAS detector is expected to collect in its lifetime which results in the limit improving to 6.3 times the SM value [33]. Observations of $H \to \mu^+\mu^-$ and $H \to Z\gamma$ will likely be possible in future LHC runs [34].

Higgs boson pair decay modes

The small HH cross sections mean that only the final states which have the largest branching ratios can be expected to have sensitivity to SM HH production. The branching ratios for pairs of Higgs bosons is shown in Figure 1.10.

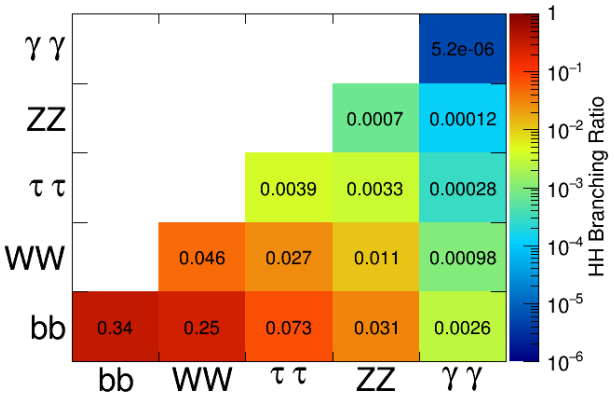

Figure 1.10 *The branching ratios for pairs of Higgs bosons.*

- $HH \to b\bar{b}b\bar{b}$ has the largest branching ratio at 34%. However similar to $H \to b\bar{b}$, this final state suffers from large QCD backgrounds and an inefficient trigger. The efficiency of the trigger is expected to get worse during the High Luminosity LHC program which makes this final state more challenging.

- $HH \to b\bar{b}\tau\tau$ has the third largest yield. This is split into two categories, one category in which both taus decay hadronically and another category where one tau decays hadronically and another leptonically. The backgrounds include $t\bar{t}$ and $Z \to \tau^+\tau^-$ + jets.

- $HH \to b\bar{b}\gamma\gamma$. This final state has a small yield, however the photons offer a clean trigger and excellent mass resolution. Similarly to $H \to \gamma\gamma$, the signal can be extracted by performing a fit to the $m_{\gamma\gamma}$ distribution. This final state is the topic of Chapter 5.

The ATLAS collaboration also searches for HH in a number of other decay modes. For example $b\bar{b}WW^*$, $WW^*\gamma\gamma$ and WW^*WW^*. However the sensitivity of these decay modes to SM HH production are significantly worse than those discussed above. The $b\bar{b}WW^*$ has the second largest yield but unfortunately this is identical to the $t\bar{t}$ final state.

It is not clear at present whether a $5\,\sigma$ observation of HH production will be achieved at the LHC. The work in this thesis sets an observed (expected) upper limit on HH production at the 95% CL of 10.1 (7.0) times the SM value. The ratio of the Higgs boson self-coupling to its SM expectation, κ_λ is constrained at 95% CL to $-4.4 < \kappa_\lambda < 9.5$ while the expected constraint was $-3.3 < \kappa_\lambda < 8.9$.

1.6.1 Beyond the Standard Model

The current LHC data set does not have sensitivity to SM HH production. However, there are a number of modifications to the SM that could enhance the rate of HH production. For example, it can be seen in Figure 1.9 that the rate of HH production increases significantly when the Higgs boson self-coupling is modified. HH production is also sensitive to the coupling of the top quark to the Higgs boson, however the current measurements of the $t\bar{t}H$ process [23] are compatible with the SM which suggests that significant modifications of the top quark coupling are already excluded.

Effective field theories (EFT) can be used to probe physics beyond the Standard Model. In an EFT approach, additional dimensional-D operators O_i^D are added to the SM Lagrangian $\mathcal{L}_{SM}$, leading to the effective Lagrangian:

$$\mathcal{L}_{EFT} = \mathcal{L}_{SM} + \sum_{i,D} \frac{c_i^{(D)}}{\Lambda^{D-4}} \mathcal{O}_i^{(D)} \tag{1.36}$$

The dimensionless coefficients $c_i^{(D)}$ specify the strength of the new interactions and are known as *Wilson* coefficients. These Wilson coefficients are the results of integrating out the heavy degrees of freedom of some new physical theory at an energy scale Λ which is much larger than the energy scales being probed. Some of these operators describe new interactions between the Higgs boson and the SM particles and therefore these operators can modify the HH production cross section as well as its kinematic properties. There are five dimension-6 operators which are directly relevant for HH production, shown in Figure 1.11.

Finally, there are also models [36, 37] that predict the existence of new particles that decay to pairs of Higgs bosons, often referred to as resonant HH production. This is not explored in this thesis.

The search for Higgs boson pair production is therefore well motivated with the

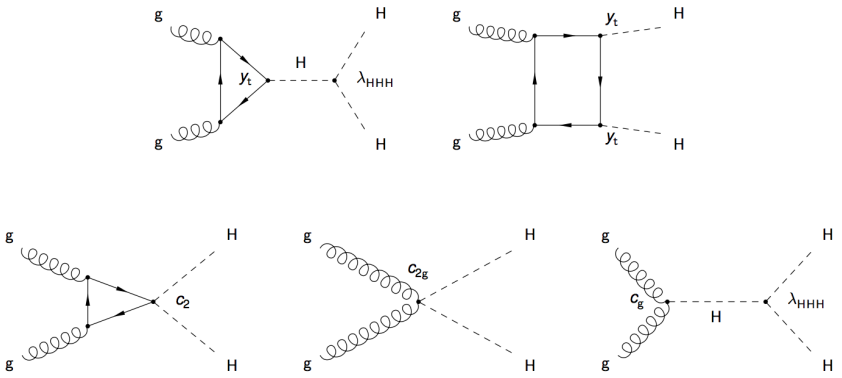

Figure 1.11 *Top: Feynman diagrams that contribute to SM HH production. Bottom: BSM Feynman diagrams that could contribute to HH production [35].*

current LHC data set despite the lack of sensitivity to SM HH production.

Chapter 2

The LHC and the ATLAS detector

In this chapter a short overview of the LHC accelerator and the ATLAS detector is presented.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's most powerful facility for particle physics research. It is a superconducting collider that is capable of accelerating counter-rotating proton beams to a centre-of-mass energy of $\sqrt{s} = 14$ TeV.

The LHC is housed in the tunnel previously used for the Large Electron Position (LEP) [38] accelerator at CERN near Geneva, Switzerland. The tunnel lies between 45 m and 170 m underground and has a circumference of about 27 km. The LEP accelerator operated as an e^+e^- collider between 1989 and 2000 and reached a centre-of-mass energy of $\sqrt{s} = 209$ GeV, crucially below the centre-of-mass energy required for $e^+e^- \to ZH$ production and therefore preventing the discovery of the Higgs boson.

The LHC currently holds the record for the highest energy collisions with a centre-of-mass energy of $\sqrt{s} = 13$ TeV, more than 13 times higher than that previously held by the Tevatron [39]. To reach these high energies, the LHC utilises the CERN accelerator complex (Figure 2.1).

Protons are produced by a duoplasmotron source at 100 keV and are then accelerated by a linear accelerator (LINAC2) to 50 MeV. This is followed by a

circular booster (PSB) where the protons attain energies of 1.4 GeV. Finally, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) accelerate the protons to 26 GeV and 450 GeV respectively. The protons enter the LHC ring and are accelerated by superconducting radiofrequency (RF) cavities which increase the beam energy by 485 keV each turn until they reach $\sqrt{s}/2$. The protons follow circular trajectories due to the magnetic fields produced from 1232 Niobium-Titanium superconducing dipole magnets. These magnets are cooled to temperatures below 2 K using superfluid Helium and produce fields above 8 T.

The LHC beams can collide at four different points, with each hosting a large experiment: ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Soldenoid), ALICE (A Large Ion Collider Experiment) [40] and LHCb [41].

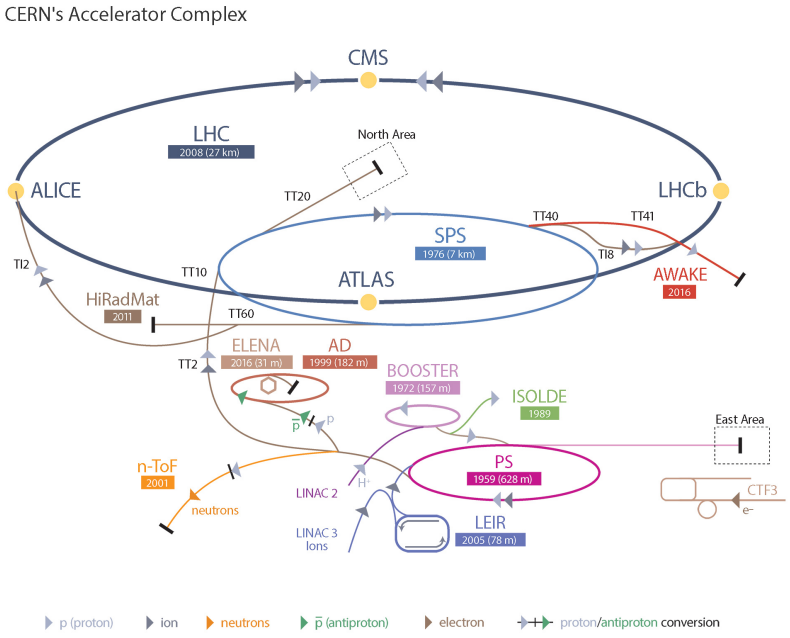

Figure 2.1 *The CERN accelerator complex [42].*

2.1.1 Luminosity and pile-up

The number of events expected from a process with a cross section σ is given by the product

$$N = L\sigma \tag{2.1}$$

where L is the integrated luminosity given by $\int_0^T \mathcal{L}dt$. Here, T is the length of time the ATLAS detector has been collecting data for and $\mathcal{L}$ is the instantaneous luminosity. The instantaneous luminosity depends only on the parameters of the beam and is given by the equation

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F \tag{2.2}$$

where (the nominal parameters for the LHC are given in parenthesis):

- N_b is the number of particles per bunch (~10^{10} - ~10^{11})
- n_b is the number of bunches per beam (2808)
- f_{rev} is the revolution frequency (11,245 Hz)
- γ_r is the relativistic gamma factor (~7000)
- ϵ_n is the transverse normalised beam emittance, related to the spread of particles in the beam (3.75 μm)
- β^* is the beta focus function at the collision point, related to how tightly the magnets are focused at the interaction point (0.55 m)
- F is the geometric luminosity reduction factor if the beams do not collide head-on. A crossing angle of 285 μrad is introduced to prevent collisions outside the nominal interaction points.

Using the nominal parameters, the LHC was expected to operate at 10^{34} cm^{-2}s^{-1} which it has now exceeded. Figure 2.2 shows the cumulative integrated luminosity delivered to the ATLAS detector in the years 2011 - 2018.

Another important parameter related to the instantaneous luminosity is the mean number of inelastic interactions per bunch crossing, known more colloquially as the *pile-up*, $\langle\mu\rangle$. The pile-up degrades the performance of the reconstructed

physics objects and makes the identification of vertices and the reconstruction of tracks more difficult. To compound this further, the nominal 25 ns between each bunch crossing is faster than the response time of some of the sub-detector systems. This means effects of events from adjacent bunch crossings can also be present in the event of interest. This is known as out-of-time pile-up. The luminosity recorded by ATLAS in the years 2015 - 2018 is shown as a function of $\langle\mu\rangle$ in Figure 2.2.

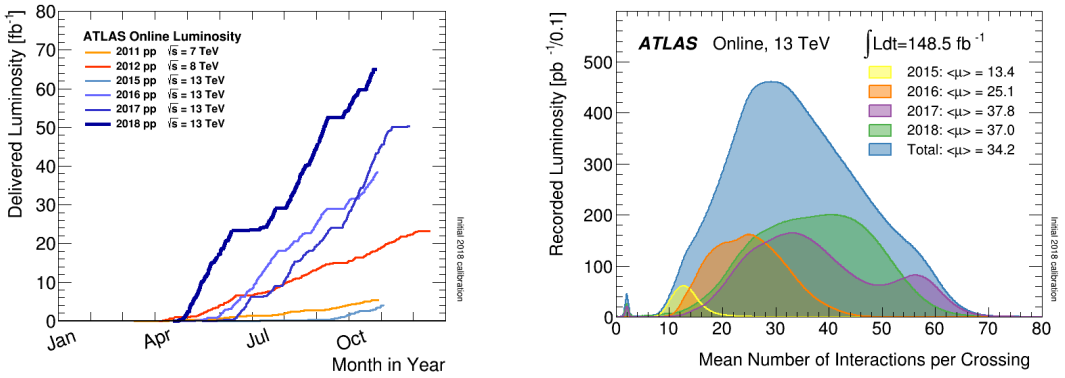

Figure 2.2 *Left: Cumulative luminosity delivered to the ATLAS detector as a function of day of the year, separately for each year. Right: The luminosity recorded by ATLAS shown as a function of $\langle\mu\rangle$ [43].*

In the years 2011 and 2012, the LHC succesfully delivered proton-proton collisions to the ATLAS detector at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV with ATLAS recording data sets of $5.2\,\text{fb}^{-1}$ and $21.7\,\text{fb}^{-1}$ respectively. This period of data taking receives the name *Run 1*. In 2015 - 2018, the ATLAS detector recorded $148.5\,\text{fb}^{-1}$ of proton-proton collisions at $\sqrt{s} = 13$ TeV. This period of data taking is referred to as *Run 2*.

2.2 The ATLAS detector

The ATLAS detector is one of four detectors at the LHC and one of two general purpose detectors, the other being CMS. ATLAS is the biggest in physical size, being 25 m in height and 44 m in length as shown in Figure 2.3. The detector is installed approximately 100 m underground at point 1 in Figure 2.1.

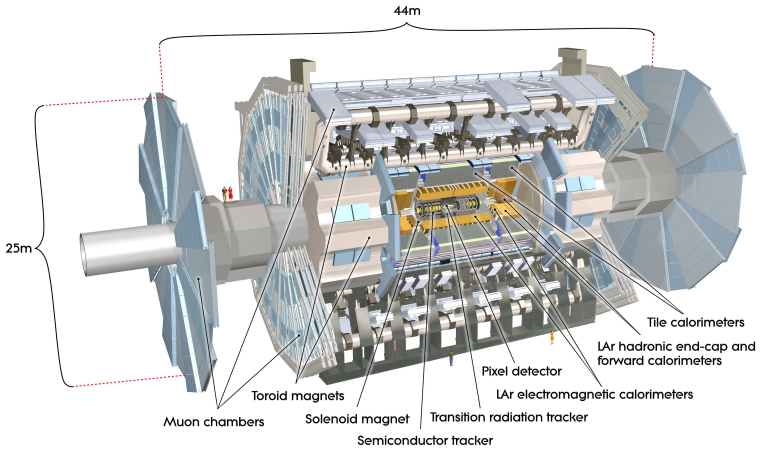

Figure 2.3 *The ATLAS detector, 25 m in height and 44 m in length. The weight of the detector is ~ 7000 tonnes [44].*

The desire to discover or exclude the SM Higgs boson guided the design of the ATLAS detector alongside ensuring the detector was suited for a wide range of beyond the Standard Model scenarios. These searches require that the detector is able to measure final state objects like photons, electrons, muons, hadronic jets and missing transverse energy over a huge range of energies.

General purpose detectors like the ATLAS detector use a cylindrical design. Nested barrel cylinders are installed at increasing radii around the interaction point (IP), followed by end-caps that cover the regions further along the beam axis. A tracking detector (named Inner Detector in ATLAS) is located closest to the IP and measures particle trajectories. The ID is surrounded by calorimeters designed to stop and measure particle energies through their interactions with detector material. The ATLAS detector is designed such that all particles are absorbed by the calorimeters with the exception of muons and neutrinos. Muons have a low interaction probability and can easily penetrate calorimeters, as a result a dedicated muon detector is employed outside all the sub-detectors. The neutrino interaction probability is too low to build a detector that would be small enough for collider physics. The neutrinos therefore escape detection and

instead they are inferred through an imbalance of momentum. In order to provide an accurate measurement of this imbalance, hermetic calorimeters are of vital importance. In what follows, the coordinate system used by ATLAS will be detailed followed by descriptions of each of the sub-detector systems. A comprehensive description of the whole detector, including all technical aspects can be found in Reference [44].

The coordinate system used to describe ATLAS and the particles it detects is summarised in Figure 2.4. The centre of the detector corresponds to the interaction point and defines the origin of the coordinate system. The beam direction defines the z-axis while the positive x-axis points towards the centre of the LHC ring and the positive y-axis points upwards. The polar angle θ is defined as the angle from the beam axis and the azimuthal angle ϕ is around the beam axis. The pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$, often it is preferred to describe the direction of a particle in η rather than θ as differences in pseudorapidity are Lorentz invariant under boosts in the z-direction for massless particles. The angular separation in $\eta - \phi$ space is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. The transverse quantities (p_{T}, E_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$) are defined in the $x - y$ plane.

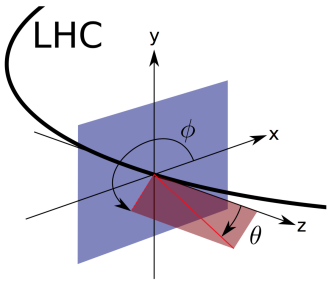

Figure 2.4 *Illustration of the coordinate system used in ATLAS.*

2.2.1 Magnet System

A strong magnetic field is necessary to measure the momenta and direction of charged particles emerging from the interaction point. In ATLAS, this field is provided by a small 2 T solenoidal magnet which bends the particles in the $x - y$ plane. A strong magnetic field is also required outside the calorimeters in order to trigger on and possibly measure the momenta of muons. ATLAS implemented huge toroidal magnets which has enabled triggering on muons up to an $|\eta|$ of 2.4 and momentum measurements independent of the Inner Detector.

The configuration of the magnet system is shown in Figure 2.5.

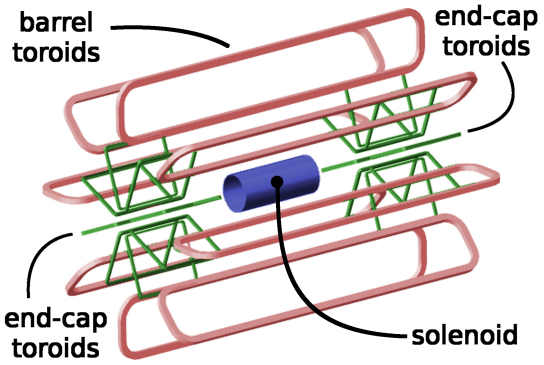

Figure 2.5 *Schematic diagram of the ATLAS magnet system [45].*

2.2.2 Inner Detector

The Inner Detector (ID) is composed of three sub-detector systems that exploit different technologies. The granularity of the ID sub-detectors decrease while moving away radially from the interaction point. Closest to the interaction point is the high granularity pixel detector that utilises silicon pixels. This is followed by the Semi-Conductor Tracker (SCT) detector that utilises silicon microstrips. The silicon detectors are surrounded by a transition radiation tracker (TRT) that uses straw-tubes. The ID is shown in Figure 2.6. It is immersed in the 2 T magnetic field produced by the solenoid which is placed between the ID and the calorimeter.

The inner detector is divided into two sections, the barrel which is parallel to the beam axis and two end-caps perpendicular to the beam axis. In the barrel region, both the Pixel and SCT detectors are arranged as concentric cylinders around the beam axis while in the end-cap regions, the detectors are disks. This set up means the ID is capable of reconstructing charged particle tracks up to $|\eta| \leq 2.5$.

Pixel Detector

The pixel detector is the first detector a particle from the interaction point will traverse, the first layer is 50.5 mm away from the interaction point. The pixel detector has 3 layers in the barrel and in the end-cap. Each module has 46080 pixels, resulting in a total of 80.4 million pixels. The pixel size is 50 μm in the $R-\phi$ direction and 400 μm in the z direction.

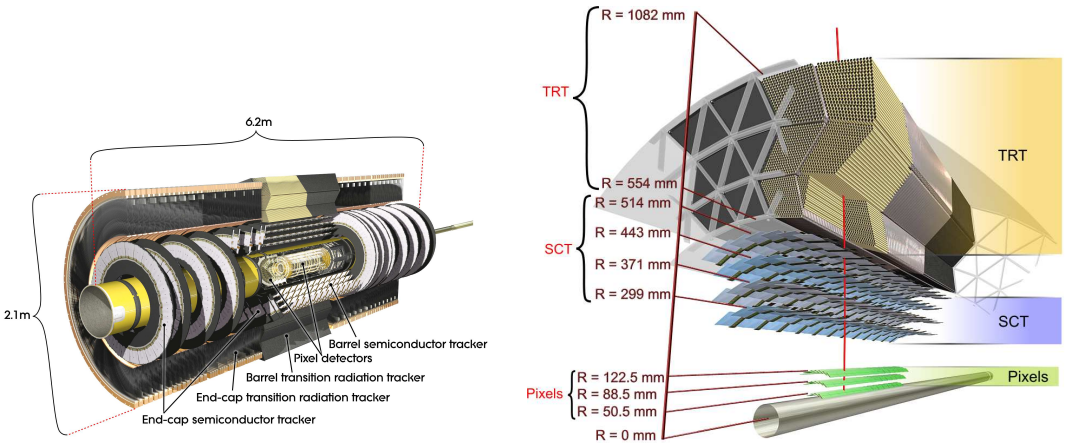

Figure 2.6 *Left: The ATLAS Inner Detector. Right: Vertical cut out of the ATLAS Inner Detector. The drawing shows a charged track with* p_T = 10 GeV *at an* η *of 0.3 [44].*

At the end of Run 1, the *insertable b-layer* (IBL) was added to the pixel detector. This layer was inserted at $R = 33.25$ mm in order to improve the tracking and b-tagging efficiency for Run 2 since the first layer of the pixel detector had deteriorated during Run 1 due to radiation damage. The pixel size in the IBL is 50 μm in the $R-\phi$ direction and 250 μm in the z direction. Figure 2.7 shows the improvement of the transverse impact parameter, d_0 (the distance in the $x-y$ plane between the tracks closest point to the z-axis and the z-axis itself) and the longitudinal impact parameter, z_0 (the distance between the z-position on the track where d_0 has been defined and the primary vertex) with the IBL.

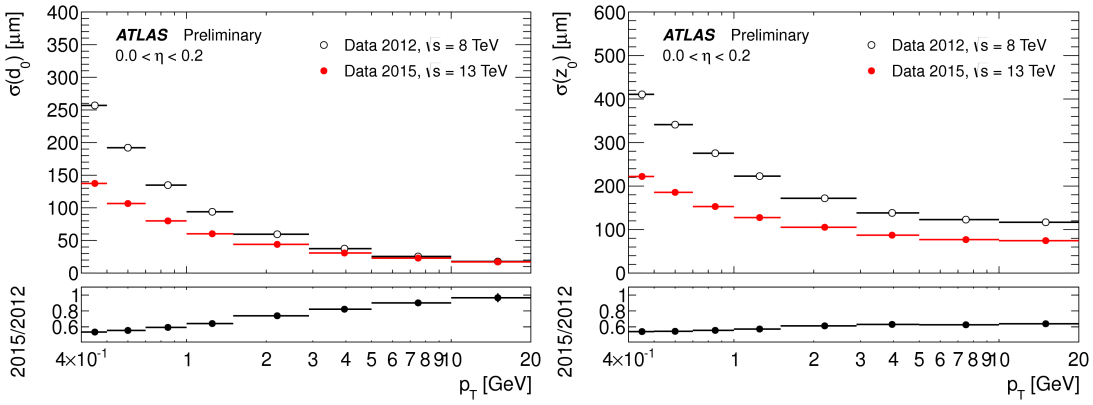

Figure 2.7 *Left: The transverse impact parameter resolution as a function of* p_T *measured in 2015 with the IBL and in 2012 without the IBL. Right: The longitudinal impact parameter resolution as a function of* p_T *measured in 2015 with the IBL and in 2012 without the IBL [46].*

SCT Detector

The SCT detector is placed between $R = 299$ mm and 514 mm. There are eight strip layers in the barrel and nine layers in the end-cap. In the barrel, the eight layers are arranged into four cylinders. The modules in the two layers per cylinder are glued back-to-back with a 40 mrad stereo angle between them. This method significantly improves the z measurement of each space-point. In total there are 6.2M read out channels in the SCT Detector.

TRT Detector

The TRT detector is composed of straw tubes filled with a Xenon based gas mixture which are separated by polypropylene radiator foils. Transition radiation (TR) is emitted when a relativistic charged particle traverses boundaries between materials of different refractive indices. This radiation is absorbed by the Xenon gas mixture inducing a signal in addition to the charged track ionisation. The TR signal is proportional to the γ factor of the charged particle which allows the mass of the charged particle to be deduced, given a measurement of its energy. The TRT can therefore discriminate between pions and electrons.

The TRT only provides $R-\phi$ information in the barrel and $z-\phi$ in the end-cap. At the LHC design luminosity 10^{34} cm^{-2}s^{-1}, the TRT straw occupancy can reach 60% [47]. The high luminosity program of the LHC plans to reach luminosities of $\geq 7.5 \times 10^{34}$ cm^{-2}s^{-1} which results in the TRT performance being completely inadequate since its occupancy will approach 100%. Before the high luminosity program begins, the ID including the TRT will be completely replaced with a new all silicon inner tracker, a project known as the ITk and described in detail in the next chapter.

The amount of material used by the ID is important to understand the behaviour of particles before they reach the calorimeters. For example photons can convert into e^+e^- pairs and electrons can lose energy through bremsstrahlung emissions which impacts the energy measurements of these particles in the calorimeter.

The characteristics of a material interacting with high energy particles is quantified using the radiation length, X_0 and the hadronic interaction length, λ. The radiation length X_0 is defined as the mean path length over which a high energy electron loses all but $1/e$ of its energy due to bremsstrahlung. Similarly,

the hadronic interaction length λ is the mean path length required in order to reduce the flux of primary hadrons by a factor of $1/e$. Both of these quantities are shown as a function of $|\eta|$ in Figure 2.8. For $|\eta| \leq 0.6$, the radiation length from the ID only is approximately ~ 0.5. However in the so called *crack region*, where the service infrastructure is located, the radiation length can exceed ~ 2.0. Usually in a physics analysis, electron and photon candidates in the interval $1.37 \leq |\eta| \leq 1.52$ are excluded. However, in the measurement of the W boson mass which used a data set of $4.6\,\text{fb}^{-1}$ at $\sqrt{s} = 7$ TeV [48], the ATLAS collaboration excluded the $|\eta|$ region [1.2, 1.82] for electrons.

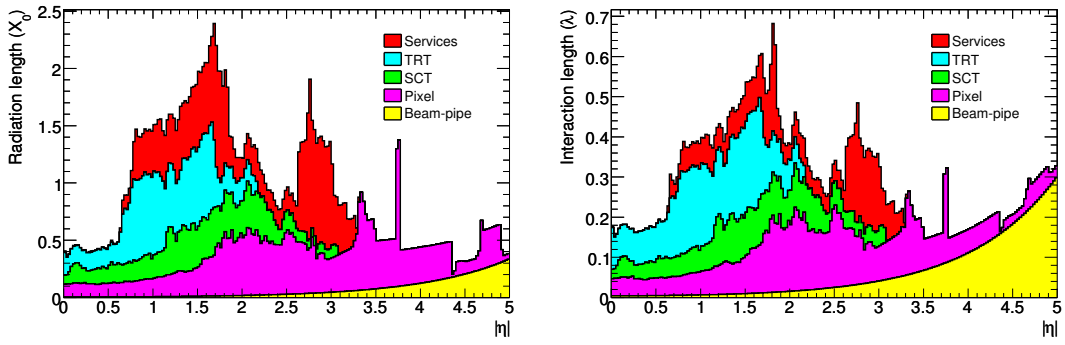

Figure 2.8 *The material description at the exit of the ID in units of* X_0 *and* λ *as a function of* $|\eta|$ *and averaged over* ϕ*. The material description is broken down into the individual sub-detectors and the services that support them [44].*

The inner detector provides the measurement of tracks with their direction, momentum and impact parameters for charged particles. The ID momentum resolution σ_{p_T} that was achieved during Run 1 can be approximated by the formula:

$$\frac{\sigma_{p_T}}{p_T} = 0.05\% \cdot p_T \oplus 1\% \tag{2.3}$$

2.2.3 Calorimeters

The ATLAS calorimeter system, shown in Figure 2.9, is placed after the inner detector and covers a pseudorapidity range up to $|\eta| \leq 5$. The calorimeter consists of three sub-detector systems which exploit different technologies due to the different functionality required from each sub-detector. The main purpose of the calorimeter is to measure the energy and direction of electrons, photons and hadrons. This is done by collecting the energy of the particles produced in the interaction of the incident particle with the material in the calorimeter. This

cascade of secondary particles is referred to as the *shower*. This process keeps repeating, producing more particles until the particles have low enough energy that they are absorbed by the material. The energy of these particles is measured by detecting light produced in a scintillating material or the charge produced by ionisation in a gas or a liquid.

Calorimeters can be separated into two different types. ATLAS uses a sampling calorimeter, where the *passive* material (lead, copper or iron in ATLAS) degrades the energy of the particle and the *active* material (Liquid Argon (LAr) or polystyrene scintillator in ATLAS) measures the signal. The other type of calorimeter is a homogeneous calorimeter where the the material can degrade and measure the energy of the particle simultaneously. This is used by CMS in their electromagnetic calorimeter, made from lead tungstate crystals.

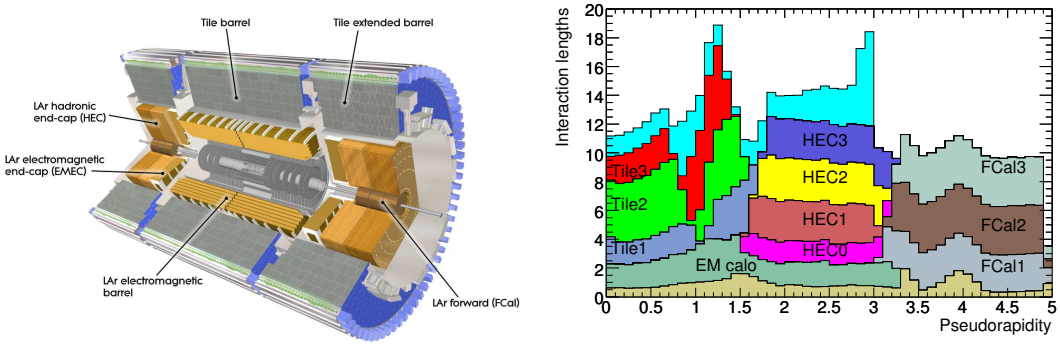

Figure 2.9 *Left: Cut-away view of the ATLAS calorimeter system Right: Cumulative amount of material, in units of hadronic interaction length, as a function of $|\eta|$ in front of the calorimeters, in the calorimeters themselves and after the calorimeters. Also shown for completeness is the amount of material in front of the first active layer of the muon spectrometer (light blue) [44].*

Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is a sampling calorimeter made from 2 mm layers of Liquid Argon (LAr) acting as the active material, interleaved with copper electrodes which collect the ionisation charge and with lead absorber plates that act as the passive material. The electrodes and absorbers are arranged in an accordion geometry in order to provide full coverage in ϕ.

The ECAL is divided into two separate systems, the electromagnetic barrel (EMB) in the central region ($|\eta| \leq 1.475$) and the electromagnetic end-cap

(EMEC) which covers the region ($1.375 \leq |\eta| \leq 3.2$). The LAr detectors require an operating temperature of around 88 K. The EMB shares the cryostat with the solenoid whereas the end-caps are hosted in their own cryostat.

The ECAL has three sampling layers up to $|\eta| \leq 2.5$ (where the ID extends to) and two layers in the region $2.5 \leq |\eta| \leq 3.2$. In the region $|\eta| \leq 1.8$, the sampling layers have a presampler placed in front of them. The presampler is a thin layer of liquid argon with no lead absorber in front of it and is used to correct for the energy loss that has occured either in the ID or the solenoid magnet. The first sampling layer is finely granulated in $|\eta|$. For $|\eta| \leq 1.4$, $\Delta\eta$ between the cells is 0.003 and $\Delta\phi$ is 0.01. This layer is used for a precision measurement of electromagnetic showers and can discriminate between electron / photon candidates and hadrons as well as resolving two nearby photons separately, such as those that occur in a $\pi^0 \rightarrow \gamma\gamma$ decay. It is not as valuable to have the same granularity in the ϕ direction since converted photons such as those from a decay of a Higgs boson, typically extend further in the ϕ direction and look similar to $\pi^0 \rightarrow \gamma\gamma$ decays. The granularity description of the EM calorimeter is summarised in Table 2.1 and shown in Figure 2.10.

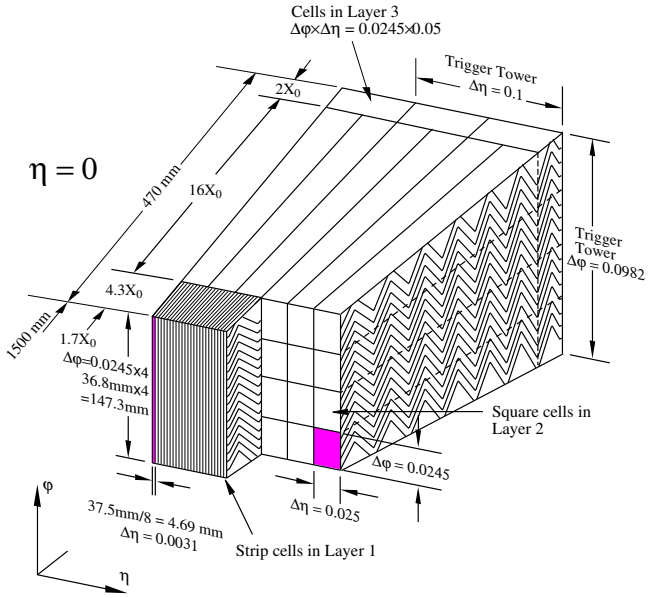

Figure 2.10 *Sketch of the barrel module of the ATLAS electromagnetic calorimeter [44].*

For $|\eta| \leq 1.4$, the second layer has a constant cell size, 0.025×0.025 which enables a precise measurement of the position and energy of the electron or photon candidate. This is the thickest part of the calorimeter (Figure 2.11) with up to 17 X_0. A final layer with cell granularity 0.1×0.1 completes the system and

Layer	Granularity ($\Delta\eta \times \Delta\phi$)
Presampler	0.025×0.1
1	0.003×0.1 for $\lvert\eta\rvert \leq 1.4$ 0.025×0.0025 for $1.4 \leq \lvert\eta\rvert \leq 1.475$ 0.003-0.025×0.01 for $1.475 \leq \lvert\eta\rvert \leq 2.5$ 0.1×0.1 for $2.5 \leq \lvert\eta\rvert \leq 3.2$
2	0.025×0.025 0.075×0.025 for $1.4 \leq \lvert\eta\rvert \leq 1.475$ 0.1×0.1 for $2.5 \leq \lvert\eta\rvert \leq 3.2$
3	0.050×0.025

Table 2.1 *Granularity of the Electromagnetic Calorimeter.*

ensures the electromagnetic showers are contained within the calorimeter system.

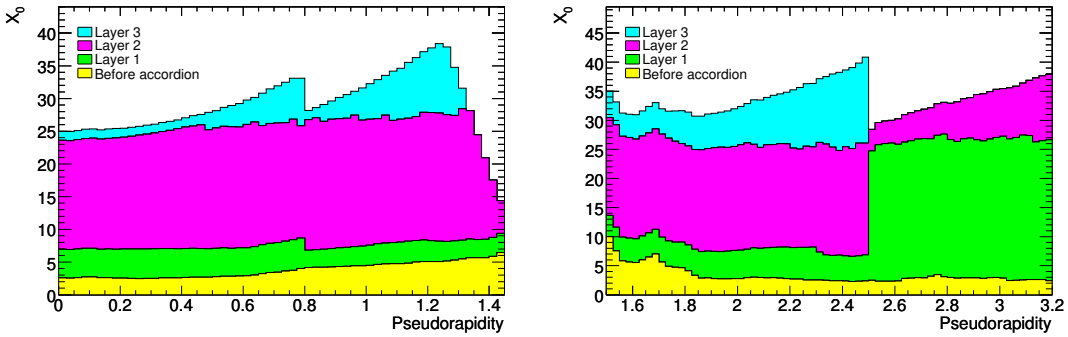

Figure 2.11 *Cumulative amounts of material, in units of radiation length X_0 as a function of $|\eta|$, in front of and in the electromagnetic calorimeters. Shown is the thickness of each layer as well as the amount of material in front of the accordion for the barrel (Left) and the end-cap (Right) [44].*

Hadronic calorimeter

Hadrons only leave a fraction of their energy in the ECAL systems so the hadronic calorimeter system (HCAL) surrounds the ECAL system in both the barrel and the end-cap. The HCAL relies on the strong interaction to measure the energy of the hadrons and to stop them reaching the muon system.

In the barrel region a steel-scintillator-sampling calorimeter (called TileCal) is used. It consists of three parts, the barrel which extends to $|\eta| \leq 1.0$ and two extended barrels which cover $0.8 \leq |\eta| \leq 1.7$. The steel absorber causes hadrons to

produce lower momentum hadrons, electrons and photons. These particles excite atoms within the scintillator material which produce scintillation light which is read out by photomultiplier tubes.

In the hadronic end-cap calorimeter (HEC), liquid argon is used as the active material and copper and tungsten as the passive material. It shares the same cryostat as the EMEC and covers the region $1.5 \leq |\eta| \leq 3.2$.

Forward calorimeter

The forward calorimeter (FCal) provides both electromagnetic and hadronic energy measurements in the region $3.1 \leq |\eta| \leq 4.9$. These measurements are important as they contribute to the measurement of missing transverse energy and allow the reconstruction of forward jets which are essential for VBF measurements. The detector must be extremely radiation tolerant in this region since the particle fluxes are much higher here. The FCal is constructed as a liquid argon sampling calorimeter, separated by copper absorbers in the first layer and tungsten absorbers for layers two and three. In order to completely contain the electromagnetic showers, the total amount of material used by the EM calorimeter exceeds 22 X_0 in the barrel and 24 X_0 in the end-cap (Figure 2.8). The HCAL contributes approximately 10 λ in the barrel and in the end-caps (Figure 2.11). This amount of material ensures good resolution for high energy jets and also stops particles reaching the muon spectrometer to below the irreducible level of muons from pion and kaon decays.

The energy resolution σ_E/E for a calorimeter is given by:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{2.4}$$

where a is the stochastic term, b includes the contribution from electronic noise and is usually negligible at the energy ranges studied by the ATLAS detector. The final term c is the constant term that depends on detector non-uniformity and electronic calibration uncertainties.

The design energy resolution for the ECAL is

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\% \tag{2.5}$$

where E is expressed in GeV. The energy resolution for hadronic jets (combined with the ECAL) for $|\eta| \leq 3.2$ is:

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \tag{2.6}$$

Finally, in the FCAL which covers the region $3.2 \leq |\eta| \leq 4.9$, the energy resolution is:

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\% \tag{2.7}$$

2.2.4 Muon Spectrometer

The only particles that are not absorbed by the calorimeters are muons and neutrinos, hence the muon spectrometer (Figure 2.12) is the outer most part of the ATLAS detector. The muon spectrometer consists of a magnet system of three air-core super conducting magnets that produce a magnetic field of 0.3 T and high-precision tracking chambers. Precise momentum measurements are made independently of the inner detector up to $|\eta| \leq 2.7$ and it is possible to trigger on muons up to $|\eta| \leq 2.4$. The magnetic field forces the muons to move on a helix in the $r - z$ plane where hits for track reconstruction are measured by two different systems. Three layers of monitored drift tube chambers (MDTs) which cover the central region up to $|\eta| \leq 2.7$. In the region $2.0 \leq |\eta| \leq 2.7$, the innermost layer of MDTs is replaced with Cathode Strip Chambers (CSCs) due to high radiation levels. Muons are particularly clean signatures to trigger on since almost all other particles are absorbed by the calorimeter system. In order to provide this trigger, the muon system is equipped with resistive plate chambers (RPCs) in the barrel and thin gap chambers (TGCs) in the end-cap. These provide a less precise measurement than that of the MDT and CSC but a faster measurement. The expected resolution of the muon system ranges from 3 to 11% in the p_{T} interval between 10 GeV and 1 TeV. At low p_{T}, the resolution is dominated by energy losses and multiple scattering while at high momentum, the MDT tube resolution dominates.

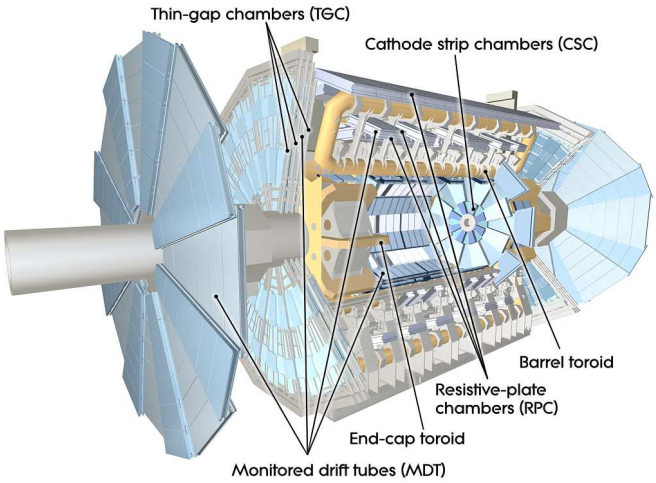

Figure 2.12 *Cut-away view of the ATLAS Muon system [44].*

2.2.5 Trigger system

One of the biggest challenges to the ATLAS detector is filtering out the events of interest given the huge amount of data produced. During Run 2 the ATLAS detector has been exposed to one bunch crossing every 25 ns. The trigger system must decrease the event rate from a nominal bunch crossing rate of 40 MHz to 100 kHz and then for storage down to 1 kHz, approximately a factor of 40,000. The ATLAS trigger system is composed of two levels which have access to different levels of detector information. A hardware based Level One (L1) trigger is followed by the software based High-Level Trigger (HLT).

Level One Trigger

The L1 Trigger is based on custom built electronics that have a processing time of 2.5 μs. It uses information from the calorimeters and the muon system (RPCs and TGCs) only. The calorimeter information is read out in the trigger towers which have separate electronics paths and correspond to a reduced granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. At this granularity, L1 objects are built which are electromagnetic clusters, hadronic decays of τ leptons, jets and missing transverse energy. The data corresponding to the regions of the detector where the L1 decision is taken serve as a *region of interest* (ROI) for the L2 Trigger. A maximum of 100 kHz is accepted by the L1 trigger and handed over to the HLT.

High Level Trigger

The HLT has access to the full granularity of the calorimeter, but only in the ROIs given by the L1 Trigger. The tracks from the inner detector are also available at this stage of the chain. The tracking and more detailed calorimeter information allow particle identification. The L1 trigger decision is refined and the rate is reduced to 1 kHz. If the event is accepted, a copy of the event is stored in the CERN computing centre and a second copy is copied to one of the ATLAS Tier One grid computing centres.

2.3 Definition of Physics Objects

Particles created from the proton collisions in the LHC leave different signatures in the ATLAS detector. These include hits in the inner detector and energy deposits in cells of the calorimeter which are used to reconstruct physics objects. A brief summary of the reconstruction algorithms are described below.

2.3.1 Tracks

A charged particle leaves a trace of hits as it traverses the inner detector. The task of track reconstruction is finding the collection of hits that have been created by a single particle and avoiding any other spurious hits. Including spurious hits will degrade the track resolution or could even lead to particle misidentification.

There exists a number of different techniques that can be applied in order to reconstruct tracks from hits. The technique used most often in ATLAS is the *inside-out* algorithm, which is described below.

Space-points are identified from hits in the silicon (Pixel and SCT) detectors. *Seeds* are then formed from the combination of three space points, the minimum number of measurements required to constrain a trajectory. The trajectory of a charged particle in a uniform magnetic field is a helix, which can be parameterised by a set of five track parameters. In the ATLAS experiment, the following parameterisation is used:

$$\mathbf{x} = (d_0, z_0, \phi, \theta, q/p) \tag{2.8}$$

where d_0 is the distance of closest approach of the helix to the z-axis, z_0 is the distance between the z-position of the track where d_0 has been defined and the primary vertex, ϕ and θ are the azimuthal and polar angles respectively and q/p is the inverse momentum multipled by the charge q.

In order to reduce the number of potential seeds, selections are applied on the momentum and impact parameters of the initial trajectory. The seeds that pass these requirements are used as an input to a combinatorial Kalman filter [49] that builds track candidates by including additional space-points in the remaining silicon layers that are compatible with the initial trajectory. The next step is called *ambiguity solving*, it aims to reject tracks which have been formed from spurious hits, referred to as *fakes*. The ambiguity solver uses a scoring function which applies positive scores for track fit quality and unique hit measurements and negative scores for missing hits and hits that are shared between more than one track. Tracks that pass the ambiguity solver are extended into the TRT where the increased track length improves the momentum resolution.

Track finding due to its combinatorial nature becomes computationally expensive. This will become particularly challenging during the high luminosity phase of the LHC. In light of this, CERN has recently launched a machine learning challenge for finding tracks on the Kaggle platform with prize money of $25,000 [50].

2.3.2 Vertices

Finding the trajectories of the particles is the only way to determine their production origin, also referred to as their vertex. Vertices along the beamline associated with the interactions of incoming protons are called primary vertices while vertices displaced from the beamline, caused by particle decays are called secondary vertices.

Vertex reconstruction associates the tracks with vertices (vertex finding) and thereby reconstructs the vertices (vertex fitting). ATLAS employs an approach called *finding through fitting* [51] which is briefly described below.

A vertex seed is selected from the mode in the distribution of z_0 (computed with respect to the beam spot center) of all reconstructed tracks. The vertex position and its uncertainty is determined using an *adaptive vertex fitting algorithm* [52] that takes the seed position and the tracks compatible with the seed position as

an input. Tracks that are incompatible with the vertex by more than $7\,\sigma$ are discarded and the above procedure is repeated with those tracks available again.

2.3.3 Muons

Muons are the only charged particles not absorbed by the calorimeters. As a result, muons leave tracks not only in the inner detector but in the muon spectrometer (MS) as well. The information from the different sub-detectors can be combined to reconstruct muons that can be used in physics analyses. Four different types of muons can be reconstructed:

- Combined (CB) muons: Tracks are reconstructed independently in the ID and the MS and then a combined track is formed with a global fit that uses the hits from both the inner detector and the MS sub-detectors. CB muons have the highest reconstruction and identification efficiency of all types of muons.
- Segment-tagged (ST) muons: A track in the ID can be classified as a muon if once extrapolated to the MS, it is associated with at least one local track segment in the MDT or CSC chambers.
- Calorimeter-tagged (CT) muons: Muons can be reconstructed from the association of track in the ID to an energy deposit in the calorimeter compatible with a minimum-ionising particle. This type of muon has the lowest reconstruction efficiency but helps recover acceptance in the region where there is no MS coverage.
- Extrapolated (ME) muons: Muons can be reconstructed based only on a track in the MS and a loose requirement on that track being compatible with originating from the interaction point. The track is defined to start at the interaction point, taking into account the effect of multiple scattering and energy loss in the material. The ME muons can extend the acceptance for muons into the region $2.5 \leq |\eta| \leq 2.7$.

2.3.4 Topological clusters

Jets, electrons and photons are all reconstructed from *topological clusters* in the calorimeter.

Topological clusters are clusters seeded by cells which have more than four times the energy of the noise threshold of that cell. The clusters are then expanded by adding neighbouring cells if they have an energy more than two times above the noise threshold. A final layer of neighbouring cells is added to the cluster providing the energy in each cell is greater than the noise level. This is known as *4-2-0* topological cluster reconstruction.

2.3.5 Electrons and Photons

Electron and photon reconstruction is based on tracks reconstructed in the ID and on topological clusters measured in the calorimeter. The reconstruction is designed to separate electrons and photon candidates and to classify the latter as unconverted or converted.

If the topological clusters measured in the calorimeters pass loose requirements that are consistent with electrons and photon energy deposits, a search begins for tracks in the inner detector to begin electron reconstruction. Tracks with silicon hits are considered to be matched to electrons if they are nearby in ϕ (up to 0.2) and η (up to 0.05). If there are multiple tracks, tracks with pixel hits are preferred and if there is a number of these, the track closest in ΔR to the cluster is taken as the best matched track. Clusters that do not have any matching tracks with silicon hits are considered as photon candidates and are not used to reconstruct electrons.

For central ($|\eta| \leq 2.5$) electrons, the four-momentum is computed using the cluster energy measured in the calorimeter while the η and ϕ directions are taken from the best matched track.

For photons that have converted into an e^+e^- pair, a search is performed for two tracks that form a vertex consistent with that of a massless particle. The track pairs can be classified into three categories (Si-Si / Si-TRT / TRT-TRT) depending on whether each track has hits in the silicon or TRT detector. Single track conversions are also permitted in case one of the tracks has failed to be reconstructed.

There is some ambiguity between electron and photon reconstruction. This is resolved somewhat in the following way:

- Photon - The candidate particle is reconstructed as a photon if no tracks

with at least four hits in the silicon detector are matched to the cluster **or** if a Si-Si conversion vertex is found and the tracks associated to it have no hits in the pixel detector.

- Electron - The candidate particle is reconstructed as an electron if no conversion vertex is found and the track has at least two hits in the pixel detector and at least four hits in the silicon detector. If the best matched track to the electron is part of a conversion vertex, it is still reconstructed as an electron unless it has a Si-Si vertex.

- Both - The candidate particle is reconstructed as both an electron and a photon if it does not fall into either of the above categories. Each physics analysis can then apply their own criteria to separate the electrons and photons further.

Superclusters

Only recently, in ATLAS, have topological clusters been used to reconstruct electron and photon candidates, previously a *sliding window algorithm* [53] was used which had the disadvantage of the clusters being fixed in size. The topological cluster reconstruction is improved further by grouping multiple clusters together into so called *superclusters* which are described below.

The topological clusters now act as seeds and are sorted according to descending p_{T}. Starting from the highest p_{T} seed cluster, all clusters in a window $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ are added. For electrons only, an additional search is performed in a window $\Delta\eta \times \Delta\phi = 0.125 \times 0.3$ and, again, the cluster is added if it shares the same track as the seed cluster. For converted photons only, a cluster is added if it is matched to a track that belongs to the conversion vertex of the seed cluster or it has the same conversion vertex as the seed cluster. This algorithm is shown in Figure 2.13.

This supercluster approach improves the measurement of electron and photon energies, in particular for cases where the electron has radiated a photon or the photon has converted. Figure 2.14 shows the invariant mass of a Higgs boson decaying to four electrons and the invariant mass of a Higgs boson decaying to two photons which have converted. There is an improvement in terms of the ratio of the resolution to the mean of ~ 5% for four electrons and ~ 9% for two converted photons compared to the sliding window algorithm previously used by

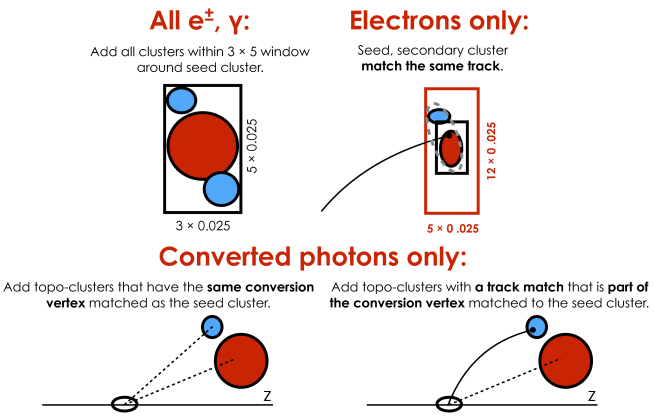

Figure 2.13 *Diagram of the supercluster algorithm for electrons and photons [54].*

ATLAS.

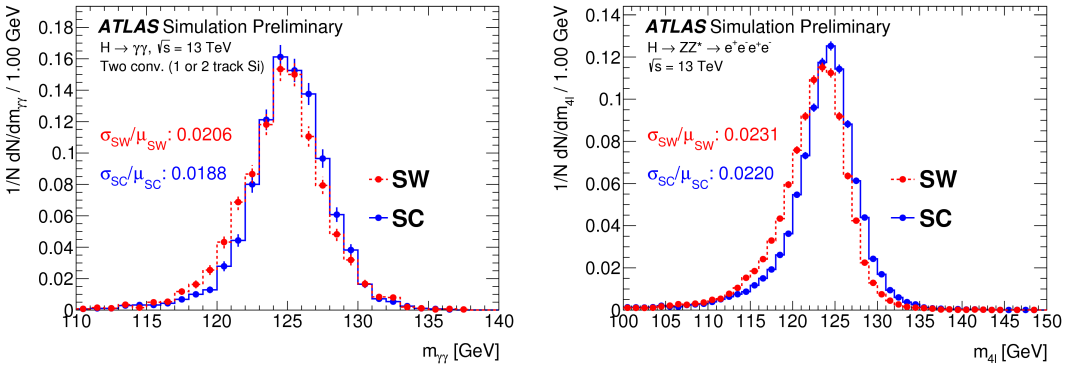

Figure 2.14 *Invariant mass distributions of $H \to ZZ^* \to 4e$ and $H \to \gamma\gamma$ (both photons converted) using both the supercluster (SC) and sliding window (SW) algorithms [54]. The resolution and mean of each distribution are determined from a fit using a double-sided Crystal Ball function.*

2.3.6 Jets

Quarks and gluons hadronise immediately after production and produce a spray of hadrons which are referred to as jets. All jets used in the physics analyses presented in this thesis are reconstructed from topological clusters of energy deposits in the calorimeter cells using the anti-k_t algorithm [55] with a radius parameter of 0.4.

The energy of the jets is corrected for pile-up using a subtraction method which is applied on an event-by-event basis [56]. The jets are calibrated using a

combination of simulation based and data-driven correction factors [57].

Jets originating from pile-up with $|\eta| \leq 2.4$ are suppressed using the jet vertex tagger (JVT) [58] which combines tracking information into a multivariate likelihood. A selection on the JVT is used such that hard-scatter jets with 20 GeV $< p_{\mathrm{T}} <$ 60 GeV and $|\eta| < 2.4$ have an efficiency of approximately 92%.

Flavour tagging

Determining whether a jet originated from a specific quark flavour or gluon is of vital importance in many physics analyses, for example in processes like $H \to b\bar{b}$, $t\bar{t}H$ or in $HH \to b\bar{b}\gamma\gamma$. Correctly identifying ("tagging") jets originating from bottom quarks is essential in measuring the couplings of the Higgs boson.

When bottom quarks are produced, the hadronisation process creates b-hadrons. The b-hadrons typically have longer lifetimes than the other hadrons, for example the lifetime of $\bar{B}^0(b\bar{d})$ is approximately 1.5×10^{-12} s. This results in a measurable distance (3 mm) between the primary and secondary vertex. The relatively large mass of the b-quark can also result in the decay products being produced at large angles to the original b-quark direction. These features of b-jets are exploited in a multivariable discriminant to distinguish b-jets from other types of jets.

Identifying jets other than b-jets with the ATLAS detector is very challenging and large misidentification rates are common. There have been recent developments in charm-tagging algorithms which have been used in the search for $ZH(H \to c\bar{c})$. Figure 2.15 shows the c-tagging efficiency as a function of b-jet and light jet rejection. For a c-tagging efficiency of 41%, rejection factors of 4 and 20 are obtained for b-jets and light jets. This can be compared to b-tagging algorithms, where for a 70% b-tagging efficiency, rejection factors of 8 and 313 for c-jets and light jets can be achieved [59]. This difference arises because c-hadrons have shorter lifetimes and decay to fewer charged particles than b-hadrons.

Quark versus gluon jet tagging is another area of active development. On average, jets originating from gluons have more constituent particles than those originating from quarks. The number of tracks inside the jet can therefore be exploited to discriminate between the two types of jet. Figure 2.15 shows the track multiplicity, n_{track} in the highest p_{T} jet in dijet Monte Carlo events for different p_{T} intervals. For a 50% quark tag efficiency, the mistag rate for gluon jets is approximately 10% for $p_{\mathrm{T}} = 50$ GeV [60]. Quark gluon tagging could be

useful in VBF measurements, however typically jets in VBF production are in the forward direction ($|\eta| \geq 2.5$) while the ATLAS inner detector only extends to $|\eta| \leq 2.5$. ATLAS plans to extend the coverage of its inner detector to ($|\eta| = 4.0$) for the HL-LHC program.

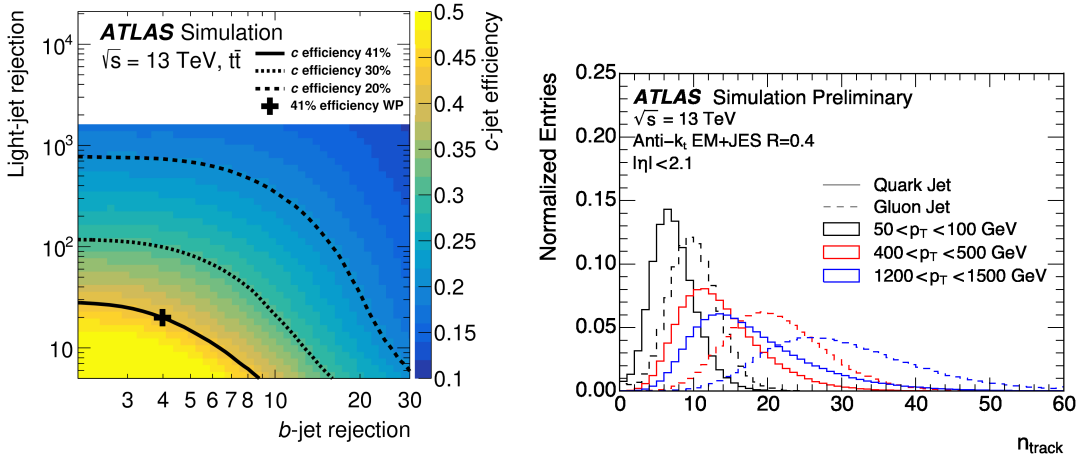

Figure 2.15 *Left: The c-jet tagging efficiency (z-axis) as a function of the b-jet and light-jet rejection as obtained from simulated $t\bar{t}$ events. The cross denotes the 41% c-tagging efficiency used in the $ZH(H \to c\bar{c})$ analysis [32] Right: Distribution of the track multiplicity n_{track} in different p_T ranges in dijet Monte Carlo events [60].*

Chapter 3

The HL-LHC and the ITk

In this chapter the motivation for the High Luminosity LHC program will be presented and my contributions towards the development of the new ATLAS inner tracker, the ITk. Many of the results shown in this chapter were included in the Technical Design Report for the ATLAS Inner Tracker Strip Detector [61].

3.1 HL-LHC

In the near future the LHC will undergo a number of upgrades in order to increase the instantaneous luminosity. Run 2 at the LHC finished with ATLAS recording a data set of $148.5\,\text{fb}^{-1}$. The LHC plans to collide protons at the LHC design energy ($\sqrt{s} = 14$ TeV) for the first time in the years 2021 - 2023, a period of data taking referred to as *Run 3* where the LHC expects to run at twice the design luminosity and deliver $300\,\text{fb}^{-1}$. After Run 3, the LHC will be shutdown for a further 3 years in order to upgrade it to the High Luminosity LHC (HL-LHC). In this final period, the LHC plans to run at 7 - 10 times the design luminosity and aims to deliver $3000\,\text{fb}^{-1}$. The pile-up, $\langle\mu\rangle$, during this period is estimated to be 200, extremely challenging conditions for the ATLAS detector.

3.2 Physics motivation

At present, the biggest physics motivation for the HL-LHC is a precision Higgs physics program. In the Standard Model, once the mass of the Higgs boson is known, all properties of the particle can be predicted. Precision measurements of the properties of the Higgs boson therefore open a window into new physics. Figure 3.1 shows the expected precision on the main Higgs boson production cross sections and the expected precision on the main branching ratios of the Higgs boson. These results assume a scenario where the current experimental and theoretical uncertainties are reduced by a factor of two. The large HL-LHC data set will enable measurements with a precision of a few percent on the main Higgs production cross sections and branching ratios. It will also be possible to observe the rare $H \to Z\gamma$ and $H \to \mu^+\mu^-$ decays at the HL-LHC assuming the Standard Model decay rates.

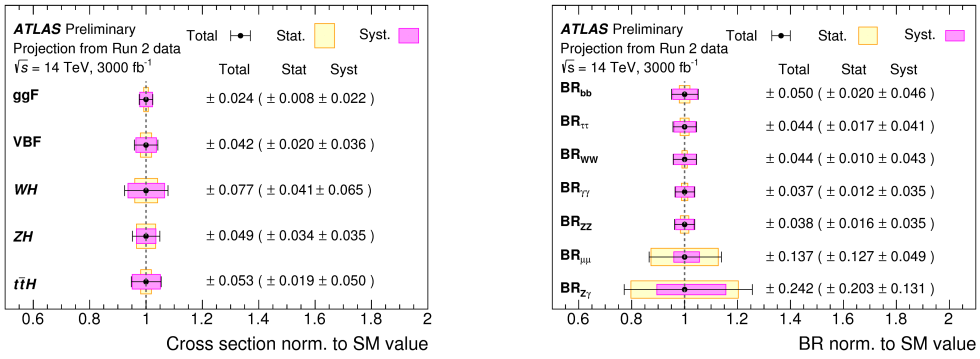

Figure 3.1 *The expected precision on the cross sections for the main Higgs boson productions modes (left) and the main Higgs boson branching ratios (right), normalised to their SM predictions [62].*

Another strong motivation for the HL-LHC program is measuring the Higgs boson self-coupling which can be measured directly by studying Higgs boson pair production. The prospects for this have been studied in detail by the ATLAS and CMS collaborations. At present, the most promising final states appear to be the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau\tau$ and $b\bar{b}b\bar{b}$ final states. The ATLAS collaboration combined the prospects for these final states assuming $3000\,\text{fb}^{-1}$ and obtained an expected significance of $3.0\,\sigma$ [63] while the Higgs boson self-coupling relative to its SM value ($\kappa_\lambda = \lambda_{HHH}/\lambda_{SM}$) is expected to be constrained to $[-0.4, 7.3]$ at the 95% CL. The ATLAS and CMS collaborations have also combined their prospective

analyses [64]. An expected significance of $4.0\,\sigma$ is obtained, with κ_λ expected to be constrained to [0.1, 2.3] at the 95% CL and [0.5, 1.5] at the 68% CL. Figure 3.2 shows the likelihood scan as a function of κ_λ for the ATLAS only results and the ATLAS and CMS results combined.

The prospects for observing $t\bar{t}HH$ production at the HL-LHC is presented in Chapter 4 and the search for HH in the $b\bar{b}\gamma\gamma$ final state with the Run 2 data set is presented in Chapter 5.

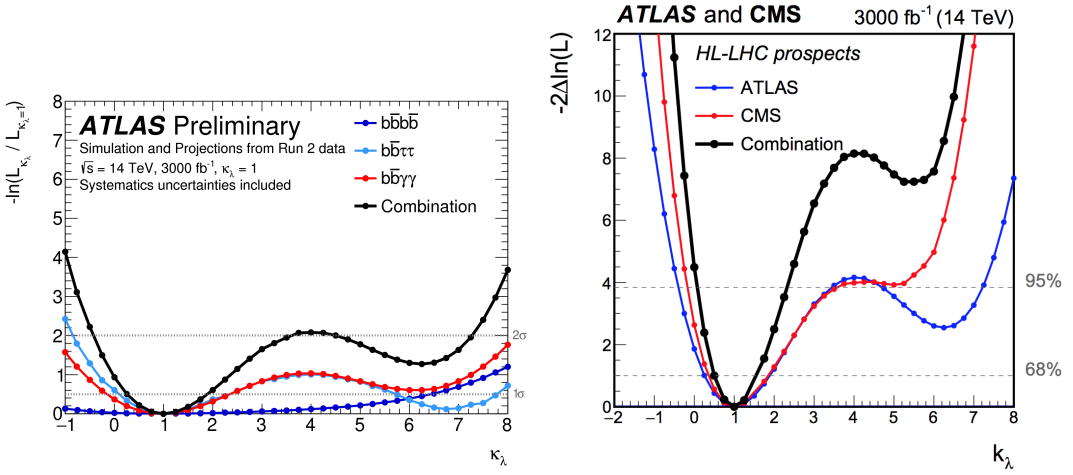

Figure 3.2 *The likelihood scan as a function of κ_λ for ATLAS (left) [63] and for ATLAS and CMS combined (right) [64]. The dashed lines correspond to the $1\,\sigma$ and $2\,\sigma$ confidence intervals.*

3.3 Upgraded LHC

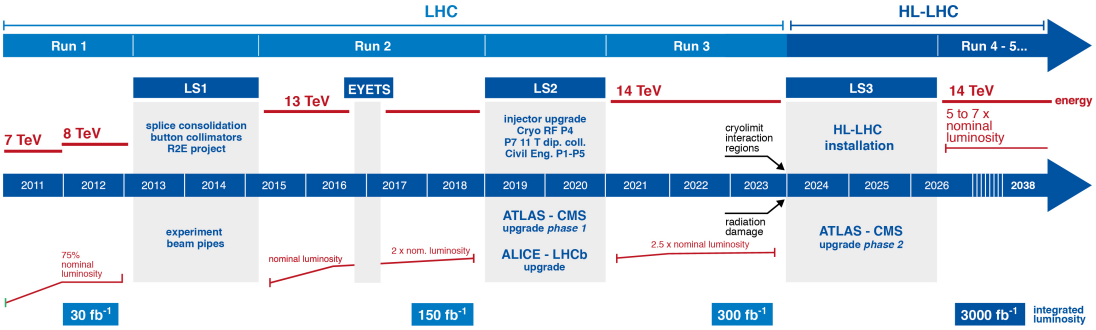

Figure 3.3 *Timeline of the LHC with planned shutdowns, projected instantaneous and integrated luminosity levels, and expected pile-up conditions [65].*

The timeline of the LHC and its upgrade to the HL-LHC is shown in Figure 3.3. Two long shutdown periods (LS) are planned to upgrade the LHC and the

detectors before the HL-LHC program begins. During this period, a number of upgrades are planned to the injector systems of the LHC in order to obtain higher instantaneous luminosities. For example LINAC2 which currently accelerates the protons to 50 MeV before they reach the circular booster will be replaced with LINAC4 which will accelerate the protons to 160 MeV. In LS3, the HL-LHC will be installed. This will include replacing some of the Niobium-Titanium dipole magnets with Niobium-Tin magnets which can produce stronger magnetic fields and installing new collimators to mitigate radiation from the proton beams. Crab cavities will also be installed. These help increase the luminosity (approximately an increase of 15% is expected) without increasing the pile-up, at present there is a small crossing angle (285 μrad) at the LHC. The crab cavities tilt the proton bunches off-axis which allow for a larger geometrical overlap, see Figure 3.4.

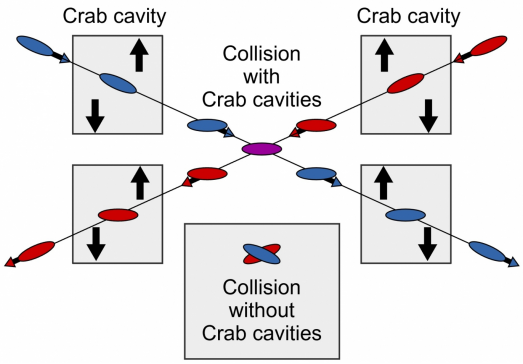

Figure 3.4 *An illustration of the effect of crab cavities on the proton bunches [66].*

3.4 Upgraded ATLAS

The harsher radiation environment and increased pile-up require that the ATLAS detector is upgraded in order to maintain the current performance of the ATLAS detector. In the years 2016 - 2018, the ATLAS collaboration published six Technical Design Reports (TDRs). These TDRs are for the upgraded Strip detector [61], Pixel detector [67], Liquid Argon calorimeter [68], Muon spectrometer [69], Tile calorimeter [70] and TDAQ system [71]. ATLAS has also published a technical proposal for a high granularity timing detector (HGTD) [72].

This is a huge commitment by the ATLAS collaboration to ensure that the ATLAS detector can perform well in HL-LHC conditions.

3.4.1 New Muon Small Wheel

The Monitored Drift Tubes (MDT) chambers currently installed in the Muon Spectrometer can operate without problems during the HL-LHC period with the exception of the inner layer in the end-cap. The ATLAS collaboration therefore plans to replace the inner end-cap of the Muon Spectrometer with two new small wheels (NSWs). The NSWs consist of eight layers of Micromegas and small-strip Thin Gap Chambers (sTGC). The sTGC will act as the primary trigger and the Micromegas will make precision measurements. The NSWs were planned to be installed during LS2, however the project has been delayed significantly and at the time of writing, it is unclear whether 0, 1 or 2 NSWs will be installed during LS2.

3.4.2 Fast Tracker

As discussed in Chapter 2, the first ATLAS trigger system (L1) has access to coarse calorimeter information and the muon system. The data corresponding to the regions of the detector where the L1 decision is taken (ROIs) is then passed to the L2 Trigger. The L2 Trigger has access to the tracks from the inner detector but only in the ROIs. The purpose of the Fast Tracker (FTK) is to give the L2 Trigger access to all the tracks from the inner detector. The FTK receives hits from all the layers of the inner detector and groups nearby hits into clusters. It then compares the cluster pattern to millions of pre-calculated track patterns from simulation. At the time of writing, the FTK is significantly behind schedule and the physics motivation for the project is currently in question.

3.4.3 High Granularity Timing Detector

ATLAS has published a technical proposal for a High Granularity Timing Detector (HGTD) [72]. The challenge for tracking detectors is to reconstruct tracks and correctly assign them to their production vertices. However due to the large amount of pile-up activity, it is impossible to unambigously assign tracks to vertices. A powerful new method to address this is to exploit the spread in time of different interactions during a single bunch crossing. This is done by measuring the time of individual tracks with a precision smaller than the spread of the interaction times. The time spread of interactions in a bunch crossing at

the HL-LHC has an RMS of ~ 180 ps. The HGTD has a timing resolution of 30 ps and will therefore help to mitigate the impact of pile-up.

The HGTD plans to cover the pseudorapidity region $2.4 \leq |\eta| \leq 4.0$ in order to improve the performance in the forward region since the performance of the upgraded inner tracker deteriorates here.

3.4.4 Calorimeter and Trigger Upgrades

The ATLAS calorimeter system does not need replaced for HL-LHC running. However the readout electronics for both the LAr and Tile calorimeters need to be replaced because of radiation tolerance limits and their inability to be able to operate with the increased trigger rates at the HL-LHC.

The current ATLAS TDAQ system will also be replaced for the HL-LHC. At present the L1 trigger accepts an event rate of 100 kHz followed by the L2 trigger that reduces this to 1 kHz. The upgraded TDAQ system will be a single-level Level 0 (L0) hardware trigger that will accept an event rate of 1 MHz. This single-level L0 hardware trigger will have access to the full granularity of the calorimeter, rather than the reduced granularity the L1 trigger has at present and will continue to have access to the same information from the muon system the L1 trigger has at present. After this the event filter will use hardware-based tracking systems for track finding. The trigger decision is refined based on the tracking information to reduce the event rate down to 400 kHz.

3.5 ITk

The current ATLAS inner detector will not be able to withstand the conditions of the HL-LHC. The TRT will reach 100% occupancy, and, as a result, the tracking efficiency will decrease. The Pixel and SCT detectors would also have significant radiation damage that would limit their performance. The current inner detector also only extends to an $|\eta|$ of 2.5. There is more physics potential for the ATLAS detector if this is increased. For example, the acceptance will increase for a Higgs boson decaying into four charged leptons and it will also help reject jets which originate from pile-up in the forward direction, extremely important for VBF measurements. The ATLAS collaboration has therefore decided to completely

rebuild the inner detector. The TRT and the current silicon detector is removed in favour of a new all-silicon detector, a project known as the inner tracker (ITk).

There were two different proposed layouts for the ITk. The layouts (Inclined and Extended) both have coverage up to an $|\eta|$ of 4.0. In the inclined layout, all pixel barrel layers have a section where the sensors are inclined with respect to the z-axis (Figure 3.5) while in the extended layout, there is no inclined sensors. For all designs, the innermost 2 pixel layers are designed to be insertable as it may be necessary to replace them to ensure high performance for the entire HL-LHC data set.

At the beginning of 2017, the ATLAS collaboration decided to pursue an ITk layout which resembled that of the inclined layout. The rest of this chapter documents studies which were used as an input to this decision.

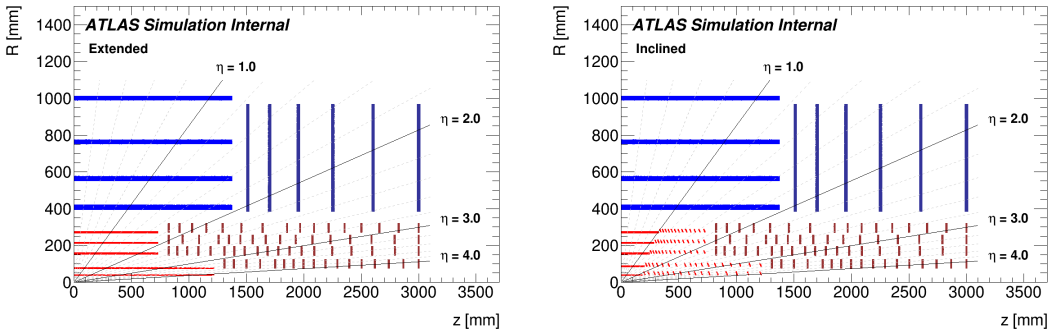

Figure 3.5 *The Extended (left) and Inclined ITk layouts (right) in the $R-z$ plane [61].*

3.6 Occupancy studies

The ITk layout is constantly evolving with feedback from performance studies and engineering. The ITk layout used in the simulation at the time these studies were performed is detailed in Appendix A.

The terms channel occupancy, hit density and cluster density are poorly defined in the literature. In the studies presented here, the following definitions are used:

$$\text{Channel Occupancy} = \frac{1}{\text{N}_{\text{Events}}} \frac{\text{Number of pixels [strips] hit in a detector section}}{\text{Total number of pixels [strips] in a detector section}} \tag{3.1}$$

$$\text{Hit Density} = \frac{1}{\text{N}_{\text{Events}}} \frac{\text{Number of pixels [strips] hit in a detector section}}{\text{Area of pixels [strips] in a detector section}} \tag{3.2}$$

$$\text{Cluster Density} = \frac{1}{\text{N}_{\text{Events}}} \frac{\text{Number of pixel [strip] clusters in a detector section}}{\text{Area of pixels [strips] in a detector section}} \tag{3.3}$$

In order to reduce the number of tracks that are formed from spurious hits, the channel occupancy must be kept low. From the current understanding of the Pixel and SCT detectors, a channel occupancy of less than 0.1% in the pixel layers and less than 1% in the strip layers leads to an efficient and stable pattern recognition suitable for tracking.

The channel occupancy, hit density and cluster density of the ITk will vary from event to event, depending on the activity of a particular event. These studies presented show results for these quantities in the ITk strip and pixel detectors for both the inclined and extended layouts. Simulated minimum bias events (inelastic pp collisions) with $\langle\mu\rangle = 200$ are used. The results are averaged over 1000 events and averaged over ϕ.

In the following, it is assumed that the detector is read out in the so called *01X* mode, as is currently the case for Run 2. This means that a hit (1) is only counted if there is no hit in that same channel in the previous bunch crossing (0), and regardless of whether or not there is a hit in the subsequent bunch crossing (X).

A detailed specification of the ITk detector is given in Appendix A.

3.7 The ITk Strip Detector

The building blocks of the strip barrel and end-caps are called staves and petals, shown in Figure 3.6. Each barrel stave houses 28 modules. In the first two layers, each module has four rows of short (24.10 mm) strips and in the outer two layers, each module has two rows of longer strips (48.20 mm). Each row has 1280 strips. The decision to use short or long strips is one of the most important design decisions in the strip detector, a compromise between detector performance and cost.

In each end-cap, there are six disks and each disk is populated with 32 petals with each petal hosting six modules. The strip lengths vary in the modules, increasing in size moving away from the beam axis (details in Appendix A). The shorter strips are used close to the beam axis in order to minimise the channel occupancy.

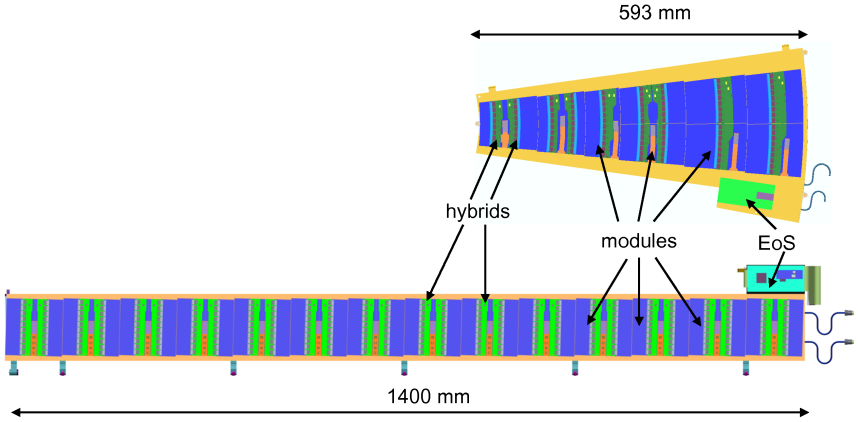

Figure 3.6 *End-cap petal (top) and barrel stave (bottom) components [61].*

Strip Results

Each data point in the results presented represents a module as discussed above, with the result averaged over ϕ.

The average cluster size (number of strips hit) increases as a function of z in the barrel and as a function of R in the end-caps. The cluster size in the barrel is typically 2.3 - 2.8 and 1.5 - 2.1 in the end-caps.

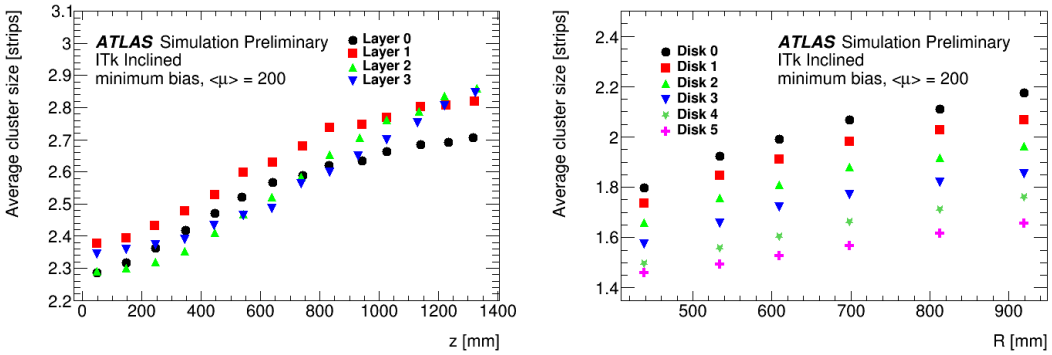

Figure 3.7 *The average cluster size as a function of z in the barrel (Left) and as a function of R in the end-cap (Right).*

Figure 3.8 shows the average hit density in the barrel and in the end-cap. The hit density is almost constant as a function of z in the barrel while in the end-

cap, the hit density is highest near the the centre of the end-caps and falls off quadratically with R. The maximum hit density for the strip detector in the barrel is 0.005 hits / mm^2 in the innermost layer and 0.0058 hits / mm^2 in the end-cap.

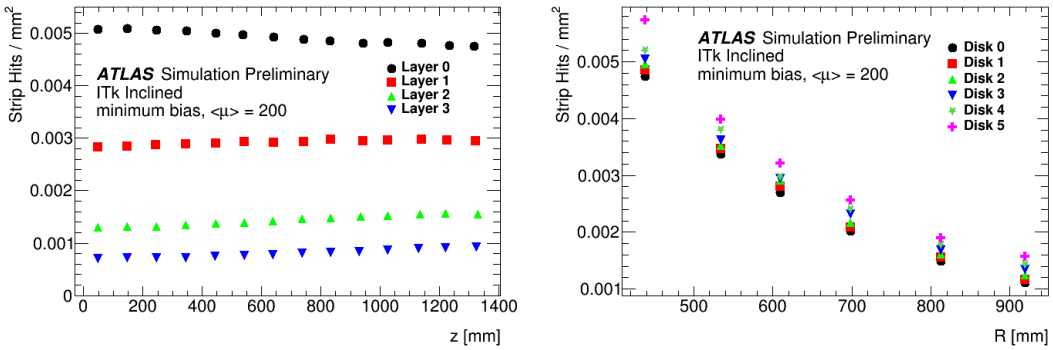

Figure 3.8 *The average hit density as a function of z in the barrel (Left) and as a function of R in the end-cap (Right).*

The channel occupancy is below 1% in all layers in the barrel. In the end-cap disks the channel occupancy is typically 0.6 - 0.8% except for nearest the centre where it is up to 1.2%, slightly above the desired 1%.

The channel occupancy is similar for layers 1 and 2 despite layer 1 being closer to the interaction point. This is a result of the use of short strips in layer 1 and longer strips in layer 2. These results motivate the use of short strips in the first two layers, otherwise the channel occupancy would exceed the desired 1% and it is clear that short strips are not necessary in layers 2 and 3 since long strips are used in simulation at present and the channel occupancy is significantly less than 1%.

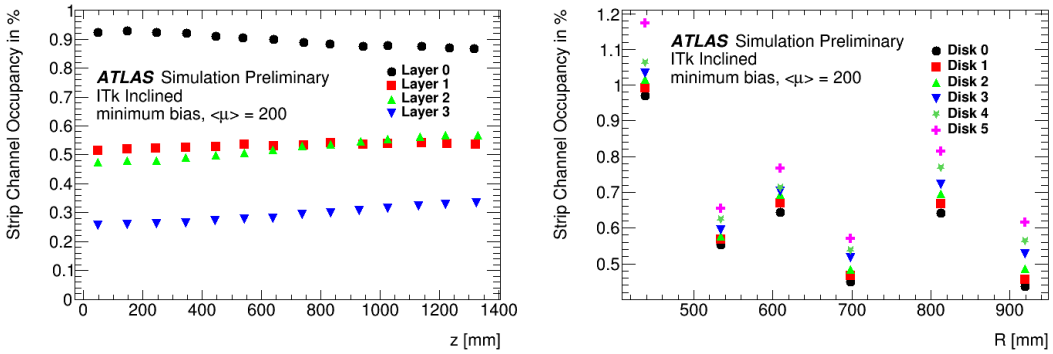

Figure 3.9 *The average channel occupancy as a function of z in the barrel (Left) and as a function of R in the end-cap (Right).*

3.8 The ITk Pixel Detector

Pixel Barrel

The pixel size for the ITk is chosen to be 50×50 μm^2 at present, in comparison to the size in the current pixel detector of 50×400 μm^2. The chip is 20.0 mm by 16.8 mm and has 336×400 pixels in it.

In the extended layout, there are two chips per sensor in the innermost layer and four chips per sensor in the other layers. In the inclined layout, the flat section of the pixel barrel is identical to the extended layout while in the inclined sections, there is only one chip per sensor in the innermost layer and only two chips per sensor in the other layers.

The pixel end-caps are composed of rings. The rings are composed of sensors with four chips per sensor. A detailed description of both the pixel barrels and the pixel ring systems are given in Appendix A.

Pixel Results

Each data point in the results presented represent a sensor as defined above, with the result averaged over ϕ. In the strip results, the barrel results were shown beside the end-cap results for the inclined layout. The layout for the strip detector is identical in both layouts so results were not repeated. In this section the results for the pixel detector will be compared between the two layouts.

For a charged particle incident on a sensor, the number of pixels hit increases as the track incidence angle increases. In the extended pixel barrel, the average cluster size (number of pixels hit) shown in (Figure 3.10) reaches more than 20 pixels for $|z| \geq 600$ mm. In the inclined pixel barrel, the average cluster size increases as a function of $|z|$ until the sensors are inclined, then the average cluster size drops to approximately 2.0 - 4.0.

The average channel occupancy (Figure 3.11) increases as a function of $|z|$ in the extended layout due to the larger cluster sizes. In the inclined layout, it increases as a function of $|z|$ until the sensors are inclined where it drops then increases as a function of $|z|$ again.

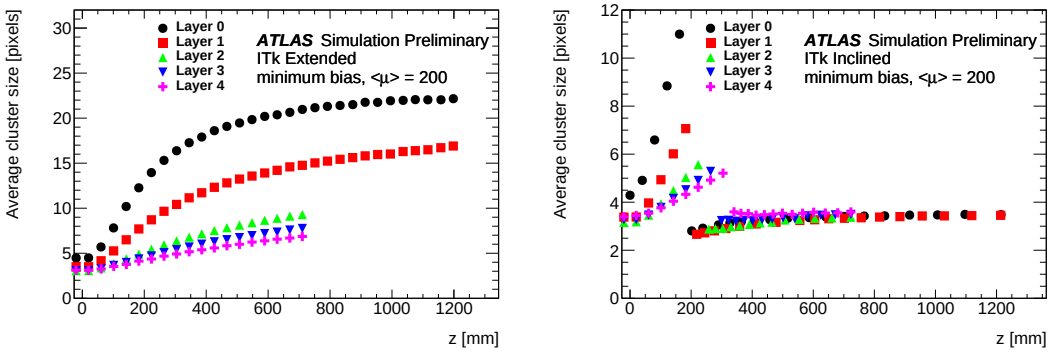

Figure 3.10 *The average cluster size as a function of z in the extended barrel (Left) and in the inclined barrel (Right).*

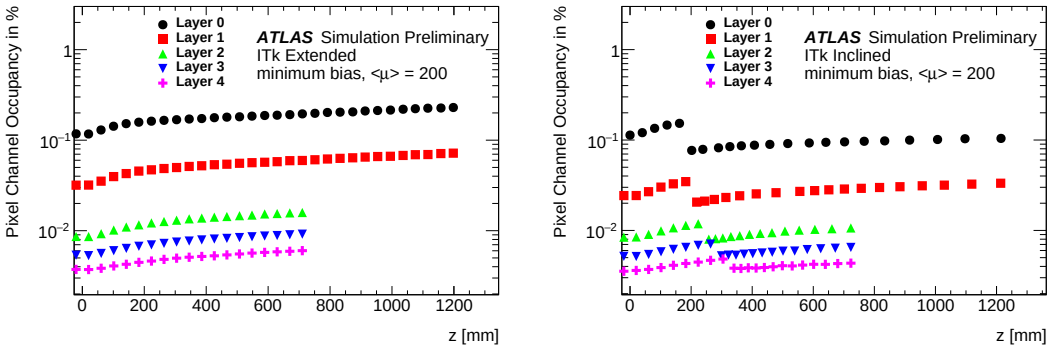

Figure 3.11 *The average channel occupancy as a function of z in the extended barrel (Left) and in the inclined barrel (Right).*

In the extended layout, the maximum cluster density (Figure 3.12) is approximately 0.1 clusters / mm^2 in the centre of the innermost pixel barrel layer. The cluster density decreases with z since in the forward direction, the particles are traversing less sensors and creating larger clusters as shown in Figure 3.16. In the inclined layout, the results are the same in the flat section of the barrel while for the inclined section, the cluster density increases as a function of z as a result of the incident particles creating smaller clusters in more sensors.

The channel occupancy and cluster density results suggest that different readout strategies would be optimal for each layout. The extended layout would benefit from a cluster based on-chip data compression scheme due to its smaller cluster density while the inclined layout could read out the individual pixel hits as normal.

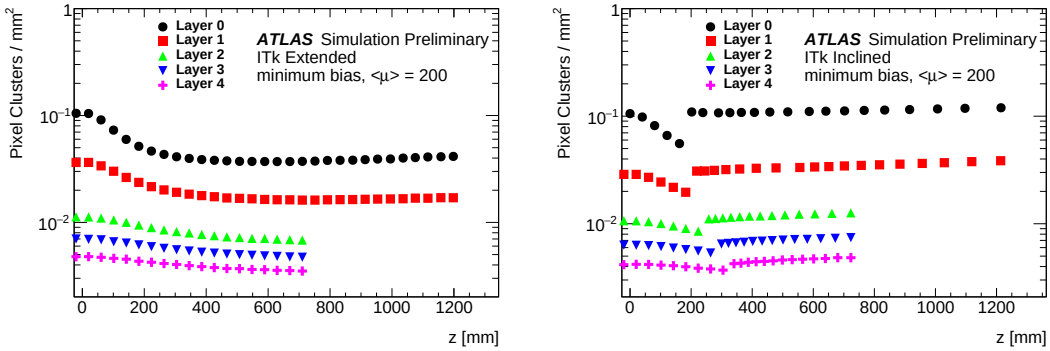

Figure 3.12 *The average cluster density as a function of z in the extended barrel (Left) and in the inclined barrel (Right).*

The pixel rings are the same in both the inclined and extended layouts, so only small differences are expected here. Results are shown for both for completeness.

The average cluster size is typically 2.2 - 4.0 in the pixel rings.

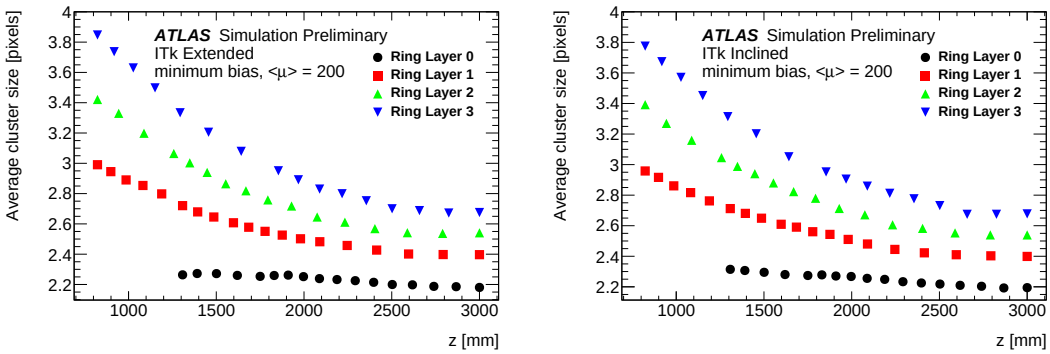

Figure 3.13 *The average cluster size as a function of z in the pixel rings in the extended layout (Left) and in the inclined layout (Right).*

The channel occupancy and cluster density in the pixel rings are shown in Figure 3.14. Both of these quantities are slightly higher in the extended layout due to the higher number of secondary particles as a result of the additional material in the extended layout.

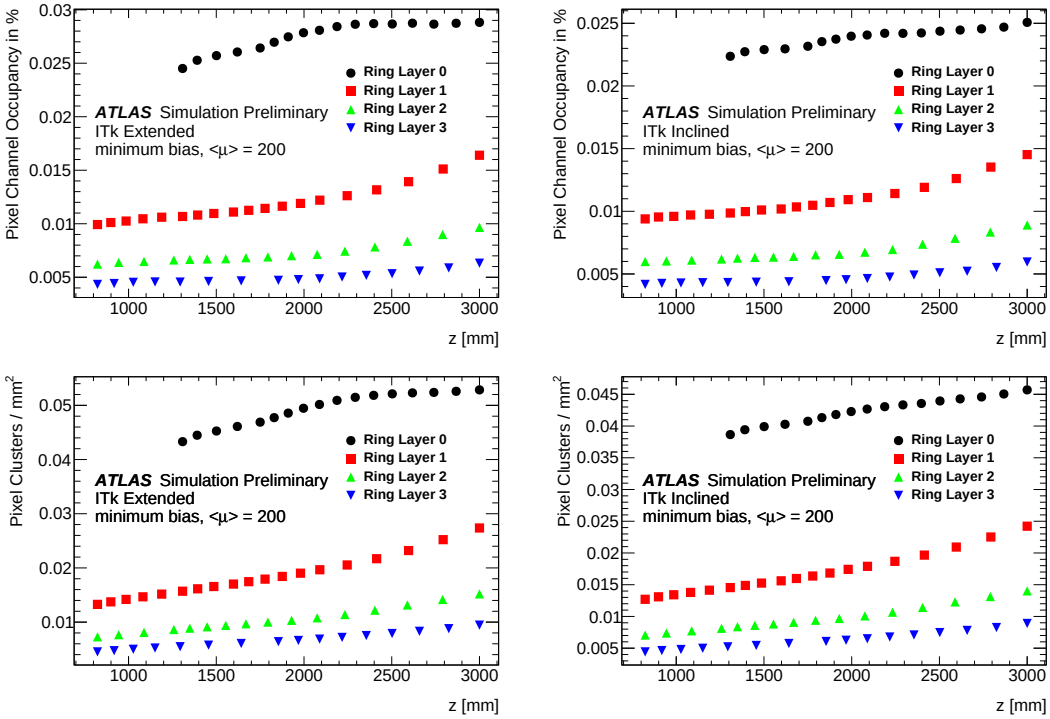

Figure 3.14 *The average channel occupancy as a function of z in the pixel rings in the extended layout (Top Left) and in the inclined layout (Top Right). The average cluster density as a function of z in the pixel rings in the extended layout (Bottom Left) and in the inclined layout (Bottom Right).*

2D Results

The channel occupancy results for the strip and pixel detectors were placed in the $R - z$ plane to produce a 2D result (Figure 3.15) for the entire ITk.

Innermost radius and sensor thickness

At present, the innermost pixel layer for the ITk is set to be placed at $R = 39$ mm. The IBL that is used in Run 2 is located at $R = 33$ mm. Being closer to the interaction point has the advantage of more precise impact parameter

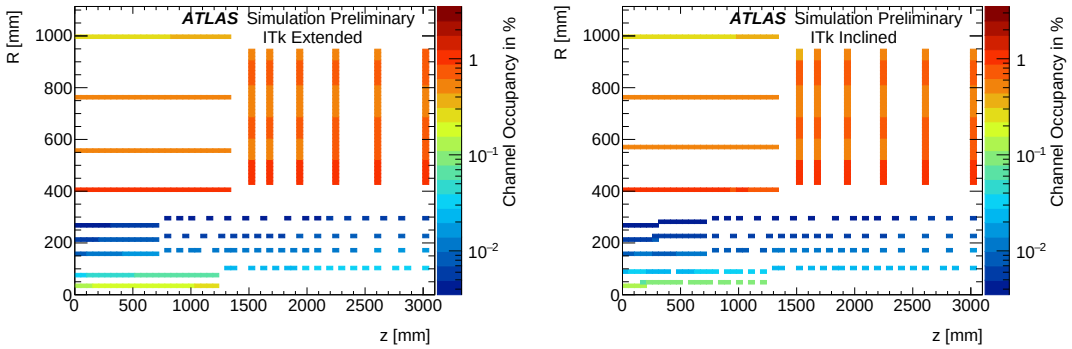

Figure 3.15 *The average channel occupancy in the inclined layout (Left) and extended layout (Right) in a 2D plane of R & z. The sensor positions are the average of the cluster positions associated to that sensor.*

measurements but also the disadvantage of increased channel occupancy, cluster density and more radiation damage.

The sensor thickness in the innermost layer was also modified. In the results that follow, the label (Step 1.5) refers to sensors with thickness 150 μm and (Step 1.6) refers to sensors with thickness 100 μm. These studies were only performed for the extended layout. There was also no minimum bias samples available for the $R = 33$ mm setup so $t\bar{t}$ samples were used and compared to minimum bias samples in Step 1.5 to check if the use of $t\bar{t}$ samples would cause any significant differences.

The average cluster size is shown in Figure 3.16 for pixel layer 0 in the barrel and ring layer 0 for the different samples. It can be seen that the cluster sizes decrease when the sensor thickness decreases and there is little dependence on the innermost radius.

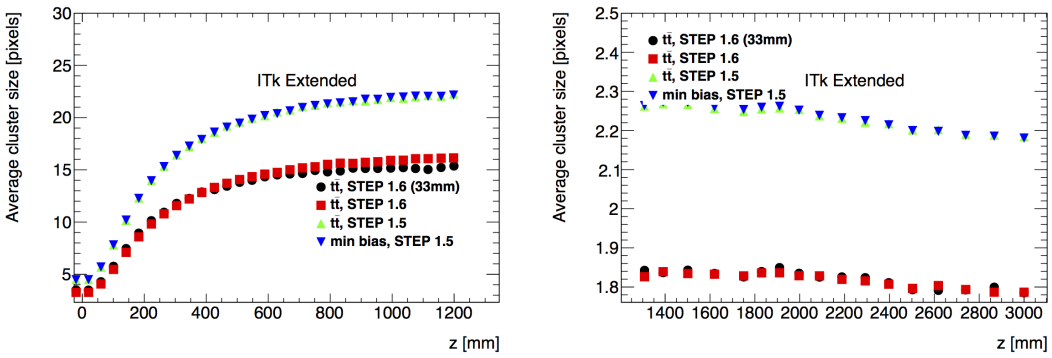

Figure 3.16 *The average cluster size as a function of z in layer 0 in the pixel barrel (Left) and pixel rings (Right) in different samples.*

The channel occupancy and cluster densities are both smaller when the sensor thickness is decreased due to the smaller cluster sizes. An increase in both quantities is observed as expected when R is moved from 39 mm to 33 mm.

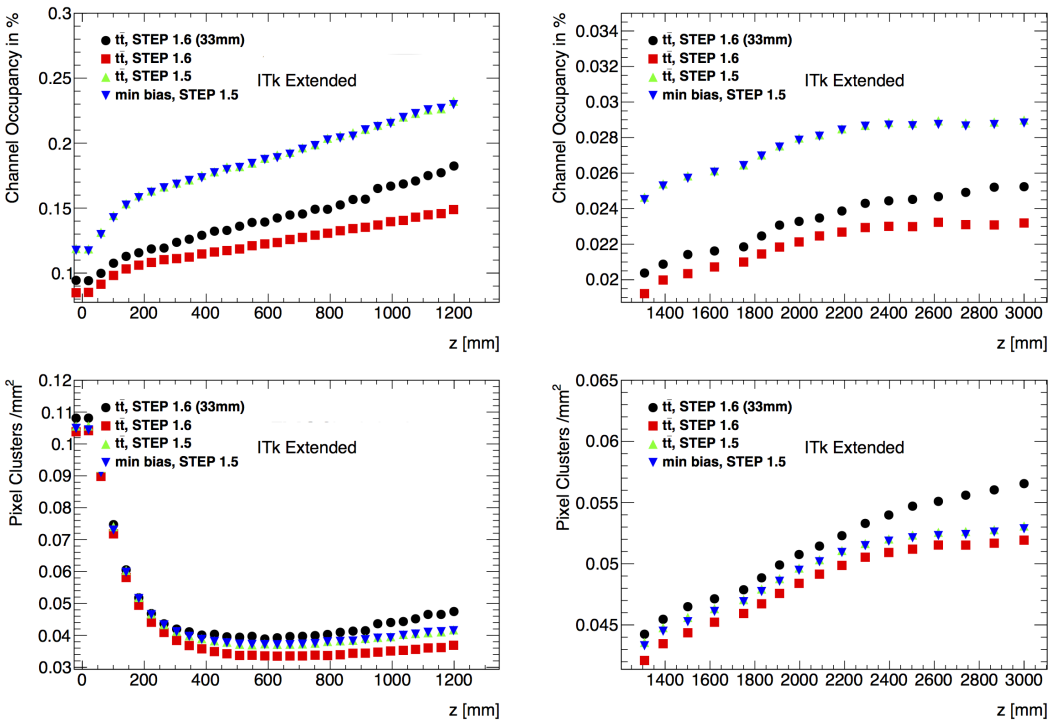

Figure 3.17 *The average channel occupancy as a function of z in layer 0 in the pixel barrel (Top Left) and pixel rings (Top Right) in different samples. The average cluster density as a function of z in layer 0 in the pixel barrel (Bottom Left) and pixel rings (Bottom Right) in different samples.*

3.9 Other ITk studies

Many other studies were performed by the ATLAS collaboration in order to reach the decision of building an ITk that resembles the inclined layout. Some of these studies are shown here to put this decision into context.

3.9.1 Material description

Since the amount of material has a large impact on tracking performance, particular care was taken to describe the material that was used in simulation. The material used in terms of radiation length X_0 is shown as a function of $|\eta|$ for both layouts in Figure 3.18. In the central region, both layouts are similar and have very little material. In the forward region, $|\eta| \geq 2.2$, the X_0 can be up to 30% greater in the extended layout because it has more silicon sensors where the pixel barrel is inclined in the inclined layout.

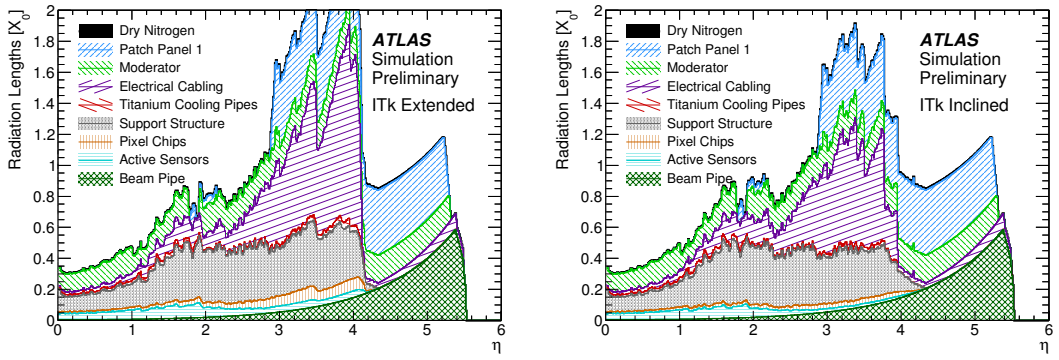

Figure 3.18 *The material description at the exit of the ITk in units of* X_0 *as a function of* $|\eta|$ *in the extended layout (Left) and for the inclined layout (Right) [61].*

3.9.2 Impact parameter resolutions

Inclined vs Extended

Results are shown in Figure 3.19 for the impact parameter resolutions for single muons with $p_{\mathrm{T}} = 1$, 10 and 100 GeV for the inclined and extended layouts. In the central region, $|\eta| \leq 2.2$, the layouts have almost identical performance for all p_{T}. In the forward region, for $|\eta| \geq 2.2$ and for p_{T} 1 - 10 GeV, the impact parameter resolutions are significantly better in the inclined layout. This is due to the inclined sensors, the particles traverse less material and the effect of multiple scattering on the track decreases. For $|\eta| \geq 2.2$ and $p_{\mathrm{T}} = 100$ GeV, the extended layout performs better. As shown in Figure 3.10, particles in the forward direction can create long clusters (≥ 20 pixels) containing angular information which can help improve the impact parameter resolutions for particles with $p_{\mathrm{T}} = 100$ GeV. However, high p_{T} muons with $|\eta| \geq 2.2$ are rarely produced in the decays of Standard Model paricles and are therefore unlikely to be important for physics analyses. The resolution from the inclined layout is also already adequate.

Change to innermost radius

Results are shown in Figure 3.20 for the impact parameter resolutions for single muons with $p_{\mathrm{T}} = 1$, 10 and 100 GeV for the extended layout with the innermost pixel layer set to $R = 39$ mm (nominal) and with it set to $R = 33$ mm. For muons with $p_{\mathrm{T}} = 1$ GeV, the impact parameter resolutions improve by approximately 10% for all η. For muons with $p_{\mathrm{T}} \geq 10$ GeV, there is only improvements at very high η.

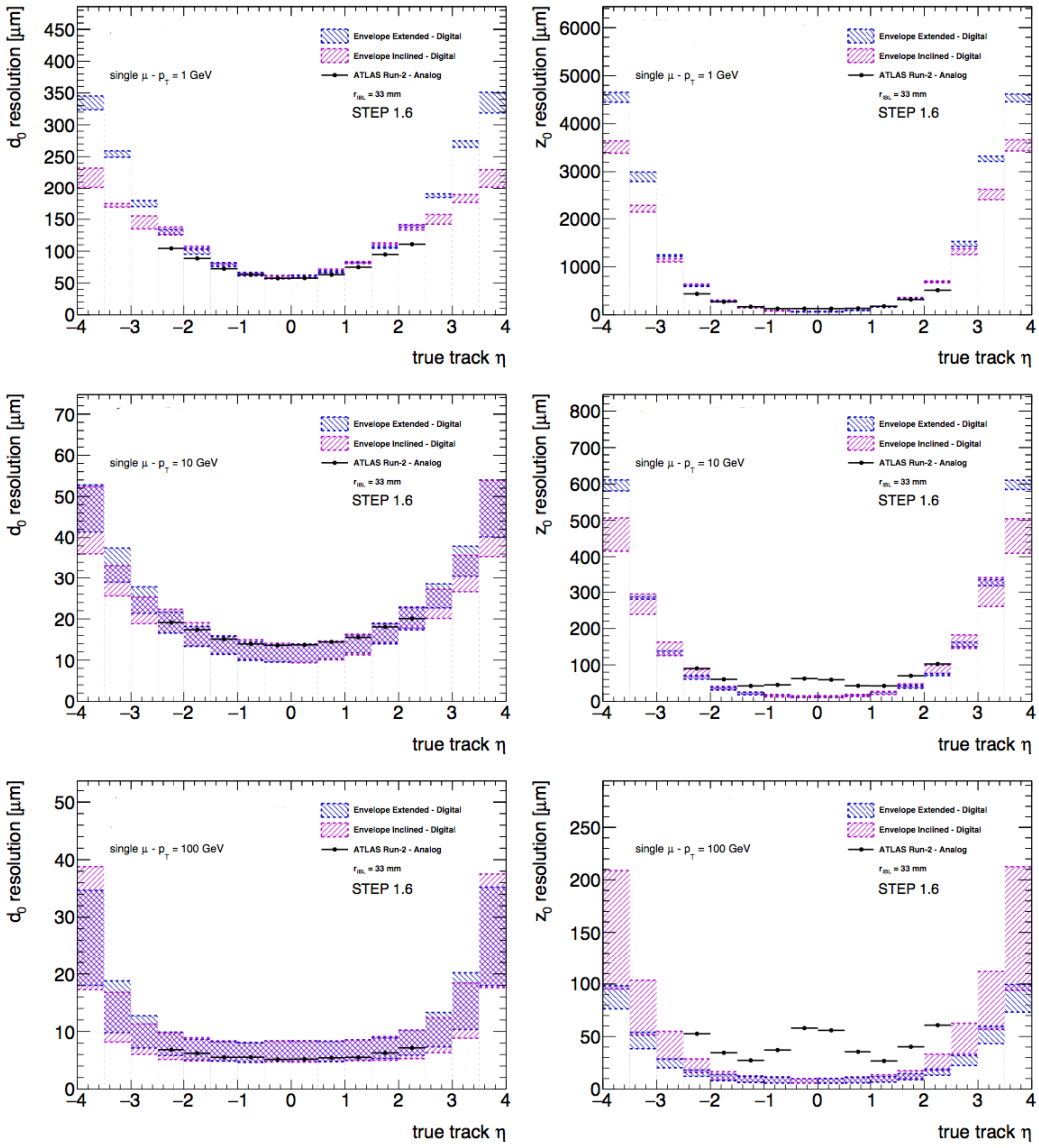

Figure 3.19 *Comparison of the transverse impact parameter resolutions $\sigma(d_0)$ (left column) and longitudinal impact parameter resolutions $\sigma(z_0)$ (right column) versus $|\eta|$ for muons with transverse momentum $p_T = 1, 10$ and 100 GeV comparing the inclined and extended layouts. The ATLAS Run-2 results are overlaid. [73]*

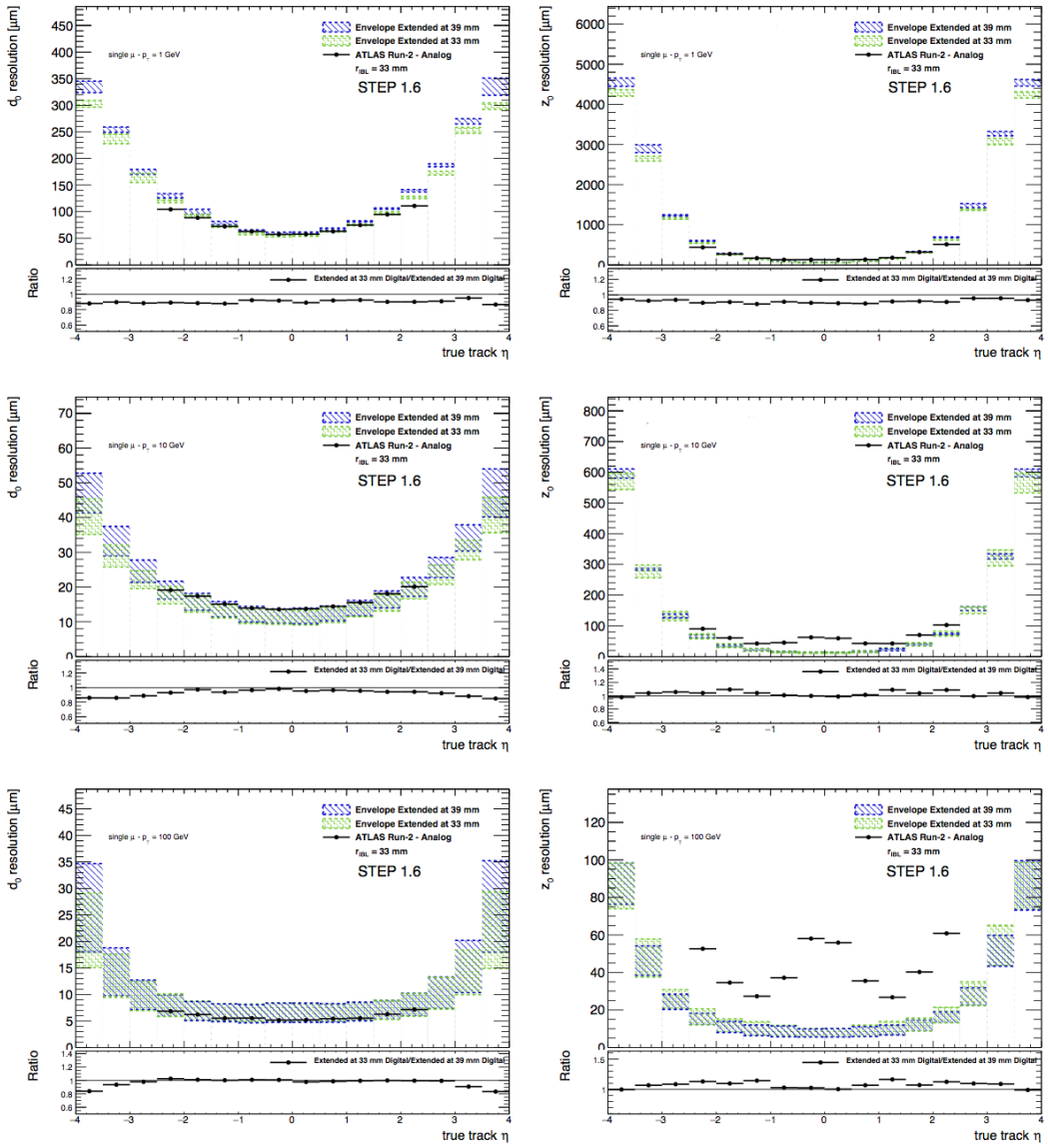

Figure 3.20 *Comparison of the transverse impact parameter resolutions $\sigma(d_0)$ (left column) and longitudinal impact parameter resolutions $\sigma(z_0)$ (right column) versus $|\eta|$ for muons with transverse momentum $p_T = 1, 10$ and 100 GeV comparing extended layouts where the innermost pixel layer has a different radius. [73]*

3.9.3 Jets

In this section, the rejection of jets which originate from pile-up is studied in both layouts. One of the most common techniques for doing this, is by considering the discriminant R_{pT}. The R_{pT} is defined as the scalar p_{T} sum of the tracks that are associated with the jet and originate from the hard-scatter vertex divided by the fully calibrated jet p_{T}:

$$R_{pT} = \frac{\sum_k p_T^{trk_k}(\mathrm{PV}_0)}{p_T^{jet}} \tag{3.4}$$

The tracks used in the calculation of R_{pT} must pass various quality requirements as defined in Ref. [74] and must have $p_{\mathrm{T}}^{trk} \geq 1$ GeV and $|\eta^{trk}| \leq 4.0$. Jets that have large values of R_{pT} are more likely to originate from the hard-scatter vertex. In this simple study, tracks are associated to the hard-scatter vertex by requiring the difference between the z_0 impact parameter of the track and the z-coordinate of the hard-scatter vertex to be less than a given value. The chosen value is two times the parameterised RMS of the $|z_0 - z_{HS}|$ distribution as a function of $|\eta^{trk}|$ computed from di-jet events, shown in Figure 3.21. The larger values of $|z_0 - z_{HS}|$ in the forward region from the extended layout reflects the larger z_0 resolution for low p_{T} tracks as shown in Figure 3.19. This larger window gives a higher contamination of tracks coming from pile-up that are near the hard-scatter vertex.

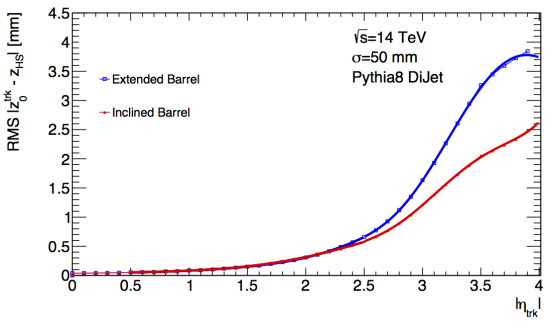

Figure 3.21 *RMS of $|z_0 - z_{HS}|$ as a function of $|\eta|$ of the track computed from di-jet events with $\langle\mu\rangle = 200$ using the inclined and extended layouts. [73]*

In the following, a reconstructed jet is defined as a hard-scatter jet if a truth jet associated with the hard-scatter vertex has $p_{\mathrm{T}}^{true} \geq 10$ GeV and is found within $\Delta R = 0.3$ of the reconstructed jet. Reconstructed jets which are separated by

more than $\Delta R = 0.6$ from any truth jet with $p_{\mathrm{T}}^{true} \geq 4$ GeV are labelled as pile-up jets. For events used in this study, the reconstructed hard-scatter vertex is required to be within 0.1 mm of the true hard-scatter vertex.

The distribution of R_{pT} for hard-scatter and pile-up jets is shown in Figure 3.22 for the $|\eta|$ region, $2.9 \leq |\eta| \leq 3.8$. In the other $|\eta|$ regions, the performance between the two layouts is almost identical. The R_{pT} distribution for pile-up jets has a larger tail for the extended layout as a result of the worse z_0 resolution for low p_{T} tracks.

The efficiency for pile-up jets as a function of the efficiency for hard-scatter jets for passing an R_{pT} selection is shown in Figure 3.22. It can be seen that for $|\eta| \leq 2.9$, the performance of the two layouts (triangles and rectangles) is very similar while there is a significant difference in the forward region, $|\eta| \geq 2.9$ (circles). In the region, $|\eta| \geq 2.9$, for a fixed efficiency of pile-up jets of 5%, the efficiency for hard-scatter jets is 76% for the extended layout and 86% for the inclined layout.

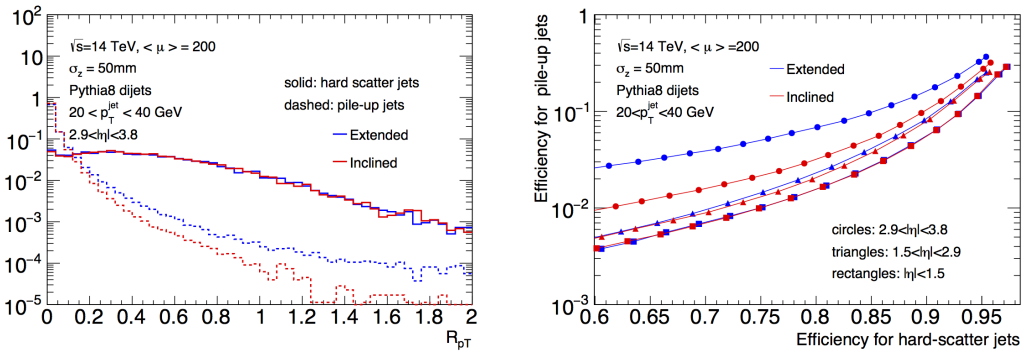

Figure 3.22 *The R_{pT} distribution for hard-scatter and pile-up jets for jets with $20 \leq p_T \leq 40$ GeV in the region $2.9 \leq |\eta| \leq 3.8$ (Left). The efficiency for pile-up jets as a function of the efficiency for hard-scatter jets with $20 \leq p_T \leq 40$ GeV using the R_{pT} discriminant (Right). Both quantities are calculated from Monte Carlo dijet events with $\langle\mu\rangle = 200$ using the inclined and extended layouts [73].*

3.9.4 Summary

The studies shown in this chapter led to the following recommendations:

- Keep the innermost radius of the innermost pixel layer at $R = 39$ mm instead of $R = 33$ mm. There are small improvements to the impact parameter resolutions in the Extended layout when moving the innermost pixel layer from $R = 39$ mm to $R = 33$ mm, however the channel occupancy increases by 10% to 30% depending on position, and the data transmission rate is already expected to be challenging. There was no radiation damage studies performed for the $R = 33$ mm setup which was also a significant concern.

- The ITk community will continue to study and build an ITk which resembles the inclined layout. In the forward region, the impact parameter resolutions are better for the inclined layout at low p_{T}. This low p_{T} regime is what is expected from the HL-LHC in the forward region. The superior impact parameter resolution produces better rejection of forward pile-up tracks which translates into better suppression of pile-up jets.

Chapter 4

Prospects for Observing $t\bar{t}HH$ Production at the HL-LHC

In this chapter, the prospects for observing $t\bar{t}HH$ production with both Higgs bosons decaying to $b\bar{b}$ are presented. This study assumes HL-LHC conditions, an upgraded ATLAS detector, and 3000 fb^{-1} of proton-proton collisions at $\sqrt{s} =$ 14 TeV.

This study has been approved as an ATLAS public note [75].

4.1 Motivation

There are a large number of studies, both phenomenological and by the ATLAS and CMS collaborations studying Higgs boson pair production at the HL-LHC using the ggF production mechanism. The other production mechanisms have largely been ignored due to their smaller cross sections, however they do offer additional handles for rejecting background processes. For example, charged leptons from a vector boson decay or b-jets originating from the decay of a top quark. It is important to examine these other production mechanisms and understand if they could improve the sensitivity to the search for HH production.

The predicted cross section for $t\bar{t}HH$ production at $\sqrt{s} = 14$ TeV is 0.98 fb [25], in comparison to 37 fb [24] for ggF HH production and 55000 fb [19] for ggF single Higgs production. The following phenomenological publications [76, 77]

suggest that $t\bar{t}HH$ production is a promising channel to study Higgs boson pair production:

Englert, C., F. Krauss, M. Spannowsky, and J. Thompson. "Di-Higgs phenomenology in tthh: The forgotten channel." Phys. Lett. B743: (2015) 93-97. [76]

Liu, T., and H. Zhang. "Measuring Di-Higgs Physics via the tthh → ttbbbb Channel." [77]

Both publications study $t\bar{t}HH$ production with both Higgs bosons decaying to $b\bar{b}$ assuming a data set of $3000\,\text{fb}^{-1}$. Ref. [76] does not study the impact of a detector apart from assuming that b-jets can be tagged with an efficiency of 70% for a 1% mistag rate for other jets while Ref. [77] uses the DELPHES 3 [78] simulation to emulate an ATLAS-style detector. Both studies find that approximately $2\,\sigma$ statistical significance can be achieved for the $t\bar{t}HH$ process by targeting the semi-leptonic final state of the $t\bar{t}$ system, that is: $t\bar{t} \to WWb\bar{b} \to l\nu_l qqb\bar{b}$ where l is an electron, muon or tau particle.

The most promising HH HL-LHC study by the ATLAS collaboration is in the $b\bar{b}\gamma\gamma$ final state. The expected statistical significance is $2\,\sigma$ and the ratio of the Higgs boson self-coupling to its SM expectation, κ_λ is expected to be constrained to $-1.1 \leq \kappa_\lambda \leq 8.1$ at 95% CL [63].

If $2\,\sigma$ statistical significance for $t\bar{t}HH$ production can be achieved with the upgraded ATLAS detector and the pile-up conditions, the $t\bar{t}HH$ process would be one of the strongest motivations for the HL-LHC program. The unique signature of six b-jets would also make this a benchmark process to optimise the design of the upgraded inner detector.

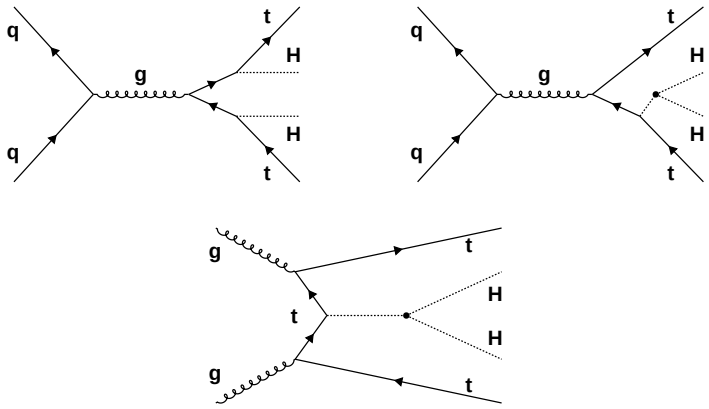

Figure 4.1 *Leading order Feynman diagrams for $t\bar{t}HH$ production.*

4.2 Signal and Background Generation

The signal and background Monte Carlo samples used for this analysis are summarised in Table 4.1. In all samples a Higgs boson of $m_H = 125.0\,\text{GeV}$ is used. The branching ratio for $H \to b\bar{b}$ is taken from Ref. [79].

Sample	Generator	σ (fb)	Filter	Events in $3\,\text{ab}^{-1}$	Events Generated
$t\bar{t}HH(HH \to b\bar{b}b\bar{b})$	MadGraph/Pythia8	0.33	-	990	20,000
$t\bar{t}b\bar{b}$ + jets	Sherpa	3750	0.52	5,850,000	6,000,000
$t\bar{t}H(H \to b\bar{b})$ + jets	Sherpa	371	0.55	612,000	600,000
$t\bar{t}Z(Z \to b\bar{b})$ + jets	Sherpa	163	0.55	269,000	300,000

Table 4.1 *Summary of the signal and background samples used in this analysis. The background samples are generated with a filter requiring a charged lepton (e, μ or τ) with $p_T > 20$ GeV. Additional filters are placed on the $t\bar{t}b\bar{b}$ process, b-quarks are required to have $p_T > 15$ GeV and m_{bb} is required to be > 30 GeV at the matrix element level.*

The signal $t\bar{t}HH(HH \to b\bar{b}b\bar{b})$ is generated using the MadGraph [12] generator at leading order. Showering and hadronisation are simulated using Pythia8 [15]. The A14 tune [80] of shower and multiple parton interactions parameters is used together with the NNPDF2.3 PDF set [81].

Following the approach in the phenomenological studies, the following backgrounds are considered: $t\bar{t}b\bar{b}$+jets, $t\bar{t}Z(Z \to b\bar{b})$+jets and $t\bar{t}H(H \to b\bar{b})$+jets. It should be noted that in Ref. [76], $t\bar{t}b\bar{b}b\bar{b}$ is considered as a background, while this analysis considers $t\bar{t}b\bar{b}$+jets.

Each background process is generated using Sherpa2.2 [9] at leading order, using massless b-quarks, and with up to two additional jets using the NNPDF3.0 PDF

set [82]. This analysis considers only final states containing electrons or muons, therefore a filter is applied to select events with at least one electron, muon or tau with $p_{\mathrm{T}} > 20$ GeV. Additionally, for the $t\bar{t}b\bar{b}$ sample, a selection at the matrix element level is applied to the b-quarks of $p_{\mathrm{T}} > 15$ GeV and on the invariant mass $m_{bb} > 30$ GeV.

The cross section for the $t\bar{t}HH$ sample is normalised to the next-to-leading-order prediction [25] of 0.98 fb (before the Higgs boson branching ratio is applied). The background samples are normalised to leading order, using the cross section calculated by the generator.

4.3 Analysis

4.3.1 Object selection

Jets

This analysis uses stable particles from Monte Carlo events. The stable particles, excluding muons and neutrinos are clustered into truth jets using the anti-k_t algorithm with a radius parameter of 0.4 via the FastJet package [55].

The p_{T} of the truth jets are smeared by an η- and p_{T}-dependent parameterisation as described in Ref. [83]. For central jets with p_{T} around 30 GeV the smearing has a standard deviation of approximately $50\% \times p_{\mathrm{T}}$; for central jets with $p_{\mathrm{T}} >$ 100 GeV, the smearing has a standard deviation of approximately $12\% \times p_{\mathrm{T}}$.

All jets are required to have $p_{\mathrm{T}} > 30$ GeV and $|\eta| \leq 4.0$. Jets that are within ΔR of 0.1 of an electron or muon candidate are removed from consideration.

Each generated event is also overlaid with simulated pile-up jets, these jets are taken from a library which are produced from fully simulated $\langle\mu\rangle = 200$ events.

Pile-up jets can be suppressed by using information from the inner detector. The discriminating variable R_{pT} as defined in Chapter 3 is used. The efficiency for pile-up jets to pass an R_{pT} selection is shown as a function of the efficiency for hard-scatter jets to pass the same selection in dijet $\langle\mu\rangle = 200$ events in Figure 4.2. A working point is used so that 98% of pile-up jets with $p_{\mathrm{T}} \geq 25$ GeV and $|\eta| \leq$ 2.5 are rejected. The efficiency for the hard scatter jets to pass the same selection

is $\geq 86\%$. The average number of pile-up jets as a function of $|\eta|$ before and after the R_{pT} requirement is shown in Figure 4.2.

The truth flavour of the jets is determined by matching the truth b/c-hadron to a truth jet within a ΔR of 0.2. If the truth hadron matches more than one jet, only the jet closest in ΔR is labelled as a b/c jet. Any truth jets which are not matched to a b/c hadron are labelled as light jets and have originated from other quark flavours or a gluon.

In this analysis, the p_{T}, $|\eta|$ and truth flavour of each jet is used as an input to a parameterisation of the ATLAS MV1 b-jet tagger algorithm [84]. Working points are defined which correspond to given b-jet tagging efficiencies and associated rejection rates of non b-jets. The 70% working point corresponds to 70% of truth b-jets being correctly identified in $t\bar{t}$ events, the working point has a background rejection rate of 5 for c-jets and 500 for light jets.

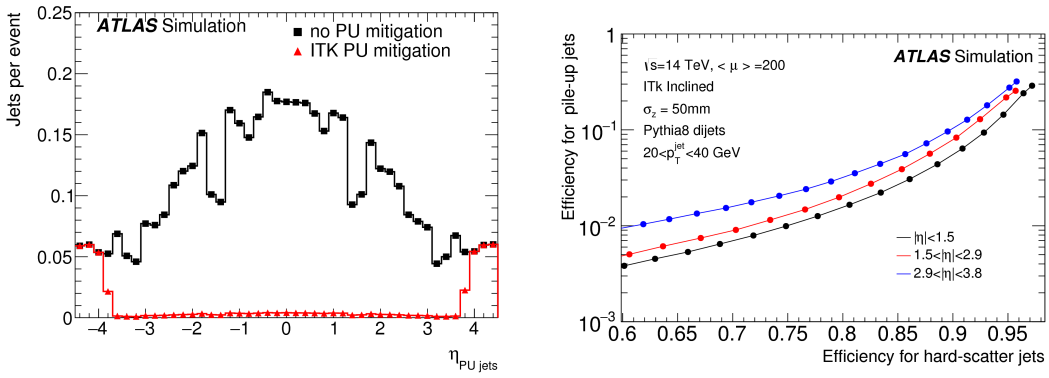

Figure 4.2 *Left: The η distribution of the number of pile-up jets before and after the R_{pT} selection. Right: The efficiency for pile-up jets as a function of the efficiency for hard-scatter jets.*

Electrons and muons

Electrons and muons are taken from the stable particle record of the Monte Carlo events. The energy and p_{T} of electrons and muons are smeared using p_{T} and η-dependent functions, as described in Ref. [83].

Electrons and muons are required to have $p_{\mathrm{T}} > 25$ GeV. Electrons are required to have $|\eta| \leq 4.0$ while muons are required to have $|\eta| \leq 2.5$. An isolation requirement is placed on the electrons and muons, the sum of the transverse energy in a cone of ΔR of 0.2 must be less than $0.2 \times p_{\mathrm{T}}$. The isolation requirement removes around 2% of electrons and muons.

4.3.2 Event selection

This analysis follows a similar strategy to the phenomenological studies and targets the semi-leptonic top quark decay final state. The event selection is optimised to maximise the statistical significance for a counting experiment, $S/\sqrt{B}$, where S is the number of signal events and B is the number of background events.

- Events must have exactly one isolated electron or muon with $p_{\mathrm{T}} > 25$ GeV and $|\eta| \leq 4.0$ for electrons and $|\eta| \leq 2.5$ for muons.
- A single lepton trigger requirement for HL-LHC running is emulated: each event must have an electron or muon consistent with trigger requirements [85].
- Events must have ≥ 7 jets with $p_{\mathrm{T}} > 30$ GeV and $|\eta| \leq 4.0$.
- ≥ 5 b-tag selection: at least 5 jets are required to be b-tagged.
- The average separation in pseudorapidity between two b-tagged jets: $\langle \eta(b_i, b_j) \rangle$ is required to be less than 1.25.
- Events are sorted into 5 b-tag and ≥ 6 b-tag categories.
- No requirement is made on the missing transverse momentum.

The single electron trigger selects electrons with $|\eta| < 2.5$ with an efficiency of 95% for 22 GeV $< p_{\mathrm{T}} <$ 35 GeV and 100% for $p_{\mathrm{T}} > 35$ GeV; for electrons with $2.5 < |\eta| < 4.9$ and $p_{\mathrm{T}} > 35$ GeV, the efficiency is 90%. The single muon trigger selects muons with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$ with an efficiency of around 96%.

4.3.3 Event shape variables

Several event shape variables are studied in order to discriminate the $t\bar{t}HH$ signal from the background processes. The following variables are found to have the most discriminating power:

1. The average separation in pseudorapidity between two b-tagged jets: $\langle \eta(b_i, b_j) \rangle$.

2. Centrality, defined as the scalar sum of p_{T} for all jets, divided by the energy sum of all jets.

3. The scalar sum of p_{T} for b-tagged jets, H_B.

These variables are shown in Figures 4.4 and 4.5.

A scan is performed on the above variables in order to optimise the statistical significance $S/\sqrt{B}$. A single selection criteria of $\langle \eta(b_i, b_j) \rangle < 1.25$ is found to maximise the significance. No selection on either the centrality or H_B is made as these do not improve the statistical significance.

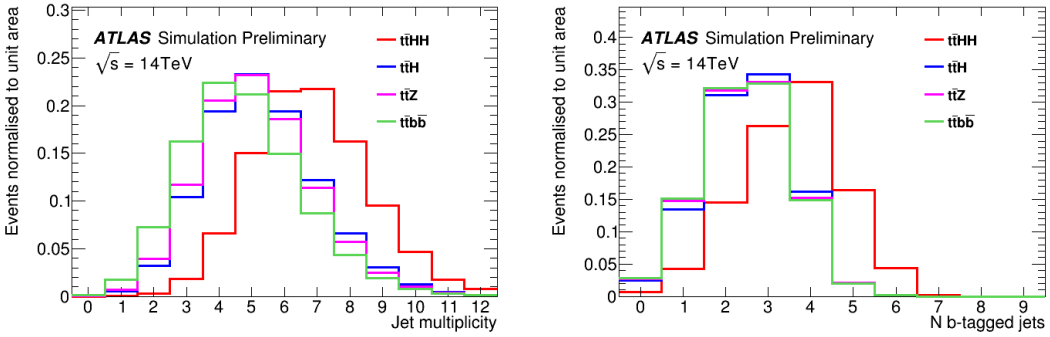

Figure 4.3 *Left: Jet multiplicity, after the selections for trigger and electron/muon. Right: Number of b-tagged jets in events that have passed the ≥ 7 jets selection.*

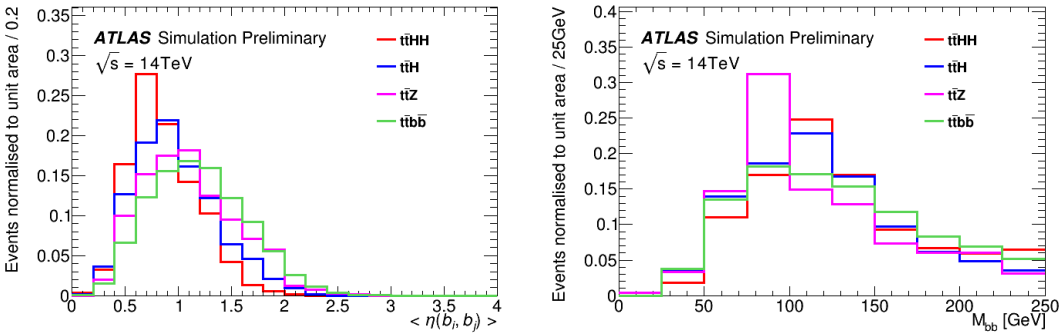

Figure 4.4 *Left: The average separation in pseudorapidity between two b-tagged jets $\langle\eta(b_i, b_j)\rangle$, after trigger, lepton and jet selections. Right: Higgs boson candidate mass, m_{bb}, found from selecting the b-tagged jets which have the largest vector sum p_T, shown for events that have passed the ≥ 5 b-tag selection, normalised to unity.*

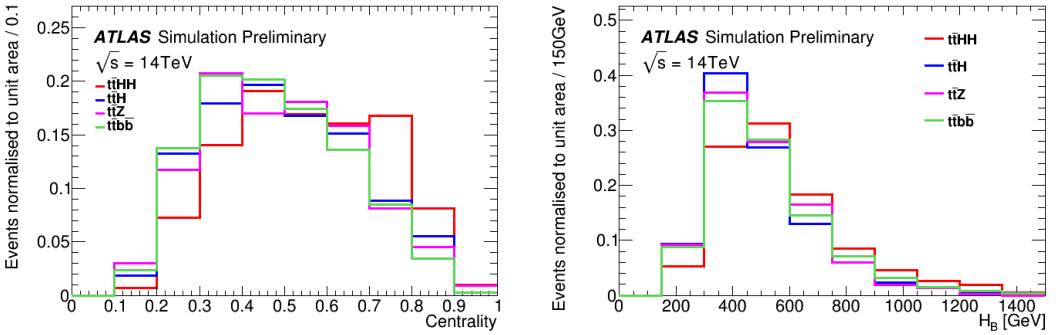

Figure 4.5 *Left: Centrality; Right: H_B. Both variables are plotted after trigger, lepton and number of jets requirements with ≥ 5 b-jets.*

4.3.4 Higgs Boson Candidate Reconstruction

Higgs boson candidates can be reconstructed from the b-jets. However, due to the large number of b-jets in the event, there is a combinatorial problem. Two different methods are studied to assign the b-tagged jets to the Higgs boson candidates.

The first method follows the procedure performed in Refs [76, 77] by selecting the b-tag jet pairs that minimise the following quantity:

$$\chi^2 = (m_{b_1 b_2} - m_H)^2 + (m_{b_3 b_4} - m_H)^2 \tag{4.1}$$

where m_H is set to 120 GeV. The choice of 120 GeV, not the 125 GeV used for simulation is motivated by the loss of the jet energy through decays to neutrinos. The distributions for $m_{b_1 b_2}$ and $m_{b_3 b_4}$ are shown in Figure 4.6. It can be seen that there is little discriminating power in these variables.

A second method finds the pair of b-tagged jets which have the largest vector sum p_{T} and assigns this pair as a Higgs boson candidate; this is shown in Figure 4.4.

In contrast with Refs [76, 77], it is found that making a requirement on m_{bb} does not improve the statistical significance in either case and therefore no requirements are made on the masses of the Higgs boson candidates. It is thought that this is due to the inclusion of a realistic detector response.

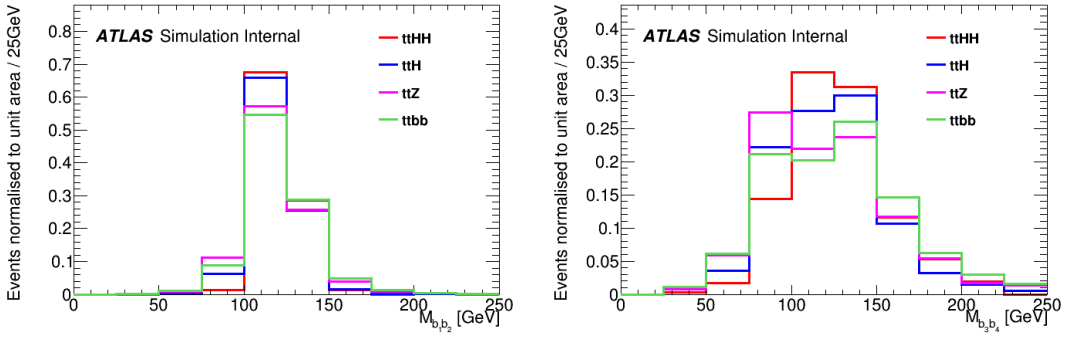

Figure 4.6 *$m_{b_1 b_2}$ and $m_{b_3 b_4}$ found by minimising the χ^2.*

4.3.5 Systematic uncertainties

This analysis is most similar to the search for $t\bar{t}H(H \to b\bar{b})$ production which ATLAS has performed in Run 1 [86] and Run 2 [87]. The dominant systematic

uncertainties in these results arise from the understanding of the $t\bar{t}b\bar{b}$ process. In Run 1, a 50% uncertainty was assigned to the normalisation of $t\bar{t}b\bar{b}$ due to differences in the cross section predicted from different event generators. In Run 2, the normalisation of $t\bar{t}b\bar{b}$ was allowed to float in the fit and the fit favoured a normalisation scale factor of $k = 1.24 \pm 0.10$.

The HL-LHC is set to finish in approximately 2035 so theory calculations are expected to evolve and the data from the LHC can be used to improve the predictions of event generators further. The 95% CL exclusion limit on the cross section for $t\bar{t}HH$ production is therefore computed as a function of the same normalisation systematic uncertainty applied to all backgrounds. No other systematic uncertainties are considered.

4.3.6 Results

The event yields for the $t\bar{t}HH$ signal and the background processes considered are shown for each selection in Table 4.2. For the ≥ 5 b-tag selection, the number of signal and background events in $3000\,\text{fb}^{-1}$ is 25 and 7,100 respectively, with the largest background contribution coming from the $t\bar{t}b\bar{b}$ + jets process.

For exactly 5 (6) b-tags, the number of signal events is 19 (6) and the number of background events is 6,600 (510), resulting in a statistical significance of $0.23\,\sigma$ ($0.27\,\sigma$). These two categories can be combined in quadrature to give an overall statistical significance of $0.35\,\sigma$.

Sample	None	Trigger	One lepton	$\geq$7 jets	$\geq$5 b-tags	$\eta(b_i, b_j)$	=5 b-tags	$\geq$6 b-tags
$t\bar{t}HH(HH \to b\bar{b}b\bar{b})$	990	513	253	139	29	25	19	6
$t\bar{t}H(H \to b\bar{b})$ + jets	610,000	500,000	290,000	69,000	1,580	1,200	1110	90
$t\bar{t}Z(Z \to b\bar{b})$ + jets	270,000	220,000	125,000	26,000	600	390	360	30
$t\bar{t}b\bar{b}$ + jets	5,900,000	4,800,000	2,800,000	460,000	9,700	5,500	5100	400
total background	6,800,000	5,500,000	3,200,000	550,000	11,900	7,100	6580	520

Table 4.2 *Summary of event selection criteria applied to signal and background processes for* $3000\,\text{fb}^{-1}$*. The background samples are filtered to require a charged lepton with* $p_T > 20$ GeV*, whereas no filter is required on the signal sample; this leads to the appearance of a smaller trigger efficiency for the signal sample.* $\eta(b_i, b_j)$ *refers to the* $\langle \eta(b_i, b_j) \rangle < 1.25$ *requirement.*

For this analysis, the likelihood is a product of two Poisson functions, one for the

5 b-tag category and one for the 6 b-tag category given by:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-\mu s_j + b_j} \cdot G(\boldsymbol{\theta}) \tag{4.2}$$

where μ is the signal strength defined as the ratio of the measured cross section to the SM expectation, s_j is the number of signal events expected, b_j is the number of background events expected and the total number of events expected is given by n_j.

The systematic uncertainties are managed by introducing nuisance parameters $\boldsymbol{\theta}$ where $G(\boldsymbol{\theta})$ is a set of unit Gaussian constraints. In this analysis, the only systematic uncertainties considered are normalisation uncertainties on the backgrounds. These are modelled using log-normal distributions.

The expected 95% CL upper limit on the signal strength is shown as a function of the systematic uncertainties in Figure 4.7. A description of the statistical tests used at the LHC is given in Appendix B.

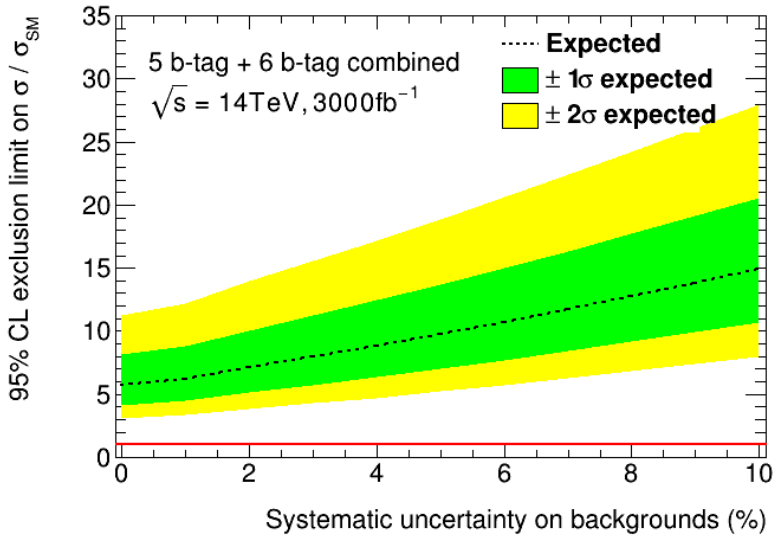

Figure 4.7 *The expected 95% CL upper limit on $\sigma(t\bar{t}HH)/\sigma_{\mathrm{SM}}$ as a function of a normalisation systematic uncertainty placed on all the backgrounds.*

4.4 Comparisons with phenomenological publications

Both Refs. [76, 77] found that $2\,\sigma$ statistical significance could be achieved for the $t\bar{t}HH$ process at the HL-LHC assuming a dataset of $3000\,\text{fb}^{-1}$ whereas this study finds that only $0.35\,\sigma$ can be achieved. The main difference between this study and Refs. [76, 77] are the background processes considered. For example, Ref. [76] uses a smaller background of $t\bar{t}b\bar{b}b\bar{b}$ while this study uses $t\bar{t}b\bar{b}$ with up to an additional two jets. The event selection between the two studies is very similar which faciltates comparisons. In Ref. [76] they find 134 $t\bar{t}b\bar{b}b\bar{b}$ events after requiring 5 b-tagged jets at the 70% working point, while this study finds 5,500 events for $t\bar{t}b\bar{b}$ + jets.

The composition of the events in the $t\bar{t}b\bar{b}$ + jets sample passing the event selection was examined. For every $t\bar{t}b\bar{b}$ + jets event that passes the 5 b-tag jet selection, the event is categorised using truth information. The anti-k_T truth jets without smearing applied are used. A jet is only counted if it has $p_{\text{T}} > 30$ GeV and $|\eta| < 2.5$ to match the jet definition used in Ref. [76]. The truth flavour of the the jet is found by matching the truth b/c-hadrons to the truth jet within a ΔR of 0.2. For jets which originate from a charm hadron, the truth record of the charm hadron was examined to identify if it was the result of the decay of a W boson. The results are shown in Table 4.3. It is found that the dominant background in this study is $t\bar{t}b\bar{b}$ + jets with mistagging a c-jet from the decay of a W boson as a b-jet and that the $t\bar{t}b\bar{b}b\bar{b}$ component which was considered by Refs. [76, 77] only makes up a small fraction of the background events.

Classification		Number of events		
truth-b-jets	truth-c-jets	all events	with c-jet from W	with c-jets not from W
2	0	15		
2	≥ 1	100	90	10
3	0	200		
3	≥ 1	1170	1020	150
4	0	1130		
4	≥ 1	4630	3890	740
5		1950		
6		410		

Table 4.3 *Classification of truth jets in the $t\bar{t}b\bar{b}$ sample for the $\geq$ 5 b-jet selection. The main component of the background is due to real charm jets from the decay of W-bosons, that are mistagged as b-jets.*

4.5 Future improvements

A multivariable analysis approach which used a Boosted Decision Tree with the variables shown in Figures 4.3 and 4.4 was briefly studied, however the improvement to the significance was modest (approximately 10%). It is found that the dominant background is from mistagging c-jets in the $t\bar{t}b\bar{b}$ + jets process. For a b-tag working point of 70%, the rejection rates obtained for c-jets and light jets were 5 and 500 respectively. Since this study was carried out, the design of the ITk detector has evolved and the b-tagging algorithm has been re-optimised for HL-LHC conditions. The rejection rates for c-jets and light jets are now 20 and 700 respectively [67].

A two bin counting experiment is used to determine the significance and to establish the upper limit on the cross section. This could be improved upon by performing a binned fit to a variable which discriminates between the signal and background processes, exploiting the fact that each bin has different S/B. An example of a variable that could be used is the scalar sum of jet p_{T}, the H_T shown in Figure 4.8.

4.6 Summary

This is the first time the $t\bar{t}HH$ process has been studied by an LHC collaboration. A cut based selection has been employed targeting the semi-leptonic decay of the

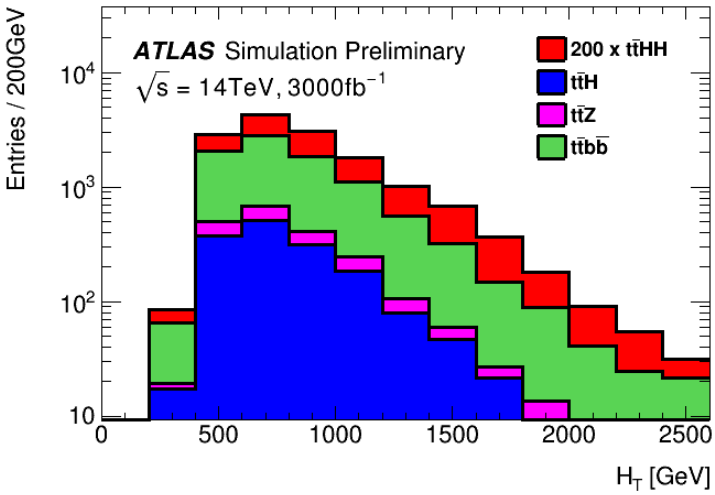

Figure 4.8 *The scalar sum of jet p_T is shown for the signal and background processes normalised to* $3000\,\text{fb}^{-1}$ *after the* $\geq$ *5 b-tag selection. The* $t\bar{t}HH$ *signal is scaled by a factor of 200 so it is visible.*

top quark pair as studied in Refs [76, 77]. The event selection is optimised for the statistical significance. It is found that by combining the 5 b-tag and 6 b-tag signal regions together, a statistical significance of $0.35\,\sigma$ can be achieved. The expected 95% CL exclusion limit on $\sigma(t\bar{t}HH)/\sigma_{\text{SM}}$ is determined as a function of the normalisation systematic uncertainties placed on all backgrounds. The exclusion limit varies between 6 and 12 times the SM value depending on the systematic uncertainties assumed as shown in Figure 4.7.

The result shown here is markedly different from that in Refs. [76, 77] which found that approximately $2\,\sigma$ could be achieved for the $t\bar{t}HH$ process. The origin of this difference is due to Refs. [76, 77] underestimating the backgrounds, in particular the mis-tag rate for c-jets.

The statistical significance for $t\bar{t}HH$ production is approximately a factor of 6 smaller than the most promising HH HL-LHC ATLAS projection ($b\bar{b}\gamma\gamma$). The $t\bar{t}HH$ production cross section also has a weaker dependence on λ_{HHH} as can be seen in Figure 1.9. It can therefore be concluded that at best it will make a small contribution to the combined search for Higgs boson pair production and the measurement of the Higgs boson self-coupling, λ_{HHH}.

Chapter 5

The search for $HH \to b\bar{b}\gamma\gamma$

This chapter presents the search for Standard Model $HH \to b\bar{b}\gamma\gamma$ production with the full Run 2 data set.

5.1 Introduction

The search for $HH \to b\bar{b}\gamma\gamma$ aims to measure the cross section for HH production, σ_{HH} and ultimately measure the Higgs boson self-coupling, λ_{HHH}. However, the Run 2 data set is not sensitive to the SM, as the cross section for HH production is approximately three orders of magnitude smaller than single Higgs boson production. The small HH cross section is a result of the reduced phase space from requiring an additional Higgs boson, an additional interaction term and destructive interference between two Feynman diagrams shown in Figure 5.1.

Many BSM theories predict enhanced rates of Higgs boson pair production as discussed in Chapter 1. For example, simply setting λ_{HHH} to 0, increases the cross section of HH production by more than a factor of 2. In what follows, κ_λ is used to denote the ratio of the Higgs boson self-coupling to its Standard Model expectation, that is $\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{SM}$.

In order to measure κ_λ, it is necessary to understand how the kinematics of HH production are modified when κ_λ varies. The Higgs boson propagator in the *triangle* diagram, shown in Figure 5.1 (right), is probed off-mass-shell and as a result the triangle diagram is suppressed in the SM compared to the box

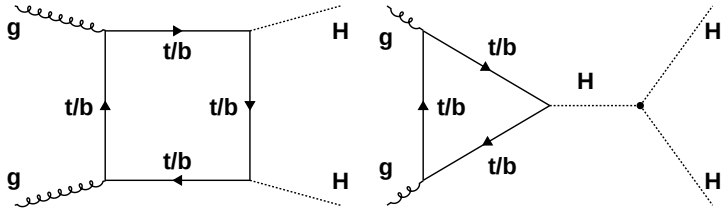

Figure 5.1 *Examples of Feynman diagrams for Higgs boson pair production via gluon-fusion. In the Standard Model, there is destructive interference between the quark loop (left) and the Higgs boson self-coupling (right) which reduces the cross section [75].*

diagram, shown in Figure 5.1 (left). Figure 5.2 shows the m_{HH} distribution for a number of different κ_λ hypotheses. For $\kappa_\lambda = +6$, the triangle diagram dominates HH production and the distribution peaks at $2m_H$ which is the kinematic region where the Higgs boson propagator in the triangle diagram is least suppressed. For $\kappa_\lambda = 0$, there is no contribution from the triangle diagram and the m_{HH} shape is characterised by an increase in cross section when $m_{HH} > 2m_{top}$. Hence, to target large BSM values of κ_λ, it is necessary to search at the m_{HH} threshold while SM HH production is characterised by large values of m_{HH}.

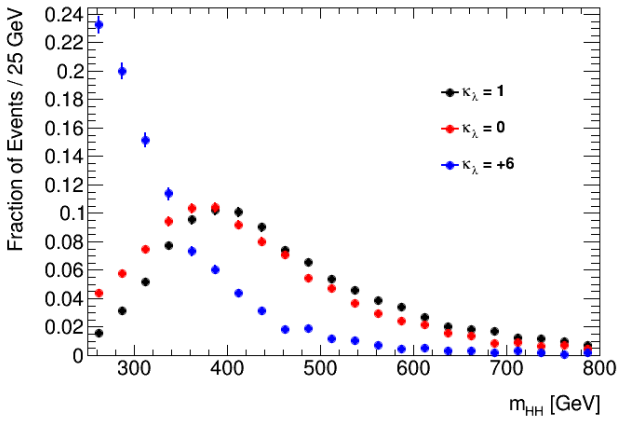

Figure 5.2 *The m_{HH} distribution is shown for a number of different κ_λ hypotheses.*

The $b\bar{b}\gamma\gamma$ final state is a particularly attractive channel for studying HH production. It combines the large branching ratio of $H \to b\bar{b}$ with the low backgrounds, clean trigger and excellent mass resolution of $H \to \gamma\gamma$. In Standard Model $H \to \gamma\gamma$ analyses, a narrow peak appears in the diphoton invariant mass spectrum on top of a smoothly falling background and the signal can be extracted by performing a fit to the $m_{\gamma\gamma}$ distribution. A $HH \to b\bar{b}\gamma\gamma$ signal would also appear in the $m_{\gamma\gamma}$ spectrum, on top of a $H \to \gamma\gamma$ peak from single Higgs boson

production and a smoothly falling background. The $b\bar{b}\gamma\gamma$ channel also contains a second mass peak with the $H \to b\bar{b}$ decay. This could prove to be particularly powerful in constraining the single Higgs background since they have the same resonant shape in $m_{\gamma\gamma}$ but not in m_{bb}.

The analysis strategy is to use multiple categories in order to maximise the sensitivity and then to extract the signal by performing a fit to the $m_{\gamma\gamma}$ distribution as traditionally done in $H \to \gamma\gamma$ analyses and then to study extracting the signal by performing a 2D fit to both the $m_{\gamma\gamma}$ and m_{bb} distributions.

5.2 Data and Monte Carlo samples

This analysis uses a data set of proton-proton collisions at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector from 2015 to 2018. This corresponds to an integrated luminosity of $139.0\,\text{fb}^{-1}$. Table 5.1 shows the integrated luminosity obtained in each year and the average number of pp interactions in a bunch crossing in each year.

Year	Integrated luminosity pb^{-1}	$<\mu>$
2015	3219.56	13.4
2016	32995.4	25.1
2017	44307.4	37.8
2018	58450.1	37.0
Total	138972	34.2

Table 5.1 *The integrated luminosity obtained and the average number of pp interactions in each data taking year.*

Standard Model gluon-gluon fusion HH production is simulated at NLO accuracy in QCD using an effective field theory approach with form factors for the top-quark loop from HPAIR [88] to approximate the finite top quark mass effects. The simulated events are reweighted to the m_{HH} spectrum obtained in Ref. [89] which calculated the process at NLO while fully accounting for the top-quark mass. This simulated sample is then normalised to a cross section of 33.41 fb which has been calculated at NNLO [90]. HH production with varied κ_λ is simulated at leading order accuracy in QCD for eleven values of κ_λ.

For the background processes, the Sherpa 2.2.4 [9] program is used to generate continuum $\gamma\gamma$ production with up to one additional jet at NLO and up to three

additional jets at LO. The $t\bar{t} + \gamma$ process is also considered as a background and is generated using MadGraph with Pythia8 performing the shower and hadronisation.

An important source of background in this analysis is from single Higgs boson production. The dominant single Higgs boson backgrounds in this analysis are the ggF, ZH and $t\bar{t}H$ production mechanisms. All single Higgs boson production mechanisms are considered and the samples used to model them are described in Table 5.2.

The HH signal samples, the continuum $\gamma\gamma$ background and the $t\bar{t} + \gamma$ samples are passed through a fast parametric simulation of the ATLAS detector [91] while the single Higgs boson background samples are passed through a detailed Geant4 [92] simulation of the ATLAS detector.

Process	Generator	Showering	PDF set	σ [fb]	Order of calculation of σ	Simulation
ggF HH	MadGraph5_aMC@NLO	Herwig++	CT10 NLO	33.41	NNLO+NNLL	Fast
BSM ggF HH	MadGraph5_aMC@NLO	Pythia	NNPDF 2.3 LO	-	LO	Fast
$\gamma\gamma$ + jets	Sherpa	Sherpa	CT10 NLO	-	NLO	Fast
$t\bar{t} + \gamma$	MadGraph5_aMC@NLO	Pythia	NNPDF 2.3 LO	-	LO	Fast
ggF	Powheg-Box NNLOPS	Pythia	PDF4LHC15	48520	NNLO(QCD) + NLO(EW)	Full
VBF	Powheg-Box	Pythia	PDF4LHC15	3780	approximate-NNLO(QCD)+NLO(EW)	Full
WH	Powheg-Box	Pythia	PDF4LHC15	1370	NNLO(QCD) + NLO(EW)	Full
$q\bar{q} \to ZH$	Powheg-Box	Pythia	PDF4LHC15	760	NNLO(QCD)+NLO(EW)	Full
$t\bar{t}H$	Powheg-Box	Pythia	NNPDF3.0	510	NNLO(QCD) + NLO(EW))	Full
$b\bar{b}H$	Powheg-Box	Pythia	NNPDF3.0	490	NNLO(QCD) + NLO(EW)	Full
$gg \to ZH$	Powheg-Box	Pythia	PDF4LHC15	120	NNLO(QCD) + NLO(EW)	Full
t-channel tH	MadGraph5_aMC@NLO	Pythia	CT10 NLO	70	4FS(LO)	Full
W-associated tH	MadGraph5_aMC@NLO	Herwig++	CT10 NLO	20	5FS(NLO)	Full

Table 5.2 *Summary of the event generator programs and PDF sets used to model the signal and background processes. The SM cross sections which are used for the normalisation of Higgs boson production are also shown. The event generators used are:* Pythia8 *[15],* Herwig++ *[13],* Powheg-Box *[93],* MadGraph5_*aMC@NLO [12],* Sherpa *2.2.1 [9]. The PDF sets used are: CT10 NLO [94], NNPDF2.3 LO [81], NNPDF 3.0 LO [82] and PDF4LHC15 [95]*

5.3 Object selection

5.3.1 Photons

The two highest E_{T} photons with $|\eta| < 2.37$, excluding the transition region between the barrel and endcap calorimeters of $1.37 < |\eta| < 1.52$ are selected as the diphoton Higgs candidate. The leading (subleading) photon is required to

satisfy $E_{\mathrm{T}}/m_{\gamma\gamma} > 0.35$ (0.25).

Identification

The photon identification used by ATLAS relies on selections applied to calorimetric variables which separate prompt photons from background photons. Background photons are not only jets faking photons but also genuine photons such as those from a π^0 decay.

Ten variables are used in total for photon identification. These variables can be grouped into two categories: Hadronic leakage variables which use information from the hadronic calorimeter, and electromagnetic variables which use information from the first and second layers of the EM calorimeter.

Loose and tight identification selections are defined. The loose selection only uses information from the hadronic calorimeter and the second layer in the EM calorimeter, while the tight selection exploits information from all three layers. The identification efficiency for the loose (tight) selection for a photon with E_{T} = 40 GeV is 99% (85%) which corresponds to a background rejection factor (jets faking photons) of 1000 (5000) [96].

Both photons are required to pass the tight identification selection.

Isolation

Photons from a Higgs boson decay are typically well isolated from jet activity so requirements are placed on the isolation using information from the calorimeter and the inner detector.

To calculate calorimeter isolation, the transverse energies of topological clusters are summed up in a cone of $\Delta R = 0.2$ around the photon. The transverse energy of the photon and contributions from pile-up are removed from the cone in this computation.

To calculate track isolation, the p_{T} of the tracks are summed up in a cone of $\Delta R = 0.2$ around the photon. For converted photons, tracks that are associated with the photon conversion are removed.

The calorimeter (track) isolation is required to be 6.5% (5%) less than the

transverse energy of the photon.

Both photons are required to pass the calorimeter and track isolation requirements. After requiring the tight identification selection, the isolation requirements offer a further rejection factor of 1.5 for jets faking photons [96].

5.3.2 Jets

Jets are reconstructed from topological clusters of energy deposits in the calorimeter cells using the anti-k_t algorithm with a radius parameter of 0.4 via the FastJet package [55].

Jets originating from pile-up with $|\eta| \leq 2.4$ are suppressed using the jet vertex tagger (JVT) [58] which combines tracking information into a multivariate likelihood.

All jets used in this analysis are required to be central ($|\eta| \leq 2.5$) and to have p_{T} > 25 GeV.

Flavour Tagging

A multivariable discriminant is used to classify b-jets from other flavours of jets [59]. The discriminant uses impact parameter information, reconstructed secondary vertex position and decay information as inputs. Working points are defined by requiring that the discriminant is greater than a certain value that corresponds to a specific b-tagging efficiency in $t\bar{t}$ events. The working points and their corresponding background rejection factors are shown in Table 5.3.

b-tag working point	charm rejection	τ rejection	light rejection
60%	22	151	1150
70%	8	38	301
77%	4	15	109
85%	2	6	27

Table 5.3 *The rejection rates for charm, τ and light jets for each b-tag working point.*

5.3.3 Leptons

Electron candidates are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.47$, excluding the region $1.37 < |\eta| < 1.52$. Electrons must also satisfy identification and isolation requirements.

Muon candidates are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.7$. Muons must satisfy identification and isolation requirements.

Hadronic tau candidates are not used in this analysis.

5.4 Data / MC comparisons

It is not essential in this analysis for MC simulation to describe data well across a range of kinematic variables since the background in the analysis is estimated using data-driven methods described later. However if the MC is to be used for analysis optimisation, it is necessary for there to be at least a reasonable agreement between data and MC. The MC is also used to assess the bias from the choice of functional form used to model the background which is discussed in more detail later.

The following loose preselection is used to compare data and simulation: N_{cen} jets < 6, $N_{lep} = 0$, $N_{85\%\,b-tags} \geq 2$, $N_{70\%\,b-tags} < 3$, $m_{bb} \in$ [60,180] GeV.

Figure 5.3 shows comparisons between data and MC for a number of variables. In all figures, events in the signal region $m_{\gamma\gamma} \in$ [120,130] GeV are removed in both data and MC with the exception of the $m_{\gamma\gamma}$ figure where the events are removed in data only.

In general, the shapes between MC and data are in reasonable agreement but the normalisation is not. This is likely due to a missing γ + jets component. It is not practical to simulate this component but nor is it necessary since the background is estimated using data-driven methods.

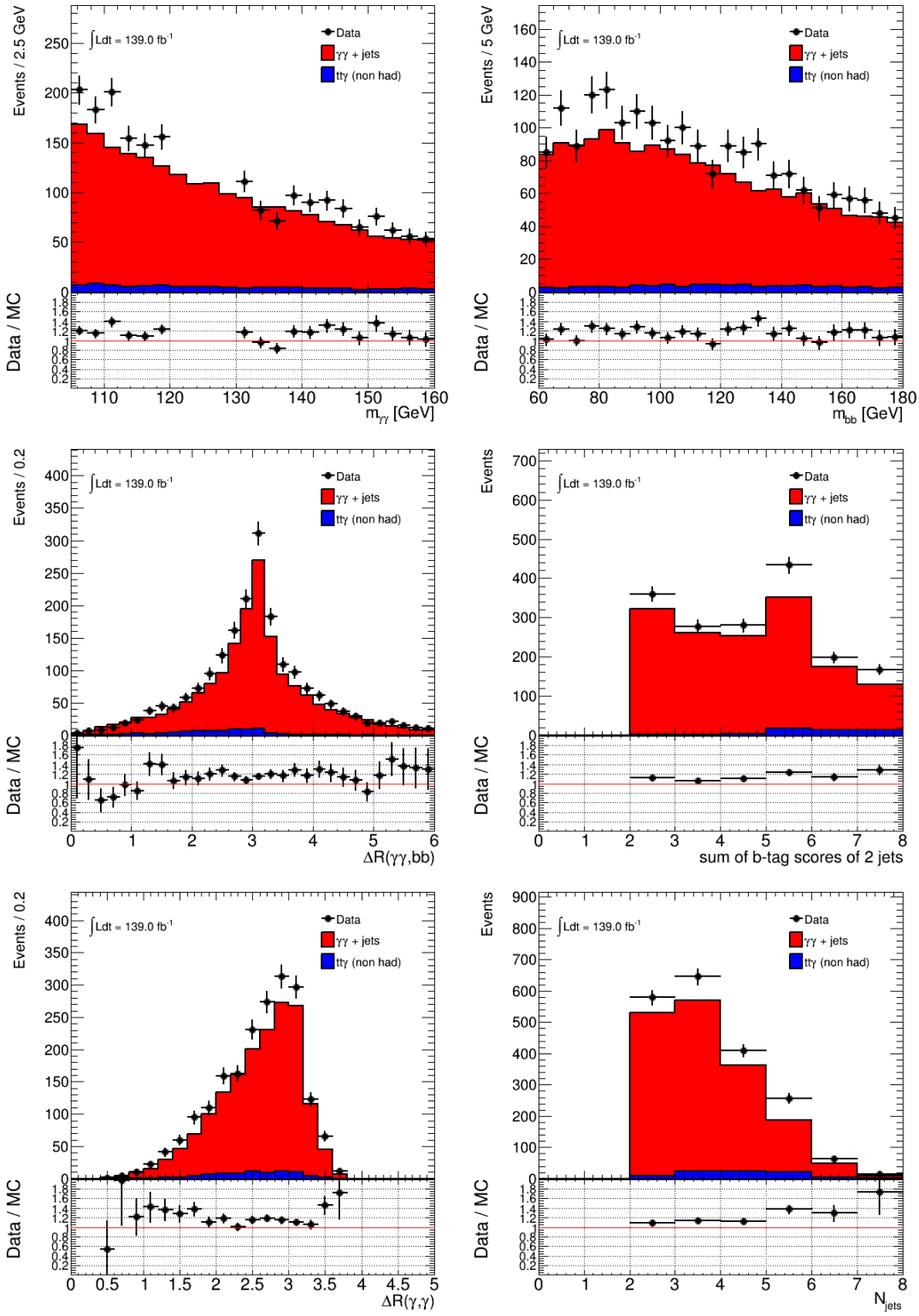

Figure 5.3 *Comparison between data and simulation for $m_{\gamma\gamma}$, m_{bb}, $\Delta R(\gamma\gamma, bb)$, the sum of the b-tag scores of the two jets that make up the $H \to b\bar{b}$ candidate, $\Delta R(\gamma, \gamma)$ and the jet multiplicity.*

5.5 Event selection

5.5.1 $H \to \gamma\gamma$ selection

Multiple triggers are used at the High Level Trigger (HLT) for $H \to \gamma\gamma$ candidate events during Run 2. The use of multiple triggers is a result of the different pile-up conditions in each data taking year. All triggers require two photons to be reconstructed at the HLT passing a loose photon identification criteria with the isolation criteria varying depending on data taking year. The leading (subleading) photon must have $E_{\mathrm{T}} > 35(25)$ GeV.

This analysis uses the nominal $H \to \gamma\gamma$ selection which is common to all $H \to \gamma\gamma$ physics analyses in ATLAS. Events are required to have the following:

- The event must pass one of the diphoton triggers used.
- At least one primary vertex is required to be reconstructed in the event.
- Two photons passing the tight identification selection.
- Two photons are required to pass the isolation criteria, that is the calorimeter (track) isolation is required to be 6.5% (5%) less than the transverse energy of the photon.
- The leading and subleading photons must have $E_{\mathrm{T}}/m_{\gamma\gamma} > 0.35$ and $E_{\mathrm{T}}/m_{\gamma\gamma} > 0.25$.

The acceptance $\times$ efficiency for the $H \to \gamma\gamma$ selection is shown as a function of κ_λ in Figure 5.4. As discussed in the introduction, the kinematics of HH production become significantly softer for large BSM values of κ_λ and therefore the efficiency of the $H \to \gamma\gamma$ selection decreases at these values.

5.5.2 Primary vertex selection

In most physics analyses, the primary vertex is usually selected by choosing the vertex which has the highest $\sum p_{\mathrm{T}}^2$ of tracks associated to it and is often referred to as the *hardest* vertex. However this selection is not optimal for $H \to \gamma\gamma$ since unconverted photons do not leave tracks in the inner detector. A neural network

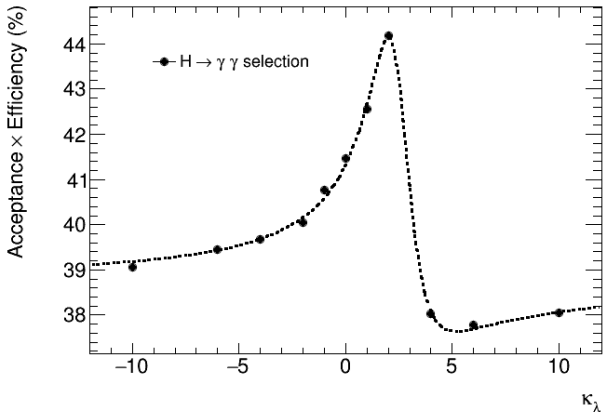

Figure 5.4 *The acceptance × efficiency for the $H \to \gamma\gamma$ selection as described in 5.5.1 as a function of κ_λ.*

trained on a ggF $H \to \gamma\gamma$ sample is used to select the primary vertex using the following inputs:

- The combined z-position of the intersections of the extrapolated photon trajectories (reconstructed by exploiting the longitudinal segmentation of the calorimeter) with the beam axis.
- $\sum p_{\mathrm{T}}$, the scalar sum of transverse momenta of the tracks associated to the vertex.
- $\sum p_{\mathrm{T}}^2$, the sum of the squared transverse momenta of the tracks.
- $\Delta\phi(\gamma\gamma, vertex)$: the azimuthal angle between the diphoton system and the system defined by the vector sum of the tracks associated to the vertex.

The selected vertex in each event is the one that maximises the output value of the neural network [30]. The vertex selection efficiency is shown as a function of $\langle\mu\rangle$ for ggF $H \to \gamma\gamma$ and SM $HH \to b\bar{b}\gamma\gamma$ production in Figure 5.5. The primary vertex selected is deemed correct if the z position of the selected vertex is within 0.3 mm of the true vertex. The neural network improves the selection efficiency significantly for ggF $H \to \gamma\gamma$ while in $HH \to b\bar{b}\gamma\gamma$, the additional jet activity results in the hardest vertex actually being a better choice. Since the vertex selection efficiency in either case is already very good ($\geq 90\%$) for $b\bar{b}\gamma\gamma$, the vertex selected by the neural network is used in order to be compatible with other $H \to \gamma\gamma$ physics analyses in ATLAS.

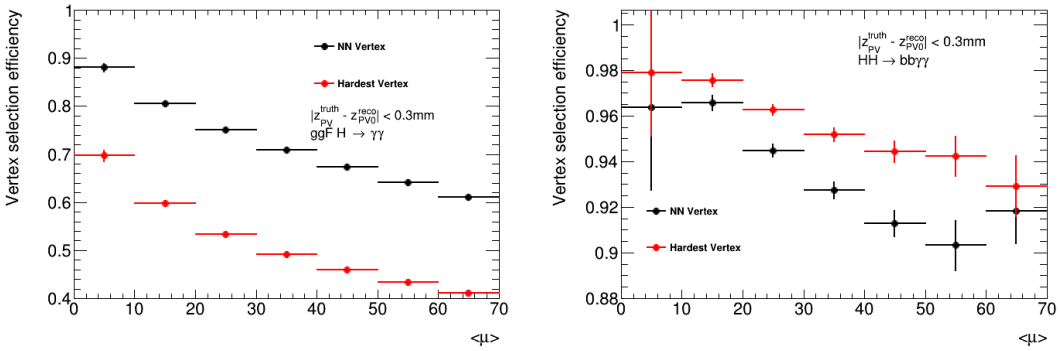

Figure 5.5 *The vertex selection efficiency as a function of $\langle\mu\rangle$ for ggF $H \to \gamma\gamma$ (Left) and $HH \to b\bar{b}\gamma\gamma$ (Right).*

After the neural network has selected the vertex, the position of the photons is updated by redefining the η of the photon using a straight line that connects the impact point in the first layer of the EM calorimeter with the vertex selected. This results in an improvement to the $m_{\gamma\gamma}$ resolution in ggF $H \to \gamma\gamma$, shown in Figure 5.6 (left) and in $HH \to b\bar{b}\gamma\gamma$, the resolution is approximately the same, shown in Figure 5.6 (right).

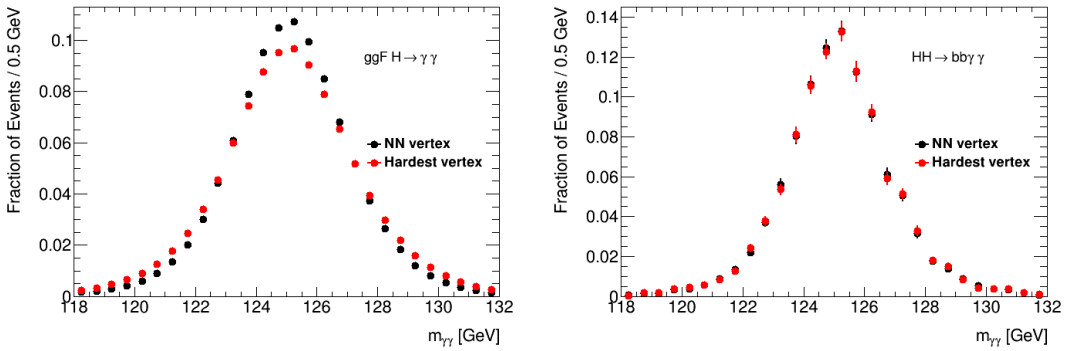

Figure 5.6 *The $m_{\gamma\gamma}$ distribution in ggF $H \to \gamma\gamma$ (left) and $HH \to b\bar{b}\gamma\gamma$ (right) when the hardest vertex is selected (red) and when the neural network vertex is selected (black).*

5.5.3 κ_λ optimisation

The kinematics of HH production change significantly for $\kappa_\lambda \neq 1$. For large values of κ_λ, the kinematics become particularly soft while SM HH production is characterised by high transverse momentum kinematics. Figure 5.7 shows the

so called modified $b\bar{b}\gamma\gamma$ mass defined as

$$M_{\mathrm{X}}^{*} = M_{b\bar{b}\gamma\gamma} - M_{b\bar{b}} - M_{\gamma\gamma} + 250.0\,\mathrm{GeV} \tag{5.1}$$

for different κ_λ signal hypotheses and the $\gamma\gamma$ + jets background. The analysis is therefore separated into different categories using this variable: $M_{\mathrm{X}}^{*} < 350$ GeV to target BSM values of κ_λ and $M_{\mathrm{X}}^{*} \geq 350$ GeV to target SM HH production.

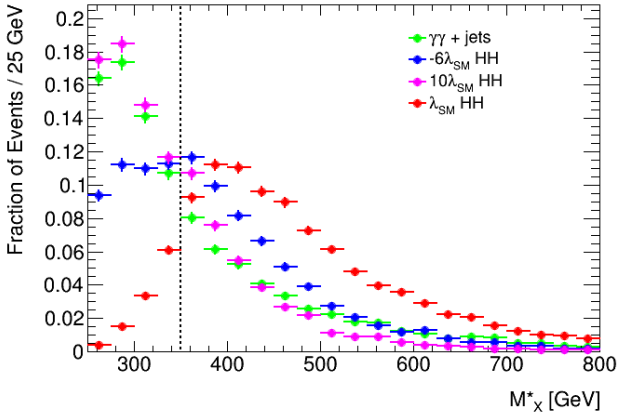

Figure 5.7 *M_{X}^{*} for a number of different κ_λ hypotheses and the $\gamma\gamma$ + jets background.*

5.5.4 $H \to b\bar{b}$ candidate

The $H \to b\bar{b}$ candidate is constructed by ranking the jets depending on the b-tag working point passed and then by p_{T}. The first two jets are taken as the $H \to b\bar{b}$ candidate.

5.5.5 Angular variables

The angles between the photons and the b-jets are found to be powerful in rejecting the $\gamma\gamma$ + jets continuum background in the $M_{\mathrm{X}}^{*} \geq 350$ GeV category while not so useful in the $M_{\mathrm{X}}^{*} < 350$ GeV category.

Figures 5.8 show the angular variables $\Delta R(bb, \gamma\gamma)$, $\Delta R(\gamma, \gamma)$ and $\Delta R(b, b)$ in both categories.

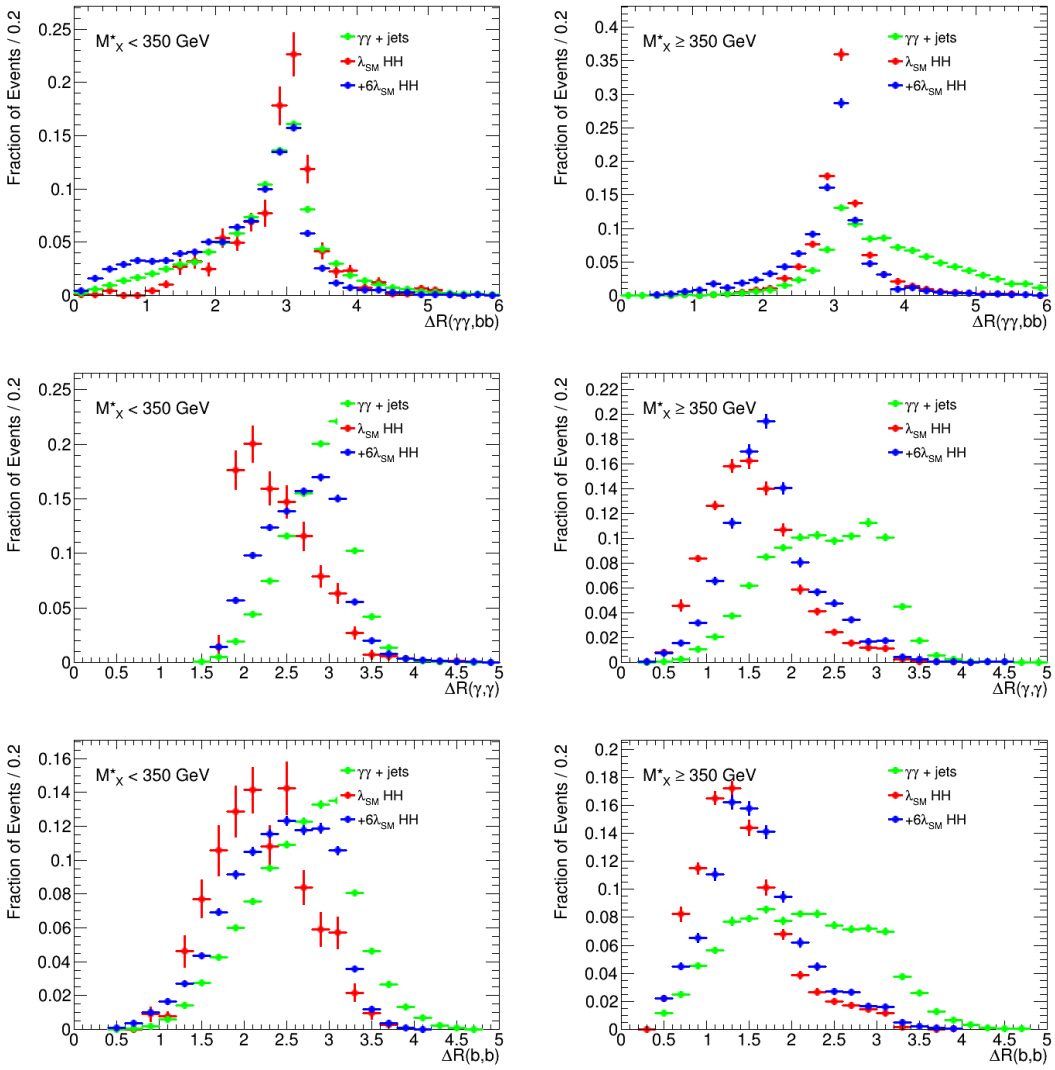

Figure 5.8 *Top:* $\Delta R(bb,\gamma\gamma)$ *Middle:* $\Delta R(\gamma,\gamma)$ *Bottom:* $\Delta R(b,b)$ *for SM* HH*,* $\kappa_\lambda = 6$ *signal and the* $\gamma\gamma$ *+ jets background in the* $M_X^* \geq 350$ GeV *(right) and* $M_X^* < 350$ GeV *(left) categories.*

5.5.6 Selection

Events are separated into categories using M_X^* and then further sorted into *b-tag tight* categories which require 2 jets to have passed the 70% b-tag working point. If this requirement is failed, events can be accepted into *b-tag loose* categories which has a looser requirement of 2 jets passing the 85% b-tag working point resulting in a total of four categories.

Figure 5.8 illustrates that angular variables are powerful in separating signal and background in the $M_X^* \geq 350$ GeV categories. In the b-tag tight, $M_X^* \geq 350$ GeV category, the analysis is optimised for the SM signal by performing a 3D scan on the angular variables in order to optimise for the statistical significance Z for a counting experiment, given by [97]

$$Z = \sqrt{2((s+b)\ln(1+s/b) - s} \tag{5.2}$$

where s is the number of SM HH events and b is sum of the number of $\gamma\gamma$ + jets ($m_{\gamma\gamma} \in [120,130]$ GeV) and single Higgs events. The angular selections found ($\Delta R(\gamma,\gamma) < 2.0$, $\Delta R(b,b) < 2.0$ and $\Delta R(bb,\gamma\gamma) < 3.4$) are then harmonised with the b-tag loose category.

In order to suppress the $t\bar{t}H$ background, 0 leptons (where a lepton is an electron or muon as defined in Section 5.3.3) are required and the number of central jets is required to be less than six.

In order to remain orthogonal to the ATLAS search for $HH \to b\bar{b}b\bar{b}$, events must have less than 3 b-tagged jets at the 70% working point.

Two different analysis strategies are employed. One strategy fits the $m_{\gamma\gamma}$ distribution while another strategy performs a 2D fit to the $m_{\gamma\gamma}$ and m_{bb} distributions. In the first strategy, the $H \to b\bar{b}$ candidate mass is required to be in the interval [90,140] GeV. While in the second strategy, the selection is loosened to [80,180] GeV since the shape of m_{bb} is exploited. In what follows, these selections will be referred to as the 1D and 2D selections respectively. The choice of [80,180] GeV is made in the 2D selection, since as seen in Figure 5.3, there is a small turn on at approximately 70 GeV which is difficult to model.

Table 5.4 summarises the 1D selection.

Common selection	N_{cen} jets < 6, $N_{lep} = 0$, $N_{85\%\, b-tags} \geq 2$, $N_{70\%\, b-tags} < 3$, $m_{bb} \in$ [90,140] GeV
Category	Further Selections
b-tag tight, $M_X^* \geq 350$ GeV	$M_X^* \geq 350$ GeV, $N_{70\%\, b-tags} = 2$, $\Delta R(\gamma,\gamma) < 2.0$, $\Delta R(b,b) < 2.0$, $\Delta R(bb,\gamma\gamma) < 3.4$
b-tag loose, $M_X^* \geq 350$ GeV	$M_X^* \geq 350$ GeV, $\Delta R(\gamma,\gamma) < 2.0$, $\Delta R(b,b) < 2.0$, $\Delta R(bb,\gamma\gamma) < 3.4$
b-tag tight, $M_X^* < 350$ GeV	$M_X^* < 350$ GeV, $N_{70\%\, b-tags} = 2$
b-tag loose, $M_X^* < 350$ GeV	$M_X^* < 350$ GeV

Table 5.4 *Summary of the 1D selection.*

The event yields and the acceptance $\times$ efficiencies ($A \times \epsilon(\%)$) for each Higgs boson production mechanism are shown in Tables 5.5 and 5.6 for the 1D and 2D selections respectively.

Higgs	$M_X^* \geq$ 350 GeV *b*-tag tight		$M_X^* \geq$ 350 GeV *b*-tag loose		$M_X^* <$ 350 GeV *b*-tag tight		$M_X^* <$ 350 GeV *b*-tag loose	
	Yield	$A \times \epsilon(\%)$	Yield	$A \times \epsilon(\%)$	Yield	$A \times \epsilon(\%)$	Yield	$A \times \epsilon(\%)$
SM HH	0.668	5.45	0.285	2.32	0.104	0.847	0.0463	0.378
ggH	0.184	0.0012	0.326	0.00213	0.235	0.00154	0.719	0.0047
VBF	0.0176	0.00148	0.0332	0.00279	0.0258	0.00217	0.0771	0.00646
WH	0.00305	0.000706	0.0479	0.0115	0.0101	0.00251	0.119	0.0281
ZH	0.289	0.12	0.212	0.0883	0.261	0.109	0.281	0.117
ggZH	0.122	0.314	0.0927	0.239	0.0137	0.0354	0.0203	0.0523
ttH	0.356	0.223	0.412	0.258	1.06	0.664	0.994	0.622
bbH	0.00774	0.00505	0.00667	0.00435	0.0657	0.0428	0.0802	0.0523
tHW	0.009	0.188	0.0152	0.317	0.00966	0.202	0.0182	0.38
$tHjb$	0.0208	0.0887	0.0315	0.134	0.0614	0.262	0.0579	0.247

Table 5.5 *The yields and efficiences in each category for the HH signal and the single Higgs boson backgrounds for the 1D selection.*

As discussed in 5.5.3, when $\kappa_\lambda \neq 1$, both the kinematics and the cross section for HH production are also modified. In order to interpret the analysis in terms of κ_λ, the acceptance $\times$ efficiency has to be parameterised as a function of κ_λ. This is done using a function of the form:

$$f(\kappa_\lambda) = \frac{A + B\kappa_\lambda + C\kappa_\lambda^2}{D + E\kappa_\lambda + F\kappa_\lambda^2} \tag{5.3}$$

The use of multiple categories results in the total acceptance $\times$ efficiency being approximately flat as a function of κ_λ as shown in 5.9.

Higgs	$M_X^* \geq$ 350 GeV b-tag tight		$M_X^* \geq$ 350 GeV b-tag loose		$M_X^* <$ 350 GeV b-tag tight		$M_X^* <$ 350 GeV b-tag loose	
	Yield	$A \times \epsilon(\%)$	Yield	$A \times \epsilon(\%)$	Yield	$A \times \epsilon(\%)$	Yield	$A \times \epsilon(\%)$
SM HH	0.766	6.26	0.346	2.82	0.13	1.07	0.0608	0.497
ggH	0.31	0.00202	0.547	0.00357	0.381	0.00249	1.25	0.00818
VBF	0.0298	0.0025	0.0556	0.00466	0.0495	0.00415	0.145	0.0121
WH	0.0083	0.00199	0.11	0.0262	0.0161	0.00395	0.216	0.0511
ZH	0.541	0.226	0.392	0.164	0.489	0.204	0.504	0.21
ggZH	0.239	0.618	0.174	0.45	0.0322	0.0833	0.0424	0.11
ttH	0.601	0.376	0.656	0.411	1.95	1.22	1.71	1.07
bbH	0.013	0.00849	0.0105	0.00681	0.127	0.083	0.138	0.0897
tHW	0.0161	0.337	0.0243	0.508	0.0168	0.352	0.029	0.607
$tHjb$	0.0464	0.198	0.0485	0.207	0.103	0.439	0.0978	0.418

Table 5.6 *The yields and efficiences in each category for the HH signal and the single Higgs boson backgrounds for the 2D selection.*

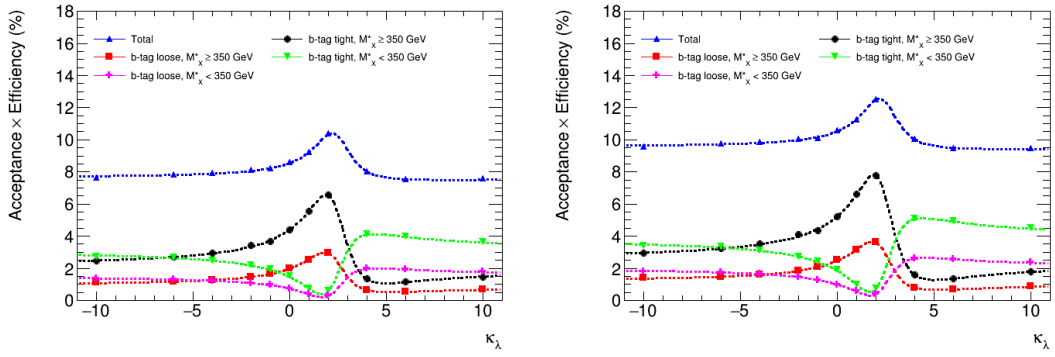

Figure 5.9 *The acceptance × efficiency as a function of κ_λ for the 1D (left) and 2D (right) selections.*

5.6 Signal and background modelling

An unbinned maximum likelihood fit is performed to the mass of the $H \to \gamma\gamma$ candidates in the case of the 1D fit strategy, and to the mass of the $H \to \gamma\gamma$ and $H \to b\bar{b}$ candidates for the 2D fit. Probability density functions (PDFs) are therefore required to model the signal and background processes.

In this section, the following notation is used on a number of figures:

- C4: b-tag tight, $M_X^* \geq 350$ GeV
- C3: b-tag loose, $M_X^* \geq 350$ GeV
- C2: b-tag tight, $M_X^* < 350$ GeV
- C1: b-tag loose, $M_X^* < 350$ GeV

5.6.1 Signal modelling

A Higgs boson decaying to $\gamma\gamma$ results in a narrow signal peak in $m_{\gamma\gamma}$ due to the excellent photon energy resolution of the ATLAS detector. The $m_{\gamma\gamma}$ signal distribution is modelled with a double-sided Crystal Ball (DSCB) function which consists of a Gaussian core and two independent power-law tails. The double-sided Crystal Ball function is defined as:

$$f_i^{\text{sig}}(m_{\gamma\gamma}; \Delta\mu_{\text{CB},i}, \sigma_{\text{CB},i}, \alpha^{\pm}_{\text{CB},i}, n^{\pm}_{\text{CB},i}) = \mathcal{N}_c \begin{cases} e^{-t^2/2} & -\alpha^{-}_{\text{CB},i} \leq t \leq \alpha^{+}_{\text{CB},i} \\ \dfrac{\left(\frac{n^{-}_{\text{CB},i}}{|\alpha^{-}_{\text{CB},i}|}\right)^{n^{-}_{\text{CB},i}} e^{-|\alpha^{-}_{\text{CB},i}|^2/2}}{\left(\frac{n^{-}_{\text{CB},i}}{\alpha^{-}_{\text{CB},i}} - \alpha^{-}_{\text{CB},i} - t\right)^{n^{-}_{\text{CB},i}}} & t < -\alpha^{-}_{\text{CB},i} \\ \dfrac{\left(\frac{n^{+}_{\text{CB},i}}{|\alpha^{+}_{\text{CB},i}|}\right)^{n^{+}_{\text{CB},i}} e^{-|\alpha^{+}_{\text{CB},i}|^2/2}}{\left(\frac{n^{+}_{\text{CB},i}}{\alpha^{+}_{\text{CB},i}} - \alpha^{+}_{\text{CB},i} - t\right)^{n^{+}_{\text{CB},i}}} & t > \alpha^{+}_{\text{CB},i} \end{cases}, \tag{5.4}$$

where $t = (m_{\gamma\gamma} - m_H - \Delta\mu_{\text{CB},i})/\sigma_{\text{CB},i}$, and $\mathcal{N}_c$ is a normalization factor. The non-Gaussian parts are parameterised by $\alpha^{\pm}_{\text{CB},i}$ and $n^{\pm}_{\text{CB},i}$ separately for the low- ($-$) and high-mass ($+$) tails. The two independent tails make this a suitable choice of functional form since the $H \to \gamma\gamma$ mass peak has asymmetric tails.

The fits are performed simultaneously to the HH signal and the single Higgs backgrounds in each category since they have the same resonant $H \to \gamma\gamma$ shape. The fits performed in each category are shown in Figure 5.10. The resolution, σ_{CB} is significantly better in the $M_X^* \geq 350$ GeV categories since the photons have higher p_T in these categories.

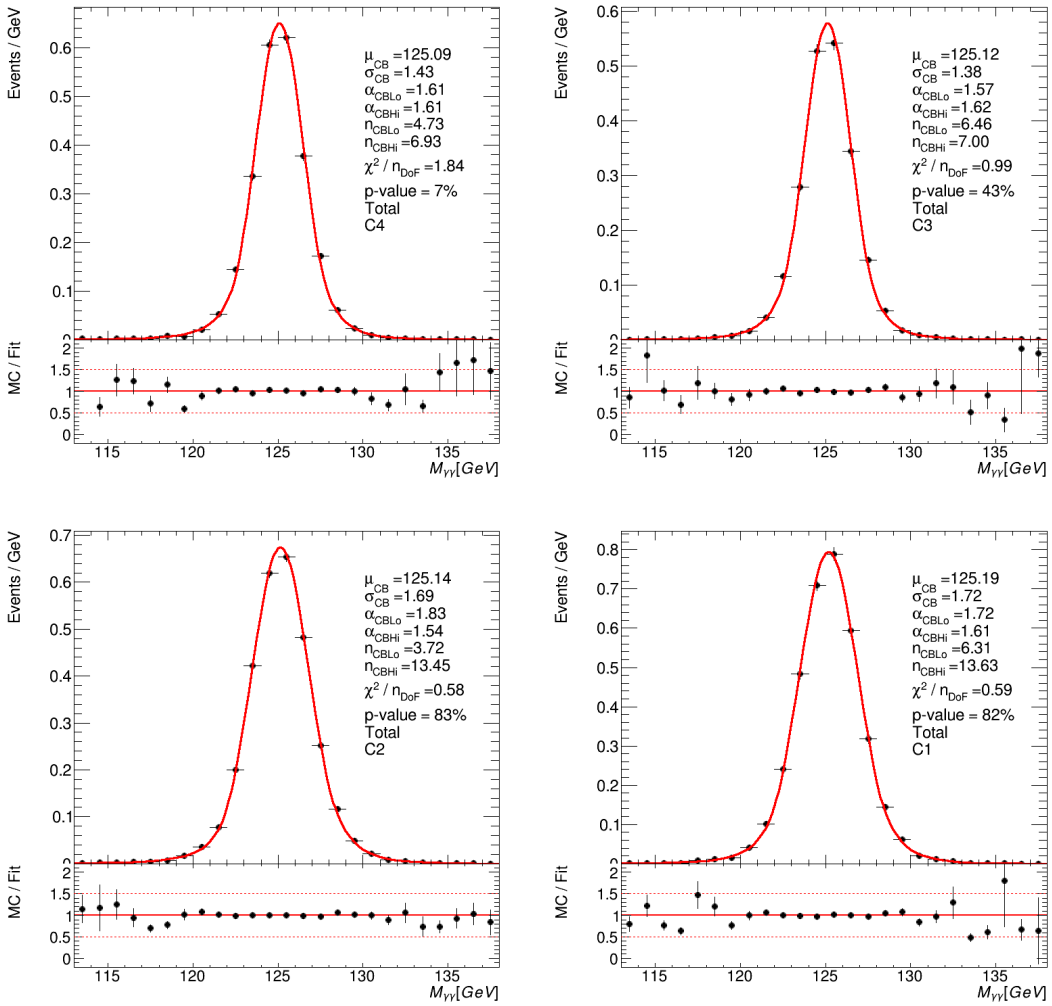

Figure 5.10 *The fitted models used to model $m_{\gamma\gamma}$ in the HH signal and the single Higgs backgrounds in each category.*

For the 2D fit strategy, PDFs are also required to model the invariant mass of the $H \to b\bar{b}$ candidates. The DSCB function is used to model this distribution in the HH signal. The fits performed in each category are shown in Figure 5.11.

For the $M_X^* < 350$ GeV categories, there is a lack of events in the SM HH MC sample which results in the fit parameters having large uncertainties. The lack of events is due to the low signal efficiency (approximately 1%) in these categories. Some of the BSM κ_λ samples have significantly higher signal efficiency in these

categories. The fit parameters in the $M_X^* < 350$ GeV categories are therefore obtained by performing fits to the $\kappa_\lambda = +10$ MC sample.

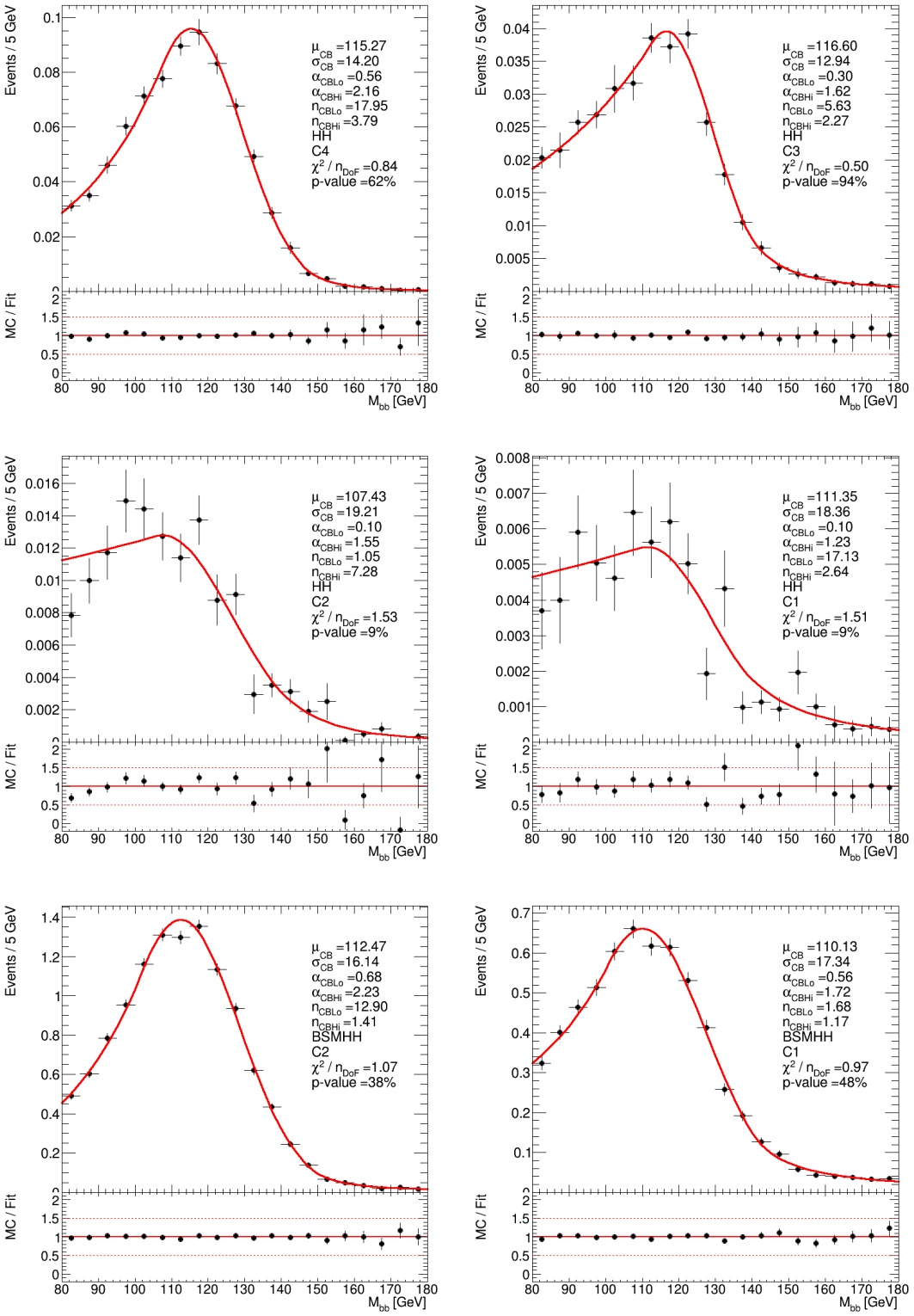

Figure 5.11 *The fits performed to m_{bb} in the HH signal samples in all categories. Due to the large statistical uncertainties in the middle figures, the fit parameters for the $M_X^* < 350$ GeV categories are determined by performing fits to a BSM sample with $\kappa_\lambda = +10$ (Bottom).*

5.6.2 Background modelling

Continuum (1D)

The continuum background is largely made up of $\gamma\gamma$ + jets production; a leading order Feynman diagram for this process is shown in Figure 5.12. There are also small contributions from γj + jets and multijets with jets faking photons. There is also a small $t\bar{t}\gamma$ contribution with the second photon arising from an electron faking a photon.

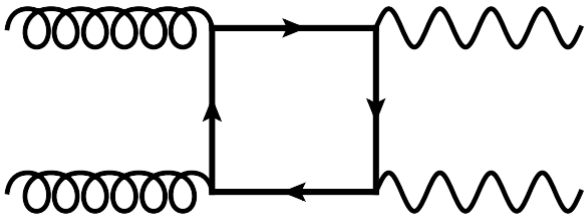

Figure 5.12 *A leading order Feynman diagram for $gg \to \gamma\gamma$ production [98]*

Figure 5.13 shows the comparison between data and simulation for $m_{\gamma\gamma}$ for the 1D selection in all categories. The agreement in terms of the normalisation is poor due to the missing γj + jets component and there are also large statistical uncertainties.

Functional forms are used to model the continuum background. The background shape will not be completely flexible due to the choice of functional form and this will induce a bias in the estimation of the background below the signal peak. This bias can be made smaller with more complex (and therefore more flexible) functional forms but this comes at the price of increased statistical uncertainties and it can also result in overfitting when there is a small number of events in the data.

The bias can be estimated by constructing background only templates with small statistical uncertainties. These templates are fitted to a signal plus background model where the background model is the functional form under study and the signal model is as described in Section 5.6.1. The fits are perfomed by modifying the signal peak position, μ_{CB}, in 1 GeV steps from 121 to 129 GeV ($\pm$ 4 GeV around the measured mass of the Higgs boson) and performing a fit at each step. The maximum signal fitted over this range is taken as the bias due to the choice of functional form, often referred to as the *spurious signal*.

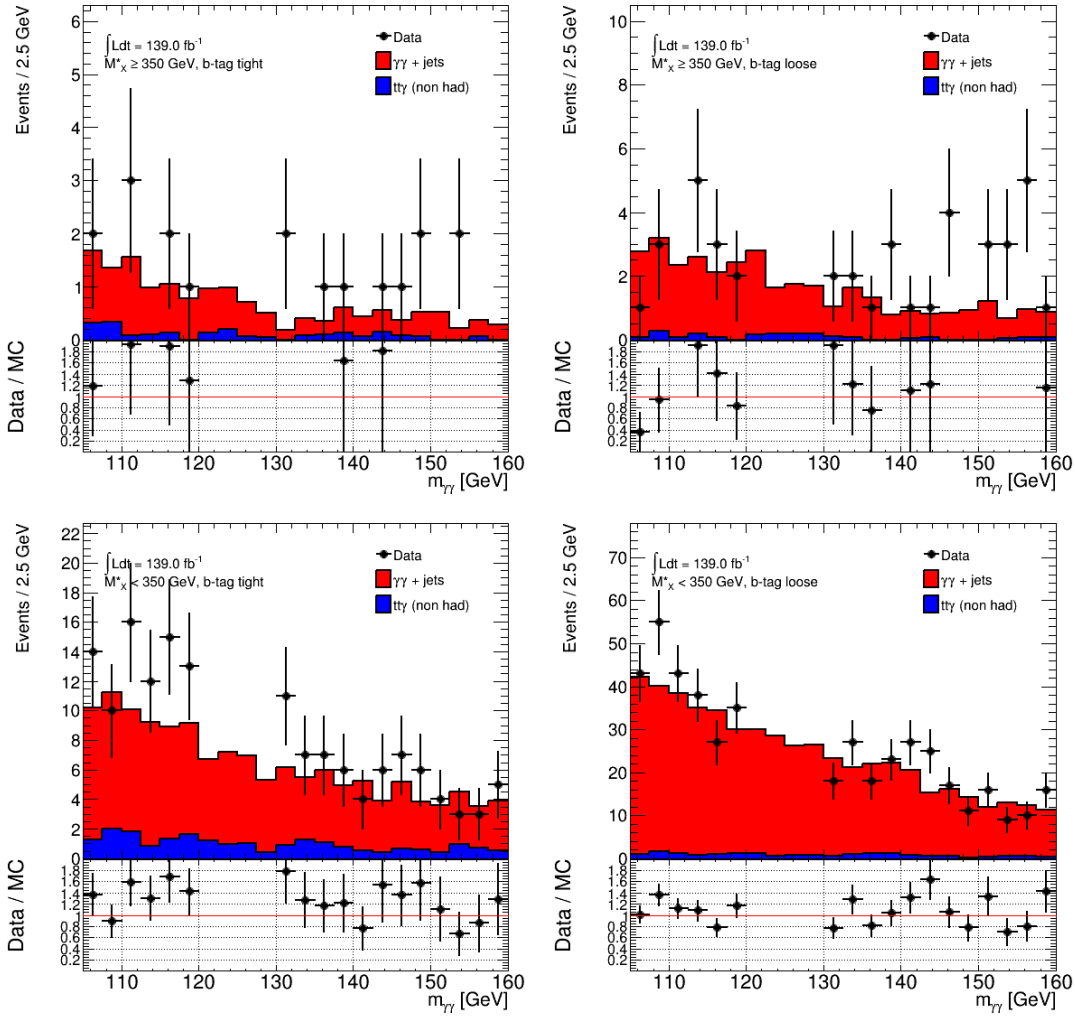

Figure 5.13 *Comparison between data and simulation for $m_{\gamma\gamma}$ in all categories.*

The function selected is required to have the ratio of the spurious signal to the statistical uncertainty on the fitted background, Z_{spur} as less than 20%. In the event more than one function satisfies this criteria, the function with the fewest number of degrees of freedom is selected and in the event no function satisfies this criteria, the function with the smallest Z_{spur} is selected. The criteria of 20% is tight enough so that it should lead to small modelling uncertainties but it should also be loose enough so that reasonably simple functions with smaller statistical uncertainties can still pass.

The background template using the MC directly has statistical uncertainties which are too large for it to be useful for the spurious signal test. However, it is expected that there is little correlation between the $m_{\gamma\gamma}$ shape and the b-tag working points. The $\gamma\gamma$ MC sample is therefore used with the same event selection but without the b-tag requirements and the sample is then normalised to the yield in the data sidebands ($m_{\gamma\gamma} \in [105,160]$ GeV excluding $m_{\gamma\gamma} \in [120,130]$ GeV) with an additional scale factor of (55.0/45.0) since 10 GeV of data is blind. The agreement between data and the templates are shown in Figure 5.14.

Since the number of events in data is small in all categories, to reduce the risk of overfitting, only the following functions are considered:

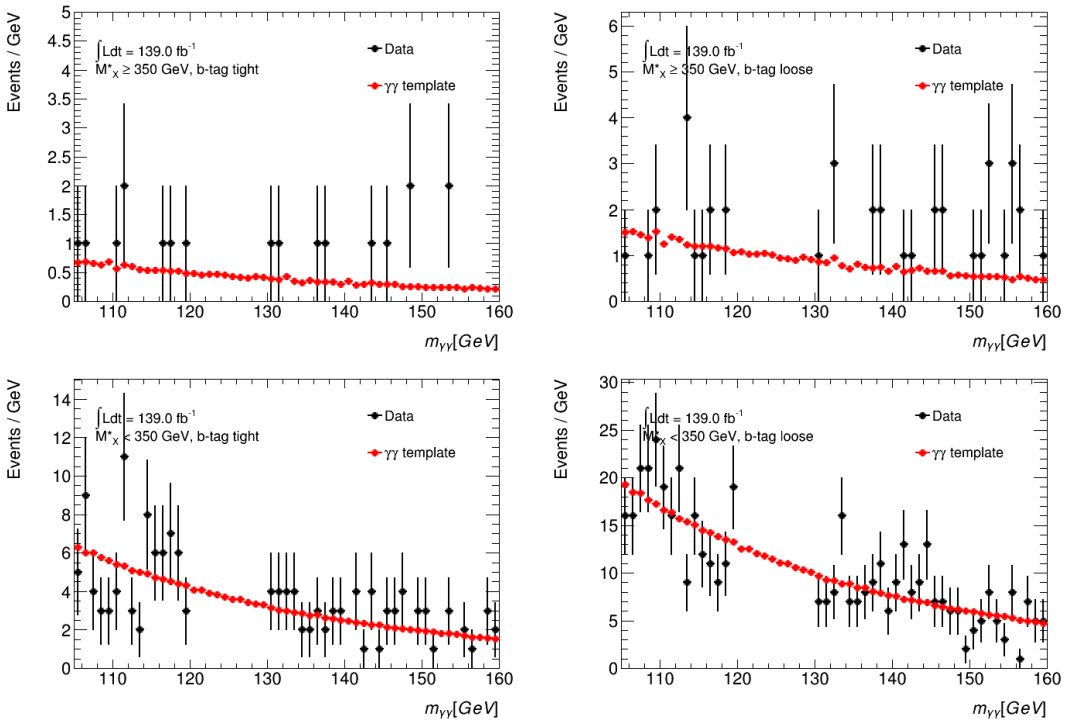

Figure 5.14 *Comparison between data and the templates constructed from MC for $m_{\gamma\gamma}$ in all categories.*

b-tag tight, $M_X^* \geq 350$ GeV				
Model	N_{spur}	Z_{spur} [%]	$nPars$	χ^2/ndof
Exponential	0.024	18.15	1	1.35
Exp. of 2$^{\text{nd}}$ order	0.02	15.17	2	1.35
Power Law	0.07	54.05	1	1.34
b-tag loose, $M_X^* \geq 350$ GeV				
Model	N_{spur}	Z_{spur} [%]	$nPars$	χ^2/ndof
Exponential	0.04	12.07	1	1.35
Exp. of 2$^{\text{nd}}$ order	0.04	11.97	2	1.35
Power Law	0.15	52.80	1	1.34
b-tag tight, $M_X^* < 350$ GeV				
Model	N_{spur}	Z_{spur} [%]	$nPars$	χ^2/ndof
Exponential	-0.11	-41.62	1	1.80
Exp. of 2$^{\text{nd}}$ order	0.12	40.62	2	0.82
Power Law	0.45	169.26	1	2.05
b-tag loose, $M_X^* < 350$ GeV				
Model	N_{spur}	Z_{spur} [%]	$nPars$	χ^2/ndof
Exponential	-0.29	-40.48	1	1.80
Exp. of 2$^{\text{nd}}$ order	0.39	39.97	2	0.82
Power Law	1.43	179.68	1	2.05

Table 5.7 *Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of background events, Z_{spur}, for each category. The χ^2/ndof from performing a background-only fit is also shown. The number of degrees of freedom for each function is given by $nPars$.*

- Exponential, including exponential of 2$^{\text{nd}}$ order polynomial. These functions take the form:
  $\text{Exp}^N(m_{\gamma\gamma};\theta^{bkg}) = \exp\left(\sum_{j=0}^{N}\theta_j^{bkg}\cdot m_{\gamma\gamma}^j\right)$
- First-order Power Law function: $\text{Pow}(m_{\gamma\gamma};\theta^{bkg}) = m_{\gamma\gamma}^{\theta_1^{bkg}}$

Table 5.7 shows the results for these tests in all categories. In the $M_X^* \geq 350$ GeV categories, an exponential with one degree of freedom is used while in the $M_X^* < 350$ GeV categories, an exponential of 2$^{\text{nd}}$ order polynomial is used.

Continuum (2D)

For the 2D fit, the background only templates are constructed using the same methodology as in the 1D case. Figure 5.15 shows the 2D templates used in each category while Figures 5.16 and 5.17 show comparisons between the projections of the 2D templates and data for $m_{\gamma\gamma}$ and m_{bb} respectively.

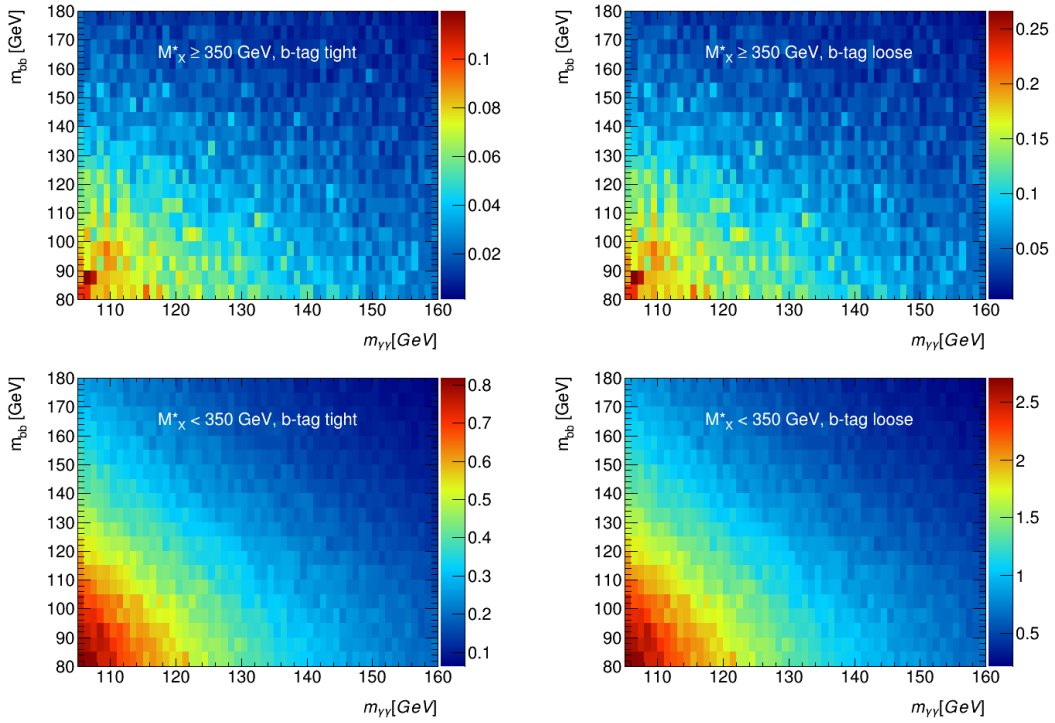

Figure 5.15 *The background only templates used in each category.*

It is assumed to begin with that the 2D continuum PDF can be modelled using a product of 1D PDFs and the correlation is then studied later. The functions tested are limited to products of exponentials of first and second order polynomials. Table 5.8 shows the results for these tests in all categories. In all categories, an exponential of first order is chosen to model the $m_{\gamma\gamma}$ spectrum while an exponential of second order polynomial is used to model the m_{bb} spectrum.

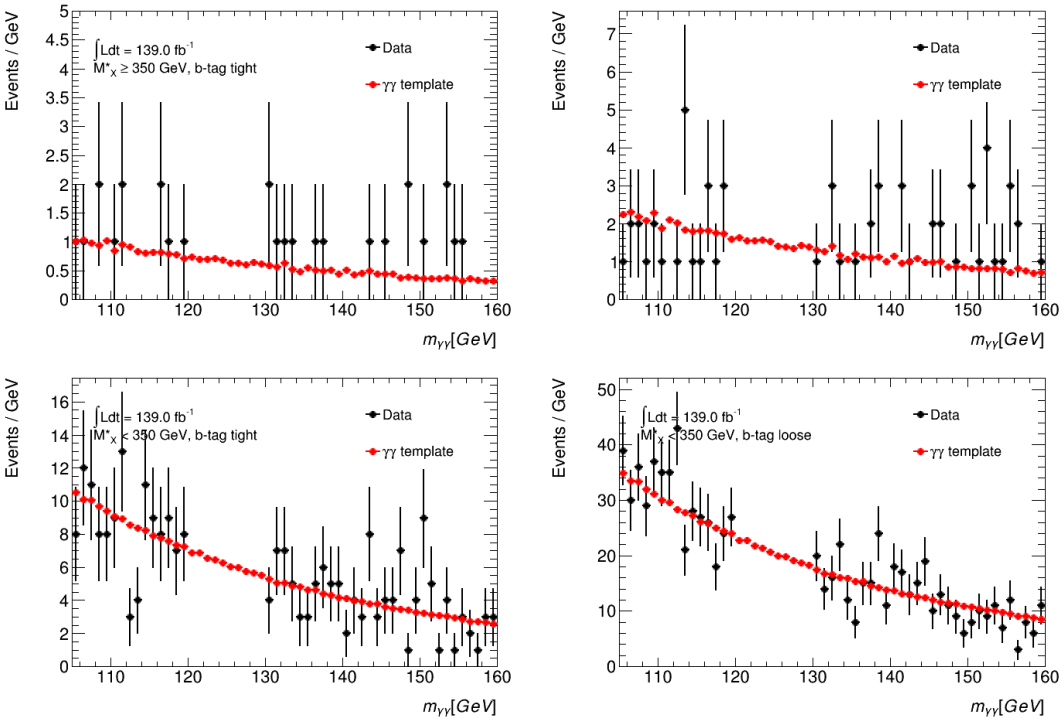

Figure 5.16 *Comparison between data and the templates constructed from MC for $m_{\gamma\gamma}$ in all categories.*

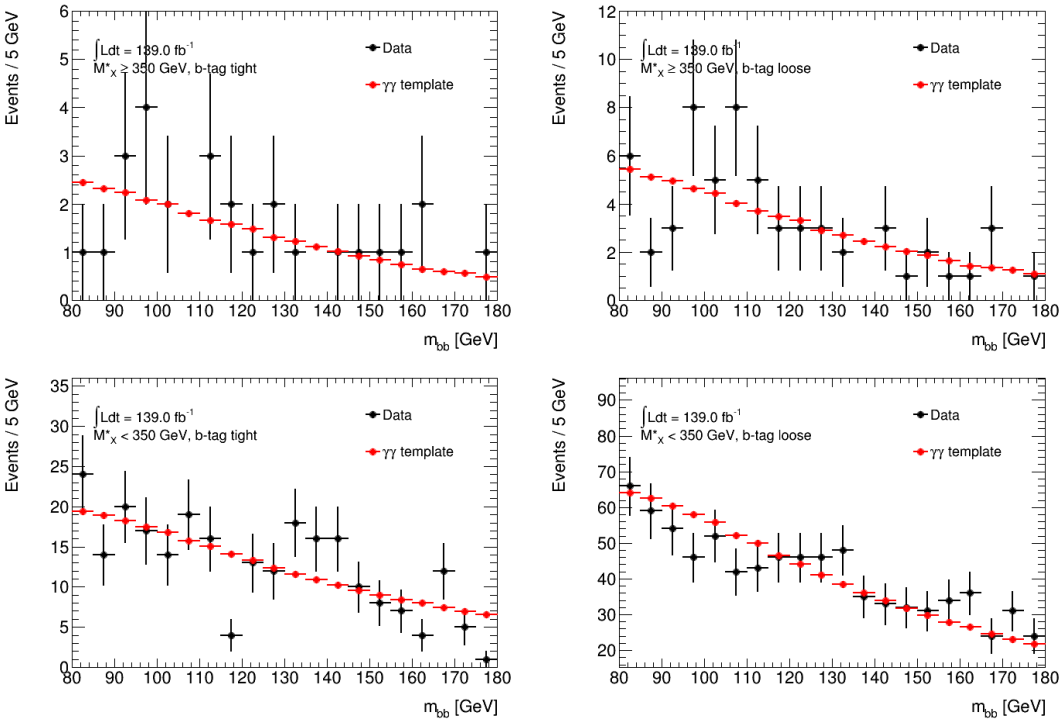

Figure 5.17 *Comparison between data and the templates constructed from MC for m_{bb} in all categories.*

b-tag tight, $M_X^* \geq 350$ GeV				
Model	N_{spur}	Z_{spur} [%]	$nPars$	χ^2/ndof
E1E1	0.07	38.14	2	1.52
E1E2	−0.006	−2.79	3	1.36
E2E1	0.094	45.8	3	1.52
E2E2	0.004	1.83	4	1.36
b-tag loose, $M_X^* \geq 350$ GeV				
E1E1	0.12	27.16	2	1.52
E1E2	−0.06	−13.37	3	1.36
E2E1	0.157	35.58	3	1.52
E2E2	−0.02	−4.72	4	1.36
b-tag tight, $M_X^* < 350$ GeV				
E1E1	0.44	103.16	2	1.43
E1E2	0.04	9.31	3	1.18
E2E1	0.70	162.25	3	1.38
E2E2	0.33	79.32	4	1.13
b-tag loose, $M_X^* < 350$ GeV				
E1E1	1.41	100.33	2	1.43
E1E2	0.17	12.05	3	1.18
E2E1	2.50	149.28	3	1.38
E2E2	1.28	74.55	4	1.13

Table 5.8 *Number of spurious signal events, N_{spur}, and its ratio to the statistical uncertainty on the fitted number of signal events, Z_{spur}, for each category. The χ^2/ndof from performing a background-only fit is also shown. The number of degrees of freedom for each function is given by $nPars$.*

Single Higgs

The single Higgs boson backgrounds have the same resonant shape in $m_{\gamma\gamma}$ as the signal but have a different shape to both the HH signal and the continuum background for m_{bb} as can be seen in Figure 5.18. For the 2D fit, it is necessary to derive PDFs to model these contributions.

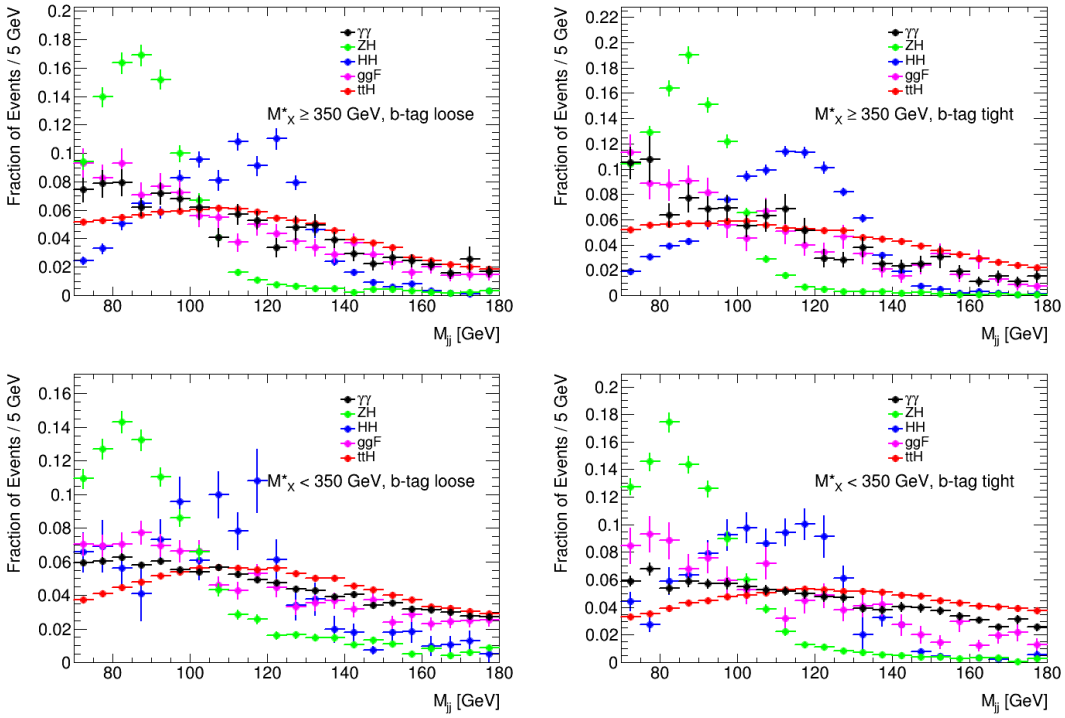

Figure 5.18 *The m_{bb} distributions in the HH signal and the $\gamma\gamma$ + jets, ggF, ZH and $t\bar{t}H$ backgrounds in each category.*

For the $q\bar{q} \to ZH$ and the $gg \to ZH$ processes, the DSCB function is used to model the m_{bb} spectrum. The $t\bar{t}H$ process is modelled with Chebychev polynomials of the third order, defined as:

$$1 + a_0 x + a_1(2x^2 - 1) + a_2(4x^3 - 3x) \tag{5.5}$$

The ggF process is modelled with an exponential with one degree of freedom. Figures 5.19, 5.20 and 5.21 show the fits to m_{bb} in these samples for all categories.

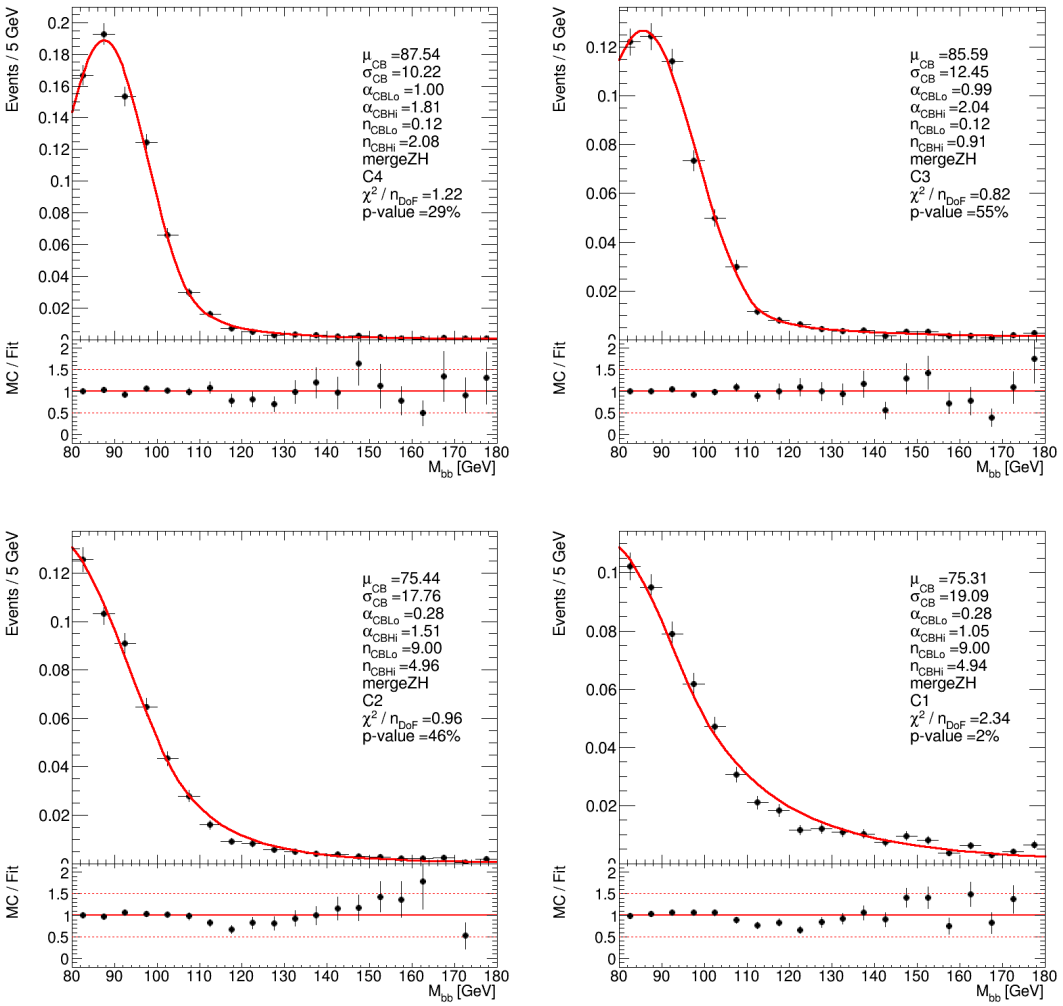

Figure 5.19 *Fits to m_{bb} in the ZH samples for all categories.*

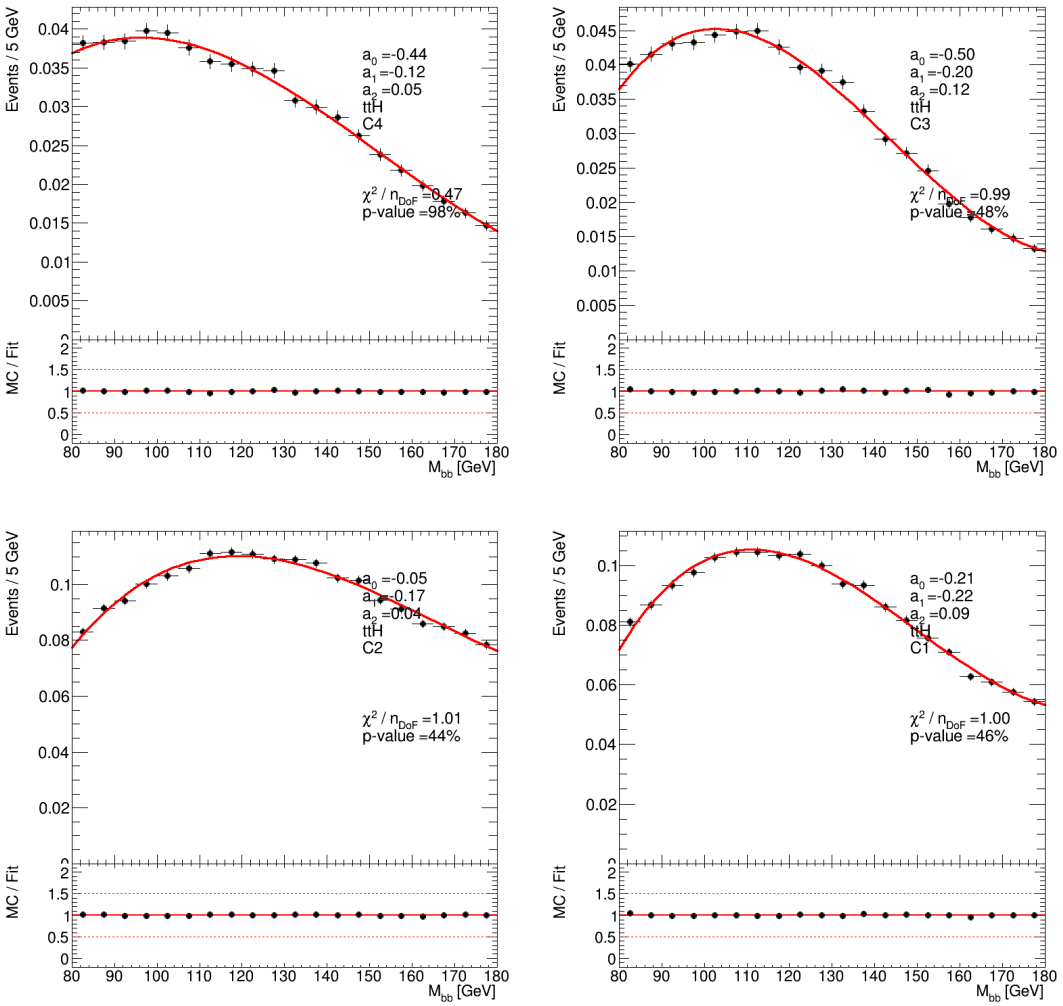

Figure 5.20 *Fits to m_{bb} in the $t\bar{t}H$ samples for all categories.*

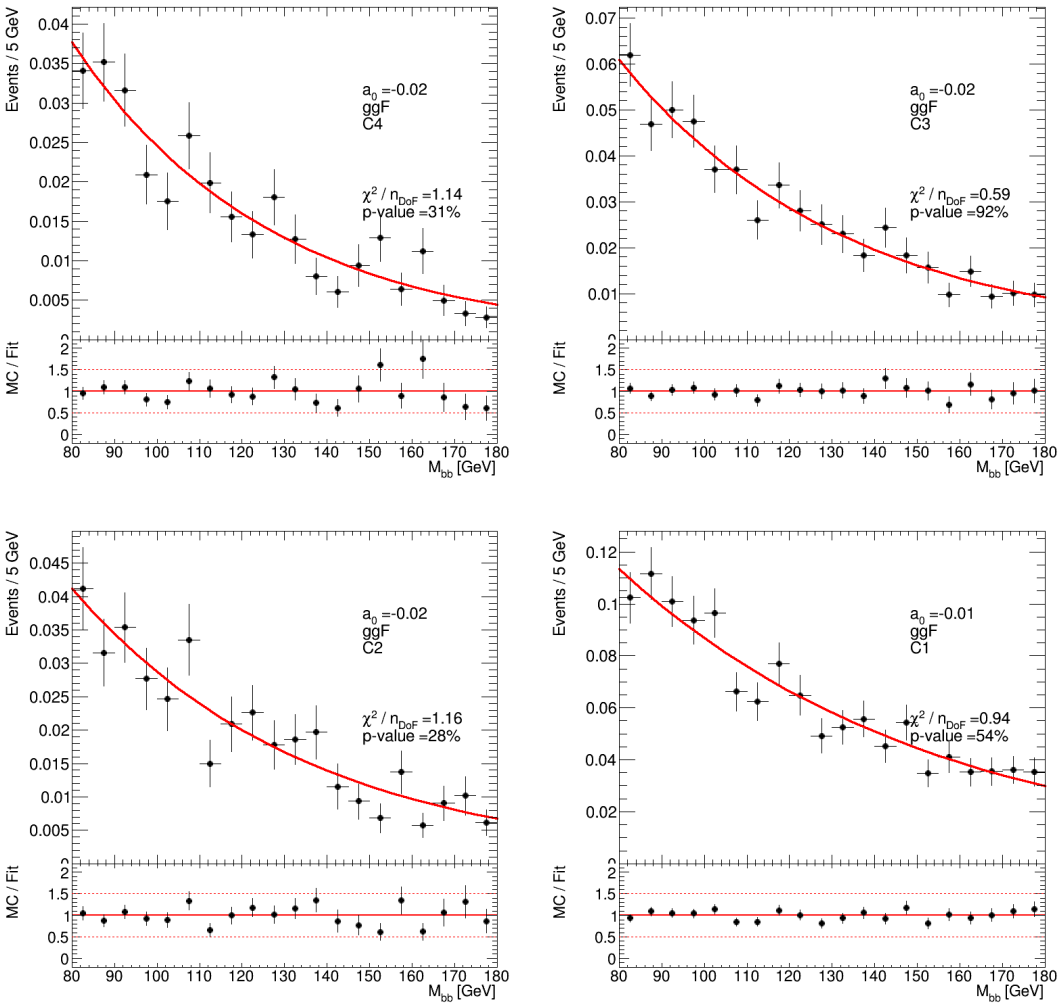

Figure 5.21 *Fits to m_{bb} in the ggF samples for all categories.*

Correlations

So far it has been assumed that it is possible to model the 2D PDFs by using a product of 1D PDFs. To study this assumption, the difference between the 2D MC prediction and the 2D PDFs derived are studied with the following quantity:

$$R_{ij} = \frac{\text{Yield}_{ij}^{MC} - \text{Yield}_{ij}^{PDF}}{\Delta_{ij}^{MC}} \tag{5.6}$$

where i and j are bins in the $m_{\gamma\gamma} - m_{bb}$ plane and Δ^{MC} is the statistical uncertainty from the MC sample.

This computation is done for the HH signal and the ggF, ZH, $t\bar{t}H$ and the $\gamma\gamma$ + jets backgrounds. Figures 5.22 to 5.26 show the results for the b-tag tight, $M_X^* \geq 350$ GeV category. It can be seen there are no correlations or structures within the uncertainties and therefore the assumption that the 2D PDFs can be modelled using the product of the 1D PDFs is sufficient for the precision of the current analysis. The same conclusion is found for all other categories in the analysis.

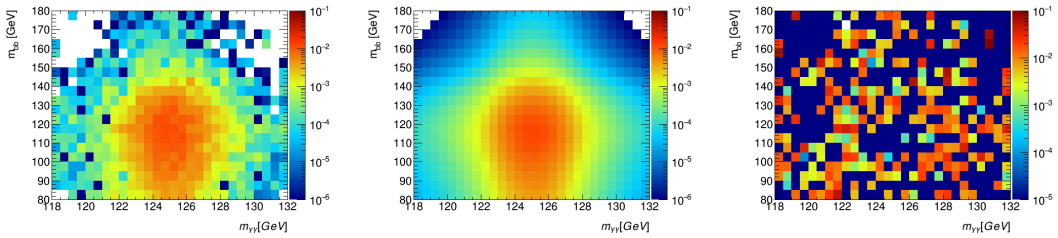

Figure 5.22 *The 2D distribution of $m_{\gamma\gamma}$ - m_{bb} for the HH MC (left), the PDF model derived (middle) and the residuals between the MC and the PDFs (right).*

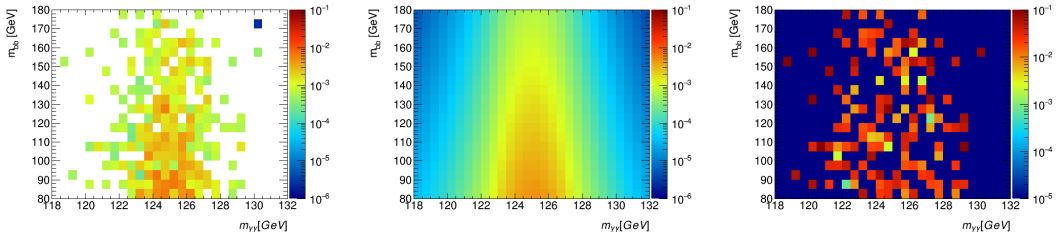

Figure 5.23 *The 2D distribution of $m_{\gamma\gamma}$ - m_{bb} for the ggF MC (left), the PDF model derived (middle) and the residuals between the MC and the PDFs (right).*

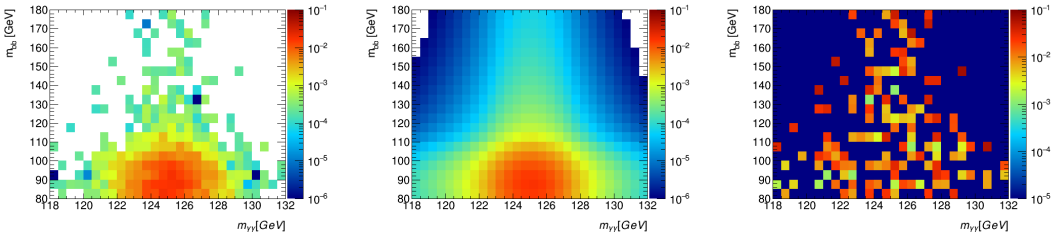

Figure 5.24 *The 2D distribution of $m_{\gamma\gamma}$ - m_{bb} for the ZH MC (left), the PDF model derived (middle) and the residuals between the MC and the PDFs (right).*

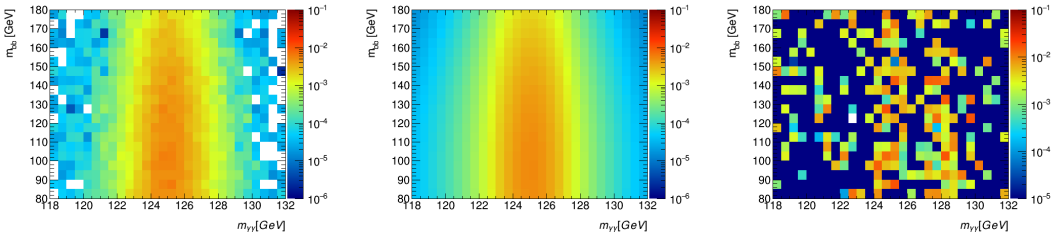

Figure 5.25 *The 2D distribution of $m_{\gamma\gamma}$ - m_{bb} for the $t\bar{t}H$ MC (left), the PDF model derived (middle) and the residuals between the MC and the PDFs (right).*

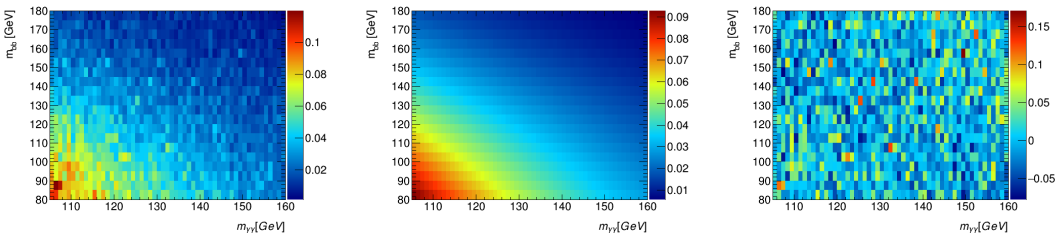

Figure 5.26 *The 2D distribution of $m_{\gamma\gamma}$ - m_{bb} for the $\gamma\gamma$ MC (left), the PDF model derived (middle) and the residuals between the MC and the PDFs (right).*

5.7 Systematic uncertainties

Theoretical uncertainties are applied to the normalisation of the single Higgs boson backgrounds. The ggF, VBF and WH processes are assigned a 100% uncertainty due to the radiation of additional heavy flavour jets (*Heavy Flavour*). This is supported by measurements of the $t\bar{t}+b\bar{b}$ process [99]. This heavy flavour uncertainty is not assigned to the other production mechanisms. For the other production mechanisms (except bbH), two sources of uncertainty are considered: the choice of QCD scales and PDF + α_S uncertainties. For the bbH process, a single source of uncertainty is considered which combines the QCD scales, PDF + α_S and an uncertainty on the b-quark mass. An uncertainty on the $H \to \gamma\gamma$ branching ratio is also applied to these processes.

Theoretical uncertainties are applied to the normalisation of the SM HH signal. The impact from the QCD scale and the PDF + α_S uncertainties is $^{+4.5\%}_{-6.0\%}$ and $\pm 2.3\%$ respectively [79]. An uncertainty related to the inclusion of m_t into the cross section calculations is included, as well as uncertainties on the $H \to \gamma\gamma$ and $H \to b\bar{b}$ branching ratios. All theoretical uncertainties are taken from Ref. [79].

The choice of functional forms to model the continuum background introduces a potential bias (*spurious signal*) into the fit. The background modelling studies (Section 5.6.2) aim to minimise the spurious signal and quantify the impact of a given background model on the number of fitted signal events. The spurious signal is not corrected for but is included in the likelihood as a term with the same shape as the signal and a Gaussian constraint. The values of the spurious signal in each category are listed in Table 5.7 for the 1D selection and Table 5.8 for the 2D selection.

The experimental uncertainties from the photons and jets are provided by the ATLAS Combined Performance groups. All uncertainties are applied to the normalisation with the exception of the photon energy scale and photon energy resolution which are applied to the peak and resolution of $m_{\gamma\gamma}$ for the 1D selection. The photon energy scale and resolution uncertainties are computed by fitting the DSCB function to the $m_{\gamma\gamma}$ distribution with the scale and resolution changed by $\pm 1\ \sigma$ and the relative difference of the fit parameters is taken as the uncertainty. Figure 5.27 shows the nominal $m_{\gamma\gamma}$ distribution and the distributions with the photons scaled by the scale and resolution uncertainties. For the 2D selection, the jet energy scale and jet energy resolution are applied to the peak and resolution

of m_{bb}. The experimental uncertainties vary depending on the analysis category and signal or background process. In Table 5.9, all experimental uncertainties are quoted on the SM HH signal and the maximum across all four categories is taken for the 1D selection.

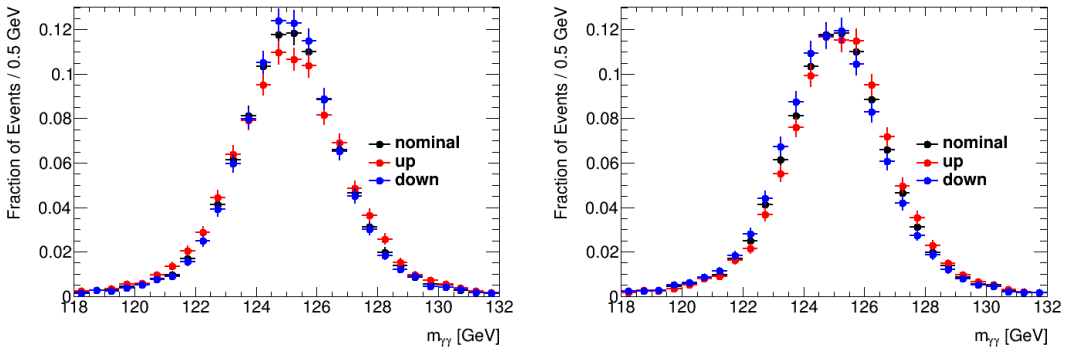

Figure 5.27 *Black curve shows the nominal $m_{\gamma\gamma}$ distribution and the coloured curves show the same distribution but with the photon resolution (Left) and photon scale (Right) uncertainties applied after the $H \to \gamma\gamma$ selection.*

5.8 Statistical analysis

Two different strategies are employed for the statistical analysis. For the 1D selection, a fit to the $m_{\gamma\gamma}$ distribution is performed while for the 2D selection a fit is performed to the $m_{\gamma\gamma}$ and m_{bb} distributions. Extended likelihood functions are built from the number of events observed and the $m_{\gamma\gamma}$ ($m_{\gamma\gamma}$ and m_{bb}) values observed for the 1D (2D) selection. In what follows the likelihood is defined for the 1D selection but it can be generalised to 2D. The likelihood for a given category c in the analysis is given by:

$$\mathcal{L}_c = \text{Pois}(n_C | N_c(\boldsymbol{\theta})) \cdot \prod_{i=1}^{n_c} f_c(m^i_{\gamma\gamma}, \boldsymbol{\theta}) \cdot G(\boldsymbol{\theta}) \tag{5.7}$$

where $\boldsymbol{\theta}$ are nuisance parameters representing systematic uncertainties which are constrained by Gaussian constraint terms $G(\boldsymbol{\theta})$, n_c is the observed number of data events and N_C is defined as the sum of the HH signal, single Higgs background and continuum background:

$$N_C = \mu_{HH} \cdot N_{\text{HH},c} + N_{\text{Higgs},c} + N_{\text{bkg},c} + N_{spur,c} \cdot \boldsymbol{\theta}_{\text{spur,c}} \tag{5.8}$$

Source of systematic uncertainty		% effect *w.r.t.* nominal
Luminosity		± 1.7
Trigger		± 0.4
Pile-up modelling		± 2.6
Photon	identification	± 1.3
	isolation	± 1.4
	energy resolution	± 10.6
	energy scale	± 0.4
Jet	energy resolution	± 4.8
	energy scale	± 1.8
Flavour tagging	b-jets	± 2.4
	c-jets	± 1.6
	light-jets	± 1.1
Theory (SM HH)	PDF+α_S	± 2.3
	Scale	$^{+4.3}_{-6.0}$
	EFT	± 5.0
	$\mathcal{BR}(H \to \gamma\gamma)$	± 2.1
	$\mathcal{BR}(H \to b\bar{b})$	± 1.3
Theory (SM Higgs)	Heavy Flavour	± 100.0
	Scale	$\pm 0.3 - 25.1$
	PDF+α_S	$\pm 1.9 - 9.2$
	bbH uncertainty	$^{+20.1}_{-23.9}$
	$\mathcal{BR}(H \to \gamma\gamma)$	± 2.1

Table 5.9 *Summary of the systematic uncertainties in the analysis. For the experimental uncertainties, only uncertainties on the HH signal are presented with the maximum uncertainty taken over all categories for the 1D selection for presentation purposes.*

where μ_{HH} is the signal strength defined as the ratio of the measured cross section to the SM expectation and $N_{spur,c}$ is the number of spurious signal events for a category c.

Apart from the spurious signal, systematic uncertainties are incorporated into the likelihood by multiplying the relevant parameter of the statistical model by a response function. In the case of a Gaussian PDF for an uncertainty of size σ, it is given as

$$F_G(\sigma, \theta) = (1 + \sigma \cdot \theta) \tag{5.9}$$

and for cases where a negative model parameter does not make physical sense,

the log-normal PDF is used instead, given as

$$F_{LN}(\sigma, \theta) = e^{\ln(1+\sigma)\theta} \tag{5.10}$$

In both cases the corresponding constraint product $G(\boldsymbol{\theta})$ is a unit Gaussian centered at zero for θ.

In this analysis, multiple categories are used. Many of the systematic uncertainties are correlated across the categories: the uncertainties share the same nuisance parameter and also the same constraint PDF in the likelihood. The magnitude and sign of the uncertainties are allowed to differ in each category.

The HH signal and the single Higgs background models are parameterised by the DSCB function as described in section 5.6.1. The single Higgs boson backgrounds are all fixed to their SM expectations in the fit. Functional forms are used to model the continuum background in each category. The value of $f_c(m^i_{\gamma\gamma})$ is the full signal and background probability density function evaluated for each $m^i_{\gamma\gamma}$ candidate. It can be written as:

$$\begin{aligned} f_c(m^i_{\gamma\gamma}) =& [(\mu \cdot N_c^{\mathrm{HH}} + N_c^{\mathrm{Higgs}} + N_c^{\mathrm{spur}} \cdot \theta_c^{\mathrm{spur}}) \cdot f_c^{\mathrm{sig}}(m^i_{\gamma\gamma}, \theta_c^{\mathrm{sig}}) \\ &+ N_c^{\mathrm{bkg}} \cdot f_c^{\mathrm{bkg}}(m^i_{\gamma\gamma}, \theta_c^{\mathrm{bkg}})]/N_c, \end{aligned} \tag{5.11}$$

The full likelihood for the analysis is a product of four likelihoods, one for each category.

A description of the statistical tests performed in the analysis are given in Appendix B.

5.8.1 Results

1D Fit Results

Figure 5.28 shows the $m_{\gamma\gamma}$ spectra observed in data and the background only fits in each category.

The best fit signal strength is $\mu_{HH} = 3.3 \pm 3.9$ where all single Higgs backgrounds are fixed to their SM expectation. The significance for SM HH production with respect to the background only expectation is observed (expected) to be $0.94\,\sigma$

$(0.33\,\sigma)$. 95% CL upper limits on μ_{HH} and κ_λ are determined using the CL_s prescription [100] using the asymptotic approximation [97].

Figure 5.29 shows the expected and observed upper limits in each category and from combining all categories for SM HH production. The expected and observed upper limits are also shown as a function of κ_λ. The observed (expected) limit for the SM HH cross section is 11.4 (7.7) times the SM value while the ratio of the Higgs boson self-coupling to its SM prediction, κ_λ is observed (expected) to be constrained at 95% CL to $-5.3 < \kappa_\lambda < 10.5$ $(-4.3 < \kappa_\lambda < 10.4)$.

For κ_λ in the interval [4,7], the observed limit is better than the expected while the observed limit is worse than the expected limit for all other values of κ_λ. This is a result of a downwards fluctuation in data in the $M_X^* < 350$ GeV categories and the particularly soft kinematics that HH production has when κ_λ is in the interval [4,7].

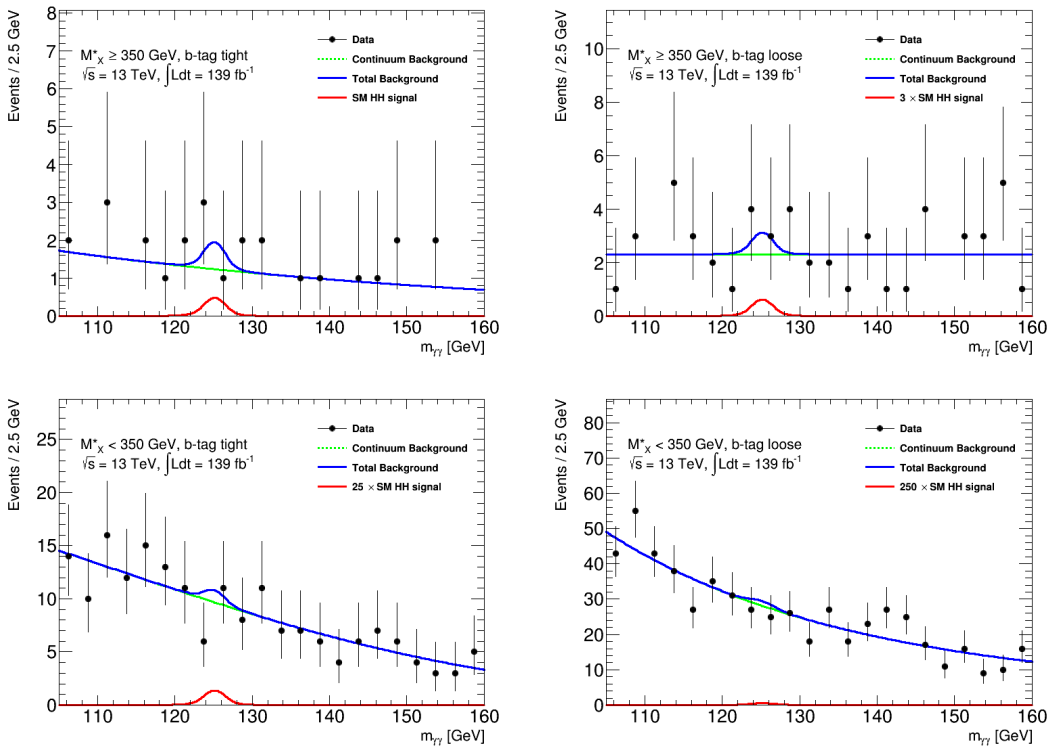

Figure 5.28 *The $m_{\gamma\gamma}$ spectrum observed in each category for the 1D selection. The background only fits are shown with a solid blue line.*

Figure 5.35 shows the observed and expected 95% CL upper limits on μ_{HH} with and without systematic uncertainties. It can be seen there is only a small impact on the sensitivity of the analysis when including systematic uncertainties.

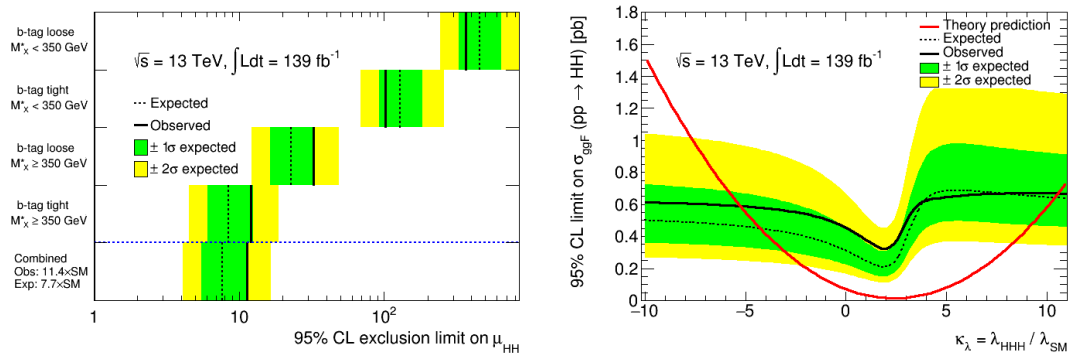

Figure 5.29 *Left: Observed and expected 95% CL upper limits on the SM HH production cross section relative to the SM expectation, μ_{HH}. Right: The observed and expected 95% CL upper limits on HH production as a function of κ_λ. The predicted cross section is overlaid in red.*

The impact of each systematic uncertainty is evaluated by computing the best fit signal strength, $\hat{\mu}_{HH}$ when varying the nuisance parameter associated with a systematic uncertainty $\hat{\theta}$ within its confidence interval, that is:

$$\Delta_{\hat{\mu},\pm} = \hat{\mu}(\hat{\theta} \pm \Delta_\theta) - \hat{\mu}(\hat{\theta}) \tag{5.12}$$

Figure 5.35 shows the five systematic uncertainties which have the highest impact, $\Delta_{\hat{\mu},\pm}$, and their central values and uncertainties for the observed data. The central values and uncertainties are all close to their nominal values (0 and $\pm$ 1), demonstrating that the fit does not have the power to constrain any systematic uncertainties yet.

The systematic uncertainty with the largest impact is the conservative (100%) heavy flavour systematic uncertainty placed on a number of the single Higgs backgrounds. In the future, a measurement of ggF single Higgs production plus additional b-jets would help to justify reducing this heavy flavour systematic uncertainty. Other large uncertainties include the theory uncertainties on HH production, the jet energy scale and the photon identification efficiency.

2D Fit Results

Figures 5.31 and 5.32 show the $m_{\gamma\gamma}$ and m_{bb} spectra observed in data and the background only fits in each category.

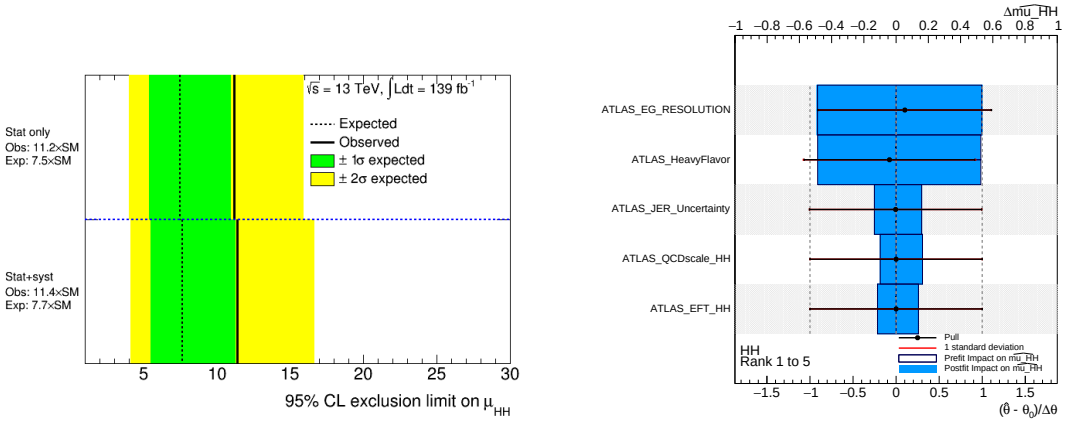

Figure 5.30 *Left: The observed and expected 95% CL upper limits on the SM HH production cross section with and without systematic uncertainties. Right: The five systematic uncertainties with the highest impact on the best fit signal strength, $\hat{\mu}_{HH}$ in order.*

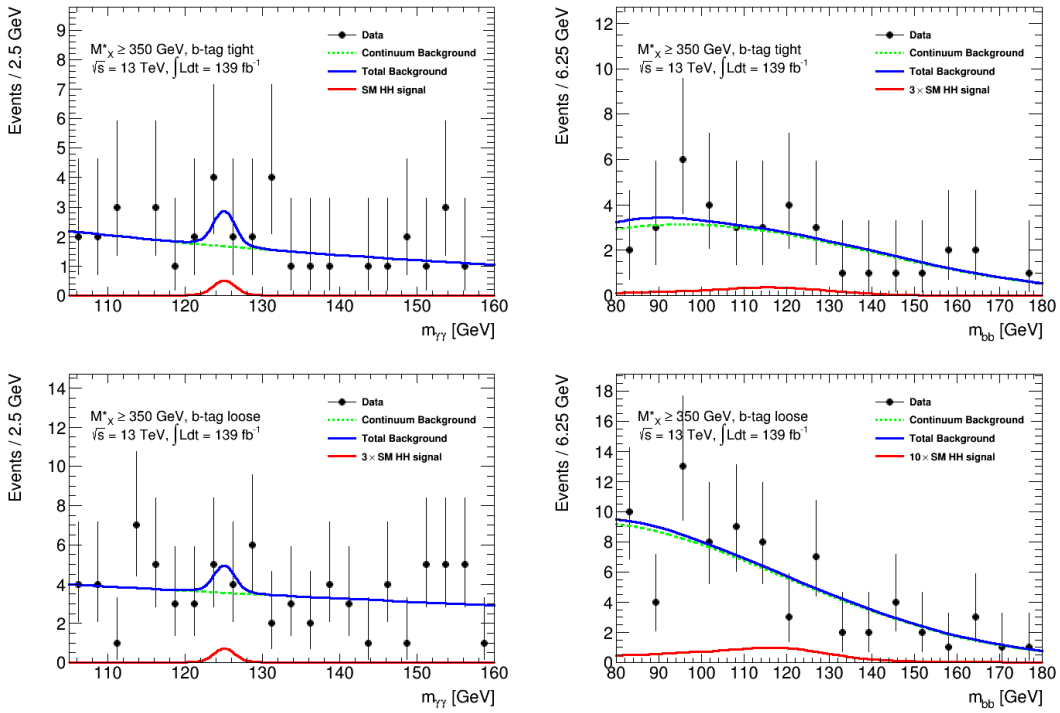

Figure 5.31 *The $m_{\gamma\gamma}$ and m_{bb} spectra in the $M_X^* \geq 350$ GeV categories. The b-tag tight (loose) categories are shown on the top (bottom).*

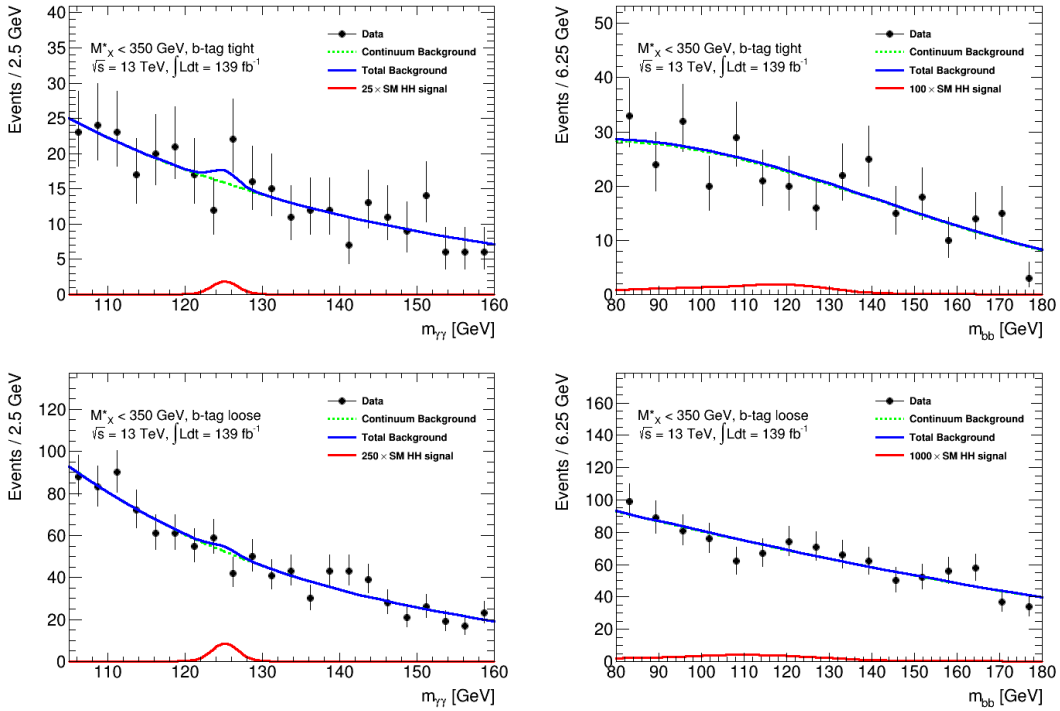

Figure 5.32 *The $m_{\gamma\gamma}$ and m_{bb} spectra in the $M_X^* < 350$ GeV categories. The b-tag tight (loose) categories are shown on the top (bottom).*

The best fit signal strength is $\mu_{HH} = 2.7 \pm 3.5$ where all single Higgs backgrounds are fixed to their SM expectations. The significance for SM HH production with respect to the background only expectation is observed (expected) to be $0.84\,\sigma$ $(0.37\,\sigma)$.

Figure 5.33 shows the expected and observed upper limits in each category and from combining all categories for SM HH production. The expected and observed upper limits are also shown as a function of κ_λ. The observed (expected) limit for the SM HH cross section is 10.1 (7.0) times the SM value. The ratio of the Higgs boson self-coupling to its SM prediction, κ_λ is observed (expected) to be constrained at 95% CL to $-4.4 < \kappa_\lambda < 9.5$ $(-3.3 < \kappa_\lambda < 8.9)$.

The use of a 2D fit is found to improve the expected significance and the upper limit on the SM HH cross section by approximately 10%. Figure 5.34 shows the expected 95% CL upper limit on HH production as a function of κ_λ. The improvement as a function of κ_λ varies between 5 and 35%. As discussed previously, HH production with values of κ_λ close to the SM value is characterised by high momentum kinematics and the backgrounds can be largely rejected using angular variables as shown in Section 5.5.5. For values of κ_λ far away from the SM, the kinematics are much softer and exploiting the invariant mass of the $H \rightarrow$

$b\bar{b}$ candidate as much as possible is vitally important.

Figure 5.35 shows the observed and expected 95% CL upper limits on μ_{HH} with and without systematic uncertainties and the ranking of the top five systematic uncertainties by impact. As in the 1D case, the systematic uncertainties have only a small impact on the analysis.

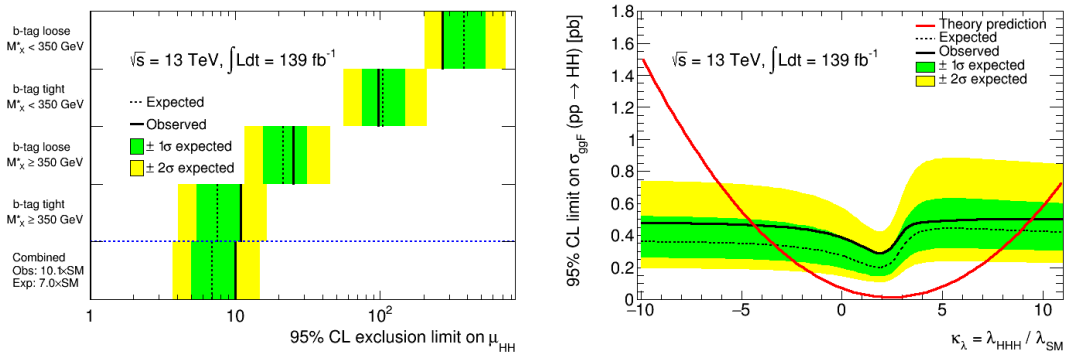

Figure 5.33 *Observed and expected 95% CL upper limits on the SM HH production cross section relative to the SM expectation, μ_{HH}. Right: The observed and expected 95% CL upper limits on HH production as a function of κ_λ. The predicted cross section is overlaid in red.*

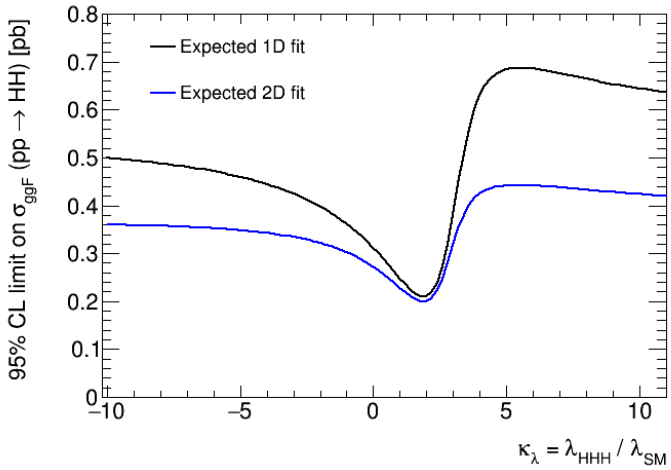

Figure 5.34 *The expected 95% CL upper limits on HH production as a function of κ_λ for the 1D fit (black) and the 2D fit (blue).*

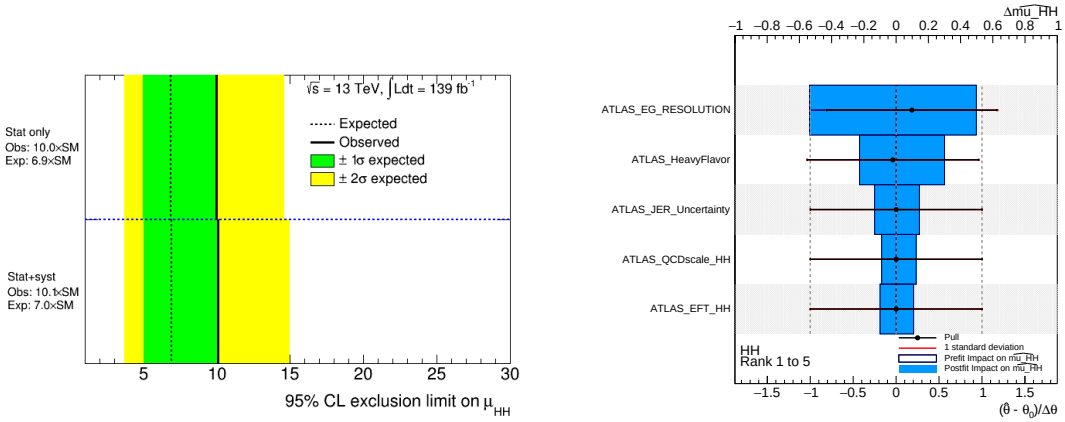

Figure 5.35 *Left: The observed and expected 95% CL upper limits on the SM HH production cross section with and without systematic uncertainties. Right: The five systematic uncertainties with the highest impact on the best fit signal strength, $\hat{\mu}_{HH}$ in order.*

5.8.2 HL-LHC projection

There are no signs of BSM physics in HH production in the Run 2 data set. In order to have sensitivity to SM HH production, a larger data set is required. The results in this chapter are projected to 3000 fb^{-1}, the data set that ATLAS expects to record during the HL-LHC program.

This simple projection is performed by taking the best fit continuum PDFs from the Run 2 data set and scaling the normalisation from 139 fb^{-1} to 3000 fb^{-1}. An expected significance of 1.7 σ is obtained for SM HH production.

For the κ_λ constraint in the Run 2 analysis, it is assumed that no HH production exists and the upper limit on the cross section is computed as a function of κ_λ with κ_λ fixed. The κ_λ constraint is found by finding where the theory prediction for the cross section intersects the expected limit. With the HL-LHC data set, it is thought that in an analysis which combines multiple final states, the projected significance for SM HH production will exceed 3 σ. With this level of sensitivity expected, it is preferable to assume that SM HH production does exist when measuring κ_λ. In Figure 5.36, HH production is fixed and the likelihood is re-parameterised with κ_λ as the parameter of interest. It is projected that κ_λ is constrained to [0.2, 2.5] at the 68% CL and [−0.5, 5.7] at the 95% CL.

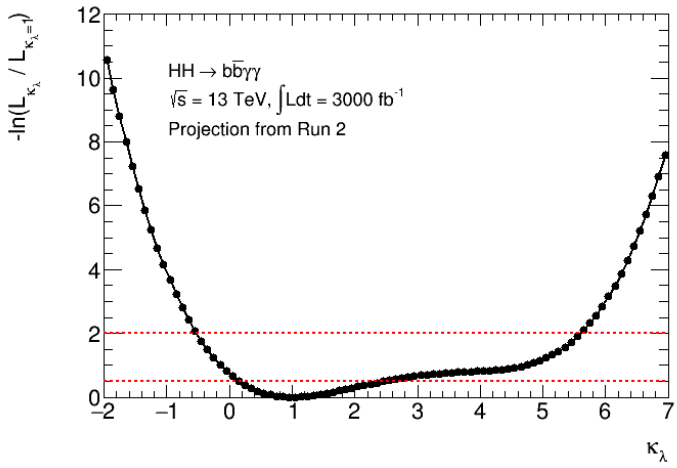

Figure 5.36 *The likelihood scan as a function of κ_λ from projecting the Run 2 analysis to a data set of* $3000\,\mathrm{fb}^{-1}$*. The dashed lines correspond to the 1σ and 2σ confidence intervals.*

5.8.3 Further improvements

In the b-tag tight, $M_X^* \geq 350$ GeV category, the largest single Higgs backgrounds are the ZH and $t\bar{t}H$ processes which have yields of 0.78 and 0.60 events respectively for the 2D selection. These yields can be compared to the SM HH signal which has an event yield of 0.77. It is likely that the ZH background is largely constrained by the 2D fit due to the $Z \to b\bar{b}$ resonance peaking around the Z mass. The $t\bar{t}H$ background shape in m_{bb}, however, peaks near the Higgs mass so it is likely that this background contributes significantly. In order to suppress the $t\bar{t}H$ background, the event selection requires less than six central jets and zero electrons or muons. In the future, it would be useful to study to what extent requiring zero hadronic tau candidates can further suppress the $t\bar{t}H$ background. Hadronic tau candidates are currently not available in the ATLAS $H \to \gamma\gamma$ framework so this study cannot easily be performed.

The modified $b\bar{b}\gamma\gamma$ mass, M_{X}^* is a variable particularly sensitive to the value of κ_λ. In the analysis presented here, the analysis was split into two M_{X}^* categories. It could be interesting to study splitting the analysis into additional M_{X}^* categories and studying the interplay between improving the constraint on κ_λ further but possibly degrading to sensitivity to SM HH production.

Multivariate algorithms such as boosted decision trees could also be used for the event selection. However care is necessary to avoid sculpting the $m_{\gamma\gamma}$ and m_{bb} distributions so that the background can still be obtained by fitting

monotonically falling distributions. Another issue is that simply training a multivariate algorithm on the SM HH sample significantly reduces the sensitivity to BSM values of κ_λ.

5.8.4 Comparisons to other results

Comparison to $36\,\text{fb}^{-1}$ results

The ATLAS collaboration has combined HH results in the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau\tau$ and $b\bar{b}b\bar{b}$ final states using a data set of $36\,\text{fb}^{-1}$ [101]. The expected and observed upper limits on the SM HH cross section and κ_λ are shown in Figure 5.37. For the $b\bar{b}\gamma\gamma$ final state, the expected limit is 26.3 times the SM value. For a single bin counting experiment with no systematic uncertainties, the expected limit will scale as $1 / \sqrt{L}$, where L is the integrated luminosity of the data set. If this scaling is applied to the previous $b\bar{b}\gamma\gamma$ analysis, an expected limit of 13.4 times the SM is obtained. The expected limit obtained in the analysis presented here is 7.0 times the SM. This improvement has been found by updating the categorisation and by performing a 2D fit instead of a 1D fit. The expected sensitivity obtained in this analysis is better than the full $36\,\text{fb}^{-1}$ combination, however due to the deficit in the combination and the upwards fluctuation here, the observed limit is worse. In the combination, κ_λ is expected to be constrained in the interval $[-5.8, 12.0]$ while the analysis presented has an expected constraint of $[-3.3, 8.9]$.

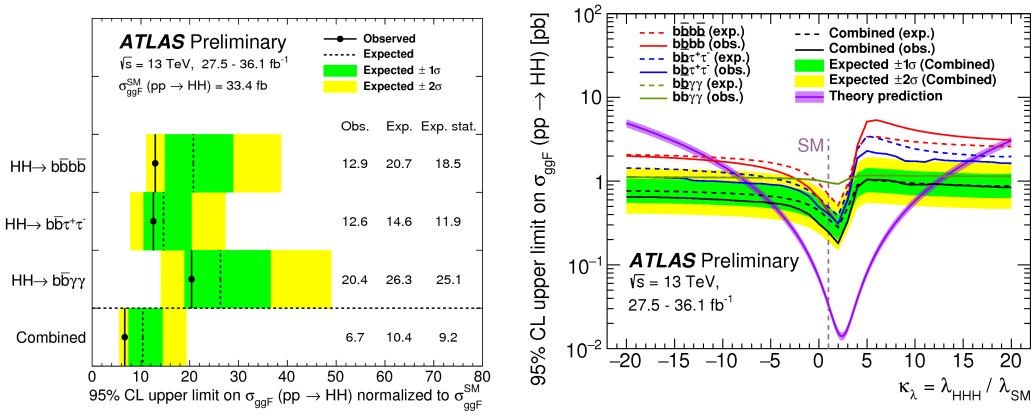

Figure 5.37 *Left: The observed and expected 95% CL upper limits on the SM HH production cross section. Right: The observed and expected 95% CL upper limits on HH production as a function of κ_λ. The predicted cross section is overlaid in purple.*

Comparison to ATLAS HL-LHC projection

In the projection presented, a significance of $1.7\,\sigma$ is obtained for SM HH production while a significance of $2.0\,\sigma$ is obtained in the HL-LHC study. For the ratio of the Higgs boson self-coupling to its SM expectation, this study finds κ_λ can be constrained at the 95% CL to $[-0.5, 5.7]$ while the HL-LHC study finds $[-1.1, 8.2]$.

There are a large number of caveats that must be applied to the HL-LHC projection in section 5.8.2. For example, the photon identification and isolation efficiency will be worse at the HL-LHC as a result of the higher pile-up. More stringent selections will be necessary in order to suppress jets faking photons in the ATLAS detector.

However, the b-tagging performance is expected to be superior at the HL-LHC due to the ITk detector replacing the current inner detector. For a b-tag working point of 70%, it is expected that a charm rejection of a factor of 20 and a light jet rejection of a factor of 700 [67] can be achieved. This can be compared to the Run 2 performance of 9 and 412 respectively.

The cross sections and the kinematics for the signal and background processes will also change with the increase in centre-of-mass energy of the LHC from 13 TeV to 14 TeV. This has also not been accounted for.

As discussed in Chapter 3, there is a study from the ATLAS collaboration which does take the above changes into account [63]. One of the disadvantages of this study is that the background is taken from MC simulation which and not validated against any data and as shown in this chapter, the agreement between data and simulation in the Run 2 analysis is not good.

The large difference in the κ_λ constraint arises due to the different analysis strategies employed. In the HL-LHC study, a boosted decision tree algorithm trained on $\kappa_\lambda = 1$ is used and a 1D fit is performed to the $m_{\gamma\gamma}$ spectrum for one category only. The acceptance $\times$ efficiency and the likelihood scan are shown as a function of κ_λ in Figure 5.38 for the HL-LHC study. The acceptance $\times$ efficiency for $\kappa_\lambda = 6$ is approximately 0.5% which results in a second minimum in the likelihood scan near this value and a weak κ_λ constraint overall. In the Run 2 analysis, the acceptance $\times$ efficiency for $\kappa_\lambda = 6$ is approximately 9% which has been achieved by splitting the analysis into multiple categories. It can be seen in Figure 5.36 that the second minimum in the likelihood scan has largely been

lifted in the projection from the Run 2 analysis.

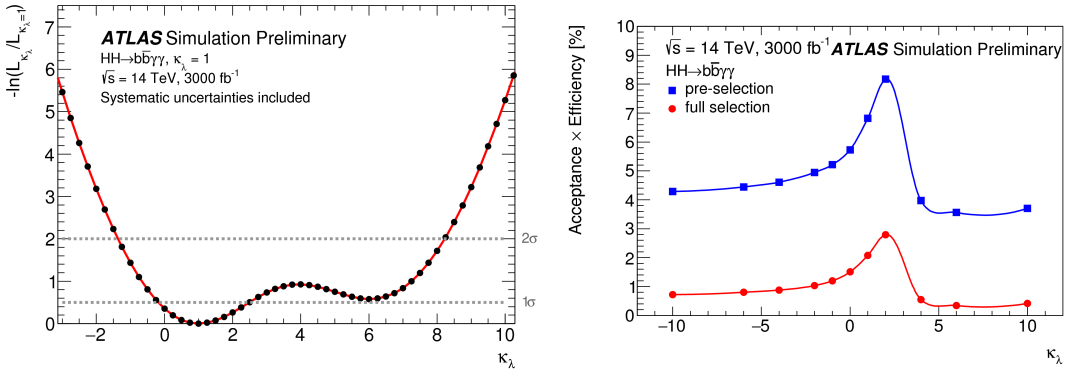

Figure 5.38 *The likelihood scan (Left) and the acceptance × efficiency (Right) as a function of κ_λ for the HL-LHC study.*

5.9 Summary

No significant excess is observed in the search for HH production in the $b\bar{b}\gamma\gamma$ final state with the Run 2 data set. The observed (expected) limit at 95% CL for the SM HH cross section is 10.1 (7.0) times the SM value. The ratio of the Higgs boson self-coupling to its SM prediction, κ_λ is observed (expected) to be constrained at 95% CL to $-4.4 < \kappa_\lambda < 9.5$ ($-3.3 < \kappa_\lambda < 8.9$). At the time of writing, this constraint on the Higgs boson self-coupling is the best in the world.

For the first time in ATLAS in the $HH \to b\bar{b}\gamma\gamma$ channel, a 2D fit has been developed and compared to the 1D fit typically used in $H \to \gamma\gamma$ physics analyses. It is found to be a powerful tool, in particular for BSM values of the Higgs boson self-coupling.

The studies in this chapter lay the foundations for a potential observation of HH and a measurement of the Higgs boson self-coupling at the LHC.

Chapter 6

Conclusion

The discovery of a new particle consistent with the Standard Model Higgs boson in 2012 by the ATLAS and CMS collaborations completed the Standard Model of particle physics. A huge amount of work has now been done to study this new particle. In Run 1 at the LHC, the experiments observed the Higgs boson decaying into photons, W and Z bosons. In Run 2, the experiments have now observed the Higgs boson coupling to top and bottom quarks and tau fermions.

The next observations in Higgs physics assuming the SM will be the $H \to \mu^+\mu^-$ and the $H \to Z\gamma$ decays. The $H \to \mu^+\mu^-$ decay is particularly important as it will be the only observation of the Higgs boson coupling to second generation fermions at the LHC assuming the SM. After these goals have been reached, one of the remaining goals will be to observe Higgs boson pair production and to measure the Higgs boson self-coupling.

Due to the small cross section for Higgs boson pair production, it is necessary to upgrade the LHC so it can deliver higher luminosities and therefore a larger data set. The higher luminosities will present more challenging detector conditions and therefore in order to maintain the current detector performance, the ATLAS detector must be upgraded. Contributions to the development of the future ATLAS Inner Detector (ITk) are presented. Results are presented for the expected silicon cluster sizes, channel occupancies and cluster densities that would be expected in two proposed layout designs of the ITk. It is demonstrated that the channel occupancy can be kept to approximately less than 1% throughout the ITk Strip detector and to less than approximately 0.1% throughout the ITk Pixel detector for the inclined design. These studies combined with many other

studies performed by members of the ATLAS collaboration led the collaboration to pursue the construction of an ITk that resembled the inclined design.

It has been proposed in a number of phenomenological studies that the $t\bar{t}HH(HH \to b\bar{b}b\bar{b})$ channel could be a promising final state for studying HH production. However in this thesis, this final state is examined for the first time with a parameterised detector response and realistic backgrounds. It is found that the backgrounds are much larger than those presented in the phenomenological studies. The origin of this is from mis-tagging c-jets which have originated from the decay of a W boson in the $t\bar{t}b\bar{b}$ + jets process. The phenomenological studies only included the irreducible $t\bar{t}bb\bar{b}\bar{b}$ process which led to this large difference. It is likely that in a combined HH analysis, the $t\bar{t}HH$ production mechanism will only contribute a small amount.

It is feasible that BSM physics could enhance the rate of HH production and therefore it is already interesting to study this process with the current data set despite having no sensitivity to the SM. The search for HH production in the $b\bar{b}\gamma\gamma$ final state with the full Run 2 data set is presented in Chapter 5. For the first time in this analysis within ATLAS, a 2D fit is developed and used. It is shown to be a powerful tool, improving the sensitivity by approximately 10% to SM HH production and by more than 30% to some BSM values of κ_λ. No significant deviations from the SM are observed in this analysis and 95% CL upper limits are established. The observed (expected) limit for the SM HH cross section was 10.1 (7.0) times the SM value while the ratio of the Higgs boson self-coupling to its SM prediction, κ_λ was observed (expected) to be constrained at 95% CL to $-4.4 < \kappa_\lambda < 9.5$ ($-3.3 < \kappa_\lambda < 8.9$). A simple projection is performed to a data set of $3000\,\text{fb}^{-1}$, the size of data set that ATLAS hopes to obtain during the HL-LHC program. It is found that a significance of $1.7\,\sigma$ can be achieved in this final state at the HL-LHC while κ_λ can be constrained to $[-0.5, 5.7]$ at the 95% CL.

The ATLAS and CMS collaborations have recently combined the prospects for HH at the HL-LHC and found that a significance of $4.0\,\sigma$ can be achieved, short of the $5\,\sigma$ observation threshold that particle physicists hold themselves to. An observation will only be possible by examining as many HH final states as possible and by improving the sensitivity in the most promising final states with advanced analysis techniques. The work presented in this thesis lays the foundation for a potential observation of HH which would represent the pinnacle of the LHC program.

Appendix A

ITk detector specification

A specification of the ITk detector used for the studies shown in Chapter 3 is given.

Strip Barrel

In the first two layers in the strip barrel, each module has four rows of short (24.10 mm) strips and in the outer two layers, each module has two rows of longer strips (48.20 mm). Each row has 1280 strips.

The number of strips over ϕ is given by the number of strips in a segment (1280) $\times$ the number of rows of segments (four in layers 0 and 1, two in layers 2 and 3) $\times$ the number of staves $\times$ 2 (the modules are double sided). The area over ϕ is given by the area of a module $\times$ the number of staves $\times$ 2.

Layer	Type	Staves	read out per sensor	read out over ϕ	area (mm^2)	area over ϕ (mm^2)
0	Short	28	5120	286,720	9316.096	521,701
1	Short	40	5120	409,600	9316.096	745,288
2	Long	56	2560	286,720	9316.096	1,043,403
3	Long	72	2560	368,640	9316.096	1,341,518

Table A.1 *Specification of the strip barrel for the ITk layouts.*

Strip End-cap

There are six disks in each end-cap with each disk having 6 modules. The area over ϕ is calculated as:

$$A = A_{Outer} - A_{Inner} = \pi(r^2_{Outer} - r^2_{Inner}) \tag{A.1}$$

multiplied by a factor of 2 because the sensors are double sided. The area for each sector is calculated by dividing the area over ϕ by the number of sectors.

Module	Sectors	read out per sensor	read out over ϕ	Inner radius (mm)	Outer radius (mm)	area (mm^2)	area over ϕ (mm^2)
0	32	1024+1024+1152+1152	278,528	384.5	488.423	8906.1	569,990
1	32	1280+1280+1408+1408	344,064	489.823	574.194	8813.4	564,055
2	32	1536+1536	196,608	575.594	637.209	7336.3	469,522
3	64	896+896+896+896	458,752	638.609	755.501	7999.3	1,023,909
4	64	1024+1024	262,144	756.901	866.062	8696.5	1,113,155
5	64	1152+1152	294,912	867.462	967.785	9037.8	1,156,844

Table A.2 *Specification of the strip end-cap for the ITk layouts*

Pixel Barrel

The pixel size for the ITk is chosen to be 50 × 50 μm^2 at present, in comparison to the size in the current pixel detector of 50 × 400 μm^2. The chip is 20.0 mm by 16.8 mm and has 336 × 400 pixels in it.

In the extended layout, there are two chips per sensor in the innermost layer and four chips per sensor in the other layers. In the inclined layout, the flat section of the pixel barrel is identical to the extended layout but in the inclined sections, there is only one chip per sensor in the innermost layer and only two chips per sensor in the other layers.

Table A.3 shows the specification for the pixel barrel. For the inclined layout, Table A.3 is only applicable to the central part of the barrel. The specification for the inclined barrel is shown in Table A.4.

Layer	Sensor Size [mm^2]	Chips per sensor	Number of sensors around ϕ	Radius [mm]
0	40.2×16.8	2	18	39
1	40.2×33.8	4	18	85
2	40.2×33.8	4	32	155
3	40.2×33.8	4	44	213
4	40.2×33.8	4	54	271

Table A.3 *Specification of the pixel barrel for the ITk layouts. For the inclined layout, the above refers only to the central part of the barrel.*

Pixel Rings

The specification for the pixel rings is shown in Table A.5. The pixel ring system is the same in both layouts.

Layer	Sensor Size [mm^2]	Positions in Barrel [mm]
0	20.0×16.8	197.8 234.1 285.8 322.8 359.7 403.1 454.0 513.9 584.2 638.4 696.4 760.9 832.4 911.9 1000.2 1098.1 1206.9
1	20.0×33.8	214.4 240.7 272.2 309.9 355.1 409.2 473.9 551.5 599.9 646.5 697.3 752.7 813.0 878.8 950.5 1028.6 1113.7 1206.5
2	20.0×33.8	254.1 275.9 300.0 326.4 355.4 387.3 422.4 461.0 503.4 550.0 601.3 657.6 719.6
3	20.0×33.8	295.7 318.6 343.2 369.7 398.1 428.8 461.8 497.2 535.4 576.5 620.7 668.3 719.5
4	20.0×33.8	336.7 359.5 383.5 409.1 436.1 464.8 495.1 527.3 561.4 597.5 635.8 676.4 719.4

Table A.4 *Specification of the inclined pixel barrel in the inclined layout*

Ring Layer	Sensor Size [mm^2]	Sensors per Ring	Inner Radius [mm]	Ring Positions [mm]
0	40.2×33.8	24	80	1308 1391 1501 1620 1750 1830 1910 1997 2088 2188 2292 2397 2503 2618 2740 2867 3000
1	40.2×33.8	36	150	823 899 986 1082 1189 1308 1394 1486 1598 1685 1778 1876 1980 2090 2246 2414 2596 2793 3000
2	40.2×33.8	48	212.5	823 944 1088 1258 1349 1448 1554 1669 1794 1929 2075 2233 2404 2589 2790 3000
3	40.2×33.8	60	275	823 918 1027 1151 1294 1456 1642 1854 1968 2089 2217 2355 2502 2658 2825 3000

Table A.5 *Specification for the pixel rings.*

Appendix B

LHC Statistics

Well defined statistical tests are necessary to make statements on discovery, exclusion and measurement at the LHC. This section describes a number of the statistical tests that are used at the LHC.

This text is based upon Refs. [97, 100] and lectures (Course d'Hiver 2018, Statistics in High Energy Physics) given by Nicolas Berger (LAPP, Annecy).

Introduction

The discovery of a new signal can be claimed in particle physics when the background only hypothesis can be rejected from the data. When it is not possible to reject the background only hypothesis and claim a discovery, the extent to which the hypothesised signal can be rejected is typically reported instead. Statistical hypothesis testing is therefore at the centre of physics results that are reported from the LHC.

In the case of searching for new physics, the p-value is defined as the fraction of outcomes that are as signal-like (H_1 hypothesis) as the data when the background only (H_0 hypothesis) is true. Obtaining a p-value of 10^{-7} from an experiment therefore corresponds to obtaining that particular outcome once in 10^7 experiments providing the background only hypothesis is true. Instead of quoting the p-value, it is often preferred to report the equivalent number of

Gaussian quantiles, using the formula:

$$Z = \Phi^{-1}(1 - p) \tag{B.1}$$

where Φ^{-1} is the inverse of the cumulative distribution of the Gaussian. It is conventional in particle physics, that *evidence* of a signal is claimed when the p-value is less than 1.35×10^{-3} ($3\,\sigma$) and an *observation* is claimed when the p-value is less than 2.87×10^{-7} ($5\,\sigma$).

The p-value can be computed from a choice of test statistic. At the LHC, a profile likelihood ratio test statistic is used. The profile likelihood ratio is given by:

$$\lambda(\mu) = -2 \ln \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{B.2}$$

where $\hat{\hat{\boldsymbol{\theta}}}$ are the values of the nuisance parameters that maximise the likelihood on the condition that μ (usually the signal strength modifier) is fixed to a specified value. In the denominator, $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the values of the signal strength and nuisance parameters that unconditionally maximise the likelihood.

The profile likelihood ratio test statistic, t_μ is given by

$$t_\mu = -2 \ln \lambda(\mu) \tag{B.3}$$

and from this the p-value can be computed as

$$p_\mu = \int_{t_\mu, \text{obs}}^{\infty} f(t_\mu|\mu) dt_\mu \tag{B.4}$$

where t_μ, obs is the value of the test statistic observed from the data and $f(t_\mu|\mu)$ is the PDF of t_μ for a given value of μ. The relationships between the p-value and the observed t_μ and the significance Z are shown in Figure B.1.

Discovery

For a discovery in particle physics, two hypotheses are required. The background only hypothesis (H_0, $\mu = 0$) is tested against a signal hypothesis (H_1, μ != 0).

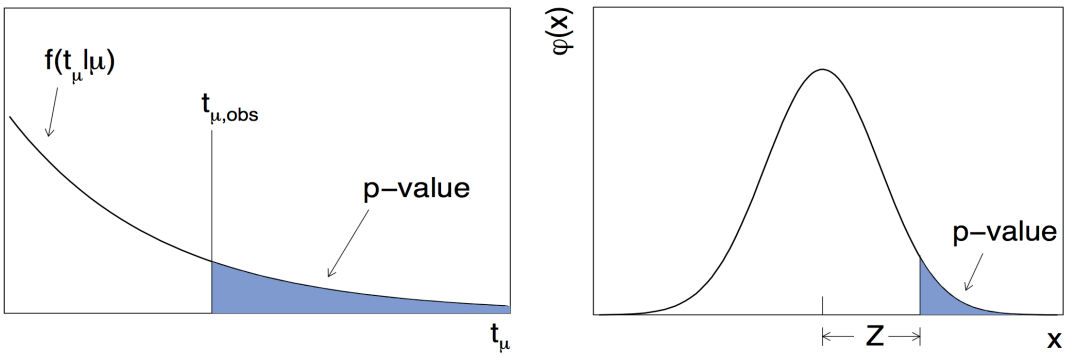

Figure B.1 *Left: The relationship between the p-value and the test statistic t_μ. Right: The relationship between the p-value and the significance, Z.*

For the H_1 hypothesis, the signal that is used is the best fit value from the data, $\hat{\mu}$. The test statistic is therefore:

$$t_0 = \begin{cases} -2\ln\frac{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}(0))}{\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{B.5}$$

In collider physics, it is assumed that only $\hat{\mu} \geq 0$ represents a signal and potential discovery, hence $t_0 = 0$ for $\hat{\mu} < 0$. This is known as a one-sided test statistic.

If $\hat{\mu}$ follows a Gaussian distribution, it can be shown that the test statistic, t_0 is distributed as a χ^2 distribution, more precisely due to the choice of a one-sided test statistic a half-χ^2 distribution. It can also be shown that the significance is given by

$$Z = \sqrt{t_0} \tag{B.6}$$

The above assumptions and equations can be used to derive a result that is commonly used in particle physics. For a single bin counting experiment with no systematic uncertainties, the likelihood is given by

$$\mathcal{L}(S) = e^{-\frac{1}{2}\left(\frac{n-(S+B)}{\sqrt{S+B}}\right)^2} \tag{B.7}$$

where the background B is known, n is the observed number of events and S is the number of signal events used as the parameter of interest. Using equations

(B.5 - B.7) and the fact that $\hat{S} = n - B$, it is possible to obtain the known result:

$$Z = \frac{\hat{S}}{\sqrt{B}} \tag{B.8}$$

Exclusion

If there is no significant excess in data, it is more useful to report upper limits on the signal hypothesised. In order to establish an upper limit, the hypothesis for a given signal (H_0, $\mu = \mu_0$) is tested against the alternative hypothesis, (H_0, $\mu < \mu_0$). When establishing an upper limit, an upwards fluctuation of the data is not regarded as representing incompatibility with the hypothesised μ_0 value, therefore the test statistic is set to 0 when $\hat{\mu} > \mu_0$. If one considers signal models that have $\mu_0 \geq 0$ and if the best fit value $\hat{\mu}$ is negative, the closest physical model has $\mu_0 = 0$. The test statistic is therefore given by:

$$q_{\mu_0} = \begin{cases} -2\ln \frac{\mathcal{L}(\mu_0, \hat{\hat{\boldsymbol{\theta}}}(\mu_0))}{\mathcal{L}(0, \hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \\ -2\ln \frac{\mathcal{L}(\mu_0, \hat{\hat{\boldsymbol{\theta}}}(\mu_0))}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} \leq \mu_0 \\ 0 & \hat{\mu} > \mu_0 \end{cases} \tag{B.9}$$

The value of μ_0 is adjusted to give a pre-defined p-value, conventionally in ATLAS this is chosen as 5% which corresponds to a 95% confidence level.

Similar to the discovery case, it can be shown that q_μ is distributed as a half-χ^2 distribution and the p-value is given by

$$p_{\mu_0} = 1 - \Phi(\sqrt{q_{\mu_0}}) \tag{B.10}$$

The procedure to establish the upper limit is then to estimate μ_0, compute q_{μ_0} and find the p-value. The value of μ_0 can then be adjusted until the pre-defined p-value (usually 5%) is reached.

For the case of a single bin counting experiment a simple equation can be derived

for the upper limit. Substituting equation B.7 into equation B.9 gives

$$q_{S_0} = \left(\frac{S_0 - \hat{S}}{\sigma_S} \right)^2 \tag{B.11}$$

where S_0 is the tested number of signal events and $\hat{S}$ is the best fit number of signal events. From equation B.10, if the p-value is set to 0.05, q_{μ_0} is equal to 2.70. Equation B.11 can then be re-arranged to give

$$S_{\text{up}} = \hat{S} + 1.64\sigma_S \tag{B.12}$$

where S_{up} is the number of signal events excluded at 95% CL.

CLs method

If there is a deficit in data with respect to the background only expectation, the value of $\hat{S}$ will be negative. If the deficit is large enough, it can be seen from equation B.12 that it could be possible that the upper limit on the number of signal events is also negative.

A procedure is developed (CL_s, [100]) is developed to avoid this situation. The solution is to use a modified p-value, given by

$$p_{CL_s} = \frac{p_{\mu_0}}{p_0} \tag{B.13}$$

where p_0 is the p-value computed under the hypothesis $\mu = 0$ rather than μ_0. This has the property desired that the upper limit on μ is always greater than zero. However there is the disadvantage that for some small values of p_0, the upper limit is greater than the 95% CL that is quoted.

Asimov Datasets

In order to obtain the expected sensitivity (either discovery significance or upper limit) of experiments it is necessary to compute the test statistic distributions $f(q_\mu|\mu)$ for a given μ. There are two options for computing this, the first option

involves generating pseudo-experiments and calculating q_μ for each experiment and determining the full distribution $f(q_\mu|\mu)$. The second option involves generating a so called *Asimov* dataset in which all the parameters contained in the likelihood are set to their expected values. From this dataset it is possible to use formulas based on Gaussian approximations and determine the expected sensitivity of the experiment. The use of the Asimov dataset avoids generating large sets of pseudo-experiments and significantly reduces the computation time needed to find the sensitivity of an experiment.

The disadvantage of the Asimov dataset approach is that it relies on the Gaussian approximation. For a small number of observed events, this approximation breaks down. This can be seen in Figure B.2 which are the results from ATLAS for the search for a resonant particle decaying to two photons [102]. After $m_{G^*} > 2.5$ TeV, the expected and observed limits are calculated both with Asimov datasets (grey lines) and pseudo-experiments (blue lines). There is clear disagreement between the two approaches.

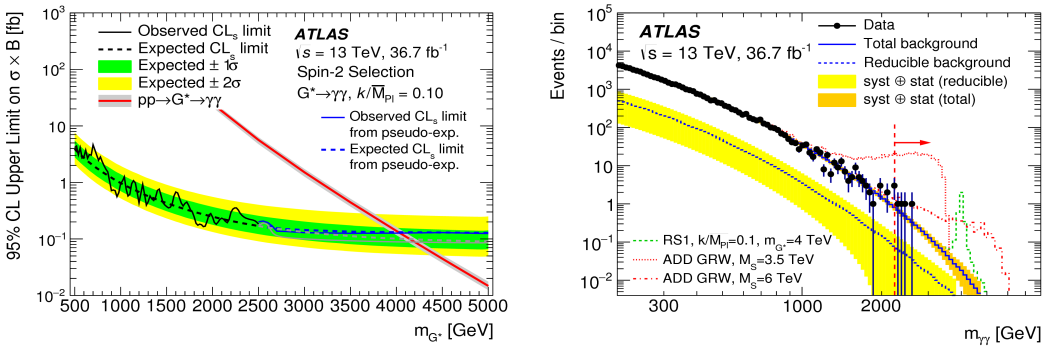

Figure B.2 *Left: Upper limits on the production cross section times branching ratio for a spin-2 particle decaying to two photons. For $m_{G^*} > 2.5$ TeV, the observed and expected limits are determined with pseudo-experiments shown by the blue solid and dashed lines, respectively. Right: The observed invariant mass distribution of diphoton events with the spin-2 selection with the predicted SM background.*

Measurement

If there is the prior assumption that $\hat{\mu} > 0$ then the test statistic is defined as

$$t_{\mu_0} = \begin{cases} -2\ln\frac{\mathcal{L}(\mu_0,\hat{\hat{\boldsymbol{\theta}}}(\mu_0))}{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \\ -2\ln\frac{\mathcal{L}(\mu_0,\hat{\hat{\boldsymbol{\theta}}}(\mu_0))}{\mathcal{L}(\hat{\mu},\hat{\theta})} & \hat{\mu} \geq 0 \end{cases} \tag{B.14}$$

where for $\hat{\mu} < 0$, the least compatibility the data can have with any value of μ is $\mu = 0$. In the Gaussian approximation, t_{μ_0} is distributed as a χ^2 and as a result the uncertainties can be constructed by finding the values of the test statistic that are equal to Z^2 to determine the $\pm$ Zσ uncertainties.

Systematic uncertainties

In the discussion below, the $HH \to b\bar{b}\gamma\gamma$ detailed in Chapter 5 will serve as an example.

Systematic uncertainties are described by nuisance parameters which are implemented into the likelihood that describes the analysis.

They are implemented into the likelihood by multiplying the relevant parameter of the statistical model by a response function. In the case of a Gaussian PDF for an uncertainty of size σ, it is given as:

$$F_G(\sigma, \theta) = (1 + \sigma \cdot \theta) \tag{B.15}$$

and for cases where a negative model parameter does not make physical sense, the log-normal PDF is used instead, given as:

$$F_{LN}(\sigma, \theta) = e^{\ln(1+\sigma)\theta} \tag{B.16}$$

In both cases, the corresponding constraint product $G(\boldsymbol{\theta})$ is a unit Gaussian centered at zero for θ.

For example, in the $H \to \gamma\gamma$ signal model, the signal mean μ_{CB} and resolution σ_{CB} become:

$$\mu_{CB} \to \mu_{CB}(1 + \sigma_{PES}\theta_{PES}) \tag{B.17}$$

$$\sigma_{CB} \to \sigma_{CB}(1 + \sigma_{PER}\theta_{PER}) \tag{B.18}$$

where PES and PER are the photon energy scale and resolution uncertainties respectively.

The number of HH signal events N_{HH}, considering only HH theory systematic uncertainties θ_T, is given by:

$$N_{HH} \to N_{HH} \prod_j e^{\ln(1+\sigma^j_{T,HH})\theta^j_T} \tag{B.19}$$

The impact of each systematic uncertainty on an analysis can be evaluated by computing the best fit signal strength, $\hat{\mu}_{HH}$ when varying the nuisance parameter associated with a systematic uncertainty $\hat{\theta}$ within its confidence interval, that is:

$$\Delta_{\hat{\mu},\pm} = \hat{\mu}(\hat{\theta} \pm \Delta_\theta) - \hat{\mu}(\hat{\theta}) \tag{B.20}$$

Figure B.3 shows the ten highest ranked systematic uncertainties for the 1D fit for the Asimov data set (left) and the observed data (right). Figure B.4 shows the same quantities for the 2D fit.

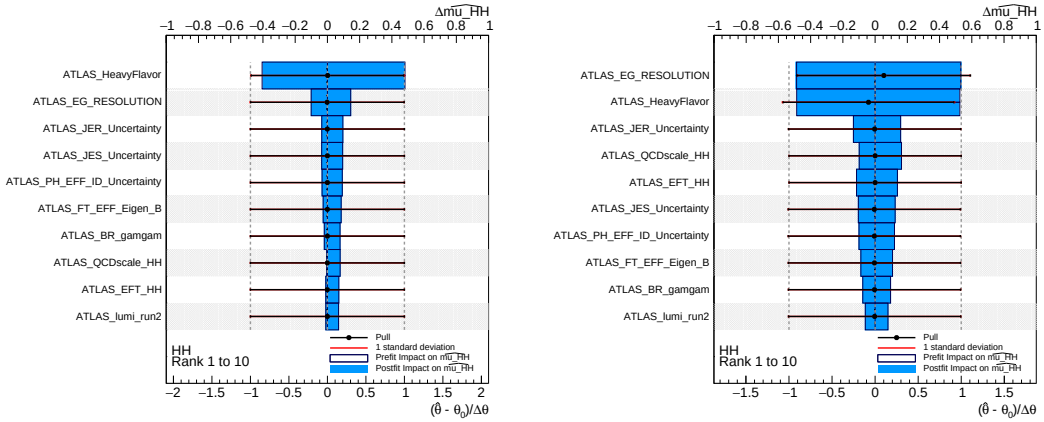

Figure B.3 *The ten systematic uncertainties with the highest impact on the signal strength, μ_{HH} are shown in order for the Asimov data set (left) and the observed data (right).*

Before the fit, each nuisance parameter (NP) has a value of 0 ± 1 by construction because $G(\boldsymbol{\theta})$ are unit Gaussian constraints. Post-fit, each NP can have a non-zero central value since the NP can absorb some feature(s) in the data. The NP uncertainties can also drop below unity because some feature(s) in the data can constrain the NPs. If the central value of the NP is pulled far from zero or its uncertainty is notably constrained, the features in the data responsible for this should be understood.

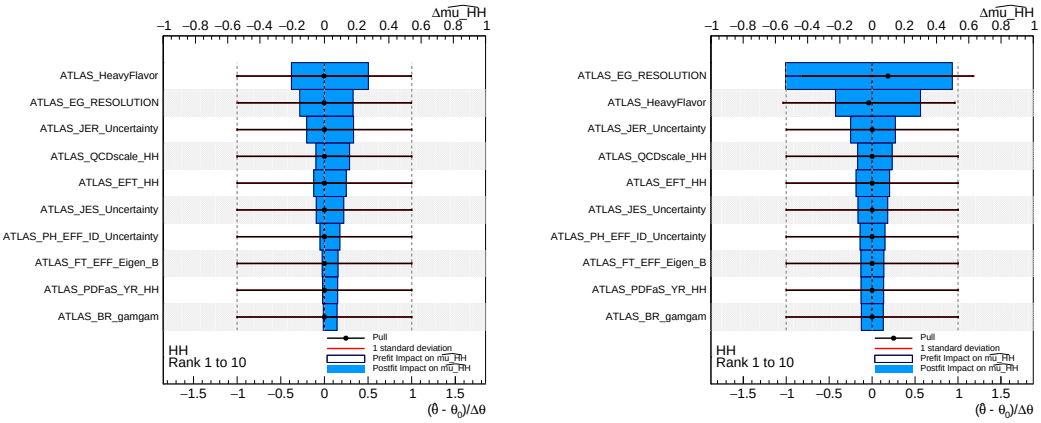

Figure B.4 *The ten systematic uncertainties with the highest impact on the signal strength, μ_{HH} are shown in order for the Asimov data set (left) and the observed data (right).*

Bibliography

[1] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[2] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[3] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.

[4] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[5] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

[7] K. A. Olive el al. Review of Particle Physics, 2014-2015. Review of Particle Properties. *Chin. Phys. C*, 38:090001, 2014. All tables, listings, and reviews (and errata) are also available on the Particle Data Group website: http://pdg.lbl.gov.

[8] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne. Parton distributions in the LHC era: MMHT 2014 PDFs. *Eur. Phys. J.*, C75(5):204, 2015.

[9] Stefan Hoche, Silvan Kuttimalai, Steffen Schumann, and Frank Siegert. Beyond Standard Model calculations with Sherpa. *Eur. Phys. J.*, C75(3):135, 2015.

[10] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, and Michele Treccani. Matching matrix elements and shower evolution for top-pair production in hadronic collisions. *Journal of High Energy Physics*, 2007(01):013, 2007.

[11] S. Catani, Yuri L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally invariant k_T clustering algorithms for hadron hadron collisions. *Nucl. Phys.*, B406:187–224, 1993.

[12] 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. *Journal of High Energy Physics*, 2014(7):79, Jul 2014.

[13] J. Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *The European Physical Journal C*, 76(4):196, Apr 2016.

[14] T. Gleisberg et al. 2009(02):007–007, feb 2009.

[15] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852 – 867, 2008.

[16] ATLAS collaboration. Measurements of the production cross section of a Z boson in association with jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J.*, C77(6):361, 2017.

[17] ATLAS simulation of boson plus jets processes in Run 2. Technical Report ATL-PHYS-PUB-2017-006, CERN, Geneva, May 2017.

[18] Image of main Higgs boson production mechanisms. `https://www-cdf.fnal.gov/physics/new/hdg/Plain_English.html`. Accessed: 2018-09-30.

[19] C. Anastasiou et al. High precision determination of the gluon fusion Higgs boson cross-section at the LHC. *Journal of High Energy Physics*, 2016(5):58, May 2016.

[20] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1, 2012.

[21] Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector. Technical Report ATLAS-CONF-2018-036, CERN, Geneva, Jul 2018.

[22] Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment. Technical Report ATLAS-CONF-2018-031, CERN, Geneva, Jul 2018.

[23] ATLAS Collaboration. Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector. *Phys. Lett.*, B784:173–191, 2018.

[24] M. Grazzini et al. Higgs boson pair production at NNLO with top quark mass effects. *JHEP*, 05:059, 2018.

[25] R. Frederix et al. Higgs pair production at the LHC with NLO and parton-shower effects. *Phys. Lett.*, B732:142–149, 2014.

[26] *Technical proposal*. LHC Tech. Proposal. CERN, Geneva, 1994. CMS, the Compact Muon Solenoid : technical proposal.

[27] CMS Collaboration. Inclusive Search for a Highly Boosted Higgs Boson Decaying to a Bottom Quark-Antiquark Pair. *Phys. Rev. Lett.*, 120:071802, Feb 2018.

[28] ATLAS Collaboration. Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector. *Phys. Rev. D*, 92:012006, 2015.

[29] ATLAS Collaboration. Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Phys. Rev. D*, 91:012006, 2015.

[30] ATLAS Collaboration. Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Phys. Rev. D*, 90:112015, 2014.

[31] Cross-section measurements of the Higgs boson decaying to a pair of tau leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2018-021, CERN, Geneva, Jun 2018.

[32] ATLAS Collaboration. Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment. *Phys. Rev. Lett.*, 120(21):211802, 2018.

[33] Prospects for $H \to c\bar{c}$ using Charm Tagging with the ATLAS Experiment at the HL-LHC. Technical Report ATL-PHYS-PUB-2018-016, CERN, Geneva, Aug 2018.

[34] Prospects for the measurement of the rare Higgs boson decay $H \to \mu\mu$ with 3000 fb^{-1} of pp collisions collected at $\sqrt{s} = 14$ TeV by the ATLAS experiment. Technical Report ATL-PHYS-PUB-2018-006, CERN, Geneva, May 2018.

[35] CMS Collaboration. Search for Higgs boson pair production in the $b\bar{b}\gamma\gamma$ final state in $p-p$ collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 120:071802, Feb 2018.

[36] Lisa Randall and Raman Sundrum. Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, Oct 1999.

[37] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo. Phenomenology of the Randall-Sundrum Gauge Hierarchy Model. *Phys. Rev. Lett.*, 84:2080, 2000.

[38] *LEP design report*. CERN, Geneva, 1984. Copies shelved as reports in LEP, PS and SPS libraries.

[39] Design Report Tevatron 1 project. Technical Report FERMILAB-DESIGN-1984-01, 1984.

[40] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, 2008.

[41] *LHCb : Technical Proposal*. Tech. Proposal. CERN, Geneva, 1998.

[42] Image of CERN accelerator complex. `https://stfc.ukri.org/research/particle-physics-and-particle-astrophysics/large-hadron-collider/cern-accelerator-complex/`.

[43] Website of ATLAS public luminosity results. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`.

[44] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[45] Image of ATLAS magnet system. `http://www.jetgoodson.com/images/thesisImages/magnetSystems.png`.

[46] ATLAS Collaboration. Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV. ATL-PHYS-PUB-2015-018, 2015.

[47] ATLAS Collaboration. Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties. *JINST*, 12:P05002, 2017.

[48] ATLAS Collaboration. Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 78:110, 2017.

[49] R Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Methods Phys. Res., A*, 262(HEPHY-PUB-503):444. 19 p, Jun 1987.

[50] TrackML Particle Tracking Challenge. `https://www.kaggle.com/c/trackml-particle-identification`. Accessed: 2018-10-30.

[51] ATLAS Collaboration. Performance of primary vertex reconstruction in proton–proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment. ATLAS-CONF-2010-069, 2010.

[52] R. Fruhwirth, W. Waltenberger, and P. Vanlaer. Adaptive vertex fitting. *J. Phys.*, G34:N343, 2007.

[53] Walter Lampl et al. Calorimeter Clustering Algorithms: Description and Performance. ATL-LARG-PUB-2008-002, 2008.

[54] ATLAS Collaboration. Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach. ATL-PHYS-PUB-2017-022, 2017.

[55] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[56] ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in *pp* collisions with the ATLAS detector. *Nucl. Instrum. Meth.*, A824:367, 2016.

[57] ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 96:072002, 2017.

[58] ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector. ATLAS-CONF-2014-018, 2014.

[59] ATLAS Collaboration. Performance of b-Jet Identification in the ATLAS Experiment. *JINST*, 11:P04008, 2016.

[60] Quark versus Gluon Jet Tagging Using Charged Particle Multiplicity with the ATLAS Detector. Technical Report ATL-PHYS-PUB-2017-009, CERN, Geneva, May 2017.

[61] ATLAS Collaboration. Technical Design Report for the ATLAS ITk Strip Detector. Technical Report ATL-COM-UPGRADE-2017-006, CERN, Geneva, Mar 2017.

[62] Projections for measurements of Higgs boson cross sections, branching ratios, coupling parameters and mass with the ATLAS detector at the HL-LHC. Technical Report ATL-PHYS-PUB-2018-054, CERN, Geneva, Dec 2018.

[63] Measurement prospects of the pair production and self-coupling of the Higgs boson with the ATLAS experiment at the HL-LHC. Technical Report ATL-PHYS-PUB-2018-053, CERN, Geneva, Dec 2018.

[64] M. Cepeda et al. Higgs Physics at the HL-LHC and HE-LHC. 2019.

[65] Image of the timeline of the HL-LHC program. `https://project-hl-lhc-industry.web.cern.ch/content/project-schedule`.

[66] Image of the crab cavities. `https://home.cern/about/updates/2017/12/crab-cavities-colliding-protons-head`.

[67] ATLAS Collaboration. Technical Design Report for the ATLAS ITk Pixel Detector. Technical Report ATL-COM-ITK-2017-073, CERN, Geneva, Dec 2017.

[68] ATLAS Collaboration. Technical Design Report for the Phase-II Upgrade of the ATLAS LAr Calorimeter. Technical Report CERN-LHCC-2017-018. ATLAS-TDR-027, CERN, Geneva, Sep 2017.

[69] ATLAS Collaboration. Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer. Technical Report CERN-LHCC-2017-017. ATLAS-TDR-026, CERN, Geneva, Sep 2017.

[70] ATLAS Collaboration. Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter. Technical Report CERN-LHCC-2017-019. ATLAS-TDR-028, CERN, Geneva, Sep 2017.

[71] ATLAS Collaboration. Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System. Technical Report CERN-LHCC-2017-020. ATLAS-TDR-029, CERN, Geneva, Sep 2017.

[72] ATLAS Collaboration. Technical Proposal: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade. Technical Report CERN-LHCC-2018-023. LHCC-P-012, CERN, Geneva, Jun 2018.

[73] ATLAS Collaboration. Final Report of the ITk Layout Task Force. Technical Report ATL-COM-UPGRADE-2016-042, CERN, Geneva, Sep 2017.

[74] ATLAS Collaboration. Expected Performance of the ATLAS Inner Tracker at the High-Luminosity LHC. ATL-PHYS-PUB-2016-025, 2016.

[75] ATLAS Collaboration. Prospects for Observing $t\bar{t}HH$ Production with the ATLAS Experiment at the HL-LHC. ATL-PHYS-PUB-2016-023, 2016.

[76] Christoph Englert, Frank Krauss, Michael Spannowsky, and Jennifer Thompson. Di-Higgs phenomenology in $t\bar{t}hh$: The forgotten channel. *Phys. Lett.*, B743:93–97, 2015.

[77] Tao Liu and Hao Zhang. Measuring Di-Higgs Physics via the $t\bar{t}hh \to t\bar{t}b\bar{b}b\bar{b}$ Channel. 2014.

[78] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi. Delphes 3: a modular framework for fast simulation of a generic collider experiment. *Journal of High Energy Physics*, 2014(2):57, Feb 2014.

[79] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016.

[80] ATLAS Run 1 Pythia 8 tunes. Technical Report ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov 2014.

[81] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys.*, B867:244–289, 2013.

[82] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.

[83] ATLAS Phase-II Upgrade Scoping Document. Technical Report CERN-LHCC-2015-020. LHCC-G-166, CERN, Geneva, Sep 2015.

[84] ATLAS Collaboration. Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data. ATLAS-CONF-2014-046, (2014).

[85] ATLAS Collaboration. Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System. Technical Report CERN-LHCC-2017-020. ATLAS-TDR-029, CERN, Geneva, Sep 2017.

[86] ATLAS Collaboration. Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 75:349, 2015.

[87] ATLAS Collaboration. Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, (CERN-EP-2017-291):44 p, Dec 2017.

[88] S. Dawson, S. Dittmaier, and M. Spira. Neutral higgs-boson pair production at hadron colliders: QCD corrections. *Phys. Rev. D*, 58:115012, Nov 1998.

[89] S. Borowka, N. Greiner, G. Heinrich, S.P. Jones, M. Kerner, J. Schlenk, and T. Zirke. Full top quark mass dependence in Higgs boson pair production at NLO. *Journal of High Energy Physics*, 2016(10):107, Oct 2016.

[90] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016.

[91] ATLAS Collaboration. The ATLAS Simulation Infrastructure. *Eur. Phys. J. C*, 70:823, 2010.

[92] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[93] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.

[94] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C. P. Yuan. New parton distributions for collider physics. *Phys. Rev.*, D82:074024, 2010.

[95] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys.*, G43:023001, 2016.

[96] ATLAS Collaboration. Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data. 2016.

[97] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].

[98] John M. Campbell, R. Keith Ellis, Ye Li, and Ciaran Williams. Predictions for diphoton production at the LHC through NNLO in QCD. *JHEP*, 07:148, 2016.

[99] ATLAS Collaboration. Measurements of fiducial cross-sections for $t\bar{t}$ production with one or two additional b-jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *Eur. Phys. J. C*, 76:11, 2016.

[100] A L Read. Presentation of search results: the CLs technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 2002.

[101] ATLAS Collaboration. Combination of searches for Higgs boson pairs in pp collisions at 13 TeV with the ATLAS experiment. Technical report, CERN.

[102] ATLAS Collaboration. Search for new phenomena in high-mass diphoton final states using 37 fb^{-1} of proton–proton collisions collected at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Lett.*, B775:105–125, 2017.