Observation of the Higgs boson in the WW^* channel and search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ channel with the ATLAS detector

A DISSERTATION PRESENTED
BY
Tomo Lazovich
TO
The Department of Physics

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy
IN THE SUBJECT OF
Physics

Harvard University
Cambridge, Massachusetts
May 2016

Dissertation advisors: Professors João Guimarães da Costa and Melissa Franklin

Tomo Lazovich

Observation of the Higgs boson in the WW^* channel and search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ channel with the ATLAS detector

Abstract

This dissertation presents the observation and measurement of the Higgs boson in the $H \to WW^* \to \ell\nu\ell\nu$ channel at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV and a search for Higgs pair production in the $HH \to b\bar{b}b\bar{b}$ channel at $\sqrt{s} = 13$ TeV with the ATLAS detector in pp collisions at the Large Hadron Collider.

First, the discovery of a particle consistent with the Higgs boson in $4.8\,\text{fb}^{-1}$ at $\sqrt{s} = 7$ TeV and $5.8\,\text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV is discussed. Then, the measurement of the Higgs boson signal strength and cross section in both the gluon fusion and vector boson fusion (VBF) production modes using $20.3\,\text{fb}^{-1}$ of $\sqrt{s} = 8$ TeV data combined with $4.8\,\text{fb}^{-1}$ of 7 TeV data is shown. The combined signal strength is measured to be $\mu = 1.09^{+0.23}_{-0.21}$. The total observed significance of the $H \to WW^*$ process is observed to be 6.1σ (with 5.8σ expected). Advanced methods for background reduction and estimation, particularly in same-flavor lepton final states, are shown. The VBF signal strength is measured to be $\mu_{\text{VBF}} = 1.27^{+0.53}_{-0.45}$ with an observed significance of 3.2σ (with 2.7σ expected). In the VBF channel, a selection requirement based method, the precursor to the final multivariate technique used for the result, is detailed.

Finally, a search for Higgs pair production in the $b\bar{b}b\bar{b}$ final state with $3.2\,\text{fb}^{-1}$ at $\sqrt{s} = 13$ TeV is presented. A particular focus is placed on a tailored signal region for resonant production of Higgs pairs at high masses. No significant excesses are observed, and upper limits on cross sections are placed for spin-2 Randall Sundrum gravitons (RSG) and narrow spin-0 resonances. The cross section of $\sigma(pp \to G^*_{\text{KK}} \to hh \to b\bar{b}b\bar{b})$ with $k/\bar{M}_{\text{Pl}} = 1$ is constrained to be less than 70 fb for masses in the range $600 < m_{G^*_{\text{KK}}} < 3000$ GeV. The cross section upper limits for $\sigma(pp \to H \to hh \to b\bar{b}b\bar{b})$ ranges from 30 to 300 fb in the mass range of $500 < m_H < 3000$ GeV.

Contents

Listing of figures

Listing of tables

For my family

Acknowledgments

I have been a member of the Harvard ATLAS group for many years now, first as an undergraduate and then as a graduate student. As a result, I have had the privilege of interacting with many amazing people there over the years and have accumulated a large list of folks to thank.

First and foremost, I must thank the two people who have effectively been my academic parents since I started in the Harvard group: João Guimarães da Costa and Melissa Franklin. Melissa Franklin and João Guimarães da Costa. They have both been so important to both my academic and personal development that I can't even put one before the other. João has been an excellent PhD advisor, showing me how to look at the big picture and helping me navigate the sometimes complicated politics of ATLAS. He got me started on my first projects with ATLAS as a young college sophomore before there was even beam in the LHC (go cosmic ray muons!). He has also been a constant source of advice and support, even when we have been on different continents. Melissa gave me my start in HEP as a summer student on CDF and has been an unbelievable mentor throughout my time at Harvard. I still remember our weekly chalkboard particle physics lessons after that first summer. She also graciously took me on as a co-advisee after João moved on to his new position at IHEP. I am incredibly lucky to have had both of them as advisors.

Another mentor who was essential to my development as a graduate student is Paolo Giromini. His uncanny knowledge and intuition about detectors is unmatched and I am very grateful to have had the chance to work with him on the micromegas for the ATLAS New Small Wheel upgrade project. I owe essentially all my practical knowledge about detectors (and building things in general) to him. I also appreciated his unique sense of humor which made sometimes difficult tasks much easier to get through.

I am grateful to John Huth and Masahiro Morii for their helpful advice as the other professors in the Harvard ATLAS group as well. I especially thank John for helping me get started on the micromegas trigger project and being a great professor to TF particle physics for. Additionally I thank Howard Georgi for serving as my third committee member and offering me feedback throughout my graduate career.

I also owe enormous thanks to Hugh Skottowe, the postdoc that I worked most closely with in my early years as a graduate student. He was always able to help me through complicated tasks in everything from writing code to understanding difficult physics concepts. I particularly enjoyed walking down to his office in Palfrey at random times and talking through whatever problem I was tackling on that day.

Alex Tuna, the second postdoc that I worked closely with at Harvard, deserves great thanks as well. He helped me push through to the end of my graduate career and offered great advice along the way.

Being at Harvard, I have seen an incredible array of graduate students graduate before me: Ben Smith, Verena Martinez Outschoorn, Srivas Prasad, Michael Kagan, Giovanni Zevi Della Porta, Laura Jeanty, Kevin Mercurio, William Spearman, and Andy Yen. I want to thank them all for showing me what a good physicist looks like and for patiently answering my questions and offering insightful advice about physics and life.

Getting through graduate school would not have been possible without the support and friendship of the other students in our group. Thanks to Emma Tolley for geeking out with me about cool computing stuff, going to taste delicious beers with me, and helping start the Palfrey tradition of Taco Tuesdays. Thanks to Brian Clark for being a great friend and housing companion both in Kirkland House and in our tiny summer apartment in Geneva (and thanks to his partner Allison Goff for the same reasons!). Thanks to Siyuan Sun for giving me my first aikido lesson and always being there for great conversations, big and small. Tony (Baojia) Tong deserves special recognition for working with me on the $4b$ analysis and putting up with my sometimes strange requests (and giving me rides to the Val Thoiry Migros so I wouldn't have to pay exorbitant Geneva grocery prices!). Stephen Chan is probably the only student in the group who both understands my references to the Sopranos and makes some of his own. To the younger graduate students - Karri Di Petrillo, Jennifer Roloff, Julia Gonski, and Ann Wang - I want to say thank you for making the group a fun and lively place to be and giving all of us energy that the older graduate students like myself can sometimes lack.

I'd like to thank Annie Wei and Gray Putnam, the two undergraduates I have worked with as a graduate student. Their unbelievable intuition and quickness in picking up difficult particle physics concepts is inspiring.

I would also like to thank all of the postdocs that I have interacted with in my time in the Harvard group: Kevin Black, Alberto Belloni (who would always ask me "Do you have it?"...I can now say that I do!), Shulamit Moed, Corrinne Mills, Geraldine Conti, David Lopez Mateos, Chris Rogan, Valerio Ippolito, and Stefano Zambito.

There are many people on ATLAS who have helped me get to this point as well. In the WW group, I have to thank Jonathan Long, Joana Machado Miguens, Ben Cerio, Philip Chang, Bonnie Chow, Richard Polifka, Heberth Torres, Tae Min Hong, and Jennifer Hsu for being wonderful colleagues and making the entire analysis run smoothly. In the $4b$ group, I have to thank Qi Zeng, Tony Tong, Alex Tuna, Michael Kagan, Max Bellomo, John Alison, and Patrick Bryant.

Kirkland House was my home for the last three years of graduate school and was an wonderful environment and support system. I want to thank my fellow tutors, especially Brian Clark and Allison Goff (again), Zach Abel, Kelly Bodwin, Alex Lupsasca, John and Pam Park, Luke and Erin Walczewski, and Philip Gant for their friendship and support. I also want to thank Kate Drizos Cavell, Bob Butler, and the Faculty Deans Tom and Verena Conley.

There are still a few friends that haven't been covered yet and deserve great thanks. Jake Connors and Meredith MacGregor have been absolutely wonderful friends and I thank them in particular for the many home-cooked meals and great conversations we've had in their apartment. Nihar Shah has been my friend and confidant since we were both wee freshmen in Harvard Yard. Gareth Kafka, though he sits on the "neutrino" side of Palfrey House, has made days there more fun and has also been an enthusiastic participant in the Palfrey Taco Tuesdays.

Being at Harvard necessarily means having to navigate through bureaucracy at some point or another. I thank Lisa Cacciabaudo, Carol Davis, and Jacob Barandes for always having open doors and being the most kind, helpful people in the Physics department.

I thank Venky Narayanamurti for putting on a great SPU course that I was proud to be a part of and TF for. I'd also like to thank Jim Waldo for offering me much advice about working in Computer Science and giving me a fun data project to be a part of in my free time.

I grew up in a very tight knit Serbian community on the south side of Chicago which helped make me

the person I am today. I would like to thank all of the people at St. Simeon Mirotochivi Serbian Orthodox Church who have always been sources of enthusiasm and support in my life.

I would not be here without the unconditional love and constant support and encouragement of my family. To my pokojni Dedas Branko and Miloje, my pokojni Baba Milka, and my Baba Desa, I want to say thank you for instilling in me at an early age the love of curiosity and storytelling that I have carried throughout my life. To my sister Angelina, I want to say thank you for always loving me and being my partner in crime throughout our childhoods. To my parents, Miroljub and Nada, Tata and Mama, I really cannot express how grateful I am to you and how much I owe you. As I look back now I see how I am a combination of both of your best qualities and every day I am in situations where I understand more and more the lessons you made sure to teach me and the sacrifices you made to make sure I got the best possible education. I love you all.

Finally, I have to thank my soul mate, the one person in my life who understands me more than anyone else, my fiancée Kelly Brock. You are my sounding board, my support system, my cheerleader (figuratively and literally!), my best friend, my role model, and my everything. I would not have gotten through graduate school without you and my life would not be the same without you. I cannot wait to start our new lives together as the married doctors, tackling whatever comes our way with the same zeal with which we tackled graduate school. I love you with all my heart and soul.

0

Introduction

The Higgs boson is often described as one of the cornerstones of particle physics. When the Standard Model was first developed as a theory to describe the fundamental particles and forces of nature, physicists were faced with a dilemma. The electroweak theory beautifully characterized both electromagnetism and the weak force with a single underlying framework. However, the mass of the weak W and Z bosons was puzzling given the fact that their electromagnetic counterpart, the photon, is massless. The Higgs mechanism was developed as the leading theory for the origin of this electroweak symmetry breaking. It predicted the existence of an additional spin-0 boson in the Standard Model, the Higgs boson. Generations of collider experiments searched for this elusive particle. This dissertation presents research work on the Higgs boson from its discovery to its use as a tool in the search for physics beyond the Standard Model with the ATLAS detector at the Large Hadron Collider (LHC).

One of the first priorities for the LHC when it began colliding proton beams in 2010 was the search for the Higgs boson. This search was initially tackled in three main channels: $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^* \rightarrow 4\ell$, and $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$. Each channel has its own merits, but the WW^* mode is particularly suited to searching over a wide range of masses. The $H \rightarrow WW^*$ branching ratio is large and it is the primary decay channel above the $2m_W$ mass threshold. Despite the fact that the full Higgs invariant mass cannot be reconstructed in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ channel, its signal to background ratio makes it ideal for measurement of Higgs properties such as the production cross section and couplings.

In 2012, the ATLAS and CMS experiments announced the discovery of a new particle consistent with the Higgs boson [1, 2]. In ATLAS, this discovery was made with $4.8\,\mathrm{fb}^{-1}$ collected at $\sqrt{s} = 7\,\mathrm{TeV}$ and $5.8\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 8\,\mathrm{TeV}$. This dissertation first presents the search for gluon fusion production of the Higgs in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ channel, which played an important role in this discovery. Selection requirements which were optimized to maximize the discovery significance in this channel, as well as background estimation procedures, are discussed.

After its discovery, interest in the Higgs shifted to focus on the measurement of its properties. As a result, extensions of the initial discovery analysis in larger datasets had two main goals. Improvement of signal to background ratio was important to allow for precision measurements. Also, searches for rarer production modes of the Higgs were a priority. The first such extension presented in this dissertation is a tailored selection for $\ell\nu\ell\nu$ final states with same flavor leptons. Novel variables for the reduction of the Z+jets background that could remain robust under increasing LHC instantaneous luminosities are shown. The second post-discovery result shown is the first evidence of Vector Boson Fusion (VBF) production of the Higgs boson.

VBF production of the Higgs boson is particularly interesting in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ final state. In this combination of production and decay modes, the Higgs boson couples exclusively to W bosons, allowing for precise measurement of the Higgs-W coupling constant. However, it is challenging to observe VBF Higgs production because its cross section at the LHC is an order of magnitude lower than gluon fusion production. The large $H \rightarrow WW^*$ branching ratio thus presents another advantage over other final states. Additionally, VBF production of the Higgs boson creates two forward jets in addition to the Higgs, and these jets can be used to isolate VBF Higgs events from other production modes. The VBF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ analysis first created a selection requirement based signal region using variables constructed specifically for the VBF Higgs production topology. This "cut-based" analysis is presented in detail in this dissertation. These VBF topology variables, once validated in the cut-based analysis, were then input into a multivariate boosted decision tree discriminant to achieve the first evidence of VBF Higgs production with the full $20.3\,\text{fb}^{-1}$ of $\sqrt{s} = 8$ TeV data in ATLAS.

After a two year shutdown, the LHC restarted in 2015 with a center of mass energy of $\sqrt{s} = 13$ TeV. This increase improved the LHC's ability to probe for physics beyond the Standard Model, and the Higgs sector remained one of the largest regions of unprobed phase space where such new physics could be discovered. Production of high mass resonances benefited most from the center of mass energy increase. In particular, the cross section for a generic gluon-initiated 2 TeV resonance increased tenfold with the increase from 8 to 13 TeV. Therefore, a natural next step in studies of the Higgs was a search for a new heavy resonance which decays into a pair of Higgs bosons. The final result shown in this dissertation is a search for resonant di-Higgs production in the $X \to HH \to b\bar{b}b\bar{b}$ final state with $3.2\,\text{fb}^{-1}$ recorded by ATLAS at $\sqrt{s} = 13$ TeV. This search has the unique advantage that it can both probe physics beyond the Standard Model and gain further understanding of the Standard Model through constraints on SM pair production of the Higgs.

As mentioned above, this dissertation begins by discussing the discovery of the Higgs and the role of the $H \to WW^* \to \ell\nu\ell\nu$ channel. It then presents the first evidence for the VBF production mode using the $H \to WW^* \to \ell\nu\ell\nu$ channel with the full ATLAS Run 1 dataset. It also shows the final combined Run 1 measurements of gluon fusion Higgs production from this channel. Finally, it presents a search for Higgs pair production in the $HH \to b\bar{b}b\bar{b}$ channel. It is organized into four parts.

Part 1 presents the theoretical and experimental background required for the subsequent parts. Chapter 1 gives an overview of Higgs physics, particularly single and double Higgs production in the Standard Model and beyond. Chapter 2 presents details regarding the Large Hadron Collider and the ATLAS experiment. The evolution of machine conditions, descriptions of the ATLAS sub-detectors, and an overview of object reconstruction in ATLAS are all shown. A brief interlude on the ATLAS Muon New Small Wheel upgrade is also given, as this upgrade has been a focus of my graduate work and will have an important impact on ATLAS' ability to study the Higgs at the High Luminosity LHC.

Part 2 discusses the observation and measurement of the Higgs in the $H \to WW^* \to \ell\nu\ell\nu$ channel in the ATLAS Run 1 dataset at $\sqrt{s} = 7$ and 8 TeV. Because I worked in this channel from before the discovery through to the final analysis of the Run 1 dataset, Part 2 is organized in such a way to allow easy presentation of multiple analyses on different subsets of the full Run 1 dataset. Chapter 3 presents a general overview of the $H \to WW^*$ analysis strategy and defines many of the variables and common elements used in the rest of Part 2. Chapter 4 presents the discovery of the Higgs boson, focusing on the role of the WW^* channel in this discovery. Chapter 5 presents the first evidence for the VBF production mode of the Higgs, a result from the WW^* channel in the full Run 1 ATLAS dataset. In this chapter, the focus is mainly on the cut-based VBF analysis. The cut-based analysis was an important first step to the final VBF result which used a boosted decision tree. Where appropriate, connections between the cut-based and BDT analyses are shown and their compatibility is discussed. Finally, the VBF analysis was an important input into the combined Run 1 $H \to WW^* \to \ell\nu\ell\nu$ result, which used both the gluon fusion and VBF channels in a combined fit to infer properties of the Higgs, including its couplings to the gauge bosons and its production cross section. This is the topic of Chapter 6.

Part 3 presents a search for Higgs pair production in the $HH \to b\bar{b}b\bar{b}$ channel. Chapter 7 presents an overview of this search in the boosted regime, where the Higgs pairs are the result of the decay of a heavy resonance. Chapter 8 shows the combined results between the boosted regime and the resolved regime, which is sensitive to lower mass resonances and non-resonant Higgs pair production. Finally, Part 4 presents a conclusion and brief outlook of future Higgs physics with ATLAS.

Part I

Theoretical and Experimental Background

In modern physics, there is no such thing as "nothing."

Richard Morris

1

The Physics of the Higgs Boson

This chapter presents an overview of the Standard Model of Particle Physics and in particular the physics of the Higgs boson. First, a brief overview of the Standard Model is presented. Then, a description of the Higgs mechanism of electroweak symmetry breaking is given. Next, the physics of single Higgs boson production and decay is described. The Standard Model also allows for production of two Higgs bosons and this is detailed as well. Finally, di-Higgs production in two beyond the Standard Model (BSM) theories - Randall-Sundrum gravitons (RSG) and Two Higgs Doublet Models (2HDM) - is shown.

1.1 The Standard Model of Particle Physics

The Standard Model (SM) of Particle Physics is a quantum field theory describing the fundamental particles of nature and the forces that govern their interactions. Several comprehensive treatments of the SM already exist in the literature [3–8] and this section will not rehash those. Rather, this section presents a brief overview of the SM particles and forces in order to define them for subsequent discussions.

The Standard Model consists of two primary categories of fundamental particles: fermions (spin $1/2$ particles) and bosons (integer spin particles). The SM also describes three forces: electromagnetism, the weak nuclear force, and the strong nuclear force. Gravity is not included in the theory and is largely irrelevant at the scales currently probed by collider experiments. Within the fermions, there are both quarks (which interact via all three forces) and the leptons. The charged leptons interact via electromagnetic and weak interactions, while neutrinos (neutral leptons) interact only via the weak force. Within the bosons, there are the $W^{\pm}$ and Z bosons (the mediators of the weak force), the gluon (g, the mediator of the strong force), and the photon (γ), the mediator of the electromagnetic force. Finally, there is the Higgs boson, a fundamental spin-0 particle resulting from the Higgs mechanism of electroweak symmetry breaking. Figure 1.1 summarizes the fermions and bosons of the SM.

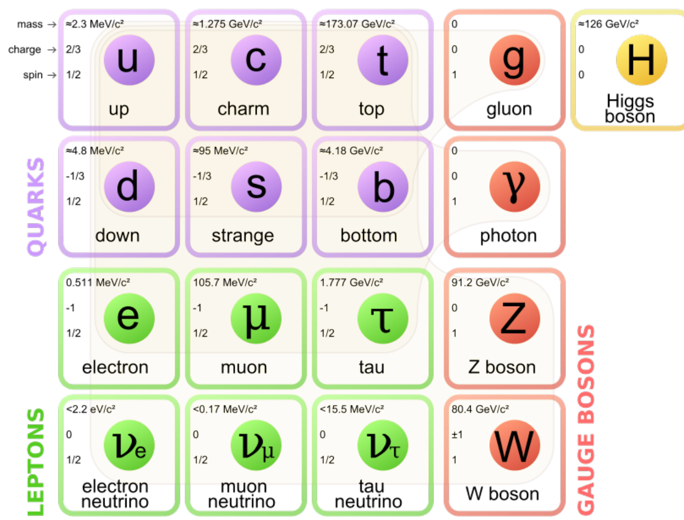


Figure 1.1: The particles of the Standard Model and their properties [6].

The Standard Model coalesced into a unified theoretical framework in the 1960s through the work of Glashow, Weinberg, Salam, and others on the theory of electroweak interactions [9–12]. This theory characterized both the electromagnetic and weak interactions as unified under a single gauge symmetry group, namely $SU(2) \times U(1)$. At low enough energy scales (on the order of the W and Z masses, the electroweak symmetry is broken, as evidenced by the fact that the weak bosons have mass while the photon does not. The discovery of the Higgs boson in 2012 confirmed the Higgs mechanism as the most likely candidate for this electroweak symmetry breaking [1, 2]. The electroweak theory is then combined with the theory of quantum chromodynamics (which models the strong sector as a non-Abelian $SU(3)$ gauge group) to form the complete SM [13].

1.2 Electroweak Symmetry Breaking and the Higgs

In the Standard Model Lagrangian, it is difficult to include mass terms for the W and Z bosons without breaking the fundamental gauge symmetry of the Lagrangian. A traditional mass term does not preserve the $SU(2) \times U(1)$ symmetry. Additionally, scattering of massive W and Z bosons violate unitarity and these diagrams diverge at high energy scales. In the 1960s, Higgs, Brout, Englert, Guralnik, Kibble, and Hagen developed a mechanism for spontaneous symmetry breaking via the additional of a complex scalar doublet to the SM. Three of the four real degrees of freedom of this complex field would go to the longitudinal modes of the $W^{\pm}$ and Z, thus allowing them to have mass [14–17]. The remaining degree of freedom would manifest as an additional scalar, known now as the Higgs boson.

The mechanism works by introducing a Lagrangian for the newly introduced field that still respects the symmetry of the Standard Model inherently, but with a minimum at a non-zero vacuum expectation value for the field. In this minimum of the potential, the electroweak symmetry is broken. Specifically, consider a complex scalar doublet $\mathbf{\Phi}$ with four degrees of freedom, as shown in equation 1.1.

$$\mathbf{\Phi} = \begin{pmatrix} \phi^{+} \\ \phi^{0} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1^{+} + i\phi_2^{+} \\ \phi_1^{0} + i\phi_2^{0} \end{pmatrix} \tag{1.1}$$

The minimal potential of a self-interacting Higgs that still respects the SM symmetry is given in equation 1.2.

$$V(\mathbf{\Phi}) = \mu^2 \mathbf{\Phi}^{\dagger}\mathbf{\Phi} + \lambda(\mathbf{\Phi}^{\dagger}\mathbf{\Phi})^2 \tag{1.2}$$

If the μ^2 term of this potential is positive, then the potential has a minimum at $\mathbf{\Phi} = 0$ and the SM symmetry is preserved. However, if instead $\mu^2 < 0$, then the minimum is at a finite value of $\mathbf{\Phi}$, namely

$$\mathbf{\Phi}_{\text{min}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.3}$$

where $v = \sqrt{\mu^2/\lambda}$. Because this is the location of the minimum, it corresponds to the vacuum expectation value for the field ($\langle\mathbf{\Phi}\rangle = \mathbf{\Phi}_{\text{min}}$). The excitations of the Higgs can then be parameterized as

$$\mathbf{\Phi} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H \end{pmatrix} \tag{1.4}$$

The full scalar Lagrangian, including the kinetic term, is then given as

$$\mathcal{L}_s = (D^\mu\mathbf{\Phi})^\dagger(D_\mu\mathbf{\Phi}) - V(\mathbf{\Phi}) \tag{1.5}$$

where the covariant derivative is defined as

$$D_\mu = \partial_\mu + \frac{ig}{2}\tau^a W_\mu^a + ig'YB_\mu \tag{1.6}$$

and W^1, W^2, W^3 and B are the $SU(2)$ and $U(1)$ gauge fields of the electroweak theory, respectively. g and g' are the corresponding coupling constants. With the scalar Lagrangian in place, the physical gauge fields can then be written as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \tag{1.7}$$

$$Z_\mu = \frac{-g'B_\mu + gW_\mu^3}{\sqrt{g^2 + g'^2}} \tag{1.8}$$

$$A_\mu = \frac{gB_\mu + g'W_\mu^3}{\sqrt{g^2 + g'^2}} \tag{1.9}$$

Equation 1.7 corresponds to the charged W^+ and W^- bosons, equation 1.8 corresponds to the neutral Z boson, and equation 1.9 corresponds to the neutral photon. The masses of the particles also arise from the Lagrangian. The photon has zero mass, while the masses of the W and Z bosons are given in equation 1.10.

$$\begin{aligned} M_W^2 &= \tfrac{1}{4}g^2v^2 \\ M_Z^2 &= \tfrac{1}{4}(g^2 + g'^2)v^2 \end{aligned} \tag{1.10}$$

The fermion masses also arise through a coupling with the Higgs via the Yukawa interaction (for a detailed description, see [8]). In this case the coupling between the Higgs and the fermions goes as

$$g_{Hf\bar{f}} = \frac{m_f}{v} \tag{1.11}$$

The full Lagrangian of Higgs interactions can be written as

$$\mathcal{L}_{\text{Higgs}} = -g_{Hf\bar{f}}\bar{f}fH + \frac{g_{HHH}}{6}H^3 + \frac{g_{HHHH}}{24}H^4 + \delta_V V_\mu V^\mu \left(g_{HVV}H + \frac{g_{HHVV}}{2}H^2\right) \tag{1.12}$$

with

$$\begin{aligned} g_{HVV} &= \tfrac{2m_V^2}{v} \quad & g_{HHVV} &= \tfrac{2m_V^2}{v^2} \\ g_{HHH} &= \tfrac{3m_H^2}{v} \quad & g_{HHHH} &= \tfrac{3m_H^2}{v^2} \end{aligned} \tag{1.13}$$

Here, V refers to the $W^\pm$ and Z, and $\delta_W = 1$ while $\delta_Z = 1/2$. Phenomenologically, there are a few features of this Lagrangian that are useful to note. First, note that the Higgs mass is a free parameter of the theory that must be determined experimentally. Second, note that the coupling of the Higgs to the vector bosons and fermions scales with the masses of these particles, a fact that is important when considering both the production and decays of the particle. Also note that the branching ratio of the Higgs to W bosons will be twice that of the branching ratio to Z if the Higgs mass is large enough to produce the particles on shell because of the extra symmetry factor associated with the W coupling. Finally, note the presence of the cubic and quartic Higgs self interaction terms, which can lead to final states with multiple Higgs bosons produced.

1.3 Higgs Boson Production and Decay

This section discusses the properties of Higgs production and decay mechanisms. The details presented here will focus on the properties of a 125 GeV Higgs boson, as this is the mass closest to that of the newly discovered Higgs.

1.3.1 Higgs production

The Higgs is produced by four main production modes at the Large Hadron Collider - gluon-gluon fusion (ggF), vector boson fusion (VBF), associated production with a W or Z boson, or associated production with top quarks ($t\bar{t}H$). Figure 1.2 shows the Feynman diagrams for these four modes.

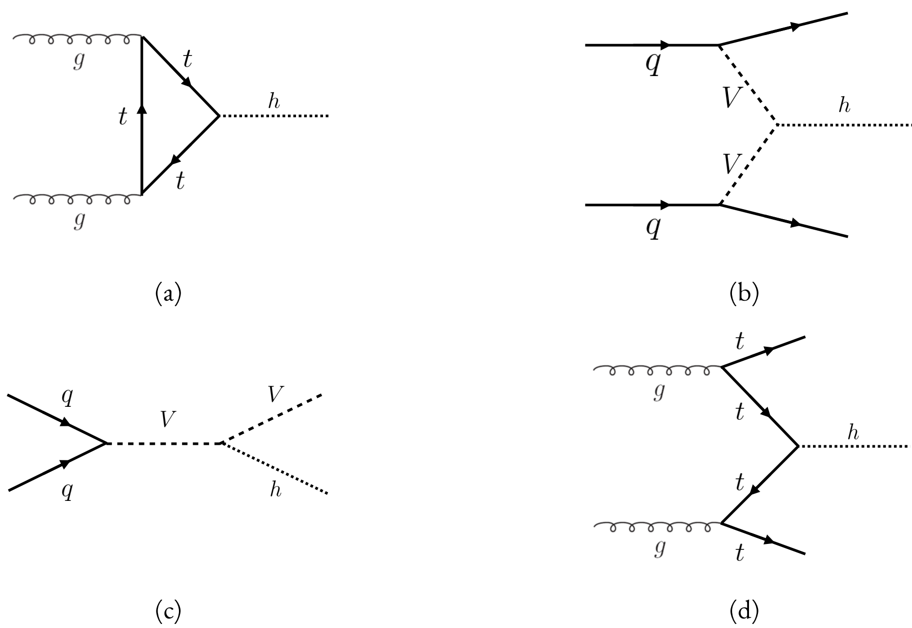


Figure 1.2: The four most common Higgs boson production modes at the LHC: (a) gluon-gluon fusion, (b) vector boson fusion, (c) $W/Z + H$ production, (d) $t\bar{t}H$ production

In gluon-gluon fusion, gluons from the incoming protons fuse via a top-quark loop to produce a Higgs. The top quark is the dominant contribution in the loop due to its heavy mass and the fact that the Higgs-fermion coupling constant scales with fermion mass. In vector boson fusion, the incoming quarks each radiate a W or Z boson which fuse to produce the Higgs. This production mode results in a final state with a Higgs boson and two additional jets which tend to be forward because they carry the longitudinal momentum of the incoming partons. The Higgs can also be produced in associated with a W or Z boson. The W/Z is produced normally and then radiates a Higgs (this mode is also sometimes known as "Higgs-strahlung"). Finally, the Higgs can be produced in association with two top quarks. Each incoming gluon splits into a $t\bar{t}$ pair, and one of the top pairs combines to create a Higgs. Figure 1.3 shows the production cross section for a 125 GeV Higgs boson in each of these modes at a pp collider as a function of center of mass energy.

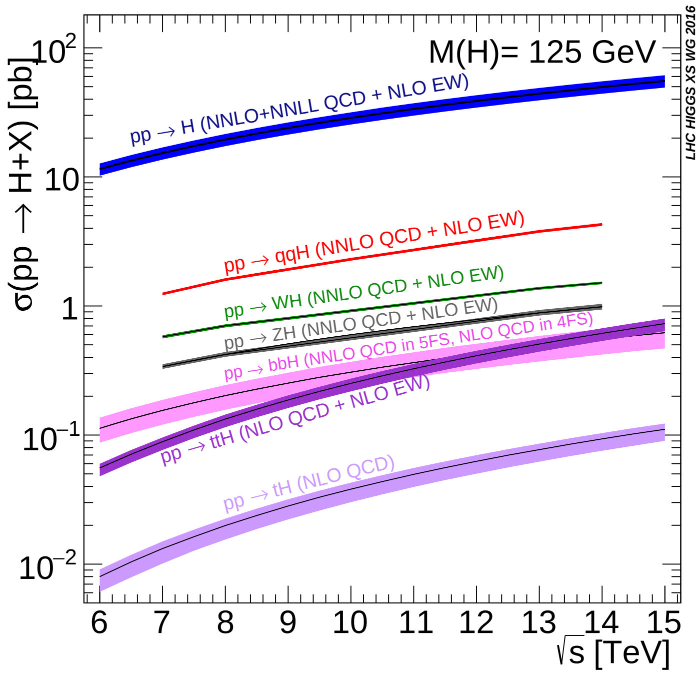


Figure 1.3: Higgs production cross sections as a function of center of mass energy ($\sqrt{s}$) at a pp collider [18].

In figure 1.3, note that gluon fusion has the largest cross section, while VBF is the second largest at approximately a factor of 10 smaller. The figure also includes the less commonly studied $b\bar{b}H$ and tH modes. While the $b\bar{b}H$ mode has a larger cross section than $t\bar{t}H$, it also has larger backgrounds and is thus less sensitive. The tH mode is not as sensitive as $t\bar{t}H$ due to its lower cross section. At $\sqrt{s} = 8$ TeV, ggF production of a 125 GeV Higgs has a cross section of 19.47 pb, while VBF has a cross section of 1.601 pb [18]. The cross sections of all of the main Higgs production modes at this center of mass energy, as well as their uncertainties from varying the renormalization and factorization scales and PDFs, are summarized in table 1.1 for a 125 GeV Higgs.

Production mode	σ (pb)	QCD scale uncert. (%)	PDF + α_s uncert. (%)
Gluon fusion	19.47	$+7.3/-8.0$	3.1
Vector boson fusion	1.601	$+0.3/-0.2$	2.2
WH	0.7026	$+0.6/-0.9$	2.0
ZH	0.4208	$+2.9/-2.4$	1.7
$b\bar{b}H$	0.2021	$+20.7/-22.3$	
$t\bar{t}H$	0.1330	$+4.1/-9.2$	4.3
tH (t-channel)	0.01869	$+7.3/-16.5$	4.6
tH (s-channel)	1.214×10^{-3}	$+2.8/-2.4$	2.8

Table 1.1: Production cross sections for a 125 GeV Higgs boson at $\sqrt{s}=8$ TeV with scale and PDF uncertainties [18].

1.3.2 Higgs branching ratios

The fact that the Higgs couples more strongly to more massive particles is crucial for understanding its branching ratios. The width for Higgs decays to fermions is given in equation 1.14 [5].

$$\Gamma(H\to f\bar{f})=\frac{N_c\sqrt{2}G_F m_f^2 m_H}{8\pi} \tag{1.14}$$

In this case, N_c is the number of colors, G_F is the Fermi constant, m_f is the mass of the fermion, and m_H is the mass of the Higgs. Note that the width scales with the square of the fermion mass. (This also assumes that the Higgs mass is large enough to decay with both the fermions on shell.)

The decay width to WW, in the case where both W bosons are produced on shell ($m_H\geq 2m_W$), is given in equation 1.15 [5].

$$\Gamma(H\to W^+W^-)=\frac{\sqrt{2}G_F M_W^2 m_H}{16\pi}\frac{\sqrt{1-x_W}}{x_W}\left(3x_W^2-4x_W+4\right) \tag{1.15}$$

where m_W is the mass of the W and $x_W=4M_W^2/m_H^2$. To get the branching ratio to ZZ (in the regime where $m_H\geq 2m_Z$), the equation is divided by 2 to account for identical particles in the final state, and x_W is replaced with $x_Z=4M_Z^2/m_H^2$. This is shown in equation 1.16 [5].

$$\Gamma(H\to ZZ)=\frac{\sqrt{2}G_F M_Z^2 m_H}{32\pi}\frac{\sqrt{1-x_Z}}{x_Z}\left(3x_Z^2-4x_Z+4\right) \tag{1.16}$$

The more general formula for Higgs branching into WW or ZZ, taking into account the case where one or both vector bosons is off-shell, is shown in equation 1.17 [19].

$$\Gamma(H \to V^*V^*) = \frac{1}{\pi^2} \int_0^{M_H^2} \frac{dq_1^2 M_V \Gamma_V}{(q_1^2 - M_V^2)^2 + M_V^2 \Gamma_V^2} \int_0^{(M_H - q_1)^2} \frac{dq_2^2 M_V \Gamma_V}{(q_2^2 - M_V^2)^2 + M_V^2 \Gamma_V^2} \Gamma_0 \quad (1.17)$$

Here, q_1^2 and q_2^2 are the invariant masses of the virtual gauge bosons, M_V is the W or Z mass, and Γ_V is the W or Z width. Γ_0 is the squared matrix element, which is given in equation 1.18 [19].

$$\Gamma_0 = \frac{G_F M_H^3}{8\sqrt{2}\pi} \delta_V \sqrt{\lambda(q_1^2, q_2^2, M_H^2)} \left[\lambda(q_1^2, q_2^2, M_H^2) + \frac{12 q_1^2 q_2^2}{M_H^4} \right] \quad (1.18)$$

The function λ is defined as $\lambda(x, y, z) = (1 - x/z - y/z)^2 - 4xy/z^2$. The integral in the general off-shell boson case is much more difficult to interpret than the simpler on-shell branching ratios, but it can be evaluated numerically. These formulas can also be visualized as a function of Higgs mass. Figure 1.4 shows the branching ratios as a function of the Higgs mass. There are a few interesting features to note in this

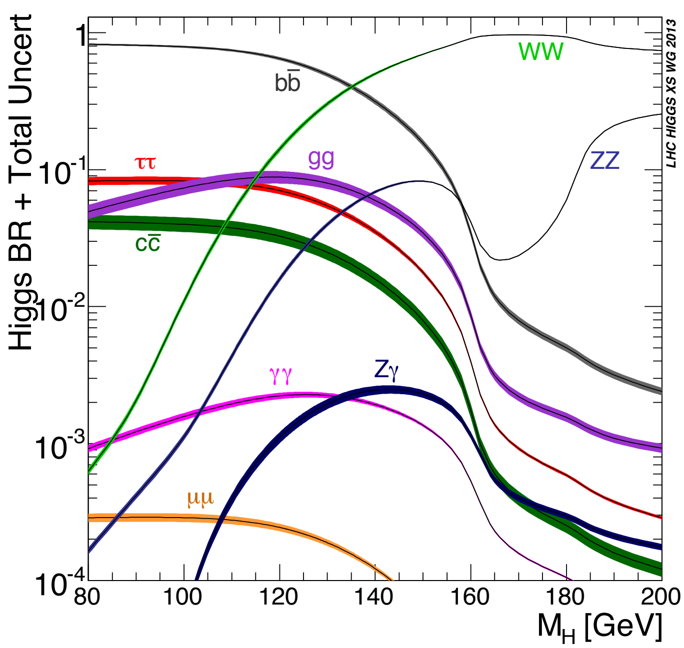


Figure 1.4: Higgs boson branching ratios as a function of m_H [18].

figure. First, note that at high Higgs masses, once on-shell production of both W and Z bosons is possible, these two decays are the dominant ones due to the large masses of the W/Z. Also note that the branching ratio to Ws is twice that of Zs at these large masses due to the δ_V symmetry factor noted previously. At 125 GeV, the Higgs is accessible through many different decay modes. The largest branching ratio is the decay $H \to b\bar{b}$ at 58.24% [18]. This branching is larger than the WW/ZZ decays because one of the two bosons must be produced off-shell for $m_H = 125$ GeV. The second largest branching ratio is to WW^* at 21.37 % (before taking into account the branching ratios of the W). Table 1.2 summarizes the branching ratios for a 125 GeV Higgs. Note that there is in fact a Higgs branching ratio to $\gamma\gamma$ even though photons are massless. This decay happens through a loop (the largest contributions to the loop are top and W) which suppresses the branching ratio.

Decay	Branching ratio (%)	Relative uncertainty (%)
$b\bar{b}$	58.24	$+0.25/-0.25$
WW^*	21.37	$+0.99/-0.99$
gg	8.187	$+3.40/-3.41$
$\tau\tau$	6.272	$+1.17/-1.16$
$c\bar{c}$	2.891	$+1.20/-1.20$
ZZ^*	2.619	$+0.99/-0.99$
$\gamma\gamma$	0.2270	$+1.73/-1.72$
$Z\gamma$	0.1533	$+5.71/-5.71$
$\mu\mu$	0.02176	$+1.23/-1.23$

Table 1.2: Branching ratios for a 125 GeV Higgs boson, quoted as a percentage of the total width of the Higgs. Uncertainties shown are relative to the branching ratio value [18].

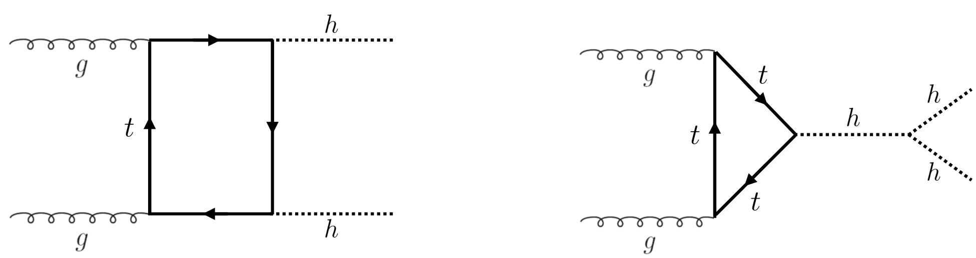
Note that the branching ratios alone do not tell the full story of which Higgs channels are the most sensitive. For example, a $H \to b\bar{b}$ search in gluon fusion production is incredibly difficult due to the large QCD dijet background at the LHC. However, in associated production of the Higgs, where a W or Z gives additional final state particles that can be used to reduce background, a search for $H \to b\bar{b}$ can be sensitive. The combinations of production and decay modes that are most commonly studied are summarized in table 1.3 [5].

Decay	Inclusive (incl. ggF)	VBF	WH/ZH	$t\bar{t}H$
$H \to \gamma\gamma$	✓	✓	✓	✓
$H \to b\bar{b}$			✓	✓
$H \to \tau^+\tau^-$		✓		
$H \to WW^* \to \ell\nu\ell\nu$	✓	✓	✓	
$H \to ZZ \to 4\ell$	✓			
$H \to Z\gamma \to \ell\ell\gamma$	very low			

Table 1.3: Possible channels for Higgs searches. Checkmarks denote the most sensitive production modes for each decay channel [5].

1.4 Higgs Pair Production in the Standard Model

The Standard Model also allows for processes that produce two Higgs bosons in the final state, known as Higgs pair production or di-Higgs production. The two main production mechanisms are shown in figure 1.5. The two diagrams in figure 1.5 interfere destructively with one another, resulting in a low overall



(a) (b)

Figure 1.5: The two leading diagrams for Standard Model di-Higgs production at the LHC: (a) box diagram, (b) Higgs self coupling

cross section for di-Higgs production at the LHC. Nevertheless, Higgs pair production is quite interesting to study because it gives direct access to the λ parameter of the Higgs potential, also known as the Higgs self coupling. The diagram in figure 1.5(b) is sensitive to this coupling through the triple Higgs vertex.

One can substitute the gluon fusion production of diagram 1.5(b) with any of the other production modes previously discussed. These other modes do not suffer from interference with the box diagram in figure 1.5(a) due to the presence of additional particles in the final state. They still have a lower cross section than the gluon fusion mode, however. The cross sections for di-Higgs production in the different modes, as well as their uncertainties, are shown in table 1.4 [20]. These are shown for $\sqrt{s} = 14$ TeV as the higher center of mass energy is more sensitive to this process. Note that the scale of cross section quoted is now in fb rather than pb.

Production mode	σ (fb)	Total uncert. (%)
Gluon fusion	33.89	$+37.2/-27.8$
Vector boson fusion	2.01	$+7.6/-5.1$
WHH	0.57	$+3.7/-3.3$
ZHH	0.42	$+7.0/-5.5$
$t\bar{t}H$	1.02	-

Table 1.4: Production cross sections for pair production of a 125 GeV Higgs boson at $\sqrt{s} = 14$ TeV with total uncertainty [20]. The uncertainties include QCD scale and PDF variations as well as uncertainties on α_S.

1.5 Higgs Pair Production in Theories Beyond the Standard Model

The Standard Model Higgs pair production cross section is rather small, and datasets on the scale of the full lifetime of the LHC will be required to obtain sensitive measurements of the Higgs self-coupling. However, the discovery of the Higgs also gives particle physicists a new tool that can be exploited in the search for new physics beyond the Standard Model. In particular, Higgs pair production is a promising channel in the search for new physics. The cross section for di-Higgs production can be altered through both resonant and non-resonant production of Higgs pairs. In non-resonant production, di-Higgs production vertices can arise from the presence of a new strong sector and additional colored particles [21–23]. Figure 1.6 shows examples of the types of vertices that can arise. In the resonant case, new heavy particle can decay to Higgs pairs. Such new particles can include heavy Higgs bosons arising in two Higgs doublet models (2HDM) or Higgs portal models as well as heavy gravitons in Randall-Sundrum theories [21, 24–30]. Figure 1.7 shows a generic diagram for a heavy resonance decaying to two Higgs bosons. In the 2HDM, X corresponds to the heavy CP-even scalar H. In the Randall-Sundrum model, X corresponds to a heavy spin-2 graviton G. The next sections provide more detail on the phenomenology of resonant Higgs production in Randall-Sundrum and 2HDM models, as these models will later be tested in a dedicated search for resonant production of boosted Higgs pairs.

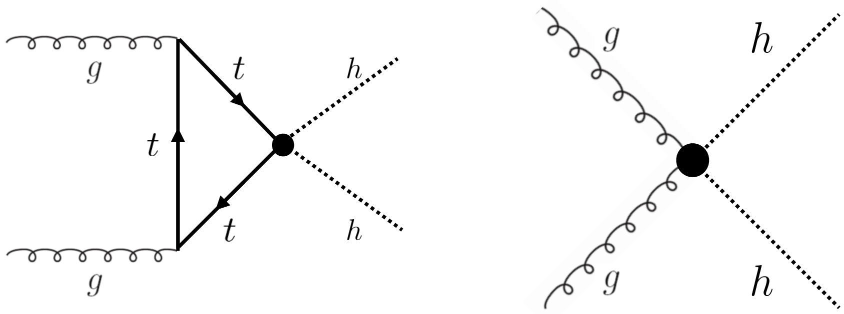


Figure 1.6: Diagrams with new vertices for non-resonant Higgs pair production arising in composite Higgs models

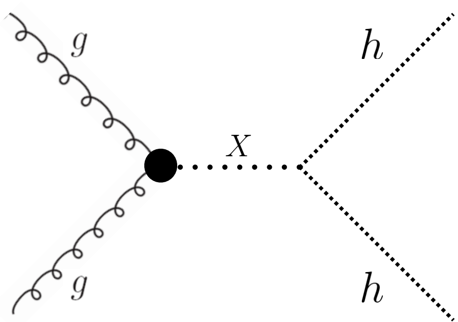


Figure 1.7: Generic Feynman diagram for resonant Higgs pair production in BSM theories

1.5.1 Randall-Sundrum Gravitons

The Randall-Sundrum model is a proposed solution to the hierarchy problem that posits a five-dimensional warped spacetime that contains two branes: one where the force of gravity is very strong and a second brane at the TeV scale corresponding to the known Standard Model sector [24]. In the theory, the branes are weakly coupled and the graviton probability function drops exponentially going from the gravity brane to the SM brane, rendering gravity weak on the SM brane. The experimental consequence of this theory is a tower of widely spaced (in mass) Kaluza-Klein graviton resonances. In theories where the fermions are localized to the SM brane, production of gravitons from fermion pairs is suppressed and the primary mode of production is gluon fusion [25]. These gravitons have a substantial branching fraction to Higgs pairs, ranging from 6.43% for gravitons with a mass of 500 GeV to 7.66% at 3 TeV. Figure 1.8 shows the branching ratios of the spin-2 Randall Sundrum graviton (RSG) as a function of its mass. The predominant decays are to $t\bar{t}$ above the mass threshold for that channel.

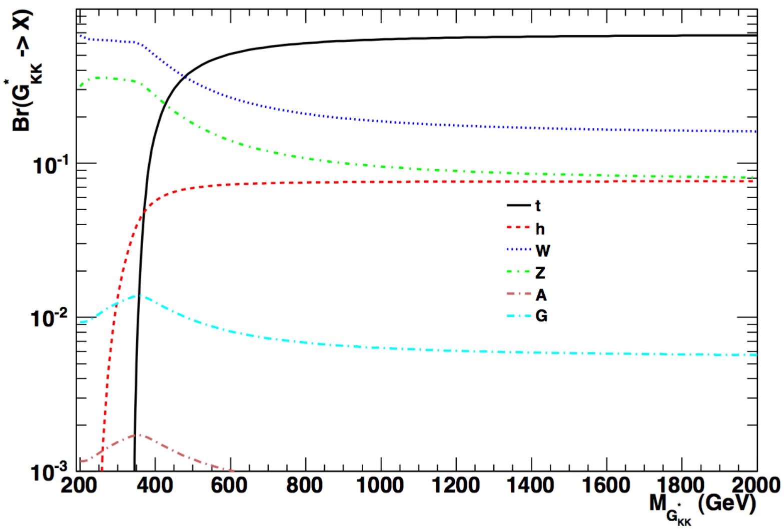


Figure 1.8: Branching ratios for a spin-2 Randall-Sundrum graviton as a function of mass computed in MadGraph with the CP3-Origins implementation [25, 31]

These models have two free parameters - the mass of the graviton and a curvature parameter k. Typically, rather than k, the theory is parameterized using $c \equiv k/\bar{M}_{\rm pl}$, where $\bar{M}_{\rm pl}$ is the reduced Planck mass. The cross section for production of the RSG decreases as a function of mass and is strongly dependent on the gluon PDF. The increase in center of mass energy from 8 to 13 TeV in LHC Run 2 greatly increases the cross section at higher mass. Figure 1.9 shows the cross section as a function of graviton mass at $\sqrt{s}$ = 13 TeV for RSG models with $c = 1.0$ and $c = 2.0$.

Another interesting feature of the theory is that the width of the graviton increases with both c and m_G. Figure 1.10 shows the graviton width for both $c = 1.0$ and $c = 2.0$ as a function of mass. In $c = 1.0$, the width starts at 8.365 GeV for a mass of 300 GeV and increases to 187.2 GeV at a mass of 3 TeV. SImilarly, with $c = 2.0$, the width starts at 33.46 GeV for $m_G = 300$ GeV and increases to 748.8 GeV at a mass of 3 TeV.

1.5.2 Two Higgs Doublet Models

In Two Higgs Doublet Models (2HDM), a second complex scalar doublet is added to SM [27–29]. In this case, all four degrees of freedom in the second doublet correspond to new particles, meaning that there

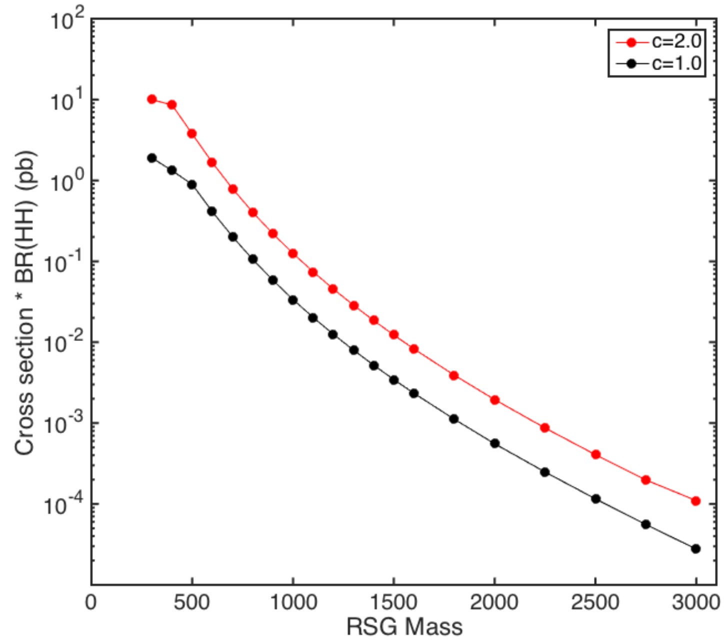


Figure 1.9: $\sigma \times \mathrm{BR}(HH)$ for RSG as a function of mass computed in MadGraph with the CP3-Origins implementation [25, 31]

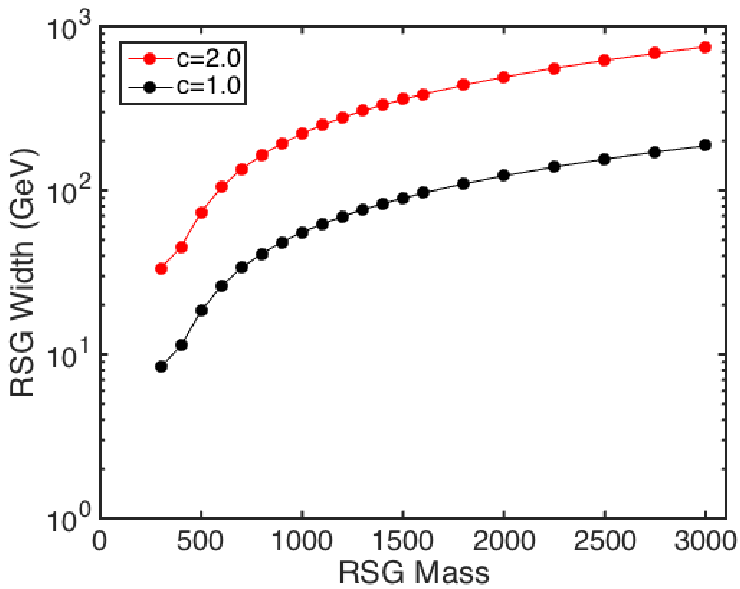


Figure 1.10: RSG width as a function of mass computed in MadGraph with the CP3-Origins implementation [25, 31]

are five total scalars from the two Higgs doublets - h (light CP-even Higgs), H (heavy CP-even Higgs), A (heavy CP-odd Higgs), and $H^{\pm}$ (charged Higgs). The model is parameterized by two main parameters. The first, $\tan\beta \equiv \frac{v_2}{v_1}$, the the ratio of the vacuum expectation values of the two Higgs doublets (where v_1 corresponds to the v in the SM Higgs model described above). The second parameter is α, a mixing angle between the heavy and light Higgs fields. Models are also often parameterized with $\cos(\beta - \alpha)$ rather than α directly. The limit where $\cos(\beta - \alpha) = 0$ is called the alignment limit, and it is in this limit that the light Higgs h has the same couplings as a Standard Model Higgs.

2HDM models are usually separated into two main types - Type I and Type II. In Type I models, the charged fermions only couple to the second Higgs doublet, leading to a fermiophobic light Higgs. In Type II models, up-type quarks couple to the first doublet while down-type quarks couple to the second doublet. One specific realization of a Type II 2HDM is the Minimal Supersymmetric Standard Model (MSSM).

Resonant di-Higgs production in this model can proceed through decays of the heavy CP-even Higgs $H \to hh$. The branching ratio for $H \to hh$ depends on the model type as well as the values of $\tan\beta$ and $\cos\beta - \alpha$. Figure 1.11 shows the branching ratios as a function of the mass of the heavy scalar H for both Type I and Type II models. Depending on the type of model hh can be a substantial fraction of the decays of H.

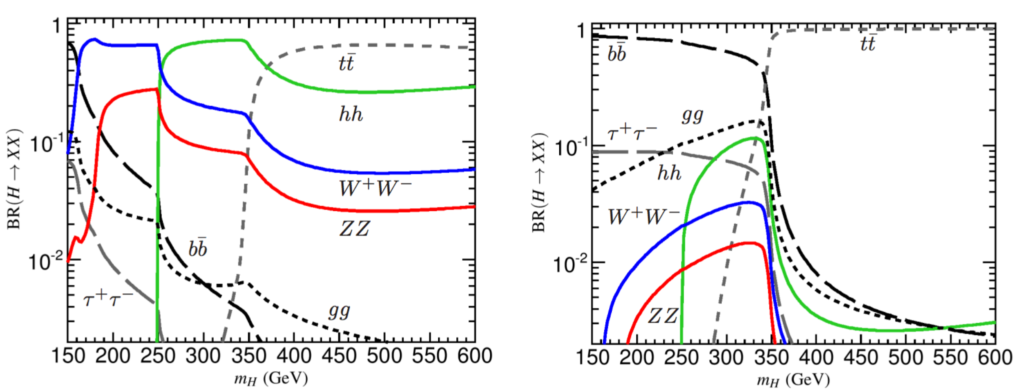


Figure 1.11: Branching ratios for heavy Higgs H in Type I (left) and Type II (right) 2HDM models with $\tan\beta = 1.5$ and $\cos(\beta - \alpha) = 0.1(0.01)$ for Type I (Type II). [29]

1.6 Conclusion

Studying the Higgs sector is essential for understanding the details of how mass arises in the Standard Model and how the electroweak symmetry is broken. The discovery of the Higgs boson also opens the door for its use as a tool to search for new physics, and Higgs pair production is an ideal candidate for this study. Even if no BSM physics is found in Higgs pair production, searches for Higgs pairs will put constraints on the Higgs self coupling and thus further knowledge of the Standard Model and the details of the Higgs potential.

The enthusiasm and motivation to explore particle physics at the high-energy frontier knows no borders between the nations and regions of the planet.

Peter Jenni

2

The ATLAS detector and the Large Hadron Collider

This chapter presents an overview of the experimental systems used to conduct the measurements presented in this thesis. First, a brief overview of the accelerator, the Large Hadron Collider, will be given. In this section, the accelerator conditions relevant to data-taking are presented as well. Next, an overview of the ATLAS experiment is given. The basics of each sub-detector's role are summarized, as well as the details of the datasets accumulated. Then, a brief interlude on the ATLAS Muon New Small Wheel upgrade is presented. While this new detector does not have a direct impact on any of the datasets taken so far, it will have an impact on future analyses and the work done on it is briefly summarized here. Finally, an overview of object reconstruction in ATLAS is given. While the details of all of the algorithms will not be presented in detail, aspects of the reconstruction performance such as object resolutions are shown as these are relevant to the two studies presented later in this thesis.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a proton-proton collider at the CERN laboratory in Geneva, Switzerland [32]. It is designed for a maximum collision center of mass energy of $\sqrt{s} = 14$ TeV and has a circumference of 26.7 kilometers. Four main experiments are located at the interaction points (IP) of the accelerator: ATLAS (A Toroidal LHC ApparatuS), CMS (the Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment), and LHCb [33–36]. The studies performed in this thesis were all completed with the ATLAS detector. Figure 2.1 shows a schematic of the LHC ring and the various experiments.

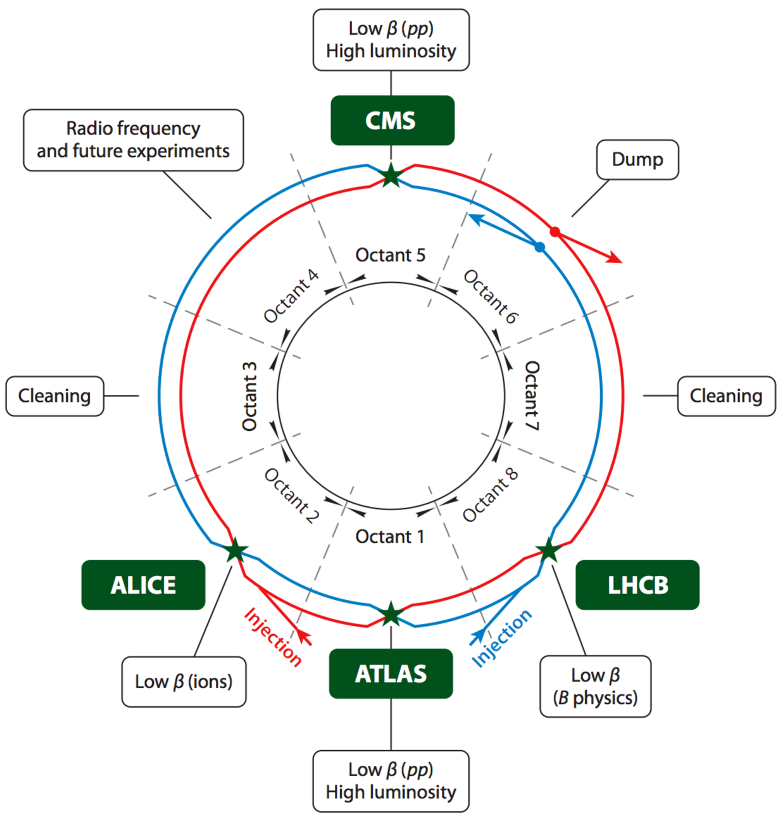


Figure 2.1: A schematic view of the LHC ring [37]

One of the most interesting features of the LHC is in its magnet design. Because the tunnel does not have room for separate superconducting magnets for each of the beam pipes, the LHC employs a twin-bore magnet design. Each magnet must hold an 8.3 Tesla magnetic field in order to bend the proton beams at $\sqrt{s} = 14$ TeV. The superconducting magnets are cooled to a temperature of 1.9 Kelvin with superfluid helium.

2.1.1 Instantaneous luminosity

The rate of physics events expected from the accelerator is dependent on the instantaneous luminosity of the machine and the cross section of the physics process, $R_{\text{events}} = L\sigma$. Here, R_{events} is the number of events per second, L is the instantaneous luminosity of the machine, and σ is the cross section for the physics process being measured. The instantaneous luminosity of the LHC is determined by numerous factors related to machine conditions. Equation 2.1 gives the equation for instantaneous luminosity of Gaussian beam profile [37].

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \tag{2.1}$$

The LHC collides protons in bunches, and in the above equation N_b is the number of protons per bunch while n_b is the number of bunches per beam. Nominally, the LHC can hold up to 2808 proton bunches. f_{rev} is the revolution frequency. ϵ_n is the normalized transverse beam emittance, a measurement of the average spread of the particles position-momentum space which has the dimension of length. β^* is the value of the $beta$ function for the beam at the interaction point. It relates the emittance to the Gaussian width of the beam with $\sigma_{\text{beam}} = \sqrt{\epsilon \cdot \beta}$. F is a reduction factor that corrects for the fact that the beams are colliding at an angle at the IP.

Another way of writing the instantaneous luminosity is shown in equation 2.2. In this case, the instantaneous luminosity is written as the ratio of the rate of inelastic collisions with the inelastic cross section [38].

$$L = \frac{R_{\text{inel}}}{\sigma_{\text{inel}}} = \frac{\mu n_b f_{\text{rev}}}{\sigma_{\text{inel}}} \tag{2.2}$$

In this case, μ is the average number of interactions per bunch crossing in the accelerator. μ is a useful parameter for characterizing the amount of activity recorded in an experiment. As the instantaneous luminosity and thus μ increase, there are more interactions per bunch crossing and more activity in the detector. This is often characterized with $\langle\mu\rangle$, the measured per bunch crossing μ value averaged over all bunch crossings. The interactions inside each bunch crossing that are not the main physics process of interest are often referred to as "pileup" interactions, and $\langle\mu\rangle$ is a measurement of the level of pileup in the detector.

2.1.2 Evolution of machine conditions

This thesis uses datasets taken at three different center of mass energies: $\sqrt{s} = 7$ TeV data taken in the year 2011, $\sqrt{s} = 8$ TeV data taken in the year 2012, and $\sqrt{s} = 13$ TeV data taken in the year 2015. In addition to increasing center of mass energy, the instantaneous luminosity and parameters that determine it were evolving. Table 2.1 summarizes that machine conditions in each of these datasets.

	2011	2012	2015	Design
$\sqrt{s}$ [TeV]	7	8	13	14
Number of bunches	1380	1380	1825	2808
Max. protons per bunch	1.45×10^{11}	1.7×10^{11}	1.2×10^{11}	1.15×10^{11}
Bunch spacing [ns]	50	50	25	25
Max. instantaneous luminosity [cm^{-2}s-1]	3.7×10^{33}	7.7×10^{33}	5×10^{33}	10^{34}
β^* [m]	1.0	0.6	0.8	0.55
$\langle\mu\rangle$	11.6	20.7	13.7	-

Table 2.1: Evolution of LHC machine conditions [39, 40]

2.2 The ATLAS Detector

The ATLAS detector is a multi-purpose particle detector experiment at the LHC's Point 1 [33]. It has nearly 4π coverage in solid angle around the interaction point. It consists of an inner detector for measuring charged particles, electromagnetic and hadronic calorimeters, and a muon spectrometer. Figure 2.2 gives an overview of the detector.

2.2.1 Coordinate system

Before defining the properties of the individual detectors, it is important to establish the coordinate system used. Figure 2.3 shows a schematic of the coordinate system. The azimuthal plane (perpendicular to the beam line) is defined as the x-y plane. The angle in this plane is referred to as ϕ. The angle relative

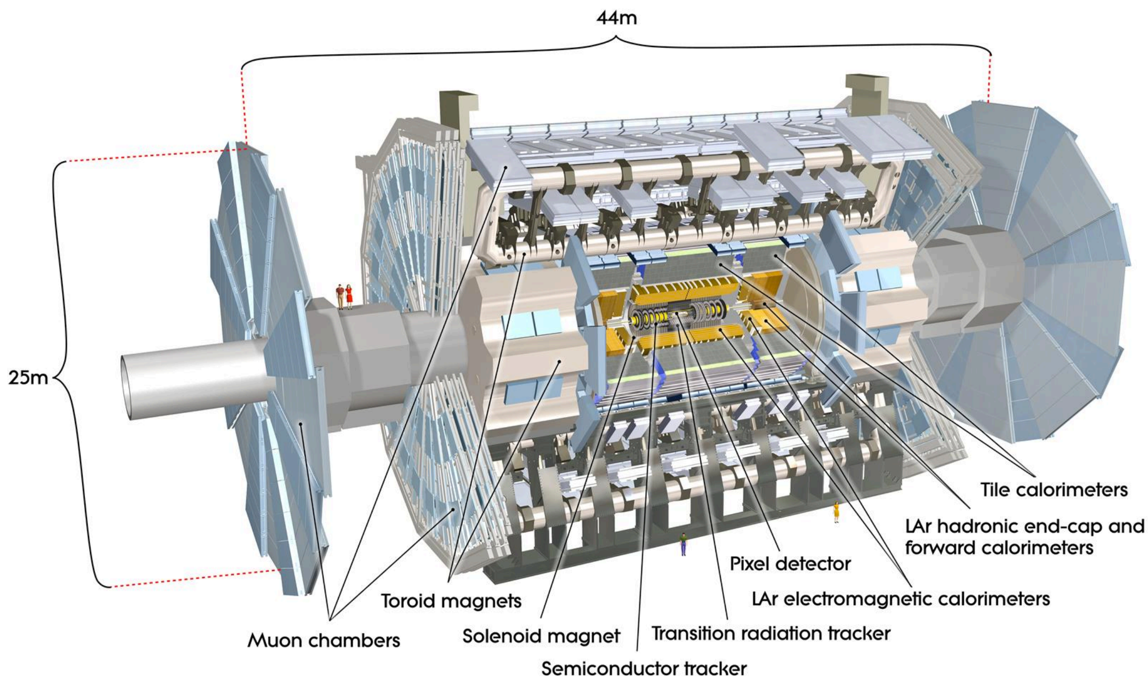


Figure 2.2: A full diagram of the ATLAS detector [33]

to the beam axis is referred to as θ. Rather than using θ directly as a coordinate, the experiment often uses the pseudorapidity η. η is defined in equation 2.3.

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{2.3}$$

Pseudorapidity is the massless approximation of rapidity, the angle used to parameterize boosts in special relativity. This is important for two reasons. First, it means that differences in η are Lorentz invariant. Second, particle production is roughly constant in pseudorapidity. Particles with η close to zero are referred to as "central", while those at high $|\eta|$ are called "forward". In general, two main detector topologies can be seen in figure 2.2. There are "barrel" elements, which surround the beam line cylindrically and are in the central region of the detector. In the forward region, there are "endcap" regions which are arranged as disks perpendicular to the beam line.

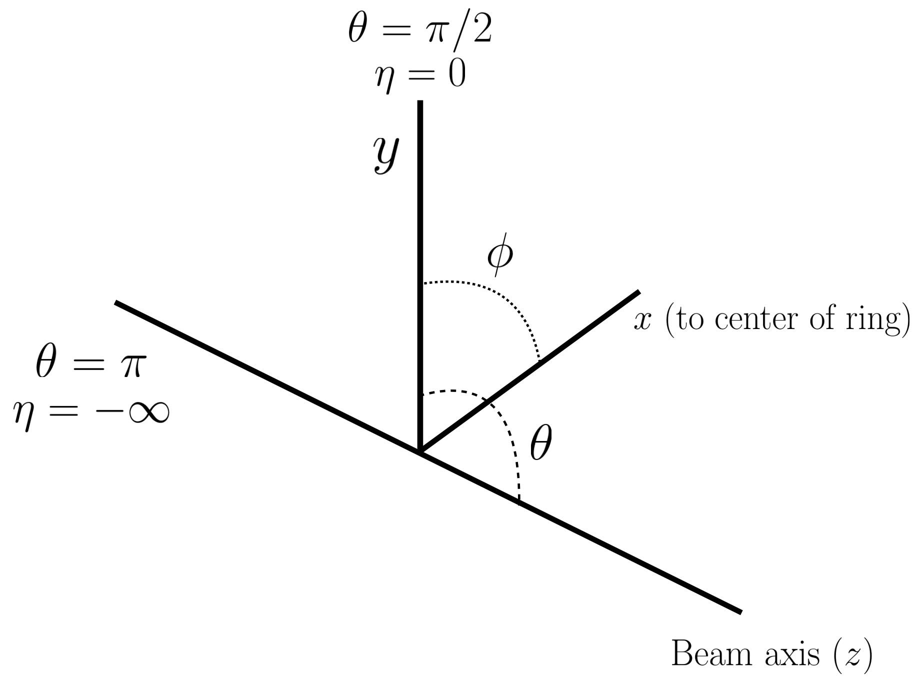


Figure 2.3: The ATLAS coordinate system

2.2.2 Inner detector

The ATLAS Inner Detector (ID) system is built for precision tracking of charged particles. It covers the range $|\eta| < 2.5$. In this range, approximately 1000 particles are generated every bunch crossing in the detector. This requires having fine granularity to achieve the resolutions required for good momentum measurement and vertex reconstruction.

The ID consists of three sub-components: the pixel detector, semiconductor tracker (SCT), and transition radiation tracker (TRT). It is surrounded by a solenoid providing a 2 T axial magnetic field which bends particles in the transverse plane to allow for momentum measurement. Figure 2.4shows the layout of each of these components.

Pixel detector

The pixel detector is the first detector particles traverse after being generated in proton collisions and is the most granular detector. Its operation is crucial for precision tracking and vertex reconstruction as

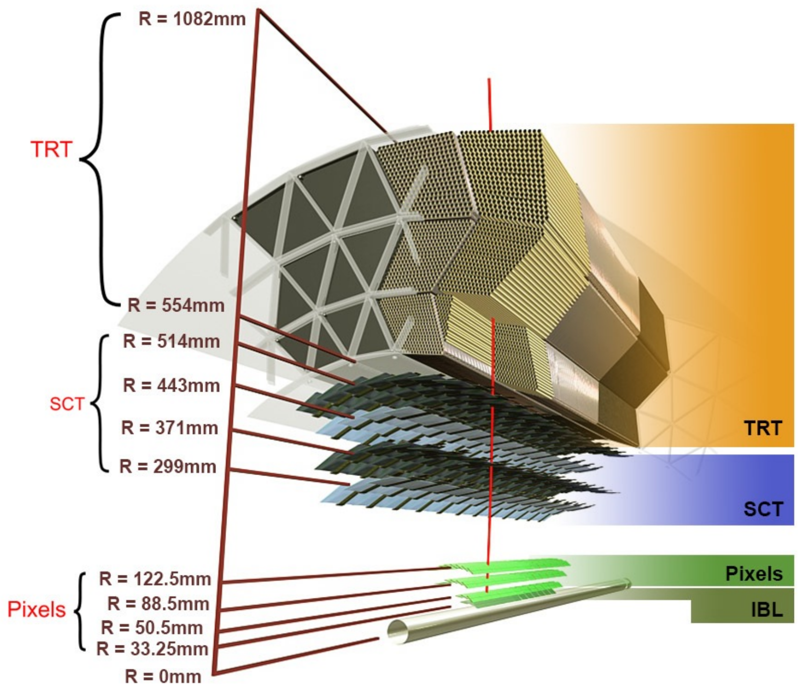


Figure 2.4: Layout of the ATLAS Inner Detector system [41]

well as higher level object reconstruction like tagging of jets from b-quarks. The basic sensing element in this subdetector is a silicon pixel detector. The operating principle for the silicon pixels is that of a p-n junction. When a charged particle passes through, it creates electron-hole pairs that are then separated by the electric field. The sensors are $250\ \mu$m thick and use oxygenated n-type wafers with readout pixels on the n^+ side of the detector [33]. Overall, the pixel detector has 1744 sensors and 80.4 million readout channels.

In the barrel region, the pixel detector has three concentric layers of sensors surrounding the beamline. In the endcap region, it consists of disks perpendicular to the beam axis. The detector is segmented in the R-ϕ plane and in z. Usually, three pixel layers are crossed by a charged particle track. The intrinsic accuracies of the sensors are $10\ \mu$m in R-ϕ and $115\ \mu$m in z (or R for the endcap).

Insertable B-layer

In Run 2, a new innermost pixel layer, known as the insertable B-layer (IBL), was added to the Inner Detector [42]. This layer was added to cope with the higher luminosities planned in LHC Run 2 and at the high luminosity HL-LHC. Additionally it improves tracking position resolution which in turn improves the vertexing and b-tagging capabilities in ATLAS. The detector sits directly on a new beam pipe, only 33.25 mm away from the collision points in the azimuthal plane.

Semiconductor Tracker (SCT)

The semiconductor tracker (SCT) consists of silicon microstrips and comprises the next four layers of the ID. This sub-detector has 6.4cm long sensors that are daisy-chained into strips with a strip pitch of $80\,\mu$m [33]. Some of the strips have a small stereo angle to allow for measurement of both angular coordinates. In total there are 6.3 million readout channels. The intrinsic accuracies are $17\,\mu$m in R-ϕ and $580\,\mu$m in z (or R in the endcap).

Transition radiation tracker (TRT)

The transition radiation tracker (TRT) serves two purposes. First, it consists of 4mm diameter straw tubes filled with a $70/27/3$% gas mixture of xenon, carbon dioxide, and oxygen to provide tracking of charged particles. Particles typically have 36 TRT straw tube hits per track. The material in between the straws is designed to induce transition radiation which can be useful for particle identification. As particles pass between media with different dielectric constants, they emit transition radiation that can cause additional showers in the TRT. In particular it is useful for discrimination between electrons and pions or other charged hadrons, as the amount of transition radiation is proportional to the Lorentz factor of the particle.

2.2.3 Calorimeters

The calorimeter system consists of two main sub-components: a fine granularity electromagnetic calorimeter tailored for the measurement of photons and electrons and multiple coarser hadronic calorimeters dedicated to the measurement of hadronic showers [33]. The calorimeter system has broader coverage than the inner detector, covering the region out to $|\eta| < 4.9$. It is also designed to deliver good containment of showers so as to limit leakage into the muon system. Figure 2.5 shows the layout of the calorimeter system.

Both the electromagnetic and hadronic calorimeters are sampling calorimeters. They alternate active material for energy measurement with passive material for energy absorption. The materials used for each purpose vary based on the type of calorimeter and its location in the detector.

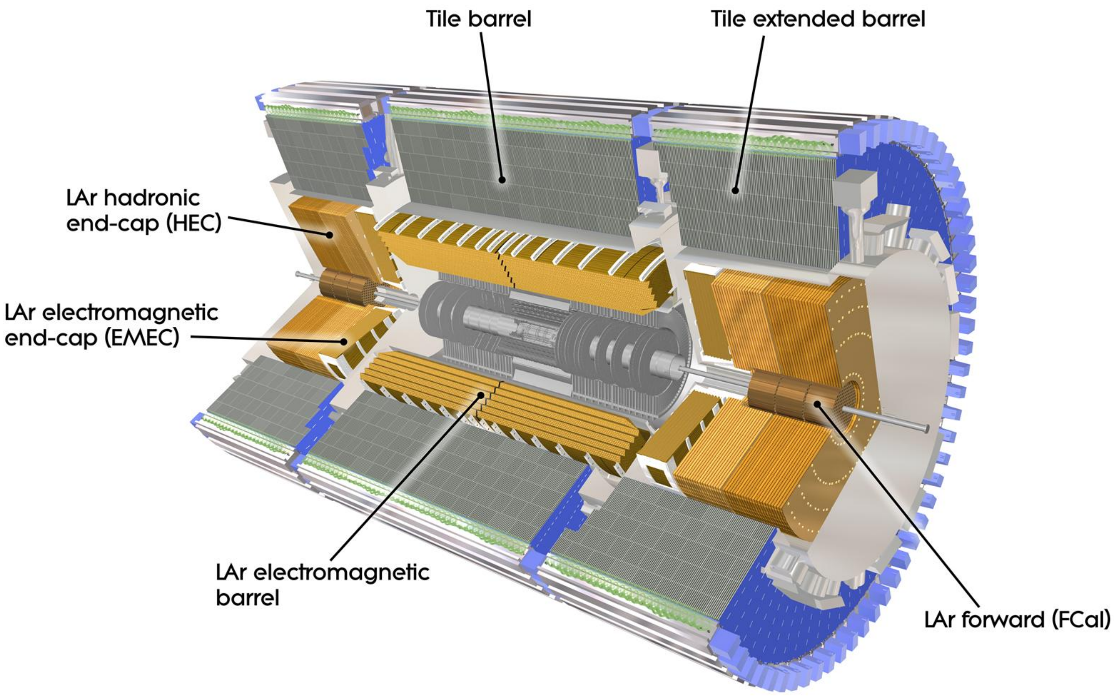


Figure 2.5: Layout of the ATLAS calorimeter system [33]

Electromagnetic calorimeter

The electromagnetic calorimeter (EM calorimeter) use liquid Argon (LAr) as its active material and lead as its passive material. It is arrange in an accordion geometry to increase the absorption area while still allowing it to have no azimuthal cracks (complete symmetry in ϕ). The EM calorimeter is divided into a barrel portion that extends to $|\eta| < 1.475$ and an endcap portion going from $1.375 < |\eta| < 3.2$. The region where these two units overlap is called the "transition region".

In order to provide good containment the calorimeter depth must be optimized. Typically, for electromagnetic calorimeters the depth is meaasured in radiation lengths. In general, the intensity of a particle beam attenuates exponentially in distance with a constant equal to the radiation length. That is,

$I(x) = I_0 e^{-x/X_0}$, where I is the intensity, x is the distance traveled, and X_0 is the radiation length. The ATLAS EM calorimeter is designed to have > 22 radiation lengths in the barrel and > 24 in the endcap [33].

Hadronic calorimeters

There are three types of hadronic calorimeters present in ATLAS: the tile calorimeter (TileCal), hadronic endcap (HEC), and forward calorimeter (FCal). Each one is optimized for stopping of hadronic showers and the materials chosen are specific to their placement in the detector.

The TileCal is a scintillating tile calorimeter placed directly outside the EM calorimeter. It uses steel as the absorber and plastic scintillator tiles as the active material. It has coverage in the barrel at $|\eta| < 1.0$ and in the "extended barrel" region of $0.8 < |\eta| < 1.7$.

The HEC had two wheels perpendicular to the beam line per endcap and is located directly behind the EM calorimeter endcap modules. The HEC covers the region from $1.5 < |\eta| < 3.2$, overlapping slightly with both the tile calorimeter and the forward calorimeter. Like the EM calorimeter, it uses liquid Argon as the active material, but it uses copper as the absorber.

The FCal covers the most forward regions of the calorimeter system, extending to the region of $3.1 < |\eta| < 4.9$. It again uses liquid argon as its active material. For absorber, it consists of an innermost module made of copper followed by a module made of tungsten.

The hadronic equivalent of radiation length is called the interaction length and is denoted as λ. In the barrel, the hadronic calorimeter depth is approximately 9.7λ, while in the endcap is is 10λ. The outer supports contribute an additional 1.3λ. This is been shown to be sufficient to limit punch-through of showers to the muon system [33].

2.2.4 Muon spectrometer

The muon spectrometer is dedicated to measuring the momentum and position of muons. It consists of tracking and trigger chambers which are unique in the barrel and endcap regions. The magnetic field for bending of muons is provided by a system of three large air-core toroid magnets (from which ATLAS derives its name.) These magnets provide 1.5 to 5.5 Tm of bending power at $0 < |\eta| < 1.4$ and approximately 1 to 7.5 Tm in the endcap region of $1.6 < |\eta| < 2.7$. The entire muon system covers the range $0 < |\eta| < 2.7$. Monitored drift tubes (MDTs) are used for tracking in the barrel and the two outer layers of the endcap, while cathode strip chambers (CSCs) are used to provide tracking in the innermost endcap wheel. In the barrel, resistive plate chambers (RPCs) are used as trigger chambers while thin gap chambers (TGCs) are used in the endcap. Figure 2.6 shows the layout of the ATLAS muon system. The entire muon system is designed with the specification of providing a 10% momentum resolution for a 1 TeV muon.

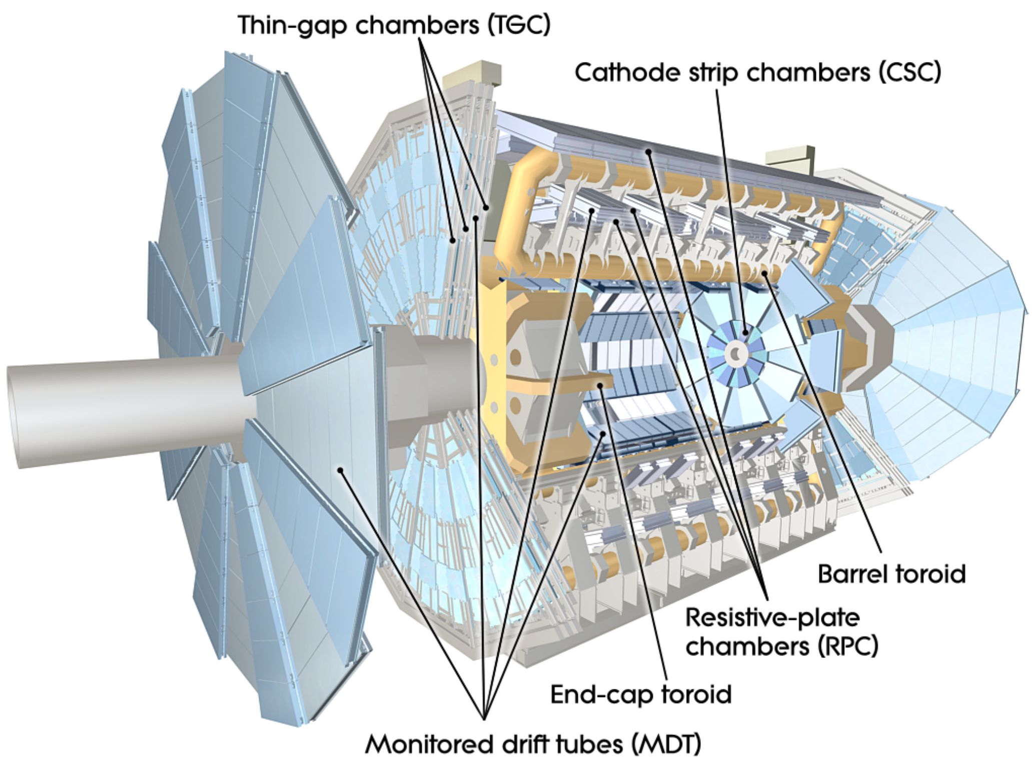


Figure 2.6: Layout of the ATLAS muon system [33]

Monitored drift tubes (MDTs)

The monitored drift tubes (MDTs) are aluminum 3cm diameter tubes filled with a 93/7 % mixture of Argon and CO_2, with trace amounts of water. As a charged particle traverses the tube, it ionizes the gas and the ions drift to a wire at the center of the tube. The radial distance of traversal of the particle in the tube is determined by the drift time of the electrons, allowing for fine position resolution. The tubes have an average resolution of 80 μm per tube and a maximum drift time of approximately 700ns. The tubes are oriented so that they give precision measurement in η and run along ϕ. They cover $|\eta| < 2.7$, except in the innermost layer of the endcap where they only go to $|\eta| < 2.0$ [33].

Cathode strip chambers (CSCs)

The cathode strip chambers cover a narrow window of the innermost endcap region at $2.0 < |\eta| < 2.7$. In this region the background rates in the cavern are particularly high and the CSCs are designed to handle these higher rates. The CSCs are multiwire proportional chambers with wires pointing in the radial direction (away from the beam pipe). The wire serves as an anode and there are two types of segmented cathode strip, one perpendicular to the wires which gives the precision measurement and one parallel which provides the transverse coordinate. It has an 80/20 gas mixture of Argon and CO_2 [33].

Resistive plate chambers (RPCs)

The resistive plate chambers (RPCs) are gaseous electrode-plate detectors covering the region $|\eta| < 1.05$. They consist of two resistive plates separated by a distance of 2 mm. The gas mixture used is a 94.7/5/0.3% mixture of $C_2H_2F_4$, Iso-C_4H_10, and SF_6. It has readout strips with a pitch of 23-35 mm for both η and ϕ measurement and thus provides measurement of the azimuthal coordinate in the barrel that the MDTs do not. The thin gas gap allows for a quick response time which makes it ideal for use in the trigger. There are three layers of RPCs which are referred to as the three trigger stations. They allow for both a low p_T and high p_T trigger. The coincidence of hits in the innermost chambers allows for triggering of muons between 6 and 9 GeV, while the outermost layer allows the trigger to select high momentum tracks in the range of 9 to 35 GeV [33].

Thin gap chambers (TGCs)

The thin gap chambers (TGCs) are multiwire proportional chambers where the wire to cathode distance (1.4mm) is smaller than the wire-to-wire distance (1.8 mm). They contain a gas mixture of CO_2 and n-pentane and use a high electric field to gain good time resolution. They serve two functions in the end-cap system. First, they serve as the trigger chambers. Second, they also provide azimuthal coordinate measurement which the MDTs do not. They sit on the inner and middle layers of the endcap. The outermost layer's azimuthal coordinate is determined by extrapolation [33].

2.2.5 Magnet system

As mentioned previously, there are two independent magnet systems in ATLAS. The first is a 2 T solenoid field in the inner detector which provides bending in the azimuthal plane. The second is an approximately 0.5 T toroidal field in the muon system which provides bending in η. Figure 2.7 shows the predicted field integral as a function of $|\eta|$ [33].

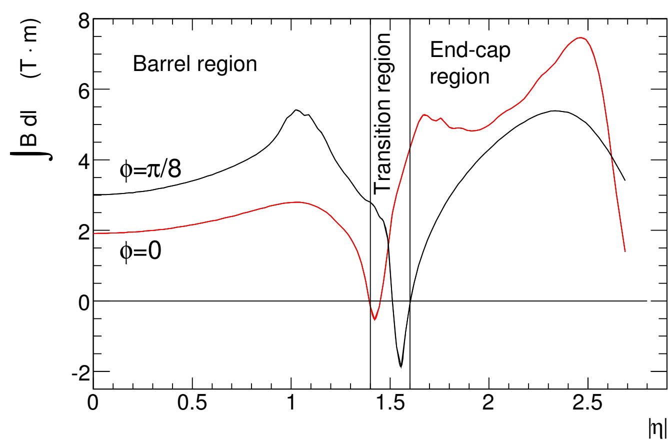
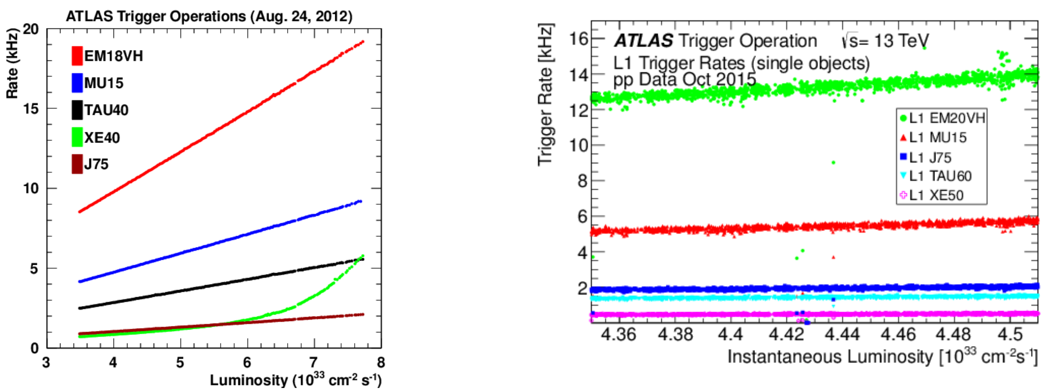


Figure 2.7: Predicted field integral as a function of $|\eta|$ for the ATLAS magnet system [33]

2.2.6 Trigger system

The ATLAS trigger system searches for signatures of muons, electrons, photons, hadronically decaying τ leptons, and jets in order to save these events for further analysis. The trigger system in ATLAS is designed to reduce the maximum LHC event rate of 40 MHz to a more reasonable rate that can be recorded. The trigger first consists of a fast, hardware based system called the Level-1 (L1) trigger. The L1 trigger consists of independent dedicated detector sub-components that can seed regions of interest (RoIs) for further analysis downstream. For muons, the RPCs and TGCs are used, while in the calorimeter coarsely grained sections of calorimeter cells called towers are used. Once regions of interest are seeded, a software based system called the High Level Trigger (HLT) is used to reconstruct objects and integrate information from different parts of the detector. In Run 1 of ATLAS, the HLT consisted of two separate stages: the level 2 (L2) trigger and the event filter (EF).

The maximum trigger rate that the L1 trigger can handle is 75 kHz. In the HLT, the rate of events written to disk is approximately 200 Hz. Figure 2.8 shows the trigger rates for different L1 triggers in 2012 and 2015 for ATLAS [43].



(a) 2012 (b) 2015

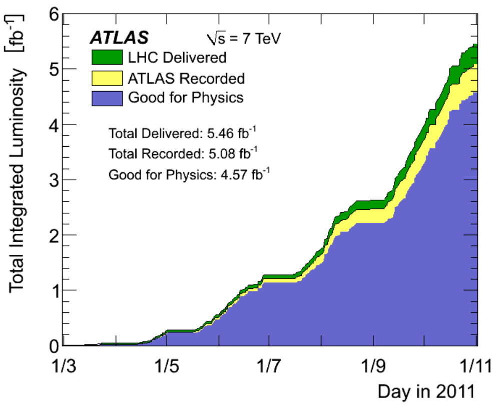
Figure 2.8: ATLAS trigger rates for Level-1 triggers as a function of instantaneous luminosity in 2012 and 2015 operation. These are single object triggers for electromagnetic clusters (EM), muons (MU), jets (J), missing energy (XE), and τ leptons (TAU). The threshold of the trigger is given in the name in GeV. [43]

2.2.7 ATLAS datasets

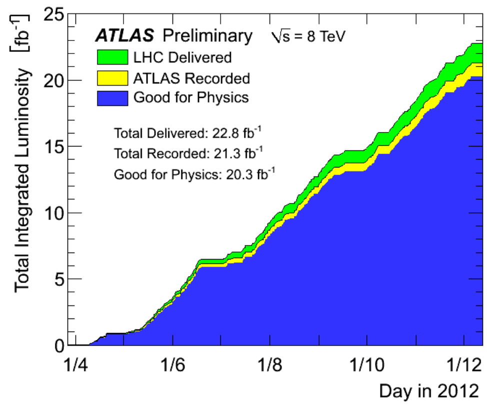
ATLAS has collected data at center of mass energies of 7, 8, and 13 TeV. Figure 2.9 shows the integrated luminosity as a function of time for each of the three collected datasets. At $\sqrt{s} = 7$ TeV, ATLAS recorded 5.08 fb^{-1}. Increased instantaneous luminosity in 2012 led to a larger dataset of 21.3 fb^{-1} recorded at $\sqrt{s} = 8$ TeV. After Long Shutdown 1 (LS1) of the LHC and a restart in 2015, ATLAS recorded 3.9 fb^{-1} of data at $\sqrt{s} = 13$ TeV. [44, 45]

2.2.8 Detector performance

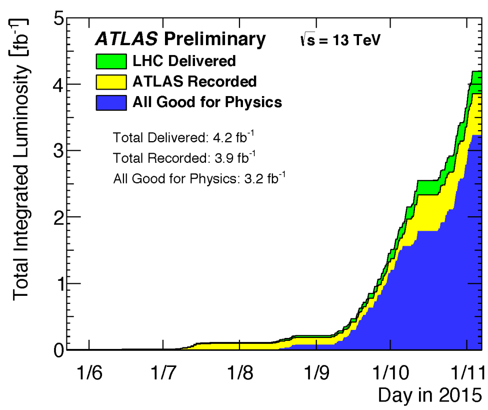
Table 2.2 summarizes the design requirements for each of the different sub-detectors. This table shows the energy and momentum resolution of each tracking, calorimetry, and muon measurements.



(a) $\sqrt{s} = 7$ TeV (2011)



(b) $\sqrt{s} = 8$ TeV (2012)



(c) $\sqrt{s} = 13$ TeV (2015)

Figure 2.9: Instantaneous luminosity as a function of time for data recorded by ATLAS at different center of mass energies [44, 45]

	Required resolution
Tracking	$\sigma_{p_T}/p_{\mathrm{T}} = 0.05\% p_{\mathrm{T}} \oplus 1\%$
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$
Hadronic calorimetry	
Barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$
Forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$
Muon spectrometer	$\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}}$ at $p_{\mathrm{T}} = 1$ TeV

Table 2.2: Performance requirements for the ATLAS detector [33].

2.3 The ATLAS Muon New Small Wheel Upgrade

As the LHC continues operation, it is scheduled to be upgraded in several phases to allow it to reach higher instantaneous luminosities and thus collect larger datasets. These conditions will open new doors for study of rare physics processes but will also present interesting challenges that must be faced. ATLAS will require new detector technologies to cope with the increased background rates in the cavern in these high luminosity conditions. One such upgrade, scheduled to be installed during Long Shutdown 2 (LS2) of the LHC in 2018, is the ATLAS Muon New Small Wheel (NSW) upgrade [46]. The NSW will replace the innermost end-cap wheel of the muon system with new technologies, as this is the part of the muon detector closest to the beam and thus suffers from the highest rates.

2.3.1 Motivation

The motivation of the NSW is two-fold. First, the objective is to alleviate the decreased tracking efficiency that comes in a high rate environment. As figure 2.10, at the LHC design luminosity both the efficiency of recording hits and reconstructing track segments in the MDTs decreases at the LHC design luminosity.

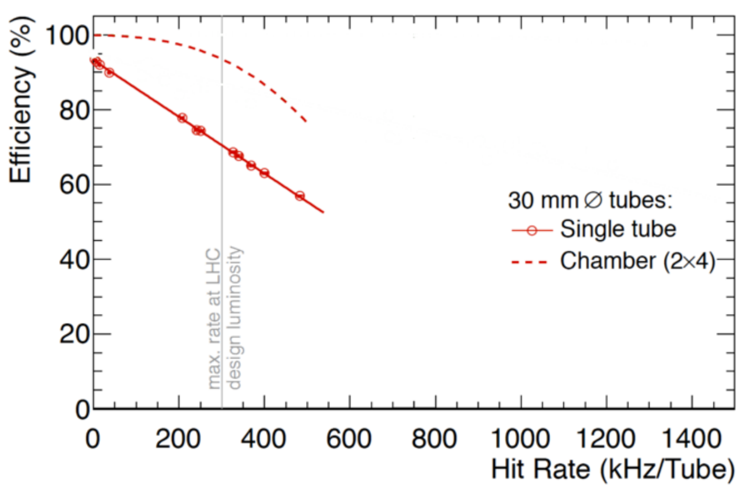


Figure 2.10: MDT tube hit (solid) and segment (dashed) efficiency as a function of hit rate per tube [46]

Second, the NSW will work to alleviate the rate of fake triggers arising in the endcap. Figure 2.11 shows the extrapolated trigger rates as a function of the p_T threshold with and without the NSW upgrade. As the figure shows, the NSW upgrade will reduce the trigger rate by an order of magnitude compared to the current endcap trigger system.

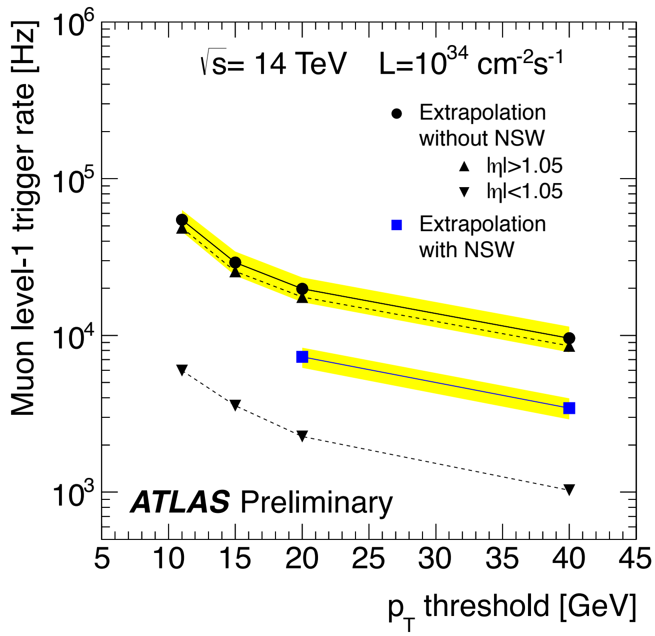


Figure 2.11: Trigger rate as a function of p_T threshold with and without the NSW upgrade [46]

2.3.2 NSW detector technologies

The NSW will use two new detector technologies - micromesh gaseous structure detectors (micromegas) and small-strip thin gap chambers (sTGCs) [46, 47]. Unlike the previous detectors, both of these detector technologies can be used for tracking or trigger. However, the micromegas is more suited to tracking because of its good spatial resolution, while the sTGCs have better time resolution and are more suited for the trigger. To maintain a fully redundant system, both technologies are used for both purposes.

Micromegas

Micromegas detectors operate using a thin metallic mesh that sits approximately 100 μm away from the readout electrodes to create the amplification region. Above this mesh, there is a drift region on the order of a few mm in length capped by a drift electrode. As a charged particle traverses the detector, it ionizes gas and the electrons drift down towards readout strips. The timing of the drift can be used to reconstruct the angle of traversal of the particle. This is illustrated in figure 2.12. The micromegas used in ATLAS will be resistive micromegas, where the readout electrodes are topped with resistive strips [48]. This alleviates the risk of sparking in the large area detectors that ATLAS will use.

In ATLAS, the micromegas drift gap will be 5 mm and the amplification gap will be 128 μm. They are filled with the same gas mixture as the MDTs. They will be stacked in an octuplet in an XXUV-UVXX geometry, where X refers to straight strips and U and V refer to stereo strips at an angle of $\pm 1.5°$. This arrangement allows for measurement of the azimuthal coordinate and gives a large lever arm between the straight strips for triggering purposes. Figure 2.12 shows the geometry of a single micromegas detector as well as its operating principle [46].

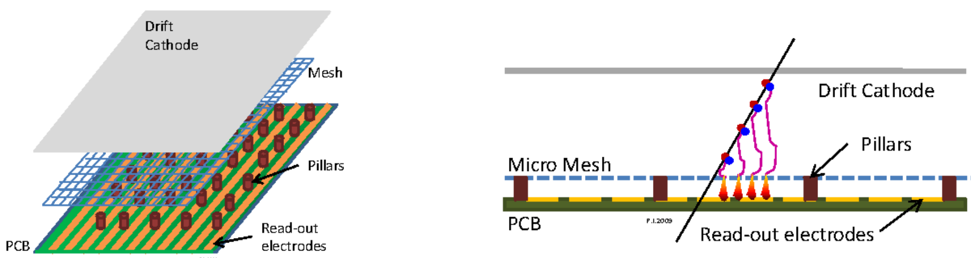


Figure 2.12: Illustrations of the geometry (left) and operating principle (right) of the micromegas detector [46]

sTGCs

The sTGCs are similar to the TGCs already described. They consist of gold-plated tungsten wires with a 1.8 mm pitch between two cathode planes 1.4 mm away from the wire plane. One cathode plane consists of strips with a 3.2 mm pitch (much smaller pitch than the TGCs), while the other consists of coarser pads that are used for defining regions of interest in the sTGC trigger algorithm. Figure!2.13 shows the basic detector geometry.

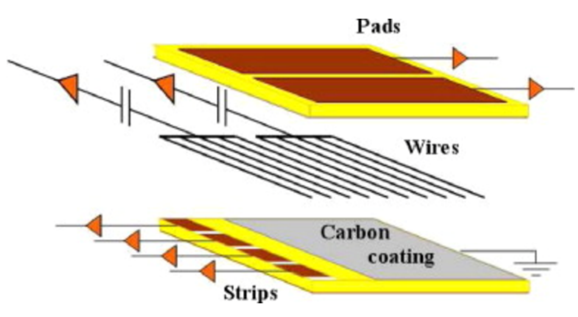


Figure 2.13: Geometry of the sTGC detector [46]

2.3.3 Physics Impact

Maintaining low p_T thresholds for muons while still staying within the trigger rate budget at Level 1 (20 kHz) for the muon system is crucial for physics analyses to be successful in high luminosity conditions. One realm where the lepton trigger threshold is especially important is in Higgs physics. In the $H \to WW^*$ analysis, one of the W bosons is off shell and tends to decay to soft leptons. In associated production of a Higgs with a W, the lepton is also important because the lepton provides the main handle which allows the event to be triggered. Table 2.3 shows the impact of increasing the trigger thresholds on these analyses. It shows that either raising the threshold or using only the barrel both have significant impacts on the signal efficiency. With the NSW, the signal efficiency is largely maintained and the triggers can be unprescaled.

Threshold	$H \to b\bar{b}$ (%)	$H \to WW^*$ (%)
$p_T > 20$ GeV	93	94
$p_T > 40$ GeV	61	75
$p_T > 20$ GeV (barrel only)	43	72
$p_T > 20$ GeV (with NSW)	90	92

Table 2.3: Signal efficiencies for WH production with $H \to b\bar{b}$ and $H \to WW^* \to \mu\nu qq$ under different trigger configurations [46].

2.4 Object Reconstruction in ATLAS

ATLAS analyses first start by requiring the presence of certain reconstructed physics objects in the event. This section will present a brief overview of the algorithms used to reconstruct electrons, muons, jets (including b-jets), and missing energy[1]. The performance of object reconstruction and measurement will also be discussed as these are relevant to the analyses presented later. Figure 2.14 gives an overview of the different sub-detectors that each type of particle will interact with in ATLAS.

[1] Reconstruction algorithms for other objects, such as photons and τ leptons, are not detailed here as these objects are not used in the presented studies.

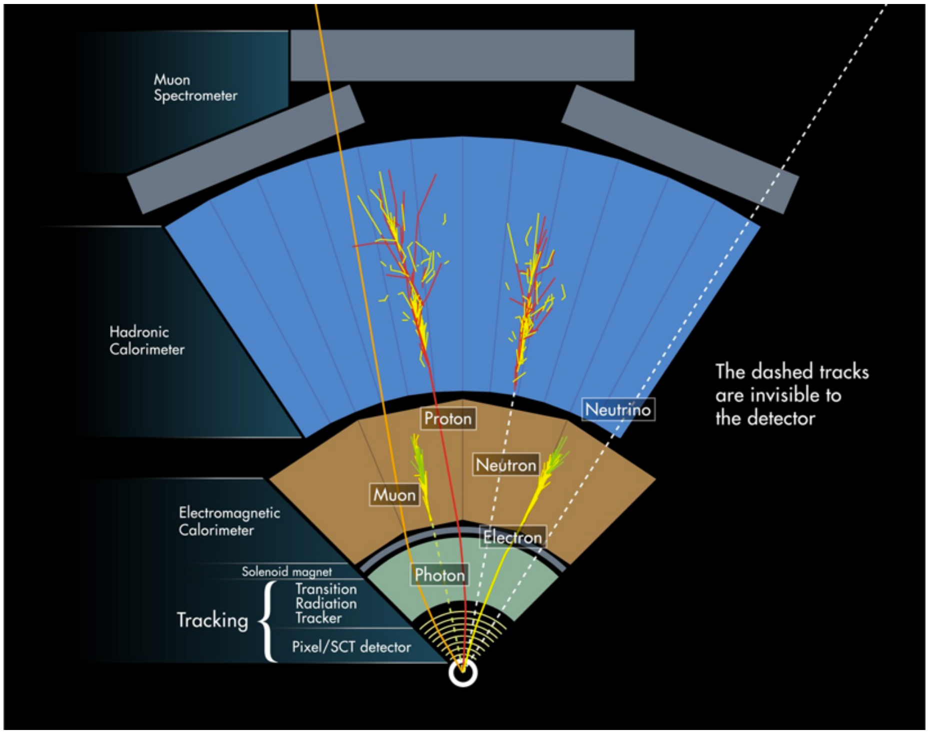


Figure 2.14: Illustration of particle interactions in ATLAS [49]

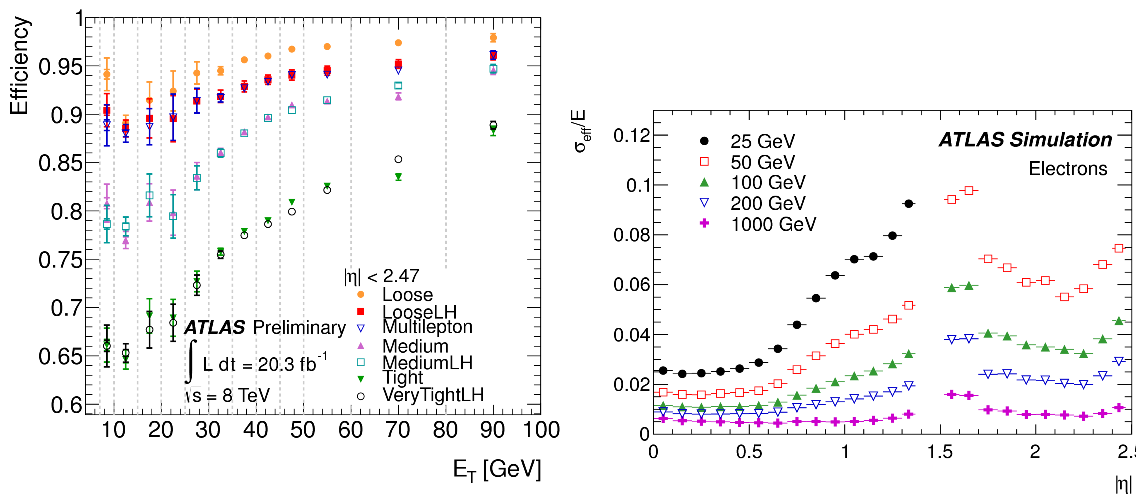
2.4.1 Electrons

Electrons in ATLAS will leave tracks in the inner detector and energy deposits in the electromagnetic calorimeter. The algorithm for recognizing the signature of electrons proceeds in two steps: reconstruction and identification.

In reconstruction, an electron candidate is formed by matching EM calorimeter deposits with ID tracks. The algorithm first chooses seed clusters in the EM calorimeter by using a sliding window algorithm that searches for towers with transverse energy larger than 2.5 GeV. In addition to seed clusters, track candidates must be identified in the ID. The algorithm selects seed tracks with $p_{\mathrm{T}} > 1$ GeV that do not fit well with a pion hypothesis. Once candidate tracks are selected, they are re-fit with a Gaussian Sum Filter (GSF) algorithm to estimate electron parameters [50]. Finally, an electron candidate is formed if at least one track matches to a seed cluster in the calorimeter. The full details of the reconstruction algorithm can be found in reference [51].

Once an electron candidate is present, identification criteria must be applied in order to reject fake electrons from background. Many different variables are used for this identification, most of them related to the shower shape in the EM calorimeter and the amount of leakage into the hadronic calorimeter, as well as information from the ID and in particular the TRT. There are both cut-based and likelihood-based criteria that range from "loose" to "very tight". For details, see reference [51].

Figure 2.15 shows the algorithm's reconstruction efficiency of true electrons for different identification criteria as well as the electron energy resolution in simulation [51, 52]. The reconstruction efficiency is measured using both Z and J/ψ tag and probe techniques.



(a) (b)

Figure 2.15: Electron performance: (a) reconstruction efficiency as a function of electron E_T [51] (b) energy resolution in simulation as a function of $|\eta|$ for different energy electrons [52]

2.4.2 Muons

The ATLAS detector is designed to stop most particles before they reach the muon spectrometer. Muons, however, are minimum ionizing particles, meaning that they will not lose a significant amount of energy through interactions with the detector and will thus pass through. Therefore, the muon reconstruction works to match tracks in the muon spectrometer with tracks in the inner detector.

The first step of reconstruction is to reconstruct local straight line tracks, called segments, in each muon chamber. Segments are then fit to larger tracks that traverse the entire muon spectrometer. Such muon tracks are referred to as "standalone" tracks (SA) as they only use information from the muon spectrometer.

The standalone tracks are then matched to tracks in the inner detector to form “combined” (CB) muons, where the combined ID and MS fit are used to determine the momentum and direction of the muon. To improve acceptance, segment-tagged and calorimeter-tagged muons are also reconstructed. In these cases, ID tracks are matched to segments in the MS and calorimeter deposits consistent with a minimum ionizing particle, respectively. The details of the reconstruction can be found in reference [53].

As with electrons, once muon candidates are reconstructed they have identification criteria applied to reduce background. These criteria include the χ^2 match between the ID and MS tracks, the number of hits in the ID, overall ID and MS track fit quality, and additional variables [53]. The criteria range from “loose” to “tight” as with electrons.

Figure 2.16 shows the muon reconstruction efficiency (measured with Z and J/ψ tag and probe) and invariant mass resolution [53].

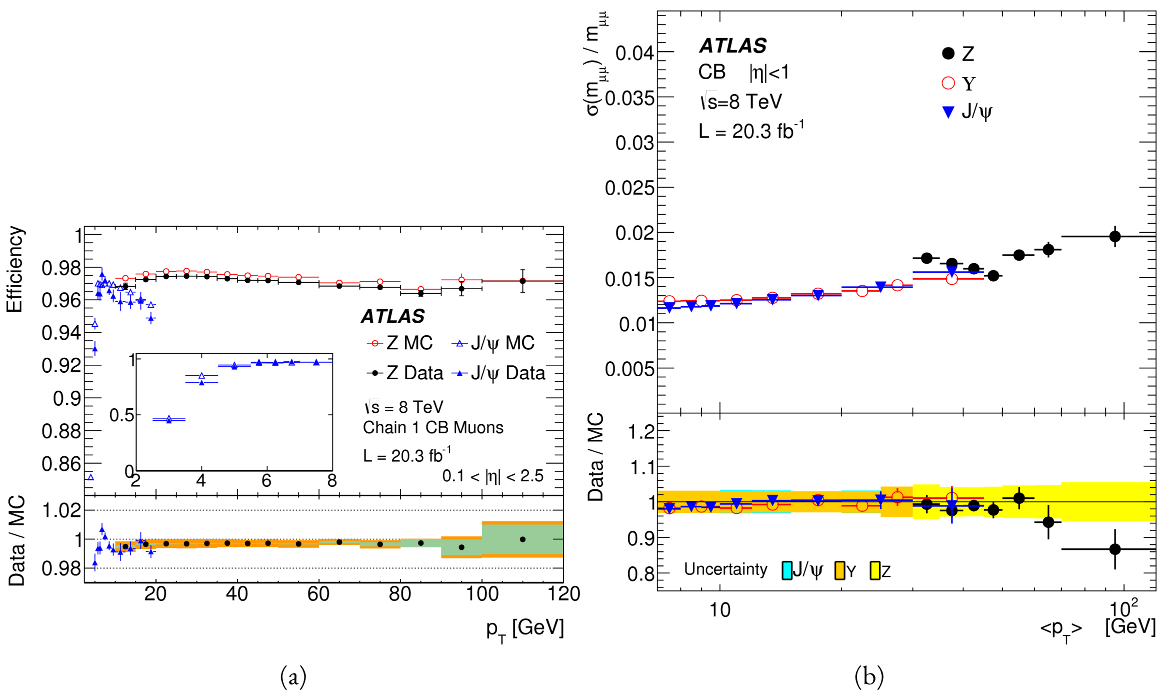


Figure 2.16: Muon performance: (a) reconstruction efficiency as a function of muon p_T (b) dimuon mass resolution as a function of average p_{T} [53]

2.4.3 Jets

When a quark or gluon is produced in collisions, it is not measured directly in ATLAS. Rather, due to QCD effects, it produces a collimated spray of hadrons in the direction of the original parton, which is known as a jet. Jets are reconstructed in ATLAS using energy deposits in the hadronic calorimeter. The first step is build "topological clusters" out of energy deposits in calorimeter cells [54, 55]. This is done using strategy where seed cells are chosen by picking cells whose energy measurements are four times the amount of noise expected for that cell. Adjacent cells with at least 2σ energy measurements are added to the cluster, then a final layer of clusters with energy above 0σ are added. Once calorimeter clusters are formed, they are clustered further into jet candidates using the anti-k_T jet clustering algorithm [56]. This algorithm uses a parameter R that appears in the denominator of the clustering distance metric and defines the radial size of the jet in η-ϕ space.

The energy response of the calorimeter must be properly characterized in order to reconstruct jet energy. Calorimeter clusters can be calibrated either with the EM calibration, where each cluster is assumed to have come from the energy deposit of an electron or photon, or the LCW calibration, where local cluster weights are computed to allow for local calibration of clusters as hadronic or electromagnetic. The details of the jet energy calibration are not detailed here and are discussed in reference [57].

Figure 2.17 shows the jet energy response after calibration in Monte Carlo as a function of the true p_T of the jet [57].

2.4.4 b-tagging

One important aspect of jet physics is the task of identifying the flavor of parton that produced the measured jet. While in general this is very difficult, jets from b-quarks offer an interesting case where such identification is possible. B mesons have a lifetime on the order of 10^{-12} seconds, which makes a $c\tau$ on the order of millimeters [6]. This type of displaced decay vertex can be identified in detectors like ATLAS and allows b-jets to be distinguished from other flavors of jets[2].

[2] Jets from charm quarks can also be detected in this way but they do not live quite as long so the displacement of the vertex is harder to distinguish

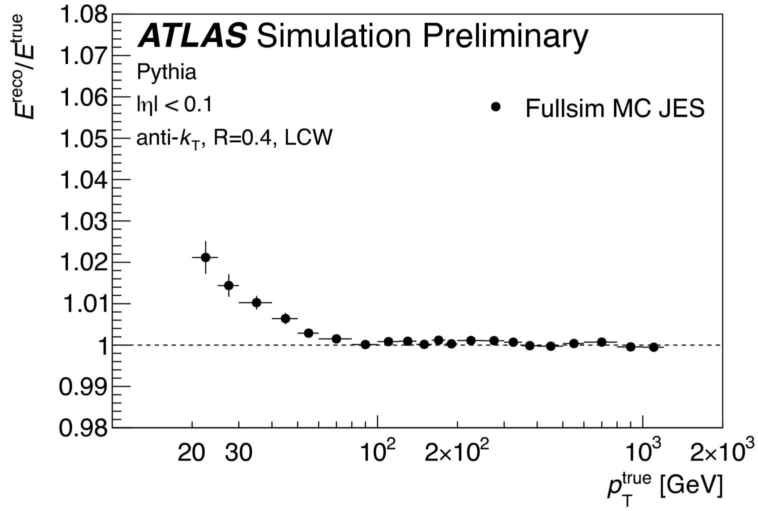


Figure 2.17: Jet energy response after calibration as a function of true p_T in simulation [57]

ATLAS uses a multivariate machine learning algorithm to identify jets from b-quarks. The inputs to this algorithm are determined from lower level reconstruction algorithms. There are three distinct algorithms that reconstruct variables which are used as input to the multivariate technique.

The first family is referred to as IPxD (where the x can either be 2 or 3). These algorithms use the transverse and longitudinal impact parameters d_0 and z_0 of the tracks inside a jet to determine their consistency with the primary vertex. They two or three dimensional (hence the x) templates for light flavor, charm, and bottom jets and then evaluate the likelihood of the jet coming from each of these types. The likelihood ratios are used as inputs to the multivariate algorithm.

The next two algorithms used as input are referred to as the secondary vertex (SV) and JetFitter (JF) algorithms. The SV algorithm uses tracks inside the jet to fit for vertices that are displaced from the primary vertex. The JF algorithm attempts to reconstruct the full flight path of the b by looking for multiple displaced vertices along the same line (as B decays often result in subsequent c decays).

In Run 1, the multivariate b-tagging algorithm used a neural network and was referred to as MV1. The details of this algorithm and its inputs are given in reference [58]. In Run 2, the number of inputs was simplified and a boosted decision tree with 24 input variables was used, referred to as MV2 [59]. The MV2 algorithm is a boosted decision tree incorporating twenty-four input variables constructed from three lower level input algorithms described above. Figure 2.18 summarizes the inputs to MV2. Figure 2.19 shows the performance of each of these algorithms.

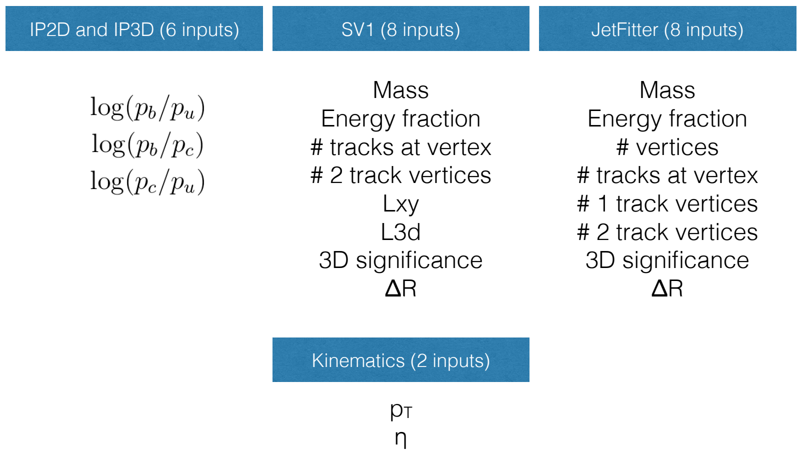
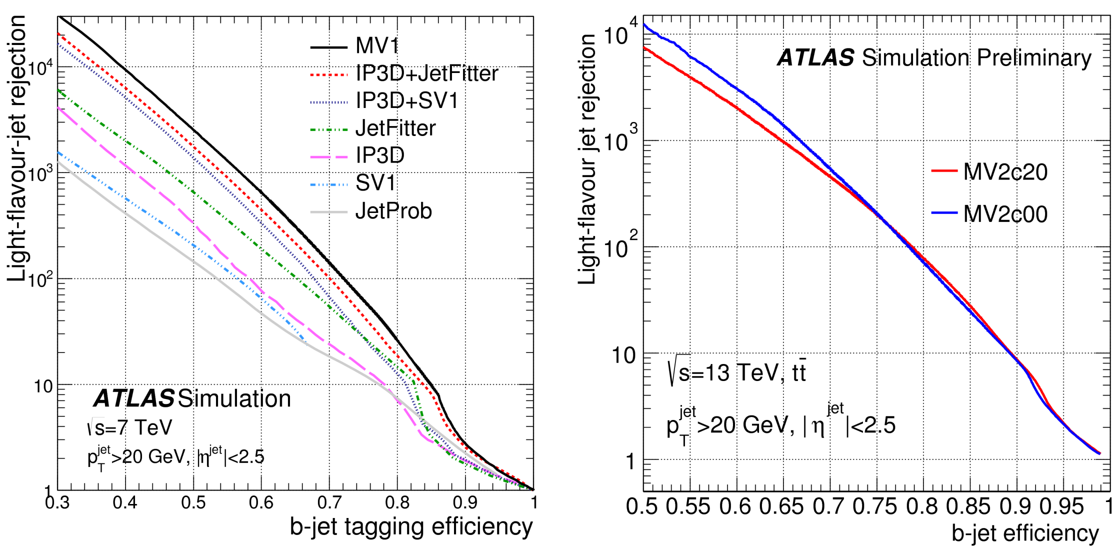


Figure 2.18: Summary of the inputs to the MV2 *b*-tagging algorithm



(a) (b)

Figure 2.19: Light jet rejection (1/efficiency) vs. *b*-jet efficiency for MV1 and its input algorithms (a) [58] and MV2 (b) [59] in simulated $t\bar{t}$ events. The numbers in the algorithm names in (b) refer to the fraction of charm events used in the MV2 training.

2.4.5 Missing transverse energy

As noted in figure 2.14, neutrinos produced in ATLAS will pass through the detector without interacting. The only way of detecting the presence of particles like neutrinos (or BSM particles that are long-lived) is to use missing transverse momentum. The basic principle of missing transverse energy is to use the momentum balance of the incoming protons to infer the presence of missing particles. The net longitudinal momentum of the incoming partons that collide is not known (since each carries an unknown fraction of the proton's momentum). However, the protons (and thus incoming partons) have no net momentum in the plane transverse to the beam line (the x-y) plane. Therefore, if there are no un-measured particles in the final state, the transverse momenta of all of the final state particles should balance. The magnitude of this imbalance is known as missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$).

The basic calculation of missing transverse momentum from calorimeter cells is given in equation 2.4 [60].

$$\begin{aligned} E_x^{\mathrm{miss}} &= -\sum_{i=1}^{N_{\mathrm{cell}}} E_i \sin\theta_i \cos\phi_i \\ E_y^{\mathrm{miss}} &= -\sum_{i=1}^{N_{\mathrm{cell}}} E_i \sin\theta_i \sin\phi_i \end{aligned} \tag{2.4}$$

The $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is separated into different terms based on the objects that the calorimeter clusters are associated with. This way, each cell's constribution is calibrated appropriately according to the object. This separation of terms is shown in equation 2.5 [60].

$$\begin{aligned} E_{x(y)}^{\mathrm{miss,calo}} = E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss},\gamma} + E_{x(y)}^{\mathrm{miss},\tau} + E_{x(y)}^{\mathrm{miss,jets}} \\ + E_{x(y)}^{\mathrm{miss,softjets}} + E_{x(y)}^{\mathrm{miss},\mu} + E_{x(y)}^{\mathrm{miss,CellOut}} \end{aligned} \tag{2.5}$$

The CellOut term of the above equation corresponds to calorimeter cells with energy deposits that are not associated with other objects. The soft jets term comes from cells associated to jets with p_{T} between 7 and 20 GeV, while the jets term comes from jets with $p_{\mathrm{T}} > 20$ GeV. Because muons do not deposit significant energy in the calorimeter, the muon momentum is used for the muon term [60]. The final $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated using equation 2.6.

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{\left(E_x^{\mathrm{miss}}\right)^2 + \left(E_y^{\mathrm{miss}}\right)^2} \tag{2.6}$$

Figure 2.20 shows the resolution of the components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ under different pileup suppression techniques [61].

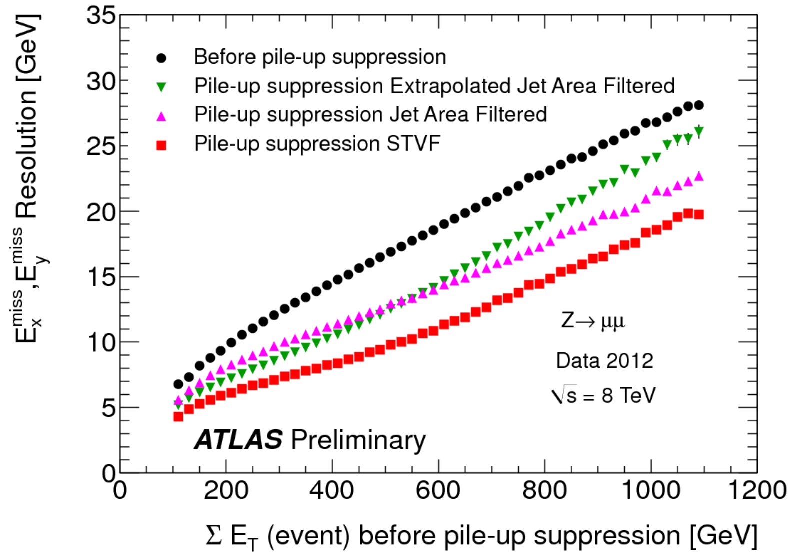


Figure 2.20: Resolution of $E_{\mathrm{T}}^{\mathrm{miss}}$ components as a function of $\sum E_T$ before pileup suppression with different pileup techniques [61]

Part II

Observation and measurement of Higgs boson decays to WW^* in LHC Run 1 at $\sqrt{s} = 7$ and 8 TeV

Basic research is what I am doing when I don't know what I am doing.

Wernher von Braun

3

$H \to WW^* \to \ell\nu\ell\nu$ Analysis Strategy

3.1 Introduction

This chapter presents an overview of the strategy for searching for a Higgs boson in the $H \to WW^* \to \ell\nu\ell\nu$ decay topology. Its purpose is to define in broad terms how the search and measurement are undertaken, before going into details on the specific sub-categories within the larger analysis. First, the properties of the Higgs signal are discussed and the associated backgrounds are presented. Next, the observables used to enhance the signal to background ratio are defined. Finally, the parameters of interest in the search and measurement will be shown, along with a brief overview of the statistical treatment of the final Higgs candidates.

Following this chapter, the results of three different studies within the $H \to WW^* \to \ell\nu\ell\nu$ channel are shown. Chapter 4 presents a search for Higgs boson production in gluon fusion mode and the role of the $H \to WW^*$ channel in its discovery. Chapter 5 shows the search and first observation in ATLAS of the Vector Boson Fusion (VBF) production mode of the Higgs in the $H \to WW^*$ decay channel. Finally, chapter 6 shows the combined Run 1 $H \to WW^*$ results for the measurement of the Higgs cross section and relative coupling strengths to other SM particles.

3.2 The $H \to WW^* \to \ell\nu\ell\nu$ signal in ATLAS

The signal studied in this and subsequent chapters is the Higgs boson in the WW^* final state, where each W boson subsequently decays into a charged lepton and a neutrino. In the simplest decay path, the final state consists of two neutrinos and two charged leptons, each of which can be either an electron or a muon. If a W decays to a τ lepton, only leptonic decays of the τ are considered. The τ lepton produce additional neutrinos in the final state but still yield two charged leptons (where each lepton is an electron or muon). Neutrinos are not detected in ATLAS, so the final state ultimately consists of two reconstructed leptons and missing transverse momentum (denoted as $E_{\mathrm{T}}^{\mathrm{miss}}$). Final states where both of the charged leptons are electrons or muons are referred to as the "same flavor" ($ee/\mu\mu$) final states, while those with one electron and one muon are referred to as "different flavor" ($e\mu$ or μe).

There can be additional jets produced in association with the Higgs boson. As described in detail in Chapter 1, if the Higgs is produced via vector boson fusion production, there will be two additional forward jets in the event. In gluon fusion, one or more jets can be produced through initial state radiation from the incoming gluons. Because of the varying background composition as a function of jet multiplicity, each bin in this variable has its own dedicated requirements applied in the search and measurement. The $n_j = 0$ and $n_j = 1$ bins are dedicated to gluon fusion production, while the $n_j \geq 2$ bin has separate dedicated searches for ggF and VBF production.

Figure 3.1 shows the relative branching fractions for the $H \to WW^*$ process, calculated from the Particle Data Group values for the W and τ branching ratios [6]. The largest branching ratio corresponds to both W bosons decaying to quark pairs at 45.44%. The second largest ratio is for one W decaying leptonically and the other decaying to quarks, a branching ratio of 34.18%. In all cases, ℓ denotes either an electron or muon, and the leptonic branching ratios of the τ are included. For example, the $\ell\nu qq$ final state includes one W decaying to $e\nu$, $\mu\nu$, or $\tau\nu$. In the case of the $W \to \tau\nu$ decay, the τ lepton then decays to an electron or muon via $\tau \to \nu_\tau \ell \nu_\ell$. Final states with a τ_h refer to hadronic decays of the τ. The

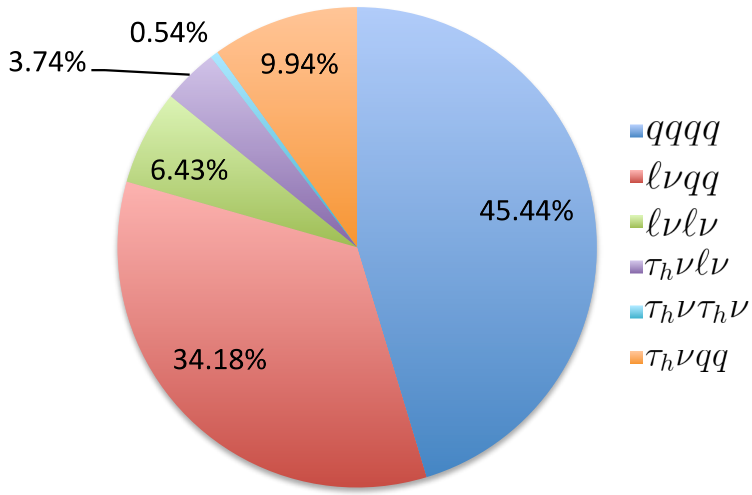


Figure 3.1: Branching ratios for a WW system. q refers to quarks. ℓ can be either an electron or muon, and the leptonic branching ratios of the τ are included. For example, the $\ell\nu qq$ final state includes one W decaying to $e\nu$, $\mu\nu$, or $\tau\nu$. τ_h refer to hadronic decays of the τ.

branching ratio of the $\ell\nu\ell\nu$ final state is 6.43%.

While the $\ell\nu\ell\nu$ final state is not a large fraction of the branching ratio, there are significant advantages to using this channel in an analysis. First, both the $qqqq$ and $\ell\nu qq$ channels suffer from a large QCD multijet background, which is often difficult to model. Second, events in the the $\ell\nu\ell\nu$ channel in data can be triggered more efficiently due to the presence of two leptons.

3.3 Background processes

Many processes from the Standard Model can also produce a final state with two leptons and missing transverse momentum. This section describes the dominant backgrounds to Higgs production and further explains how they can be reduced. Table 3.1 summarizes the different background processes.

3.3.1 Standard Model WW production

Non-resonant Standard Model diboson production, as shown in figure 3.2, is an irreducible background to Higgs boson production in the WW final state. It produces the same exact final state objects, namely

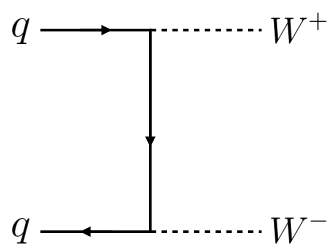


Figure 3.2: Feynman diagram for Standard Model WW production

leptonically decaying W bosons. There are no additional objects in the final state that allow for background reduction. Therefore the analysis solely relies on the correlations between the leptons to reduce this background.

3.3.2 Top quark production

Top quark production can mimic the Higgs in the WW^* final state as well. Top quarks can be produced either in pairs ($t\bar{t}$ production) or singly (s-channel, t-channel, or associated production Wt). The dominant top background are $t\bar{t}$ and Wt production.

Because top quarks decay via $t \to Wb$, top pair production can produce a final state with two W bosons that then decay leptonically. In Wt production, there are two real W bosons produced, as with $t\bar{t}$. In both cases, there is at least one b-jet in the final state. By vetoing on the presence of b-jets, these top quark backgrounds can be reduced. Figure 3.3 shows the Feynman diagrams for $t\bar{t}$ and Wt production.

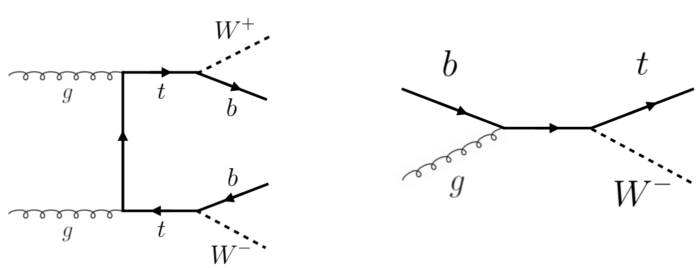


Figure 3.3: Feynman diagrams for top pair production (left) and Wt production (right)

3.3.3 W+jets background

Single W boson production in association with jets is a unique background to Higgs production. The other backgrounds considered thus far have all included two prompt leptons, each decaying from a W boson, in the final state. In W+jets production, however, only one reconstructed lepton originates from a W. The second reconstructed lepton is either an algorithmic "fake" or the result of non-prompt decays. In the first case, the lepton is a jet misidentified as a lepton by either the electron or muon reconstruction algorithms. In the second case, the lepton may be a real lepton but coming from semi-leptonic decays of particles inside the shower of the jet. This background can be reduced by requiring that the reconstructed lepton have little activity in the surrounding region of the calorimeter (also known as an "isolation"). Figure 3.4 shows the Feynman diagram for W+jets production.

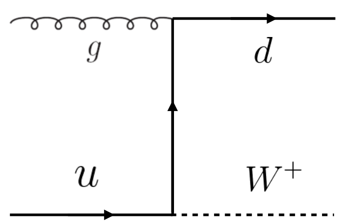


Figure 3.4: An example Feynman diagram of W+jets production

3.3.4 Z/γ^*+jets background

Production of a Z boson or virtual photon (also known as Drell-Yan and denoted with Z/γ^*) in association with jets is also a background to Higgs production. The Z boson decays to two leptons of the same flavor. When the Z/γ^* decays directly to electrons or muons, the background enters the same flavor final state. When the Z decays to two τ leptons the background can enter the different flavor final state as well. Figure 3.5 shows the production of a Z in association with one jet. Because there are no neutrinos in this final state, variables like $E_{\mathrm{T}}^{\mathrm{miss}}$ can be used to reduce the background[1].

[1]The $E_{\mathrm{T}}^{\mathrm{miss}}$ cut is much more effective for the reduction of Z/γ^* production in the same flavor final state. If the background enters the different flavor final state through τ decays, there will be neutrinos present. Other requirements on the lepton invariant mass are made to reduce the $Z/\gamma^* \to \tau\tau$ background.

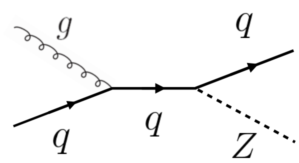


Figure 3.5: An example Feynman diagram of Z+jets production

3.3.5 Subdominant backgrounds

There are additional processes which contribute to the background composition. These backgrounds are subdominant and contribute less to the total background estimate than those discussed previously. The first process is referred to as VV or "Other diboson" processes and includes multiple Standard Model diboson processes, including WZ, ZZ, $W\gamma$, $W\gamma^*$, and $Z\gamma$ production. Additionally, there is a background contribution from QCD multijet production. While the cross section for this process is large, its contribution to the WW^* final state is small because two jets must be misidentified as leptons.

Category	Process	Description
SM WW	$WW \to \ell\nu\ell\nu$	Real leptons and neutrinos
Top quark production	$t\bar{t} \to WbW\bar{b} \to \ell\nu b\ell\nu\bar{b}$	Real leptons, untagged bs
	$tW \to WbW \to \ell\nu\ell\nu b$	Real leptons, untagged b
	$t\bar{b}$, $tq\bar{b}$	Untagged b, jet misidentified as lepton
Drell-Yan	$Z/\gamma^* \to ee, \mu\mu$	"Fake" $E_{\mathrm{T}}^{\mathrm{miss}}$
	$Z/\gamma^* \to \tau\tau \to \ell\nu\nu\ell\nu\nu$	Real leptons and neutrinos
Other dibosons	$ZZ \to \ell\ell\nu\nu$	Real leptons and neutrinos
	$W\gamma^*, WZ \to \ell\nu\ell\ell, ZZ \to \ell\ell\ell\ell$	Unreconstructed leptons
	$W\gamma, Z\gamma$	γ reconstructed as e, unreconstructed lepton
W+jets	$Wj \to \ell\nu j$	Jet reconstructed as lepton
QCD multijet	jj	Jets reconstructed as leptons

Table 3.1: A summary of backgrounds to the $H \to WW^* \to \ell\nu\ell\nu$ signal

3.4 Shared signal region selection requirements

As presented in section 3.2, there are many different combinations of physics objects that can define a $H \to WW^* \to \ell\nu\ell\nu$ final state. The multiplicity of jets and the flavor combinations of the leptons both lead to many potential signal regions. Additionally, signal regions can be optimized separately to be sensitive to the distinct production modes of the Higgs. Gluon fusion, vector boson fusion, and associated production of a Higgs all lead to unique final state topologies. Figure 3.6 delineates the different signal regions used in the gluon fusion and vector boson fusion $H \to WW^*$ analyses. While there are different optimizations possible in each signal region, there are also some commonly shared selections that will be described here.

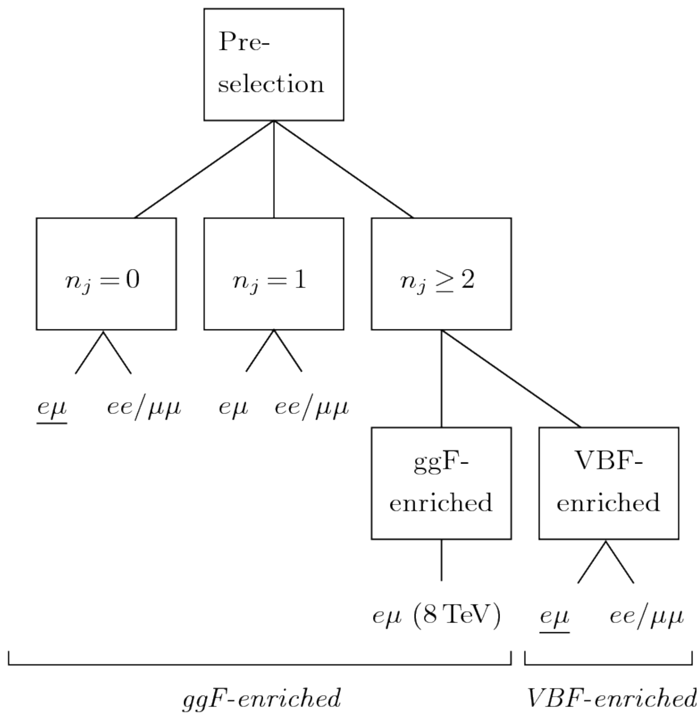


Figure 3.6: An illustration of the unique analysis signal regions [62]. The most sensitive regions for both gluon fusion and vector boson fusion production are underlined.

3.4.1 Event pre-selection

Before being sorted into the distinct signal regions, basic requirements are applied to the reconstructed objects in the event to select Higgs-like event candidates. First, two oppositely charged leptons are required. Once the leptons are selected, the last requirement for event pre-selection is the presence of neutrinos. As neutrinos cannot be detected directly in ATLAS, $E_{\mathrm{T}}^{\mathrm{miss}}$ can be used as a proxy for the combined neutrino momentum in the transverse plane.

In general, it is expected that the signal should have a harder $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum than backgrounds, especially if these backgrounds do not contain neutrinos in the final state. When using $E_{\mathrm{T}}^{\mathrm{miss}}$, it is possible mis-measurements of objects in the detector can lead to imbalances in the transverse plane. When such a mis-measurement occurs, the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector in the transverse plane will often point in the same direction as the mis-measured object. Therefore, a new variable, $E_{\mathrm{T,rel}}^{\mathrm{miss}}$, is used in the pre-selection. $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ is defined in equation 3.1.

$$E_{\mathrm{T,rel}}^{\mathrm{miss}} = \begin{cases} E_{\mathrm{T}}^{\mathrm{miss}} \sin \Delta\phi_{\mathrm{near}} & \text{if } \Delta\phi_{\mathrm{near}} < \pi/2 \\ E_{\mathrm{T}}^{\mathrm{miss}} & \text{otherwise,} \end{cases} \tag{3.1}$$

If the closest object to the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is within $\pi/2$ radians in the transverse plane, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is projected away from this object. Otherwise, the normal $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is used. Figure 3.7 shows a graphical illustration of this concept.

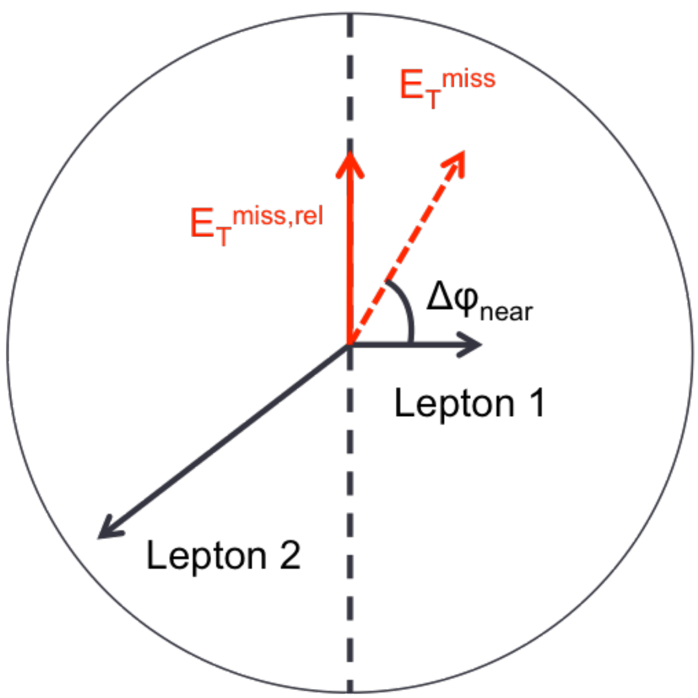


Figure 3.7: A graphical illustration of the $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ calculation

Once the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ pre-selections are made, the analysis is divided into different regions according to jet multiplicity.

3.4.2 Jet multiplicity

Jet multiplicity, denoted as n_j, is used to sub-divide the analysis into distinct signal regions. By creating separate signal regions, each bin in jet multiplicity becomes sensitive to different modes of Higgs production and different backgrounds.

For example, the $n_j \geq 2$ region is more sensitive to VBF production because of the two high momentum jets produced at matrix element level. For gluon fusion production to enter this bin, two initial state radiation jets must be emitted.

Figure 3.8 shows the jet multiplicity in both the different flavor and same flavor regions after the pre-selection. It also shows the background composition in the bins of n_b. A few trends from this distribution are worth noting. The first is that the Drell-Yan background dominates in the same flavor channels for $n_j \leq 1$. Second, the top background becomes a clear contributor to the total background for $n_j \geq 1$. Lastly, the SM WW production dominates in the $n_j = 0$ bin, as it is an irreducible background to $H \rightarrow WW^*$ production. Because of these distinct features, each jet multiplicity bin is treated separately.

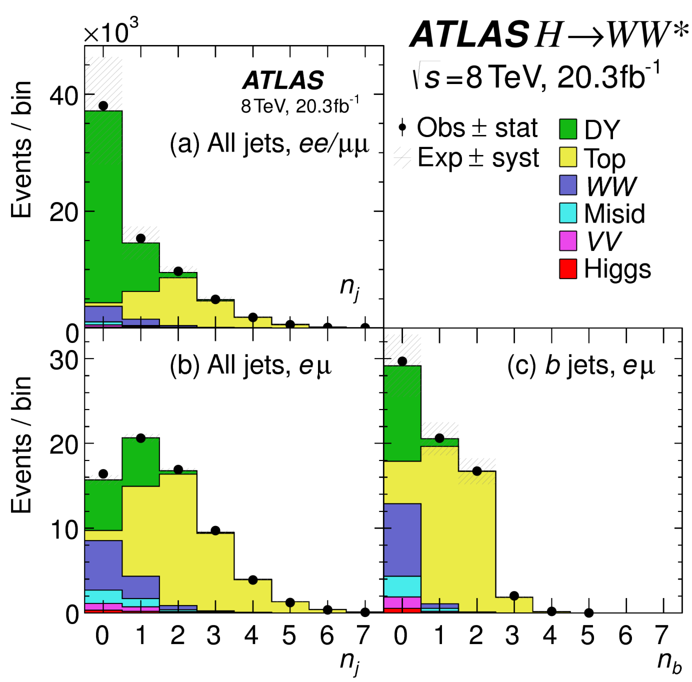


Figure 3.8: Predicted backgrounds (compared with data) as a function of n_j (a and b) and n_b (c) after pre-selection requirements

3.5 Background reduction in same-flavor final states

As described in section 3.4.2, the background composition of the same flavor final states is different from that of the different flavor states. In particular, Drell Yan processes play a much larger role because the Z/γ^* decays to same flavor leptons. Because real neutrinos are absent in the Z/γ^* decays to ee and $\mu\mu$, a requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$ should largely reduce the background. However, as this section will demonstrate, with increasing pileup conditions the resolution of the calorimeter-based $E_{\mathrm{T}}^{\mathrm{miss}}$ degrades greatly. Therefore, two new variables for Z/γ^* background reduction are constructed and described in this section.

3.5.1 Pileup and $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution

Secondary interactions of protons in the colliding bunches of the LHC (known as pileup interactions, described in detail in Chapter 2) deposit energy into the ATLAS calorimeter in addition to the energy that comes from the hard scatter process of interest. The calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$ is fundamentally Poissonian. Summing up all of the energy deposits in individual calorimeter cells or clusters is similar to a counting experiment. The error on a mean of N in a Poisson distribution is $\sqrt{N}$, so the energy resolution scales as $\sqrt{E}$. As more energy is deposited in the calorimeter, the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution degrades, meaning that the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is particularly sensitive to LHC instantaneous luminosity conditions.

Figure 3.9 shows an event display of a Z/γ^* + jets event candidate with the twenty-five reconstructed primary vertices. This display illustrates that while the interaction of interest only has tracks coming from the hardest primary vertex, all of the secondary interactions deposit energy in the calorimeter as well.

Figure 3.10 shows the RMS of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in $Z \to \mu\mu$ events (where there are no real neutrinos) as a function of the number of the average number of interactions. Under 2011 LHC conditions, this RMS was approximately 9 GeV, while under 2012 running conditions the resolution worsened to 12 GeV. The increase in pileup dilutes the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable's ability to reduce the Z/γ^* background.

3.5.2 Track-based definitions of missing transverse momentum

Because the increasing number of secondary proton-proton interactions degrades calorimeter-based $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution, a new variable using only contributions from the primary interaction vertex is necessary

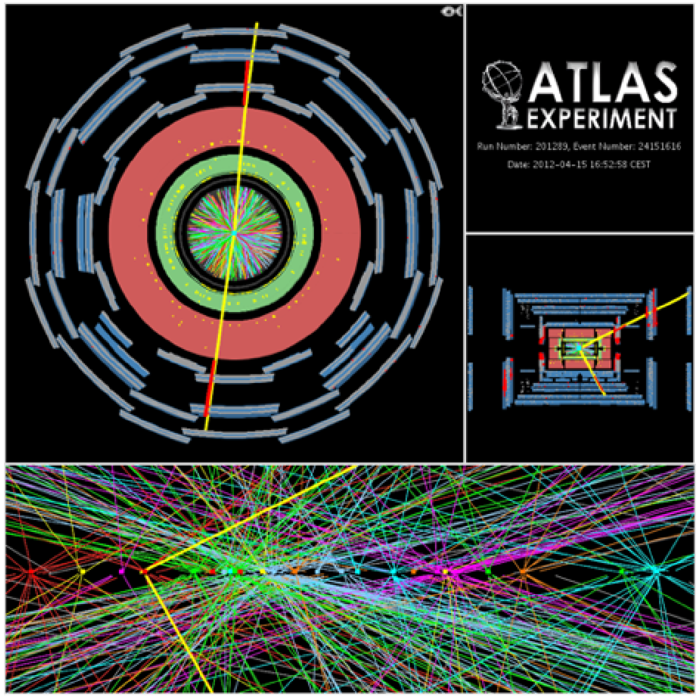


Figure 3.9: An event display of a Z/γ^* + jets event illustrating the effect of pileup interactions

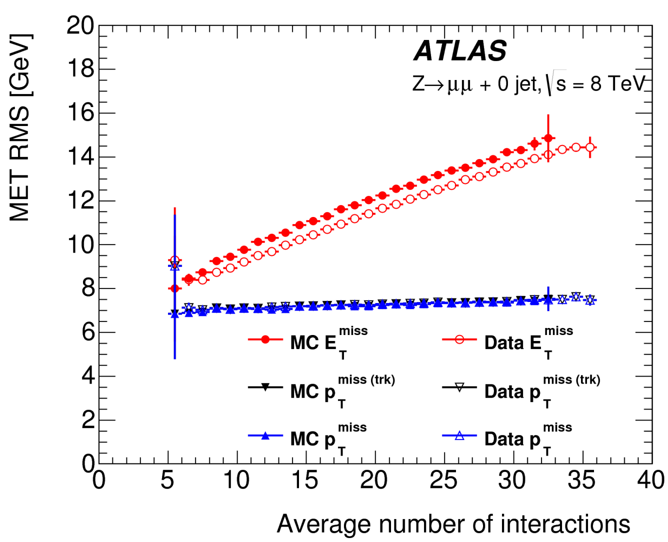


Figure 3.10: The RMS of different missing transverse momentum definitions as a function of the average number of interactions per bunch crossing

to further reduce the Z/γ^* background. While it is not possible to associate calorimeter energy deposits with a particular vertex, individual charged particle tracks in the Inner Detector are associated to unique vertices. Thus, two track-based definitions of missing transverse momentum , using only tracks coming from the primary vertex in the event, are used in the analysis. The simplest variable, $p_{\mathrm{T}}^{\mathrm{miss\,(trk)}}$, is the vectorial sum of the p_{T} of all of the tracks from the primary vertex and the selected leptons (excluding the tracks associated with the selected leptons to avoid double counting). Equation 3.2 defines $p_{\mathrm{T}}^{\mathrm{miss\,(trk)}}$.

$$\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss\,(trk)}} = -\left(\sum_{\substack{\text{selected}\\\text{leptons}}} \boldsymbol{p}_{\mathrm{T}} + \sum_{\substack{\text{other}\\\text{tracks}}} \boldsymbol{p}_{\mathrm{T}} \right), \tag{3.2}$$

To further improve the resolution on the missing transverse momentum, the variable $p_{\mathrm{T}}^{\mathrm{miss}}$ is used as defined in equation 3.3. For selected leptons and jets, the nominal p_{T} measurements are used. Tracks are used to estimate the soft component of the missing transverse momentum instead of calorimeter measurements.

$$\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} = -\left(\sum_{\substack{\text{selected}\\\text{leptons}}} \boldsymbol{p}_{\mathrm{T}} + \sum_{\substack{\text{selected}\\\text{jets}}} \boldsymbol{p}_{\mathrm{T}} + \sum_{\substack{\text{other}\\\text{tracks}}} \boldsymbol{p}_{\mathrm{T}} \right), \tag{3.3}$$

Figure 3.10 illustrates that these two new variables accomplish their intended purpose. The resolution as a function of mean number of interactions for both $p_{\mathrm{T}}^{\mathrm{miss\,(trk)}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$ is much flatter than the dependence for $E_{\mathrm{T}}^{\mathrm{miss}}$. Figure 3.11a shows the difference between the true and reconstructed values of missing transverse momentum using both the track-based $p_{\mathrm{T}}^{\mathrm{miss}}$ and calorimeter based $E_{\mathrm{T}}^{\mathrm{miss}}$. The RMS of the distribution improves by 3.5 GeVwhen using $p_{\mathrm{T}}^{\mathrm{miss}}$.

3.5.3 Distinguishing Z/γ^* +jets and $H \to WW^*$ topologies

In addition to measuring missing transverse momentum, another variable can be constructed to exploit kinematic and topological differences between the Z/γ^* background and $H \to WW^*$ signal. Because there are no real neutrinos in the final state (in the case of $Z/\gamma^* \to ee, \mu\mu$ decays), the dilepton system will be balanced with the jets produced in the hard scatter. A new variable, f_{recoil}, is constructed to estimate the balance between the dilepton system and recoiling jets and is defined in equation 3.4. The transverse plane is divided into four sections, or quadrants, with one quadrant centered on the dilepton vector. The numerator of f_{recoil} is the magnitude of the vectorial sum of the p_{T} of jets in the quadrant opposite the dilepton system, weighted by each jet's Jet Vertex Fraction (JVF, described in chapter 2). The denominator

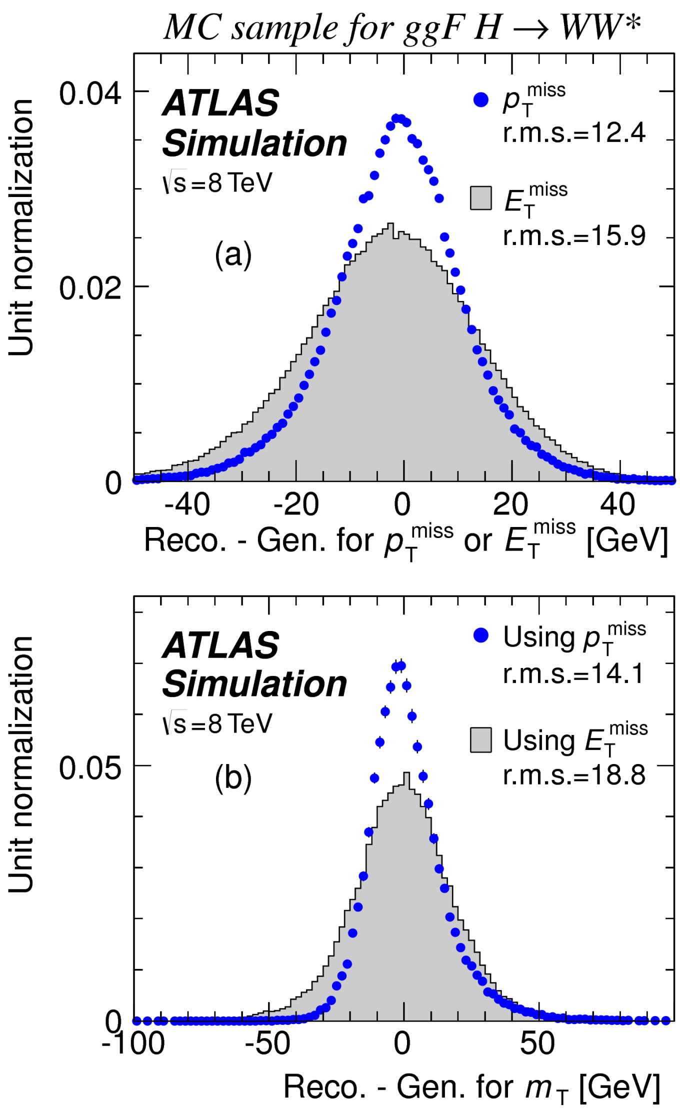


Figure 3.11: The difference between the true and reconstructed values of the missing transverse momentum (a) and m_{T} (b) in a gluon fusion signal sample

is the magnitude of the dilepton p_{T}.

$$f_{\text{recoil}} = \left| \sum_{\text{jets } j \text{ in } \wedge} \mathrm{JVF}_j \cdot \boldsymbol{p}_{\mathrm{T}}^{j} \right| \Big/ p_{\mathrm{T}}^{\ell\ell}. \tag{3.4}$$

Figure 3.12 shows a shape comparison of the f_{recoil} distribution in a simulated Z/γ^* + jets sample, a $H \to WW^*$ signal sample, and other backgrounds that contain real neutrinos. The Z/γ^* + jets events tend to be more balanced between the dilepton system and recoiling jets, while the processes containing real neutrinos are less balanced in the transverse plane. Thus, a requirement on f_{recoil} will reduce the Z/γ^* + jets background while maintaining a good signal efficiency.

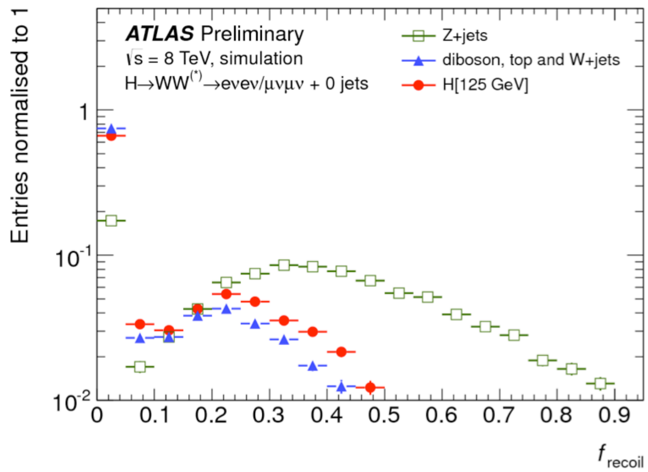


Figure 3.12: Comparison of f_{recoil} distributions for Z/γ^*+jets, $H \to WW^*$, and other backgrounds with real neutrinos.

3.5.4 Optimizing background reduction selection requirements

The requirements on $p_{\text{T}}^{\text{miss (trk)}}$ and f_{recoil} used to reduce the Z+jets background must be optimized to maximize expected signal significance in the same flavor channels. Figure 3.13 shows an optimization of the combination of the two requirements in the gluon fusion zero jet bin. Each bin shows the expected signal significance if the $p_{\text{T,rel}}^{\text{miss (trk)}}$ is required to be greater than the left edge of the bin and the f_{recoil} is required to be less than the top edge of the bin. The figure shows that the best signal significance comes from requiring low values of f_{recoil} (< 0.05) and $p_{\text{T,rel}}^{\text{miss (trk)}}$ values greater than 45 GeV.

3.6 Parameters of interest and statistical treatment

As with any search or measurement, there are particular parameters of the Higgs that the $H \to WW^*$ analysis is interested in measuring. In this case, the parameters of interest are the mass of the Higgs boson and its production cross section. Because the $H \to WW^* \to \ell\nu\ell\nu$ process does not have a closed final

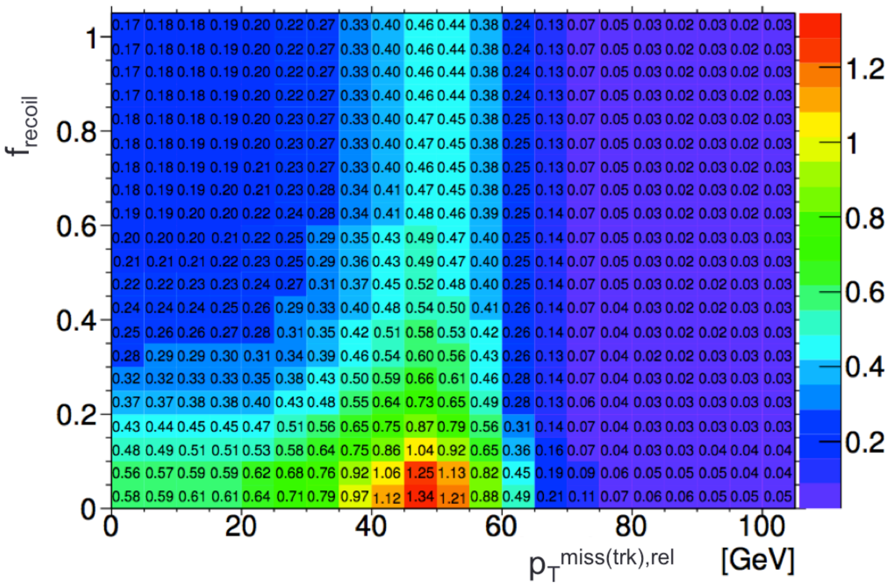


Figure 3.13: Signal significance as a function of required value for f_{recoil} and $p_{\text{T,rel}}^{\text{miss (trk)}}$ in the ggF $H \to WW^*$ with $n_j = 0$

state, it is not possible to measure the full invariant mass of the particle that may have produced the final state. However, a proxy for the invariant mass is defined using transverse plane information and detailed in section 3.6.1. The second parameter of interest is the ratio of the measured cross section to that expected from the Standard Model Higgs, which is denoted a μ. This is defined in equation 3.5.

$$\mu = \frac{\sigma}{\sigma_{\text{SM}}} \tag{3.5}$$

All of the likelihoods used in the statistical analysis of the final signal region events are parameterized as a function of μ. μ is a natural variable for hypothesis testing, as $\mu = 0$ corresponds to a background only hypothesis and $\mu = 1$ corresponds exactly to a Standard Model Higgs.

3.6.1 Transverse mass

The $H \to WW^* \to \ell\nu\ell\nu$ analysis cannot reconstruct the full invariant mass of the Higgs because of the neutrinos in the final state. The transverse mass serves as a proxy for the full invariant mass by exploiting information from the transverse plane. The transverse mass is defined in equation 3.6.

$$m_{\mathrm{T}} = \sqrt{\left(E_{\mathrm{T}}^{\ell\ell} + p_{\mathrm{T}}^{\mathrm{miss}}\right)^2 - \left|\,\boldsymbol{p}_{\mathrm{T}}^{\ell\ell} + \boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}\,\right|^2}, \tag{3.6}$$

Here the $E_{\mathrm{T}}^{\ell\ell}$ and $p_{\mathrm{T}}^{\ell\ell}$ are the transverse energy and momentum of the dilepton system, while $p_{\mathrm{T}}^{\mathrm{miss}}$ is a proxy for the transverse momentum of the di-neutrino system. The track-based $p_{\mathrm{T}}^{\mathrm{miss}}$ is used in the m_{T} rather than the calorimeter based $E_{\mathrm{T}}^{\mathrm{miss}}$ because it has a better resolution on the true transverse mass. Figure 3.11b shows the improvement in the RMS of the difference between the true and reconstructed transverse mass in a ggF signal sample. The RMS improves by 4.7 GeVusing $p_{\mathrm{T}}^{\mathrm{miss}}$ in the m_{T} calculation.

3.6.2 Statistical treatment[2]

Likelihood function

The statistical analysis of final event candidates is framed as a hypothesis test, where the null hypothesis is background-only (no Standard Model Higgs). The first step in the analysis is to form a likelihood function for the data. In its simplest form, this likelihood is the probability of observing the number of events seen in the final signal region given knowledge of the signal strength. Because observation of events is fundamentally a Poisson counting experiment, this simple likelihood can be expressed as a Poisson probability of observing N events given a total number of predicted signal and background events. This basic likelihood is shown in equation 3.7.

$$\mathcal{L}(\mu) = P\left(N|\mu S + B\right) \tag{3.7}$$

Here, P is the Poisson probability density function, N is the total number of observed events, μ is the signal strength, S is the predicted number of signal events, and B is the predicted number of background events.

In particle physics, certain background estimates are commonly normalized in so-called "control" regions and those predictions are scaled by the same normalization factor in the signal region. This leads to a

[2]Many thanks to Aaron Armbruster, whose thesis [63] inspired parts of this section.

slightly more complicated likelihood, which is a function of both the signal strength and the background normalization. This is shown in equation 3.8.

$$\mathcal{L}(\mu, \theta) = P\left(N|\mu S + \theta B\right) P\left(N_{\mathrm{CR}}|\theta B_{\mathrm{CR}}\right) \tag{3.8}$$

Here, θ serves as a "nuisance parameter", or a parameter that is not of primary interest but still enters the likelihood. The second Poisson term enforces that the background normalization be consistent with the number of observed events in data in the control region, N_{CR}.

So far, these two formulations of likelihoods have assumed a single signal region and do not take into account any shape information of potential discriminating variables. The $H \to WW^*$ analysis is divided into many different categories, the counting experiment described above can be performed in each individual category. As mentioned in section 3.6.1, the transverse mass is used as the primary discriminating variable in many of the $H \to WW^*$ sub-analyses. The same counting experiment can be performed in each bin of the m_{T} distribution to incorporate some shape information. Thus, the total likelihood becomes a product over signal regions and bins of the m_{T} distribution. Finally, there are usually many backgrounds that are normalized in control regions. The new formulation of the likelihood takes this into account by including a product over control regions in the second Poisson term. All of these modifications are shown in equation 3.9.

$$\mathcal{L}\left(\mu, \boldsymbol{\theta}\right) = \prod_{\substack{\mathrm{SRs\,i} \\ \mathrm{bins\,b}}} P\left(N_{ib} \middle| \mu S_{ib} + \sum_{\mathrm{bkg\,k}} \theta_k B_{kib}\right) \prod_{\mathrm{CRs\,l}} P\left(N_l \middle| \sum_{\mathrm{bkg\,k}} \theta_k B_{kl}\right) \tag{3.9}$$

The final step to obtain the full likelihood used in the analysis is to add nuisance parameters for the systematic uncertainties. In cases where the uncertainty does not affect the shape of m_{T} bin-by-bin, each systematic uncertainty ϵ is allowed to affect the expected event yields through a linear response function of the nuisance parameter, namely $\nu(\theta) = (1+\epsilon)^{\theta}$. If instead the uncertainty does affect the shape, the effect is instead parameterized by $\nu_b(\theta) = 1 + \epsilon_b \theta$. The value of the nuisance parameters for the systematic uncertainty are constrained with a Gaussian term that is added to the likelihood as well. This is of the form $g(\delta|\theta) = e^{-(\delta-\theta)^2/2}/\sqrt{2\pi}$, where δ is the central value and θ is a nuisance parameter. Finally, a last term is added to account for the statistical uncertainty in the Monte Carlo samples used, which adds an additional poisson term. The full likelihood used in the final statistical analysis is defined in equation 3.10.

$$\begin{aligned}\mathcal{L}\left(\mu,\boldsymbol{\theta}\right) = \prod_{\substack{\text{SRs i}\\ \text{bins b}}} P &\left(N_{ib} \middle| \mu S_{ib} \cdot \prod_{\substack{\text{sig.}\\ \text{syst.}\\ r}} \nu_{br}(\theta_r) + \sum_{\text{bkg k}} \theta_k B_{kib} \cdot \prod_{\substack{\text{bkg.}\\ \text{syst.}\\ s}} \nu_{bs}(\theta_s) \right) \\ &\cdot \prod_{\text{CRs l}} P \left(N_l \middle| \sum_{\text{bkg k}} \theta_k B_{kl} \right) \\ &\cdot \prod_{\substack{\text{syst}\\ t}} g\left(\delta_t | \theta_t\right) \cdot \prod_{\text{bkg k}} P\left(\xi_k | \zeta_k \theta_k\right)\end{aligned} \tag{3.10}$$

The fourth term of the equation quantifyies the uncertainty due to finite Monte Carlo sample size. Here, ξ represents the central value of the background prediction, θ is the associated nuisance parameter, $\zeta = (B/\delta B)^2$, where δB is the statistical uncertainty of B.

The best fit value of the signal strength μ is determined by finding the values of μ and $\boldsymbol{\theta}$ that maximize the likelihood, while setting $\delta = 0$ and $\xi = \zeta$. Once the likelihood is defined, a test statistic must be built for use in hypothesis testing.

Test statistic

To distinguish whether the data match a background only or background and signal hypothesis, a test statistic must be used. The $H \to WW^*$ analysis uses the profile likelihood technique [64]. The first step in formulating this test statistic is to define the profile likelihood ratio, shown in equation 3.11.

$$\lambda(\mu) = \frac{\mathcal{L}\left(\mu, \hat{\theta}_\mu\right)}{\mathcal{L}\left(\hat{\mu}, \hat{\theta}\right)} \tag{3.11}$$

Here $\hat{\theta}_\mu$ is the value of θ that maximizes the likelihood for the choice of μ being tested. Additionally, $\hat{\theta}$ and $\hat{\mu}$ represent the values of θ and μ that gives the overall maximum value of the likelihood.

Once this is defined, a test statistic q_μ is constructed. This is shown in equation 3.12.

$$q_\mu = -2\ln\lambda(\mu) \tag{3.12}$$

A higher value of q_μ indicates that the data are more incompatible with the hypothesized value of μ, and q_0 then corresponds to the value of the test statistic for the background only hypothesis. A p_0 value is then defined to quantify the compatibility between the data and the null hypothesis. The p_0 value is the probability of obtaining a value of q_0 larger than the observed value, and this is shown in equation 3.13.

$$p_0 = \int_{q_0^{\mathrm{obs}}}^{\infty} f(q_\mu|\mu = 0)dq_\mu \tag{3.13}$$

Here $f(q_\mu)$ is the probability distribution function of the test statistic. Finally, the p_0 value can be converted into a signal significance, using the formula in equation 3.14, or the one-sided tail of the Gaussian distribution.

$$Z_0 = \sqrt{2}\mathrm{erf}^{-1}(1 - 2p_0) \tag{3.14}$$

The threshold for discovery used in particle physics is $Z_0 \geq 5$, more commonly known as a value of 5σ.

The real voyage of discovery consists not in seeking new landscapes, but in having new eyes.

Marcel Proust

4

The discovery of the Higgs boson and the role of the $H \to WW^* \to \ell\nu\ell\nu$ channel

4.1 Introduction

This chapter presents the results of the search for the Higgs boson in $4.8\,\mathrm{fb}^{-1}$ collected at $\sqrt{s} = 7$ TeV and $5.8\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 8$ TeV. The results of three searches at $\sqrt{s} = 8$ TeV in the $H \to WW^* \to \ell\nu\ell\nu$, $H \to \gamma\gamma$, and $H \to ZZ \to 4\ell$ channels are combined with results of searches at $\sqrt{s} = 7$ TeV in the same search channels (as well as the $H \to \tau\tau$ production and associated production searches for $H \to b\bar{b}$). The results of this combination are a 5.9σ detection of a new particle consistent with a Higgs boson. Rather than going into detail for all of the different Higgs decay searches, this chapter will discuss the three most sensitive channels and in particular focus on $H \to WW^* \to \ell\nu\ell\nu$. While the focus is on WW^*, some of the ZZ^* and $\gamma\gamma$ results are shown for completeness. The results not discussed here can be found in the ATLAS Higgs discovery publication [1].

4.2 Data and simulation samples

The data sample used for the following results was taken in 2011 and 2012 at center of mass energies of 7 and 8 TeV, respectively, with $4.8\,\text{fb}^{-1}$ collected at 7 TeV and $5.8\,\text{fb}^{-1}$ collected at 8 TeV. Higgs production in the gluon fusion and vector boson fusion modes is modeled with POWHEG for the hard scattering event and PYTHIA for the showing and hadronization. Associated production of a Higgs with a vector boson or top quarks is modeled via PYTHIA . Table 4.1 shows the Monte Carlo generators used for modeling the signal and background processes relevant for the three analyses to be discussed.

Process	Generator
ggF, VBF H	POWHEG + PYTHIA
$WH, ZH, t\bar{t}H$	PYTHIA
W+jets, Z/γ^* + jets	ALPGEN + HERWIG
$t\bar{t}, tW, tb$	MC@NLO + HERWIG
tqb	ACERMC + PYTHIA
$q\bar{q} \to WW$	MC@NLO + HERWIG
$gg \to WW$	GG2WW+ HERWIG
$q\bar{q} \to ZZ$	POWHEG + PYTHIA
$gg \to ZZ$	GG2ZZ+ HERWIG
WZ	MADGRAPH+ PYTHIA , HERWIG
$W\gamma$ + jets	ALPGEN + HERWIG
$W\gamma^*$	MADGRAPH+ PYTHIA
$q\bar{q}/gg \to \gamma\gamma$	SHERPA

Table 4.1: Monte carlo generators used to model signal and background for the Higgs search [1].

4.3 $H \to WW \to e\nu\mu\nu$ search

The $H \to WW \to e\nu\mu\nu$ search is unique compared to the ZZ and $\gamma\gamma$ channels. The Higgs mass cannot be fully reconstructed due to the presence of neutrinos in the final state, so the transverse mass m_{T} is used as the final discriminating variable. Compared to the other channels, there are more backgrounds here as well, as discussed in chapter 3. The same flavor final states are excluded from this search due to high pileup in the 8 TeV dataset.

4.3.1 Event selection

The analysis requires to opposite charge isolated leptons, with the leading (sub-leading) lepton required to have $p_{\mathrm{T}} > 25(15)$ GeV. The events are separated into different signal regions depending on which flavor of lepton is leading ($e\mu$ for leading electron, μe for leading muon). Strict lepton quality cuts are applied to the sample to reduce backgrounds from fake leptons.

Jets are reconstructed with the anti-k_T algorithm with a radius parameter $R = 0.4$. The jets are required to have $p_{\mathrm{T}} > 25$ GeV and $|eta| < 4.5$, with jets in the tracking volume required to have a jet vertex fraction of 0.5 and jets in the forward region required to have $p_{\mathrm{T}} > 30$ GeV. The analysis is separated into three different signal regions based on jet multiplicity: $n_j = 0, 1, \geq 2$.

To indicate the presence of neutrinos in the event, a requirement of $E_{\mathrm{T,rel}}^{\mathrm{miss}} > 25$ GeV is made[1]. This requirement significantly reduces the QCD multijet and Z/γ^* + jets backgrounds. Figure 4.1 shows the distribution of n_j in data and simulation after applying these "pre-selection" requirements.

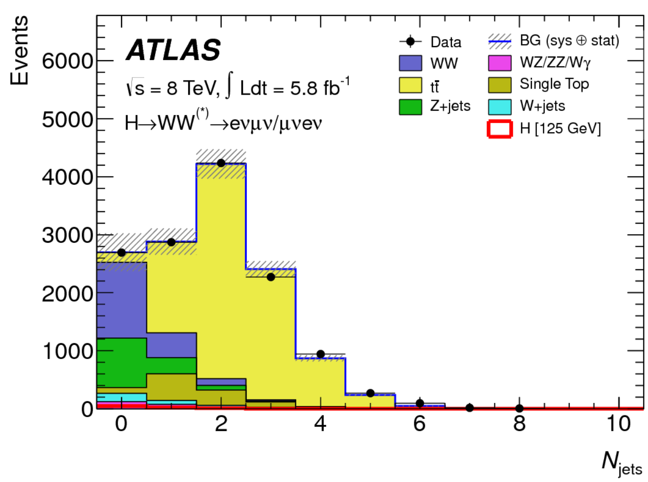


Figure 4.1: Jet multiplicity distribution in data and MC after applying lepton, jet, and $E_{\mathrm{T,rel}}^{\mathrm{miss}}$ selections. The WW and top backgrounds have been normalized using control samples, and the hashed band indicates the total uncertainty on the prediction. [1]

Additional selections are applied to require the dilepton topology to correspond to that of a SM Higgs. The requirements are presented here - more detailed discussion on the motivation for each requirement is saved for Chapter 5. In all of the jet multiplicity channels, the dilepton system is required to have a small

[1] For the definition of $E_{\mathrm{T,rel}}^{\mathrm{miss}}$, see chapter 3

gap in azimuthal angle, $\Delta\phi_{\ell\ell} < 1.8$. Similarly, the $m_{\ell\ell}$ is required to be less than 50 GeV in the lower jet multiplicity channels and less than 80 GeV in the $n_j \geq 2$ channel. In the $n_j = 0$ channel, the magnitude of the dilepton p_{T}, $p_{\mathrm{T}}^{\ell\ell}$, is required to be greater than 30 GeV.

In the higher jet multiplicity channels ($n_j \geq 1$), the top background is a more important component and must be reduced. The total transverse momentum $p_{\mathrm{T}}^{\mathrm{sum}}$ is thus required to be less than 30 GeV. Additionally, the di-τ invariant mass $m_{\tau\tau}$ (dilepton mass computed under the assumption that the neutrinos from the τ decay are emitted collinear to the charged leptons) is used to reject $Z \rightarrow \tau\tau$ events by requiring $|m_{\tau\tau} - m_Z| > 25$ GeV. These variables are also discussed in more detail in Chapter 5.

In the $n_j \geq 2$ channel, requirements are made to isolate the VBF contribution to Higgs production. The kinematics of the two leading jets are used to make these requirements. In particular, the event must have $\Delta y_{jj} > 3.8$ and $m_{jj} > 500$ GeV, along with a veto on having any additional jets with rapidity between the to leading jets.

4.3.2 Background estimation

The details of the background estimation techniques used in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ analysis are discussed in section 5.5. As that section refers to a later iteration of the analysis, a general discussion is given here for completeness. The dominant backgrounds are SM WW production and top (both pair and single) production, and these backgrounds have their normalizations estimated from dedicated control regions while their shapes are taken from simulation.

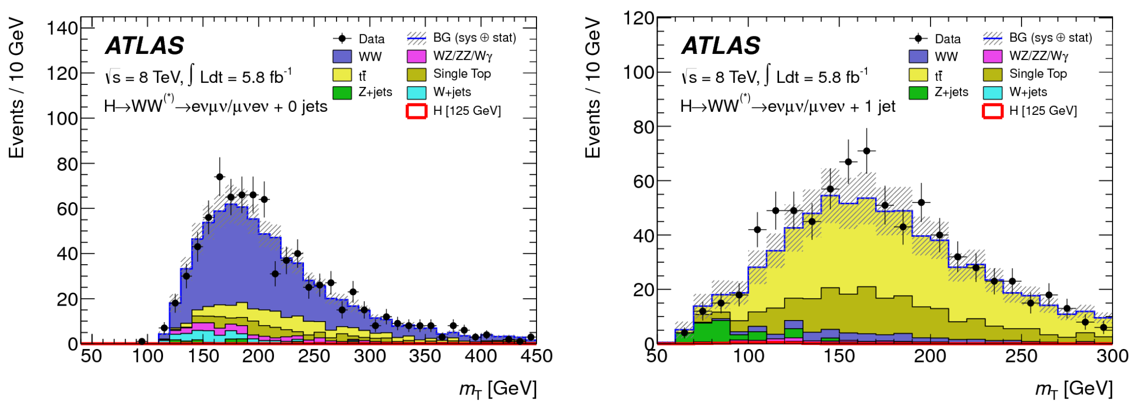
The control sample for the Standard Model WW background is defined by making the same requirements as the signal region with the $m_{\ell\ell}$ requirement inverted (now requiring $m_{\ell\ell} > 80$ GeV) and removing the $\Delta\phi_{\ell\ell}$ requirement. This creates a control sample that is 70% (40%) pure in the 0(1)-jet region. The correction to the pure MC-based background estimate is quantified by defining a normalization factor β which is the ratio of the data yield to the MC yield ($N_{\mathrm{data}}/N_{\mathrm{MC}}$) in this control sample. Table 4.2 shows the WW normalization factors in the $n_j = 0$ and $n_j = 1$ bins (the $n_j \geq 2$ estimate is taken directly from MC).

The top background estimate is also computed separately in each jet multiplicity bin. In the $n_j = 0$

n_j	β_{WW}	β_t
$= 0$	1.06 ± 0.06	1.11 ± 0.06
$= 1$	0.99 ± 0.15	1.11 ± 0.05
≥ 2	-	1.01 ± 0.26

Table 4.2: Normalization factors (ratio of data and MC yields in a control sample) for the Standard Model WW and top backgrounds in the $H \to WW^* \to \ell\nu\ell\nu$ analysis [1]. Only statistical uncertainties are shown.

channel, the background is first normalized using data after pre-selection requirements with no selection on n_j. Then, a dedicated b-tagged control sample is used to evaluate the ratio of one-jet to two-jet events in data. The details of this technique are shown in reference [65]. In the $n_j = 1$ and the $n_j \geq 2$ regions, the top background is normalized in a control sample where the signal region selections are applied, but the b-jet veto is reversed and the Higgs topology requirements on $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ are removed. The resulting normalization factors for these techniques are shown in table 4.2.



(a) (b)

Figure 4.2: Comparison of m_{T} between data and simulation in the $n_j = 0$ WW (a) and $n_j = 1$ top (b) control samples [1]

The control samples which are used for background normalization can also be used to validate the modeling of the m_{T} distribution for each background. Figure 4.2 shows the comparison between data and MC for the m_{T} distribution after correcting the normalization of the backgrounds in the WW and top control regions. Good agreement between data and simulation is seen in both cases.

The W+jets background estimate is taken entirely from data using a control sample with one well reconstructed lepton and one anti-identified lepton. All other backgrounds are taken purely from simulation.

4.3.3 Systematic uncertainties

The systematic uncertainties that have the largest impact on the analysis are the theoretical uncertainties associated with the signal cross section, and these are shared with the ZZ^* and $\gamma\gamma$ channels. The uncertainties resulting from variations of the QCD scale are $+7\%/-8\%$ on the final singal yield. Those coming from variations of the parton distribution function (PDF) used in the simulation add a $\pm 8\%$ uncertainty on the yield. The uncertainties on the branching ratios of the Higgs are $\pm 5\%$.

The main experimental uncertainties come from variations of the jet energy scale (JES), jet energy resolution, pile-up, $E_{\mathrm{T}}^{\mathrm{miss}}$, b-tagging efficiency, W+jets background estimate, and integrated luminosity. For more details, see reference [1].

4.3.4 Results

Table 4.3 shows the signal and background yields in the final signal region after normalizing the backgrounds according to the methods described above.

	$n_j = 0$	$n_j = 1$	$n_j \geq 2$
Signal	20 ± 4	5 ± 2	0.34 ± 0.07
WW	101 ± 13	12 ± 5	0.10 ± 0.14
Other dibosons	12 ± 3	1.9 ± 1.1	0.10 ± 0.10
$t\bar{t}$	8 ± 2	6 ± 2	0.15 ± 0.10
Single top	3.4 ± 1.5	3.7 ± 1.6	-
Z/γ^* + jets	1.9 ± 1.3	0.10 ± 0.10	-
W + jets	15 ± 7	2 ± 1	-
Total background	142 ± 16	26 ± 6	0.35 ± 0.18
Observed in data	185	38	0

Table 4.3: Data and expected yields for signal and background in the final $H \to WW^* \to \ell\nu\ell\nu$ signal region. Uncertainties shown are both statistical and systematic. [1]

Figure 4.3 shows the m_{T} distribution in the $n_j \leq 1$ channels for 8 TeV data. (No events are observed in data in the $n_j \geq 2$ channels in this dataset). The excess shown here relatively flat as a function of hypothesized Higgs mass. The combined 7 and 8 TeV data gives an excess with local significance of 2.8σ with an expected significance of 2.3σ, corresponding to a μ measurement of 1.3 ± 0.5.

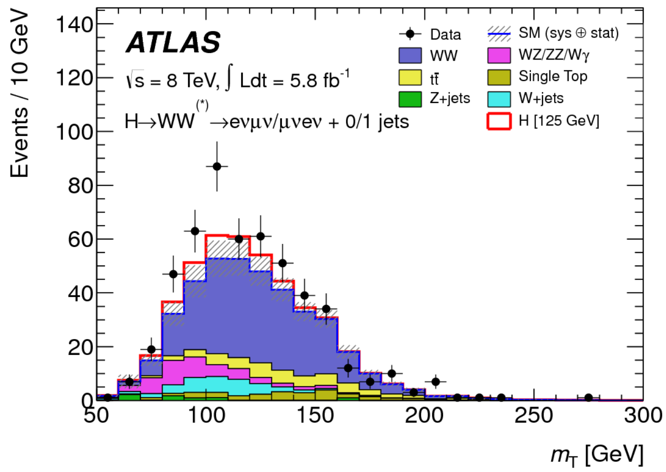


Figure 4.3: m_{T} distribution in the $H \to WW \to e\nu\mu\nu\ n_j \leq 1$ channels for 8 TeV data [1].

4.4 $H \to \gamma\gamma$ search

The $H \to \gamma\gamma$ search is in essence a search for a peaked excess above the falling SM diphoton mass spectrum, with $m_{\gamma\gamma}$ as the ultimate discriminating variable. Events are selected by requiring two isolated photons, with the leading (sub-leading) photon required to have $E_T > 40(30)$ GeV. In the 8 TeV data, the photons are required to pass cut-based identification criteria consistent with a photon in the electromagnetic calorimeter and little leakage in the hadronic calorimeter.

The main challenges for this analysis are accurate mass reconstruction and background estimation. In order to accurately reconstruct the invariant mass of the di-photon system, both the energy and direction of the photons must be measured well. Therefore, the identification of the primary vertex of the hard interaction is particularly important, and is done using a multivariate likelihood which combines information about the photon direction and vertex position. The background is modeled with a falling spectrum in $m_{\gamma\gamma}$ that is parameterized by different functions depending on the category of the event.

4.4.1 Results

The resulting diphoton mass spectrum is shown in figure 4.4. The best fit mass value in the $\gamma\gamma$ channel alone in the combined 7 and 8 TeV data is 126.5 GeV. The local significance as this point is 4.5σ, with an expected significance of 2.5σ. Therefore, the measured signal strength μ is 1.8 ± 0.5 in this channel.

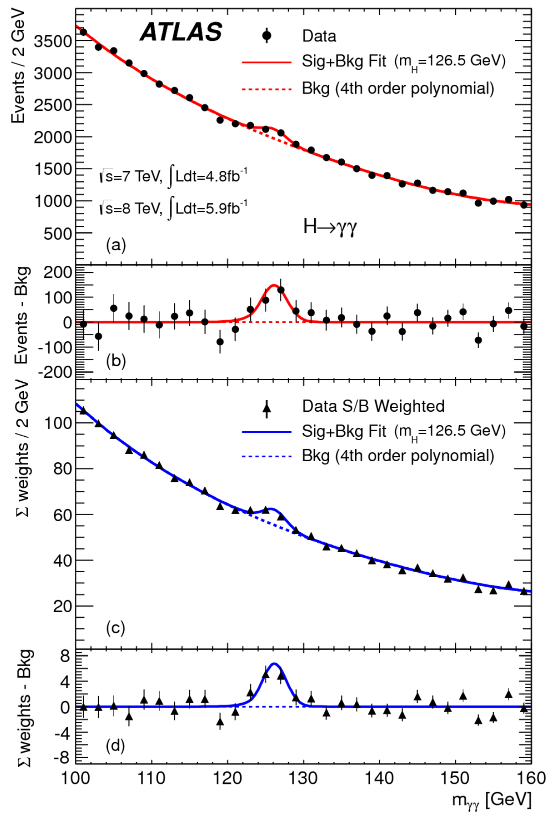


Figure 4.4: Diphoton mass spectrum in 7 and 8 TeV data. Panel a) shows the unweighted data distribution superimposed on the background fit, wile panel c) shows the data where each event category is weighted by its signal to background ratio. Panels b) and d) show the respective distributions with background subtracted [1].

4.5 $H \to ZZ \to 4\ell$ SEARCH

The $H \to ZZ \to 4\ell$ analysis searches for a Standard Model Higgs boson decaying to two Z bosons, each of which decays to a pair of same flavor, opposite charge isolated leptons. The ultimate discriminating variable is $m_{4\ell}$, or the invariant mass of the four selected leptons. The ℓ denotes an e or μ as with the $H \to WW^* \to \ell\nu\ell\nu$ analysis.

Four distinct signal regions are constructed depending on the flavors of the final state, additionally separated by the flavor of the leading lepton pair. These are referred to as $4e$, $2e2\mu$, $2\mu 2e$, 4μ.

The main backgrounds in the $H \to ZZ \to 4\ell$ search are continuum ZZ^* production, Z+ jets production, and $t\bar{t}$. The $m_{4\ell}$ distribution for background is estimated from simulation. The normalization of the SM ZZ^* background is also taken from MC simulation, while the Z+jets and $t\bar{t}$ normalizations are taken from data-driven methods.

4.5.1 Results

Figure 4.5 shows the $m_{4\ell}$ spectrum measured in the 7 and 8 TeV datasets. The total number of events observed in the window between 120 and 130 GeV is 13, with 6 events in the 4μ channel, 2 events in the $4e$ channel, and 5 events in the $2e2\mu/2\mu2e$. The best fit μ value in the combined 7 and 8 TeV data occurs at 125 GeV and is measured to be 1.2 ± 0.6. The observed significance at this mass is 3.6σ, with an expected significance of 2.7σ.

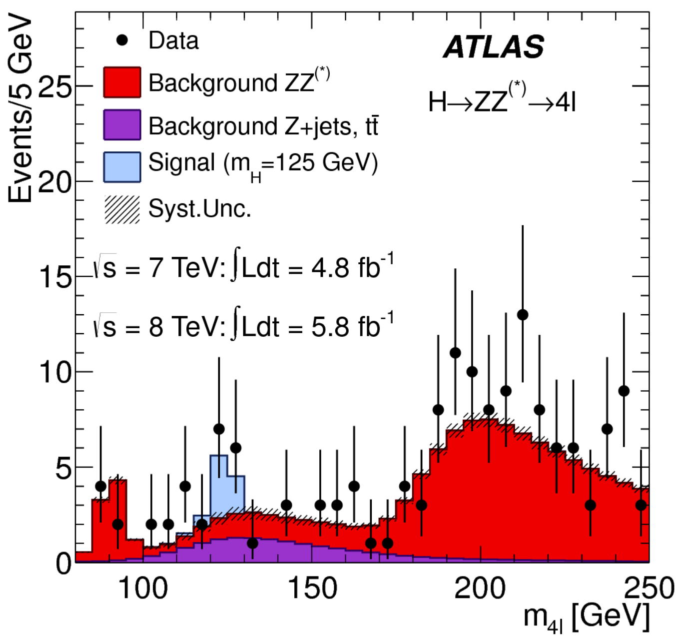


Figure 4.5: Four lepton invariant mass spectrum ($m_{4\ell}$) in 7 and 8 TeV data compared to background estimate. A 125 GeV SM Higgs signal is shown in blue [1].

4.6 Combined results

The statistical interpretation of the combined results is undertaken as described in section 3.6.2, with a hypothesis test based on a likelihood ratio parameterized by the Higgs signal strength μ. The null hypothesis corresponds to $\mu = 0$, while the SM Higgs corresponds to $\mu = 1$.

Table 4.4 summarizes the properties of the individual channels as well as the significances of the excesses seen. The most significant observed local excess comes from the $\gamma\gamma$ channel. Figure 4.6 shows a comparison of the observed local p_0 values as a function of hypothesized mass for the three different search channels. Both the ZZ^* and $\gamma\gamma$ channels have very peaked excesses, while the WW^* excess can be seen as very broad because the $m_{\rm T}$ distribution does not provide detailed information about the true Higgs mass.

Channel	Fit var.	Observed Z_l	Expected Z_l	$\hat{\mu}$
$H \to ZZ^* \to 4\ell$	$m_{4\ell}$	3.6	2.7	1.2 ± 0.6
$H \to \gamma\gamma$	$m_{\gamma\gamma}$	4.5	2.5	1.8 ± 0.5
$H \to WW^* \to e\nu\mu\nu$	$m_{\rm T}$	2.8	2.3	1.3 ± 0.5
Combined	-	6.0	4.9	1.4 ± 0.3

Table 4.4: Summary of the expected and observed significance and measured signal strengths in the combined 7 and 8 TeV datasets for the Higgs discovery analysis [1].

Figure 4.7 shows the combined exclusion limit, p_0, and signal strength. The highest local excess comes at a value of 126.5 GeV and corresponds to a 6.0σ observed excess.

Figure 4.8 shows a comparison of the measured signal strengths between the different Higgs search channels. All measured μ are consistent with unity within their uncertainty, and the combined μ measurement is 1.4 ± 0.3.

The likelihood can also be computed in a two-dimensional plane of m_H and μ, and this is shown in figure 4.9. The figure shows that while the $\gamma\gamma$ and ZZ^* channels have very good mass resolution, the excess in WW^* covers a broad mass range. The banana shape of the WW^* result is due to the fact that the excess in this channel can either be explained by increasing the signal strength or by changing the mass (and thus the cross section). The two parameters are correlated due to the lack of mass sensitivity in this channel.

Because multiple Higgs mass points are searched for, the local significance must be corrected for a look-elsewhere effect to compute a true global significance. The global significance for finding a Higgs anywhere in the mass range of 110 GeV to 600 GeV is 5.1σ. This increases slightly to 5.3σ if only mass range from 110 to 150 GeV.

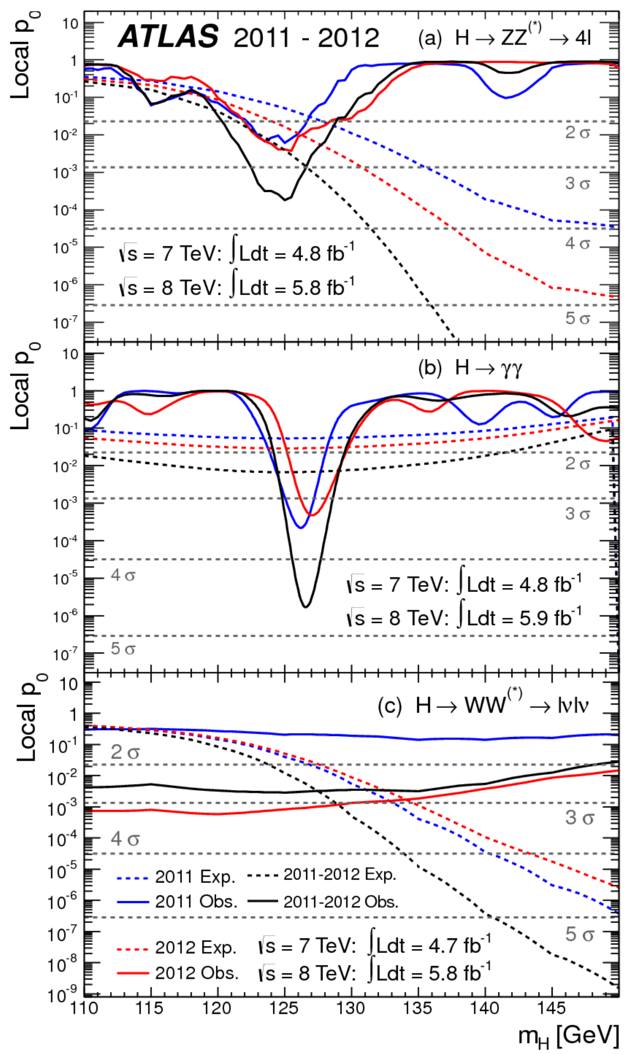


Figure 4.6: Local p_0 distribution as a function of hypothesized Higgs mass for the $H \to ZZ^* \to 4\ell$ (a), $H \to \gamma\gamma$ (b), and $H \to WW^* \to \ell\nu\ell\nu$ (c) channels. Dashed curves show expected results, while solid curves show observed. Red curves are from 7 TeV data, blue curves from 8 TeV, and black curved combined [1].

4.7 Conclusion

A search for the production of a Standard Model Higgs boson was conducted in $4.8\,\text{fb}^{-1}$ collected at $\sqrt{s} = 7$ TeV and $5.8\,\text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV. A new particle consistent with the Higgs boson was observed, with a mass of 126.5 GeV and a global (local) significance of $5.1(6.0)\sigma$. This is the first discovery level observation of a particle consistent with the Higgs.

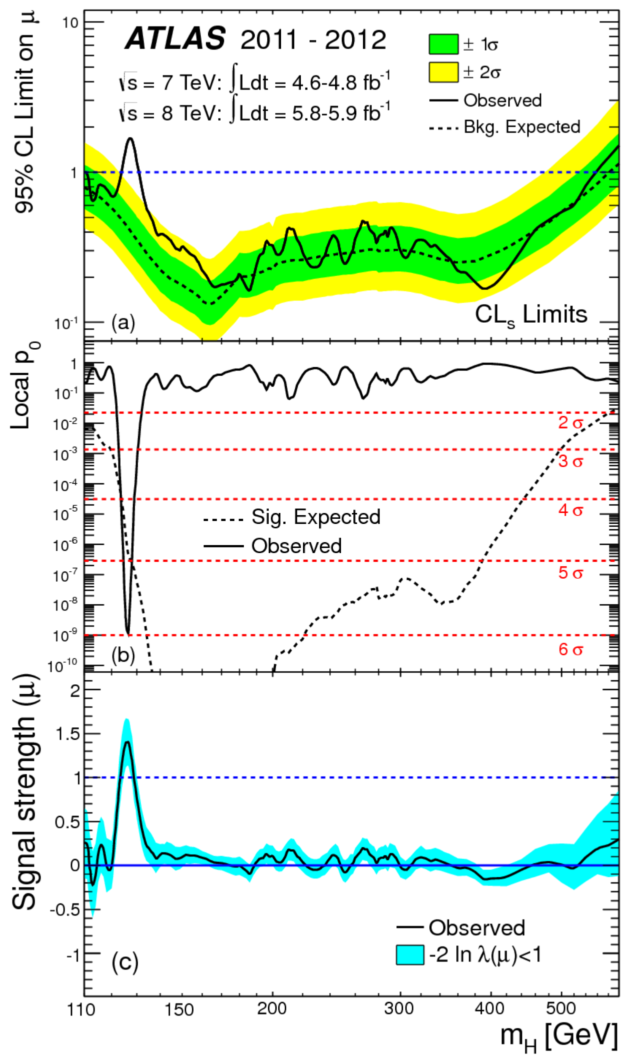


Figure 4.7: Combined 95% CL limits (a), local p_0 values (b), and signal strength measurement (c) as a function of Higgs mass [1].

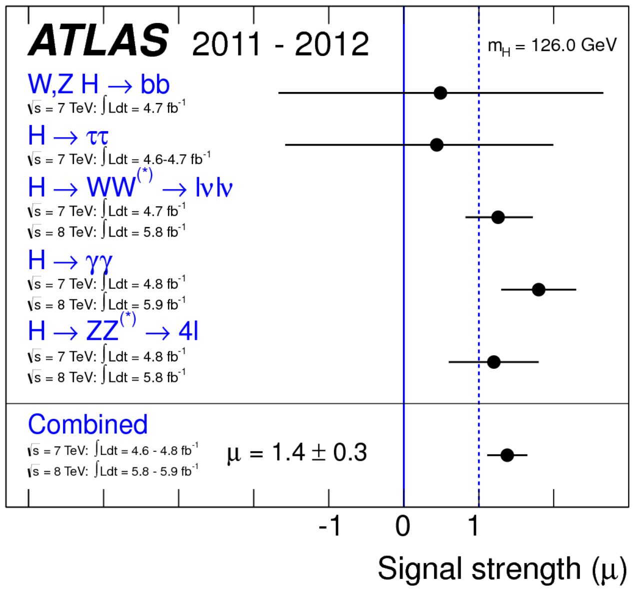


Figure 4.8: Comparison of measured signal strength μ for a 126 GeV Higgs in the 7 and 8 TeV datasets [1].

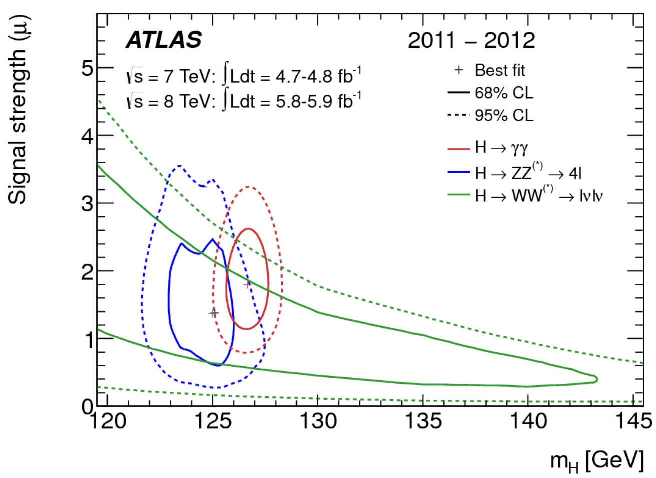


Figure 4.9: Two dimensional likelihood as a function of signal strength μ and Higgs mass m_H [1].

The imagination of nature is far, far greater than the imagination of man.

Richard Feynman

5

Observation of Vector Boson Fusion production of $H \to WW^* \to \ell\nu\ell\nu$

5.1 Introduction

After the discovery of a particle consistent with the Higgs boson, the $H \to WW^*$ analysis had two main goals. The first goal was to increase the sensitivity of the analysis to fully confirm that the $H \to WW^*$ process did indeed exist. The second goal was to characterize the particle as much as possible, including searching for the lower cross-section production modes, in order to confirm that it was indeed a Higgs boson. This chapter presents a dedicated search for Vector Boson Fusion (VBF) production of a Higgs boson decaying via the $H \to WW^* \to \ell\nu\ell\nu$ mode. First, basics of the topology of VBF production are presented. Then, the details of the analysis are shown, including signal region definition, background estimation techniques, and systematic uncertainties. Finally, the results of the analysis are presented. As will be shown, this analysis is the first and most sensitive observation of the VBF production mode of the Higgs on ATLAS.

In the VBF channel, there are both a selection requirement based signal region analysis (known as the "cut-based") and a multivariate analysis which uses a boosted decision tree (known as the BDT analysis). The focus of this chapter will be on the cut-based signal region, as this is an important component of the VBF analysis and in particular acts as strong validation for the final BDT result. Connections between the cut-based and BDT analyses will be discussed where appropriate.

5.2 Data and simulation samples

The results presented here are with 20.3 fb^{-1}taken at $\sqrt{s} = 8$ TeV and 4.5 fb^{-1}taken at $\sqrt{s} = 7$ TeV. The details of the LHC and detector conditions during this period are given in Chapter 2. The trigger selection defining the dataset is discussed in section 5.2.1. The simulation samples used for signal and background modeling are given in section 5.2.2.

5.2.1 Triggers

The analysis uses a combination of single lepton and dilepton triggers to allow lowering of the p_{T} thresholds and increased signal acceptance. The p_{T} threshold on the leptons is a particularly important consideration for this signal. Because the second W produced in the decay can be off-shell, it tends to produce lower momentum leptons. Thus, being able to lower the p_{T} threshold while still maintaining a low background rate is critical. Figure 5.1 shows an example of the subleading lepton p_{T} for a VBF $H \rightarrow WW^*$ signal compared to the corresponding $t\bar{t}$ background. Note that the lepton p_{T} spectrum is considerably softer in the signal sample.

As discussed in Chapter 2, there are multiple levels in the ATLAS trigger system, and there are different p_{T} thresholds imposed for the leptons at each level. Additionally, some triggers have a loose selection on the isolation of the lepton (looser than that applied offline in the analysis object selection). Table 5.1 shows the thresholds used for single lepton triggers, while table 5.2 shows the thresholds coming from di-lepton triggers. The single lepton trigger efficiency for muons that pass the analysis object selection is

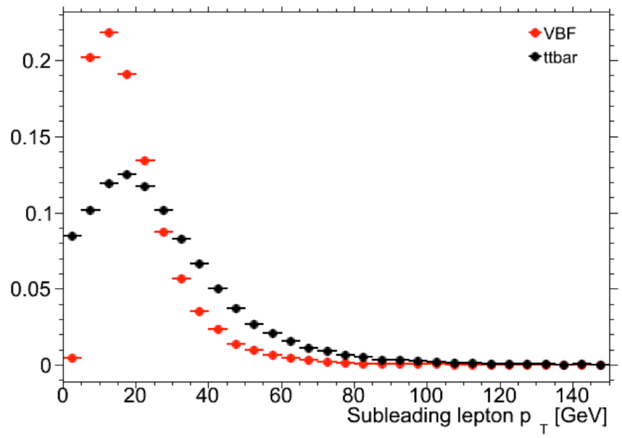


Figure 5.1: A comparison of the subleading lepton p_{T} spectrum between VBF $H \rightarrow WW^*$ production and $t\bar{t}$ background

70% for muons in the barrel region ($|\eta| < 1.05$) and 90% in the endcap region. The electron trigger efficiency increases with electron p_{T} but the average is approximately 90%. These efficiencies are measured by combined performance and trigger signature groups [66, 67].

	Level-1 threshold	High-level threshold
Electron	18	$24i$
	30	60
Muon	15	$24i$
		36

Table 5.1: Single lepton triggers used for electrons and muons. A logical "or" of the triggers listed for each lepton type is taken. Units are in GeV, and the i denotes an isolation requirement in the trigger.

	Level-1 threshold	High-level threshold
ee	10 and 10	12 and 12
$\mu\mu$	15	18 and 8
$e\mu$	10 and 6	12 and 8

Table 5.2: Di-lepton triggers used for different flavor combinations. The two thresholds listed refer to leading and sub-leading leptons, respectively. The di-muon trigger only requires a single lepton at level-1.

The combination of all triggers shown gives good efficiency for signal events. This efficiency is summarized in table 5.3. The relative improvement in efficiency by adding the dilepton triggers is also shown in the same table. The largest gain comes in the $\mu\mu$ channel. Overall the trigger selection shows a good efficiency for $H \to WW^*$ signal events.

Channel	Trigger efficiency	Gain from 2ℓ trigger
ee	97%	9.1%
$\mu\mu$	89%	18.5%
$e\mu$	95%	8.3%
μe	81%	8.2%

Table 5.3: Trigger efficiency for signal events and relative gain of adding a dilepton trigger on top of the single lepton trigger selection. The first lepton is the leading, while the second is the sub-leading. Efficiencies shown here are for the ggF signal in the $n_j = 0$ category but are comparable for the VBF signal.

5.2.2 Monte Carlo samples

Modeling of signal and background processes in the signal region, in particular for the m_{T} distribution, is an important consideration for the final interpretation of the analysis. Therefore, careful consideration must be paid to which Monte Carlo (MC) generators are used for specific processes. With the exception of the W+jet and multijet backgrounds, the m_{T} shape used as the final discriminant is taken from simulation. (Many backgrounds are normalized from data, as described in section 5.5).

Table 5.4 shows the MC generators used for the signal and background processes, as well as their cross sections. In order to include corrections up to next-to-leading order (NLO) in the QCD coupling constant α_s, the POWHEG [68] generator is often used. In some cases, only leading order generators like ACERMC [69] and GG2VV [70] are available for the process in question. If the process requires good modeling for very high parton multiplicities, the SHERPA [71] and ALPGEN [72] generators are used to provide merged calculations for five or fewer additional partons. These matrix element level calculations must then be additionally matched to models of the underlying event, hadronization, and parton shower. There are four possible generators for this: SHERPA , PYTHIA 6 [73], PYTHIA 8 [74], or HERWIG [75] + JIMMY [76]. The simulation additionally requires an input parton distribution function (PDF). The CT10 [77] PDFs are used for SHERPA and POWHEG simulated samples, while CTEQ6L1 [78] is used for ALPGEN +HERWIG and ACERMC simulations. The Drell-Yan samples are reweighted to the MRST [79] PDFs, as these are found to give the best agreement between data and simulation.

Process	MC generator	$\sigma \cdot \mathcal{B}$ (pb)
Signal		
ggF $H \to WW^*$	POWHEG +PYTHIA 8	0.435
VBF $H \to WW^*$	POWHEG +PYTHIA 8	0.0356
VH $H \to WW^*$	PYTHIA 8	0.0253
WW		
$q\bar{q} \to WW$ and $qg \to WW$	POWHEG +PYTHIA 6	5.68
$gg \to WW$	GG2VV +HERWIG	0.196
$(q\bar{q} \to W) + (q\bar{q} \to W)$	PYTHIA 8	0.480
$q\bar{q} \to WW$	SHERPA	5.68
VBS $WW + 2$ jets	SHERPA	0.0397
Top quarks		
$t\bar{t}$	POWHEG +PYTHIA 6	26.6
Wt	POWHEG +PYTHIA 6	2.35
$tq\bar{b}$	ACERMC +PYTHIA 6	28.4
$t\bar{b}$	POWHEG +PYTHIA 6	1.82
Other dibosons (VV)		
$W\gamma$ ($p_{\mathrm{T}}^{\gamma} > 8$ GeV)	ALPGEN +HERWIG	369
$W\gamma^*$ ($m_{\ell\ell} \leq 7$ GeV)	SHERPA	12.2
WZ ($m_{\ell\ell} > 7$ GeV)	POWHEG +PYTHIA 8	12.7
VBS $WZ + 2$ jets ($m_{\ell\ell} > 7$ GeV)	SHERPA	0.0126
$Z\gamma$ ($p_{\mathrm{T}}^{\gamma} > 8$ GeV)	SHERPA	163
$Z\gamma^*$ (min. $m_{\ell\ell} \leq 4$ GeV)	SHERPA	7.31
ZZ ($m_{\ell\ell} > 4$ GeV)	POWHEG +PYTHIA 8	0.733
$ZZ \to \ell\ell\,\nu\nu$ ($m_{\ell\ell} > 4$ GeV)	POWHEG +PYTHIA 8	0.504
Drell-Yan		
Z ($m_{\ell\ell} > 10$ GeV)	ALPGEN +HERWIG	16500
VBF $Z + 2$ jets ($m_{\ell\ell} > 7$ GeV)	SHERPA	5.36

Table 5.4: Monte Carlo samples used to model the signal and background processes [62].

Once the basic hard scattering process is simulated, it must be passed through a detector simulation and additional pile-up events must be overlaid. The pile-up events are modeled with PYTHIA 8, and the ATLAS detector is simulated with GEANT4 [80]. Because of the unique phase space of the $H \to WW^*$

analysis, events are sometimes filtered at generator level to allow for more efficient generation of relevant events. The efficiency of the trigger in MC simulation does not always match the measured efficiency in data, so trigger scale factors are applied to correct the MC efficiency to the data. These are derived by the combined performance groups [66, 67].

5.3 Object selection

In order to define the signal region, the analysis must first select the objects to be considered. The details of the object reconstruction algorithms are discussed in Chapter 2, while this section gives specific selection cuts used in the $H \to WW^*$ analysis.

The first step in this process is to select a primary vertex candidates. The event's primary vertex is the vertex with the largest sum of p_{T}^2 for associated tracks and is required to have at least three tracks with $p_{\mathrm{T}} > 450$ MeV. Many of the object selection cuts are then made relative to this chosen primary vertex.

5.3.1 Muons

The analysis uses combined muon candidates, where a track in the Inner Detector has been matched to a standalone track in the Muon Spectrometer. The track parameters are combined statistically in the muon reconstruction algorithm [53]. The muons are required to be within $|\eta| < 2.5$ and have a $p_{\mathrm{T}} > 10$ GeV. To reduce backgrounds coming from mis-reconstructed leptons, there are requirements on the impact parameter of the muon relative to the primary vertex. The transverse impact parameter d_0 is required to be small relative to its estimated uncertainty, the exact cut value being $d_0/\sigma_{d_0} < 3$. The longitudinal impact parameter z_0 must satisfy $|z_0 \sin\theta| < 1$ mm.

As discussed previously, the muons must also be isolated. There are two types of lepton isolations that are calculated: track-based and calorimeter-based. For muons, the track-based isolation is defined using the scalar sum $\sum p_{\mathrm{T}}$ for tracks with $p_{\mathrm{T}} > 1$ GeV (excluding the muon's track) within a cone of $\Delta R = 0.3$ (0.4) for muon with $p_{\mathrm{T}} > 15$ GeV ($10 < p_{\mathrm{T}} < 15$ GeV). The final isolation requirement is made my requiring that this scalar sum be no more than a certain fraction of the muon's p_{T}. This requirement varies with muon p_{T} and the exact cuts are defined in table 5.5.

The calorimeter-based muon isolation is defined using as a $\sum E_T$ calculated from calorimeter cells using the same cone size as the track-based isolation but excluding cells with $\Delta R < 0.05$ around the muon. This requirement is also defined as a cut on the ratio of the sum to the muon p_{T} and varies with muon p_{T}. The cut values are also given in table 5.5.

The isolation requirements loosen as a function of p_{T} to allow for larger signal acceptance. At low p_{T}, the isolation is tightened to reduce the W+jets background which arises from a misidentified lepton.

p_T range (GeV)	Calorimeter isolation	Track isolation
$10 - 15$	0.06	0.06
$15 - 20$	0.12	0.08
$20 - 25$	0.18	0.12
> 25	0.30	0.12

Table 5.5: p_{T} dependent isolation requirements for muons. Muons are required to have the amount of calorimeter or track based cone sums be less than this fraction of their p_{T}.

5.3.2 Electrons

Electrons are identified by matching reconstructed clusters in the electromagnetic calorimeter with tracks in the inner detector. The electrons are identified using a likelihood based method [50, 51] which takes into account the shower shapes in the calorimeter, the matching of tracks to clusters, and the amount of transition radiation in the TRT. The electrons are required to have $|\eta| < 2.47$, and candidates in the transition region between the barrel and endcap ($1.37 < |\eta| < 1.52$) are excluded. As the muons, the electrons are required to have transverse impact parameter significance < 3, while in the longitudinal direction they must have $|z_0 \sin\theta| < 0.4$ mm. Some electron requirements also vary with electron E_T, and these requirements are summarized in table 5.6.

The isolation for electrons are defined similarly to the muons but with unique cuts on the objects included. The track-based isolation is defined using tracks with $p_{\mathrm{T}} > 400$ MeV with cone sizes as defined previously. The calorimeter-based isolation also uses the same cone size as the muon, but here the cells within a 0.125×0.175 area in $\eta \times \phi$ around the electron cluster's barycenter are excluded. The other difference with respect to muons is that the denominator of the isolation ratio is the electron's E_T rather than p_{T}. The isolation cuts very with electron $E_T$and are defined in table 5.6.

The electron is also required to not be consistent with a vertex coming from a photon conversion.

p_T range (GeV)	Quality cut	Calorimeter isolation	Track isolation
$10-15$	Very tight LH	0.20	0.06
$15-20$	Very tight LH	0.24	0.08
$20-25$	Very tight LH	0.28	0.10
>25	Medium	0.28	0.10

Table 5.6: p_{T} dependent requirements for electrons. Electrons are required to have the amount of calorimeter or track based cone sums be less than this fraction of their E_T.

5.3.3 Jets

Jets are clustered with the anti-k_T reconstruction algorithm using a radius parameter of $R = 0.4$. They are required to have a jet vertex fraction (JVF) of at least 50%, meaning that half of the tracks associated with the jet originated from the primary vertex. Jets with no tracks associated (i.e. those outside the acceptance of the ID) do not have this requirement applied. Jets are required to have $p_T > 25$ GeV if they are within the tracking acceptance ($|\eta| < 2.4$). Jets with $2.4 < |\eta| < 4.5$ are required to have $p_{\mathrm{T}} > 30$ GeV. This tighter requirement reduces jets from pileup in the region where JVFrequirements cannot be applied. The two highest p_{T} jets in the event are referred to as the "VBF" jets and used to compute various analysis selections later.

Identification of b-jets is done using the MV1 algorithm and is limited to the acceptance of the ID ($|\eta| < 2.5$). The operating point of MV1 that is used is the one that is 85% efficient for identifying true b-jets. This operating point has a 10.3% of mis-tagging a light quark jet as a b-jet. In order to improve the rejection of b-jets, a lower threshold than the nominal p_{T} threshold described above is used. For the purposes of counting the number of b-jets, jets with p_{T} down to 20 GeVare used.

5.3.4 Overlap removal

There are some cases where certain reconstructed objects will overlap and one will have to be chosen (for example, an electron and a jet in the calorimeter). First, the case of lepton overlap is dealt with. If an electron candidate extends into the muon spectrometer, it is removed. If a muon or electron have a $\Delta R < 0.1$, the electron is removed and the muon is kept. If two electron candidates overlap within the same radius, then the higher E_T electron is kept. Next, the overlap between leptons and jets is considered. If an electron and jet are within $\Delta R < 0.3$ of one another, the electron is kept and the jet is removed. However, if a muon and jet overlap within $\Delta R < 0.3$, the jet is kept (as it is likely that the muon is the result of a semileptonic decay inside the jet).

Once the overlap removal is complete, the final set of objects used in the analysis is defined.

5.4 Analysis selection

The VBF analysis uses two distinct selections. The first is a more standard selection, referred to as "cut-based", that applies requirements on the VBF variables and uses m_{T} as the final discriminating variable. The second is a looser selection that uses a Boosted Decision Tree (BDT) score as the final discriminator in order to take advantage of the detailed correlations between the VBF variables. While the BDT analysis is ultimately more sensitive, the cut-based serves as an important component of the analysis. First, the cut-based allows for confirming the modeling and validity of many variables used as input to the BDT. Second, because this is the first use of such an MVA technique in the $H \to WW^*$ analysis, the cut-based selection allows confirmation of the final BDT result with a more traditional analysis. The cut-based techniques are the focus of this chapter, but connections to the BDT result will be illustrated when appropriate.

One important note is that because this analysis is dedicated to the measurement of the VBF production mode of the Higgs, events coming from gluon fusion production with the Higgs decaying via $H \to WW^* \to \ell\nu\ell\nu$ are treated as background events. This will be seen throughout the various predictions shown.

5.4.1 Common pre-selection

Both the cut-based and BDT analyses have a common pre-selection that is applied before the main signal region requirements. The requirements on leptons are common to all n_j bins. The analysis requires two oppositely charged leptons, with the leading lepton required to have $p_{\mathrm{T}} > 22$ GeVwhile the subleading lepton must have $p_{\mathrm{T}} > 10$ GeV. Next, to remove low mass Z/γ^* events, a cut on the dilepton mass $m_{\ell\ell} > 10$ (12) GeVis applied in the different (same) flavor channel. In the same flavor channels, there is an additional veto placed on the region around the Z peak, requiring that $|m_{\ell\ell} - m_Z| > 15$ GeV.

There are also requirements on the amount of missing transverse momentum in the event. These are only applied in the same flavor channels, as in the different flavor channels $t\bar{t}$ is the dominant background in $n_j \geq 2$. The BDT analysis requires $p_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$. The cut-based analysis must select more tightly on these variables to have maximal sensitivity and thus requires $p_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{\mathrm{miss}} > 55\,\mathrm{GeV}$.

Finally, because this analysis is focused on VBF, a requirement on the jet multiplicity is placed, with $n_j \geq 2$. Additionally, the analysis requires that there are no jets identified as b-quarks in the event, or $n_b = 0$.

5.4.2 Cut-based selection

The cut-based selection places sequential requirements on variables reconstructed from the VBF jets in order to increase the signal to background ratio.

General background reduction

Top pair production is the primary background in the $n_j \geq 2$ bin. Even though $n_b = 0$ is required, an additional variable is constructed to further suppress the top background. There is often additional QCD radiation that accompanies the $t\bar{t}$ system when it is produced. Therefore, a variable which tests for the presence of this additional radiation, $p_{\mathrm{T}}^{\mathrm{sum}}$, is constructed. It is defined in equation 5.1.

$$\boldsymbol{p}_{\mathrm{T}}^{\mathrm{sum}} = \boldsymbol{p}_{\mathrm{T}}^{\ell\ell} + \boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} + \sum \boldsymbol{p}_{\mathrm{T}}^{j} \tag{5.1}$$

The first cut after pre-selection in the cut-based analysis requires $p_{\mathrm{T}}^{\mathrm{sum}} < 15\,\mathrm{GeV}$ to further suppress $t\bar{t}$ production.

In the different flavor channels, a cut is made to reduce the contamination from $Z \to \tau\tau$ decays. The di-τ invariant mass, $m_{\tau\tau}$. is constructed by assuming that the neutrinos from the τ decays were collinear with the leptons [81]. The analysis requires that this mass not be consistent with a Z by requiring $m_{\tau\tau} < m_Z - 25$ GeV.

VBF topological cuts

The characteristic feature of VBF production of the Higgs is the presence of two additional forward jets coming from the incoming partons which radiate the vector bosons that make the Higgs. These jets are forward because the outgoing partons still carry the longitudinal momentum of the incoming partons. Figure 5.2 shows the distribution of the η for the leading jet in a VBF event compared to a background top pair production event. As can be seen, the VBF jets tend to be more forward in η, while the $t\bar{t}$ jets are more central.

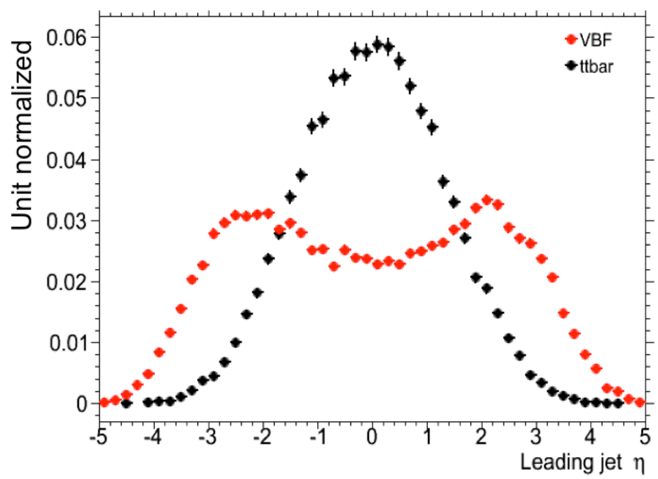


Figure 5.2: Leading jet η in VBF $H \to WW^*$ (red) and $t\bar{t}$ (black)

Because the cross section for VBF production is an order of magnitude smaller than gluon fusion production, these forward jets must be used in order to better reduce background and achieve a good signal to background ratio. The dedicated VBF search selection requirements are constructed to maximally exploit the features of the unique VBF topology.

Requirements on the VBF jets are collectively referred to as the "VBF topological cuts". First, a requirement on the dijet invariant mass of the VBF jets, m_{jj}, is placed, requiring $m_{jj} > 600$ GeV. Next, the event is required to have a large gap in rapidity between the two VBF jets, or $\Delta y_{jj} > 3.6$. Both of these cuts put tight requirements on the presence of two forward, high p_{T} jets moving in opposite directions in the longitudinal plane.

Beyond requiring the presence of the two forward VBF jets, the analysis also vetoes on the presence of any additional jets that fall between the two VBF jets. This cut is referred to as the central jet veto, or CJV. Any events with a third jet with $p_{\mathrm{T}} > 20$ GeV whose rapidity is between the region defined by the two VBF jets are vetoed. This can be expressed in terms of a variable called the jet centrality, defined in equation 5.2.

$$C_{j3} = \left| \eta_{j3} - \frac{\eta_{j1} + \eta_{j2}}{2} \right| \Big/ \frac{|\eta_{j1} - \eta_{j2}|}{2}, \tag{5.2}$$

Here, η_{j1} and η_{j2} are the pseudorapidities of the leading and subleading jets, respectively, while η_{j3} is the pseudorapidity of the extra jet in the event (if one exists). Intuitively, C_{j3} is zero when η_{j3} is directly centered between the two jets and unity when η_{j3} is aligned with either of the VBF jets. Thus, the CJV can be expressed as a requirement that $C_{j3} > 1$.

The decay products of the Higgs tend to be central as well. Thus, the analysis also requires that both leptons in the analysis fall within the rapidity gap defined by the jets. This cut is referred to as the outside lepton veto, or OLV. A quantitative way to define the cut is to require that the centrality of each lepton (defined analogously to that of the third jet in equation 5.2) correspond to the lepton being within the jet rapidity gap, or $C_{\ell} < 1$ for both leptons.

Figure 5.3a-c shows the m_{jj}, Δy_{jj}, and $C_{\ell 1}$ variables at the stage where all previous cuts in the sequence have been made. The agreement between data and Monte Carlo is good, and the bottom panels show their power in discriminating the VBF signal from the background processes.

The final signal region is also split into two bins of m_{jj}, with the first bin corresponding to 600 GeV $<$ $m_{jj} < 1$ TeV and the second bin corresponding to $m_{jj} > 1$ TeV. The first bin has more statistics but also a larger contribution from background, while the second bin has lower statistics but a 1:1 signal to background ratio.

Higgs topological cuts

The final state leptons will exhibit unique correlations due to the fact that they are arising from the decay of a spin zero resonance. In particular, the spins of the final state leptons and neutrinos must all

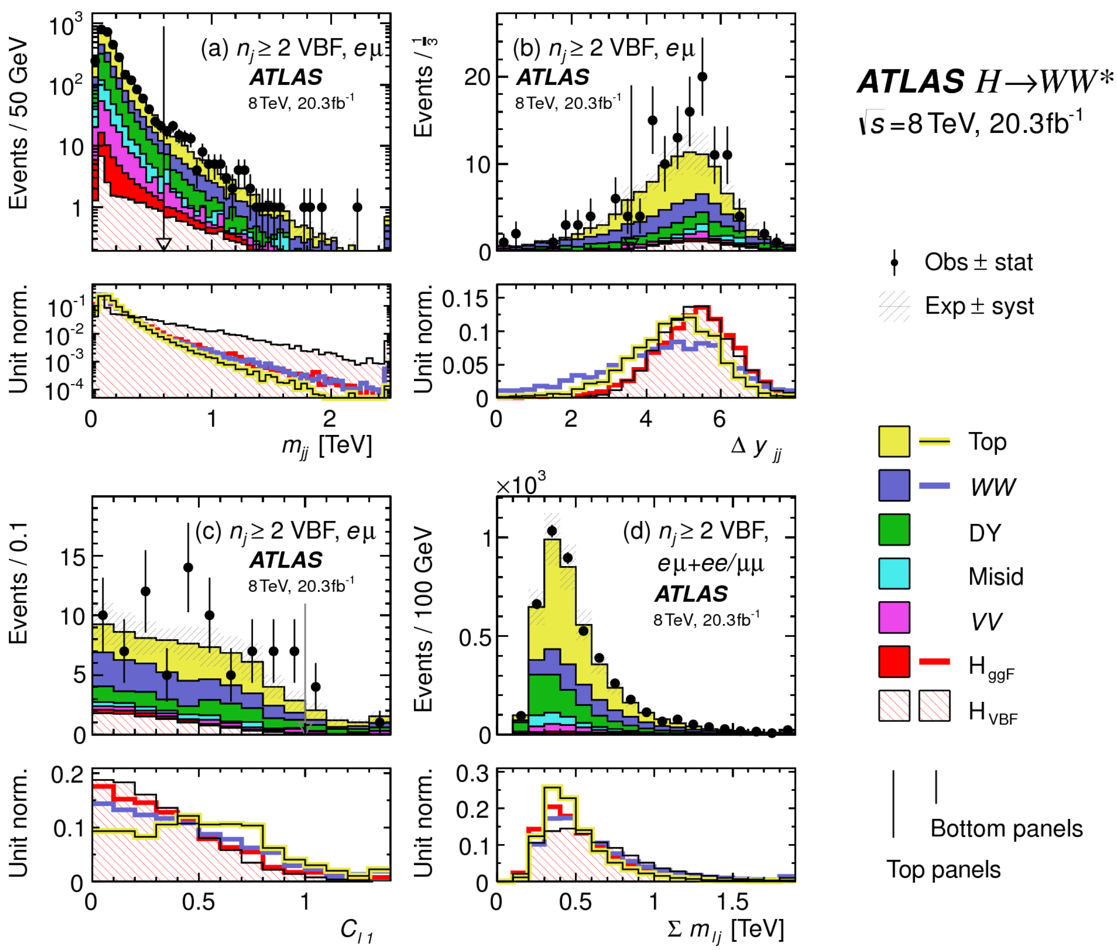


Figure 5.3: Distributions of (a) m_{jj}, (b) Δy_{jj}, (c) $C_{\ell 1}$, and (d) $\Sigma\, m_{\ell j}$, for the VBF analysis. The top panels compare simulation and data, while the bottom panels show normalized distributions for all background processes and signal [62].

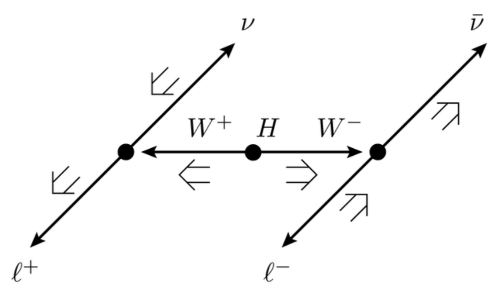


Figure 5.4: A cartoon of the WW final state. Momenta are represented with thin arrows, spins with thick arrows. [62]

cancel, as shown in figure 5.4. Because the neutrino has a left handed chirality and the anti-neutrino has a right handed chirality (in the massless neutrino approximation), the spin and momentum of the particles will be anti-aligned and aligned, respectively. In the transverse plane, the momenta of all four final state objects must cancel as well. With the constraint of having both the momenta and the spin alignments cancel, the final state kinematics strongly prefer having a small angle between the leptons in the transverse plane (low $\Delta\phi_{\ell\ell}$). This angular correlation will also lead to low values of the di-lepton invariant mass $m_{\ell\ell}$. These unique signal final state kinematic correlations will be exploited to define the ultimate signal region.

The analysis places additional requirements on the final state leptons. Two requirements on dilepton kinematics are made that are common with lower multiplicity jet bins as well. The angle between leptons in the transverse plane, $\Delta\phi_{\ell\ell}$, is required to be less than 1.8 radians. Additionally, the dilepton mass $m_{\ell\ell}$ is required to be less than 50 GeV.

The cut-based analysis uses m_{T} as the final discriminating variable as in the ggF focused analysis. The optimal number of bins in m_{T} was found to be three bins, with the bin boundaries at 80 and 130 GeV.

Table 5.7 shows a summary of the data and estimated signal and background yields from simulation as each requirement described above is made. The table shows how the overall signal to background ratio grows through the various selection requirements. Table 5.8 shows the background composition after each selection requirement, illustrating which backgrounds are reduced most by certain requirements. Figure 5.5 shows an ATLAS event display of a candidate event in the final signal region.

Selection	Summary					
	$N_{\text{obs}}/N_{\text{bkg}}$	N_{obs}	N_{bkg}	N_{signal}		
				N_{ggF}	N_{VBF}	N_{VH}
$e\mu$ sample	1.00 ± 0.00	61434	61180	85	32	26
$n_b = 0$	1.02 ± 0.01	7818	7700	63	26	16
$p_{\text{T}}^{\text{sum}} < 15$	1.03 ± 0.01	5787	5630	46	23	13
$m_{\tau\tau} < m_Z - 25$	1.05 ± 0.02	3129	2970	40	20	9.9
$m_{jj} > 600$	1.31 ± 0.12	131	100	2.3	8.2	–
$\Delta y_{jj} > 3.6$	1.33 ± 0.13	107	80	2.1	7.9	–
$C_{j3} > 1$	1.36 ± 0.18	58	43	1.3	6.6	–
$C_{\ell 1} < 1, C_{\ell 2} < 1$	1.42 ± 0.20	51	36	1.2	6.4	–
$m_{\ell\ell}, \Delta\phi_{\ell\ell}, m_{\text{T}}$	2.53 ± 0.71	14	5.5	0.8	4.7	–
$ee/\mu\mu$ sample	0.99 ± 0.01	26949	27190	31	14	10.1
$n_b, p_{\text{T}}^{\text{sum}}, m_{\tau\tau}$	1.03 ± 0.03	1344	1310	13	8.0	4.0
$m_{jj}, \Delta y_{jj}, C_{j3}, C_{\ell}$	1.39 ± 0.28	26	19	0.4	2.9	0.0
$m_{\ell\ell}, \Delta\phi_{\ell\ell}, m_{\text{T}}$	1.63 ± 0.69	6	3.7	0.3	2.2	0.0

Table 5.7: Summary of event selection for the $n_j \geq 2$ VBF analysis in the 8 TeV cut-based analysis [62].

	Composition of N_{bkg}									
	N_{WW}		N_{top}		N_{misid}		N_{VV}	$N_{\text{Drell-Yan}}$		
	N_{WW}^{QCD}	N_{WW}^{EW}	$N_{t\bar{t}}$	N_t	N_{Wj}	N_{jj}	N_{VV}	$N_{ee/\mu\mu}$	$N_{\tau\tau}^{\text{QCD}}$	$N_{\tau\tau}^{\text{EW}}$
$e\mu$ sample	1350	68	51810	2970	847	308	380	51	3260	46
$n_b = 0$	993	43	3000	367	313	193	273	35	2400	29
$p_{\text{T}}^{\text{sum}} < 15$	781	38	1910	270	216	107	201	27	2010	23
$m_{\tau\tau} < m_Z - 25$	484	22	1270	177	141	66	132	7.6	627	5.8
$m_{jj} > 600$	18	8.9	40	5.3	1.8	2.4	5.1	0.1	15	1.0
$\Delta y_{jj} > 3.6$	11.7	6.9	35	5.0	1.6	2.3	3.3	–	11.6	0.8
$C_{j3} > 1$	6.9	5.6	14	3.0	1.3	1.3	2.0	–	6.8	0.6
$C_{\ell 1} < 1, C_{\ell 2} < 1$	5.9	5.2	10.8	2.5	1.3	1.3	1.6	–	5.7	0.6
$m_{\ell\ell}, \Delta\phi_{\ell\ell}, m_{\text{T}}$	1.0	0.5	1.1	0.3	0.3	0.3	0.6	–	0.5	0.2
$ee/\mu\mu$ sample	594	37	23440	1320	230	8.6	137	690	679	16
$n_b, p_{\text{T}}^{\text{sum}}, m_{\tau\tau}$	229	12.0	633	86	26	0.9	45	187	76	1.5
$m_{jj}, \Delta y_{jj}, C_{j3}, C_{\ell}$	3.1	3.1	5.5	1.0	0.2	0.0	0.7	3.8	0.7	0.1
$m_{\ell\ell}, \Delta\phi_{\ell\ell}, m_{\text{T}}$	0.4	0.2	0.6	0.2	0.2	0.0	0.1	1.5	0.3	0.1

Table 5.8: Background composition after each requirement in the $n_j \geq 2$ VBF analysis in the 8 TeV cut-based analysis [62].

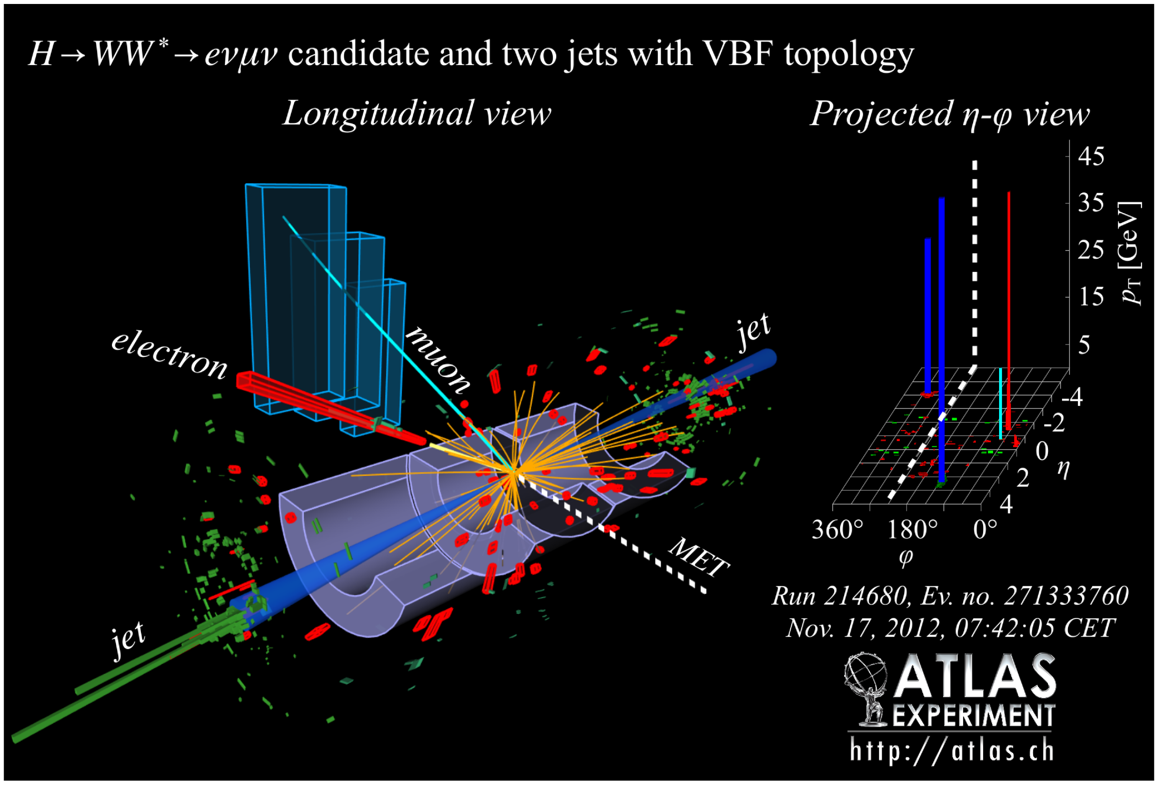


Figure 5.5: Event display of a VBF candidate event [62].

5.4.3 BDT-based selection

The boosted decision tree based analysis takes a different philosophy compared to the cut-based. Rather than cutting sequentially on many variables, the BDT analysis uses many of these variables as inputs to the BDT and the output BDT score (O_{BDT}) as the final discriminant. The BDT is trained with the VBF $H \rightarrow WW^*$ simulation as the signal samples and all other processes as background, including ggF $H \rightarrow WW^*$ production. While the BDT based analysis is treated as a separate result, it has significant overlap with the cut-based selection.

Pre-training selection and BDT inputs

Before training, the common pre-selection cuts described in section 5.4.1 are applied. Additionally, the central jet veto and outside lepton veto described in section 5.4.2 are applied. The BDT has eight input variables, six of which are also variables that are used in the cut-based analysis. The six shared variables are $p_{\mathrm{T}}^{\mathrm{sum}}$, m_{jj}, Δy_{jj}, $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, and m_{T}. The seventh variable input in the BDT is a combination of the variables used to do the OLV in the cut-based analysis. The BDT uses as input the sum of lepton centralities, or $\sum C_\ell = C_{\ell 1} + C_{\ell 2}$. The final BDT input variable, $\Sigma\, m_{\ell j}$, is constructed to account for the correlations between the jets and leptons in the event. It is the sum of the invariant masses of all four possible lepton-jet combinations.

Figure 5.3d shows the agreement between data and simulation for the $\Sigma\, m_{\ell j}$ variable, as well as showing its discriminating power. Figure 5.6 shows the distributions of the Higgs topological variables that are shared between the cut-based and BDT analyses. Figure 5.7 shows the distributions of the VBF topological variables shared between the cut-based and BDT analyses. In both cases, the VBF yield has been scaled by a factor of 50 to better show the shape difference compared to the backgrounds.

Table ?? summarizes the cuts applied for the cut-based and analyses, as well as which variables are used as input to the BDT.

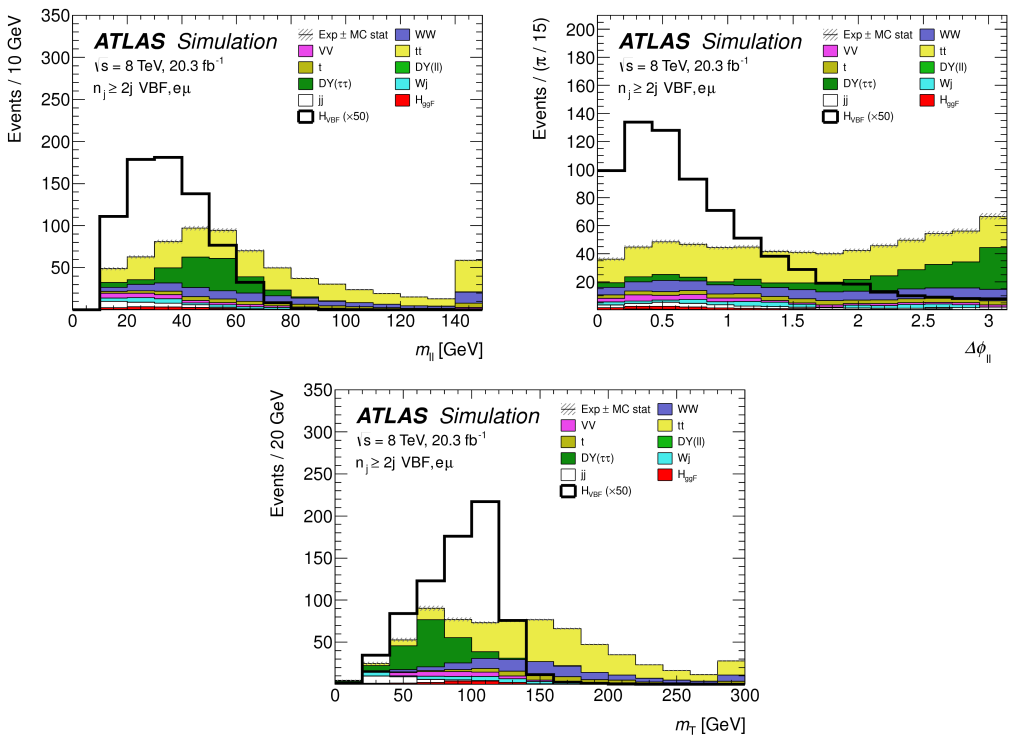


Figure 5.6: Distributions of $m_{\ell\ell}$ (top left), $\Delta\phi_{\ell\ell}$ (top right), and m_{T} (bottom), Higgs topology variables used in the selection requirements of the cut-based signal region and as inputs to the BDT result. These are plotted after all of the BDT pre-training selection cuts [62].

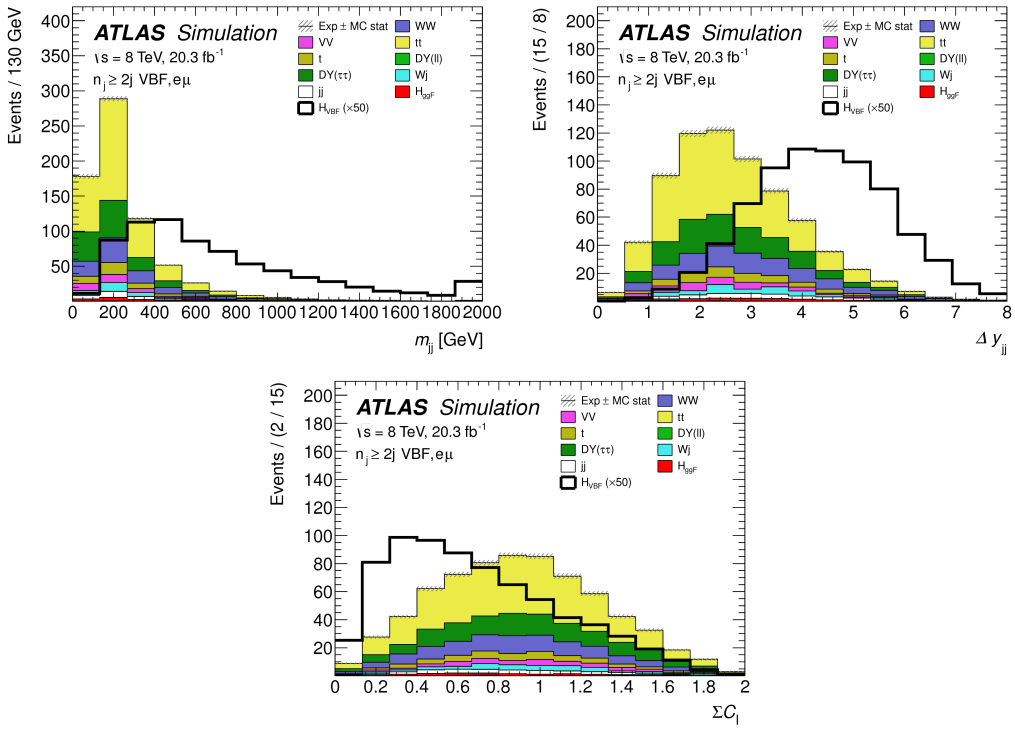


Figure 5.7: Distributions of m_{jj} (top left), Δy_{jj} (top right), $\sum C_\ell$ (bottom), VBF topology variables used in the selection requirements of the cut-based signal region and as inputs to the BDT result. These are plotted after all of the BDT pre-training selection cuts [62].

5.5 Background estimation

This section describes the procedures used to estimate backgrounds for the VBF analysis in both the cut-based and BDT analyses.

5.5.1 General strategy

Most of the backgrounds in the VBf analysis have shapes estimated from Monte Carlo simulation but normalizations derived from control regions in data. In essence, a normalization factor (denoted with β or abbreviated as NF) is derived by scaling the MC yield in the control region to the corresponding yield in data. Once this factor is derived, it can be used to scale the MC estimate of the background in the signal region. This is illustrated in equation 5.3.

$$B_{\text{SR}}^{\text{est}} = B_{\text{SR}} \times \frac{N_{\text{CR}}}{B_{\text{CR}}} \equiv B_{\text{SR}} \times \beta \tag{5.3}$$

Here, B denotes the MC yield prediction in the denoted region, while N denotes the observed number of events in data in the denoted region.

Another way of writing the same equation, in terms of an extrapolation factor α rather than a normalization factor β. The overall calculation is exactly the same. However, when phrased in this way, it shows hwo the uncertainty on the background estimation can be reduced. This is shown in equation 5.4.

$$B_{\text{SR}}^{\text{est}} = N_{\text{CR}} \times \frac{B_{\text{SR}}}{B_{\text{CR}}} \equiv N_{\text{CR}} \times \alpha \tag{5.4}$$

Phrased this way, the equation shows that with enough statistics in the control region, a large theoretical uncertainty on the overall background yield in the signal region can be replaced by a small statistical uncertainty coming from the number of data events in the CR and a smaller theoretical uncertainty on the extrapolation from the control region to the signal region.

5.5.2 Top background

The normalization factor β_t for the top background in the VBF analysis is derived in a region required to have one b-tagged jet, or $n_b = 1$. In the cut-based analysis, normalization factors are computed at every stage of the cutflow by applying the appropriate cuts in the CR. These NF are then applied to the $t\bar{t}$ and single top event yields in the SR. In the BDT analysis, a single normalization factor is computed for each bin of O_{BDT} after applying the BDT pre-training cuts described previously. The computed normalization factors are derived with all flavor combinations combined in order to decrease statistical uncertainty. Additionally, in the BDT analysis, BDT bins 2 and 3 are merged for the same reason.

Table 5.9 shows the evolution of the β_t through the cut-based selection. Table 5.10 shows the value of the β_t in each bin of O_{BDT}. In all cases, the computed factors are relatively consistent with unity, with the largest discrepancy coming in bin 1 of O_{BDT}. The normalization factors in the bins of O_{BDT} are also consistent with those derived in teh cut-based sisgnal region, increasing confidence in the BDT estimation.

Cut	β_t
$p_{\mathrm{T}}^{\mathrm{sum}} < 15\,\mathrm{GeV}$	1.03 ± 0.01
$m_{\tau\tau} < m_Z - 25$	1.05 ± 0.01
$m_{jj} > 600\,\mathrm{GeV}$	0.96 ± 0.06
$\Delta y_{jj} > 3.6$	1.02 ± 0.08
CJV	1.13 ± 0.16
OLV	1.01 ± 0.19
$m_{jj} < 1\,\mathrm{TeV}$	0.94 ± 0.19
$m_{jj} > 1\,\mathrm{TeV}$	1.48 ± 0.66

Table 5.9: Top normalization factors computed at each stage of the cut-based selection. Uncertainties are statistical only.

O_{BDT}	β_t
Bin0	1.09 ± 0.02
Bin1	1.58 ± 0.15
Bin2	0.95 ± 0.31
Bin3	0.95 ± 0.31

Table 5.10: Top normalization factors computed for each bin of O_{BDT}. Uncertainties are statistical only.

Figure 5.8 shows the m_{jj} and O_{BDT} distributions in the top control region. Overall the modeling looks consistent with the data.

While these normalization factors can be computed and applied to the expected background yields listed in tables like table ??, in the end the normalization of the top background is profiled (meaning there is a dedicated Poisson constraint) and allowed to float in the final statistical fit.

5.5.3 $Z/\gamma^* \rightarrow \tau\tau$ background

In the different flavor channels, the $Z/\gamma^* \rightarrow \tau\tau$ background is an important one. Di-tau production can produce an $e\mu$ final state if each τ lepton decays to a different flavor lepton.

In the BDT analysis, a single normalization factor for the background is derived. A control region is defined using the pre-training selection cuts, except requiring that $|m_{\tau\tau} - m_Z| < 25\,\mathrm{GeV}$ so that the region is enriched in $Z/\gamma^* \rightarrow \tau\tau$ background. Additional requirements of $m_{\ell\ell} < 80(75)\,\mathrm{GeV}$

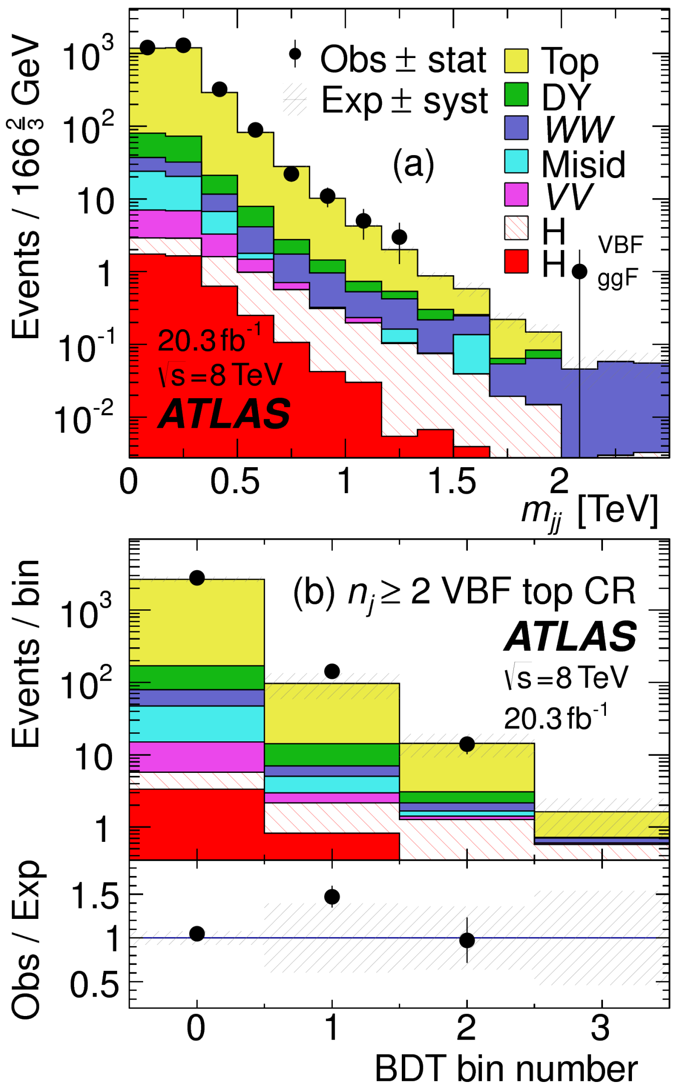


Figure 5.8: Distributions of m_{jj} (a) and O_{BDT} (b) in the VBF $n_b = 1$ top CR [62].

in the different (same) flavor channel, as well as $O_{\text{BDT}} > -0.48$ are applied to increase the purity of the region. The final $\beta_{Z/\gamma^* \to \tau\tau}$ is calculated to be 0.9 ± 0.3 (statistical uncertainty only). Because of the small contribution of this background in the BDT analysis and the large statistical uncertainty, no additional systematics are calculated. The final SR estimate is scaled by this β and not allowed to float in the fit.

The cut-based corrections are a bit more involved because they need to be applied selection by selection, as well as in the final signal region for the fit. The region is defined including all SR cuts up to the $Z/\gamma^* \to \tau\tau$ veto, which is instead made into a Z mass peak requirement as for the BDT region. The $m_{\ell\ell}$ cut from the BDT region is included as well. The cut-based approach aims to correct the normalization of the $Z/\gamma^* \to \tau\tau$ background in two ways. First, an overall normalization factor is computed from the control region. However, the VBF topological cuts are not included in this region, and applying them as is done in the top CR is not feasible due to limited statistics. So, instead, correction factors (CF) to the cut efficiencies of the VBF cuts are derived in a same flavor $Z \to \ell\ell$ control region, which has significantly more statistics. The CF is simply the ratio of the cut efficiencies in data and MC derived in this region. In the end, the overall background estimate is given by equation 5.5.

$$N^{\mathrm{est}}_{Z/\gamma^* \to \tau\tau} = B^{\mathrm{SR}}_{Z/\gamma^* \to \tau\tau} \times \beta_{\tau\tau} \times \frac{\epsilon^{\mathrm{data}}_{\mathrm{VBF\,cuts}}}{\epsilon^{\mathrm{MC}}_{\mathrm{VBF\,cuts}}} \tag{5.5}$$

The hypothesis is that while the normalization correction must be derived in a dedicated region, the efficiency of the VBF cuts should not be sensitive to the type of Z/γ^* process and thus the larger control region can be exploited to derive the CF. Figure 5.9 shows a shape comparison for the m_{jj} variable in $Z \to \tau\tau$ events in the signal region and $Z \to \ell\ell$ events in the control region. The figure shows that the shapes are indeed comparable and thus any CF derived in the same flavor control region can reliably be applied in the signal region.

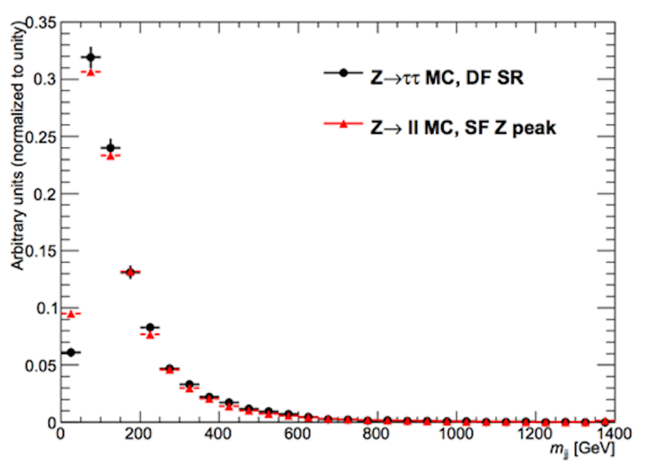


Figure 5.9: Comparison of m_{jj} shape in a same flavor $Z \to \ell\ell$ control region and the VBF cut-based signal region.

Table 5.11 shows the overall normalization factor $\beta_{\tau\tau}$ and the efficiency correction factors for the various VBF topological cuts. In general, the statistical uncertainties on the cut efficiency corrections are quite good, and the MC tends to underestimate the efficiency of the VBF cuts for the $Z/\gamma^* \to \tau\tau$ background. The overall normalization factor is also consistent with that calculated for the BDT analysis.

$\beta_{\tau\tau}$	0.97 ± 0.04
Cut	Correction factors
$m_{jj} > 600\,\text{GeV}$	1.09 ± 0.01
$\Delta y_{jj} > 3.6$	1.14 ± 0.02
CJV	1.20 ± 0.02
OLV	1.17 ± 0.03
$m_{jj} < 1\,\text{TeV}$	1.17 ± 0.06
$m_{jj} > 1\,\text{TeV}$	1.18 ± 0.13

Table 5.11: $Z/\gamma^* \to \tau\tau$ correction factors for the VBF cut-based analysis. Uncertainties are statistical only.

5.5.4 $Z/\gamma^* \to \ell\ell$ background

In the same flavor channels, the $Z/\gamma^* \to \ell\ell$ background is dominant and thus must be estimated correctly. In both the BDT and cut-based analyses, the background is estimated using the so-called "ABCD" method. The ABCD method creates four different regions by defining cuts on two variables. One of the regions (A) is the signal region, while the other regions are defined by inverting one of both of the cuts. in this case, the two variables used are $m_{\ell\ell}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$, because inverting either of the SR cuts on these variables will give regions rich in the $Z/\gamma^* \to \ell\ell$ background. Figure 5.10 illustrates the general strategy for each region.

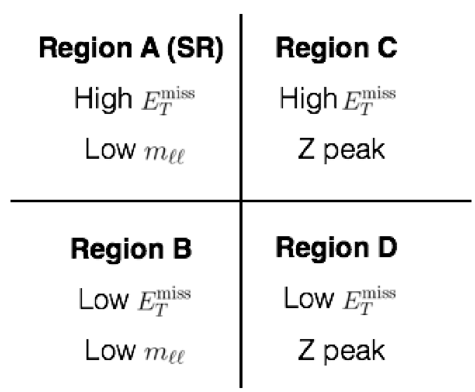


Figure 5.10: General illustration of the ABCD region definitions for $Z/\gamma^* \to \ell\ell$ background estimation.

In both of the cut-based and BDT analyses, the Z peak region is defined with $|m_{\ell\ell} - m_Z| < 15$ GeV.

In the cut-based analysis, low $m_{\ell\ell}$ corresponds to $m_{\ell\ell} < 50\,\mathrm{GeV}$ (this defines the cut-based SR) while in the BDT it is $m_{\ell\ell} < 75\,\mathrm{GeV}$. In the cut-based, high and low $E_{\mathrm{T}}^{\mathrm{miss}}$ are defined as opposite ends of the $55\,\mathrm{GeV}$ cut applied for the signal region definition. The BDT low $E_{\mathrm{T}}^{\mathrm{miss}}$ region is between 25 and $45\,\mathrm{GeV}$, while the high $E_{\mathrm{T}}^{\mathrm{miss}}$ region is $E_{\mathrm{T}}^{\mathrm{miss}} > 45\,\mathrm{GeV}$.

Once the regions are defined, the final signal region background estimate is done by taking the estimate in region B and extrapolating it to the signal region (A) by multiplying it by the ratio of regions C and D. Effectively, the Z peak region is used to estimate the efficiency of the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut in data, and then this efficiency is applied in the low $m_{\ell\ell}$ region. An additional correction is also applied for the non-closure of the method in MC. This is summarized in equations 5.6 and 5.7.

$$N^{\mathrm{SR}}_{Z/\gamma^*\to\ell\ell} = N^{\mathrm{B}}_{Z/\gamma^*\to\ell\ell} \times \frac{N^{\mathrm{C}}_{Z/\gamma^*\to\ell\ell}}{N^{\mathrm{D}}_{Z/\gamma^*\to\ell\ell}} \times f_{\mathrm{corr}} \tag{5.6}$$

$$f_{\mathrm{corr}} = \frac{B^{\mathrm{A}}_{\mathrm{MC}}/B^{\mathrm{B}}_{\mathrm{MC}}}{B^{\mathrm{C}}_{\mathrm{MC}}/B^{\mathrm{D}}_{\mathrm{MC}}} \tag{5.7}$$

Here, the N refer to data yields in each region with the non Z/γ^* backgrounds subtracted, while B refer to the Z/γ^* yields in MC in each region.

A normalization factor $\beta_{\ell\ell}$ is computed for each analysis as the ratio of the predicted data yield to the MC yield in the SR. The shape of the BDT distribution is taken from data region B, while the shape of the m_{T} distribution in the cut-based analysis is taken from Z/γ^* MC in the SR. The values of the $\beta_{\ell\ell}$ in the cut-based and BDT analyses from this method are summarized in table 5.12. They are quite consistent with one another within the statistical uncertainties. In the cut-based analysis, the same cut efficiency correction factors shown in table 5.11 are also applied (in product with the $\beta_{\ell\ell}$) in the same flavor channels to this background, as they were derived in the Z peak region.

	β_t
BDT Bin 1	1.01 ± 0.15
BDT Bin 2	0.89 ± 0.28
Cut-based	0.81 ± 0.21

Table 5.12: $Z/\gamma^* \to \ell\ell$ normalization factors for cut-based and BDT analyses. Uncertainties are statistical only.

5.5.5 WW and other diboson backgrounds

The WW and other diboson backgrounds have both their shape and normalization taken from MC simulation. They are validated in dedicated control regions and found to agree with data well.

As WW is the largest of these backgrounds and is irreducible, validating the estimate is of particular importance. The validation region is constructed by requiring the pre-selection cuts on leptons and $m_{\ell\ell}$, $n_b = 0$, and $m_{\mathrm{T}} > 100\,\mathrm{GeV}$. The m_{T2} variable [82] is an additional discriminant that will isolate the WW background, and a requirement of $m_{\mathrm{T2}} > 160\,\mathrm{GeV}$ is placed to define the WW validation region. This cut gives a 60% purity for the validation region. The derived normalization factor in the region is 1.15 ± 0.19 and is thus consistent with unity. Figure 5.11 shows the m_{T2} distribution and how it distinguishes the WW background.

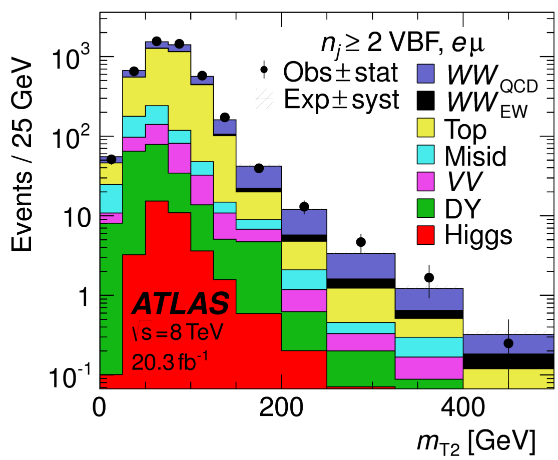


Figure 5.11: Distribution of m_{T2} in the WW validation region of the VBF analysis [62].

5.5.6 Higgs production via gluon-gluon fusion

Because this analysis is dedicated to measuring the VBF contribution to Higgs production, the component of Higgs production from gluon-gluon fusion is treated as a background. The shape is taken directly from simulation, using the generators described in table 5.4. In the final combined fit of all different signal regions, the normalization is controlled by either a combined signal strength parameter μ, which controls the normalization of both ggF and VBF production, or a separate parameter μ_{ggF} depending on the interpretation being presented in the final results.

5.5.7 Backgrounds with misidentified leptons

As discussed previously, the W+jets and QCD multijet backgrounds are derived with fully data-driven methods. These backgrounds do not make a large contribution to the final VBF signal region but their estimation methods are discussed briefly here.

W+jets background

The W+jets background enters the signal region by having one of the jets mis-reconstructed as a lepton. The background is estimated by constructing a control sample with two leptons, where one lepton passes the usual lepton quality cuts but the second lepton fails one of those cuts (also knows as the "anti-identified" lepton). This control region is rich in the W+jets contribution because if a second lepton is reconstructed in a W+jets event it is likely to be poor quality. The purity of this W+jets control sample is 85% to 90% depending on the exact configuration of leptons in the final state.

The signal region estimate of W+jets is estimated by extrapolation from the control sample to the signal region using extrapolation factors derived in a Z+jets control sample in data. The extrapolation factor is the ratio of the number of lepton candidates satisfying all quality criteria to the number of lepton candidates anti-identified. This ratio is measured in bins of p_{T} and η. Thus, the final signal region estimate (binned as the extrapolation factor is binned) is simply the number of events in the anti-identified lepton control sample multiplied by the extrapolation factor derived from the Z+jets control sample. Figure 5.12 shows the extrapolation factors derived for electrons and muons.

QCD multijet background

The method for estimating the multijet background is very similar to the W+jets estimation method. The control sample in this case has two anti-identified leptons but otherwise satisfies all signal region requirements. The extrapolation factor is estimated from a multijet sample and applied twice to the control sample.

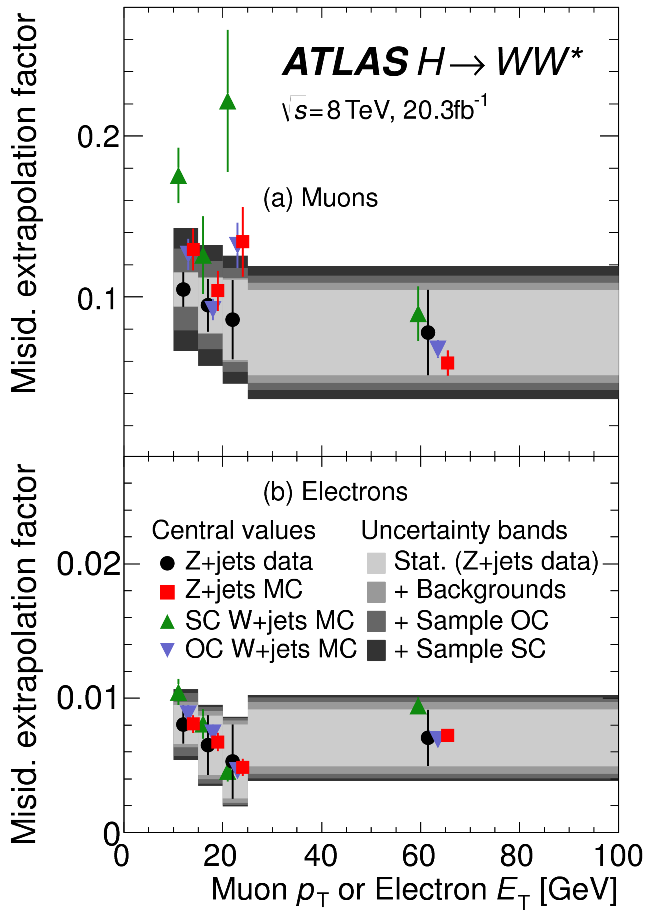


Figure 5.12: Extrapolation factors for the W+jets estimate derived for muons (a) and electrons (b) as a function of lepton p_T [62].

5.5.8 Background composition in final signal region

After all of these estimation procedures, the final signal region composition can be calculated. The estimated yields are all shown in table ??. Figure 5.13 shows the relative percentages of the different background for the different flavor and same flavor final states. In $e\mu$, the leading backgrounds are top backgrounds, ggF Higgs, and SM WW production. In $ee/\mu\mu$, the leading background is Drell-Yan, followed by top and ggF Higgs.

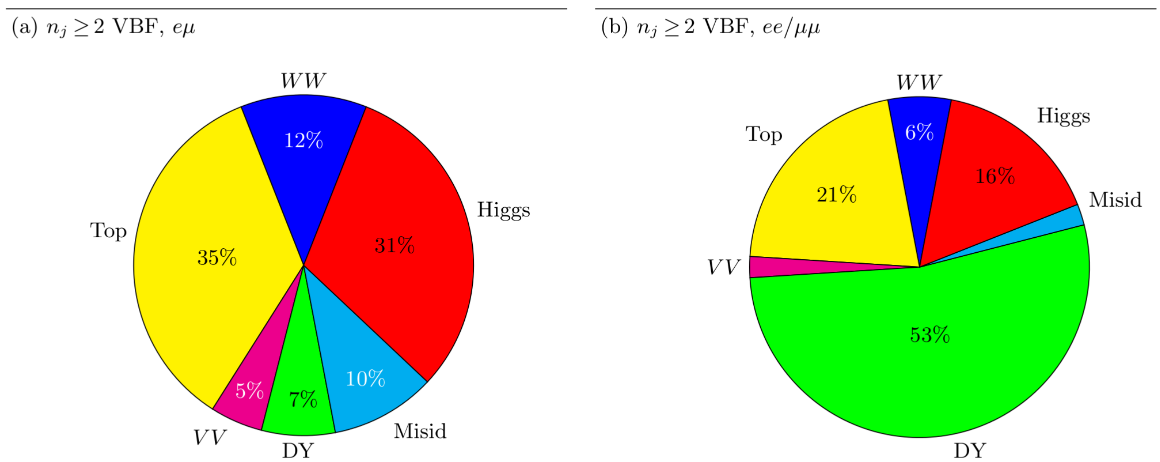


Figure 5.13: Background composition in final VBF signal region [62].

5.6 Systematic uncertainties

There are two main types of systematic uncertainties that are assessed for the analysis. First, theoretical uncertainties associated with the various signal and background yield estimates are discussed. Then, experimental uncertainties due to detector effects are shown. Normalization uncertainties refer to uncertainties that affect the cross section of the process in question in the signal region being probed. Shape uncertainties refer to systematic uncertainties that affect the shape of the final discriminating variable (either m_{T} or O_{BDT}).

5.6.1 Theoretical uncertainties

There are four main components to theoretical uncertainties assigned to signal and background processes taken from Monte Carlo. Each one is a different source of variation in the overall acceptance for that process. The first involves variation of the QCD renormalization and factorization scales used in the calculation. In this case, the two scales are varied independently and simultaneously by factors of two high or low and quantifying the resulting variation in normalization and shape for the process. This approximates the correction to the cross section that would come from including the next order of the QCD calculation (referred to as scale uncertainty). Next, there is an uncertainty associated with the PDF set used in generating the events. The uncertainty eigenvectors for the given PDF set are studied, and the envelope of maximal variation is taken as an uncertainty. Finally, there are two uncertainties associated with the choice of MC software (referred to as PDF uncertainty). An uncertainty associated with the generator chosen for the hard scattering process is evaluated by keeping the parton showering software constant but varying the matrix element generator and taking the maximal variation as an uncertainty (referred to as the generator uncertainty). The converse variation can also be done, where the matrix element generator remains constant and the generator used for the underlying event/parton shower modeling is varied (referred to as the UE/PS uncertainty). In cases where the background is normalized in a control region, the systematic uncertainty arises from variations of the extrapolation factor α between the CR and the SR, which can affect the normalization of the background in the SR.

There are two additional uncertainties that are applied to the Higgs processes as well. First, there are uncertainties assigned to the Higgs total production cross section. Then, there are uncertainties assigned based on the fact that the analysis is done in exclusive jet bins and it is possible for signal events to migrate from one bin to the next depending on the presence or absence of jets. These are assigned using the Jet Veto Efficiency (JVE) procedure [18, 83] for ggF events and the Stewart-Tackmann (ST) method [84] for VBF production.

Table 5.13 shows the total theory uncertainties on the backgrounds in the cut-based analysis. These are the sum in quadrature of the uncertainties from each of the variations described above.

Process	Theory syst. (%)
ggF H	48
Top	26
QCD WW	37
$Z/\gamma^* \rightarrow \tau\tau$	6.1

Table 5.13: Systematic uncertainties for various processes in the cut-based VBF analysis, given in units of % change in yield. Values are given for the low m_{jj} signal region.

Figures 5.14 and 5.15 show the variations in the extrapolation factor from the PDF and QCD uncertainties on the top background estimate, binned in m_{T}, for the cut-based analysis. In both cases, there was no significant shape uncertainty but normalization uncertainties were assigned according to the maximal variation. These uncertainties enter into the 26% total uncertainty on top quoted in table 5.13

While the estimate for the same-flavor $Z/\gamma^* \to \ell\ell$ background is data-driven, there is still a systematic uncertainty taken for the non-closure of the method in Monte Carlo. This is taken as the maximum of the deviation of the non-closure factor f_{corr} from unity and its uncertainty, or $\max(|1 - f_{\mathrm{corr}}|, \delta f_{\mathrm{corr}})$. For the cut-based analysis this non-closure uncertainty 23%, while for the BDT analysis it is 17%.

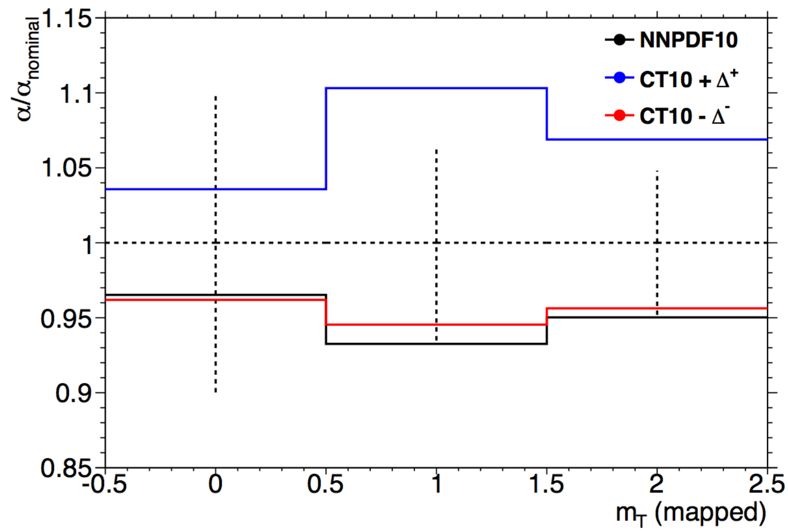


Figure 5.14: Variations in the top background extrapolation factor in the cut-based analysis due to PDF uncertainties, binned in m_{T}.

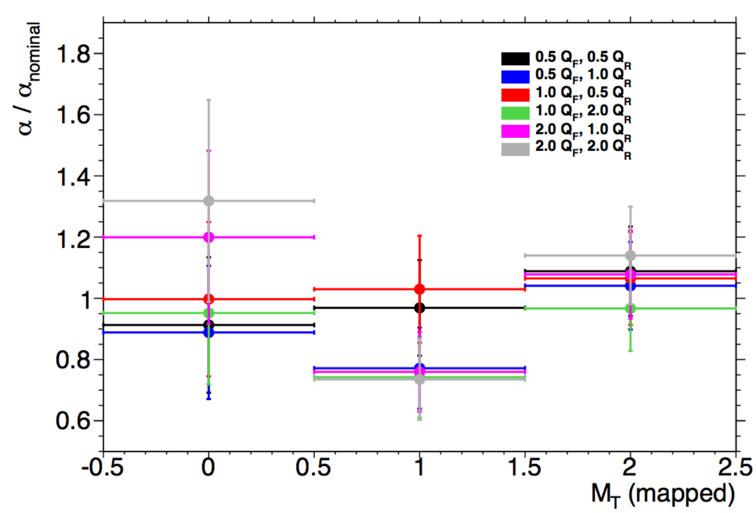


Figure 5.15: Variations in the top background extrapolation factor in the cut-based analysis due to QCD scale uncertainties, binned in m_{T}.

5.6.2 Experimental uncertainties

In this analysis, the theoretical uncertainties end up being the most dominant, but there are some experimental uncertainties that make a contribution as well. The first is the uncertainty on the measured integrated luminosity, which affects backgrounds whose normalization is taken from MC and is measured to be 2.8% in the 8 TeVdataset [85]. The dominant sources of uncertainty overall are uncertainties on the jet energy scale and resolution and the b-tagging efficiency. Additional sources include lepton uncertainties on identification, resolution, and trigger efficiency, as well as uncertainties on the missing transverse momentum .

The jet energy scale uncertainty is split into several independent components, including jet-flavor dependent calorimeter response uncertainties, uncertainties on modeling of pile-up interactions, uncertainties on extrapolation from the central to forward detector regions, and MC non-closure [86]. The uncertainty on energy scale for jets used in this analysis ranges from 1% to 7% depending on the jet p_{T} and η. The jet energy resolution varies from 5% to 20%, with uncertainties ranging from 2% to 40% (the largest uncertainties occurring at the selection threshold).

The b-tagging efficiency is independently measured in data samples enriched in dileptonic decays of $t\bar{t}$ events or in events where a muon is reconstructed in the vicinity of a jet [87, 88]. The efficiencies and their uncertainties are binned in p_{T} and decomposed into uncorrelated components using an eigenvector method [89]. Uncertainties on the efficiency range from 1% to 7.8%. The uncertainty on the rate of misidentification of c-jets as b-jets ranges from 6-14%, while the uncertainty on the rate of light jet mistagging ranges from 9-19% depending on p_{T} and η.

The total experimental uncertainties on different signal and background components are summarized in table 5.14. They are compared to the level of other statistical and systematic uncertainties as well. OVerall, the experimental uncertainties are sub-dominant compared to the statistical and theoretical uncertainties.

5.7 Results

While the combined results of all the $H \to WW^*$ sub-analyses will be discussed in the next chapter, this section presents the results of the VBF specific analysis and interpretations. As table ?? shows, the final

Sample	Total error	Stat. error	Expt. syst. err.	Theo. syst. err.
$n_j \geq 2$ VBF-enriched				
N_{sig}	13	–	6.8	12
N_{bkg}	9.2	4.7	6.4	4.5
N_{WW}	32	–	14	28
N_{top}	15	9.6	7.6	8.5
N_{misid}	22	–	12	19
N_{VV}	20	–	12	15
$N_{\tau\tau}$ (DY)	40	25	31	2.9
$N_{ee/\mu\mu}$ (DY)	19	11	15	–

Table 5.14: Composition of the post-fit uncertainties (in %) on the total signal (N_{sig}), total background (N_{bkg}), and individual background yields in the VBF analysis [62].

cut-based signal region contains 20 events in data with $m_{\text{T}} < 150$ GeV, 14 coming from the $e\mu$ channel and 6 coming from the $ee + \mu\mu$ channel. The BDT analysis has many more candidates due to its looser selection, and the yields in each bin of O_{BDT} are shown in table 5.15.

(a) Before the BDT classification

Selection	Summary						Composition of N_{bkg}									
	$N_{\text{obs}}/N_{\text{bkg}}$	N_{obs}	N_{bkg}	N_{signal}			N_{WW}		N_{top}		N_{misid}		N_{VV}	$N_{\text{Drell-Yan}}$		
				N_{ggF}	N_{VBF}	N_{VH}	N_{WW}^{QCD}	N_{WW}^{EW}	$N_{t\bar{t}}$	N_t	N_{Wj}	N_{jj}		$N_{ee/\mu\mu}$	$N_{\tau\tau}^{\text{QCD}}$	$N_{\tau\tau}^{\text{EW}}$
$e\mu$ sample	1.04 ± 0.04	718	689	13	15	2.0	90	11	327	42	29	23	31	2.2	130	2
$ee/\mu\mu$ sample	1.18 ± 0.08	469	397	6.0	7.7	0.9	37	3	132	17	5.2	1.2	10.1	168	23	1

(b) Bins in O_{BDT}

Selection	$N_{\text{obs}}/N_{\text{bkg}}$	N_{obs}	N_{bkg}	N_{ggF}	N_{VBF}	N_{VH}	N_{WW}^{QCD}	N_{WW}^{EW}	$N_{t\bar{t}}$	N_t	N_{Wj}	N_{jj}	N_{VV}	$N_{ee/\mu\mu}$	$N_{\tau\tau}^{\text{QCD}}$	$N_{\tau\tau}^{\text{EW}}$
$e\mu$ sample																
Bin 0 (not used)	1.02 ± 0.04	661	650	8.8	3.0	1.9	83	9	313	40	26	21	28	2.2	126	1
Bin 1	0.99 ± 0.16	37	37	3.0	4.2	0.1	5.0	1.0	17	3.1	3.3	1.8	2.6	–	4.0	0.2
Bin 2	2.26 ± 0.63	14	6.2	1.2	4.2	–	1.5	0.5	1.8	0.3	0.4	0.3	0.8	–	0.3	0.3
Bin 3	5.41 ± 2.32	6	1.1	0.4	3.1	–	0.3	0.2	0.3	0.1	–	–	0.1	–	0.1	0.1
$ee/\mu\mu$ sample																
Bin 0 (not used)	1.91 ± 0.08	396	345	3.8	1.3	0.8	33	2	123	16	4.1	1.1	8.8	137	20.5	0.5
Bin 1	0.82 ± 0.14	53	45	1.5	2.2	0.1	3.0	0.5	10.4	1.8	0.8	0.2	0.9	26	1.7	0.1
Bin 2	1.77 ± 0.49	14	7.9	0.6	2.5	–	0.8	0.3	1.1	0.2	0.2	–	0.3	4.4	0.3	0.1
Bin 3	6.52 ± 2.87	6	0.9	0.2	1.7	–	0.1	0.2	0.2	–	–	–	–	0.7	–	–

Table 5.15: Event selection for the VBF BDT analysis. The event yields in (a) are shown after the pre-selection and the additional requirements applied before the BDT classification (see text). The event yields in (b) are given in bins in O_{BDT} after the classification [62].

Figure 5.16(a) shows the final distribution of data candidates compared to the expected m_{T} distribution for signal and background. The data are very consistent with a VBF Higgs hypothesis. Figure 5.16(b) shows where the data candidates fall in the two-dimensional binning of m_T and m_{jj} used in the fit for the cut-based analysis.

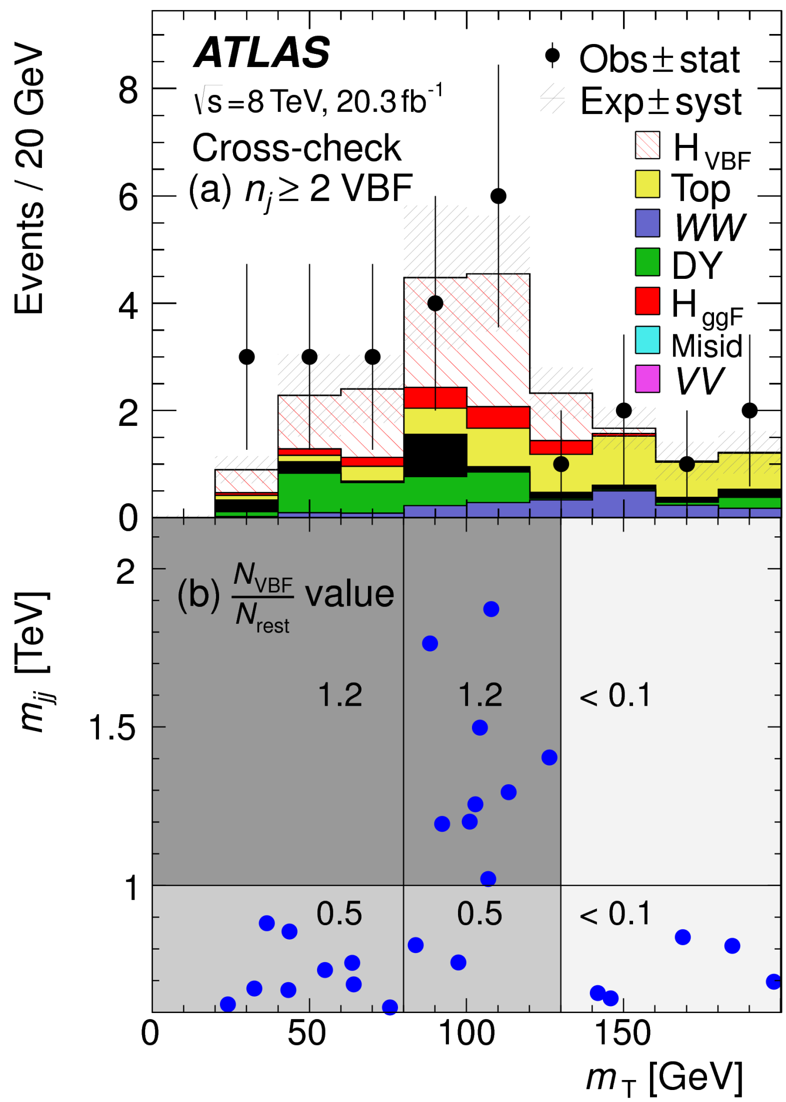


Figure 5.16: Post-fit distributions in the cut-based VBF analysis. Panel (a) shows the one-dimensional m_T distribution, while (b) shows the data candidates split into the bins of m_T and m_{jj} used in the final fit [62].

Figure 5.17 shows the distributions of O_{BDT} and m_T in the VBF BDT analysis. Again the data are quite consistent with a VBF Higgs hypothesis.

Because the cut-based result is used as a validation for the BDT analysis and the two signal regions are not fully orthogonal, it is interesting to explore which events overlap between the two analyses. Of the

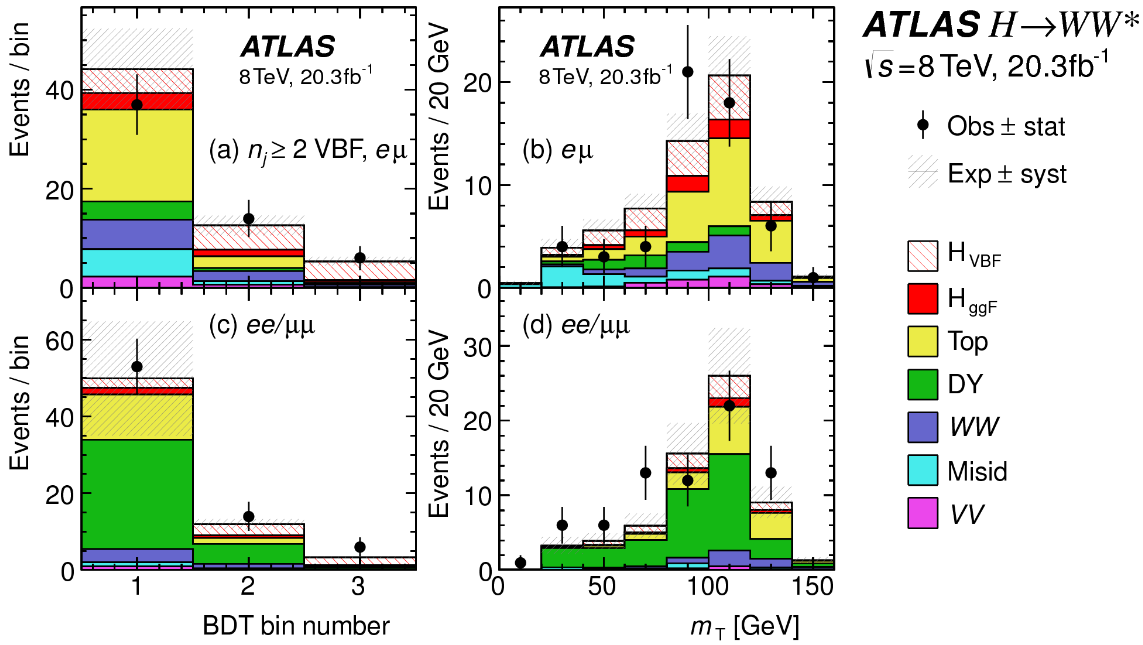


Figure 5.17: Postfit distributions in the BDT VBF analysis [62].

twenty events in the cut-based signal region, only seven were not selected by the BDT analysis, while the other thirteen also enter the BDT signal region. Figure 5.18 shows where the different analysis candidates lie in the m_{jj}-m_{T} plane. This shows clearly that the advantage of the BDT analysis is that it can extract signal candidates lower m_{jj} region due to its ability to recognize correlations with other variables.

While the context of these results in the broader $H \to WW^*$ statistical analysis will be presented in the next chapter, the significance of the VBF observation can be shown here. In the BDT analysis, the expected signal significance was 2.7σ, while the observed significance was 3.1σ. In the cut-based analysis, the expected significance was 2.1σ and the observed significance was 3.0σ. The compatibility between these two results can be evaluated by computing the probability of observing a larger difference in Z_0 values than the one measured. Using toy Monte Carlo with the ggF signal strength fixed to unity and considering only statistical uncertainties, this probability is computed to be 79%, indicating good agreement between the analyses. This result represents the first observation of the vector boson fusion production of a Higgs boson.

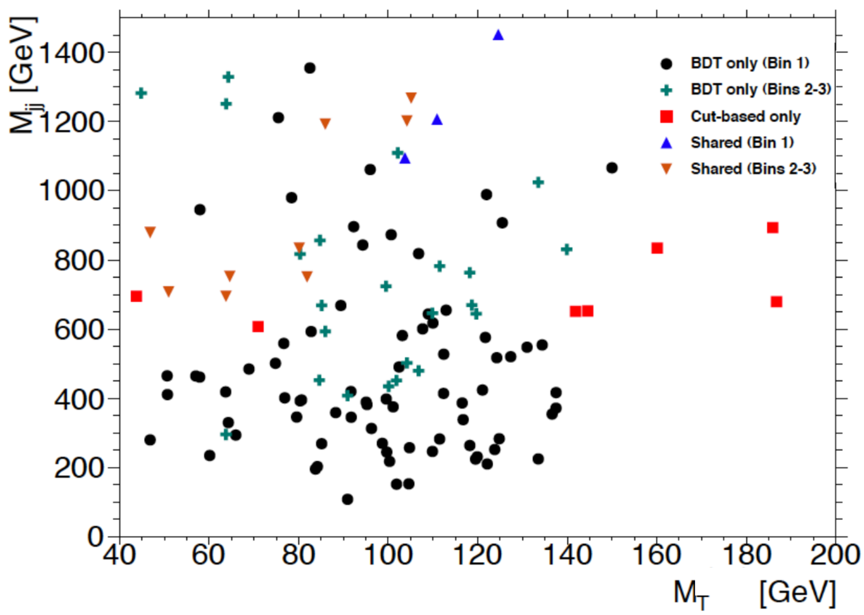


Figure 5.18: Overlap between cut-based and BDT VBF signal region candidates in the m_{jj}-m_{T} plane.

The feeling is less like an ending than just another starting point.

Chuck Palahniuk

6

Combined Run 1 $H \to WW^* \to \ell\nu\ell\nu$ results

6.1 Introduction

In the final statistical analysis of $H \to WW^* \to \ell\nu\ell\nu$, the dedicated gluon-gluon fusion and vector boson fusion sensitive signal regions are all combined into a single fit to determine the main parameters of interest, the Higgs signal strength μ and mass m_H. Therefore, while the specific requirements applied for the VBF sensitive analysis are discussed in chapter 5, the final measurement of these parameters can only be discussed in combination with the results of the ggF dedicated analysis. For example, because ggF Higgs production is considered a background in the VBF analysis, the ggF dedicated signal regions can actually constrain the normalization of this background in the VBF dedicated region.

This chapter presents the combined interpretation of results in the $H \to WW^* \to \ell\nu\ell\nu$ analysis

	SR category i			Fit var.
n_j, flavor	$\otimes\, m_{\ell\ell}$	$\otimes\, p_{\mathrm{T}}^{\ell 2}$	$\otimes\, \ell_2$	
$n_j = 0$				
$e\mu$	$\otimes\,[10, 30, 55]$	$\otimes\,[10, 15, 20, \infty]$	$\otimes\,[e, \mu]$	m_{T}
$ee/\mu\mu$	$\otimes\,[12, 55]$	$\otimes\,[10, \infty]$		m_{T}
$n_j = 1$				
$e\mu$	$\otimes\,[10, 30, 55]$	$\otimes\,[10, 15, 20, \infty]$	$\otimes\,[e, \mu]$	m_{T}
$ee/\mu\mu$	$\otimes\,[12, 55]$	$\otimes\,[10, \infty]$		m_{T}
$n_j \geq 2$ ggF				
$e\mu$	$\otimes\,[10, 55]$	$\otimes\,[10, \infty]$		m_{T}
$n_j \geq 2$ VBF				
$e\mu$	$\otimes\,[10, 50]$	$\otimes\,[10, \infty]$		O_{BDT}
$ee/\mu\mu$	$\otimes\,[12, 50]$	$\otimes\,[10, \infty]$		O_{BDT}

Table 6.1: All signal regions definitions input into final statistical fit [62].

for gluon fusion and vector boson fusion Higgs production. First, the results of the dedicated gluon fusion search are presented. Then, a comparison of the individual production mode signal strengths (μ_{ggF} and μ_{VBF} and a measurement of the combined signal strength (μ) are shown. Subsequently, the measured values of the Higgs couplings to fermions and vector bosons is presented. Finally, the cross section measurement for ggF and VBF production are shown.

6.2 Results of dedication gluon fusion $H \to WW^* \to \ell\nu\ell\nu$ search

The details of the dedicated gluon fusion $H \to WW^* \to \ell\nu\ell\nu$ search are not discussed in this thesis and instead left to more comprehensive sources [62]. However, a brief summary of the results are essential for describing the results of the full analysis and interpreting the results of the dedicated VBF search in this broader context.

Table 6.1 shows the individual signal regions that were input into the final statistical fit. The ggF dedicated bins use m_{T} as their discriminating variable and are separated into bins of p_{T} of the subleading lepton as well. The VBF dedicated bin uses the O_{BDT} distribution as its final discriminant.

Table 6.2 shows the yields in the various signal regions in both data and expected signal and backgrounds. The yields for signal and background are all scaled according to the final normalizations calculated in the fit.

	N_{obs}	N_{bkg}	N_{ggF}	N_{VBF}
$n_j = 0$	3750	3430 ± 90	300 ± 50	8 ± 4
$n_j = 1$	1596	1470 ± 40	102 ± 26	17 ± 5
$n_j \geq 2$, ggF $e\mu$	1017	960 ± 40	37 ± 11	13 ± 1.4
$n_j \geq 2$, VBF	130	99 ± 9	7.7 ± 2.6	21 ± 3

Table 6.2: Post-fit yields in the different ggF and VBF dedicated signal regions [62].

Figure 6.1 shows the final post-fit m_{T} distribution in the $n_j \leq 1$ regions. The data are very consistent with the hypothesis of ggF Higgs production. These yields are used as input, along with the VBF results

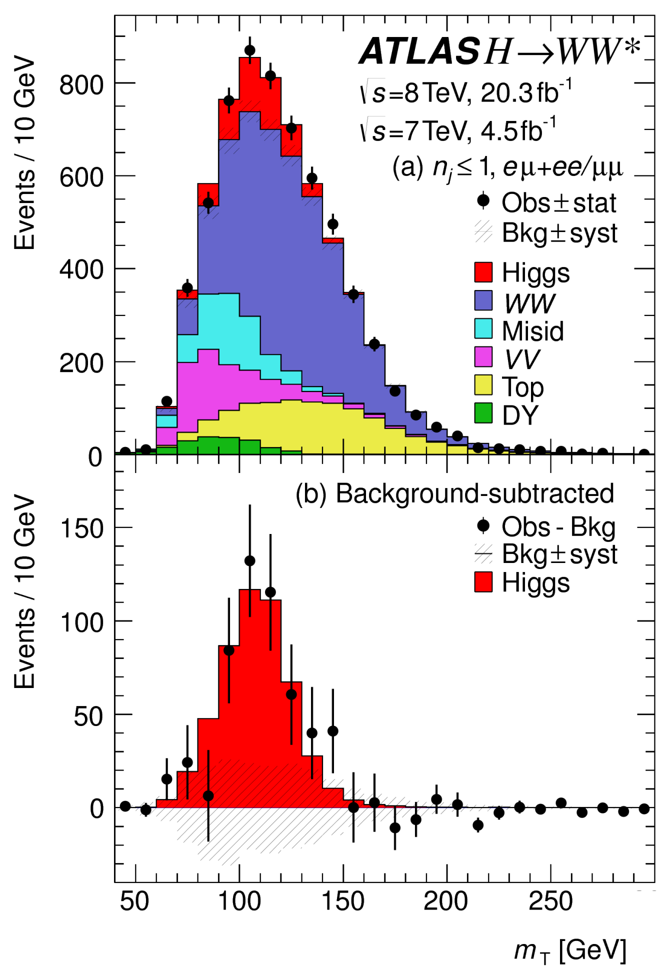


Figure 6.1: Post-fit m_{T} distribution in the $n_j \leq 1$ regions [62].

in chapter 5, for the physical interpretation of results presented in subsequent sections.

6.3 Signal strength measurements in ggF and VBF production

When all of the signal regions are combined in the fit, there can be a combined measurement of the signal strength as well as the individual ggF and VBF signal strengths. The combined signal strength is the ratio of the sum of the gluon fusion and VBF cross sections to the theory prediction, or a signal strength for the total Higgs production cross section that this analysis is sensitive to. The final measured combined signal strength μ is measured shown in equation 6.1.

$$\begin{aligned}
\mu &= 1.09 \quad {}^{+0.16}_{-0.15}\,(\text{stat.}) \quad {}^{+0.08}_{-0.07}\begin{pmatrix}\text{expt}\\ \text{syst}\end{pmatrix} \quad {}^{+0.15}_{-0.12}\begin{pmatrix}\text{theo}\\ \text{syst}\end{pmatrix} \quad \pm 0.03\begin{pmatrix}\text{lumi}\\ \text{syst}\end{pmatrix} \\
&= 1.09 \quad {}^{+0.16}_{-0.15}\,(\text{stat}) \quad {}^{+0.17}_{-0.14}\,(\text{syst}) \\
&= 1.09 \quad {}^{+0.23}_{-0.21}.
\end{aligned} \tag{6.1}$$

Figure 6.2 gives the best fit signal strength $\hat{\mu}$ as a function of the hypothesized Higgs mass. The value at 125.36 GeVcorresponds to the μ quoted in equation 6.1. This value of the Higgs mass is used because it is the most precise mass measurement from ATLAS, a result of the combined $\gamma\gamma$ and ZZ mass measurements [90].

As explained in chapter 3, a probability p_0 can be computed using the test statistic q_0 to quantify the probability that the background could fluctuate to produce an excess at least as large as the one observed in the data. The local p_0 value is shown in figure 6.3 as a function of m_H. The minimum p_0 value is at $m_H = 130$ GeV and corresponds to a significance of 6.1σ. The curve is relatively flat and the significance is the same at 125.36 GeVwithin the quoted precision. The expected significance for a signal with strength $\mu = 1.0$ is 5.8σ. This represents the first discovery level significance measurement in the $H \to WW^* \to \ell\nu\ell\nu$ analysis.

All the results presented so far in this section have been for the combined gluon fusion and VBF production modes. However, each signal strength can be calculated separately in the likelihood as well. There are two ways to do this. First, the likelihood can be parameterized in terms of a single parameter, the ratio

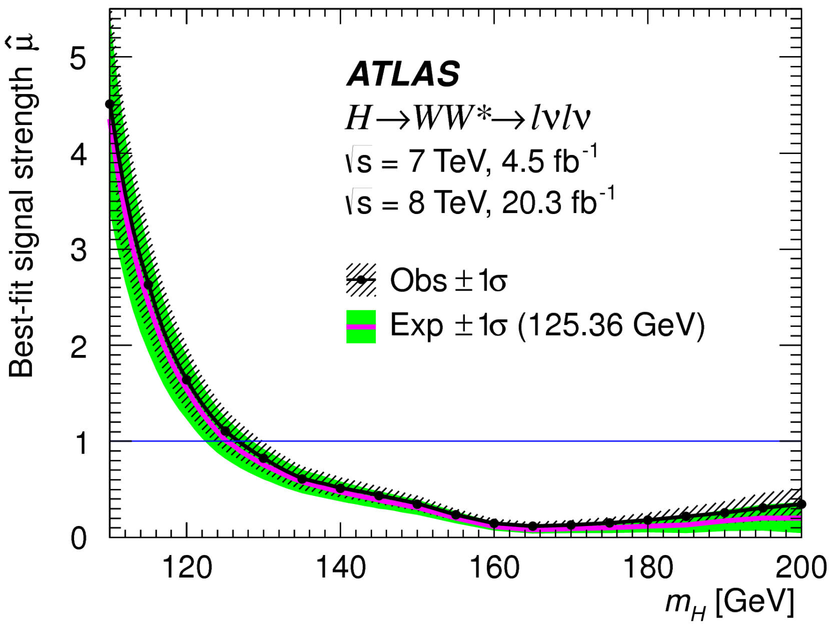


Figure 6.2: Best fit signal strength $\hat{\mu}$ as a function of hypothesized m_H [62].

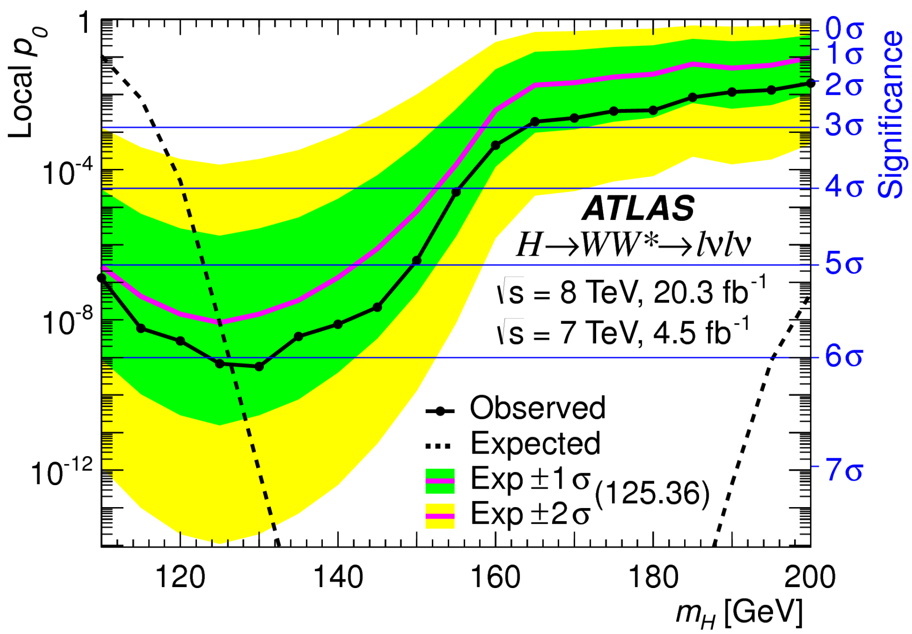


Figure 6.3: Local p_0 as a function of m_H [62].

of the VBF and gluon fusion signal strengths. With this method, the significance of the VBF observation can be evaluated. Figure 6.4 shows the likelihood as a function of the ratio $\mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}}$.

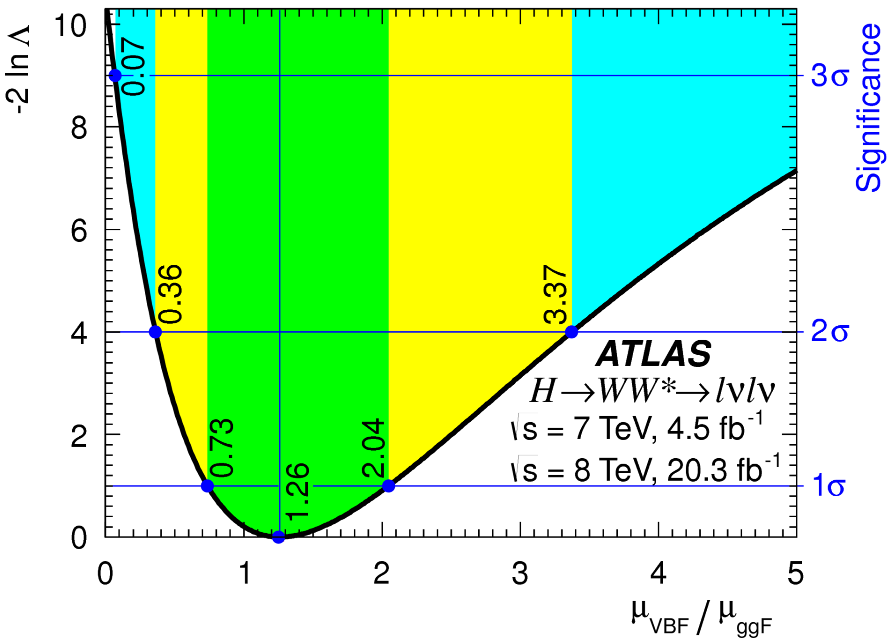


Figure 6.4: Likelihood as a function of $\mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}}$ [62].

The best fit value of the ratio of signal strengths is shown in equation 6.2. Within the quoted uncertainties, it is consistent with a ratio of unity.

$$\frac{\mu_{\mathbf{VBF}}}{\mu_{\mathrm{gg}\mathbf{F}}} = 1.26\,^{+0.61}_{-0.45}\,(\mathrm{stat.})\,^{+0.50}_{-0.26}\,(\mathrm{syst.}) = 1.26\,^{+0.79}_{-0.53} \tag{6.2}$$

The null hypothesis for VBF production corresponds to a ratio of $\mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}} = 0$. The likelihood in figure 6.4 gives a significance of 3.2σ at $\mu_{\mathrm{VBF}}/\mu_{\mathrm{ggF}} = 0$, as quoted in chapter 5.

In addition to the ratio of signal strengths, each signal strength can be varied independently in the likelihood as well. Figure 6.5 shows the two dimensional likelihood scan in the μ_{ggF}-μ_{VBF} plane. The best fit values of the two signal strengths are shown in equation 6.3. Both are consistent with unity within their uncertainties.

$$\begin{array}{lllll} \mu_{\mathrm{gg}\mathbf{F}} & = 1.02 & \pm 0.19 & {}^{+0.22}_{-0.18} & = 1.02 \quad {}^{+0.29}_{-0.26} \\ \mu_{\mathbf{VBF}} & = 1.27 & {}^{+0.44}_{-0.40} & {}^{+0.29}_{-0.21} & = 1.27 \quad {}^{+0.53}_{-0.45}. \\ & & (\mathrm{stat.}) & (\mathrm{syst.}) & \end{array} \tag{6.3}$$

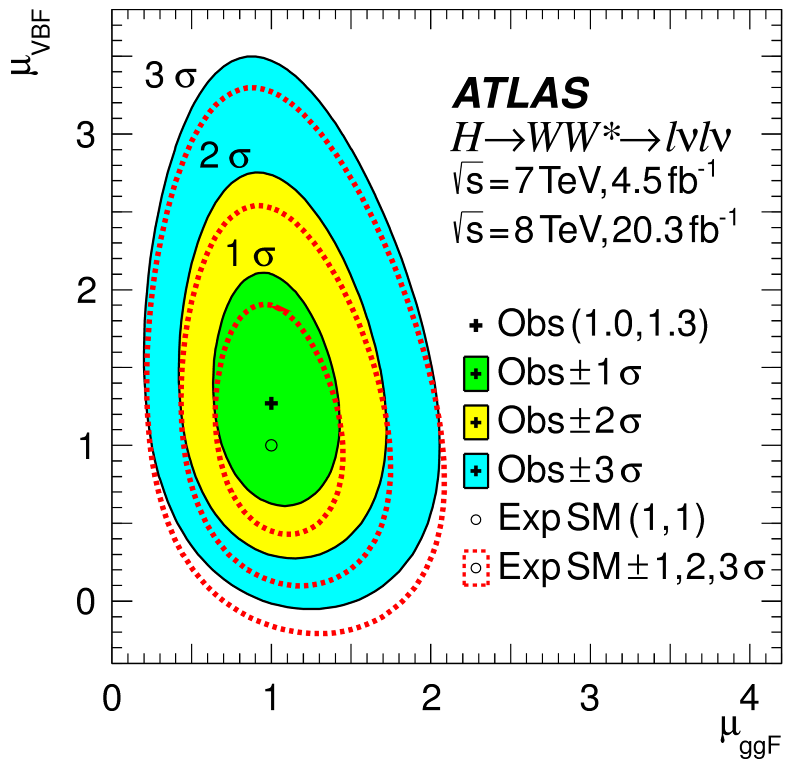


Figure 6.5: Likelihood scan as a function of μ_{VBF} and μ_{ggF} [62].

6.4 Measurement of Higgs couplings to vector bosons and fermions

Similar to the parameterization of signal strength, the couplings of the Higgs to fermions and bosons can also be parameterized. The parameter of interest in this case is κ, or the ratio of the measured coupling to the standard model expectation. Both the fermion and boson couplings have these so-called scale factors, κ_F for fermions and κ_V for bosons. Gluon fusion production is sensitive to the fermion couplings through the top quark loops in its production, while VBF production is sensitive to the vector boson couplings in its production. Both modes are sensitive to the vector boson couplings in their decays. The signal strengths will have dependence on the coupling scale factors as described in equation 6.4 [18].

$$
\begin{aligned}
\mu_{\mathbf{ggF}} &\propto \frac{\kappa_F^2 \cdot \kappa_V^2}{\left(\mathcal{B}_{H\to f\bar{f}} + \mathcal{B}_{H\to gg}\right)\kappa_F^2 + \left(\mathcal{B}_{H\to VV}\right)\kappa_V^2} \\
\mu_{\mathbf{VBF}} &\propto \frac{\kappa_V^4}{\left(\mathcal{B}_{H\to f\bar{f}} + \mathcal{B}_{H\to gg}\right)\kappa_F^2 + \left(\mathcal{B}_{H\to VV}\right)\kappa_V^2}.
\end{aligned}
\tag{6.4}
$$

Figure 6.6 shows the two-dimensional likelihood scan of κ_F and κ_V. The best-fit values are given in equation 6.5. The best-fit values are consistent with unity within their uncertainties.

$$
\begin{aligned}
\kappa_F &= 0.93 \quad {}^{+0.24}_{-0.18} \quad {}^{+0.21}_{-0.14} = 0.93 \quad {}^{+0.32}_{-0.23} \\
\kappa_V &= 1.04 \quad {}^{+0.07}_{-0.08} \quad {}^{+0.07}_{-0.08} = 1.04 \quad \pm 0.11. \\
& \qquad\quad (\text{stat.}) \quad (\text{syst.})
\end{aligned}
\tag{6.5}
$$

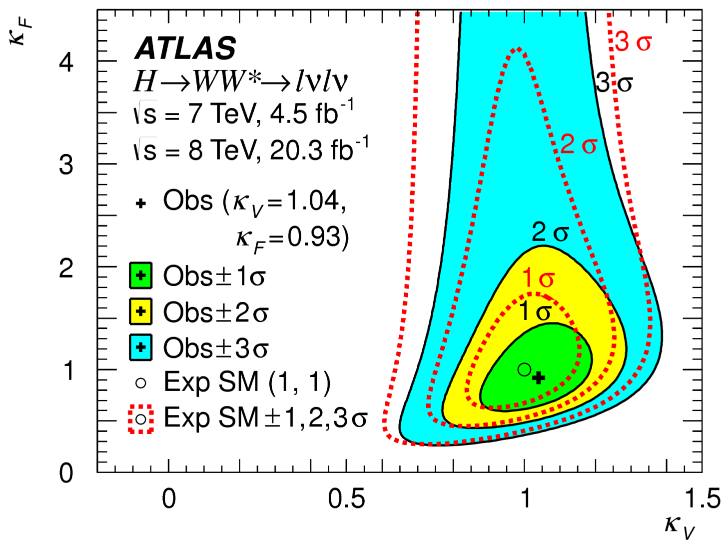


Figure 6.6: Likelihood scan as a function of κ_F and κ_V [62].

6.5 Higgs production cross section measurement

Another measurement that comes naturally from the signal strength numbers quoted earlier is the production cross section and 7 and 8 TeV for both gluon fusion and VBF production. The general equation for calculating the cross section is given in equation 6.6.

$$
\begin{aligned}
\left(\sigma \cdot \mathcal{B}_{H\to WW^*}\right)_{\text{obs}} &= \frac{(N_{\text{sig}})_{\text{obs}}}{\mathcal{A}\cdot\mathcal{C}\cdot\mathcal{B}_{WW\to\ell\nu\ell\nu}} \cdot \frac{1}{\int L\,\mathrm{d}t} \\
&= \hat{\mu} \cdot \left(\sigma \cdot \mathcal{B}_{H\to WW^*}\right)_{\text{exp}}
\end{aligned}
\tag{6.6}
$$

$(N_{\text{sig}})_{\text{obs}}$ is the number of events observed in data. $\mathcal{A}$ is the geometric and kinematic acceptance of the detector, while $\mathcal{C}$ is the efficiency of the signal region selection for events that are reconstructed in the detector. The branching ratio of a WW system to leptons must also be divided out. The production cross section depends on the center of mass energy and the production mode desired (gluon fusion or VBF), and so three separate cross section measurements are quoted in equation 6.7.

$$\begin{aligned}
\sigma^{7\text{TeV}}_{\text{ggF}} \cdot \mathcal{B}_{H\to WW^*} &= 2.0 \quad \pm 1.7 \;{}^{+1.2}_{-1.1} \;= 2.0 \;{}^{+2.1}_{-2.0} \;\text{pb} \\
\sigma^{8\text{TeV}}_{\text{ggF}} \cdot \mathcal{B}_{H\to WW^*} &= 4.6 \quad \pm 0.9 \;{}^{+0.8}_{-0.7} \;= 4.6 \;{}^{+1.2}_{-1.1} \;\text{pb} \\
\sigma^{8\text{TeV}}_{\text{VBF}} \cdot \mathcal{B}_{H\to WW^*} &= 0.51 \;{}^{+0.17}_{-0.15} \;{}^{+0.13}_{-0.08} = 0.51 \;{}^{+0.22}_{-0.17} \;\text{pb.} \\
& \qquad\quad (\text{stat.}) \;\; (\text{syst.})
\end{aligned} \tag{6.7}$$

The predicted cross section values for gluon fusion are 3.3 ± 0.4 pb at 7 TeV and 4.2 ± 0.5 pb at 8 TeV, consistent with the measured values within their uncertainties. For vector boson fusion, the predicted cross section is 0.35 ± 0.02 pb, again consistent with the measured value.

6.6 Conclusion

The combined analysis of the gluon fusion and vector boson fusion processes in $H \to WW^* \to \ell\nu\ell\nu$ in the 7 and 8 TeV datasets has yielded the first discovery level significance for Higgs production in this decay channel. Additionally, precise measurements of the couplings to vector bosons and fermions are given. Finally, signal strengths and cross sections for each production mode are measured. Figure 6.7 shows the $H \to WW^* \to \ell\nu\ell\nu$ measurements in comparison with other Higgs decay channels in ATLAS. The measurement of signal strength from this channel remains the most sensitive in both the gluon fusion and VBF production modes for the Run 1 dataset.

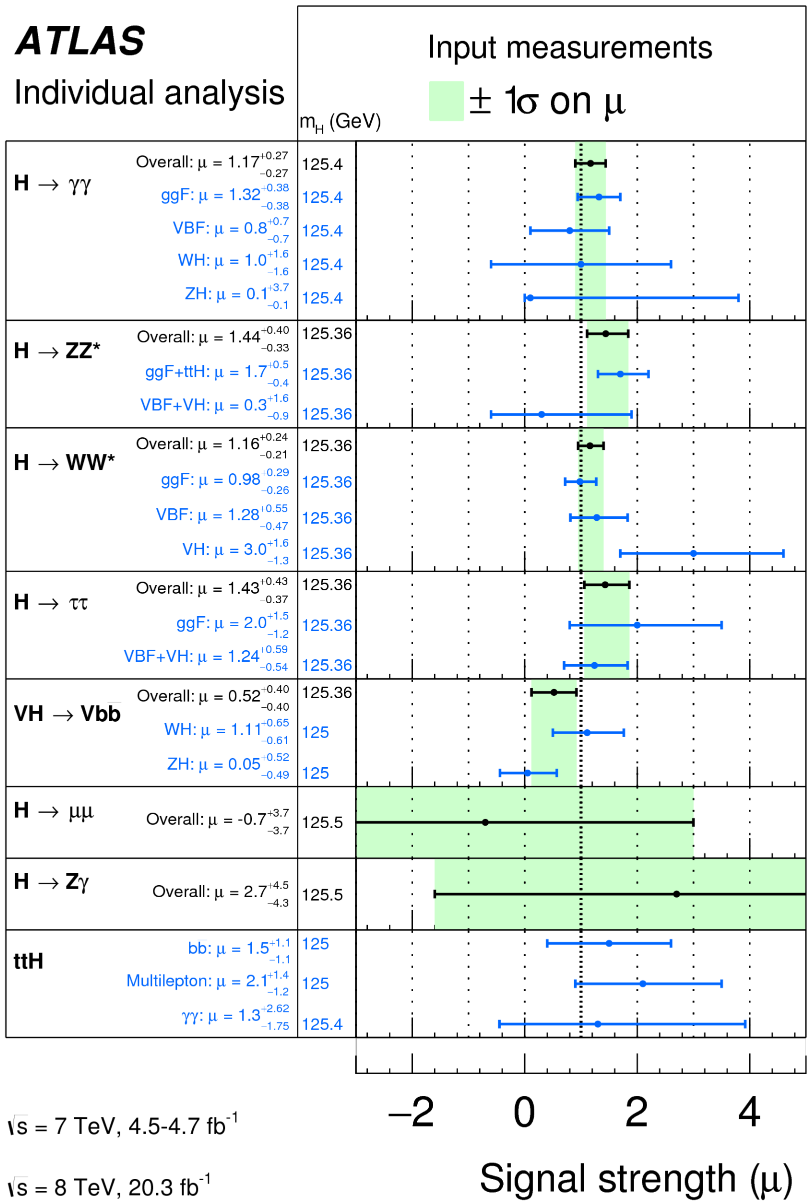


Figure 6.7: Comparison of signal strength measurements in different Higgs decay channels on ATLAS [91].

Part III

Search for Higgs pair production in the $HH \to b\bar{b}b\bar{b}$ channel in LHC Run 2 at $\sqrt{s} = 13$ TeV

Passion is in all great searches and is necessary to all creative endeavors.

W. Eugene Smith

7

Search for Higgs pair production in boosted $b\bar{b}b\bar{b}$ final states

7.1 Introduction

After the discovery of the Higgs boson in the ATLAS Run 1 dataset and the subsequent measurements of its properties, the Higgs transformed into a potential tool in searches for physics beyond the Standard Model. The pair production cross section of the Higgs can be enhanced through BSM physics. Studying di-Higgs production also probes the Higgs self-coupling, shedding light on the structure of the Higgs potential. This chapter presents a search for resonant production of a Higgs pair in the $X \to HH \to b\bar{b}b\bar{b}$ final state in $3.2\,\text{fb}^{-1}$ of data collected at $\sqrt{s} = 13$ TeV. In particular, this chapter focuses on a search for this final state in the regime where m_X is large ($\gtrsim 1$ TeV) and the Higgs bosons in the decay are significantly boosted. A tailored selection for this boosted selection, using novel techniques in jet substructure and b-tagging, is discussed. Then, the data-driven background estimate is presented. Finally, the results of the search are shown. The signal models used as benchmarks are a spin-2 Rnadall Sundrum graviton (RSG) and a narrow width spin-0 resonance. These models are described in more detail in Chapter 1. Limits on signal models are reserved for the next chapter where the results of this chapter are combined with the results of a separate selection dedicated to the lower m_X regime.

7.2 Motivation

With the center of mass energy increase from $\sqrt{s} = 8$ TeV to $\sqrt{s} = 13$ TeV, the LHC and ATLAS are able to probe new resonances at higher mass scales than previously accessible in Run 1. This is a powerful motivator for searching for a new resonance in the early 13 TeV data. Figure 7.1 shows the ratios of parton luminosities between 8 and 13 TeV for different resonance masses. For a resonance of $M_X = 2$ TeV, the cross section at $\sqrt{s} = 13$ TeV is roughly a factor of 10 larger than at $\sqrt{s} = 8$ TeV.

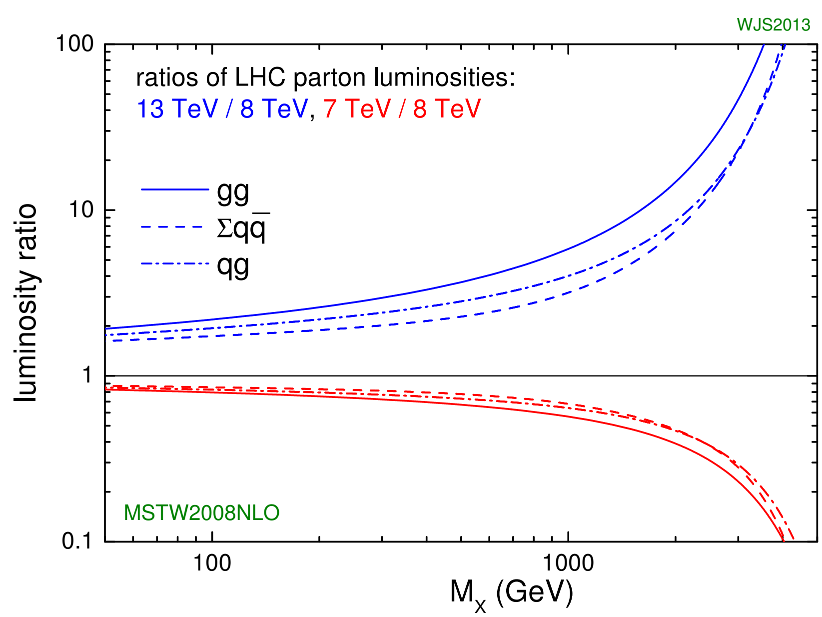


Figure 7.1: Parton luminosity ratios as a function of resonance mass M_X for 13/8 TeV and 7/8 TeV [92].

Higgs pair production offers a vast array of unprobed regions of phase space where searches for BSM physics can be made. Chapter 1 discusses some possibilities for both resonant and non-resonant enhancement of the di-Higgs production cross section. Given the increased mass reach of the LHC in Run 2, it is particularly important to focus on resonant searches at high m_X. One consideration when conducting a search in the HH final state is which decay modes of the Higgs to consider. Figure 7.2 shows the branching ratio of the HH final state for different combinations of decays of each individual Higgs. As the largest branching ratio for the 125 GeV Higgs is $H \rightarrow b\bar{b}$, the $HH \rightarrow b\bar{b}b\bar{b}$ branching ratio is also the largest at 33%.

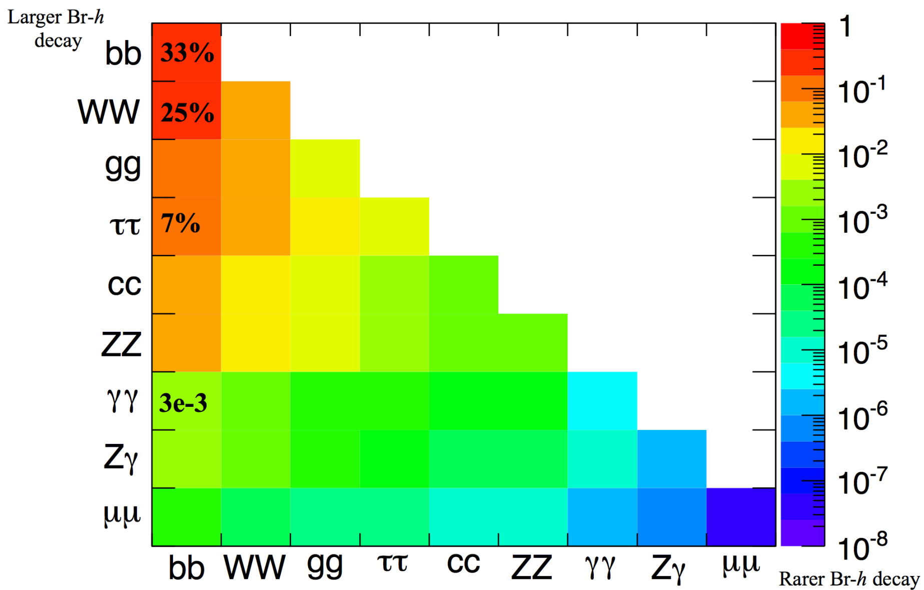


Figure 7.2: Summary of HH branching ratios [93].

At high m_X, the Higgs bosons resulting from the decay of a heavy resonance will have large p_{T}[1]. The ΔR between the decay products of the Higgs is inversely proportional to the Higgs p_{T}, as shown in equation 7.1.

$$\Delta R \approx \frac{2m}{p_{\mathrm{T}}} \tag{7.1}$$

Figure 7.3 shows the minimum ΔR between truth level B decay vertices in simulation samples for Randall-Sundrum gravitons of different masses. The figure shows that as the mass of the graviton increases, the ΔR distribution between the b quarks in the Higgs decay tends to shift to lower values. Because of this effect, it is necessary to tailor a selection to target these merged b-jets.

[1]In the limit that $m_H \ell\ell m_X$, the Higgs p_{T} is roughly $m_X/2$.

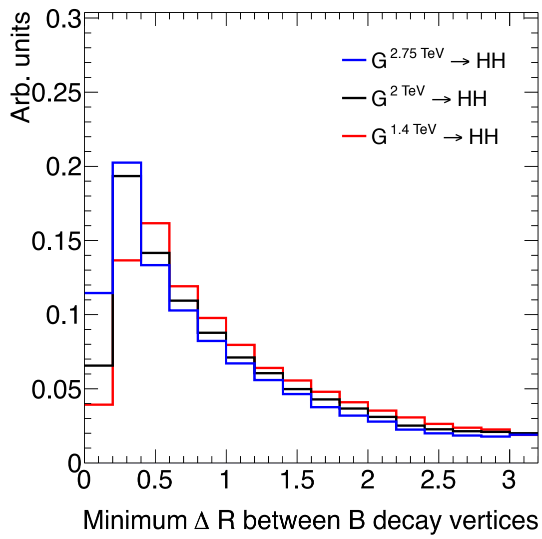


Figure 7.3: Minimum ΔR between B decay vertices for different RSG masses in a $G^*_{\rm KK} \to HH \to 4b$ sample with $c = 1$

7.3 Data and simulation samples

7.3.1 Signal models

While the resonance search is by its nature generic (as it is a simple search for a peak in the $4b$ invariant mass spectrum), there are two signal models that the selection requirements have been optimized for. The first is Randall-Sundrum (RSG) model, where a tower of massive spin-2 Kaluza-Klein gravitons is predicted. The second is a heavy narrow spin-0 resonance, the so-called "heavy Higgs". This type of resonance arises, for example, in the two Higgs doublet model (2HDM). More details about the physics of these models and their motivation is given in chapter 1.

Signal graviton ($G^*_{\rm KK}$) events are generated at leading order (LO) with MADGRAPH5 v2.2.2 [94]. The PDF set used is the NNPDF2.3 LO set [95]. For modeling parton shower and hadronization in jets, PYTHIA 8.186 is used with the A14 tune [74, 96]. The free parameters in the RSG model are the graviton mass and the coupling constant $c \equiv k/\bar{M}_{\rm Pl}$[2]. Both the production cross section and width of the graviton are proportional to c^2. Samples are generated at both $c = 1$ and $c = 2$ for a variety of mass points between

[2] k is the curvature constant for the warped extra dimension and $\bar{M}_{\rm Pl}$ is the Planck mass divided by 8π

300 GeV and 3 TeV.

The second signal sample is a heavy spin-0 resonance H with a fixed width of $\Gamma_H = 1$ GeV. This is generated with MADGRAPH5 and uses the CT10 PDF set [77]. The parton shower and hadronization are handled by HERWIG ++ with the CTEQ6L1 PDF set and the UEEE5 event tune [78, 97, 98]. Because the width and branching ratios depend on 2HDM parameters, each mass point generated with this fixed width corresponds to a different point in the 2HDM parameter phase space. Mass points are generated between 300 GeV and 1 TeV as with the RSG signal samples.

7.3.2 Background samples

While the dominant QCD multijet background is estimated with a fully data-driven method, the sub-dominant backgrounds $t\bar{t}$ and Z+jets are modeled with some input from simulation.

$t\bar{t}$ events are simulated at next-to-leading order (NLO) with the POWHEG-BOX version 1 generator using the CT10 PDF set [99]. The parton shower, hadronization, and underlying event are simulated with PYTHIA 6.428 with the CTEQ6L1 PDF set [73]. The Perugia 2012 tune is used [100]. NNLO QCD corrections to the cross sections are computed in Top++ 2.0 [101]. The top quark mass is set to 172.5 GeV. The shapes of distributions in $t\bar{t}$ are taken from MC while the normalization is taken from data.

Finally, the Z+jets background is simulated with PYTHIA 8.186 and the NNPDF2.3 LO PDF set. This background is negligible compared to the others and is taken fully from MC.

7.3.3 Data sample and trigger

This analysis is done on $3.2\,\text{fb}^{-1}$ of data taken in 2015 at $\sqrt{s} = 13$ TeV. The details of the machine conditions during this time can be found in Chapter 2. Only data which was taken during stable beam conditions with all detectors functioning is used. Events must pass a trigger which requires a single 360 GeV large radius ($R = 1.0$) jet to be reconstructed in the HLT. Figure 7.4 shows the trigger efficiency for various trigger options as a function of graviton mass. Above $m_G > 1$ TeV, the single large radius jet trigger is 99% efficient for events passing the signal selection.

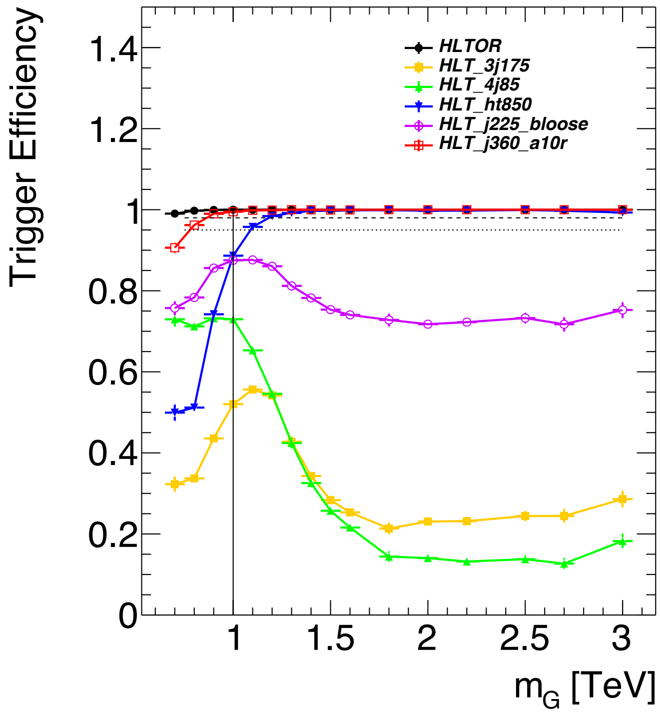


Figure 7.4: Trigger efficiency for events passing all signal region selections as a function of mass in $G^*_{\mathrm{KK}} \to HH \to 4b$ samples with $c = 1$ [102]. In the trigger names, "j" refers to a jet or jets. "ht" refers to H_T, the scalar sum of transverse momenta in the event. "bloose" refers to a loose b-tagging requirement applied to the jet. "a10r" refers to anti-k_T jets with $R = 1.0$. The numbers at the end are the thresholds on the given quantity in GeV.

7.4 Event reconstruction and object selection

The boosted selection first begins by defining a unique set of objects that can be exploited to increase signal efficiency in the kinematic regime where the final state b-jets are very merged.

7.4.1 Large radius ($R = 1.0$) jets

The first step towards reconstructing the final state is to define objects that can be used to measure the kinematics of the Higgs bosons. In the boosted selection anti-k_T jets with a radius parameter of 1.0 are used. These jets is much larger in angular size than the typical $R = 0.4$ jets and are intended to encompass both jets resulting from the Higgs decay[3]. The jets are built from clusters in the calorimeter calibrated with

[3] This is in contrast to the resolved selection, which uses two $R = 0.4$ anti-k_T jets for each Higgs

local calibration weighting [57].

Because of the large extent of these jets, great care must be taken to remove potential contributions of calorimeter clusters from pile-up. This is done using a technique called jet trimming [103]. With trimming, the constituents of the large radius jet are re-clustered with a smaller radius with the k_T algorithm. Then, these so-called subjets are removed from the larger jet if $p_T^{\text{subjet}}/p_T^{\text{jet}} < f_{\text{cut}}$. In this analysis, the subjet radius is $R = 0.2$ and $f_{\text{cut}} = 0.05$. Trimming has been shown to improve the mass resolution of large radius jets. Figure 7.5 shows the effect of trimming on the large radius jet mass (M_{J}). Because the large radius jet fully contains the higgs decay products, its invariant mass should correspond to the 125 GeV mass of the Higgs. The trimming algorithm brings the jet mass much closer to the expected Higgs mass and improves the mass resolution.

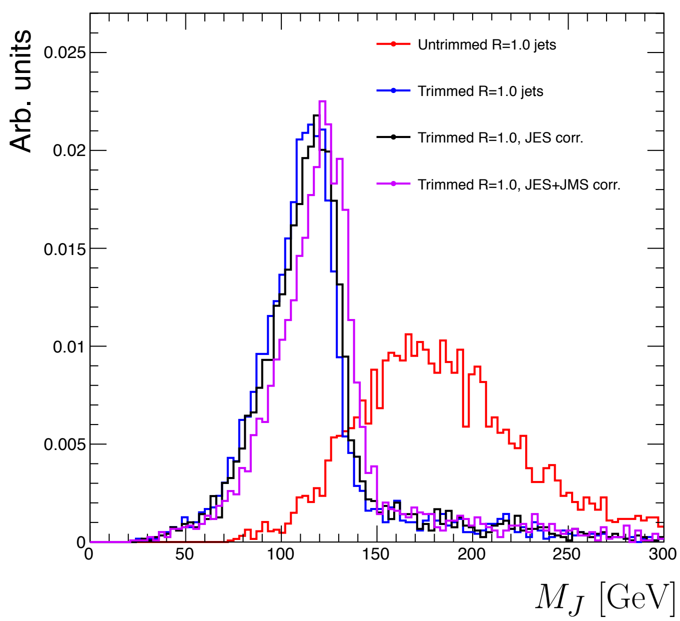


Figure 7.5: Comparison of untrimmed and trimmed jet masses for large radius jets in a RSG sample with $m_{G^*_{\text{KK}}} = 1$ TeV. JES (JMS) refers to the standard jet energy (mass) scale calibration for ATLAS [57].

The large radius jets are required to satisfy $250 < p_{\text{T}} < 1500$ GeV. They must also be within $|\eta| < 2.0$ in order to ensure that the full jet is within the inner detector tracking volume. Finally, they are required to have $M_{\text{J}} > 50$ GeV. The upper p_{T} cut and lower threshold on mass are applied to correspond to the kinematic range where uncertainties are available in ATLAS calibrations [104, 105].

7.4.2 Track jets and b-tagging

Because the b-jets from boosted Higgs decays are so close together (as illustrated in figure 7.3), narrow radius jets are required to fully resolve both b-jets. The minimum radius feasible for jets based on calorimeter deposits is determined by the calorimeter granularity. However, because b-tagging relies on information from the inner detector, it is possible to define another type of jet that can have a smaller radius and better b-tagging resolution. These jets are called "track jets" [105, 106].

Track jets are formed by applying the usual anti-k_T clustering algorithm to tracks that are required to be consistent with the primary vertex. After the jet axis has been determined using these tracks, a second step of track association is also performed to add tracks that can be useful for b-tagging [106]. In this analysis, the tracks are clustered with a radius parameter of $R = 0.2$. This radius has been shown to give good performance in boosted Higgs tagging [105, 106]. Figure 7.6 shows a comparison among different track jet radii of the efficiency for reconstructing two b-jets from each Higgs in a RSG sample as a function of mass. Track jets with radius of 0.2 give the best performance, especially at high mass. In this analysis, track jets

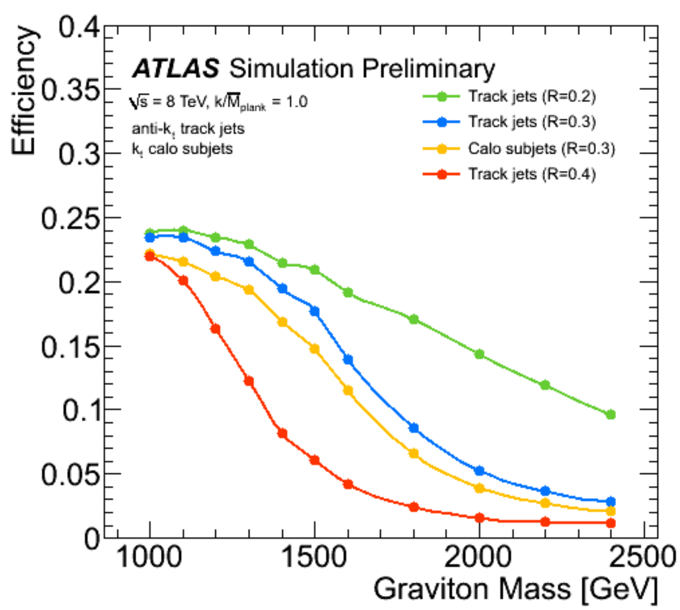


Figure 7.6: Efficiency of finding two b-jets from each Higgs in an RSG event using calorimeter jets with $R = 0.3$ or different track jet radii [106]

are required to have $p_T > 10$ GeV and $|\eta| < 2.5$. They must also have at least two tracks.

7.4.3 Muons

Muons are used in this study to correct the four-momenta of calorimeter jets by accounting for semi-leptonic b decays. The muons used are combined ID and MS muons which must satisfy tight identification requirements [53]. The muons must have $p_{\mathrm{T}} > 4$ GeV and $|\eta| < 2.5$. Table 7.1 summarizes the object requirements described in this section.

	R	p_{T}	$\lvert\eta\rvert$	M
Calorimeter jets	1.0	$250 < p_{\mathrm{T}} < 1500$ GeV	< 2.0	> 50 GeV
Track jets	0.2	> 10 GeV	< 2.5	-
Muons	-	4 GeV	< 2.5	-

Table 7.1: Summary of requirements on objects used in the $X \to HH \to b\bar{b}b\bar{b}$ search

7.5 Event selection

The first requirement in the boosted selection is for ≥ 2 large radius jets satisfying the selections outlined above. The two highest momentum large-R jets in the event are referred to as "Higgs candidates". The leading jet is required to have $p_{\mathrm{T}} > 350$ GeV.

Track jets satisfying the object selections are matched to Higgs candidate jets via ghost association [107]. Each Higgs candidate must have at least 2 track jets associated with it. These basic requirements are illustrated in figure 7.7.

The QCD multijet background produces less central jets than high mass resonances, so there is an additional requirement that the two Higgs candidates be close together in η. The large-R jets are required to satisfy $|\Delta\eta(JJ)| < 1.7$.

7.5.1 Mass requirements

The final set of requirements ensures that the Higgs candidates are consistent with expected properties of the 125.0 GeV Higgs. First, a variable (X_{hh}) is defined to measure the consistency of both of the Higgs

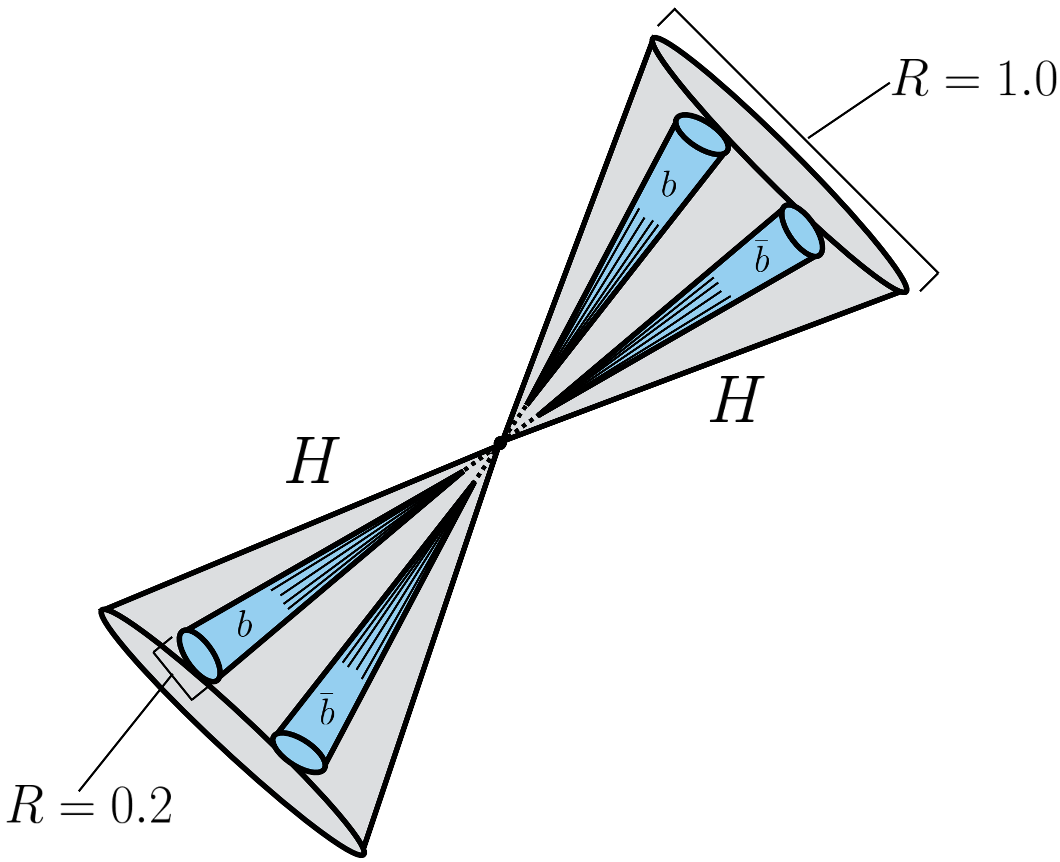


Figure 7.7: Illustration of the boosted selection requirements on Higgs candidates. Each large-radius calorimeter jet (Higgs candidate) must contain two track jets

candidate jets with the SM Higgs mass. This is shown in equation 7.2.

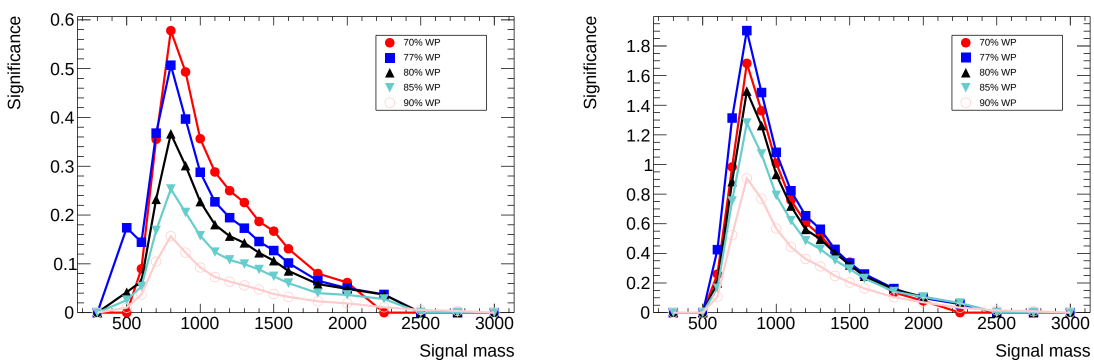
$$X_{hh} = \sqrt{\left(\frac{M_J^{\text{lead}} - 124\,\text{GeV}}{0.1 M_J^{\text{lead}}}\right)^2 + \left(\frac{M_J^{\text{sublead}} - 115\,\text{GeV}}{0.1 M_J^{\text{sublead}}}\right)^2} \tag{7.2}$$

The mass values in the X_{hh} formula are optimized to maximize signal efficiency. The sub-leading jet typically has a lower mass due to semi-leptonic b decays and final state radiation. X_{hh} effectively acts as a χ^2 measurement of the consistency of the two Higgs candidate masses with the signal hypothesis. The denominators of each term $(0.1M)$ give the uncertainty on the mass measurement for the large radius jets. Events are required to satisfy $X_{hh} < 1.6$.

Before making the requirement on X_{hh}, the masses of the Higgs candidates are corrected for semi-leptonic b decays using muons with the criteria outlined in the previous section. Any muons within a $\Delta R < 0.2$ of a b-tagged track jet (as described in the next section) have their four-momenta added to the four-momentum of the Higgs candidate. This correction does not affect the pre-selection requirements but does affect the X_{hh} requirement and the final invariant mass distribution used.

7.5.2 b-TAGGING REQUIREMENTS

The last requirement applied is on the number of b-tagged track jets. There are two signal regions defined. The first requires exactly four b-tagged track jets, two in each Higgs candidate (known as the $4b$ signal region). At high resonance masses, this requirement is inefficient, so an additional signal region requiring only three b-tagged track jets is also defined (known as the $3b$ signal region). While this has a larger background it is also more efficient for high resonance masses. For both signal regions, the MV2c20 algorithm, where the training sample for the algorithm has 20% charm events is used. More details for this algorithm can be found in Chapter 2.



(a) (b)

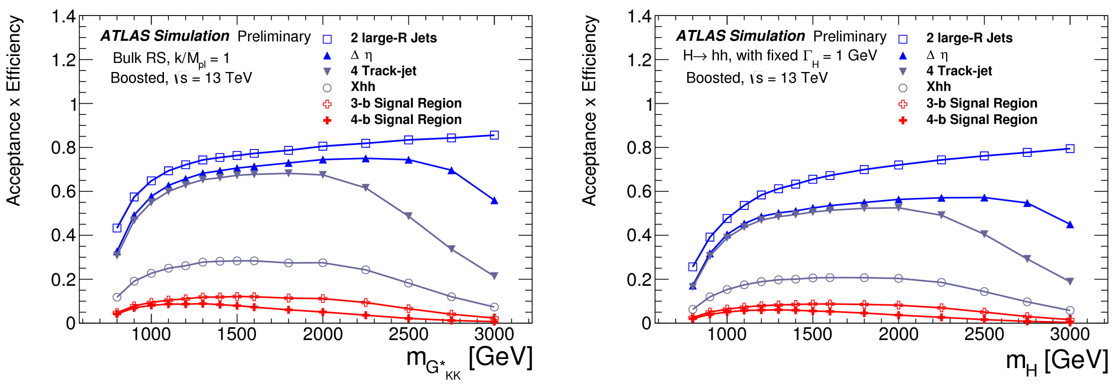
Figure 7.8: Estimated significance as a function of signal mass for RSG $c = 1$ models in the $3b$ (a) and $4b$ (b) regions for different b-tagging efficiency working points

Once the algorithm is selected, an efficiency working point must also be chosen. This working point defines the efficiency with which true b-jets are tagged and also fixes the overall background rejection of the algorithm. Higher efficiency working points accept more true b-jets but also allow for more background. Five different working points (70%, 77%, 80%, 85%, 90%) are tested. With each working point, the full data driven background estimation method is run to quantify the amount of background that will be present in the final signal region. The significance is quantified using the median discovery significance for signal and background with Poisson errors, given in equation 7.3 [108].

$$Z = \sqrt{2\left((s+b)\ln\left(1+\frac{s}{b}\right)-s\right)} \tag{7.3}$$

Here, s is the expected number of signal events and b is the expected number of background events. This formula is derived using Poisson statistics with errors on both the signal and background. It is used because it is valid in the regime where s and b are of the same order. Note that in the limit where s is much smaller than b, this equation reduces to the more well known $s/\sqrt{b}$. Figure 7.8 shows the estimated significance as a function of signal mass in RSG $c = 1$ models for the $3b$ and $4b$ signal regions. The 77% working point gives the best performance over a wide range of masses in the $4b$ signal region. As this is the region which contributes the most to the total discovery significance, the 77% efficiency working point is chosen for the analysis.

7.5.3 Selection efficiency



(a) (b)

Figure 7.9: Acceptance $\times$ efficiency as a function of mass for (a) RSG and (b) narrow heavy scalar signal models [109].

Figure 7.9 shows the product of acceptance and efficiency as a function of mass for both the RSG and narrow heavy scalar resonance signal models. After $m_X > 1$ TeV, the efficiency of the $4b$ requirement begins to decline. After $m_X > 2$ TeV, the efficiency of requiring two track jets in each Higgs candidate begins to decline as well. Both of these behaviors illustrate the difficulty of resolving the merged decay products at high mass. Figure 7.10 shows a more detailed comparison of the signal efficiency in the $3b$ vs $4b$ signal regions for the RSG model. The efficiencies shown here are relative to all prior selection requirements. It can be seen there that at high masses the $3b$ signal region is more efficient for signal.

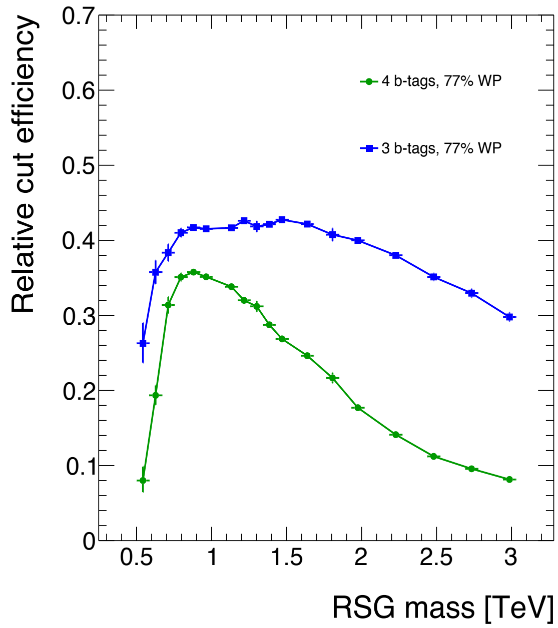


Figure 7.10: Efficiency of requiring 3 or 4 b-tagged track jets vs. RSG mass. The efficiency quoted is relative to the previous selection requirements (rather than an absolute efficiency).

To investigate the degradation of b-tagging efficiency at high p_{T}, the individual jet tagging efficiencies can be compared as a function of signal mass. This is shown in figure 7.11. The figure shows that the leading jet tagging efficiency in both calorimeter jets degrades heavily, while the sub-lead jet tagging efficiency remains relatively constant. More details on the cause of this degradation are shown in appendix A.

The final discriminating variable used in the boosted analysis is M_{2J}, the invariant mass of the two Higgs candidates. In order to improve the mass resolution, the four-momenta of each Higgs candidate are scaled by m_h/M_J. The effect of this correction is small in the boosted analysis but is done for consistency with the resolved selection. Table 7.2 shows the effect of the selection requirements on signal and background simulations as well as data.

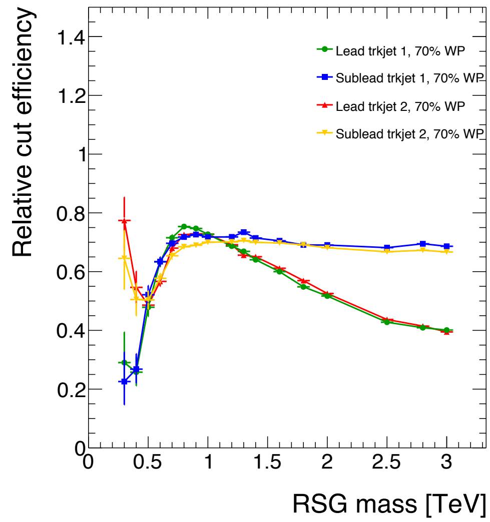


Figure 7.11: MV2c20 b-tagging efficiency for each of the four track jets in the boosted $4b$ selection as a function of RSG mass for $c = 1$ models.

Selection	Data	$m_{G^*_{\mathrm{KK}}} = 1\mathrm{TeV}$	$m_{G^*_{\mathrm{KK}}} = 2\mathrm{TeV}$	$t\bar{t}$	Z+jets
N(fiducial large-R jets)≥ 2	2202396	23.3	0.48	32345.2	4255.7
leading large-R jet $p_T > 350$ GeV	1873741	22.9	0.48	26511.7	3649.9
Both large-R jet $m > 50$ GeV	1854625	21.2	0.47	24369.8	3575.8
Both large-R jet $p_T < 1500$ GeV	1853601	21.2	0.46	24346.5	3572.9
$\lvert\Delta\eta(JJ)\rvert < 1.7$	1435273	20.8	0.44	20751.0	3265.8
≥ 2 track-jets per large-R jet	1224727	19.8	0.40	18234.5	2692.6
3 b-tags, $X_{hh} < 1.6$	316	3.4	0.067	46.7	2.0
4 b-tags, $X_{hh} < 1.6$	20	2.9	0.030	1.4	0.0

Table 7.2: Effect of boosted selection on data, RSG signal models, $t\bar{t}$, and Z+jets. The numbers from simulation are normalized with the MC generator cross section and do not take into account the data driven estimates described in section 7.6 [110].

7.6 Data-driven background estimation

The largest background to this final state is QCD multijet production, constituting 80-90% of the total background. Because of the difficulties in modeling higher order QCD processes, this background is estimated with a fully data-driven method. The only other non-negligible background is $t\bar{t}$, constituting the other 10-20%[4]. Due to the presence of $t\bar{t}$ in the sideband region where the QCD background will be

[4] The Z+jets background is a sub-percent level contribution

estimated, the normalization of the QCD and $t\bar{t}$ backgrounds are simultaneously estimated.

7.6.1 Mass region definitions

The first step in the data-driven background estimate is to define a sideband mass region where the background normalization can be derived. Additionally, a control region is defined where the background estimate can be validated. The control (CR) and sideband (SB) regions are defined using a radial distance in the two-dimensional large-R jet mass plane, R_{hh}, which is defined in equation 7.4.

$$R_{hh} = \sqrt{\left(M_J^{\text{lead}} - 124\,\text{GeV}\right)^2 + \left(M_J^{\text{sublead}} - 115\,\text{GeV}\right)^2} \tag{7.4}$$

Events in the sideband region are required to fail the signal region $X_{hh} < 1.6$ requirement and have $R_{hh} > 35.8\,\text{GeV}$. The control region consists of those events which are not in the signal or sideband regions. Figure 7.12 shows the definition of the signal, control, and sideband mass regions. Table 7.3 sum-

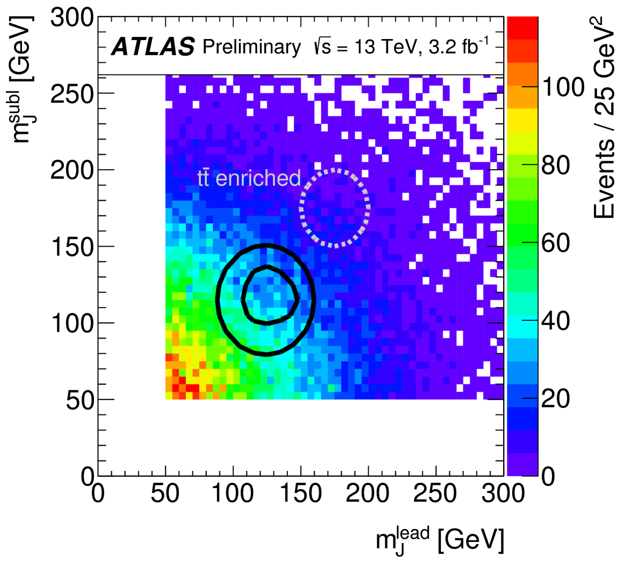


Figure 7.12: M_J^{sublead} vs. M_J^{lead} in a 2 b-tag data sample. The signal region is defined by the inner black contour ($X_{hh} < 1.6$) and the sideband region is defined by the outer contour ($R_{hh} > 35.8\,\text{GeV}$). The region between the black contours is the control region. The mass region which is enriched in $t\bar{t}$ background is also shown for illustration. [109]

marizes the mass region selections for the three different regions used in the analysis.

Region	Requirement	Notes
Signal Region (SR)	$X_{hh} < 1.6$	-
Control Region (CR)	$R_{hh} < 35.8\,\text{GeV}$ and $X_{hh} > 1.6$	Used for validation of background estimates
Sideband Region (SB)	$R_{hh} > 35.8\,\text{GeV}$	Used to derive background normalization

Table 7.3: Mass region definitions used for background estimation

7.6.2 Background estimation

The method for estimating the background in this analysis is similar to the ABCD method presented in Chapter 5. In this case, the two handles used to define different regions for the estimate are the number of b-tagged track jets and the mass regions. A region requiring exactly two b-tagged track jets in one large-R jet (referred to as the 2-tag or $2b$ region) is defined for use in the background estimate. The number of expected background events in the $3b$ and $4b$ signal regions is then given by equation 7.5.

$$N_{\text{bkg}}^{3(4)-\text{tag,SR}} = \mu_{\text{Multijet}} N_{\text{Multijet}}^{2-\text{tag,SR}} + \beta_{t\bar{t}} N_{t\bar{t}}^{3(4)-\text{tag,SR}} + N_{Z+\text{jets}}^{3(4)-\text{tag,SR}} \tag{7.5}$$

In this equation, $N_{\text{bkg}}^{3(4)-\text{tag}}$ is the expected number of background events in the $3b$ or $4b$ signal regions. $N_{\text{Multijet}}^{2-\text{tag}}$ is the number of multijet events in the 2-tag region. $N_{t\bar{t}}^{3(4)-\text{tag}}$ is the number of $t\bar{t}$ events predicted in the MC for the $3b$ or $4b$ signal region, and the variable is similarly defined for the Z+jets background. The $\beta_{t\bar{t}}$ parameter is a scale factor used to correct the normalization of the $t\bar{t}$ estimate in the signal region. μ_{Multijet} is an extrapolation factor that is derived in the sideband region and used to estimate the ratio of 2-tag events to 3(4)-tag events in the signal region. It is defined in equation 7.6.

$$\mu_{\text{Multijet}} = \frac{N_{\text{Multijet}}^{3(4)-\text{tag,SB}}}{N_{\text{Multijet}}^{2-\text{tag,SB}}} = \frac{N_{\text{data}}^{3(4)-\text{tag,SB}} - \beta_{t\bar{t}} N_{t\bar{t}}^{3(4)-\text{tag,SB}} - N_{Z+\text{jets}}^{3(4)-\text{tag,SB}}}{N_{\text{data}}^{2-\text{tag,SB}} - \beta_{t\bar{t}} N_{t\bar{t}}^{2-\text{tag,SB}} - N_{Z+\text{jets}}^{2-\text{tag,SB}}} \tag{7.6}$$

The $t\bar{t}$ scale factor ($\beta_{t\bar{t}}$) and the QCD multijet extrapolation factor (μ_{Multijet}) are estimated together in a simultaneous fit in the sideband region. Then, the number of events in the 2-tag signal region is used, along with the $t\bar{t}$ estimate in the $3b$ and $4b$ signal regions and μ_{Multijet}, to estimate the total number of background events in the two final signal regions. The shape of the final discriminant M_{2J} is also taken from the 2-tag signal region where there are more statistics. This method is illustrated graphically in figure 7.13.

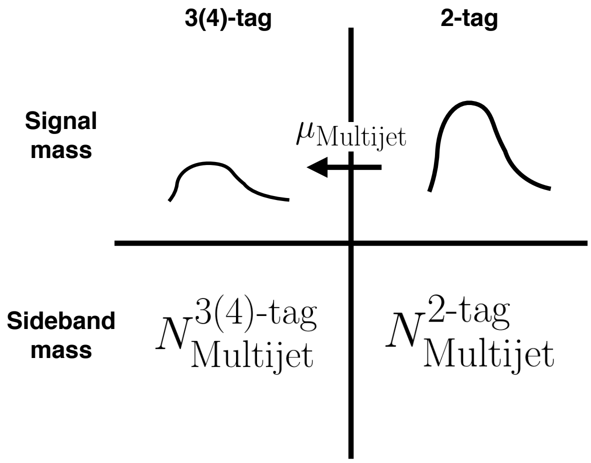


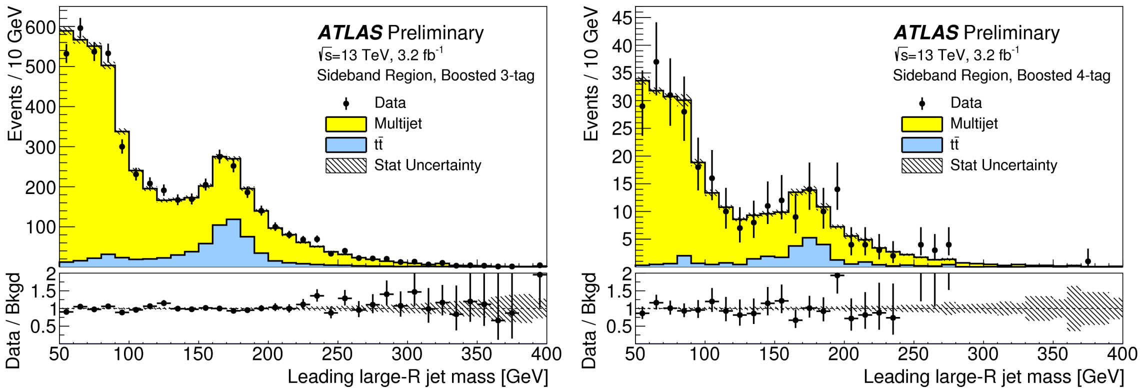
Figure 7.13: An illustration of the data-driven background estimation technique for the boosted analysis

In the $3b$ region, the fit yields values of $\mu_{\text{Multijet}} = 0.160 \pm 0.03$ and $\beta_{t\bar{t}} = 1.02 \pm 0.09$. In the $4b$ region, the fit gives $\mu_{\text{Multijet}} = 0.0091 \pm 0.0007$ and $\beta_{t\bar{t}} = 0.82 \pm 0.39$. The uncertainties quoted are statistical only. The larger uncertainties in the $4b$ values indicate the lower statistics available in that region.

Figure 7.14 shows the distributions of data and background estimates in the $3b$ and $4b$ sideband regions after the background fit has been done. The normalizations are constrained from the fit to match that of the data, but good modeling of the shape of the mass of the leading large-R jet is seen as well. The shapes of the kinematic distributions in the $4b$ region are taken from the $3b$ region due to the better MC statistics in that region.

7.6.3 Background shape fit

As mentioned in the previous section, the background shape in the 3-tag and 4-tag signal regions is taken from the 2-tag signal mass region. Due to the limited statistics available, the background shapes are additionally smoothed after being extrapolated to the 3-tag and 4-tag signal regions. Only the data in the range $900 < M_{2J} < 2000$ GeV is included in the fit due to the limited statistics available above 2 TeV. Both the $t\bar{t}$ and QCD multijet background are independently fit with an exponential shape, $y = e^{ax+b}$.



(a) (b)

Figure 7.14: Leading large-R jet mass in the $3b$ (a) and $4b$ (b) sideband regions. The multijet and $t\bar{t}$ backgrounds are estimated using the data-driven methods described above. Because their normalizations are derived in the sideband region, the total background normalization is constrained by default to match the normalization of the data [109].

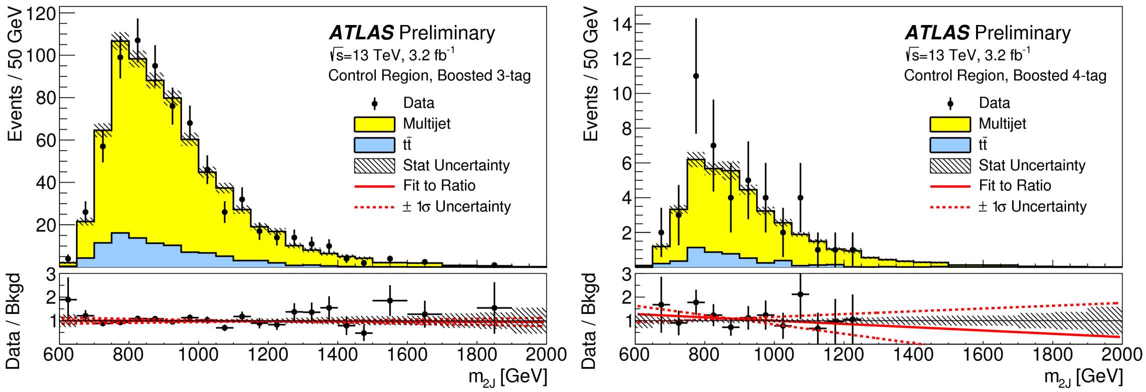
Other shapes are considered and used for the systematic uncertainties. Table 7.4 shows the fit values for the parameters. Because both the $3b$ and $4b$ QCD shapes come from the 2-tag region, the slopes derived are very similar.

	a	b
QCD $(4b)$	0.00545 ± 0.00021	5.44 ± 0.24
$t\bar{t}\,(4b)$	0.00746 ± 0.00021	4.88 ± 0.36
QCD $(3b)$	0.00545 ± 0.00021	8.30 ± 0.24
$t\bar{t}\,(3b)$	0.00746 ± 0.00021	8.58 ± 0.36

Table 7.4: Parameters derived for exponential fit to background M_{2J} shape in the $3b$ and $4b$ signal regions [110]

7.6.4 Validation of background estimate

The background estimate can be validated by using the method to estimate the number of events in the control mass region rather than the signal mass region. Figure 7.15 shows the M_{2J} distribution in the $3b$ and $4b$ control regions, comparing data and background estimates. In both cases, both the background shape and normalization are consistent with the data, indicating good agreement. The ratio of data to the background estimates is also fit to a line in the figure to test for any shape difference. The slope of the line is within 1σ (from the fit uncertainties) of flat, further indicating that the data is consistent with the background estimate in the control region.



(a) (b)

Figure 7.15: Di-jet invariant mass (M_{2J}) in the $3b$ (a) and $4b$ (b) control regions. The multijet and $t\bar{t}$ backgrounds are estimated using the data-driven methods described above [109].

Table 7.5 shows the yields in data and background estimates in the 3-tag and 4-tag sideband and control regions. Again, here, it can be seen that the total number of predicted background events from the data driven method is consistent with the number of data events in the region.

Sample (3-tag)	Sideband Region	Control Region
Multijet	4328 ± 27	607 ± 10
$t\bar{t}$	683.5 ± 8.1	99.6 ± 3.1
Z+jets	31.8 ± 3.7	7.7 ± 1.8
Total	5043 ± 28	715 ± 11
Data	5043	724
Sample (4-tag)	**Sideband Region**	**Control Region**
Multijet	247.4 ± 1.5	34.7 ± 0.6
$t\bar{t}$	28.4 ± 1.5	5.1 ± 0.7
Z+jets	3.4 ± 1.2	0.6 ± 0.5
Total	279.2 ± 2.5	40.3 ± 1.0
Data	279	45

Table 7.5: The number of events in data and predicted background events in the boosted 3-tag and 4-tag sideband and control regions. The uncertainties shown are statistical only. [109]

7.7 Systematic uncertainties

The systematic uncertainties in this analysis can be divided into two broad categories. The first type is uncertainties associated with the modeling of the signal processes. The second type of uncertainty is associated with both the shape and normalization of the background prediction.

7.7.1 Signal modeling uncertainties

The signal modeling uncertainty has three main components: theoretical uncertainty on the acceptance, experimental uncertainties on the large-R jets, and experimental uncertainties on the track jets related to b-tagging. In this analysis the experimental uncertainties are the most significant.

The first uncertainty on signal modeling is the theoretical uncertainty on the acceptance. As explained in section 5.6.1, there are four components to this uncertainty. The first is related to missing higher order terms from the matrix element calculations which is estimated by varying the QCD renormalization and factorization scales. The second is uncertainty due to the PDF set used. The third is a generator uncertainty which is estimated by modifying the generator used to model the underlying event and hadronization. Finally, there is an uncertainty associated with the modeling of the initial state and final state radiation (ISR/FSR). The total theoretical uncertainty on the signal yield is 3%, and this is dominated by the ISR/FSR modeling.

There are uncertainties on the large-R jets in both the jet energy scale (JES) and jet energy resolution (JER) as well as the jet mass scale (JMS) and jet mass resolution (JMR). These are evaluated using $\sqrt{s} = 8$ TeV data from Run 1 of ATLAS and extrapolated to the Run 2 beam and detector conditions using MC[5]. The details of these uncertainties can be found in reference [111].

Uncertainties on the track jets are related to the b-tagging efficiency. The total uncertainty on the signal yield due to b-tagging is evaluated by propagating variations of the b-tagging efficiency through the boosted selection requirements. The uncertainties are calculated jet-by-jet and parameterized as a function of b-jet p_{T} and η [89]. For high p_{T} b-jets (with $p_{\mathrm{T}} > 300$ GeV), the uncertainties are extrapolated using MC simulation from the lower p_{T} b-jets [112].

[5]The uncertainties are correspondingly larger due to the uncertainty of this extrapolation.

Table 7.6 shows the systematic uncertainties on the signal normalization for models with $m_{G^*_{\rm KK}} = 1.5$ TeV and both $c = 1$ and $c = 2$ as well as a narrow width heavy scalar. The dominant uncertainty comes from b-tagging and this uncertainty is larger in the 4-tag region than the 3-tag region.

Source	Background	$G^*_{\rm KK}$ $c=1$	$G^*_{\rm KK}$ $c=2$	H
Luminosity	-	5.0	5.0	5.0
	3-tag			
JER	<1	<1	<1	<1
JES	2	<1	<1	<1
JMR	1	12	12	11
JMS	5	14	13	17
b-tagging	1	23	22	23
Theoretical	-	3	3	3
Multijet Normalization	3	-	-	-
Statistical	2	1	1	1
Total	7	31	30	33
	4-tag			
JER	<1	<1	<1	<1
JES	<1	<1	<1	<1
JMR	4	12	13	13
JMS	5	13	13	14
b-tagging	2	36	36	36
Theoretical	-	3	3	3
Multijet Normalization	14	-	-	-
Statistical	3	1	1	1
Total	15	42	42	43

Table 7.6: Summary of systematic uncertainties in the total background and signal event yields (expressed in %) in the boosted 3-tag and 4-tag signal regions. Systematic uncertainties on the signal normalization are shown for models with $m_{G^*_{\rm KK}} = 1.5$ TeV and both $c = 1$ and $c = 2$ as well as a narrow width heavy scalar.

7.7.2 Background uncertainties

Uncertainties on the QCD multijet background normalization and shape are estimated using the control mass region. As shown previously, the background predictions in the control region match with the data yields within the statistical uncertainty in both the 3-tag and 4-tag control regions. As an additional protection, the statistical uncertainty on the background prediction in the control region is assigned as a systematic uncertainty on the normalization of the QCD background.

Additional robustness tests are done by varying the definition of the control mass region and the b-tagging requirements used to define the 2-tag sample. In all cases, the effect of the variations is found to be within the statistical uncertainties on the background normalization in the control region.

Shape uncertainties on the background are evaluated using two techniques. First, as shown in figure 7.15, the ratio between the data and background prediction is fit with a linear function. The uncertainties on the slope of this fit are assigned as shape uncertainties. An additional uncertainty is assigned by using alternate power law fit functions for the smoothing of the background shape. Table 7.7 shows the alternate shapes used. The largest difference between the nominal fit function and the alternates, taking into account the 1σ uncertainty band on each fit as well, is taken as a shape uncertainty.

Functional Form
$f_1(x) = p_0(1-x)^{p_1} x^{p_2}$
$f_2(x) = p_0(1-x)^{p_1} e^{p_2\, x^2}$
$f_3(x) = p_0(1-x)^{p_1} x^{p_2\, x}$
$f_4(x) = p_0(1-x)^{p_1} x^{p_2\, ln\, x}$
$f_5(x) = p_0(1-x)^{p_1} (1+x)^{p_2\, x}$
$f_6(x) = p_0(1-x)^{p_1} (1+x)^{p_2\, ln\, x}$
$f_7(x) = \frac{p_0}{x}(1-x)^{p_1 - p_2\, ln\, x}$
$f_8(x) = \frac{p_0}{x^2}(1-x)^{p_1 - p_2\, ln\, x}$

Table 7.7: Alternate fit functions used to model the M_{2J} distribution in the QCD multijet background. In the equations, $x = M_{2J}/\sqrt{s}$.

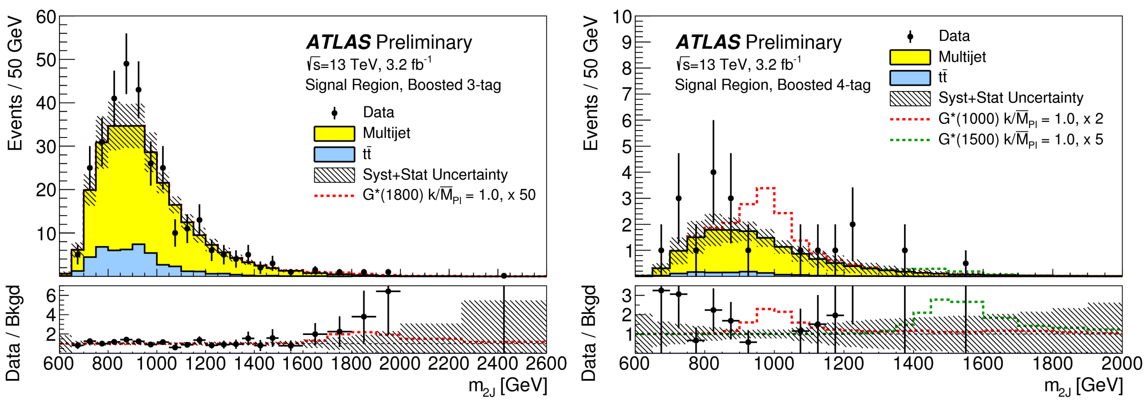
The uncertainties on the $t\bar{t}$ background are obtained by propagating the various experimental variations (JES, JER, JMS, JMR, b-tagging) through the analysis selection requirements. Table 7.6 summarizes the background uncertainties in the 3-tag and 4-tag regions.

7.8 Results

Table 7.8 shows the observed yields in the 3-tag and 4-tag signal regions for the boosted analysis compared to the predicted number of background events. In the 3-tag region, 316 events are observed with a predicted background of 285 ± 19. In the 4-tag region, 20 events are observed with a predicted background of 14.6 ± 2.4. Figure 7.16 shows the M_{2J} distribution in the 3-tag and 4-tag regions. There are some small excesses in the data, in particular in the 3-tag region around $M_{2J} \approx 900$ GeV and in the region of $1.6 < M_{2J} < 2.0$ TeV. The significance of these excesses will be evaluated in the next chapter in the statistical combination with the resolved results.

Sample	Signal Region (3-tag)	Signal Region (4-tag)
Multijet	235 ± 14	13.5 ± 2.4
$t\bar{t}$	48 ± 22	1.2 ± 1.0
Z+jets	2.0 ± 2.2	-
Total	285 ± 19	14.6 ± 2.4
Data	316	20
G^*_{KK} (1000 GeV), $c = 1$	3.4 ± 0.9	2.9 ± 1.1

Table 7.8: Observed yields in the 3-tag and 4-tag signal regions for the boosted analysis compared to the predicted number of background events Errors correspond to the total uncertainties in the predicted event yields. The yields for a graviton with $m_{G^*_{\mathrm{KK}}} = 1$ TeV and $c = 1$ are also shown. [109]



(a) (b)

Figure 7.16: Di-jet invariant mass (M_{2J}) in the $3b$ (a) and $4b$ (b) signal regions. The multijet and $t\bar{t}$ backgrounds are estimated using the data-driven methods described above. In the $3b$ region, a graviton signal with $m_{G^*_{\mathrm{KK}}} = 1.8$ TeV and $c = 1$ is overlaid, with the cross section multiplied by a factor of 50 so that the signal is visible. In the $4b$ region, signals with $m_{G^*_{\mathrm{KK}}} = 1.0$ TeV and $m_{G^*_{\mathrm{KK}}} = 1.5$ TeV are overlaid, both with $c = 1$ and the yields multiplied by factors of 2 and 5 respectively [109].

There is no real ending. It's just the place where you stop the story.

Frank Herbert

8

Combined limits from boosted and resolved searches

8.1 Introduction

In order to cover the full mass range of possible resonances decaying to di-Higgs final states, two distinct tailored selections were produced. The resolved selection is more sensitive in the mass range of $400 < m_X < 1100$ GeV while the boosted selection is more sensitive to masses in the range $1100 < m_X < 3000$ GeV. Chapter 7 presents the details of the boosted selection and results. In setting limits on spin-2 Randall-Sundrum graviton (RSG) and narrow width heavy scalar (H) models, the results of the boosted selection are combined with the results of the resolved selection to cover the full mass range.

This chapter presents limits on signal models resulting from the $X \to HH \to b\bar{b}b\bar{b}$ search in both the resolved and boosted selections. It first presents a brief overview of the resolved results that go into the limit setting. Then, an overview of the statistical methods used for the search and limit setting is given. Finally, limits on the RSG and heavy scaler models are presented.

8.2 Resolved results

The details of the resolved selection will not be presented here and can be found in reference [109]. In basic terms, the selection searches for four $R = 0.4$ b-tagged calorimeter jets (where each pair of jets is one Higgs candidate). This is distinct from the boosted methodology which searches for merged decay products. The backgrounds to the resolved selection are the same as those presented in Chapter 7 for the boosted analysis.

Table 8.1 shows the results for data yields and expected background in the resolved signal region. Figure 8.1 shows the M_{2J} distribution in the resolved signal region. The total number of events is consistent with the prediction and no significant excess is seen. One event in the boosted 4-tag signal is shared with the resolved signal region and has a mass of 852 GeV.

Sample	Signal Region Yield
Multijet	43.3 ± 2.3
$t\bar{t}$	4.3 ± 3.0
Z+jets	-
Total	47.6 ± 3.8
Data	46
SM hh	0.25 ± 0.07
$G^{*}_{\mathrm{KK}}(800\ \mathrm{GeV}), c = 1$	5.7 ± 1.5

Table 8.1: Observed yields in the resolve selection 4-tag signal region compared to the predicted number of background events Errors correspond to the total uncertainties in the predicted event yields. The yields for a graviton with $m_{G^{*}_{\mathrm{KK}}} = 800$ GeV and $c = 1$ are also shown. [109]

8.3 Search technique and results

The statistical technique used for the search in this analysis is the same as that used in the $H \to WW^{*}$ analysis presented in section 3.6.2. The test statistic q_0 is used to define the p-values which measure the

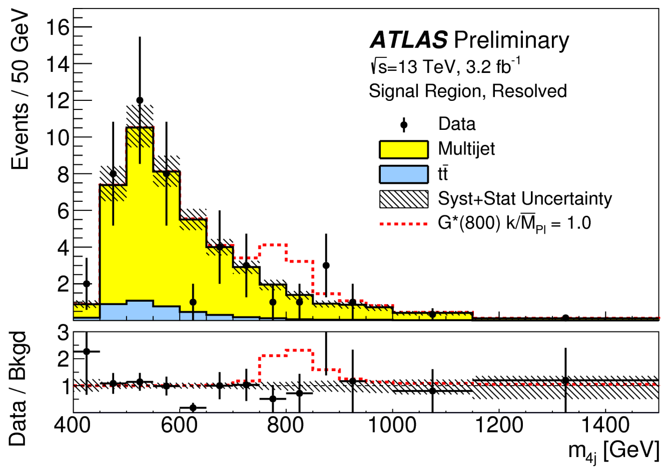


Figure 8.1: Di-jet invariant mass (M_{2J}) in the resolved signal region. A graviton signal with $m_{G^*_{\mathrm{KK}}} = 800$ GeV and $c = 1$ is overlaid. [109].

compatibility of the data with the background-only hypothesis corresponding to a signal strength $\mu = 0$.

Local p_0 values are computed to quantify the probability that the background could produce a fluctuation greater than or equal to the one observed in the data. In the resolved analysis, no significant excesses are observed. The largest discrepancy with respect to the background only hypothesis occurs near a resonance mass of 900 GeV and is found to be less than 2σ in significance.

In the boosted selection, the largest local excess is a broad excess in the $3b$ signal region that begins near $M_{2J} \approx 1.7$ GeV. Assuming a G^*_{KK} with this mass and $c = 1.0$, the local significance of this excess is 2.0σ.

8.4 Limit setting

In the absence of any significant excess observed in the data, limits on different signal models can be set. This section describes the limit setting procedure and presents combined results of the resolved and boosted analyses.

8.4.1 Limit setting procedure

The procedure used for setting exclusion limits in this analysis is the CL$_s$ method [113]. The first step in setting the limits is to define a test statistic which will be used. For limit setting, the test statistic is shown in equation 8.1.

$$\widetilde{q}_\mu = \begin{cases} -2\ln \frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(0,\hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \\ -2\ln \frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})} & 0 \leq \hat{\mu} < \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{8.1}$$

In the above equation, μ is the value of the signal strength under test, $\hat{mu}$ is the best fit μ, $th\hat{e}ta$ is the best fit value of the nuisance parameters, $\hat{\hat{\theta}}$ is the best fit value of the nuisance parameters under the fixed μ value, and L is the Poisson likelihood of the data (as described in section 3.6.2).

The test statistic $\widetilde{q}_\mu$ is constructed to protect against two interesting corner cases when setting the upper limit on the cross section. First, it protects against negative signal strengths μ which are unphysical. Second, it does not count excesses in the data larger than those expected by a signal strength μ as evidence against the μ hypothesis.

The CL$_s$ statistic is constructed by taking a ratio of two probabilities. CL$_{s+b}$ is the probability that the signal+background hypothesis would produce a value of the test statistic that is less than or equal to the observed value[1]. CL$_b$ is the probability that the that the background only hypothesis will produce a value of the test statistics less than or equal to the observed. The CL$_s$ statistic is then the ratio CL$_{s+b}$/CL$_b$. A 95% upper limit on the cross section is set at the value of μ that makes the CL$_s$ statistic less than 5%.

In practice, the limits are computed numerically within an asymptotic approximation for the distribution of the test statistic $\widetilde{q}_\mu$. The details of this approximation can be found in reference [64].

The resolved and boosted analyses are combined using a very simple procedure rather than a full statistical combination. For each mass point tested, the limit which gives the most stringent constraint is used. This means that for mass points below 1.1 TeV the resolved signal region is used, while at and above this point the combination of the orthogonal $3b$ and $4b$ boosted signal regions is used.

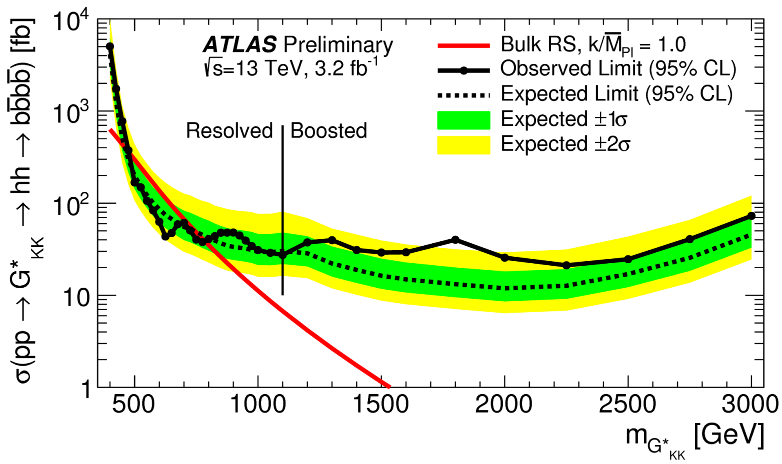
8.4.2 Limit setting results

Figure 8.2 shows the combined 95% upper bounds as a function of mass for three different models: G^*_{KK} with $c = 1$, G^*_{KK} with $c = 2$, and a narrow heavy scalar H.

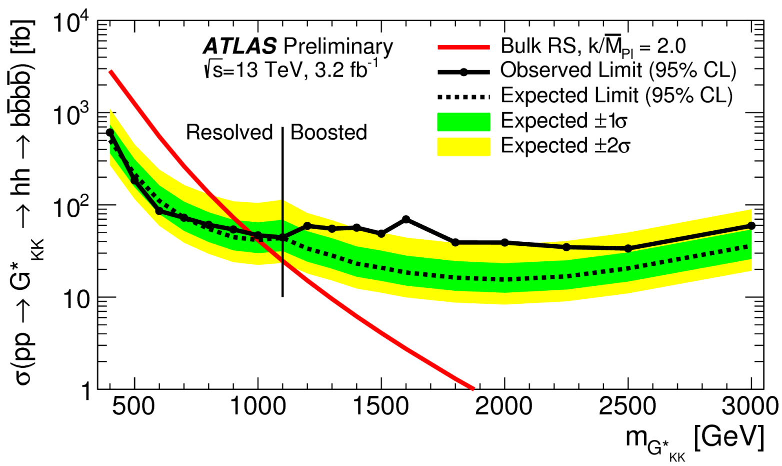
[1] Lower values of $\widetilde{q}_\mu$ mean better compatibility

The cross section of $\sigma(pp \to G^*_{\mathrm{KK}} \to hh \to b\bar{b}b\bar{b})$ with $c = 1$ is constrained to be less than 70 fb for masses in the range $600 < m_{G^*_{\mathrm{KK}}} < 3000$ GeV. For the RSG model with $c = 2$, cross sections limits between 40 fb and 200 fb are set for the mass range of $500 < m_{G^*_{\mathrm{KK}}} < 3000$ GeV. Masses in the range of $475 < m_{G^*_{\mathrm{KK}}} < 785$ GeV are excluded with $c = 1$ (with an exclusion of the range 465 to 745 GeV expected). Masses less than 980 GeV are excluded with $c = 2$ (with an exclusion for masses less than 1 TeV expected).

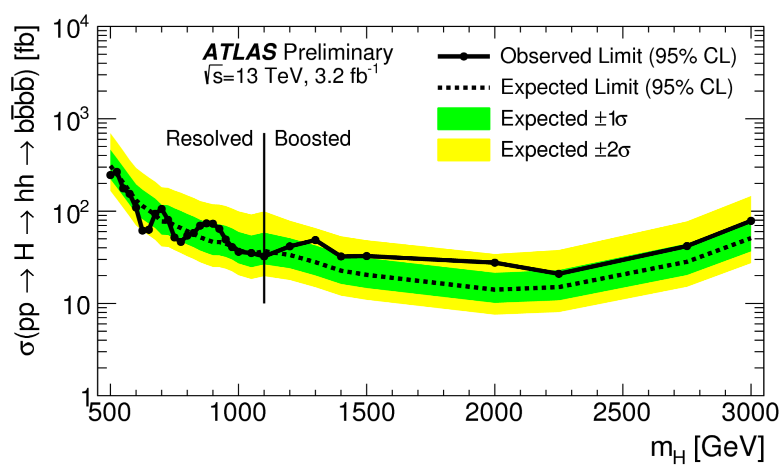
In the heavy Higgs model, the cross section upper limits for $\sigma(pp \to H \to hh \to b\bar{b}b\bar{b})$ ranges from 30 to 300 fb in the mass range of $500 < m_H < 3000$ GeV. The resolved analysis can also set an upper limit on the Standard Model di-Higgs production cross section discussed in chapter 3. The upper limit on $\sigma(pp \to hh \to b\bar{b}b\bar{b})$ in the Standard Model is constrained to be less than 1.22 pb.



(a) Bulk RS, $c = 1$



(b) Bulk RS, $c = 2$



(c) Spin-0 narrow-width H boson

Figure 8.2: Expected and observed upper limit as a function of mass for G^{*}_{KK} in the RSG model with (a) $c = 1$ and (b) $c = 2$, as well as (c) H with fixed $\Gamma_H = 1$ GeV, at the 95% confidence level in the CL$_s$ method. [109]

Part IV

Looking ahead

9

Conclusion

After being sought for many years at different collider experiments, the Higgs boson was discovered by the ATLAS and CMS experiments in 2012, confirming the leading theory for the source of electroweak symmetry breaking and filling in the last missing piece of the Standard Model. After its discovery, measurements of the particle's detailed properties and searches for new particles decaying to Higgs final states were both extremely important in constraining physics beyond the Standard Model. This dissertation presented this evolution through two results: the observation and measurement of the Higgs boson in the $H \to WW^* \to \ell\nu\ell\nu$ channel at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV and a search for Higgs pair production in the $HH \to b\bar{b}b\bar{b}$ channel at $\sqrt{s} = 13$ TeV with the ATLAS detector in *pp* collisions at the Large Hadron Collider.

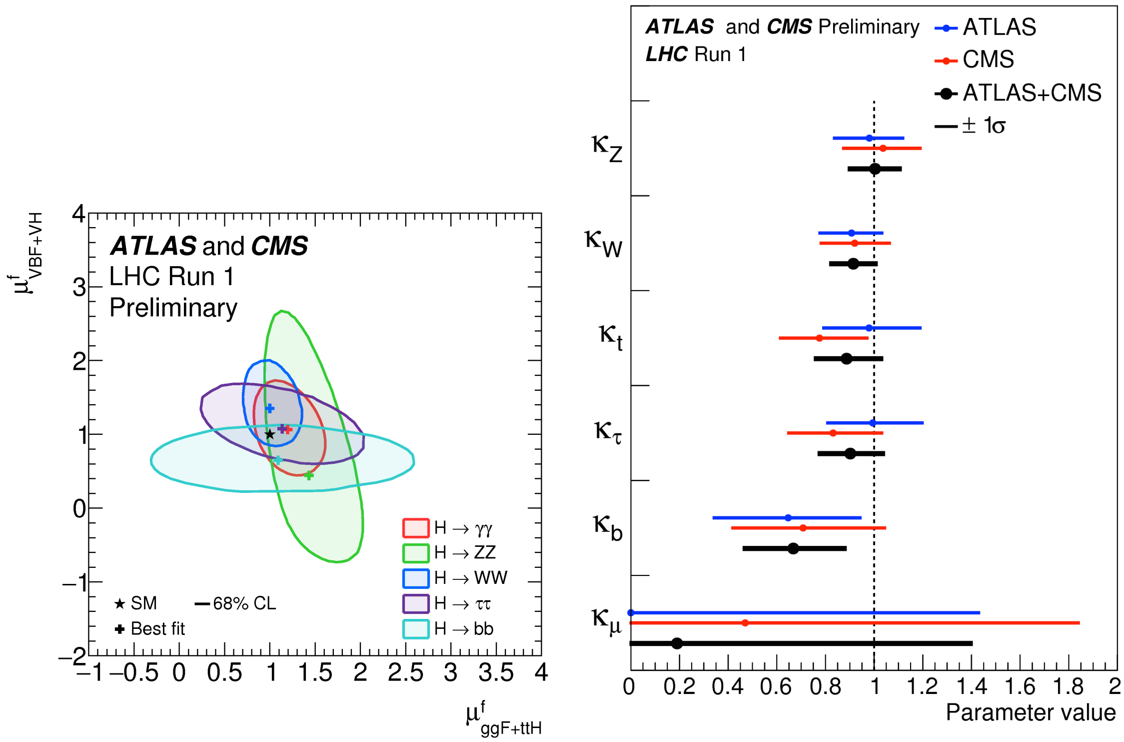
In the $H \to WW^* \to \ell\nu\ell\nu$, results from both the discovery of the Higgs boson and the full ATLAS Run 1 dataset were presented. The Higgs boson was discovered with a 6.1σ significance in a combination of the $H \to \gamma\gamma$, $H \to ZZ4\ell$, $H \to WW^* \to \ell\nu\ell\nu$ with $4.2\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 7$ TeV and

$5.2\,\text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV. With the full $20.3\,\text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV and $4.2\,\text{fb}^{-1}$ at $\sqrt{s} = 7$ TeV, ATLAS achieved discovery level significance in the $H \to WW^*$ channel alone and obtained the first observation of vector boson fusion production in that channel. The combined signal strength is measured to be $\mu = 1.09^{+0.23}_{-0.21}$. The total observed significance of the $H \to WW^*$ process is observed to be 6.1σ (with 5.8σ expected). Advanced methods for background reduction and estimation, particularly in same-flavor lepton final states, are shown. The VBF signal strength is measured to be $\mu_{\text{VBF}} = 1.27^{+0.53}_{-0.45}$ with an observed significance of 3.2σ (with 2.7σ expected).

These results required many novel innovations. The increase of pileup interactions in the higher instantaneous luminosity LHC conditions of 2012 led to a degradation of missing transverse momentum resolution. As a result, the prominent Z/γ^*+jets background of the same flavor $H \to WW^* \to \ell\nu\ell\nu$ final states increased greatly. New variables, including a track-based missing transverse momentum and a measurement of the balance between the dilepton system and recoiling jets, allowed for significant reduction of this background. In the VBF channel, selections were optimized to exploit the unique VBF final state topology. Incorporating these variables into a boosted decision tree technique allowed the analysis to exceed the 3σ observation threshold.

After the end of Run 1, the results of Higgs measurements from ATLAS were combined with those from CMS to produce the most precise measurements of the Higgs boson so far [114]. Figure 9.1 shows the combination of ATLAS and CMS data for the Higgs signal strength in and coupling measurements. In the signal strength measurements of gluon fusion and vector boson fusion, the $H \to WW^*$ channel provides the tightest constraints. Additionally, the Higgs coupling to W bosons is the most precisely measured with a relative uncertainty of 10%.

With the discovery of the Higgs firmly established and its properties measured, a natural next step was to search for new physics with Higgs final states. At $\sqrt{s} = 13$ TeV, a search for Higgs pair production in the $b\bar{b}b\bar{b}$ final state with $3.2\,\text{fb}^{-1}$ was conducted. A signal region optimized for the boosted final states arising from high mass resonances was constructed. This signal region utilized large-radius calorimeter jets and b-tagging with small radius track jets to maximize the signal acceptance. No significant excesses were observed, and upper limits on cross sections are placed for spin-2 Randall Sundrum gravitons (RSG) and



(a) (b)

Figure 9.1: Combined ATLAS and CMS measurements in Run 1 for (a) Higgs signal strength in gluon fusion and VBF and (b) Higgs couplings normalized to their SM predictions

narrow spin-0 resonances. The increase in center of mass energy in Run 2 allowed this analysis to extend upper limits up to 3 TeV, while previous results from ATLAS in Run 1 only quotes limits up to 2 TeV. The cross section of $\sigma(pp \to G^*_{\mathrm{KK}} \to hh \to b\bar{b}b\bar{b})$ with $k/\bar{M}_{\mathrm{Pl}} = 1$ is constrained to be less than 70 fb for masses in the range $600 < m_{G^*_{\mathrm{KK}}} < 3000$ GeV. For the RSG model with $k/\bar{M}_{\mathrm{Pl}} = 2$, cross sections limits between 40 fb and 200 fb are set for the mass range of $500 < m_{G^*_{\mathrm{KK}}} < 3000$ GeV. The cross section upper limits for $\sigma(pp \to H \to hh \to b\bar{b}b\bar{b})$ ranges from 30 to 300 fb in the mass range of $500 < m_H < 3000$ GeV.

While there has been a rigorous program of measurements and searches involving the Higgs, there is still much room for improvement at the High Luminosity LHC (HL-LHC) and beyond. The measured signal strength for VBF production in $H \to WW^*$ still has a relative error at the level of 40%, largely dominated by statistical uncertainty. Projections for the HL-LHC show that the uncertainty on the VBF signal strength can be reduced to approximately 15% with 3000 fb^{-1} [115, 116]. This uncertainty also assumes that theoretical uncertainties on the signal, which would be the largest contribution in this dataset, remain as they are now. Improvements in the theoretical understanding of the Higgs signal would also reduce the signal strength uncertainty dramatically. Such precision measurements allow for measurements of the Higgs coupling to vector bosons precise to the few percent level, therefore giving much power to constrain or discover new physics.

The prospects for detection of beyond the Standard Model resonant di-Higgs production at the HL-LHC are also quite promising. Figure 9.2 shows projections for the discovery significance of RSG signals at the HL-LHC in the $X \to HH \to b\bar{b}b\bar{b}$ search [116]. In all detector budget scenarios, a 1.5 TeV resonance is above or near 5σ significance, while a 2 TeV resonance is between 4-5σ except for the lowest budget.

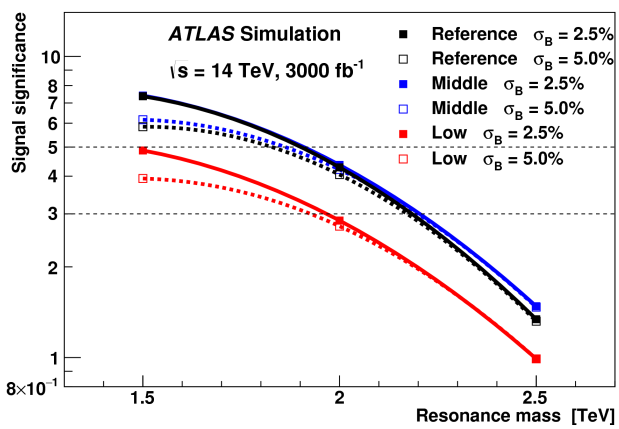


Figure 9.2: Discovery significance for RSG models at the HL-LHC in three different budget scenarios [116]. Systematic uncertainties on the background prediction (σ_B) of 2.5% and 5.0% are both tested.

The Higgs will continue to be an incredibly powerful tool in the understanding of nature at the HL-LHC and beyond. Through both precision measurements and searches, the nature of electroweak symmetry breaking will be better understood and the potential for the discovery of physics beyond the Standard Model has never been greater.

A

b-tagging performance at high p_T

One of the limiting factors of the signal acceptance in the $X \to HH \to b\bar{b}b\bar{b}$ search at high resonance masses is the degradation of the b-tagging efficiency for high p_T jets. This appendix presents a study of the underlying causes of this degradation.

A.1 Changes in MV2 score at high p_{T}

The degradation of b-tagging at high p_{T} was studied in particular in the context of RSG models at high mass. Figure A.1 shows the p_{T} of the leading track jet inside of the leading calorimeter jet in RSG events. At high $m_{G^*_{\mathrm{KK}}}$, the p_{T} spectrum of track jets is much harder than at lower masses due to the increased Higgs p_{T}.

Figure A.2 shows the MV2c20 algorithm score for the leading and subleading track jets inside of the leading calorimeter jet. In both cases, it can be seen that at higher RSG masses the MV2 score shifts towards more background like (negative) values. Additionally, this effect is more pronounced in the leading track

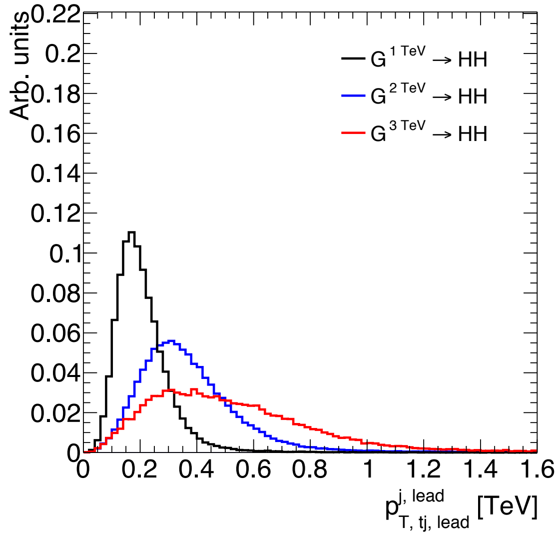
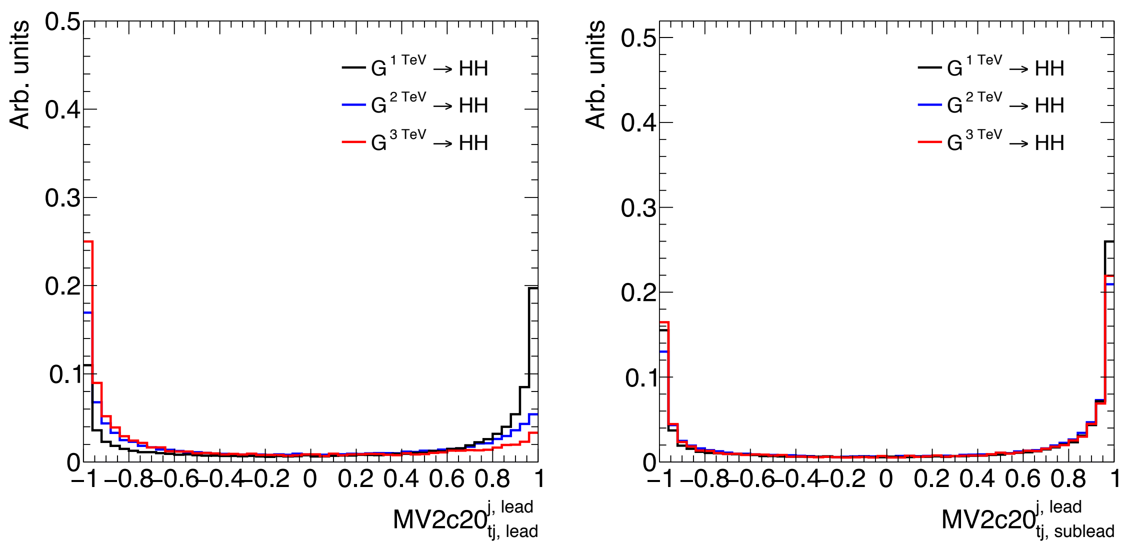


Figure A.1: p_{T} of the leading track jet in the leading calorimeter jet for different signal masses in RSG $c = 1$ models

jet than the subleading.



(a) (b)

Figure A.2: MV2c20 score for the leading track jet (a) and subleading track jet (b) of the leading calorimeter jet for different signal masses in RSG $c = 1$ models

To understand what is causing this change in the MV2c20 score, the same comparisons can be made for the input variables of MV2c20. The focus in these comparisons will be on the leading track jet as this is the one seen to have the largest difference in MV2 score. Figure A.3 shows the log likelihood ratio $\log(p_b/p_u)$

from the IP3D (three dimensional impact parameter) algorithm. At higher masses, the IP3D likelihood ratio distribution does become more background-like.

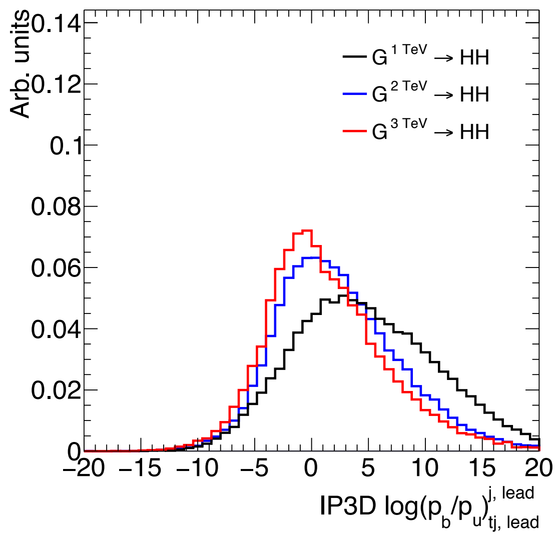


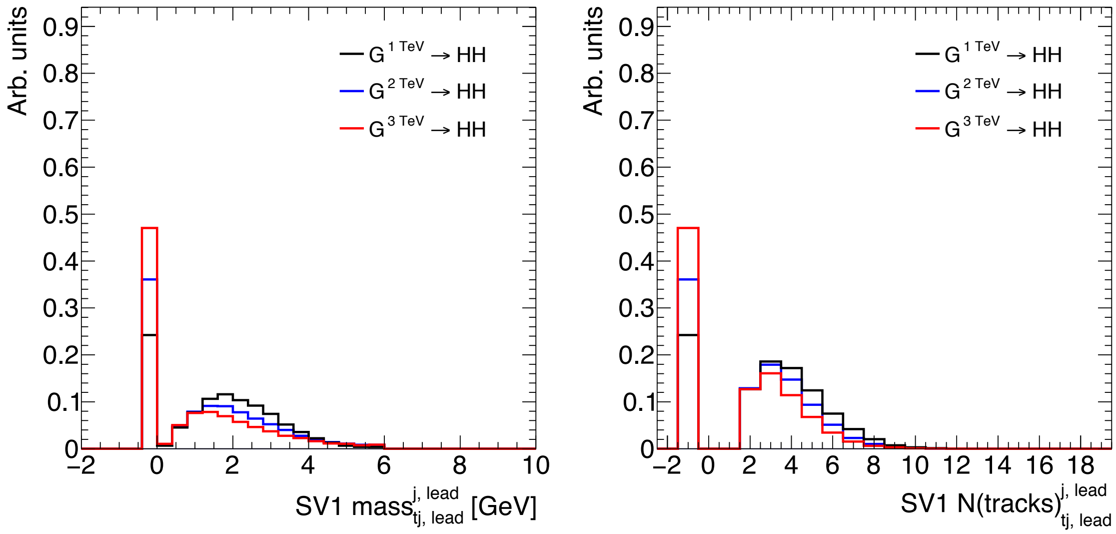
Figure A.3: IP3D log-likehood ratio ($\log(p_b/p_u)$)) of the leading track jet in the leading calorimeter jet for different signal masses in RSG $c = 1$ models

Figure A.4 shows the mass and number of tracks at the secondary vertex computed by the SV1 algorithm. When there is no secondary vertex found, the algorithm assigns a default negative value for these quantities. Both of these distributions show that there is a significantly larger fraction of jets where no secondary vertex is found in the high mass samples compared to the $m_{G^*_{\mathrm{KK}}} = 1$ TeV sample. The SV1 algorithm's inability to find a secondary vertex could be an important factor in the overall MV2 score shift, as this eliminates eight of the input variables that would normally contribute information to the algorithm.

Figure A.5 shows the same quantities for the JetFitter algorithm. In this case, there is also a change in the fraction of jets which have their secondary vertices successfully reconstructed, but this change is not as drastic as that seen in SV1. There is also an increase in the number of jets which have high values of mass.

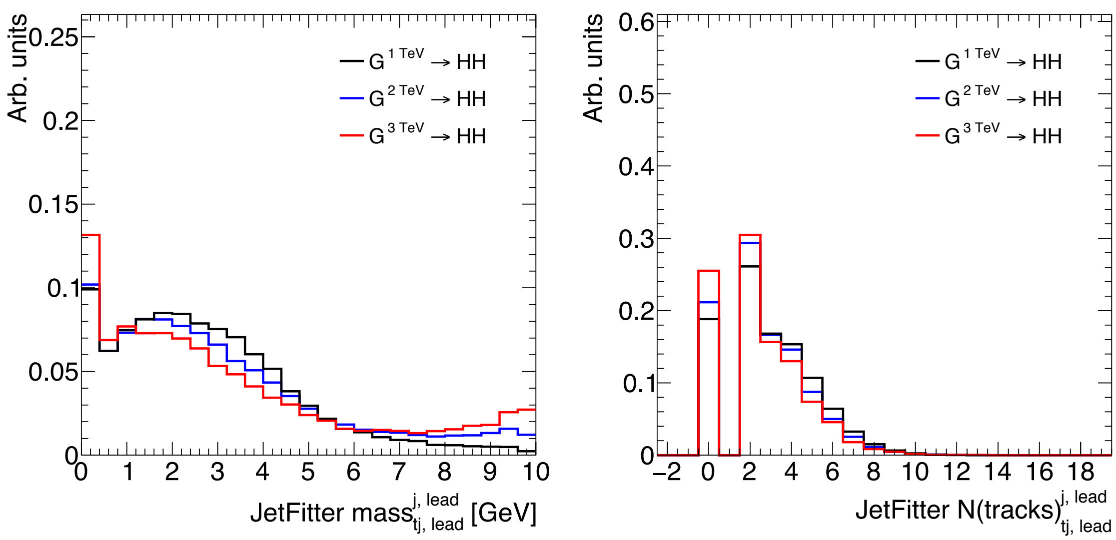
A.2 Effect of multiple b-quarks inside one jet

One hypothesis for why the efficiency of b-tagging the leading track jet degrades is that at high masses, the b quarks get close enough together that both of them are inside of the leading track jet. Because MV2



(a) (b)

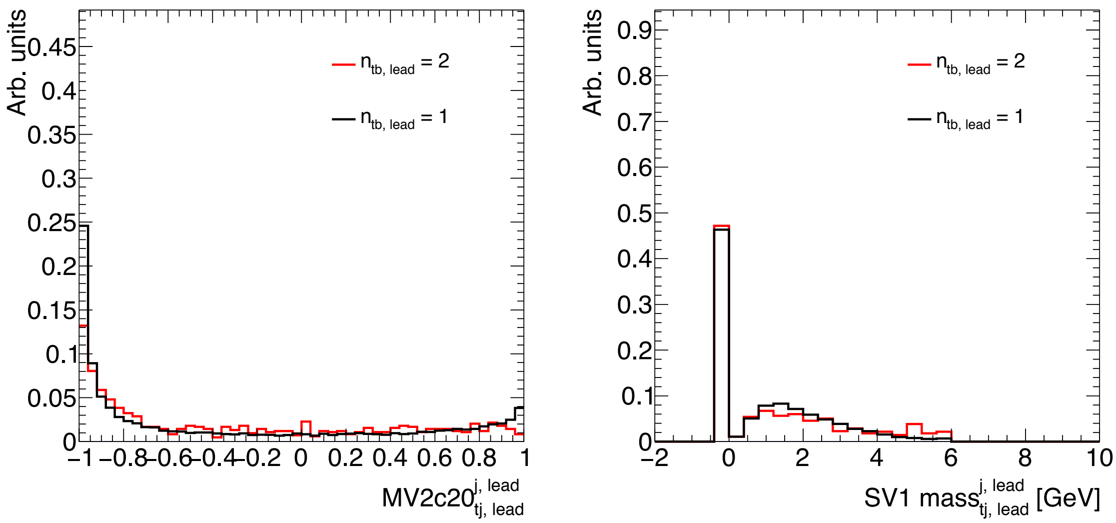
Figure A.4: Mass (a) and number of tracks (b) for the secondary vertices computed with the SV1 algorithm. When no secondary vertex is found, the quantiites are assigned to default negative values.



(a) (b)

Figure A.5: Mass (a) and number of tracks (b) for vertices computed with the JetFitter algorithm. When no vertices are found, the quantities are assigned to default negative values.

is not tuned for tagging multiple b quarks inside one jet, the tagging efficiency could degrade. Figure A.6 shows MV2 scores and SV1 mass for cases where there are two b quarks at truth level within the radius of the leading track jet compared to cases where there is only one true b[1]. This figure suggests that the presence



(a) (b)

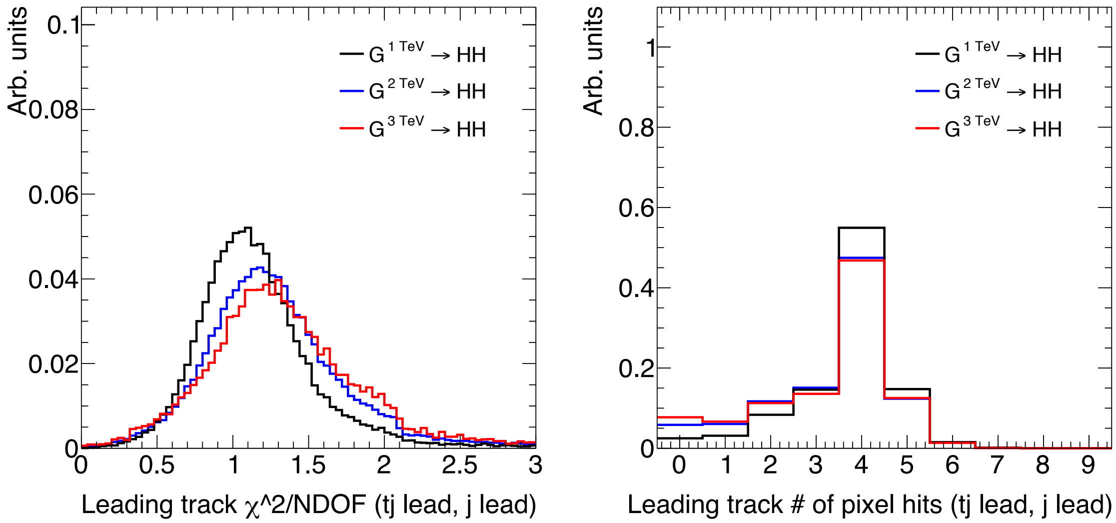
Figure A.6: MV2c20 score (a) and SV1 mass (b) for leading track jets with two truth b quarks ($n_{\mathrm{tb,lead}} = 2$) compared to those with only one truth b ($n_{\mathrm{tb,lead}} = 1$).

of two b-quarks inside the leading jet is not the cause of the degradation in efficiency. There is a change in the shape of the MV2 score distribution, but it is not nearly as pronounced as that seen in A.2 at higher masses. Additionally, the fraction of jets with no secondary vertex found is nearly identical in the track jets with two truth b-quarks.

A.3 Changes in track quality at high p_{T}

Another hypothesis for the degradation of the b-tagging efficiency is a decrease in track quality for high p_{T} b jets. One way to check the overall quality of the tracking inside the jet is to investigate quantities related to the leading track inside of the track jet. Figure A.7 shows the fit χ^2/n_{DOF} and number of hits in the pixel detector for the leading track of the leading track jet. In both cases, the figure shows that in higher mass samples, the quality of the leading track inside of the track jet degrades substantially. The fit quality is lessened and the tracks have less hits in the pixel detector. This is likely due to the fact that at higher p_{T},

[1] When two truth b quarks are required in the leading jet, the subleading jet is required to have zero. When one is required for the leading, one is also required for the subleading.

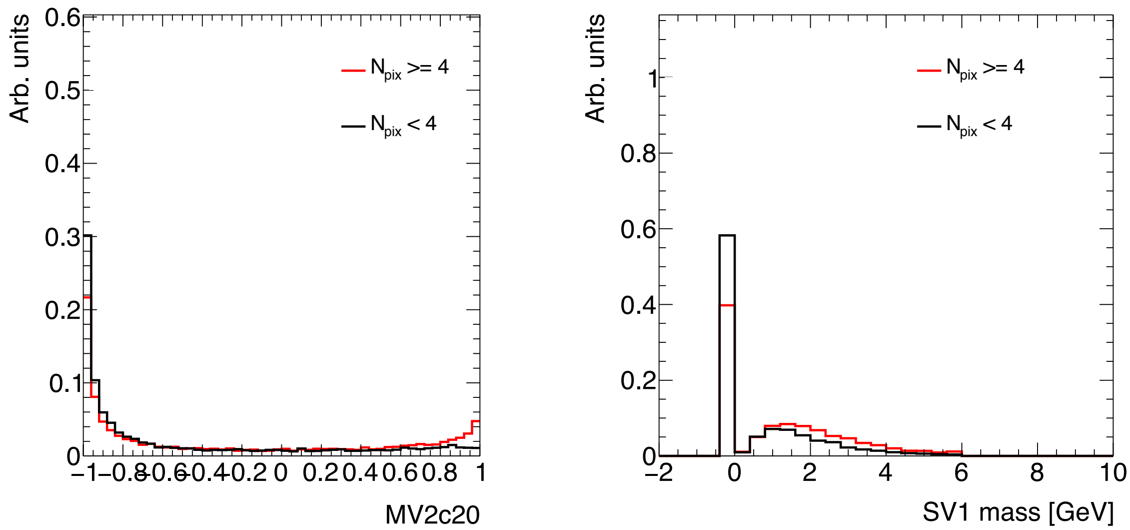


(a) (b)

Figure A.7: Track fit χ^2/n_{DOF} (a) and number of pixel detector hits (b) for the leading track of the leading track jet in different mass RSG $c = 1$ samples

the B-hadron will sometimes live long enough to miss the IBL and first pixel layer, thus decreasing the number of hits on the track.

To check whether this is the cause for the shift in the MV2 score and the higher difficulty in reconstructing secondary vertices, jets whose leading track have at least four pixel hits are compared with those whose tracks have less than four pixel hits. The results for the MV2 score and SV1 mass are shown in figure A.8. Track jets where the leading track does not have at least four pixel hits are more likely to not have a secondary vertex reconstructed. Additionally, their MV2c20 score is shifted more significantly to background-like values. This seems to confirm the hypothesis that degrading track quality is responsible for the lowered b-tagging efficiency at high p_{T}.



(a) (b)

Figure A.8: MV2c20 score (a) and SV1 mass (b) for leading track jets whose leading track jet has at least four pixel hits ($N_{\text{pix}} \geq 4$) compared to those which do not ($N_{\text{pix}} < 4$).

References

[1] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.

[2] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.

[3] David Griffiths. *Introduction to elementary particles*. 2008.

[4] F. Halzen and Alan D. Martin. *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*. 1984. ISBN 0471887412, 9780471887416.

[5] Christopher G. Tully. *Elementary particle physics in a nutshell*. 2011.

[6] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014. doi: 10.1088/1674-1137/38/9/090001.

[7] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014. ISBN 1107034736, 9781107034730. URL http://www.cambridge.org/us/academic/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-field-theory-and-standard-model.

[8] S. Dawson. Introduction to electroweak symmetry breaking. In *High energy physics and cosmology. Proceedings, Summer School, Trieste, Italy, June 29-July 17, 1998*, pages 1–83, 1998. URL http://alice.cern.ch/format/showfull?sysnb=0301862.

[9] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961. doi: 10.1016/0029-5582(61)90469-2.

[10] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967. doi: 10.1103/PhysRevLett.19.1264.

[11] A. Salam. *Elementary Particle Theory*. Almqvist and Wiksell, Stockholm, 1968.

[12] J. Iliopoulos S.L. Glashow and L. Maiani. D2:1285, 1970.

[13] R. Keith Ellis, W. James Stirling, and B. R. Webber. QCD and collider physics. *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, 8:1–435, 1996.

[14] P. W. Higgs. Broken symmetries and the masses of gauge bosons. 13:508, 1964.

[15] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. 145:1156, 1966.

[16] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. 13:321, 1964.

[17] G. S. Guralnik, C. R. Hagen, and T. W. .B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585, 1964. doi: 10.1103/PhysRevLett.13.585.

[18] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. 2013.

[19] Abdelhak Djouadi. The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model. *Phys. Rept.*, 457:1–216, 2008. doi: 10.1016/j.physrep.2007.10.004.

[20] J. Baglio, A. Djouadi, R. Gröber, M. M. Mühlleitner, J. Quevillon, and M. Spira. The measurement of the Higgs self-coupling at the LHC: theoretical status. *JHEP*, 04:151, 2013. doi: 10.1007/JHEP04(2013)151.

[21] Matthew J. Dolan, Christoph Englert, and Michael Spannowsky. New Physics in LHC Higgs boson pair production. *Phys. Rev.*, D87(5):055002, 2013. doi: 10.1103/PhysRevD.87.055002.

[22] Roberto Contino, Margherita Ghezzi, Mauro Moretti, Giuliano Panico, Fulvio Piccinini, and Andrea Wulzer. Anomalous Couplings in Double Higgs Production. *JHEP*, 08:154, 2012. doi: 10.1007/JHEP08(2012)154.

[23] R. Grober and M. Muhlleitner. Composite Higgs Boson Pair Production at the LHC. *JHEP*, 06: 020, 2011. doi: 10.1007/JHEP06(2011)020.

[24] Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999. doi: 10.1103/PhysRevLett.83.3370.

[25] Kaustubh Agashe, Hooman Davoudiasl, Gilad Perez, and Amarjit Soni. Warped Gravitons at the LHC and Beyond. *Phys. Rev.*, D76:036006, 2007. doi: 10.1103/PhysRevD.76.036006.

[26] A. Liam Fitzpatrick, Jared Kaplan, Lisa Randall, and Lian-Tao Wang. Searching for the Kaluza-Klein Graviton in Bulk RS Models. *JHEP*, 09:013, 2007. doi: 10.1088/1126-6708/2007/09/013.

[27] Julien Baglio, Otto Eberhardt, Ulrich Nierste, and Martin Wiebusch. Benchmarks for Higgs Pair Production and Heavy Higgs boson Searches in the Two-Higgs-Doublet Model of Type II. *Phys. Rev.*, D90(1):015008, 2014. doi: 10.1103/PhysRevD.90.015008.

[28] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, Marc Sher, and Joao P. Silva. Theory and phenomenology of two-Higgs-doublet models. *Phys. Rept.*, 516:1–102, 2012. doi: 10.1016/j.physrep.2012.02.002.

[29] Howard E. Haber and Oscar Stål. New LHC benchmarks for the $\mathcal{CP}$ -conserving two-Higgs-doublet model. *Eur. Phys. J.*, C75(10):491, 2015. doi: 10.1140/epjc/s10052-015-3697-x.

[30] Jose M. No and Michael Ramsey-Musolf. Probing the Higgs Portal at the LHC Through Resonant di-Higgs Production. *Phys. Rev.*, D89(9):095031, 2014. doi: 10.1103/PhysRevD.89.095031.

[31] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5:Going Beyond. *JHEP*, 1106:128, 2011. doi: 10.1007/JHEP06(2011)128.

[32] Lyndon R Evans and Philip Bryant. LHC Machine. *J. Instrum.*, 3:S08001. 164 p, 2008. URL https://cds.cern.ch/record/1129806. This report is an abridged version of the LHC Design Report (CERN-2004-003).

[33] ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider. *JINST*, 3: S08003, 2008. doi: 10.1088/1748-0221/3/08/S08003.

[34] CMS Collaboration. The cms experiment at the cern lhc. *Journal of Instrumentation*, 3(08):S08004, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[35] LHCb Collaoration. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi: 10.1088/1748-0221/3/08/S08005.

[36] ALICE Collaboration. The alice experiment at the cern lhc. *Journal of Instrumentation*, 3(08): S08002, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08002.

[37] Lyndon Evans. The Large Hadron Collider. In Holstein, BR and Haxton, WC and Jawahery, A, editor, *ANNUAL REVIEW OF NUCLEAR AND PARTICLE SCIENCE, VOL 61*, volume 61 of *Annual Review of Nuclear and Particle Science*, pages 435–466. 2011. doi: {10.1146/annurev-nucl-102010-130438}.

[38] ATLAS Collaboration. Luminosity Determination in pp Collisions at $\sqrt{s}$ = 7 TeV Using the ATLAS Detector at the LHC. *Eur. Phys. J.*, C 71:1630, 2011. doi: 10.1140/epjc/s10052-011-1630-5.

[39] Mike Lamont for the LHC team. The First Years of LHC Operation for Luminosity Production. International Particle Accelerator Conference, 2013. URL https://accelconf.web.cern.ch/accelconf/IPAC2013/talks/moyab101_talk.pdf.

[40] Paul Collier for the LHC team. LHC Machine Status. CERN Resource Review Board, 2015. URL https://cds.cern.ch/record/2063924/files/CERN-RRB-2015-119.PDF.

[41] Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s}$ = 13 TeV. Technical Report ATL-PHYS-PUB-2015-018, CERN, Geneva, Jul 2015. URL http://cds.cern.ch/record/2037683.

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

[43] ATLAS Collaboration. ATLAS Trigger Operations Public Results. 2015. URL https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults.

[44] ATLAS Collaboration. ATLAS Luminosity Public Results, Run 1. 2012. URL https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[45] ATLAS Collaboration. ATLAS Luminosity Public Results, Run 2. 2015. URL https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[46] T Kawamoto, S Vlachos, L Pontecorvo, J Dubbert, G Mikenberg, P Iengo, C Dallapiccola, C Amelung, L Levinson, R Richter, and D Lellouch. New Small Wheel Technical Design Report. Technical Report CERN-LHCC-2013-006. ATLAS-TDR-020, CERN, Geneva, Jun 2013. URL https://cds.cern.ch/record/1552862. ATLAS New Small Wheel Technical Design Report.

[47] Y. Giomataris, Ph. Rebourgeard, J.P. Robert, and G. Charpak. Micromegas: a high-granularity position-sensitive gaseous detector for high particle-flux environments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 376(1):29 – 35, 1996. ISSN 0168-9002. doi: http://dx.doi.org/10.1016/0168-9002(96)00175-1. URL http://www.sciencedirect.com/science/article/pii/0168900296001751.

[48] T. Alexopoulos, J. Burnens, R. de Oliveira, G. Glonti, O. Pizzirusso, V. Polychronakos, G. Sekhniaidze, G. Tsipolitis, and J. Wotschack. A spark-resistant bulk-micromegas chamber for high-rate applications. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 640(1):110 – 118, 2011. ISSN 0168-9002. doi: http://dx.doi.org/10.1016/j.nima.2011.03.025. URL http://www.sciencedirect.com/science/article/pii/S0168900211005869.

[49] Joao Pequenao and Paul Schaffner. An computer generated image representing how ATLAS detects particles. Jan 2013. URL https://cds.cern.ch/record/1505342.

[50] Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. Technical Report ATLAS-CONF-2012-047, CERN, Geneva, May 2012. URL https://cds.cern.ch/record/1449796.

[51] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. Technical Report ATLAS-CONF-2014-032, CERN, Geneva, Jun 2014. URL https://cds.cern.ch/record/1706245.

[52] Georges Aad et al. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *Eur. Phys. J.*, C74(10):3071, 2014. doi: 10.1140/epjc/s10052-014-3071-4.

[53] Georges Aad et al. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data. *Eur. Phys. J.*, C74(11):3130, 2014. doi: 10.1140/epjc/s10052-014-3130-x.

[54] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008. URL https://cds.cern.ch/record/1099735.

[55] Georges Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. 2016.

[56] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008. doi: 10.1088/1126-6708/2008/04/063.

[57] Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS. Technical Report ATLAS-CONF-2015-037, CERN, Geneva, Aug 2015. URL http://cds.cern.ch/record/2044941.

[58] Georges Aad et al. Performance of *b*-Jet Identification in the ATLAS Experiment. 2015.

[59] Expected performance of the ATLAS *b*-tagging algorithms in Run-2. Technical Report ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul 2015. URL http://cds.cern.ch/record/2037697.

[60] Georges Aad et al. Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS. *Eur. Phys. J.*, C72:1844, 2012. doi: 10.1140/epjc/s10052-011-1844-6.

[61] Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV. Technical Report ATLAS-CONF-2013-082, CERN, Geneva, Aug 2013. URL http://cds.cern.ch/record/1570993.

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

[63] Aaron James Armbruster. Discovery of a Higgs Boson with the ATLAS detector. 2013. CERN-THESIS-2013-047.

[64] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C 71:1554, 2011. doi: 10.1140/epjc/s10052-011-1554-0.

[65] ATLAS Collaboration. Limits on the production of the Standard Model Higgs Boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Eur. Phys. J.*, C 71:1728, 2011. doi: 10.1140/epjc/s10052-011-1728-9.

[66] ATLAS Collaboration. Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s}$ =8 TeV. *Eur. Phys. J. C*, (arXiv:1408.3179. CERN-PH-EP-2014-154):75. 19 p, Aug 2014. URL https://cds.cern.ch/record/1749694.

[67] ATLAS collaboration. Electron trigger performance in 2012 ATLAS data, 2015. ATLAS-COM-DAQ-2015-091.

[68] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[69] B. P. Kersevan and E. Richter-Was. The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5. 2004.

[70] Nikolas Kauer and Giampiero Passarino. Inadequacy of zero-width approximation for a light Higgs boson signal. 2012.

[71] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al. Event generation with SHERPA 1.1. *JHEP*, 0902:007, 2009. doi: 10.1088/1126-6708/2009/02/007.

[72] Michelangelo L. Mangano et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307:001, 2003. doi: 10.1088/1126-6708/2003/07/001.

[73] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006. doi: 10.1088/1126-6708/2006/05/026.

[74] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput.Phys.Commun.*, 178:852–867, 2008. doi: 10.1016/j.cpc.2008.01.036.

[75] G. Corcella et al. HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including super-symmetric processes) . *JHEP*, 01:010, 2001. doi: 10.1088/1126-6708/2001/01/010.

[76] J. M. Butterworth, Jeffrey R. Forshaw, and M. H. Seymour. Multiparton interactions in photoproduction at HERA. *Z. Phys.*, C 72:637, 1996. doi: 10.1007/s002880050286.

[77] Jun Gao, Marco Guzzi, Joey Huston, Hung-Liang Lai, Zhao Li, et al. The CT10 NNLO Global Analysis of QCD. *Phys.Rev.*, D89:033009, 2014. doi: 10.1103/PhysRevD.89.033009.

[78] P. M. Nadolsky. Implications of CTEQ global analysis for collider observables. *Phys. Rev.*, D 78: 013004, 2008. doi: 10.1103/PhysRevD.78.013004.

[79] A. Sherstnev and R. S. Thorne. Parton distributions for the LHC. *Eur. Phys. J*, C 55:553, 2009. doi: 10.1140/epjc/s10052-008-0610-x.

[80] S. Agostinelli et al. Geant4, a simulation toolkit. *Nucl. Instrum. Meth.*, A 506:250, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[81] R.K. Ellis, I. Hinchliffe, M. Soldate, and J.J. Van Der Bij. Higgs decay to τ+τ−a possible signature of intermediate mass higgs bosons at high energy hadron colliders. *Nuclear Physics B*, 297(2):221 – 243, 1988. ISSN 0550-3213. doi: http://dx.doi.org/10.1016/0550-3213(88)90019-3. URL http://www.sciencedirect.com/science/article/pii/0550321388900193.

[82] Eilam Gross and Ofer Vitells. Transverse mass observables for charged Higgs boson searches at hadron colliders. *Phys. Rev.*, D81:055010, 2010. doi: 10.1103/PhysRevD.81.055010.

[83] J. R. Andersen et al. Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report. 2014.

[84] I. Stewart and F. Tackmann. Theory uncertainties for Higgs mass and other searches using jet bins. *Phys. Rev.*, D 85:034011, 2012. doi: 10.1103/PhysRevD.85.034011.

[85] ATLAS Collaboration. Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC. *Eur. Phys. J.*, C 71:1630, 2011. doi: 10.1140/epjc/s10052-011-1630-5.

[86] Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ tev with atlas 2011 data. *ATLAS-CONF-2013-004*, 2013.

[87] Calibrating the b-tag efficiency and mistag rate in 35 pb^{-1} of data with the atlas detector. *ATLAS-CONF-2011-089*, 2011.

[88] ATLAS Collaboration. Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector. *ATLAS-CONF-2012-043*, 2012. URL http://cdsweb.cern.ch/record/1435197.

[89] ATLAS Collaboration. Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment. (ATLAS-CONF-2014-004), 2014. URL http://cds.cern.ch/record/1664335.

[90] Georges Aad et al. Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data. *Phys. Rev.*, D90(5):052004, 2014. doi: 10.1103/PhysRevD.90.052004.

[91] Georges Aad et al. Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment. *Eur. Phys. J.*, C76(1):6, 2016. doi: 10.1140/epjc/s10052-015-3769-y.

[92] W.J. Stirling. 7/8 and 13/8 TeV LHC luminosity ratios. 2013. URL http://www.hep.ph.ic.ac.uk/~wstirlin/plots/lhclumi7813_2013_v0.pdf.

[93] J Alison. Experimental Studies of hh. Oct 2014. URL http://cds.cern.ch/record/1952581.

[94] 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. *JHEP*, 07:079, 2014.

[95] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys. B*, 867:244, 2013.

[96] ATLAS Collaboration. ATLAS Run 1 Pythia8 tunes. (ATL-PHYS-PUB-2014-021), Nov 2014. URL https://cds.cern.ch/record/1966419.

[97] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J. C*, 58:639–707, 2008. doi: 10.1140/epjc/s10052-008-0798-9.

[98] Stefan Gieseke, Christian Rohr, and Andrzej Siodmok. Colour reconnections in Herwig++. *Eur. Phys. J. C*, 72:2225, 2012. doi: 10.1140/epjc/s10052-012-2225-5.

[99] 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.

[100] Peter Zeiler Skands. Tuning Monte Carlo Generators: The Perugia Tunes. *Phys. Rev. D*, 82:074018, 2010. doi: 10.1103/PhysRevD.82.074018.

[101] Michal Czakon and Alexander Mitov. Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders. 2011.

[102] Baojia (Tony) Tong. Private communication.

[103] D. Krohn, J. Thaler, and L.-T. Wang. Jet Trimming. *JHEP*, 02:084, 2010. doi: 10.1007/JHEP02(2010)084.

[104] ATLAS Collaboration. Identification of Boosted, Hadronically Decaying W Bosons and Comparisons with ATLAS Data Taken at $\sqrt{s} = 8$ TeV. 2015.

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

[106] Flavor Tagging with Track Jets in Boosted Topologies with the ATLAS Detector. Technical Report ATL-PHYS-PUB-2014-013, CERN, Geneva, Aug 2014. URL https://cds.cern.ch/record/1750681.

[107] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett. B*, 659:119, 2008. doi: 10.1016/j.physletb.2007.09.077.

[108] Glen Cowan, Eilam Gross. Discovery significance with statistical uncertainty in the background estimate. 2008. URL http://www.pp.rhul.ac.uk/~cowan/stat/notes/SigCalcNote.pdf.

[109] Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2016-017, CERN, Geneva, Mar 2016. URL https://cds.cern.ch/record/2141006.

[110] Qi Zeng. Private communication.

[111] ATLAS Collaboration. Identification of boosted, hadronically-decaying W and Z bosons in $\sqrt{s} = 13$ TeV Monte Carlo Simulations for ATLAS. (ATL-PHYS-PUB-2015-033), Aug 2015. URL https://cds.cern.ch/record/2041461.

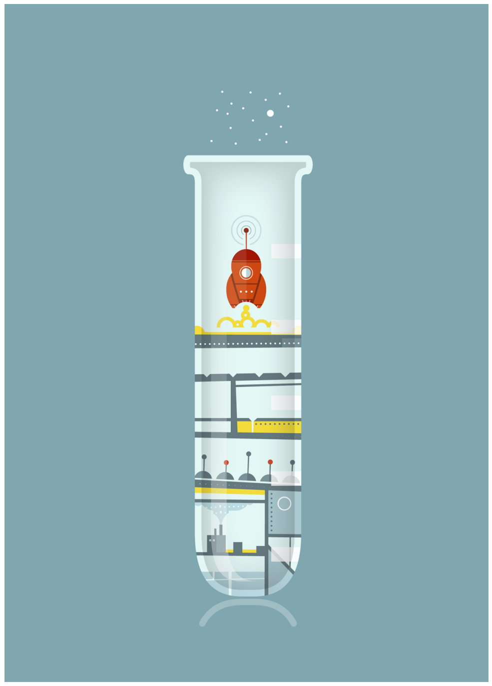
[112] ATLAS Collaboration. Performance of b-Jet Identification in the ATLAS Experiment. 2015.

[113] Alexander L. Read. Presentation of search results: The CL(s) technique. *J. Phys. G*, 28:2693, 2002. doi: 10.1088/0954-3899/28/10/313.

[114] Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV. Technical Report ATLAS-CONF-2015-044, CERN, Geneva, Sep 2015. URL http://cds.cern.ch/record/2052552.

[115] Projections for measurements of Higgs boson signal strengths and coupling parameters with the ATLAS detector at a HL-LHC. Technical Report ATL-PHYS-PUB-2014-016, CERN, Geneva, Oct 2014. URL http://cds.cern.ch/record/1956710.

[116] ATLAS Phase-II Upgrade Scoping Document. Technical Report CERN-LHCC-2015-020. LHCC-G-166, CERN, Geneva, Sep 2015. URL http://cds.cern.ch/record/2055248.



THIS THESIS WAS TYPESET using LaTeX,
originally developed by Leslie Lamport
and based on Donald Knuth's TeX. The
body text is set in 11 point Egenolff-Berner Gara-
mond, a revival of Claude Garamont's human-
ist typeface. The above illustration, *Science Ex-
periment 02*, was created by Ben Schlitter and
2962 released under CC BY-NC-ND 3.0. A template
that can be used to format a PhD dissertation
with this look & feel has been released under the
permissive AGPL license, and can be found on-
line at github.com/asm-products/Dissertate or
from its lead author, Jordan Suchow, at su-
chow@post.harvard.edu.