THE CONSTRUCTION OF A SCALAR EXTENSION TO THE STANDARD MODEL AND THE SEARCH FOR A HEAVY SCALAR AT $\sqrt{s} = 13$ TEV WITH THE ATLAS DETECTOR

STEFAN VON BUDDENBROCK

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

School of Physics

Faculty of Science

University of the Witwatersrand

20 April 2020, Johannesburg, South Africa

SUPERVISORS:

Bruce Mellado

Alan S. Cornell

"An important part of good scientific strategy is to distinguish between problem areas that might be ripe for grand synthesis, and problem areas where a more opportunistic approach will be more fruitful. A successful theory of *something* can be more valuable than an attempted Theory of *Everything*."

Frank Wilczek (*A Beautiful Question*, 2015)

DECLARATION

I, Stefan von Buddenbrock (Student number: 567245), am a student registered for the degree of Doctor of Philosophy (Physics) in the academic year 2019. I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have written about and presented many of the ideas and results in this thesis. A list of my previous works that have done so is shown on page ix. I, along with the other co-authors, have intellectual property of these works.

Johannesburg, South Africa

Stefan von Buddenbrock,

20 April 2020

ABSTRACT

Searches for physics beyond the Standard Model (SM) have always been a focus of the particle physics community, more so since the discovery of a Higgs-like boson (h) in 2012. These are typically done using either the top-down or bottom-up methodology. In this thesis, a model is constructed in order to successfully explain a variety of anomalous results from Run 1 and 2 of the Large Hadron Collider (LHC), using a hybrid methodology that iterates between theory and experiment. Initially in 2015, a new heavy scalar boson H was postulated to explain anomalies in the Run 1 LHC data. The result of this early study implied that a heavy scalar boson with a mass around 270 GeV, produced through gluon fusion, could explain these anomalies with a significance of 3σ, with a cross section comparable with that of a heavy Higgs boson as well as a dominant h-associated decay mode. Theoretical developments of the model then hinted towards the existence of an additional scalar boson S, which acted as this associated decay product. The S boson, with a mass of around 150 GeV, was likely to decay to final states comprising of multiple leptons. The culmination of these studies found that the $H \to Sh$ decay mode was dominant and could be searched for in final states with multiple leptons and b-tagged jets at the LHC. In 2017 and 2018, using the input from these initial studies, anomalies in the LHC data were successfully explained by the model, as predicted. It was found that the new physics model improved on the SM description of the data at the level of at least 8σ with just a single degree of freedom. Should these discrepancies not be resolved, this can be considered as indirect evidence for new physics processes at the LHC, since the current SM tools are unable to provide a reasonable explanation for the anomalies. The potential of correlating this result with a mild excess at 245 GeV in the search for H decaying to four leptons in the ATLAS detector is considered, and future potential developments of the model are discussed.

PUBLICATIONS

The work presented in this thesis has also been presented and published in a number of peer-reviewed journal articles and conference outputs, as well as some unpublished reports. The ideas and results in these works will therefore be of a similar nature to those presented here. A list of these publications is given below:

JOURNAL ARTICLES:

- S. Buddenbrock, A. S. Cornell, Y. Fang, A. Fadol Mohammed, M. Kumar, B. Mellado, and K. G. Tomiwa, "The emergence of multi-lepton anomalies at the LHC and their compatibility with new physics at the EW scale," *JHEP* **10** (2019) 157, arXiv:1901.05300 [hep-ph]

- S. von Buddenbrock, A. S. Cornell, E. D. R. Iarilala, M. Kumar, B. Mellado, X. Ruan, and E. M. Shrif, "Constraints on a 2HDM with a singlet scalar and implications in the search for heavy bosons at the LHC," *J. Phys.* **G46** no. 11, (2019) , arXiv:1809.06344 [hep-ph]

- **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for Higgs boson pair production in the $WW^{(*)}WW^{(*)}$ decay channel using ATLAS data recorded at $\sqrt{s} = 13$ TeV," *JHEP* **05** (2019) 124, arXiv:1811.11028 [hep-ex]

- **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for dark matter in association with a Higgs boson decaying to two photons at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *Phys. Rev.* **D96** no. 11, (2017) 112004, arXiv:1706.03948 [hep-ex]

- S. von Buddenbrock, A. S. Cornell, A. Fadol, M. Kumar, B. Mellado, and X. Ruan, "Multi-lepton signatures of additional scalar bosons beyond the Standard Model at the LHC," *J. Phys.* **G45** no. 11, (2018) 115003, arXiv:1711.07874 [hep-ph]

- S. von Buddenbrock, N. Chakrabarty, A. S. Cornell, D. Kar, M. Kumar, T. Mandal, B. Mellado, B. Mukhopadhyaya, R. G. Reed, and X. Ruan, "Phenomenological signatures of additional scalar bosons at the LHC," *Eur. Phys. J.* **C76** no. 10, (2016) 580, arXiv:1606.01674 [hep-ph]

- S. von Buddenbrock, N. Chakrabarty, A. S. Cornell, D. Kar, M. Kumar, T. Mandal, B. Mellado, B. Mukhopadhyaya, and R. G. Reed, "The compatibility of LHC Run 1 data with a heavy scalar of mass around 270 GeV," arXiv:1506.00612 [hep-ph]

CONFERENCE PROCEEDINGS:

- S. von Buddenbrock and B. Mellado, "Anomalies in the production of multiple leptons at the LHC," arXiv:1909.04370 [hep-ph]
- S. von Buddenbrock, "Di-Higgs production and the Madala Hypothesis," in *Proceedings of the HEPP Workshop 2018*, A. Cornell and B. Mellado, eds. https://indico.cern.ch/event/682259/. Accepted for publication
- S. von Buddenbrock, A. S. Cornell, M. Kumar, and B. Mellado, "The Madala hypothesis with Run 1 and 2 data at the LHC," *J. Phys. Conf. Ser.* **889** no. 1, (2017) 012020, arXiv:1709.09419 [hep-ph]
- S. von Buddenbrock, "Exploring LHC Run 1 and 2 data using the Madala hypothesis," *J. Phys. Conf. Ser.* **878** no. 1, (2017) 012030, arXiv:1706.02477 [hep-ph]
- S. von Buddenbrock, "Production of the Madala boson in association with top quarks," in *The Proceedings of SAIP2017, the 62nd Annual Conference of the South African Institute of Physics*, J. Engelbrecht, ed., pp. 187–192. ISBN: 978-0-620-82077-6
- S. von Buddenbrock *et al.*, "Constraining hypothetical extensions to the Higgs sector at the LHC," in *The Proceedings of SAIP2016, the 61st Annual Conference of the South African Institute of Physics*, S. Peterson and S. Yacoob, eds., pp. 158–163. ISBN: 978-0-620-77094-1

INTERNAL NOTES:

- S. H. Abidi *et al.*, "Search for heavy ZZ resonances in the $\ell\ell\ell\ell$ final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-ANA-HIGG-2018-09, CERN, Geneva, Sep, 2018
- J. Adelman *et al.*, "Supporting note: Selection and performance for the $h \to \gamma\gamma$ and $h \to Z\gamma$ analyses, Spring 2017," Tech. Rep. ATL-COM-PHYS-2017-357, CERN, Geneva, Apr, 2017
- S. von Buddenbrock *et al.*, "Supporting note: Search for dark matter in events with missing transverse momentum and a Higgs boson decaying to two photons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector using data in full years of 2015 and 2016," Tech. Rep. ATL-COM-PHYS-2016-1735, CERN, Geneva, Nov, 2016
- X. Ruan *et al.*, "Search for a Higgs boson with the di-photon 2 decay in association with missing energy in p p collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-1732, CERN, Geneva, Nov, 2016
- S. von Buddenbrock *et al.*, "Search for resonant pair production in $H \to SS, Sh, hh$ channels decaying to $4W$ with a four lepton final state at $\sqrt{s} = 13$ TeV in the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-1622, CERN, Geneva, Feb, 2017
- Y. Fang *et al.*, "Search for Higgs boson pair production in the final state of $\gamma\gamma WW^*(\to lvjj)$ using 36.1 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13$ TeV

with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-1406, CERN, Geneva, Sep, 2016

- L. Cerda Alberich *et al.*, "Supporting note: Selection and performance for the $h \to \gamma\gamma$ and $h \to Z\gamma$ analyses," Tech. Rep. ATL-COM-PHYS-2016-862, CERN, Geneva, Jul, 2016

- S. von Buddenbrock *et al.*, "Supporting note: Search for dark matter in events with missing transverse momentum and a Higgs boson decaying to two photons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-708, CERN, Geneva, Jun, 2016

- S. von Buddenbrock *et al.*, "Supporting note: Search for a generic heavy boson using $\gamma\gamma + X$ in p p collisions at$\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-473, CERN, Geneva, May, 2016

- E. Petit *et al.*, "Supporting note: Selection and performance for the $h \to \gamma\gamma$ and $h \to Z\gamma$ analyses," Tech. Rep. ATL-COM-PHYS-2016-109, CERN, Geneva, Feb, 2016

- X. Ruan *et al.*, "Search for a Higgs boson with the di-photon decay in association with missing energy in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2015-1589, CERN, Geneva, Dec, 2015

- J. Adelman *et al.*, "Search for resonant and enhanced non-resonant dihiggs production in the yybb channel with XXX fb-1 of data at 13 TeV," Tech. Rep. ATL-COM-PHYS-2015-1073, CERN, Geneva, Aug, 2015

ACKNOWLEDGEMENTS

First and foremost, many thanks go to Prof. Bruce Mellado and Prof. Alan S. Cornell, the supervisors of this work. In addition to providing guidance and a focus for this thesis, they were also helpful in accomplishing the additional task of completing paperwork and securing the funding necessary for a relatively comfortable time as a postgraduate student. Other members of the Wits staff in the School of Physics have also been helpful and gone out of their way to assist me. These are Dr. Deepak Kar, Dr. Mukesh Kumar, Dr. Xifeng Ruan and Dr. Jong Soo Kim.

The first part of this thesis deals with constructing a model based on a few experimental observations. For this task, a small group from the Harish-Chandra Research Institute (HRI) are acknowledged for collaborating with the team at Wits, particularly in the area of providing the theoretical physics tools and insights for this model. The group consisted of Prof. Biswarup Mukhopadhyaya, Dr. Tanumoy Mandal and Dr. Nabarun Chakrabarty. Without their helpful input and discussions, the project would not have moved forward as robustly as it did.

The second part of this thesis is a description of some experimental work done with the ATLAS collaboration. In particular, I would like to acknowledge Dr. Theodota Lagouri for allowing me a platform to join the ATLAS *HZZ* group, Dr. Simone Amoroso for helping me with my qualification task, and Dr. Rosy Nikolaidou and Dr. Lailin Xu for sub-convening the high-mass four lepton group during my time working in it.

The agencies and organisations that supported this work are also acknowledged for their willingness to fund my studies and the research at large. These are the National Research Foundation (NRF) and the National Institute for Theoretical Physics (NITheP). For funding my travels to CERN and various conferences and workshops around the world, I also am grateful to the SA-CERN consortium. With-

out the support of these funding organisations, my PhD work would have been either impossible or excruciatingly difficult.

Finally, I acknowledge the support of my family during my time as a student. It's easy to forget that life would not be possible without those closest to me: my wife, parents and brothers. Without their support I would never have been able to gain a voice for communicating the scientific ideas contained in this thesis. They don't know it, but they are the most important people in the process.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

E_{T} – transverse energy
$E_{\mathrm{T}}^{\mathrm{miss}}$ – missing transverse energy
η – pseudo-rapidity
$\langle\mu\rangle$ – average number of interactions per bunch crossing
ggF – gluon fusion
m_{T} – transverse mass
$m_{4\ell}$ – four-lepton invariant mass
p_{T} – transverse momentum
pp – proton-proton
$m_{\ell\ell}$ – di-lepton invariant mass
2HDM – two Higgs doublet model

ALICE – A Large Ion Collider Experiment
ATLAS – A Toroidal LHC Apparatus

BEH – Brout-Englert-Higgs
BR – branching ratio
BSM – beyond the Standard Model

CB – Crystal Ball
CERN – European Organisation for Nuclear Research
CL – confidence level
CMS – Compact Muon Solenoid
CR – control region

DM – dark matter

ECAL – electromagnetic calorimeter
EFT – effective field theory
EM – Electromagnetic
EW – Electroweak
EWSB – electroweak symmetry breaking

FCNC – flavour-changing neutral current
FS – full simulation
FSR – final state radiation

HCAL – hadronic calorimeter
HLT – high level trigger

ID – inner detector
ISR – initial state radiation

JO – job options

KS – Kolmogorov-Smirnov

L1 – Level 1
LAr – liquid argon
LEP – Large Electron-Positron Collider
LHC – Large Hadron Collider
LHC-HXSWG – LHC Higgs cross section working group
LO – leading order

MC – Monte Carlo
MS – Muon Spectrometer

N3LO – next-to-next-to-next-to leading order
NLL – negative log likelihood
NLO – next-to leading order
NNLL – next-to-next-to leading logarithmic
NNLO – next-to-next-to leading order
NP – nuisance parameter

OS – opposite-sign

PD – Pixel Detector
PDF – parton density function
PID – particle identification
PS – parton shower
PSB – Proton Synchrotron Booster

QCD – Quantum Chromodynamics
QED – Quantum Electrodynamics
QFT – quantum field theory

SCT – Semiconductor Tracker
SFOS – same-flavour opposite-sign
SLHA – SUSY Les-Houches Accord
SM – Standard Model
SPS – Super Proton Synchrotron
SR – signal region
SS – same-sign
SUSY – supersymmetry

TileCal – Tile Calorimeter
TRT – Transition Radiation Tracker

UFO – Universal FeynRules Output
UV – ultraviolet

VBF – vector boson fusion
VEV – vacuum expectation value
VLQ – vector-like quark

1

INTRODUCTION

1.1 THESIS CONTEXT AND METHODOLOGY

At the time of writing, the most recent major discovery in particle physics is that of the Standard Model (SM) Higgs boson (h). In July of 2012, a particle consistent with the properties of the SM Higgs boson was discovered by both the A Toroidal LHC Apparatus (ATLAS) [1] and the Compact Muon Solenoid (CMS) [2] collaborations at the Large Hadron Collider (LHC).

The discovery was an archetype for one of the natural relationships we find in science, that is, the relationship between theory and experiment. This is most succinctly seen in the practical methodology used by the collaborations in order to make the discovery. A useful variable for measuring the rate of production for the Higgs boson, used by both ATLAS and CMS, is that of a *signal strength*:

$$\mu = \frac{\sigma_{\text{obs}}}{\sigma_{\text{SM}}}. \tag{1.1}$$

Here, the numerator is proportional to the number of events experimentally measured for the Higgs production process, while the denominator is proportional to the expected theoretical prediction thereof. Therefore, the experimental collaborations could conclude that a Higgs boson exists in nature if Equation (1.1) was statistically compatible with the value of 1. In doing so, the theoretical prediction of the SM Higgs boson was brought to reality by its experimental observation.

More generally speaking, the example in Equation (1.1) reflects a fundamental principle defined by the scientific method; a process that has moved science forward over the last few centuries. The core of the idea is that developments in theoretical structures are driven by experimentally observable evidence through

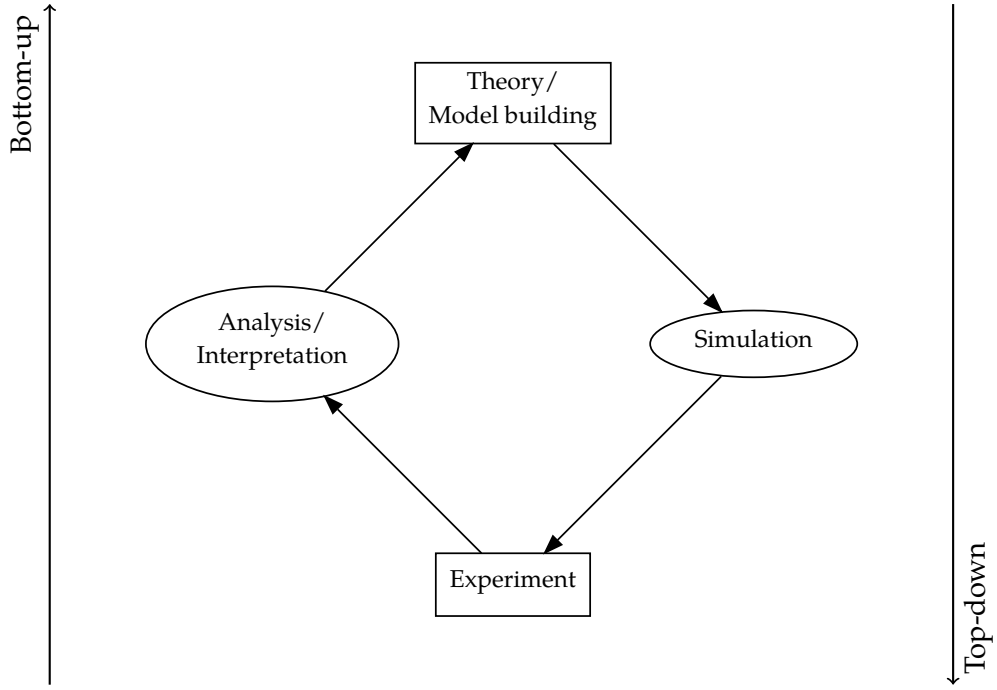


Figure 1: An illustration of the iterative process that underpins scientific progress. One typically starts the process in one of the squares, i. e. *theory* or *experiment*.

analysis and interpretation of data, while experimental observations develop in such a way as to test theoretical predictions. This principle is shown in Figure 1. The iterative process shows how to get from theory to experiment (through simulations and calculations) and from experiment to theory (through analysis and interpretation). Since science endeavours to organise and provide explanations for the observable phenomena seen in the universe, the loop shown in Figure 1 is a natural consequence for the method by which science is done. Each of the four categories shown – theory, simulation, experiment and analysis/interpretation – has a relatively well-defined structure in particle physics, and shall be discussed below.

The core *theory* that is currently accepted in particle physics is the SM. Since its structure is central to this thesis, it will be discussed in detail in Chapter 2. Its structure is well established, yet still incomplete due to the fact that it cannot explain several naturally observed phenomena. Therefore, most of the theory done in particle physics is the development of models beyond the Standard Model (BSM) in hopes that a certain new physics model can extend our current understanding

of particle physics. There exist several computational tools that assist with the development of these model building efforts. Some of these will be discussed where necessary in the main content of the thesis.

Simulation in particle physics, which is often referred to as Monte Carlo (MC) simulation, is a major aspect of the field. As mentioned above, given a model that is built (whether it is the SM or some BSM physics), one can use a variety of event generators to simulate what the model would produce in a particle collider like the LHC. These event generators predominantly rely on an MC integration technique, such that absolute and differential scattering cross sections can be calculated. Furthermore, processes like the parton shower (PS), hadronisation of unstable quarks and gluons, and radiation are simulated by certain event generators as well. As shall be seen, MC simulations play a major role in this thesis, and so the various tools used will be discussed where appropriate.

In particle physics, an *experiment* usually refers to a particle collider and/or the detector(s) that collect the data produced by the colliding particles. In the context of this thesis, the experiments that are considered are the detectors that collect data from proton-proton collisions at the LHC. Indeed, the LHC is the largest and most ambitious particle collider built to date – even arguably the biggest scientific experiment ever performed in high energy physics. Being a member of the ATLAS collaboration, most of the experimental work presented in this thesis will relate to it. However, it shall be seen that the CMS collaboration results are also often considered, especially in constraining the parameters of the hypothetical model described in this thesis. The ATLAS detector and LHC shall be described in more detail in Chapter 3.

The *analysis* of data, like simulation, is also a major part of particle physics. Traditionally, an analysis refers to the breaking up of a concept into small pieces and understanding how the small pieces behave (under the assumption that one can use the small pieces to recreate the full concept, i. e. *synthesis*). In particle physics, this most often involves separating processes into *events*, each of which usually represents a single particle collision. Working iteratively with events, quantities can be calculated and modified according to what is needed, and then the ensemble of

events can be recombined and visualised using histograms, etc. Analysis can also involve interpretation of experimental data, which makes use of statistical tools. Again, there are many computational tools that can be used to do analysis work and interpretation through statistics. This will also form a large part of the thesis, and so will be discussed where appropriate.

On closer inspection of Figure 1, one should see that the loop necessarily has a preferred direction (clockwise). One cannot use experimental data to simulate a theory, and it would be absurd to assume that analysing a theory could define experimental data! Note also that theory and experiment are starting points in this process; simulation and analysis are methods by which one can go from theory to experiment and vice versa. We can therefore define two different approaches to doing science: *top-down* and *bottom-up*. Using Figure 1 as a reference, bottom-up means starting with experimental data, and analysing and interpreting it in order to make a statement about a model or theory, including when experiments observe new phenomena not predicted by any theory. On the other hand, top-down means starting by developing a self-consistent theory or model and simulating it in order to be tested against experimental data.

Neither of these approaches is solely favoured in the pursuit of scientific advancement. Each approach is strictly one-sided and tends to treat the target as a passive element of the process, thereby damaging the connection between both approaches, which is essential to push scientific endeavours forward. A far more pragmatic method is to incorporate both approaches into a continual and iterative process around the loop in Figure 1. In doing so, one is far less likely to neglect the important connection between theory and experiment, and will remain aware of both throughout the entire procedure. The methodology pursued in this thesis follows this hybrid iterative approach, and therefore the content of the thesis is driven both by theoretical and experimental considerations.

1.2 RESEARCH QUESTION AND THESIS OUTLINE

It is well understood that a Higgs boson was the last "piece of the puzzle" for the SM, in that there are currently no major undiscovered predictions made by the SM. With the SM complete, there are currently two major classes of research (or *research questions*) in high energy physics. The first involves refining existing measurements of what is known in the SM, and the second involves searching for completely new physics that is not predicted. This thesis focuses on the latter, and an appropriate research question is:

> *Given the current reach of physics results from high energy physics experiments, to what extent can anomalous features of the data be explained and constrained by a simplified scalar extension to the Standard Model?*

The first part of the research question invokes the "current reach of physics results from high energy physics experiments," as well as "anomalous features of the data." In particular, this refers to physics results published by the experimental collaborations at the LHC. The current set of published results are mostly consistent with the SM. However, there are apparent deviations in certain corners of the phase space; these deviations have been studied at length in References [3, 4, 5]. These anomalies in the LHC data served as the motivation for the construction of a simplified scalar extension to the SM, which will henceforth be referred to as "the BSM model," unless otherwise specified.[1] Since the model was built from an interpretation of the data, the process was not straightforward and required several iterative steps to complete through the hybrid top-down and bottom-up methodology discussed in Section 1.1.

The initial stages of this process were performed in 2015, and are discussed in Chapter 4. Here, a simplified model hypothesising the existence of a heavy scalar H was developed to explain some anomalous data in Run 1 of the LHC. The result of this study constrained the mass of H to around 270 GeV, with a significance of 3σ [3, 6], where one σ represents the statistical distance of one standard deviation

1 Informally, the model and its proposed inputs have been referred to in public as the *Madala hypothesis*.

between the new physics hypothesis and the null hypothesis. The model evolved towards a more formal structure in 2016 [7], where an additional scalar S was introduced to explain theoretical problems in the model's original form. These developments are detailed in Chapter 5. One key prediction of this development was the expected production of multiple leptons at the LHC, with and without additional b-tagged jets. With this in mind, all but one of the model parameters were fixed and a set of several LHC results in the multiple lepton sector were fit with the model [4, 5, 8]. The result of this study had shown that the LHC data in the multiple lepton sector were incompatible with the SM predictions at the level of at least 8σ. The BSM model was found to perform well with just one degree of freedom and can be thought of as indirect evidence for new physics at the LHC. It was also found that the model is made to be too limited to provide a consistent explanation for certain anomalies in the LHC data, thereby indicating exciting new prospects for the future development of the model. The bulk of work that produced these results is shown in Chapter 6.

This thesis also contains work done as part of the ATLAS collaboration, which relates to the search for a heavy scalar boson. In Chapter 7, the details of the search for a high mass scalar boson decaying to the four lepton final state are presented with an integrated luminosity of 79.8 fb^{-1} of data as collected by the ATLAS detector. A mild excess of events with an invariant mass of 245 GeV can be seen in the data. This result is considered in the context of the 270 GeV excess mentioned above, where the potential of H having a large width is discussed. Concluding remarks are made in Chapter 8.

2

THE STANDARD MODEL OF PARTICLE PHYSICS

The study of particle physics is centered on the understanding of the properties of elementary particles and the means by which they interact. Our current knowledge of particle physics is neatly contained in the SM of particle physics. With the discovery of the Higgs boson in 2012 by the ATLAS [1] and CMS [2] collaborations, there are currently no particles in the SM which have not been experimentally observed. These experimental apparatus shall be described within the context of the LHC in Chapter 3, whereas this chapter will outline the theoretical structure of the SM.

Under the theoretical framework of quantum field theory (QFT), the SM describes the nature of every field (and therefore every particle by extension) which has been observed in nature. In QFT, particles are believed to be the result of propagating excitations in a number of fundamental fields that span space and time. The observation of particles therefore allows us to probe the nature of the fields from which they emerge, which allows us to further understand the fundamental constituents of the universe.

2.1 PARTICLES IN THE STANDARD MODEL

To understand the significance of the different particles in the SM, one can divide them up into two different types: fermions and bosons. These two descriptions not only make a statement about particle spin, but also the way particles arrange themselves in large ensembles and interact with each other. This has implications on the way we perceive and interact with the physical world. However, the particles in the SM are involved in much smaller scale physics, and allow us to explore the "invisible" subatomic world.

	Quarks		
Up-type	Up quark (u)	Charm quark (c)	Top quark (t)
	($2.3^{+0.7}_{-0.5}$ MeV)	(1.275 ± 0.025 GeV)	(173.21 ± 0.87 GeV)
Down-type	Down quark (d)	Strange quark (s)	Bottom quark (b)
	($4.8^{+0.5}_{-0.3}$ MeV)	(95 ± 5 MeV)	(4.66 ± 0.03 GeV)
	Leptons		
Charged	Electron (e)	Muon (μ)	Tau (τ)
	(0.51100 MeV)	(105.66 MeV)	(1776.86 ± 0.12 MeV)
Neutral	Electron neutrino (ν_e)	Muon neutrino (ν_μ)	Tau neutrino (ν_τ)
	(< 2 eV)	(< 2 eV)	(< 2 eV)

Table 1: The fermionic particle content currently explained in the SM, with their associated masses in parentheses. All quoted masses are taken from Reference [9]. The errors have been omitted for significantly precise measurements.

Fermions have half-integer spin ($\frac{1}{2}$, $\frac{3}{2}$, etc.) and have quantum mechanical wave-functions which are anti-symmetric under an exchange of two such particles, which carries the implication that two fermions cannot occupy the same quantum state concurrently. It seems intuitive then that fermionic particles make up the matter which we are familiar with on macroscopic scales. This is true in the SM, where our understanding is that matter is almost entirely made up of electrons orbiting bound states of quarks (that is, protons and neutrons).

The fermionic particles of the SM are listed with their properties in Table 1. Freely moving leptons are in principle observable in nature, but free quarks are not. The quarks have been shown to obey a principle called "confinement", and naturally cluster up into groups to make hadrons (such as protons, neutrons, pions, etc.). This is intimately linked to the fact that quarks carry a colour charge (which is separate from their electric charge), a charge that was deduced from the patterns by which the quarks tend to cluster [10].

For each fermion in the SM, there exists an equally massive anti-particle associated with it, which has inverted internal charges. The fermions in the SM are chiral in nature and can have left-handed and right-handed projections, a measure of whether or not a particle's spin is aligned with its momentum. This is

with the exception of neutrinos, which are experimentally observed only to be left-handed [11]. A correct treatment of the chirality of the leptons was made in the electroweak theory [12, 13, 14], which is discussed in Section 2.2.

Bosons, on the other hand, have integer units of spin (0, 1, 2, etc.), and have quantum mechanical wave-functions that are symmetric under particle exchange (i. e. the Pauli exclusion principle does not exist for bosons, and they can form condensates). In the SM, bosons play the role of interaction mediators.

Three of the fundamental interactions in the SM are responsible for different forces that have been identified in nature: electromagnetism, the strong nuclear force and the weak nuclear force. These interactions are associated with the exchange of spin 1 gauge bosons that arise from requiring certain symmetries to be obeyed by nature, known as gauge symmetries. These bosons are therefore also known as gauge bosons.

The photon (γ) is associated with electromagnetism and the gluon (g) to the strong interaction; these two particles are massless. The weak interaction is mediated by the massive $W^{\pm}$ and Z bosons, which have masses of 80.385 ± 0.015 GeV and 91.1876 ± 0.0021 GeV, respectively [9]. Gravity is believed to be associated with a spin 2 gauge boson called the graviton. However, there is no experimental evidence that it exists yet.

An additional fundamental interaction in the SM is the Higgs interaction, which is not mediated by a gauge field. It is rather included by hand as a field that allows particles to have mass whilst still preserving the existing symmetries imposed in the SM. The result of this is, as described in the section below, the addition of the spin zero (scalar) Higgs boson to the SM: a massive particle that has a mass of 125.09 ± 0.24 GeV [15]. The origin of the Higgs boson (and the masses of all known massive elementary particles) was made clear through the theoretical description of electroweak symmetry breaking [12, 13, 14].

2.2 ELECTROWEAK SYMMETRY BREAKING

The early development of the SM had the goal of combining different explanations of particle interactions using gauge invariance. Quantum Electrodynamics (QED) could be explained using a theory based on gauge invariance under a $U(1)$ group symmetry, through which the photon and its interactions emerge, and for Quantum Chromodynamics (QCD) an $SU(3)$ group symmetry produces eight gluons and its interactions with the quarks through colour charge. Such a theory based on an $SU(N)$ group is called a Yang-Mills theory [16]. The weak interaction is also explainable in terms of a Yang-Mills theory. However, there are some caveats which complicate the description. This section describes these caveats and their solution through the concept of electroweak symmetry breaking (EWSB).

In order to understand weak interactions, Glashow, Weinberg and Salam (GWS) introduced what is known as the Electroweak (EW) Theory [12, 13, 14]. The theory pairs the weak and electromagnetic interactions into one Yang-Mills theory. In the theory, all of the left-handed fermions are organised into weak iso-doublets (that is, the components mix with each other under transformations in an isospin space),

$$q_L = \begin{pmatrix} u \\ d \end{pmatrix}_L \text{, and } \ell_L = \begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}_L , \tag{2.1}$$

for the up- and down-type quarks, and the neutral and charged leptons, respectively. The right-handed fermions are treated as singlets, and don't transform under the EW gauge group: $SU(2)_L \times U(1)_Y$. The subscript L identifies that the $SU(2)$ group transformations act only on the left-handed fermion doublets. The Y symbol refers to weak hypercharge, which is a conserved quantity defined in terms of the Gell-Mann Nishijima equation [10, 17]:

$$Y = 2(Q - I_3), \tag{2.2}$$

where here Q refers to the electric charge of a fermion and I_3 is the third component of its isospin. The parameter Y is the generator of the $U(1)$ group.

This treatment of the fermions is necessary, however it comes with two unavoidable problems. Firstly, the requirement of gauge invariance forces the bosonic mediators of the theory to be massless, in the same way that photons and gluons are massless. It was known from previous work by Fermi, however, that the bosonic mediators of the weak interaction (the $W^{\pm}$ and Z bosons) should have a non-negligible mass [18, 19]. The second problem is related to the rewriting of the theory in terms of left-handed and right-handed projections of the fermions. Since the EW theory only applies $SU(2)$ transformations on the left-handed doublets, one cannot write down gauge invariant mass terms for the fermions, which are known to be massive. Therefore, the fermions could not have mass terms either. Remarkably, both problems are solved by the introduction of the Brout-Englert-Higgs (BEH) mechanism [20, 21, 22].

The BEH mechanism was first proposed independently in the 1960s by Peter Higgs [20, 21], Francois Englert and Robert Brout [22] who were the first to show how a gauge field can acquire mass without violating gauge invariance. This theoretical discovery was later followed by a similar work done by Guralnik, Hagen and Kibble [23]. The solution is found in postulating the existence of a scalar field which has a non-zero vacuum expectation value (VEV), which shall be briefly explained below.

For illustrative purposes, one could first consider a complex field ϕ interacting with a $U(1)$ gauge field A^{μ}. On introducing the gauge field, one needs to preserve gauge invariance by embedding A^{μ} into the Lagrangian as follows:

$$\mathcal{L} = \left(D_{\mu}\phi\right)^{\dagger}\left(D^{\mu}\phi\right) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \mu^2\phi^*\phi - \lambda\left(\phi^*\phi\right)^2, \tag{2.3}$$

where $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ and D_{μ} is the covariant derivative $D_{\mu} = \partial_{\mu} - ieA_{\mu}$. We could denote the first two terms of Equation (2.3) as the *kinetic* terms of the scalar field and gauge field, respectively. The other two terms can be thought of as the negative of a potential $V(\phi)$ (i. e. such that $\mathcal{L} = T - V$). Interactions between ϕ and

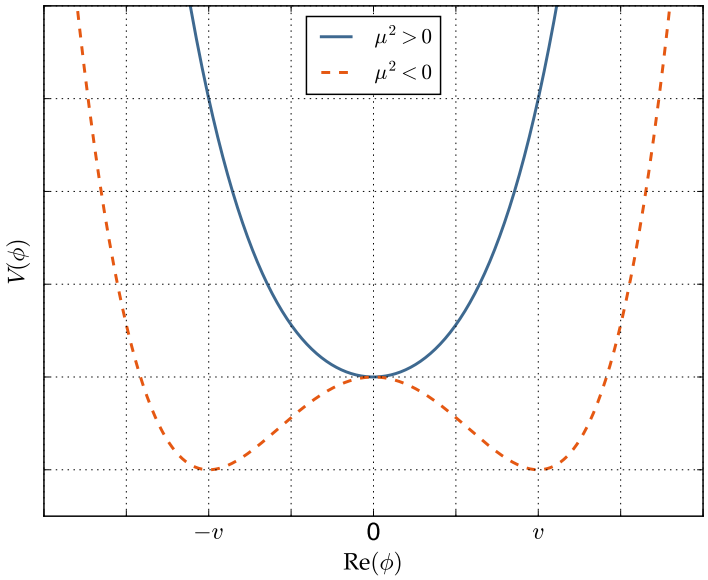


Figure 2: The potential in Equation (2.3) as a function of the real projection of the complex scalar ϕ. A mental rotation of the dashed curve around the vertical axis should yield a surface that resembles a *Mexican hat* shape. For this reason, the potential from Equation (2.3) is often referred to as the "Mexican hat potential."

A^μ arise through the covariant derivative, and gauge invariance implies that there can be no mass term for A^μ. Note that since ϕ is complex, it has two degrees of freedom. The massless gauge boson A^μ also has two degrees of freedom, adding up to a total of four degrees of freedom manifest in the theory.

From a classical perspective, one might assume that μ^2 should be positive, otherwise the theory would predict a tachyon [24]. However, from a quantum mechanical perspective, the particle spectrum produced by the theory can be approximated by looking at small fluctuations of the Lagrangian in Equation (2.3) around a stable background. This implies studying the extrema of $V(\phi)$, and becomes interesting when one considers the case when $\mu^2 < 0$, and a non-trivial solution set of vacua with a VEV $v = -\frac{\mu^2}{\lambda}$ can be explored. This can be seen in Figure 2, which shows only the real projection of the complex field ϕ. Any other projection in the complex plane would yield a similar image, and so one could mentally rotate Figure 2 around the vertical axis to understand that the set of minima for the dashed curve actually forms a circle of points with radius v.

Now, by shifting the fields such that $h = \mathrm{Re}(\phi) - v$ and $\chi = \mathrm{Im}(\phi)$, we can then make an expansion around the new vacuum of the theory to study the particle spectrum. In doing so one will find (omitting all of the interaction terms):

$$\mathcal{L}(h,\chi) = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \lambda v^2 h^2 + \frac{1}{2}(\partial_\mu \chi)(\partial^\mu \chi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2 v^2 A_\mu A^\mu. \quad (2.4)$$

One can immediately see that the last term in Equation (2.4) is a mass term for the gauge field ($m_A = \frac{ev}{\sqrt{2}}$) which arises from the fact that ϕ took on a VEV. In addition, there exists a massive scalar boson h with mass $m_h = \sqrt{\lambda}v$, and another apparent massless scalar χ. It can be shown that χ should not be treated as a true particle in the theory, since it can be removed from the Lagrangian through a simple $U(1)$ rotation after choosing a fixed gauge; χ is known as a *Nambu-Goldstone boson* [25, 26] and h is known as a *Higgs boson*. The theory as written in Equation (2.3) has four degrees of freedom, and after expanding around the vacuum there should still be four degrees of freedom. Counting the degrees of freedom in Equation (2.4), there is one for the massive scalar h and three for the now massive gauge field.

The BEH mechanism therefore explains how a gauge field can have mass without violating gauge invariance, and which keeps all of the important dynamical information introduced by the gauge symmetry in the first place. In the SM, recall that the weak gauge bosons should have mass, and that they emerge from the $SU(2)_L \times U(1)_Y$ EW gauge symmetry. In this case, the Higgs field takes the form of a complex doublet with $+1$ hypercharge, such that the vacuum is neutral. The result of applying the BEH mechanism here is that we are left with three massive gauge bosons (the $W^\pm$ and Z) and the massive Higgs boson (h). In terms of interactions, the Higgs boson can interact with the massive gauge bosons through h-V-V couplings (where $V = W^\pm$ and Z). In the SM, the value of the VEV v can be calculated using known SM parameter values with the relation $v = \frac{2m_W}{g}$. It is roughly equal to 246 GeV [9]. The photon remains massless, and what remains is the $U(1)_{\mathrm{EM}}$ symmetry that defines electromagnetism. We say that the $SU(2)_L \times U(1)_Y$ symmetry is *broken*.

Vertex	Coupling strength
h-f-f	$-\frac{m_f}{v}$
h-Z-Z	$\frac{m_Z^2}{v}$
h-W-W	$\frac{2m_W^2}{v}$
h-h-Z-Z	$\frac{m_Z^2}{2v^2}$
h-h-W-W	$\frac{m_W^2}{v^2}$
h-h-h	$-g\frac{m_h^2}{4m_W}$
h-h-h-h	$-g^2\frac{m_h^2}{32m_W^2}$

Table 2: The interactions and coupling strengths related to the Higgs boson after EWSB. Here g is the $SU(2)_L$ gauge coupling.

Recall also that the chirality of the fermions in the EW theory prohibits them from having gauge invariant mass terms. On applying the BEH mechanism to the EW theory, it can be shown that the fermions also acquire mass through the non-zero VEV of the Higgs field. The mixed chiral states have interactions with the Higgs field through the vacuum, and result in a mass term for each lepton (not the neutrinos) and the down-type quarks, as well as three-point interactions between the fermions and the Higgs boson itself. Both of these terms are multiplied by a *Yukawa coupling*, which is a free parameter in the theory. The Yukawa coupling is proportional to the mass of the fermion in question. Whereas the mass terms for the down-type fermions come naturally, the masses for the up-type fermions and neutrinos need to be added in by hand, although this can be done in a gauge invariant way.

The EW theory is said to have two phases. The phase where the $SU(2)_L \times U(1)_Y$ symmetry is exact is called the *unbroken* phase. The *broken* phase exists after the process of EWSB, and is the phase that we live in. All the possible interactions that the Higgs boson has with the SM particles in the broken phase are listed in Table 2. It is believed that the unbroken phase exists at high energy scales, and that the running of the couplings in the Higgs field causes the symmetry to be spontaneously broken when the energy scale is low enough. At the LHC, the broken phase of the EW theory is studied, and the production of the Higgs boson is one of the most important processes that can be used to study EWSB.

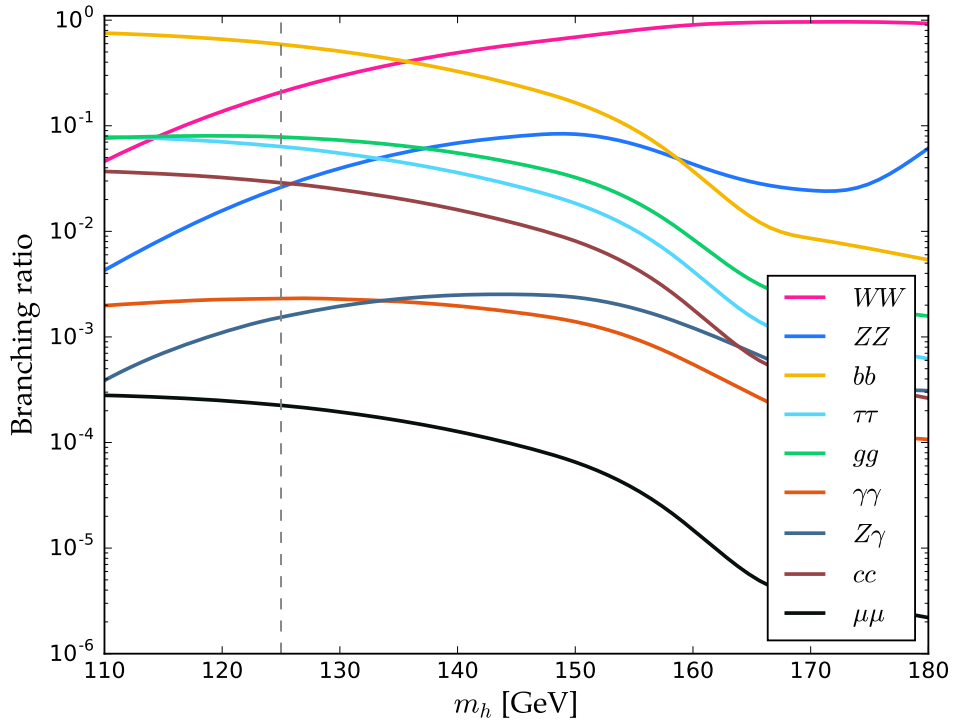


Figure 3: The BRs for a Higgs boson in the mass range 110 GeV $< m_h <$ 180 GeV. Above 180 GeV, the BRs to the weak vector bosons dominate, until $m_h > 2m_t$, where the $h \to t\bar{t}$ BR becomes dominant due to the size of the top Yukawa coupling. The grey dotted line is an indication of the BRs for the SM Higgs boson, which has a mass of around 125 GeV.

2.3 THE HIGGS BOSON

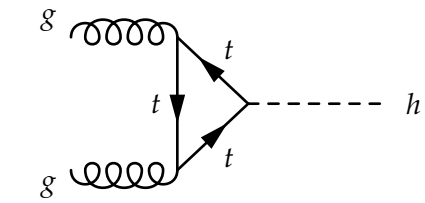
The Higgs boson is of special importance in this thesis. It is studied in the context of proton-proton (pp) collisions at the LHC, which is where it was discovered. In most phenomenological studies, the starting point for understanding a rare particle production is the calculation of its branching ratios (BRs) and the cross sections for its different production modes.

The BRs for a Higgs boson, shown in Figure 3 as a function of its mass, largely depend on the strength of the Yukawa couplings as well as phase space constraints. The fermionic decay modes all arise from a tree level Yukawa coupling, and so they are roughly proportional to the mass of the fermion in question. They are also required to be on-shell, and so the $h \to t\bar{t}$ decay mode would only be possible with $m_h > 2m_t$. The decays to WW and ZZ also arise from a tree level coupling. These

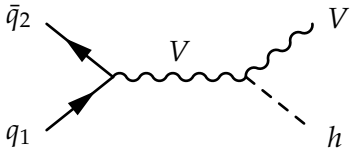
decay modes become dominant when $m_h > 2m_W$, as can be seen in Figure 3. Below the kinematic thresholds, these decays are off-shell, and need to be computed under the assumption that the W and Z bosons decay further to all of their possible decay modes. Finally, the decays to the massless gauge boson pairs as well as $Z\gamma$ are not possible at tree level since the couplings do not exist in the theory. These are loop-induced processes, and are therefore suppressed. In addition to this, the interference between Yukawa couplings and the h-V-V couplings in the loops can cause further suppression, as is the case for the $h \to \gamma\gamma$ decay mode and in the production of a single top quark in association with the Higgs boson.

At the LHC the Higgs boson is produced through processes all having partons (that is, quarks and gluons) in the initial state. This is because we understand protons as being made up of partons, the relative composition of which can be deduced from parton density functions (PDFs). At the high energies of the LHC it is well known that the gluon PDF dominates over the quark PDFs, with the gluon density being more than 10 times that of any single quark density at lower momentum fractions [9]. Therefore, the dominant Higgs production mechanism at the LHC is that of gluon fusion (ggF). The Feynman diagram for ggF can be seen in Figure 4a. It is a direct production mechanism, and so the Higgs boson is usually produced with one or two jets from initial state radiation (ISR). In spite of ggF being a one loop process, its cross section is enhanced further by the strength of the top Yukawa coupling (recall that Yukawa couplings are proportional to the mass of the fermion in question, and the top quark mass is the largest of the known fermions). The ggF production cross section also gains a small enhancement from adding contributions in which the b-quark runs in the loop in place of the top quark.

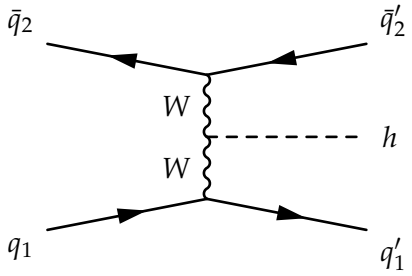
The dominant associated Higgs production modes include (in decreasing order of production cross section) vector boson fusion (VBF), vector boson associated production (Vh), and top associated production ($t\bar{t}h$). These are shown in Figure 4c, Figure 4b and Figure 4d, respectively. The former two production modes are quark initiated processes. Even though top associated Higgs production is gluon initiated and contains a strong Yukawa coupling, its production cross section is compara-



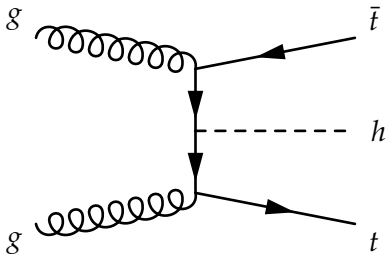
(a) Gluon fusion (ggF).



(b) Vector boson associated production (Vh).



(c) Vector boson fusion (VBF).



(d) Top associated production ($t\bar{t}h$).

Figure 4: Feynman diagrams for the dominant Higgs boson production modes in high energy pp collisions [27].

tively small due to the fact that top quarks are heavy, and so it is phase space suppressed.

The unique final state topologies of the different production mechanisms shown in Figure 4 allow for a relatively robust separation in searches for the Higgs boson at the LHC. Therefore, measurements of the production cross sections for the individual modes can be made through high energy pp collisions. However, it should be noted that potential BSM production mechanisms of the Higgs boson could have significant implications for these measurements, as shall be explored later in this thesis.

The measured mass of the SM Higgs boson is around 125 GeV [15]. For this mass and a large range of masses around it, the production cross sections and BRs have been calculated to several orders in perturbation theory and with great precision. The state of the art calculations are performed at next-to-next-to-next-to leading order (N3LO) in QCD corrections for both ggF [28, 29] and VBF [30], up to next-to-next-to leading order (NNLO) with a next-to-next-to leading logarithmic (NNLL) resummation for the Vh mode [31] and up to next-to leading order (NLO) for $t\bar{t}h$ [32]. EW corrections are usually only necessary up to NLO. The most

Branching ratios	
Mode	Branching ratio [%]
$h \to b\bar{b}$	$58.09^{+1.24\%}_{-1.26\%}$
$h \to \tau\bar{\tau}$	$6.256^{+1.65\%}_{-1.63\%}$
$h \to \mu\bar{\mu}$	$0.02171^{+1.68\%}_{-1.70\%}$
$h \to c\bar{c}$	$2.884^{+5.55\%}_{-1.97\%}$
$h \to gg$	$8.180^{+5.15\%}_{-5.08\%}$
$h \to \gamma\gamma$	$0.2270^{+2.09\%}_{-2.05\%}$
$h \to Z\gamma$	$0.1541^{+5.81\%}_{-5.83\%}$
$h \to WW$	$21.52^{+1.53\%}_{-1.52\%}$
$h \to ZZ$	$2.641^{+1.53\%}_{-1.52\%}$

Production cross sections			
Mode	$\sqrt{s}$ [TeV]	Order	Cross section [pb]
ggF	8	N3LO	$21.39^{+5.49\%}_{-7.61\%}$
	13	N3LO	$48.52^{+5.57\%}_{-7.44\%}$
VBF	8	NNLO	$1.600^{+2.22\%}_{-2.21\%}$
	13	NNLO	$3.779^{+2.14\%}_{-2.12\%}$
Wh	8	NNLO	$0.7009^{+2.14\%}_{-2.20\%}$
	13	NNLO	$1.369^{+1.93\%}_{-2.00\%}$
Zh	8	NNLO	$0.4199^{+3.28\%}_{-2.94\%}$
	13	NNLO	$0.8824^{+4.12\%}_{-3.40\%}$
$t\bar{t}h$	8	NLO	$0.1326^{+5.87\%}_{-10.2\%}$
	13	NLO	$0.5065^{+6.83\%}_{-9.88\%}$

Table 3: A summary of the SM Higgs boson production cross sections and BRs, taken from Reference [33]. All cross sections are calculated with NLO EW corrections, the "order" column refers only to the QCD corrections. BRs that are significantly negligible have been omitted. The percentage uncertainty quoted is a quadrature sum of the various components that contribute to the values' uncertainties.

extensive collection of these calculations has been done by the LHC Higgs cross section working group (LHC-HXSWG) [33]. A summary of their most updated calculations for $m_h = 125.09$ GeV at $\sqrt{s} = 8$ TeV and 13 TeV is shown in Table 3.

3

PARTICLE PHYSICS EXPERIMENTS

Controlled particle physics experiments are commonly done with particle accelerators. Fundamentally, particle physics seeks to find the structure of matter at the smallest scales possible. In terms of dimensional analysis, length is inversely proportional to energy. Therefore small distances can only be probed at high energies. It is for this reason that particle accelerators are an interesting avenue of exploration in particle physics, since colliding particles are accelerated to high energies.

In order to test the validity of predictions made in the framework of QFT, an experiment should make a measurement of cross sections. In most cases, scattering cross sections are calculated and measured from estimating and counting particle collision events in the accelerator. Therefore, particle accelerators contain collision points around which detectors are built to identify the remnants of the high energy collisions.

In this chapter, the particle accelerator and detectors relevant to this thesis are discussed. Special attention is paid to the ATLAS detector, since it is with this detector that the experimental work in this thesis is done.

3.1 THE LARGE HADRON COLLIDER

A particle accelerator that consists of two particle beams that travel in opposite directions and collide at specified points is called a *collider*. The biggest collider that has ever been built is currently the LHC [34], a ring-shaped pp and heavy ion collider based at the European Organisation for Nuclear Research (CERN) on the Franco-Swiss border.

Along the LHC lie four different collision points, each surrounded by a different detector. Two general-purpose detectors, ATLAS and CMS, are on opposite ends

of the ring, and are used to study a wide variety of physics processes. The A Large Ion Collider Experiment (ALICE) detector is designed to deal mostly with heavy ion collisions, and the LHCb (b for "beauty") is used to predominantly study B-physics processes. Each detector is designed to handle the full design specification of the LHC, the details of which are discussed below.

3.1.1 *Design*

The LHC ring is 26.7 km long and was built between 50 and 175 m underground in the concrete tunnels that previously housed the Large Electron-Positron Collider (LEP) [34]. Its sole purpose is to accelerate charged particles to high energies, including protons and heavy ions (such as lead or gold nuclei).[2]

The means by which the charged particle beams are accelerated around the LHC ring is the energy provided by a multitude of superconducting dipole magnets. A series of 1232 dipole magnets, each 15 m long, are used to bend the beams and accelerate them at a fixed orbit size through the synchronised ramping of their magnetic field (i. e. as a synchrotron would do). This twenty-minute process sees the magnetic field strength rise from 0.54 to 7.7 T. In addition to the dipole magnets, 392 quadrupole magnets (between 5 and 7 m long) are used to focus the beams. Finally, a number of higher order multipole magnets are used to "tighten" the beams around the collision points in order to enhance the probability of collisions. The magnets are cooled to around 1.85 K with liquid Helium, in order to ensure superconductivity.

For pp collisions, the source of the beams is a bottle of hydrogen gas. The hydrogen is ionised and the protons are split into bunches of an average of 115 billion protons per bunch. When the beams reach the collision points, the bunches are separated such that collisions happen every 25 ns (corresponding to a frequency of 40 MHz). The instantaneous luminosity of the beams was designed to be 10^{34} cm^{-2}s^{-1}, although during its operation the LHC has reached around twice

2 Even Xenon-Xenon collisions were studied for a short period. This was done to explore the system size dependence of physics in heavy ion collisions.

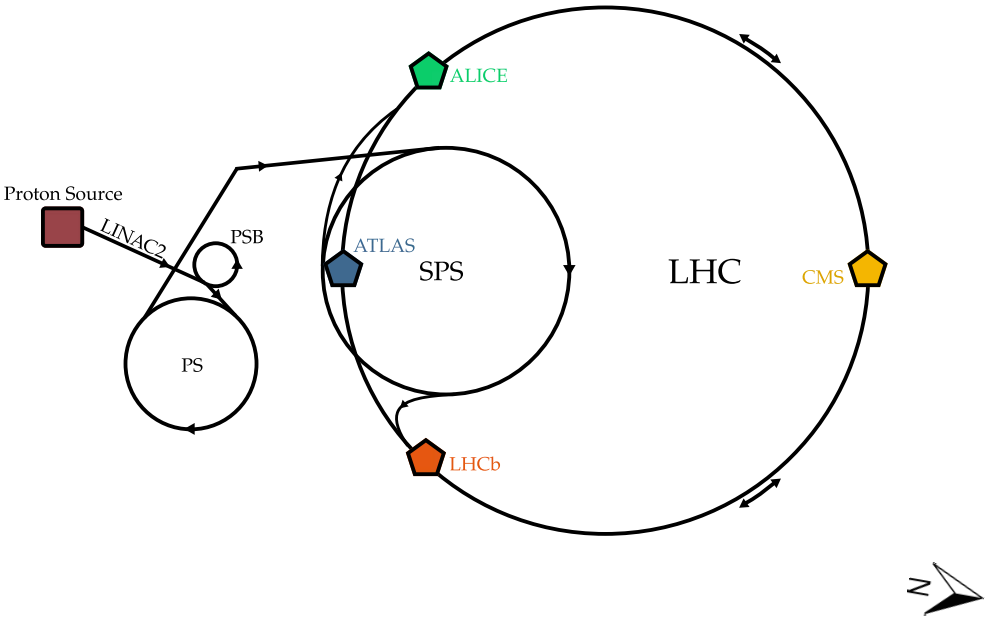


Figure 5: The systems of the CERN accelerator complex related to the acceleration of proton beams [34]. Many outlets and injection points not related to proton beams have been omitted for simplicity. The four primary detectors are represented as the four pentagons around the LHC ring.

this value. The design specification for the energy of the proton beams is 7000 GeV each, corresponding to a centre of mass energy ($\sqrt{s}$) of 14 TeV. The LHC has never reached this energy, having only been run at 7, 8 (Run 1) and 13 TeV (Run 2) since its first official activation in September 2008. However, upgrades to the collider are yet to be done that may result in it reaching its specified design energy.

3.1.2 *The CERN accelerator complex*

The procedure through which the beams travelling in the LHC reach such high energies is through a series of smaller accelerators known as the *CERN accelerator complex* [34]. The core systems of the complex that are related to pp collisions are shown in Figure 5.

The processes undertaken in the accelerator complex for the acceleration of proton beams to $\sqrt{s} = 13$ TeV are as follows [34]:

1. Bottled hydrogen is ionised and separated into bunches before being fed into the LINAC2 linear accelerator, which is designed for proton beams (LINAC3 is designed for lead ion beams).

2. LINAC2 accelerates the beams up to a kinetic energy of 50 MeV per proton.

3. The outlet of LINAC2 is the Proton Synchrotron Booster (PSB), taking the energy from 50 MeV to 1.4 GeV. The protons then enter the Proton Synchrotron (PS) and are accelerated to 25 GeV.

4. The final stage before the entering the LHC is their acceleration to 450 GeV in the Super Proton Synchrotron (SPS). The SPS has two outlets to the LHC, one for each beam pipe in the LHC. This is because two beams travel in opposite directions before meeting at the collision points.

5. In the LHC beam pipes, the proton beams are finally accelerated to 6.5 TeV over a period of 20 minutes. During this time, the magnetic field strength is ramped up from 0.54 to 7.7 T.

At this stage, the beams are allowed to meet at the collision points, and the resulting debris is studied in the detectors.

3.2 THE ATLAS DETECTOR

The ATLAS experiment is an active collaboration that studies pp and heavy ion collisions at the LHC. Over the years, the ATLAS experiment has been a collaborative effort of thousands of scientists from 182 different institutions around the globe. The combined research focus is directed towards the refinement of the SM, the understanding of EWSB, and for BSM searches in the context of the Higgs boson, supersymmetry (SUSY), and other exotic models. A large portion of manpower is also dedicated to the maintenance and operation of the ATLAS detector.

The ATLAS detector [35] is (to date) the largest particle detector ever built. It has a roughly cylindrical shape, and is 46 m long and up to 25 m in diameter. It is classed as a "general-purpose" detector, and therefore has systems measuring the

transverse energies of all visible final state particles (electrons, muons, photons and all hadrons) and tracking the trajectories of charged particles with high resolution. In addition to its identification of visible particles, the ATLAS detector can be used to reconstruct missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) from particle collisions, which can arise from the production of neutrinos or other hypothetical invisible particles. This is possible due to the fact that it is a hermetic detector, and therefore can detect particles scattered effectively up to 4π in their solid angle.[3] $E_{\mathrm{T}}^{\mathrm{miss}}$ can then be correctly calculated from the momentum imbalance of the collision.

The ATLAS detector has a modular design. Centered around the beam pipe and moving in a radial direction outwards, the detector is built from the following systems: the inner detector (ID), the calorimeters, the muon spectrometers and a magnet system. The ID is built up of trackers, which can track charged particle trajectories with negligible energy loss. A solenoid magnet (the *inner solenoid*) encompasses the ID with a magnetic field strength of 2 T; this causes the charged particles to curve in the R-ϕ plane (assuming cylindrical coordinates centered around the beam pipe). The curvature of these tracks can be used to measure the particles' momenta. The transverse energy of all hadrons, electrons and photons can then be measured in the calorimeters, where the energy is deposited into different absorbing materials. These calorimeters are coupled with specific sampling materials, which are able to sample the shape of the energy deposit as it creates a "shower" of particles during its absorption. Muons are able to pass through the calorimeters without their energy being absorbed, and so a set of muon spectrometers is built at the furthest point of the beam pipe. A collection of eight large toroidal magnets are used to again bend the trajectories of the passing muons, allowing for a high-resolution determination of the muons' momenta.

Each of these systems shall be described in more detail below. For reference, Figure 6 contains a diagram showing a "slice" of the detector in the R-ϕ plane (as though one were looking down the beam pipe), where the relative positions of each sub-system can be seen.

3 The hermeticity of the detector is of course not perfect, since particles can still escape along the beam-line. However, in regions where interesting physics is expected to produce signals, overlapping detection systems are used to ensure hermeticity.

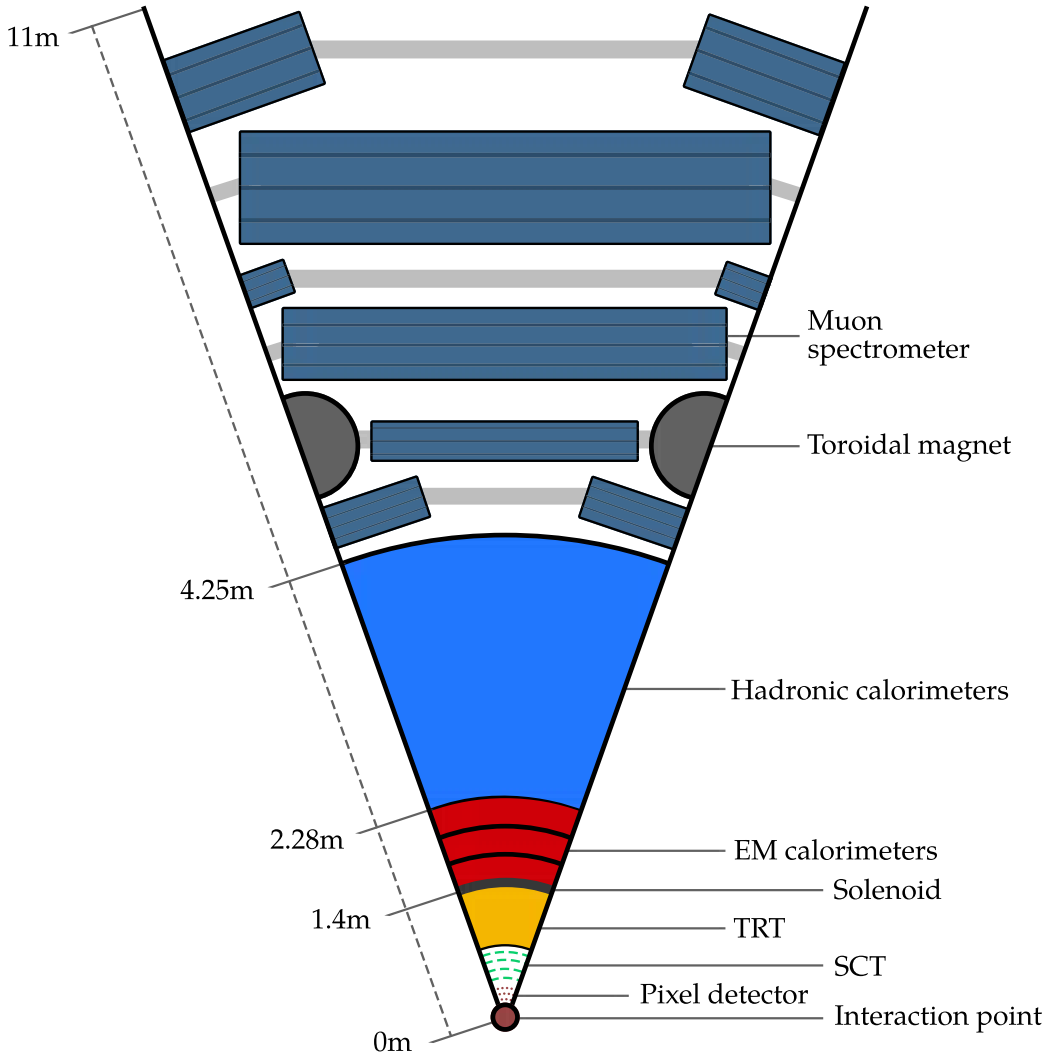


Figure 6: An R-ϕ cross-sectional slice of the ATLAS detector, showing the main components necessary for tracking and determining the transverse energies and momenta of visible particles.

3.2.1 *Trackers*

An efficient tracking system is fundamental to a particle detector, since it allows for the calculation of charged particle momentum, lifetime, charge and identification (what kind of particle it is, as based on its mass). Among other things, the tracks left by charged particles can be used to associate charged particles to reconstructed objects, determine where a primary vertex occurred in the geometry of the detector, as well as provide important information about the nature of different jets (if they originated from a B-hadron, for instance).

The final state particles that result from high energy collisions at the LHC travel close to the speed of light (with the exception of photons, which travel precisely at the speed of light). Given this fact, the equalities derived from relativistic mechanics can be used to determine particle properties. Under a magnetic field, a charged particle will be accelerated perpendicular to the directions of its motion and the magnetic field, according to the Lorentz force law:

$$\vec{F} = q\vec{v} \times \vec{B}. \tag{3.1}$$

Having measured the curvature of a charged particle's trajectory, the transverse momentum (p_{T}) can be calculated. In practice, a *sagitta* is measured, which is the distance from the arc of a particle's trajectory to the geometrical chord that subtends the full trajectory. From the sagitta, an uncertainty can be calculated and propagated for the p_{T} of the particle.

Several important parameters of a particle's creation and motion can also be determined using information from the tracking system. Using the helix equation, one can determine a set of five "helix" parameters:

- d_0 – the transverse distance of a particle's closest approach to the beam pipe,
- z_0 – the longitudinal distance from the interaction point to the particle's origin,

- ϕ_0 – the azimuthal angle of the particle's track at the point of closest approach to the beam pipe,
- $\cot\theta$ – where θ is the polar angle of the particle's motion with respect to the beam pipe, and
- q/p – the particle's charge to momentum ratio.

The distance and angle parameters listed above can be used to determine the primary vertex of a particle, and can also be used to cut out anomalously formed particle tracks based on the resolution up to which they can be measured.

The ATLAS inner detector [36, 37] contains three such tracking systems, starting as close as a few centimeters radially out from the beam pipe, and ending with the inner solenoid 1.4 m away from the beam pipe. Each of the trackers is placed under the constant magnetic field of 2 T provided by the inner solenoid [38].

Pixel Detector – The Pixel Detector (PD) [39] is the innermost detecting component of the ATLAS detector. It consists of three layers of side-by-side 2×6 cm modules, with 1744 modules in total. Each module contains a silicon board with a thickness of 250 μm with 47000 pixels of dimension 50×400 μm. The PD reads out data to 80 million different channels, which is about half of the total read-out for the entire detector.

Semiconductor Tracker – The Semiconductor Tracker (SCT) works in a similar manner to the PD, with the exception that is constructed from long strips of tracking silicon boards. These strips are 80 μm $\times$ 12 cm in dimension, and the entire SCT is made from four layers of paired strips. The purpose of the SCT's use of long strips is to enhance the precision of tracking the curvature of charged particles further away from the PD, where the particles would be more spread out.

Transition Radiation Tracker – The final tracker in the ID is the Transition Radiation Tracker (TRT). The TRT is made up of two main components: a straw tracker and a material used for the production of transition radiation. A straw tracker is a tracker made up of a large number of thin tubes (also known as "drift tubes") used to signal the presence of a passing particle in a similar fashion to a Geiger Müller tube. The tubes are filled with an inert gas and a wire passing through the

middle, with a voltage of 1500 V between the tube wall and the wire. When high energy particles pass through, they ionise the gas and the large voltage creates an avalanche of electrons directed towards the wire, thus initiating a signal. The material in between the tubes creates transition radiation, adding to the ionisation of the tubes to make a stronger signal. Only charged particles create transition radiation, and therefore the TRT only tracks charged particles. The entire TRT is made of $\sim$ 300 000 straws, each 4 mm in diameter and up to 144 cm in length depending on their location. The straws that create a signal are used to reconstruct the tracks of passing charged particles.

3.2.2 *Calorimeters*

After passing through the trackers unperturbed, all of the visible particles travelling through the detector (except for muons) are forced to deposit their energy in one of two calorimeters. Both calorimeters are "sampling calorimeters", which means that they are made of at least two materials: an absorption material for absorbing the bulk of the particle's energy, and a sampling material that absorbs the shower of particles created during the absorption. The measurement of shower shapes is an important part of identifying particles in an offline analysis. The calorimeter deposits carry no information about the momentum of the particles, and therefore can only provide a measurement of a particle's transverse energy (E_T).

Electromagnetic Calorimeter – Particles that interact predominantly through electromagnetic interactions are detected in the Electromagnetic (EM) Calorimeter [40]. These are mainly charged particles (electrons are an important example) and photons. The absorption materials used are lead and stainless steel arranged in an accordion shape. The sampling material used is liquid argon (LAr); for this reason the calorimeter is most often referred to as the LAr calorimeter. To keep the Argon in the liquid phase, a cryostat at the edges of the material is used to keep its temperature at around 90 K. The EM calorimeter has a high resolution, and can distinguish energy deposits within a polar angle of 1.4° [40].

Hadronic Calorimeter – The remaining hadrons are then absorbed in the Hadronic Calorimeter [41]. Here the absorbing material is steel and the sampling material is made of plastic scintillators. In the central region the materials are arranged as tiles, and therefore the calorimeter is called the Tile Calorimeter (TileCal). The entire TileCal is 8 m in diameter and 12 m long. At the end-caps is a forward hadronic calorimeter, situated in the EM cryostat. This is necessary because the forward calorimeter uses LAr as a sampler, while copper and tungsten are used as absorbers.

3.2.3 *Muon spectrometer*

After the Hadronic Calorimeter, only muons should still be travelling radially outwards since all other particles should have previously been absorbed. For this reason, a set of muon detectors exist on the outermost layer of the ATLAS detector, known as the Muon Spectrometer (MS) [42].

The MS acts as a large tracking system, and therefore is drenched in a non-uniform magnetic field provided by three large superconducting toroidal magnets. These magnets are made through the combination of eight barrel loops and two end-caps. Similar to the trackers in the ID, the MS provides information about muons passing through with high-resolution. There are 1 200 chambers throughout the system, each linked to specially designed triggers (this is discussed in Section 3.2.4).

3.2.4 *Trigger system*

The ATLAS detector was designed to manage around one billion pp collisions per second. If one considers all of the information recorded by the trackers and calorimeters, this adds up to an exceptionally large data volume needed to be recorded per second. Clearly the magnitude of this amount is too big to deal with for long-term storage and offline analysis. However, the number of collisions that

are of interest to the analysers of the data is a small fraction of the total number of collisions. Therefore, the collisions can be filtered, such that only the interesting collisions are kept.

The method by which this is done in ATLAS is by using a trigger system. The Run 2 trigger system [43] is made up of a Level 1 (L1) trigger [44] followed by a high level trigger (HLT) [45].

The L1 trigger operates through hardware, since it is required to make filtering decisions within 2.5 μs. It outputs "regions of interest" using information from the calorimeters and MS, using simple electronic logic gates to make decisions. These regions of interest are used in the HLT as inputs. The L1 trigger reduces the rate of collisions that pass from 30 MHz to 100 kHz [43].

The HLT, on the other hand, is largely software driven. It comprises of three sub-components: the L2 trigger, the Event Filter, and the Event Selection Software. Using the L1 regions of interest, the HLT makes use of all of the event information to make filtering decisions. Being software based, several more complicated algorithms can be used to determine what type of event the collisions are. On average, the decision process takes around 200 ms per event, and this results in a further event frequency reduction of around 100 times. At this point, events are recorded at about 1 kHz and can be stored for use in offline analysis [43].

3.2.5 *Running conditions*

One of the biggest challenges in reconstructing pp collisions is that of mitigating the effects of *pile-up*. Pile-up is an inescapable process at the LHC, and results from multiple interactions in a single bunch crossing during a collision event. This typically results in the production of additional jets that have a low p_{T} and a high pseudo-rapidity (η). The intensity of pile-up is usually given in terms of the average number of interactions per bunch crossing ($\langle\mu\rangle$), which had a value of 33.7 for the entire LHC Run 2 data taking period [46]. ATLAS has constructed novel techniques for the rejection of pile-up jets, which have rejection rates between 49%

and 67%, while maintaining an 85% efficiency for reconstructing jets from the hard scatter process [47].

In terms of the total luminosity collected for the LHC Run 2 data taking period, out of the 156 fb^{-1} that the LHC delivered, ATLAS was able to record 147 fb^{-1} [48]. Of this data volume, 139 fb^{-1} are considered good for physics analyses, thereby corresponding to an 89% efficiency for data taking during the entire Run 2 period.

3.3 THE CMS DETECTOR

While the CMS detector is not of primary importance in this thesis, it is still worthwhile discussing it. This is because many of the results that ATLAS produce are done concurrently with the CMS detector and, therefore, statements about experimental observations can be strengthened by considering a combination of both data sets.

The CMS detector [49] is another general-purpose detector at the LHC, situated at the opposite end of the LHC ring with respect to the ATLAS detector (see Figure 5). It weighs 14 000 tonnes, is 15 m in diameter and has a full length of 28.7 m. Its modular structure is designed similarly to the ATLAS detector, although emphasis was placed in its design to utilise a large solenoid magnet (the largest ever constructed) to make the detector more "compact".

The solenoid magnet produces a 3.8 T magnetic field, and surrounds both the trackers and the calorimeters (unlike the ATLAS detector, in which the inner solenoid only surrounds the trackers). This magnetic field is contained by an extremely heavy "return yoke" made of steel. This yoke also provides structural support for the entire detector.

Radially outwards, the innermost part of the detector contains silicon trackers – first in the form of pixels and afterwards as silicon strips. Thereafter, a special layer called the "preshower", which contains lead and more high-resolution silicon trackers, is used to initiate EM showers. This is used to provide information for $\pi^0 \to \gamma\gamma$ decays, such that they can be rejected in favour of prompt photon production (which more often relates to interesting physics events).

Following the preshower are the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The ECAL is made of lead tungstate ($PbWO_4$), which is primarily metal. However, its oxygen content enables it to act as a scintillator, and therefore the energy deposits are measured through the detection of the scintillated photons. The HCAL, on the other hand, is a sampling calorimeter (similar to the TileCal) where steel and brass are used as absorbers and plastic scintillators are used as samplers.

One of the hallmarks of the CMS detector is its accurate muon detection system. Inter-weaved with the steel yoke is a series of complicated tracking systems. These consist of drift tubes, cathode strip chambers and resistive plate chambers. The latter two components also act redundantly with the CMS trigger system, in order to provide information about whether or not the event is considered "interesting" and should be stored for analysis.

3.4 THE ALICE AND LHCB EXPERIMENTS

While the ALICE and LHCb experiments have no direct influence on the research done in this thesis, they are still an important part of the LHC programme. Each experiment runs operations on their own detector on the Southern side of the LHC ring.

The ALICE experiment [50] is the largest collaboration based at CERN that is focused primarily on the collisions of heavy ions. The majority of their data-taking occurs when the LHC accelerates two colliding beams of lead ions (*Pb-Pb* collisions) or one beam of lead ions and another of protons (*p-Pb* collisions). These collisions result in a high multiplicity of final state particles dominated by different types of hadrons, which are identified by ALICE's *time-of-flight* detector. Along with particle multiplicity, another key observable considered in these collisions is centrality, which is a measure of the momentum asymmetry in the direction along the beam pipe. This correlates strongly with how peripheral the individual collisions are. One of the largest research interests for ALICE is studying the quark-

gluon plasma, a state of matter that was hypothetically abundant during the early stages of the universe.

The LHCb experiment [51] also partially seeks to shed light on a question related to the origin of the universe: the question of CP-violation and therefore the matter-antimatter asymmetry in nature. Speaking more generally, the LHCb experiment focuses its research on studying the properties and decays of B hadrons. This is done with a forward-biased detector that primarily studies $b\bar{b}$ production, one of the most prevalent processes in the high energy collisions at the LHC. In general, B hadrons decay to Kaons, the properties of which first displayed CP-violation in nature through a Nobel prize winning experiment by James Cronin and Val Fitch in 1980 [52]. LHCb is also interested in the observed anomalous rates of B hadron decays, known as the "B-meson anomalies" [53].

4

INITIAL CONSTRUCTION OF THE MODEL

At the end of Run 1 of the LHC data taking, several of the features related to the production of the Higgs boson appeared to show mild deviations for the SM predictions [3, 6]. Of particular interest at the time was the apparent enhancement of the Higgs boson p_{T} spectrum in the range below $p_{\mathrm{T}} \simeq 100$ GeV, slight deviations in the search for VV and hh resonances around a mass of 275 GeV, and the systematic enhancement of the production rates for a Higgs boson in association with top quarks. None of these results on their own was significant enough to be considered as potential new physics. However, it was believed in 2015 that each of the aforementioned deviations could be explained simultaneously under a single hypothesis with only one degree of freedom [3, 6]. This was the starting point for the BSM model that is developed throughout this thesis.

The hypothesis was that a heavy scalar boson H existed with a mass of around 275 GeV (where the heavy resonant searches had shown slight excesses), and that H decayed preferentially to the Higgs boson in association with some as-yet unspecified particles X (i. e. $H \to h + X$). The specific content of X was studied under a number of different hypotheses – including $E_{\mathrm{T}}^{\mathrm{miss}}$, leptons and jets – that are explored in the next few chapters. The Higgs boson as a result of this decay would be boosted to higher values of p_{T} depending on the four momentum of X, thereby potentially explaining the excess in the Higgs p_{T} spectra. The nature of H was assumed to be somewhat Higgs-like, and therefore its primary production mode is ggF. This assumes a Yukawa coupling (at least between H and top quarks) and therefore also assumes top associated H production, as can be seen for a Higgs-like particle in Figure 4. The hypothesis could therefore shed light on the enhancement of the top associated Higgs production rates. The t-t-H Yukawa coupling was also

assigned a scaling factor to control the production rate of H, this being the only degree of freedom in the hypothesis with a fixed mass of H.

The task at hand was to first construct a simplified way of explaining the apparent deviations in the Run 1 data sets. This first approximation was a bottom-up study that involved the production of H through ggF and in association with top quarks, with the prominent decay mode producing the Higgs boson in association with dark matter (DM). The methodology and results of this procedure are discussed in Section 4.1. As more data were released by the experimental collaborations, the compatibility of the simple hypothesis was re-established with larger data sets, as discussed in Section 4.2.

With the first approximation in place, the bottom-up study turned towards a top-down study, where the new goal was to try and formalise the simple hypothesis discussed here. This was achieved by embedding H in a 2HDM and introducing another scalar mediator S. With this in place, the construction of the new physics model could be completed by laying out a set of predictive search channels. These could then be used for the purpose of studying the feasibility of the model (or extensions) which would predict similar final states. This procedure is presented in Chapter 5. The process of iterating through bottom-up and top-down studies has been an important means of developing the BSM model, as mentioned in Chapter 1.

4.1 SIMPLIFIED PRODUCTION IN ASSOCIATION WITH DARK MATTER

As stated in Section 1.2, this thesis is primarily concerned with the construction and validation of a scalar extension to the SM. The BSM model considered in this research was inspired by some of the ATLAS and CMS results from the limited data sets that were available after Run 1 of the LHC. Since its initial inspiration in 2015, the LHC has been active in producing more data available for analysis by the experimental collaborations. It is therefore important to note which experimental results served as an inspiration for the construction of the model, and which results

were used to constrain the parameter space of the model. The distinction is made in Section 4.2.

It was clear at the time that trying to study such a scenario with an ultraviolet (UV) complete theory would not be useful, since it would only be an exercise in trying to constrain a large theoretical parameter space, instead of finding a single unifying explanation for the anomalous experimental results. Therefore, a simplified effective field theory (EFT) approach was adopted to introduce the new heavy scalar H. The approach that was adopted (and which was detailed in Reference [3]) was to make the simplified assumption that H should be able to decay into the Higgs boson in association with a pair of DM particles.

In order to describe the shape of the Higgs p_{T} distribution, it was necessary to introduce decays in which at least one h is produced, $H \to hh$ and $h\chi\chi$, where χ was a DM candidate, leading to the production of the Higgs boson in association with missing energy. The considered mass range for H was between 250 GeV and 350 GeV, since anything heavier would dominantly decay to $t\bar{t}$ and anything lighter wouldn't have any decays to hh, which were not excluded at the time of this study. The DM decay was later realised through the decay of some intermediate particle, which shall be discussed in Chapter 5. In either case, a distortion of the Higgs p_{T} spectrum was an initial driving force for the construction for the model.

The enhancement of top associated Higgs production (discussed in Section 4.2.3) also has an explanation in this narrative. The new boson would naturally be produced in association with top quarks, through the assumption of it having a Yukawa coupling with the top quark. In addition, with a small h-V-V coupling, the effect of negative interference in single top associated production is suppressed, and one would expect a larger cross section for a heavy scalar being produced in association with one or two top quarks. These effects provided an explanation for the excesses which were seen in the data at the time.

4.1.1 *Simplified formalism*

The features of the data that were intended to be explained can be treated in such a way that they are purely due to BSM physics. Therefore, a scenario was proposed that is a pure BSM extension to the SM, allowing that the following could be written:

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \mathcal{L}_{\text{BSM}}, \tag{4.1}$$

where all of the new interactions and states are encoded in $\mathcal{L}_{\text{BSM}}$.

The simplest approach was to treat the new interactions as arising from effective couplings. The Higgs boson h was assumed to have normal SM interactions with the fermions and gauge bosons. The sectors of the proposed BSM Lagrangian involving the new scalars (omitting the usual mass and kinetic energy terms) include:

$$\mathcal{L}_{\text{BSM}} \supset \mathcal{L}_H + \mathcal{L}_{\text{Y}} + \mathcal{L}_{\text{T}} + \mathcal{L}_{\text{Q}}, \tag{4.2}$$

where the terms $\mathcal{L}_{\text{Y}}$, $\mathcal{L}_{\text{T}}$ and $\mathcal{L}_{\text{Q}}$ are the Yukawa, tri-linear and quartic interactions relevant for this approach, respectively. These sectors are defined as follows:

$$\mathcal{L}_H = -\frac{1}{4}\,\beta_g \kappa^{\text{SM}}_{hgg}\; G_{\mu\nu}G^{\mu\nu}H + \beta_V \kappa^{\text{SM}}_{hVV}\; V_\mu V^\mu H, \tag{4.3}$$

$$\mathcal{L}_{\text{Y}} = -\frac{1}{\sqrt{2}}\,\Big[y_{ttH}\bar{t}tH + y_{bbH}\bar{b}bH\Big], \tag{4.4}$$

$$\mathcal{L}_{\text{T}} = -\frac{1}{2}\,v\Big[\lambda_{Hhh}Hhh + \lambda_{h\chi\chi}h\chi\chi + \lambda_{H\chi\chi}H\chi\chi\Big], \tag{4.5}$$

$$\begin{aligned}\mathcal{L}_{\text{Q}} = &-\frac{1}{2}\lambda_{Hh\chi\chi}Hh\chi\chi - \frac{1}{4}\lambda_{HHhh}HHhh \\ &-\frac{1}{4}\lambda_{hh\chi\chi}hh\chi\chi - \frac{1}{4}\lambda_{HH\chi\chi}HH\chi\chi,\end{aligned} \tag{4.6}$$

where H and χ denote the heavy scalar and the DM candidate, respectively (the latter is assumed to be a scalar for illustrative purposes), and $v = 246$ GeV is the VEV that is responsible for the W- and Z-boson masses. This can be looked upon as a variant of Higgs boson portal scenarios [54, 55, 56]. This Lagrangian could in principle emerge as an effective theory after EWSB in any gauge-invariant

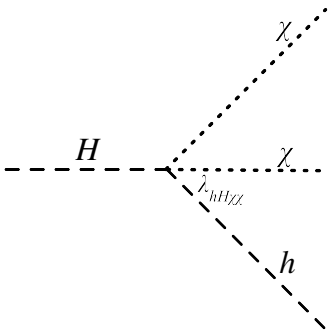


Figure 7: The leading order production mode of $h\chi\chi$ comes through the decay of the H boson. This is the representative Feynman diagram as described by Equation (4.6).

extended scalar sector. The second term in Equation (4.3) is summed over the weak vector bosons Z and $W^{\pm}$, and the κ factors are the SM-like couplings, with $\kappa_{hgg}^{\mathrm{SM}} = \alpha_s/(3\pi v)$ and $\kappa_{hVV}^{\mathrm{SM}} \simeq m_V^2/v$.

The gauge quantum numbers carried by H were deliberately not specified in this case; it was merely postulated that the above terms remain after EWSB. The parameter $\beta_g \kappa_{hgg}^{\mathrm{SM}}$ is set to be the strength of the effective g-g-H coupling. In situations where there are no additional effects over and above the top-mediated triangle diagrams contributing to this effective interaction, $\beta_g = y_{ttH}/y_{tth}$ where y_{tth} is the SM top Yukawa coupling. There would also be a similar relation for the bottom Yukawa coupling y_{bbH}, but this has been counted as negligible since the effect of bottom quarks in ggF loops is small. The production of H is made to occur through ggF and its rate can therefore be controlled by varying β_g. Likewise, the H-V-V couplings can be tuned by varying β_V. While these parameters do have deeper theoretical implications (with regards to gauge invariance, unitarity, etc.), it should be mentioned that these features can only be properly dealt with once a full theory is known. The effective model described here should by no means be treated as a full theory; rather its phenomenological consequences were used to explain what was seen in experimental data at the time.

Within this simplistic framework, one expects the process $pp \to H \to h\chi\chi$ to generate an enhanced p_{T} owing to the fact that h now recoils against a pair of invisible particles. Assuming the ggF production of H, the leading order production mode of $h\chi\chi$ comes from the decay process shown in Figure 7. The presence of an H-χ-χ coupling opens the potential for detecting invisible decays of H following

Result	Publication		N_{dof}
Higgs boson p_{T} spectra	**ATLAS**	$h \to \gamma\gamma$ [58] and $h \to ZZ^* \to 4\ell$ [59]	$8+4$
	CMS	$h \to \gamma\gamma$ [60] and $h \to ZZ^* \to 4\ell$ [61]	$6+4$
Di-Higgs boson resonance searches	**ATLAS**	$H \to hh \to b\bar{b}\tau\tau$, $\gamma\gamma WW^*$, $\gamma\gamma b\bar{b}$ and $b\bar{b}b\bar{b}$ [62]	1
	CMS	$H \to hh \to \gamma\gamma b\bar{b}$ [63], $b\bar{b}\tau\tau$ [64] and multi-leptons [65]	$1+1+1$
Top associated Higgs boson production	**ATLAS**	$h \to \gamma\gamma$ [66]	2
		Multi-leptons [67] and $h \to b\bar{b}$ [68]	$1+1$
	CMS	$h \to \gamma\gamma$, $h \to b\bar{b}$ and multi-leptons [69]	6
$H \to VV$ decays	**ATLAS**	$H \to WW$ [70] and ZZ [71]	$1+1$
	CMS	$H \to WW$ and ZZ [72]	$1+1$

Table 4: An ensemble of the experimental results which were available at the time when the relevant parameters in the effective production of $pp \to H \to h\chi\chi$ were constrained. Since the study was conducted before the first Run 2 results were published, all of the results listed here are from Run 1. In the interest of being as unbiased as possible, these results were selected regardless of whether they hint at physics beyond the SM. Note that updated results have since been released for several of these measurements, however the references here point to the results used when the study was conducted.

the methodology suggested in Reference [57], although this was not the focus of this study.

4.1.2 *Methodology for constraining the model*

In order to test the simplified model described above, an ensemble of results was constructed based on the final states in which the model would predict a signal. Many of these results served as an inspiration for the model, and are therefore discussed in Section 4.2. A list of the LHC results used to constrain the model in terms of the mass of H (m_H) and the rate of its production (controlled by β_g^2) is shown in Table 4. Several of these results have since been updated with newer developments from the experimental collaborations. However, in this section the focus is rather to present the constraints on the model at the time of its formulation in 2015; this was the starting point for the developments that are discussed later in this thesis. Tests of compatibility with the newer data sets are shown in Section 4.2.

In terms of astrophysical observables, major constraints on the Lagrangian parameters stem from the observations of the relic density of DM [73] and the DM-

nuclei inelastic scattering cross sections [74]. The computations relating to these DM constraints in the model were carried out using `micrOMEGAs` [75]. The constraints are controlled by the two model parameters m_χ and $\lambda_{h\chi\chi}$. It is found that both of the constraints can simultaneously be satisfied for a narrow choice of the parameters $m_\chi \sim [55 - 60]$ GeV for very small $\lambda_{h\chi\chi} \sim [6 \cdot 10^{-4}, 6 \cdot 10^{-3}]$ [76]. This keeps the invisible decay width of h well within the observed limits. The observed limits for the $h \to \chi\chi$ BR are 26% and 19% as measured by ATLAS and CMS, respectively [77, 78]. Other model couplings remain unconstrained by these observations.

Fitting the Run 1 ATLAS and CMS Higgs p_T spectra with the BSM prediction was accomplished as follows. For BSM predictions, the simplified model described above was built using `FeynRules` [79] and then passed to the Universal FeynRules Output (UFO) [80] such that event generation could be performed at leading order (LO) in `MadGraph5_aMC@NLO` [81]. Events with the $h\chi\chi$ and hh final states were generated from pp collisions through an H s-channel, and showered appropriately using `Pythia 8.2` [82]. Since the ATLAS and CMS Higgs boson p_T spectra were constructed from fiducial volumes in phase space, it was important that BSM predictions went through the same event selection. This was accomplished in the `Rivet` [83] analysis framework. The p_T spectrum of H was re-weighted to NLO with a re-weighting function derived from comparing an NLO calculation using `OneLOoP` [84] in `MadGraph5_aMC@NLO` with the result obtained just with the LO prediction and the shower. Overall, the LO prediction with the shower does a reasonable job, matching the NLO prediction with less than a 20% shape discrepancy over the full p_T spectrum. The effect of the these corrections on the transverse momentum of the Higgs boson from the decay of H is small and it consists of a positive shift of about 3 GeV. It is, however, important to note that the jet multiplicity of the H boson in this setup is significantly larger than that characteristic to h. This implies a significant reduction of the jet veto survival probability. The total LO cross section of the BSM prediction was enhanced to NNLL+NLO through multiplication by a k-factor determined from Reference [33].

For the SM Higgs p_{T} spectrum, the ggF Higgs p_{T} spectrum was generated at NLO using the `MINLO` HJ code [85]. The predictions of the other less prominent Higgs boson production modes (VBF, Vh and $t\bar{t}h$, which are collectively referred to as Xh) were taken directly from the ATLAS and CMS Higgs p_{T} spectrum figures. The bulk of the production cross section is in the intermediate range of Higgs p_{T}, where ggF is the dominant production mechanism. For ggF, QCD radiative corrections play a critical role in generating p_{T} for the Higgs boson. The MC used to simulate ggF describes the Higgs p_{T} differential distribution at NNLL+NLO. Results on the NNLO corrections on ggF $+1j$ production indicate that, although moderate, corrections are still significant [86, 87]. NNLO corrections with respect to NLO can be as large as 25% in the range of interest. In order to accommodate these corrections, a conservative approach was implemented. The Higgs p_{T} distribution with $p_{\mathrm{T}} > 30$ GeV was corrected with the NNLO/NLO k-factors provided in Reference [86]. The MC described above was normalised to the total ggF cross section at NNLO+NNLL from Reference [33]. It has since been shown that the total ggF cross section at N3LO is indicative of small N3LO/NNLO k-factors and scale variations [29]. For this reason the cross section with $p_{\mathrm{T}} < 30$ GeV (i.e. the region not covered in the p_{T} spectrum from Reference [86]) was re-scaled appropriately so that the total cross section matches the total cross section at NNLO+NNLL. The scale uncertainties assumed in this analysis remain at NNLO for the total cross section and at NLO for the $p_{\mathrm{T}h}$, while the PDF uncertainties were conservatively taken from Reference [33].

The results related to top associated Higgs production were also included to provide constraints to the model. Associated th production in the SM is suppressed due to the negative interference induced by the relative sign of the Yukawa and h-W-W couplings [88]. If the H-W-W coupling is relatively suppressed, this negative interference is reduced, and so the tH cross section becomes comparable to that of ttH production. For this reason, β_V was set to a small value (order of 10^{-3}) and tH cross sections were determined at LO in `MadGraph5_aMC@NLO`. These cross sections were enhanced to NNLL+NLO by multiplying by an appropriate k-factor, and were then combined with ttH cross sections from Reference [33]. For the mass

values of the heavy scalar considered in this thesis the combined cross section from tH and ttH reached a value as high as 25 fb at $\sqrt{s} = 8$ TeV.

Constraints on the parameters β_g and m_H were calculated with a statistical combination of results using Pearson's chi-square [89] as a test statistic. For results that provide measurements of a quantity μ, this is calculated as:

$$\chi^2\left(\mu_{\text{pred}}\right) = \sum_i \frac{\left(\mu^i_{\text{pred}} - \mu^i_{\text{obs}}\right)^2}{\left(\Delta\mu^i\right)^2}, \tag{4.7}$$

where here "pred" and "obs" refer to the predicted and observed quantities, respectively, and $\Delta\mu$ is the total uncertainty. The chi-square is additive, and sums over the i independent measurements. In the case of combining results that are limits, one can extract the observed and expected limits (denoted as $\mu^{95\%,i}_{\text{obs}}$ and $\mu^{95\%,i}_{\text{exp}}$, respectively) and use them in a similar way:

$$\chi^2\left(\mu_{\text{pred}}\right) = \sum_i \frac{\left(\mu^i_{\text{pred}} - \left[\mu^{95\%,i}_{\text{obs}} - \mu^{95\%,i}_{\text{exp}}\right]\right)^2}{\left(\mu^{95\%,i}_{\text{exp}}/1.96\right)^2}. \tag{4.8}$$

The details of this method will be explained in Section 4.2.1.1. Using Equations (4.7) and (4.8), a combined chi-square was constructed by adding up the contributions from all of the results presented in Table 4. In the case of the Higgs p_{T} spectra, one unit of chi-square is calculated per bin. Firstly, the BRs of $H \to hh$ and $H \to VV$ were fixed by minimising a chi-square determined from experimental results. All of these BRs were used as inputs for a combined chi-square, which was calculated while floating the free parameters β_g and m_H. For each mass point, β_g was marginalised such that the combined chi-square was minimised. Errors on marginalised parameters were calculated from identifying the points in parameter space which differ by one unit of chi-square above and below the minimised value.

The technicalities of the fitting process are as follows. In the minimised global chi-square different values of m_H are treated as different hypotheses. This was done using a scan of mass points, starting at $m_H = 260$ GeV (since no di-Higgs search from Table 4 considers values smaller than this) and going up in 5 GeV

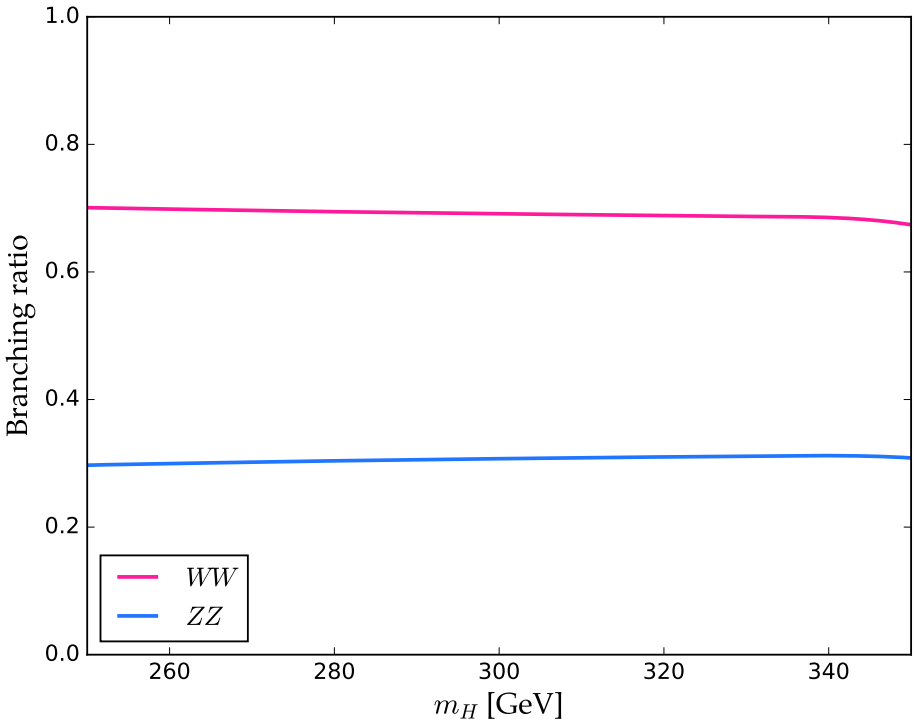


Figure 8: The dominant BRs for a Higgs-like boson in the mass range between 250 GeV and 350 GeV [33]. At these masses, the decays are dominated by $H \to VV$ decays. The slight decrease shown near to 350 GeV is a result of the mass coming close to the $2m_t$ threshold, where $H \to t\bar{t}$ decays would become dominant.

steps until 320 GeV. Points in between these were reached by an interpolation. The other parameters of the model were fixed by a number of constraints. Firstly, the branching ratio of $H \to hh$ was set to a value that is fit against the current di-Higgs boson resonance search limits set by ATLAS and CMS. Secondly, the branching ratio of $H \to VV$ was determined in the same way using ATLAS and CMS limits from searches for $H \to VV$ at high masses. An additional constraint was applied that fixes the ratio of the $H \to WW$ and $H \to ZZ$ BRs such that the ratio is SM-like (shown in Figure 8). Doing this indirectly constrains the associated couplings in Equation (4.3) and Equation (4.5). The remainder of the decay of the heavy scalar is assumed to be $H \to h\chi\chi$. Finally, the parameter β_g was constrained by fitting the ATLAS and CMS Higgs boson p_{T} spectra, as well as results from top associated Higgs boson production processes. There may exist other decay modes, such as $H \to \chi\chi$, although these other decays were not considered here. Adding in other decay modes would not change the final results of the analysis, however it would arguably allow for further constraints on the parameter β_g.

The number of degrees of freedom in the calculation of the global chi-square comes from the addition of all the channels mentioned in Table 4. The contribution from each channel is given in the third column of the table – this number refers to either the number of bins in a distribution (for example, the ATLAS $h \to \gamma\gamma$ p_{T} spectrum contributes 8 units), or the number of measurements/limits which were categorised and presented in the publication (for example, the ATLAS di-Higgs combination only contributes 1 unit since the combined limit is a single number etc.). The marginalised parameters in the fit do not contribute towards the number of degrees of freedom, and different masses are treated as different hypotheses. The only free parameter in the model (per mass hypothesis) is therefore β_g. Therefore, adding up the numbers in Table 4 and subtracting a unit gives a total of 39 degrees of freedom.

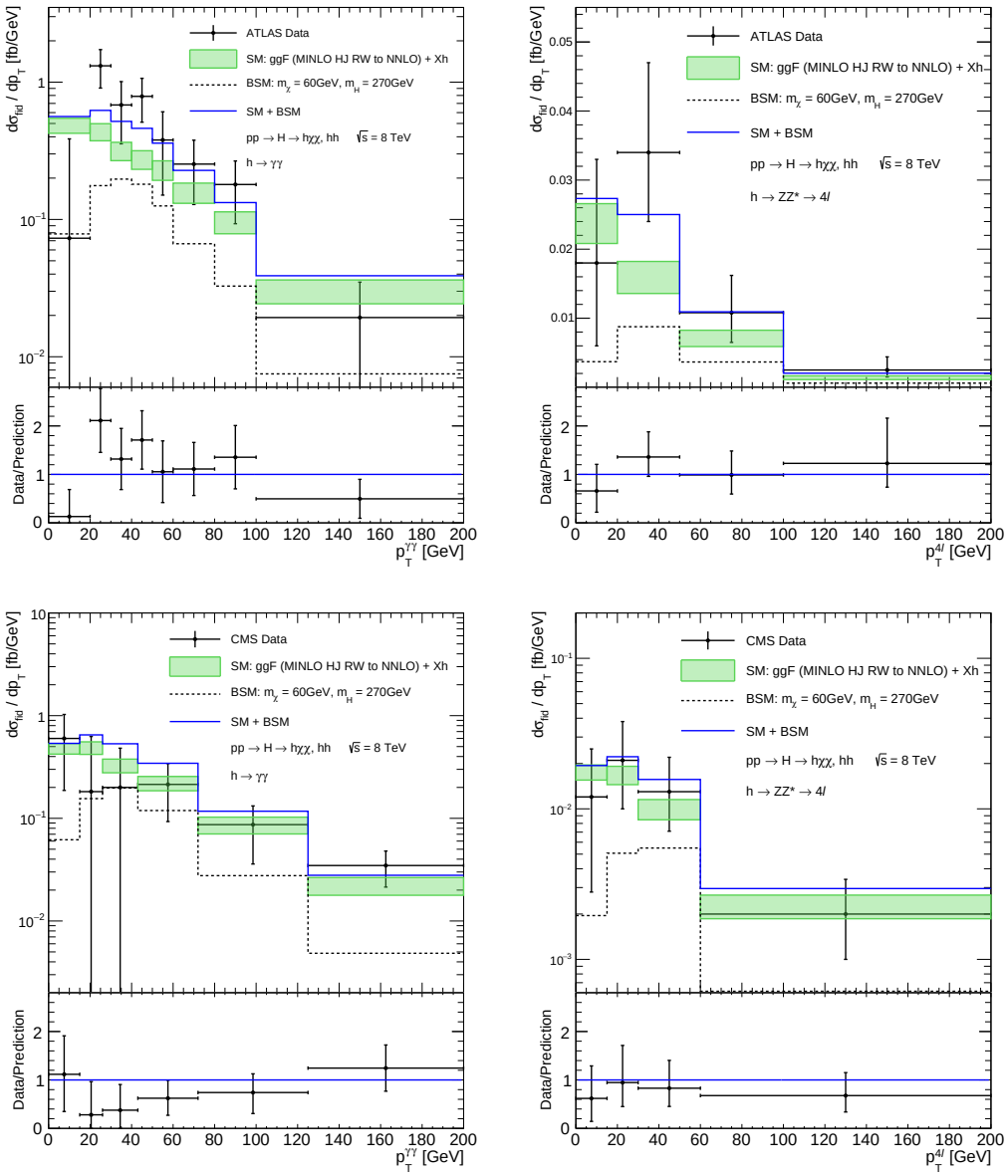


Figure 9: Fits to the fiducial differential distributions of the Higgs boson transverse momentum using the ATLAS di-photon (top left), the ATLAS $h \to ZZ^* \to 4\ell$ (top right), the CMS di-photon (bottom left) and CMS $h \to ZZ^* \to 4\ell$ (bottom right) decays (see text for detailed description). The mass points considered here are the best fit values of $m\chi = 60$ GeV and $m_H = 270$ GeV.

4.1.3 *Fit results and discussion*

When calculating and minimising the chi-square described above, it was found that the lower values of m_H fit the experimental data better than the higher values. Out of the mass points considered, the $m_H = 270$ GeV point was able to minimise the chi-square value to the smallest value. This point was determined using the best fit values of the BRs:

- $\mathrm{BR}(H \to hh) = 0.030 \pm 0.037$,
- $\mathrm{BR}(H \to ZZ) = 0.025 \pm 0.018$,
- $\mathrm{BR}(H \to WW) = 0.057 \pm 0.041$.

The parameter β_g was best fit at the value of 1.5 ± 0.6. The errors on these quantities correspond to a 1σ deviation from the mean value.

An indication of this parameter's best-fit value on the ATLAS and CMS p_{T} spectra can be seen in Figure 9. The fits to the p_{T} spectra were also able to constrain the mass of the DM candidate; for $m_H = 270$ GeV, m_χ was best fit at 60 GeV. This is relatively close to $m_h/2$, which naturally leads to the suppression of the $h \to \chi\chi$ BR and is consistent with current direct search limits. In Figure 9, the filled band corresponds to the total SM-only prediction. The width of the band indicates the size of the systematic uncertainties on the ggF process, according to the conservative scheme discussed in Section 4.1.2. These uncertainties are incorporated in the chi-square, and are folded into the statistical uncertainty by adding both in quadrature. The dotted line shows the contribution from the $H \to h\chi\chi$ and $H \to hh$ processes. The solid line corresponds to the sum of the SM and BSM components.

When interpolating between mass points, the combined minimised chi-square is found to be smallest at the value $m_H = 272$ GeV, with upper and lower errors being 12 GeV and 9 GeV, respectively. This can be seen in the top of Figure 10 where the solid blue line shows the lowest value of the minimised chi-square, and the dotted blue lines show a 1σ deviation from the value. The minimised value of chi-square has a lowest value of 0.72 per degree of freedom in the fit. In terms

of significance, the enhancement of the SM+BSM hypothesis compared to the SM-only (null) hypothesis is calculated as the difference in chi-square values for the two hypotheses,

$$\Delta\chi^2 = \chi^2_{\text{SM}} - \chi^2_{\text{SM+BSM}}. \tag{4.9}$$

For one degree of freedom, the significance is the square root of $\Delta\chi^2$. When comparing the two hypotheses, the improvement in explaining experimental data just surpasses a 3σ effect at the best-fit point, as can be seen in the bottom of Figure 10. In this figure, the large significance around $m_H = 260$ GeV can be attributed to the large $pp \to H \to hh$ cross sections in most of the ATLAS and CMS di-Higgs boson resonance search results. It is also relevant to note that results reported here do not change significantly if the NNLO corrections on $gg\text{F} + 1j$, discussed in the previous section, are not applied.

The most prominent consequence of the proposed phenomenological model considered here is the production of intermediate missing transverse energy in association with h. This is, however, mostly just a relic of the fact that the simplest assumption for a generic production mechanism of $h + X$ is that of h in association with DM. In addition, the data appear to have more jets in association with h than expected in the SM; this applies both to the inclusive production, and the production in association with top quarks. Enhanced QCD radiation in the production of H compared to that of direct h production would most likely not be sufficient to explain this effect. As a result of this reasoning, one possible extension to this simple model would be to consider the decay of a hypothetical intermediate particle S into hadronic jets (in addition to being responsible for the potential production of DM particles). This is discussed in Chapter 5.

The goal of investigating the simple model discussed in this section was to investigate whether or not some anomalies in the Run 1 LHC data were compatible with new physics processes. By and large, all of the anomalies discussed in this chapter, with the exception of the production of Higgs boson pairs ($H \to hh$), have remained to be compatible with future developments of the model and more re-

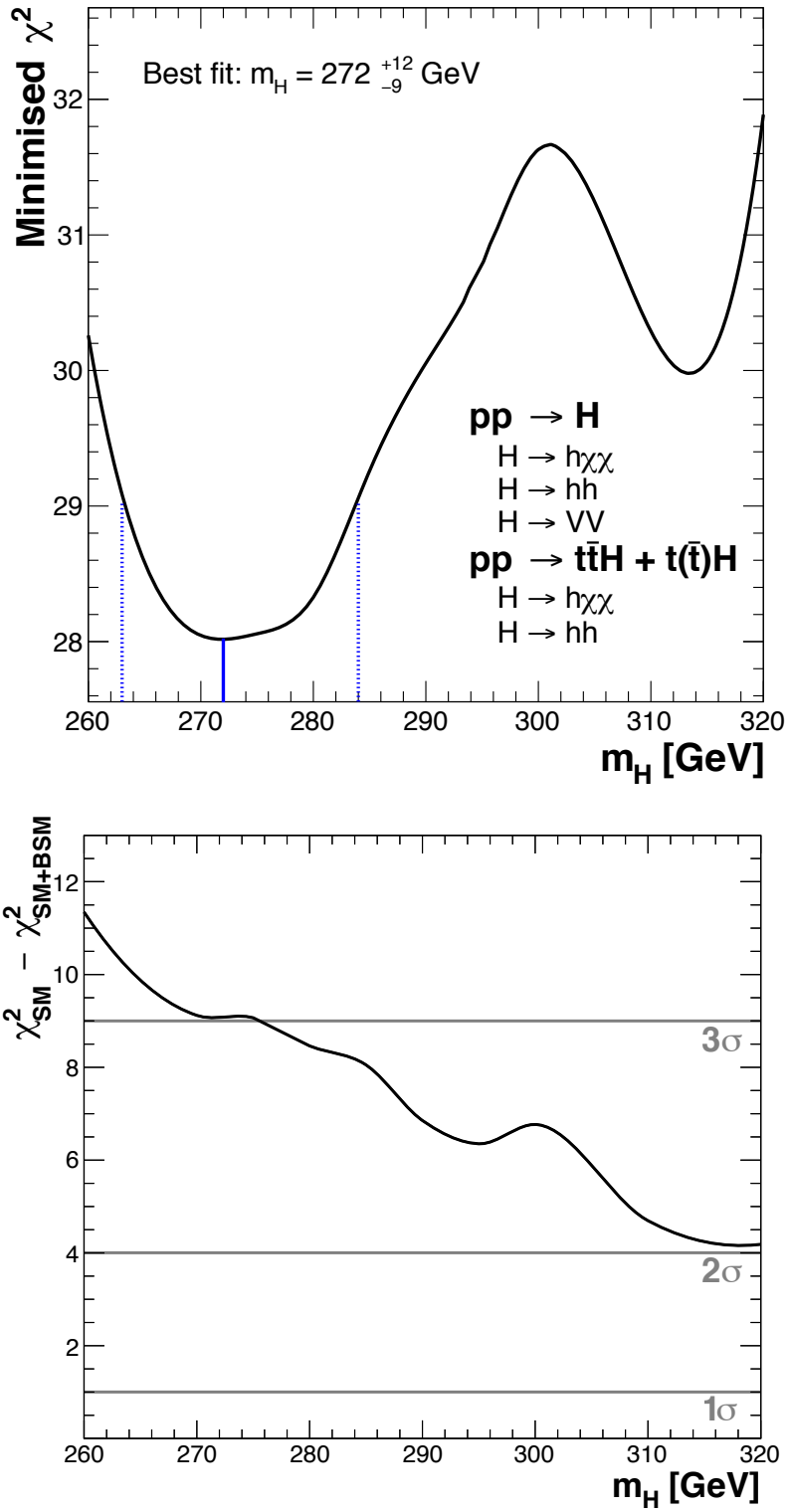


Figure 10: A scan of minimised chi-square values (top) and the $\Delta\chi^2$ test statistic (bottom) as a function of the parameter m_H. These pertain to the statistical combination of the experimental results in Table 4.

cent data sets. As newer data sets were released, the compatibility was re-checked. Several of these checks are shown in Section 4.2. The ability of this simple model to have constraints that allow for a 3σ improvement of the SM-only hypothesis for the niche ensemble of results in Table 4 warranted further studies, which are presented in the next two chapters. In particular, the following chapter expands the simple model introduced here to address certain pathological aspects of the theory, as well as provide a set of experimentally testable final states that have been used to further constrain the model's parameters.

4.2 COMPATIBILITY CHECKS WITH NEWER DATA SETS

The development of the BSM model introduced in Section 4.1 was a work in process that occurred during an active period of data-taking and analysis in the LHC collaborations. The fixed set of results that inspired and constrained the model (listed in Table 4) would continue to be updated with newer data sets and improved predictions over the years. It was therefore a constant challenge to make checks against newer data sets as they were released, and this section will show a few of the checks that were made before the model changed course towards constraints in the multiple lepton sector (this is described in Chapter 6).

The experimental results that were first deemed relevant for constraining the BSM model can be grouped into three different classes. These include the following:

- heavy resonance searches – the resonant production of vector boson pairs ($H \to WW$ and ZZ) and Higgs boson pairs ($H \to hh$) in various final states,
- the Higgs boson p_{T} spectrum, and
- the associated production of the Higgs boson with one or two top quarks.

In this section, these results will be explored in more detail, and the compatibility of the BSM model (in its initial construction from Section 4.1) with the updated data will be established.

4.2.1 *Heavy resonance searches*

Since the discovery of the Higgs boson, a large amount of focus has been placed on searches for new heavy particles not predicted by the SM. This is arguably the simplest possible way to hopefully discover BSM physics at the LHC. In principle, the observation of a resonant peak in a mass spectrum of some combination of certain final state objects would constitute direct evidence for a new particle being produced.

While searches for resonances can be made using essentially any combination of final states, in this work emphasis is placed on final states that would typically be produced by a "heavy Higgs boson." Searches for scalar extensions to the SM often make the assumption that the scalar boson in question has Higgs-like couplings to the SM particles, or in most cases that it is a component of some extended Higgs sector like a two Higgs doublet model (2HDM) (some formal extensions shall be studied in Chapter 5). In any case, the resonant searches considered in this work are all those that relate to an extended Higgs sector, and therefore the decays considered for a given heavy resonance H are all of the SM-like decay modes ($H \to ZZ, WW, Z\gamma, \gamma\gamma, bb$ etc.) and $H \to hh$. In addition to this, the searches considered are restricted to masses below $m_H = 2m_t$ (approximately 350 GeV) since any Higgs-like scalar above this mass would decay preferentially to pairs of top quarks due to the large t-t-H Yukawa coupling. A complete set of the ATLAS and CMS results pertaining to searches of this kind, which were available at the time of this study and with the appropriate mass constraints is listed in Table 5.

There is no single heavy resonance search result from either ATLAS or CMS that gives definitive evidence for the existence of a heavy BSM scalar. Because of this, each result has rather produced upper limits at the 95% confidence level (CL) for the production cross section of a hypothetical heavy resonance as a function of the resonance mass. While these limits contain useful information for constraining new physics theories, it is more useful in this work to think rather of best-fit mean cross sections and their associated standard deviation. This is useful for two reasons. Firstly, because it allows for the estimation of a combination of the independent

Collaboration	$\sqrt{s}$ [TeV]	L [fb^{-1}]	Reference	Process
		Run 1		
ATLAS	8	20.3	HIGG-2013-20 [71]	$H \to ZZ \to \ell\ell\ell\ell, \ell\ell\nu\nu, \ell\ell qq, \nu\nu qq$
	8	20.3	HIGG-2013-19 [70]	$H \to WW \to \ell\nu\ell\nu$
	8	20.3	HIGG-2014-04 [90]	$H \to \gamma\gamma$
	8	20.3	HIGG-2013-33 [62]	$H \to hh \to bb\tau\tau, WW\gamma\gamma, \gamma\gamma bb, bbbb$
CMS	7, 8	up to 24.8	HIG-13-031 [72]	$H \to ZZ, WW$ combination
	8	19.7	HIG-14-006 [91]	$H \to \gamma\gamma$
	7, 8	24.4	HIG-13-024 [92]	$H \to \mu\mu$
	8	18.3	HIG-15-013 [93]	$H \to hh \to bb\tau\tau$
	8	19.7	HIG-13-032 [94]	$H \to hh \to \gamma\gamma bb$
	8	17.9	HIG-14-013 [95]	$H \to hh \to bbbb$
	8	19.5	HIG-13-025 [65]	$H \to hh \to$ multi-leptons
		Run 2		
ATLAS	13	36.1	HIGG-2016-19 [96]	$H \to ZZ \to \ell\ell\ell\ell, \ell\ell\nu\nu$
	13	36.1	HIGG-2016-31 [97]	$H \to WW \to e\nu\mu\nu$
	13	36.7	HIGG-2016-17 [98]	$H \to \gamma\gamma$
	13	36.1	HIGG-2016-14 [99]	$H \to Z\gamma$
	13	36.1	HIGG-2016-12 [100]	$H \to \tau\tau$
	13	27.5-36.1	CONF-2018-043 [101]	$H \to hh \to bbbb, \gamma\gamma bb, bb\tau\tau$
	13	36.1	HIGG-2016-20 [102]	$H \to hh \to \gamma\gamma WW^*$
CMS	13	35.9	HIG-17-012 [103]	$H \to ZZ \to \ell\ell\ell\ell, \ell\ell qq, \ell\ell\nu\nu$
	13	35.9	HIG-17-020 [104]	$H \to \tau\tau$
	13	35.7	HIG-16-018 [105]	$H \to bb$
	13	35.9	PAS-HIG-17-030 [106]	$H \to hh \to bb\tau\tau, bbVV, bbbb, \gamma\gamma bb$

Table 5: A list of all published search results to date for heavy scalar resonances related to the Higgs sector, with the constraint that they contain information for resonance masses lower than 350 GeV.

data sets shown in Table 5, assuming that they are linked by some common physics model. Secondly, since in this work it is imperative to quantify the total production cross section of a hypothetical new scalar, the compatible values for the BRs of the new scalar and their uncertainties can be estimated more easily.

4.2.1.1 *Chi-square formalism*

One can estimate a best-fit mean cross section and standard deviation from a 95% CL relatively easily. The simple method, which follows on from the logic in Reference [107], makes the assumption that the likelihoods used to calculate 95% CLs are approximately Gaussian, in spite of the fact that Poisson probabilities are used. That is, if we observe n_{obs} events with a Poisson uncertainty of $\sqrt{n_{\text{obs}}}$, then the limiting case of the prior probability:

$$p(n_{\text{obs}}|n) \sim e^{-n} n^{n_{\text{obs}}} \xrightarrow{\text{large } n_{\text{obs}}} e^{(n-n_{\text{obs}})^2/2n_{\text{obs}}} \tag{4.10}$$

applies. In practice, this approximation is usually appropriate for $n_{\text{obs}} > 10$. Assuming that the considered limit is on a cross section times BR, we can work rather with the signal strength:

$$\mu^i = \frac{n_s^i}{(n_s^i)^{\text{SM}}} = \frac{\sum_p \sigma_p \epsilon_i^p}{\sum_p \sigma_p^{\text{SM}} \epsilon_i^p} \times \frac{\text{BR}_i}{\text{BR}_i^{\text{SM}}}, \tag{4.11}$$

for a particular channel i that is made up of the production processes p, each with a production cross section σ and analysis selection acceptance ϵ. Each channel is assumed to have a fixed BR, and μ^i is therefore a measure of the deviation from the predicted SM cross section times BR as a function of the measured cross section (or equivalently the number of predicted signal events n_s). In most cases and as given from the experimental results, one can usually assume a single production mode and equal BRs, thus simplifying Equation (4.11).

Recall that the expected and observed limits on μ ($\mu_{\text{exp}}^{95\%}$ and $\mu_{\text{obs}}^{95\%}$, respectively) are calculated using a likelihood function. Technically speaking, the limit comes from integrating the posterior probability distribution $p(\mu|n_{\text{obs}})$ as a function of μ and finds the value of $\mu = \mu_{\text{obs}}^{95\%}$ that separates the function such that 95% of the

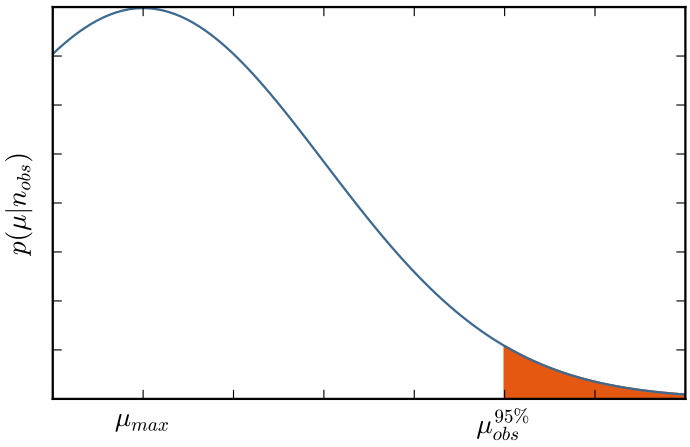


Figure 11: An illustration of the Gaussian approximation as constructed from the appropriate parameters in an observed 95% CL. The shaded area represents only 5% of the total area of the distribution, the point from which the 95% CL can be derived.

total area under the curve is contained. An illustration of this idea is shown in Figure 11. In the Gaussian approximation, the posterior probability for the observed limit takes the form:

$$p(\mu|n_{\text{obs}}) \simeq e^{-(\mu-\mu_{\text{max}})^2/2\sigma_{\text{obs}}^2}, \tag{4.12}$$

where μ_{max} is the value of μ that maximises the probability, and acts as a measure of the deviation of the observed data from the background prediction. The standard deviation of the posterior probability is σ_{obs}. Given this posterior probability and assuming it is normalised, the fundamental relation from which we can extract the mean and standard deviation of the quantity μ is [9]:

$$0.95 = \int_0^{\mu_{\text{obs}}^{95\%}} \mathrm{d}\mu \; p(\mu|n_{\text{obs}}). \tag{4.13}$$

Now, in order to calculate the approximated values of the mean and standard deviation of the quantity μ from Equation (4.13), the following steps are taken. Since μ_{max} is unknown, we first start with the expected limit. By setting n_{obs} to the background prediction, n_{b}, thereby shifting the distribution such that the max-

imum lies at $\mu = 0$. By integrating Equation (4.13) to the expected limit, we find:

$$0.95 \simeq \sqrt{\frac{2}{\pi\sigma^2_{\text{exp}}}} \int_0^{\mu^{95\%}_{\text{exp}}} \mathrm{d}\mu \; e^{-\mu^2/2\sigma^2_{\text{exp}}}. \tag{4.14}$$

Numerically integrating the normalised Gaussian, we find that:

$$\sigma_{\text{exp}} = \frac{\mu^{95\%}_{\text{exp}}}{1.96}. \tag{4.15}$$

If the expected signal (or deviation from the SM) is much smaller than the background, then $\sigma_{\text{exp}} \simeq \sigma_{\text{obs}}$, and the approximate mean of the deviation can be estimated as the shift in the distribution between the observed and expected limits, $\mu^{95\%}_{\text{obs}} - \mu^{95\%}_{\text{exp}}$.

A practical application of this extraction process is that one can combine results in the form of 95% CLs, given that they are limits on the same observable. Once the mean and standard deviation have been extracted from the limit, one can make a simple combination of other results in independent data sets with an error weighted mean (that is, a mean which uses the inverse square of the corresponding uncertainties as weights). A more formal approach, however, would be to minimise a chi-square function that combines the results. A simple chi-square function can be constructed using an adaptation of Pearson's test statistic, which was already introduced in Equation (4.7). This chi-square was written as a function of μ_{pred}, which could change depending on the parameters of the model used to calculate it. The corresponding uncertainty is calculated by varying the degrees of freedom of the prediction and seeing where the value of χ^2 is a distance of 1 from the minimum value. In the case of limits, and using the results derived above, the chi-square takes on the form of Equation (4.8), where the denominator is taken from Equation (4.15).

4.2.1.2 *Statistical combinations*

As mentioned above, this method can be used to combine compatible 95% CL results from an arbitrary number of independent data sets. For the purposes of this

thesis, this is useful in making a combination of the results in searches for resonant Higgs boson pairs (di-Higgs production). This is because these searches can involve a variety of final state configurations, based on the BRs of the Higgs boson. Only in some cases do the experimental collaborations perform a combination of di-Higgs searches. A statistical combination of the resonant di-Higgs search results shown in Table 5 was performed by minimising the χ^2 function in Equation (4.8). For the Run 1 data set, the result is shown in Figure 12 for the separate ATLAS and CMS results, as well as their combination. The same is shown for the Run 2 data set in Figure 13. The Run 1 and Run 2 results are not combined, since in all likelihood a hypothetical heavy scalar would have a different production cross section at 8 TeV compared to 13 TeV, and an estimate of how big the scaling would be is model-dependent and would introduce a bias to the study.

The Run 1 result shows that the maximum deviation from the SM occurs at a resonance mass at around 280 GeV. At this point, the best-fit cross section times BR is 654 ± 304 fb, corresponding to a 2.15σ deviation from the SM. During the time when the individual results had started coming out for the Run 1 data set, several excesses around this mass were part of the motivation for studying a hypothetical heavy scalar with a mass in the region of 250 GeV to 300 GeV. The check made on the Run 2 data set also has its maximum deviation from the SM at 280 GeV, with a corresponding best-fit cross section times BR equal to 400 ± 275 fb. The significance for the Run 2 deviation at this mass is 1.46σ. The combination of the two deviations gives a significance of 2.60σ. While this is not a striking excess in terms of significance, it is interesting to note that both the Run 1 and Run 2 data sets are consistent in their point of maximum deviation from the SM.

One issue that could be raised about the combinations discussed above is to what extent they are model-independent. The 95% CL results all include the effects of systematic uncertainties imposed on the BSM predictions. Therefore, the results are dependent on the production mechanisms considered of the hypothetical heavy BSM particle in every case. As detailed in Reference [108], these systematic uncertainties can be as big as 35%. In the same work, a modified chi-square function was developed in order to negate the systematic uncertainties on the sig-

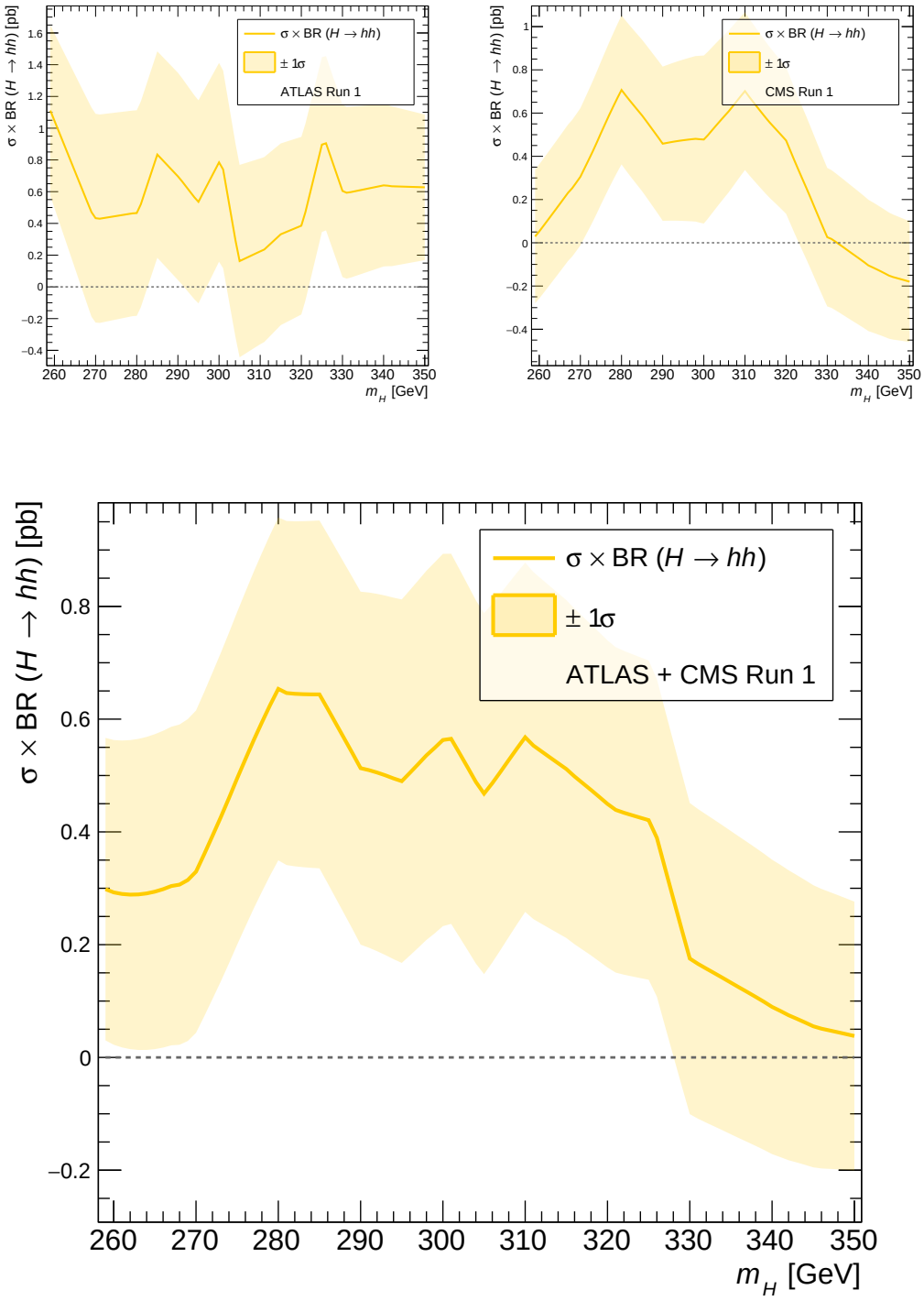


Figure 12: The statistical combination of the Run 1 di-Higgs search results that were presented in the form of 95% CLs, calculated using Equation (4.8) in terms of the cross section times BR of a resonance H decaying to hh and as a function of the resonance mass. This is done for the ATLAS only results (top-left), the CMS only results (top-right), and their combination (bottom). The solid line represents the best-fit mean cross section and the band represents a 1σ uncertainty.

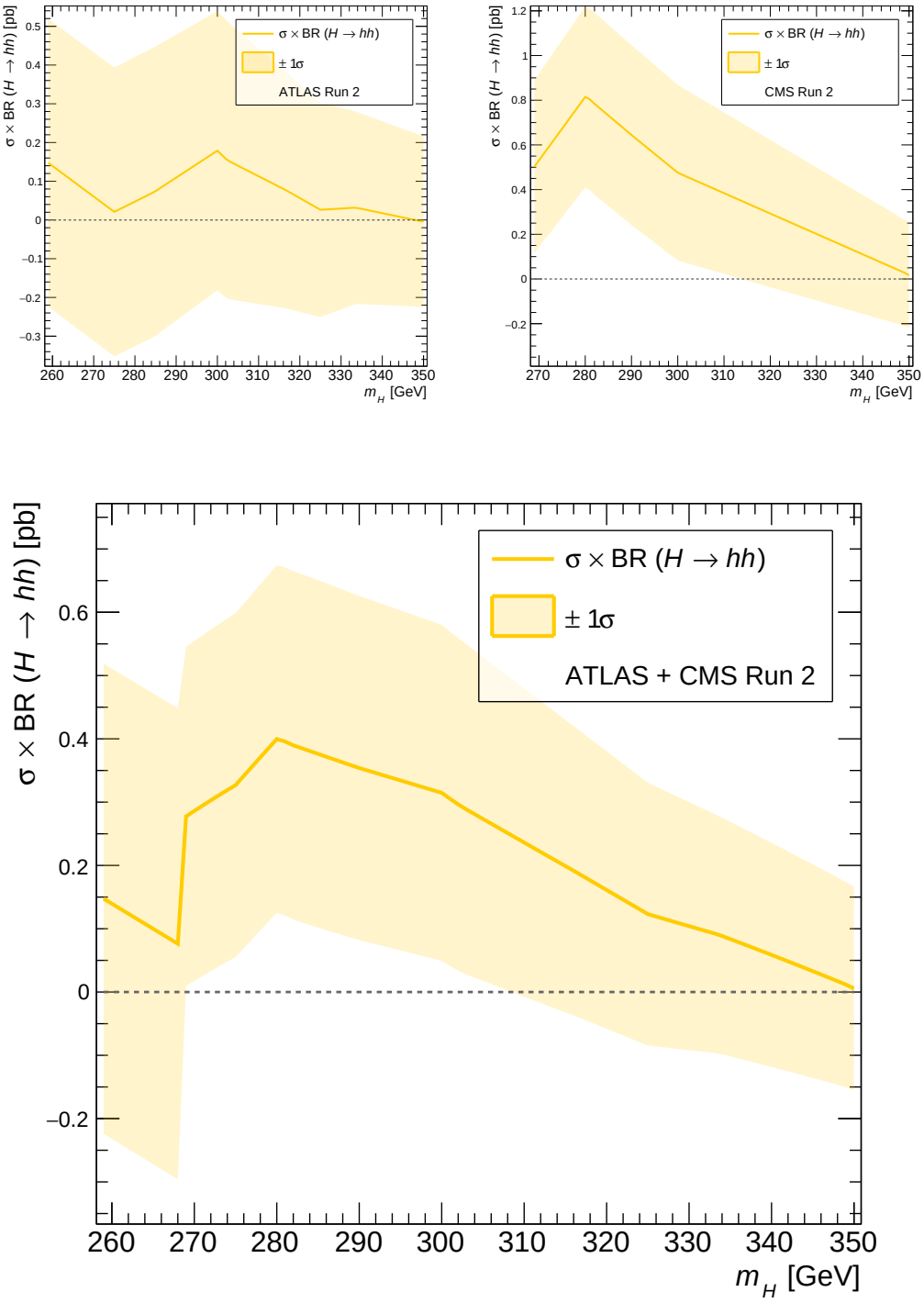


Figure 13: The statistical combination of the Run 2 di-Higgs search results that were presented in the form of 95% CLs, calculated using Equation (4.8) in terms of the cross section times BR of a resonance H decaying to hh and as a function of the resonance mass. This is done for the ATLAS only results (top-left), the CMS only results (top-right), and their combination (bottom). The solid line represents the best-fit mean cross section and the band represents a 1σ uncertainty.

nal model, to come to a more model-independent result. It was shown, however, that the effect on the significance for deviations from the SM-only hypothesis was negligible. The enhancement on the significance was less than 5% in total. Furthermore, the study assumed that the full effect of the systematic uncertainties constrained the statistical result. In practice, using a profile likelihood to compute limits would profile the systematics, greatly reducing their effect on the final result. Therefore, using the model-dependent results is an appropriate method for studying potential deviations from the SM.

4.2.2 *The Higgs p_T spectrum*

If any BSM Higgs boson production process were to be significantly observable at the LHC, it can be expected that the p_{T} spectrum of the Higgs boson would be affected characteristically according to the nature of the new physics process. Since the Higgs boson is produced dominantly through ggF at the LHC, its p_{T} spectrum as predicted by the SM peaks at a low value – since ggF is a $2 \to 1$ process, the p_{T} only deviates from 0 through the emission of ISR and due to higher order corrections. The associated production modes, commonly referred to collectively as Xh (VBF, Vh, tth, etc.), contribute to the Higgs p_{T} spectrum most significantly above ~ 20 GeV, although their combined cross section is about an order of magnitude smaller than that of ggF.

The measurement of the p_{T} of the Higgs boson is intrinsically a model-dependent procedure, due to the fact that both a signal and background model have to be assumed for purposes of signal extraction. In order to minimise on the model-dependence, the experimental results are typically presented in a *truth particle-level fiducial* region of the phase space. Designing a fiducial (or baseline) phase space region means placing cuts on the kinematic properties of the final state such that each of the known Higgs production mechanisms (i.e. those in Figure 4) have an approximately equal acceptance into the region. Therefore, no particular bias is given to any particular production mechanism, implying that any unknown production mechanisms could also be probed with the fiducial region.

The events are also presented in such a way that the final state particles are truth level particles. This means that the events in the data undergo an analysis that removes the effects of the detector, and gain the additional uncertainties that come along with this procedure. This makes it appropriate to compare the data with theoretical predictions without having to worry about simulating the response of the detector in question. The ways that this is achieved by the different experiments and in the different final states are discussed below.

The sections below discuss and show the ATLAS and CMS results for the Higgs p_T spectrum, measured in terms of differential fiducial cross sections, first for the $h \to \gamma\gamma$ decay mode, then the $h \to ZZ^* \to 4\ell$ mode, and finally the $h \to WW^* \to e\nu\mu\nu$ mode. The measurement of the data is compared with a state-of-the-art SM prediction. For the SM ggF production mode, the differential distribution is accurate up to NNLO accuracy using the `NNLOPS` method [109]. The total cross section has been scaled to the N3LO-accurate result from Reference [33]. The generated events were passed through `Rivet` routines [83] that replicate the fiducial volumes of phase space as designed by the experiments. The Xh production modes are accurate up to NLO corrections, and are taken directly from the publications themselves. These are added to the ggF distribution, and the total result is the green graph in Figure 14, Figure 15 and Figure 16. The uncertainty band corresponds to the ggF theoretical uncertainty, which was conservatively calculated to be 7.44% from Reference [33]. Each plot is accompanied by a set of statistical results showing the extent of the compatibility of the SM prediction with the data. These are the Kolmogorov-Smirnov (KS) test, a p-value and a chi-square per degree of freedom.

4.2.2.1 *In the di-photon channel*

The ATLAS experiment has performed a measurement on the differential distribution of the Higgs boson's fiducial production cross section as a function of its p_T using 20.3 fb^{-1} of the Run 1 data set [58] and 36.1 fb^{-1} of the Run 2 data set [110].The fiducial phase space is similar for both results. Two high p_T photons are selected in the η region of $|\eta| < 2.37$, excluding the gap region of the detector where $1.37 < |\eta| < 1.52$. The leading photon is required to have a p_T of at

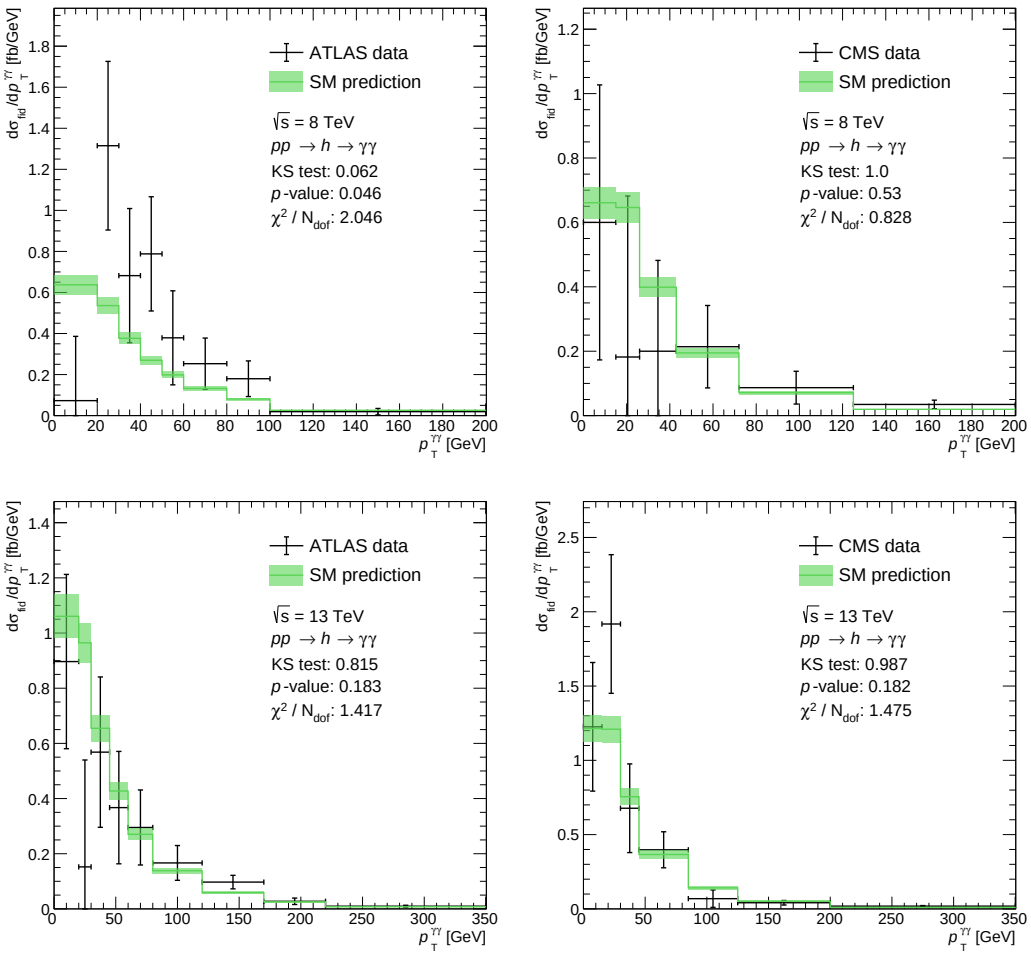


Figure 14: The SM Higgs p_{T} spectrum in the $\gamma\gamma$ decay mode, compared with the ATLAS (left) and CMS (right) data from Run 1 (top) and Run 2 (bottom).

least $0.35m_{\gamma\gamma}$, whereas for the sub-leading photon the minimum is $0.25m_{\gamma\gamma}$. The main difference between the Run 1 and Run 2 selection criteria is that different thresholds are used for photon isolation.

The CMS Run 1 (with 19.7 fb^{-1}) [60] and Run 2 (with 35.9 fb^{-1}) [111] fiducial cross sections make use of similar selection criteria to ATLAS, with some small differences. The pseudo-rapidity cuts are $|\eta| < 2.5$, where the gap region $1.44 < |\eta| < 1.57$ is excluded. The p_{T} cut on the leading photon is slightly lower for CMS, at $\frac{1}{3}m_{\gamma\gamma}$. One technical difference that CMS make (in contrast to ATLAS) is that the regularisation technique is not used for their unfolding procedure, as they claim it artificially reduces uncertainties.

For both ATLAS and CMS, the differential fiducial cross sections are constructed by binning the events into different ranges of the reconstructed Higgs p_T. In each of these bins, an un-binned signal plus background fit is made to the di-photon mass continuum background added to a signal model for the Higgs boson peak (fixing the mass of the Higgs boson). The resulting signal strength in each bin is then used to calculate the fiducial cross section in that bin. The resulting distributions can be seen in Figure 14.

The Run 1 ATLAS result in Figure 14 was the first differential Higgs boson p_T spectrum at the LHC, and at the time seemed to be quite incompatible with the most up-to-date SM predictions available. The apparent excess of cross section in the bins between 20 GeV and 100 GeV gave the impression that there may be a hint of additional BSM production mechanisms of the Higgs boson at the LHC. This appeared to have been corroborated by the Run 1 ATLAS $h \to 4\ell$ result (discussed in Section 4.2.2.2), which also showed an excess in the same range of Higgs p_T. Although the CMS result in Run 1 did not seem to indicate an excess, the combination of all the Run 1 results seemed to be consistent with the idea of a new Higgs production mechanism, as was shown in Section 4.1. The idea that was considered is that the Higgs boson was produced in association with some unidentified final state particles through the decay of a heavy boson H, i.e. $H \to h + X$. This is explored in depth in this thesis, particularly in Section 4.1 and Chapters 5 and 6.

4.2.2.2 *In the $ZZ^* \to 4\ell$ channel*

In the same spirit as the di-photon results shown above, the differential Higgs p_T spectrum for the $h \to ZZ^* \to 4\ell$ has been shown for both ATLAS in Run 1 [59] and Run 2 [112], as well as CMS in Run 1 [61] and Run 2 [113]. Due to the multiplicity of leptons in the searches (and furthermore the need to associate the leptons to their originating Z boson), the exact fiducial regions of phase space can be quite detailed. While the details can be found in the respective publications, the general principle in these searches is to analyse the events with an algorithm that finds a lepton quadruplet and pairs it with the leading or sub-leading Z boson candidate.

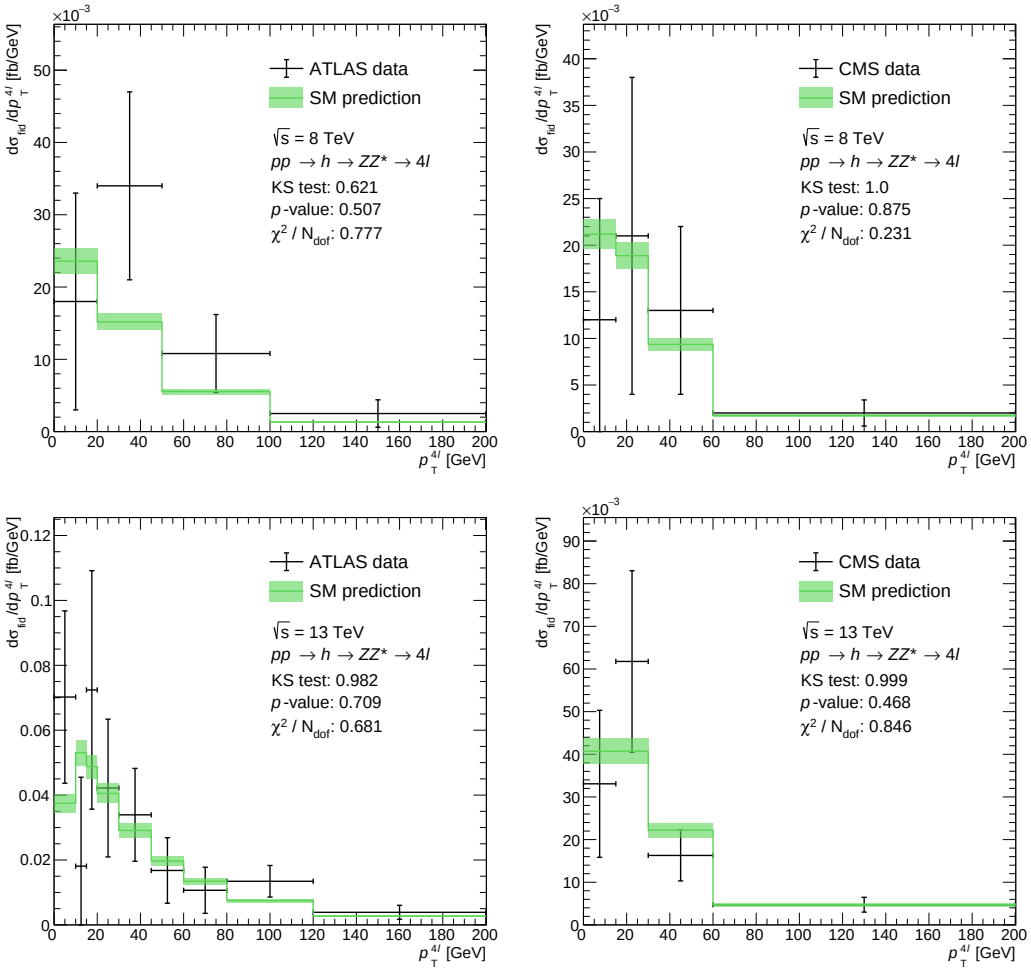


Figure 15: The SM Higgs p_{T} spectrum in the $ZZ^* \to 4\ell$ decay mode, compared with the ATLAS (left) and CMS (right) data from Run 1 (top) and Run 2 (bottom).

As for performing the actual measurement of the Higgs boson p_{T}, ATLAS use template fits in different bins of p_{T}, whereas CMS construct analytic signal models and perform an un-binned fit in different bins of p_{T}. Similarly to the di-photon results, the data undergoes an unfolding such that one can compare the data distribution with that of particle level MC predictions. The set of distributions for ATLAS and CMS in both runs can be seen in Figure 15.

As mentioned for the di-photon results above, the early excess in the range of p_{T} between 20 GeV and 100 GeV was apparent in the ATLAS Run 1 result shown in Figure 15. At first, this appeared to be an upward fluctuation of the data. As more data was taken, the structure gained more resolution and narrowed to an excess at around 30 GeV (as opposed to the initial broad fluctuation). This would

seem like an interesting feature, since it is supported by the CMS Run 2 di-photon result in Figure 14. The ATLAS Run 2 di-photon result in fact has a depletion of cross section at 30 GeV, but the overall combination of Higgs p_T results is still consistent with the BSM model as it stands from Section 4.1, and as it is developed throughout the next two chapters.

4.2.2.3 *In the $WW^* \to e\nu\mu\nu$ channel*

Measurements of the Higgs p_T spectrum in its WW^* decay mode were performed on the ATLAS [114] and CMS [115] Run 1 data sets. Both of these searches make use of the $e\nu\mu\nu$ final state channel, due to the fact that significant backgrounds can be suppressed by requiring E_T^{miss} in the final state and different flavour leptons. For events to be accepted into the fiducial volume of phase space required by these measurements, there needs to be exactly two leptons in the final state (one electron and one muon) with opposite charge. The details of the selection criteria for the leptons can be seen in the experimental papers, although it should be noted that while ATLAS required an explicit cut on E_T^{miss} (that it should be greater than 20 GeV), CMS applied an implicit cut on E_T^{miss} via a cut on the p_T of the di-lepton system and a transverse mass cut on the full set of final state particles. The latter approach is extremely model-dependent and biased towards the direct production of an SM-like Higgs boson.

Similar to the measurement in the ATLAS $ZZ^* \to 4\ell$ search, the differential p_T spectrum here is constructed by making binned template fits using MC background distributions (some of which are normalised using data-driven methods). This is done using relatively coarse bins in p_T for both ATLAS and CMS, due to the fact that the large systematic uncertainties introduced by requiring a more diverse final state can be made worse when considering potential bin migrations. The data compared with the SM prediction for both measurements can be seen in Figure 16.

One obvious observation from the measurements is that the CMS data distribution is far more compatible with the SM prediction than that of ATLAS. The ATLAS result seems consistent with the idea that there exists a small enhancement of Higgs p_T in the region around 30-40 GeV. Attempting to interpret these results

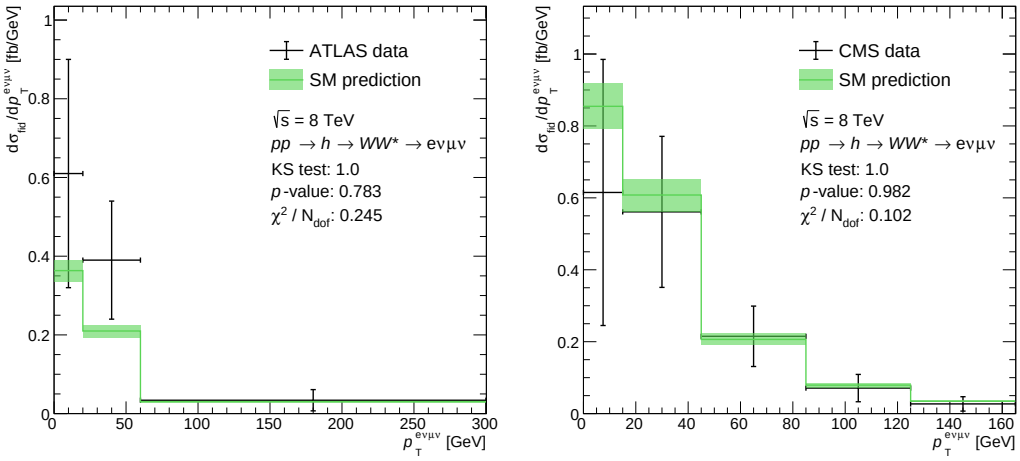


Figure 16: The SM Higgs p_{T} spectrum in the $WW^* \to e\nu\mu\nu$ decay mode, compared with the ATLAS (left) and CMS (right) data from Run 1.

in terms of possible new physics processes is relatively difficult for this final state. One reason is that if we consider a new production mechanism of the Higgs boson in association with some as-yet unspecified physics (i.e. $h + X$), it could be argued that X is a source of leptons or $E_{\mathrm{T}}^{\mathrm{miss}}$. Doing this would distort any distributions relating to $E_{\mathrm{T}}^{\mathrm{miss}}$, and potentially cause the new process to be vetoed due to having more than 2 leptons in the final state. Since the di-photon and four lepton search channels are able to reconstruct the Higgs boson in a less model-dependent way, they are certainly more useful in searches for new physics via the Higgs p_{T} spectrum.

4.2.3 *Top associated Higgs production*

Searches for the Higgs boson at the LHC are most sensitive to its dominant production mode, ggF. This direct production mode arises through the t-t-h Yukawa coupling, which is large due to the large mass of the top quark (the heaviest particle in the SM). In order to probe this Yukawa coupling more dynamically, one should rather look at the associated production of the Higgs boson with top quarks. The representative Feynman diagram for the tth production mode can be seen in Figure 4d.

In 2018, both the ATLAS [116] and CMS [117] collaborations reported an observation of production of the Higgs boson in association with top quarks at the LHC. This observation (which favoured the signal plus background hypothesis over the background-only hypothesis at the level of 5σ) was expected to have been made, although both results had shown a small enhancement on the tth production cross section compared with the SM expectation. The CMS measurement of the signal strength μ_{tth}, which made use of the Run 1 and Run 2 data sets, was found to be $1.26^{+0.31}_{-0.26}$. The ATLAS measurement, which made use of the Run 2 data set only (albeit with more data than CMS), was found to be $1.32^{+0.28}_{-0.26}$.

While the excesses noted above are quite insignificant, the results pertaining only to multi-lepton final states seem to show a more significant excess. In this thesis, the idea of introducing a new Higgs production mechanism $pp \to h + X$ is considered. As shall be discussed in detail in the following chapters, it is interesting to explore the associated production of a Higgs boson with multiple leptons. When one considers all of the tth results from ATLAS and CMS, and make a simple combination by calculating an error weighted mean of μ_{tth} for each result, one finds that $\mu_{tth} = 1.92 \pm 0.38$ [118, 119]. This is a far more significant deviation from the SM expectation of $\mu_{tth} = 1$. A breakdown of all these results and their approximate combination are shown in Table 6. It should be noted that the combination here does not account for potential correlations between systematic uncertainties between the different measurements, although it is unclear whether or not this will have a significant effect. What is more important is a comprehensive treatment of the background estimations, which do not always appear to have conservative constraints and can be artificially enhanced to improve the SM description of the data. This will be explored in a future work.

4.2.4 *Further comments*

The compatibility checks noted in this section were performed up until 2017, where new developments in the model led to a detailed study of the multiple lepton sector. The motivation for this is presented in Chapter 5, and the study is shown

Reference	Channel	Measured μ_{tth}
CMS Run 1 [69]	Same-sign 2ℓ	$5.3^{+2.1}_{-1.8}$
	3ℓ	$3.1^{+2.4}_{-2.0}$
	4ℓ	$-4.7^{+5.0}_{-1.3}$
	Combination	$2.8^{+1.0}_{-0.9}$
ATLAS Run 1 [67]	$2\ell 0\tau_{\text{had}}$	$2.8^{+2.1}_{-1.9}$
	3ℓ	$2.8^{+2.2}_{-1.8}$
	$2\ell 1\tau_{\text{had}}$	$-0.9^{+3.1}_{-2.0}$
	4ℓ	$1.8^{+6.9}_{-2.0}$
	$1\ell 2\tau_{\text{had}}$	$-9.6^{+9.6}_{-9.7}$
	Combination	$2.1^{+1.4}_{-1.2}$
CMS Run 2 [120]	Same-sign 2ℓ	$1.7^{+0.6}_{-0.5}$
	3ℓ	$1.0^{+0.8}_{-0.7}$
	4ℓ	$0.9^{+2.3}_{-1.6}$
	Combination	$1.5^{+0.5}_{-0.5}$
ATLAS Run 2 [121]	$2\ell 0\tau_{\text{had}}$	$4.0^{+2.1}_{-1.7}$
	3ℓ	$0.5^{+1.7}_{-1.6}$
	$2\ell 1\tau_{\text{had}}$	$6.2^{+3.6}_{-2.7}$
	4ℓ	< 2.2
	Combination	$2.5^{+1.3}_{-1.1}$
Error weighted mean		1.92 ± 0.38

Table 6: The measured μ values for tth production in multi-leptonic analysis channels. A combination is estimated as the error weighted mean of each quoted combined result.

in Chapter 6. Therefore, dedicated studies of the model as applied to the Higgs p_{T} spectrum, di-Higgs searches, and tth results were not a focus. However, one can still consider the most recent data sets and make note of their compatibility with the BSM model.

The most recent and sensitive search for resonant di-Higgs production is the ATLAS Run 2 combination of searches [122]. From this result, it appears that the excesses seen at around 280 GeV in the earlier data sets (shown in Figures 12 and 13) have diminished. However, all the other anomalies discussed in this chapter have remained as noticeable deviations from the SM in recent results. In terms of heavy resonance searches, the ATLAS search for a high mass scalar decaying to four leptons is presented in Chapter 7. In this search, a mild excess is seen at an invariant mass of around 245 GeV. An independent study on the ATLAS and

CMS four lepton and CMS di-muon searches [123] has shown that these hints are consistent over independent data sets and are also compatible with a large width scalar with a mass of around 265 GeV; this can be seen in Figure 39. This is also discussed in the conclusions of Chapter 7.

As for the Higgs p_{T} spectrum, the most recent measurements [124, 125] corroborate the observation that the early Run 1 ATLAS results were likely an upward fluctuation of the data, but remain compatible with the results obtained in Section 4.1. The shape of the Higgs p_{T} spectrum has become more narrow and has a peaked structure at around 30 GeV. In the studies performed in References [126, 118], it was shown that the SM normalisation of Higgs production should necessarily be lowered under the assumption of a new Higgs production mechanism. Therefore, the Higgs p_{T} spectrum remains compatible with the BSM model in its current form [5]. The most recent measurement of the Higgs boson signal strength from ATLAS indicates that the rate of Higgs boson production is around 1.11 times the predicted SM rate [127], which is in agreement with the statements above and in References [126, 118]. Future studies could also consider the more robust variable of the Higgs boson rapidity, for which the data distribution appears elevated in the central region.

Finally, the measurement of tth production in multiple lepton final states has yet to be updated in a dedicated result. At this point, compatibility with the initial anomalous LHC results that inspired the model have become less important in the face of newer studies, such as the one presented in Chapter 6. A separate study of multi-lepton results using the Run 1 ATLAS and CMS data sets was performed in Reference [4] (this publication is a pre-cursor to the study shown in Chapter 6). This study concluded that an independent set of LHC results could also be improved under the BSM hypothesis introduced in this chapter and Chapter 5. Therefore, the 3σ result obtained in Section 4.1 does not fully encapsulate the deviations that were already present in the Run 1 data. The actual deviation from the SM in Run 1 is more significant than what has been discussed in the main content of this chapter, an example of which can be seen in Reference [4].

5

THE PRODUCTION OF H IN A COMPLETE THEORY

The model described in Section 4.1 was a starting point that could be used to explore several anomalous features of the LHC data at the time of its inception. However, the overly simplified way in which it was constructed meant that it was not treated appropriately in terms of its theoretical underpinning. Arbitrarily adding fields as an extension of the SM is not always justifiable theoretically. In the following sections, the idea of embedding H into a 2HDM is explored, and thereafter the necessary details for adding additional scalars to the theory are described. In doing so, a set of model dependent final states are presented as signatures for the specific model that is chosen. Several of these signatures are studied in detail so as to present a few key distributions that could be studied further in experimental searches [7].

5.1 H IN A TWO HIGGS DOUBLET MODEL

This section briefly covers the 2HDM and its basic particle content. A fairly recent review of 2HDM phenomenology can be found in Reference [128]. The generalised Lagrangian for the 2HDM can be written as follows:

$$\mathcal{L}_{\text{2HDM}} = \left(D_\mu \Phi_1\right)^\dagger \left(D_\mu \Phi_1\right) + \left(D_\mu \Phi_2\right)^\dagger \left(D_\mu \Phi_2\right) - \mathcal{V}\left(\Phi_1, \Phi_2\right) + \mathcal{L}_{\text{int}}, \tag{5.1}$$

where Φ_1 and Φ_2 are two complex $SU(2)_{\text{L}}$ doublet scalar fields and D_μ is the covariant derivative necessary for interactions with the gauge fields. The sector $\mathcal{L}_{\text{int}}$ contains all possible interaction terms, including the SM Lagrangian, and

$\mathcal{V}(\Phi_1, \Phi_2)$ is the most general renormalisable scalar potential of the 2HDM. The potential can be written in the following way:

$$\begin{aligned}\mathcal{V}(\Phi_1,\Phi_2) = {} & m_1^2\Phi_1^\dagger\Phi_1 + m_2^2\Phi_2^\dagger\Phi_2 - m_{12}^2\Phi_1^\dagger\Phi_2 + \frac{1}{2}\lambda_1\left(\Phi_1^\dagger\Phi_1\right)^2 + \frac{1}{2}\lambda_2\left(\Phi_2^\dagger\Phi_2\right)^2 \\ & + \lambda_3\left(\Phi_1^\dagger\Phi_1\right)\left(\Phi_2^\dagger\Phi_2\right) + \lambda_4\left|\Phi_1^\dagger\Phi_2\right|^2 + \frac{1}{2}\lambda_5\left(\Phi_1^\dagger\Phi_2\right)^2 \\ & + \left[\lambda_6\left(\Phi_1^\dagger\Phi_1\right) + \lambda_7\left(\Phi_2^\dagger\Phi_2\right)\right]\Phi_1^\dagger\Phi_2 + \text{h.c.} \end{aligned} \tag{5.2}$$

The potential terms that cross multiply the two doublets and multiply the parameters m_{12}, λ_5, λ_6 and λ_7 are complex. They can therefore act as sources of CP violation. All the other terms in the potential are real. Note that the parameters in Equation (5.2) can be modified via a change of basis. The basis shown here is known as the *Higgs basis*.

After a spontaneous EWSB, the 2HDM generates a spectrum of five physical Higgs bosons. Namely, these are one charged Higgs pair ($H^\pm$), one CP-odd scalar (A), and two CP-even scalars (h and H), where by convention $m_H > m_h$. The heavy scalar considered in this thesis can be mapped onto the heavy Higgs boson H, while the SM Higgs boson can be mapped onto h. The potential can be transformed to the *physical basis* (that is, the basis in which the relevant parameters are mass eigenstates) through the mixing angles α, which diagonalises the CP-even Higgs squared-mass matrix, and β, which diagonalises both the CP-odd and charged Higgs sectors. One typically considers the value $\tan\beta$, which is the ratio of the VEVs for each Higgs doublet, v_2/v_1. The physical VEV is $v = \sqrt{v_1^2 + v_2^2}$, and it is approximately equal to 246 GeV.

Further choices of symmetries and couplings to the quarks and leptons can be made, which lead to different types of 2HDMs. Models which lead to natural flavour conservation are called the Type-I, Type-II, lepton-specific or flipped 2HDMs. These choices are detailed in Reference [128]. In this work a Type-II 2HDM has been used as a baseline, into which the additional scalars are added.

Several generic constraints can be imposed on any new BSM model so as not to impose on what has already been observed in nature. Therefore, the following constraints apply to the Type-II 2HDM considered here:

- Vacuum stability: the Higgs potential must be bounded from below such that the vacuum has a well defined minimum in the theory. Requiring this implies that the following conditions for the λ parameters must be simultaneously satisfied [128]:

$$\lambda_1 > 0;\ \lambda_2 > 0;\ \lambda_3 > -\sqrt{\lambda_1\lambda_2};\ \lambda_3 + \lambda_4 - |\lambda_5| > -\sqrt{\lambda_1\lambda_2}. \tag{5.3}$$

- Perturbativity: in order for the theory to predict calculable quantities in perturbative QFT, the bare quartic couplings in the Higgs potential need to have the upper bounds $|\lambda_i| < 4\pi$ for $i = 1, 2, .., 7$. This also implies that the magnitudes of the quartic couplings among the scalars after symmetry breaking should also be smaller than 4π.

- EW precision tests: the EW precision observables, such as the S, T and U oblique parameters [129], obtain small contributions from the extra scalars in 2HDMs due to quantum corrections. These contributions are ΔS, ΔT and ΔU, and their magnitudes have limits [130].

- Other experimental constraints, such as the LEP bounds, flavour-changing neutral current (FCNC) constraints [131], limits from the Higgs data at the LHC etc. can stringently restrict the model's parameters.

Various phenomenological studies and constraints on 2HDMs have been explored in the literature (sometimes with special emphasis on the Type-II 2HDM), a few examples of which can be found in References [132, 133, 134]. Any multi-Higgs-doublet model can, in principle, be severely constrained by tree level FCNCs. These potentially dangerous interactions can be avoided by imposing a variety of different discrete symmetries. One such discrete symmetry that avoids anomalous FCNCs is a $\mathbb{Z}_2$ symmetry, which demands an invariance of the general scalar potential under the transformations $\Phi_1 \to -\Phi_1$ and $\Phi_2 \to -\Phi_2$. This discrete $\mathbb{Z}_2$ symmetry can be imposed in three different ways. Firstly, it could be *exact* if m_{12}, λ_6 and λ_7 all vanish, thus implying a CP conserving scalar potential. Secondly, it could be broken *softly* if it is chosen to be violated in the quadratic terms only (i.e.,

in the limit where λ_6, λ_7 vanish, but m_{12} remains non-zero). Thirdly, one could impose a *hard* breaking if it is broken by the quadratic terms too, where the parameters m_{12}, λ_6 and λ_7 are all non-vanishing [128].

In a Type-II 2HDM, the discrete $\mathbb{Z}_2$ symmetry applies to both $\Phi \to -\Phi$ as well as to the charged leptons and down-type quarks $\psi_R^a \to -\psi_R^a$ (where a represents the generation index). However, in this work the terms associated with λ_6 and λ_7 are neglected, and m_{12} is taken to be real. In doing so, the quadratic couplings in terms of the physical masses of the 2HDM scalars can be expressed as follows:

$$\lambda_1 = \frac{1}{v^2 \cos^2\beta}\left(m_H^2 \cos^2\alpha + v^2 m_h^2 \sin^2\alpha - m_{12}^2 \tan\beta\right), \tag{5.4}$$

$$\lambda_2 = \frac{1}{v^2 \sin^2\beta}\left(m_H^2 \sin^2\alpha + v^2 m_h^2 \cos^2\alpha - m_{12}^2 \cot\beta\right), \tag{5.5}$$

$$\lambda_3 = \frac{1}{v^2}\left[2m_{H^+}^2 + \frac{\sin(2\alpha)}{\sin(2\beta)}\left(m_H^2 - m_h^2\right) - \frac{m_{12}^2}{\sin\beta\cos\beta}\right], \tag{5.6}$$

$$\lambda_4 = \frac{1}{v^2}\left[\left(m_A^2 - 2m_{H^+}^2\right) + \frac{m_{12}^2}{\sin\beta\cos\beta}\right], \tag{5.7}$$

$$\lambda_5 = \frac{1}{v^2}\left(\frac{m_{12}^2}{\sin\beta\cos\beta} - m_A^2\right). \tag{5.8}$$

5.2 ADDING THE DARK MATTER CANDIDATE χ

The simplified study shown in Section 4.1 requires that a Higgs boson is produced through the decay of a heavy scalar H in association with a DM candidate χ. In this section, we shall consider the possibility of accommodating χ in a complete theory. The addition of χ as a real scalar in a 2HDM requires additional terms in the potential defined in Equation (5.2). The χ field can act as a gauge-singlet scalar as well as a stable DM candidate if its mixing with the doublets Φ_1 and Φ_2 is prevented by introducing a discrete symmetry.

One possible symmetry is a $\mathbb{Z}_2$ symmetry under which χ is odd and all other fields are even. This ensures the stability of χ, as long as it is not too heavy. The most general potential consistent with the gauge and $\mathbb{Z}_2$ symmetries is therefore:

$$\begin{aligned}\mathcal{V}(\Phi_1,\Phi_2,\chi) = \mathcal{V}(\Phi_1,\Phi_2) + \frac{1}{2}m_\chi^2\chi^2 + \frac{\lambda_{\chi_1}}{2}\Phi_1^\dagger\Phi_1\chi^2 + \frac{\lambda_{\chi_2}}{2}\Phi_2^\dagger\Phi_2\chi^2 \\ + \frac{\lambda_{\chi_3}}{4}\Phi_1^\dagger\Phi_2\chi^2 + \frac{\lambda_{\chi_4}}{8}\chi^4 + \text{h.c.}\end{aligned} \tag{5.9}$$

The *hard* breaking of this $\mathbb{Z}_2$ symmetry is considered here, with λ_{χ_3} being real. If a *soft* breaking of the symmetry were considered instead, the term multiplying λ_{χ_3} and the corresponding terms in $\mathcal{V}(\Phi_1,\Phi_2)$ with λ_6 and λ_7 would disappear.

In spite of the fact that any scalar in addition to a 2HDM potential may acquire a VEV, here it is explicitly chosen that χ does not. Therefore, the Lagrangian describing the dynamics of χ, including the interactions with the 2HDM scalars after symmetry breaking is:

$$\begin{aligned}\mathcal{L}_\chi = &-\frac{1}{2}m_\chi^2\chi^2 - \frac{1}{2}v\lambda_{h\chi\chi}h\chi^2 - \frac{1}{2}v\lambda_{H\chi\chi}H\chi^2 - \lambda_{hh\chi\chi}hh\chi^2 - \lambda_{HH\chi\chi}HH\chi^2 \\ &- \lambda_{Hh\chi\chi}hH\chi^2 - \lambda_{AA\chi\chi}AA\chi^2 - \lambda_{H^+H^-\chi\chi}H^+H^-\chi^2,\end{aligned} \tag{5.10}$$

where the couplings have been re-written in terms of those in Equation (5.9) as follows:

$$\lambda_{h\chi\chi} = \lambda_{\chi_1}\cos\beta\sin\alpha - \lambda_{\chi_2}\sin\beta\cos\alpha - \frac{1}{2}\lambda_{\chi_3}\cos(\beta+\alpha), \tag{5.11}$$

$$\lambda_{H\chi\chi} = -\lambda_{\chi_1}\cos\beta\cos\alpha - \lambda_{\chi_2}\sin\beta\sin\alpha - \frac{1}{2}\lambda_{\chi_3}\sin(\beta+\alpha), \tag{5.12}$$

$$\lambda_{hh\chi\chi} = \frac{1}{4}(\lambda_{\chi_1}\sin^2\alpha + \lambda_{\chi_2}\cos^2\alpha - \lambda_{\chi_3}\sin\alpha\cos\alpha), \tag{5.13}$$

$$\lambda_{HH\chi\chi} = \frac{1}{4}(\lambda_{\chi_1}\cos^2\alpha + \lambda_{\chi_2}\sin^2\alpha + \lambda_{\chi_3}\sin\alpha\cos\alpha), \tag{5.14}$$

$$\lambda_{Hh\chi\chi} = \frac{1}{4}(-\lambda_{\chi_1}\cos\alpha\sin\alpha + \lambda_{\chi_2}\cos\alpha\sin\alpha + \lambda_{\chi_3}\cos^2\alpha - \lambda_{\chi_3}\sin^2\alpha), \tag{5.15}$$

$$\lambda_{AA\chi\chi} = \frac{1}{4}(\lambda_{\chi_1}\sin^2\beta + \lambda_{\chi_2}\cos^2\beta - \lambda_{\chi_3}\sin\beta\cos\beta), \tag{5.16}$$

$$\lambda_{H^+H^-\chi\chi} = \frac{1}{4}(\lambda_{\chi_1}\sin^2\beta + \lambda_{\chi_2}\cos^2\beta - \lambda_{\chi_3}\sin\beta\cos\beta). \tag{5.17}$$

Note that $\mathcal{L}_\chi$ does not include tree-level A-χ-χ interaction terms due to CP violation issues, although such an interaction could arise as a suppressed loop-induced process. However, the CP-odd scalar A would still play an important role in determining the DM relic density through the creation and/or annihilation process $\chi\chi \leftrightarrow AA$.

The introduction of the χ field is also subject to constraints. Perturbativity conditions imply $|\lambda_{\chi_i}| < 4\pi$ for $i = 1, 2, 3$. The coupling λ_{χ_4} is also required to be positive in order to ensure the stability of χ. Vacuum stability requires the following conditions (in addition to those given in Equation (5.3)), so that the potential $\mathcal{V}(\Phi_1, \Phi_2, \chi)$ is bounded from below:

$$\lambda_{\chi_4} > 0, \tag{5.18}$$

$$\lambda_{\chi_1} > -\sqrt{\frac{1}{12}\lambda_{\chi_4}\lambda_1}, \tag{5.19}$$

$$\lambda_{\chi_2} > -\sqrt{\frac{1}{12}\lambda_{\chi_4}\lambda_2}, \tag{5.20}$$

$$\lambda_{\chi_3} > -\sqrt{\frac{1}{12}\lambda_{\chi_4}\lambda_3}. \tag{5.21}$$

Additionally, if $\lambda_{\chi_1}, \lambda_{\chi_2}, \lambda_{\chi_3} < 0$, then the following conditions should also be satisfied:

$$-2\lambda_{\chi_1}\lambda_{\chi_2} + \frac{1}{6}\lambda_{\chi_4}\lambda_3 > -\sqrt{4\left(\frac{1}{12}\lambda_{\chi_4}\lambda_1 - \lambda_{\chi_1}^2\right)\left(\frac{1}{12}\lambda_{\chi_4}\lambda_2 - \lambda_{\chi_2}^2\right)}, \tag{5.22}$$

$$-2\lambda_{\chi_1}\lambda_{\chi_2} + \frac{1}{6}\lambda_{\chi_4}\left(\lambda_3 + \lambda_4 - |\lambda_5|\right) > -\sqrt{4\left(\frac{1}{12}\lambda_{\chi_4}\lambda_1 - \lambda_{\chi_1}^2\right)\left(\frac{1}{12}\lambda_{\chi_4}\lambda_2 - \lambda_{\chi_2}^2\right)}. \tag{5.23}$$

In order to ensure a stable DM candidate χ, an additional condition needs to be imposed, such that the VEV $\langle\chi\rangle$ should vanish at the global minimum of the scalar potential in Equation (5.9). This can be obtained numerically such that $\langle\chi\rangle = 0$ when $\langle\Phi_1\rangle \neq 0$ and $\langle\Phi_2\rangle \neq 0$. An example of this can be seen in Reference [135].

The simple assumption that χ is a scalar may or may not be appropriate. It is instructive to characterise χ in terms of other possible theories. One interesting consequence of doing so is the potential of altering the production mechanisms

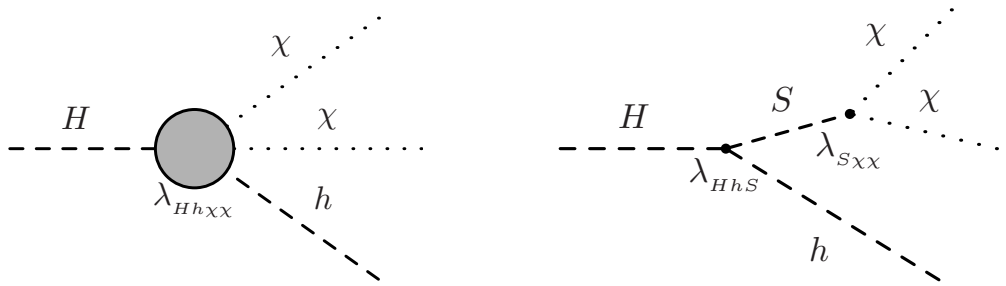


Figure 17: Representative Feynman diagrams to study the production of the Higgs boson in association with DM. On the left, the quartic $\lambda_{Hh\chi\chi}$ vertex used for the study in Section 4.1, and on the right due to the cascaded decay of the mediator S.

and decay modes for H and A through ggF and $\gamma\gamma$ fusion, since they are loop induced processes. It may be possible for χ to run in these loops, and therefore contribute in a correction to the production and decay rates. This would imply that χ is a massive coloured fermion. While this idea is not explicitly explored in this thesis, additional fermions in the context of the anomalous muon $(g-2)$ result are considered in Reference [136], with strong links to the BSM model considered in this thesis.

5.3 ADDING THE SCALAR MEDIATOR S

One of the problems with the simplified model introduced in Section 4.1 is the necessity for the $\lambda_{Hh\chi\chi}$ coupling to be significantly larger than the others. Theoretically, it is not clear why the 3-body $H \to h\chi\chi$ decay should be dominant over the possible 2-body decays. One simple way to explain this is to introduce a real scalar S, which is chosen to have similar couplings to the SM Higgs boson within the mass range $m_S \in [m_h, m_H - m_h]$. S then acts as a mediator, such that the $H \to h\chi\chi$ decay mode is a cascade of two different decay processes instead of being due to one effective vertex. This is shown in Figure 17.

Introducing S can also be used to probe more interesting physics, apart from just being a mechanism to enhance the p_{T} of the Higgs boson. For simplicity, a $\mathbb{Z}_2$ symmetry for $S \to -S$ transformations can be imposed. Since S is SM Higgs-like, it is allowed to couple with all of the SM particles as well as χ. This is phenomenologically interesting for two reasons. Firstly, S can be thought of as a portal between

which the SM particles can interact with DM. Secondly, the Higgs-like nature of S drastically reduces the number of free parameters in the theory, since all of the couplings to SM particles (and therefore also the BRs) are fixed to those of the SM Higgs boson, globally re-scaled down appropriately by the introduction of the invisible decay mode $S \to \chi\chi$. Since a large invisible BR is not experimentally observed for h, potential DM interactions can instead be studied with S.

In the absence of the interactions between S and the SM particles, one should not expect any interesting physics. But mixing with SM particles, and along with the other 2HDM scalars, has two consequences. Firstly, S could be observed as a resonance through its decays to pairs of vector bosons Z, $W^{\pm}$ and γ. For a Higgs-like S, a search for these decays would be similar to generic Higgs boson searches, albeit at higher masses. However, it should be noted that in this study the direct production of S is considered to be small due to the global re-scaling of its couplings to the SM particles. Note also that S is produced dominantly through the decay of H. Secondly, it alters the coupling strengths of known interactions in the theory. For example, in a 2HDM there exists a sum rule for the neutral scalar gauge couplings, $g^2_{hWW} + g^2_{HWW}$, which is the same as the SM coupling squared [137]. This sum rule will be violated if there is any mixing between S and the doublets Φ_1 and Φ_2, which will directly alter the expected projected bounds of 2HDM couplings.

In light of this, a real [4] scalar S is added to the theory, considering the possibility of a discrete symmetry under $S \to -S$. The parameters are arranged in such a way so that S acquires a VEV. Without the discrete symmetry, the most general potential for S can be written as:

$$\begin{aligned}\mathcal{V}(\Phi_1, \Phi_2, S) = \mathcal{V}(\Phi_1, \Phi_2) &+ \frac{1}{2} m^2_{S_0} S^2 + \frac{\lambda_{S_1}}{2} \Phi_1^\dagger \Phi_1 S^2 + \frac{\lambda_{S_2}}{2} \Phi_2^\dagger \Phi_2 S^2 \\ &+ \frac{\lambda_{S_3}}{4} \Phi_1^\dagger \Phi_2 S^2 + \frac{\lambda_{S_4}}{4!} S^4 + \mu_1 \Phi_1^\dagger \Phi_1 S + \mu_2 \Phi_2^\dagger \Phi_2 S \\ &+ \mu_3 \Phi_1^\dagger \Phi_2 S + \mu_S S^3 + \text{h.c.} \end{aligned} \tag{5.24}$$

4 One can also introduce a complex scalar in theory, the consequence of which alters the choice of symmetry. The $\mathbb{Z}_2$ symmetry would then be promoted to a global $U(1)$ and its spontaneous breaking would lead to a massless pseudo-scalar [138].

Now, if a $\mathbb{Z}_2$ symmetry is imposed for transformations of the form $S \to -S$ (and all other fields are even), then the terms with the coefficient μ_i ($i = 1, 2, 3, S$) will vanish in the above general potential. If a further $\mathbb{Z}_2^{'}$ symmetry is imposed for the transformations $h \to h$, $H \to -H$ and $S \to S$, then the λ_{S_3} term will also vanish. This also eliminates λ_6 and λ_7 from $\mathcal{V}(\Phi_1, \Phi_2)$. However, a soft breaking of $\mathbb{Z}_2^{'}$ is assumed, which implies $m_{12}^2 \neq 0$. In the case where S does not acquire a VEV, the S related interactions in the potential are given by:

$$\begin{aligned} V_S =& \frac{1}{2} m_S^2 S^2 + \lambda_{hSS} v\, hS^2 + \lambda_{HSS} v\, HS^2 - \lambda_{HHSS}\, H^2S^2 - \lambda_{hHSS}\, hHS^2 \\ & - \lambda_{hhSS}\, h^2S^2 - \lambda_{AASS}\, A^2S^2 - \lambda_{H^+H^-SS}\, H^+H^-S^2. \end{aligned} \tag{5.25}$$

The couplings in the potential in terms of λ_{S_1}, λ_{S_2}, α and β are as follows:

$$m_S^2 = m_{S_0}^2 + \left(\frac{\lambda_{S_1}}{2} \cos^2\beta + \frac{\lambda_{S_2}}{2} \sin^2\beta \right) v^2 \tag{5.26}$$

$$\lambda_{hSS} = -\frac{\lambda_{S_1}}{2} \sin\alpha \cos\beta + \frac{\lambda_{S_2}}{2} \cos\alpha \sin\beta \tag{5.27}$$

$$\lambda_{HSS} = \frac{\lambda_{S_1}}{2} \cos\alpha \cos\beta + \frac{\lambda_{S_2}}{2} \sin\alpha \sin\beta \tag{5.28}$$

$$\lambda_{hhSS} = \frac{\lambda_{S_1}}{4} \sin^2\alpha + \frac{\lambda_{S_2}}{4} \cos^2\alpha \tag{5.29}$$

$$\lambda_{HHSS} = \frac{\lambda_{S_1}}{4} \cos^2\alpha + \frac{\lambda_{S_2}}{4} \sin^2\alpha \tag{5.30}$$

$$\lambda_{hHSS} = \frac{1}{4} (\lambda_{S_2} - \lambda_{S_1}) \sin 2\alpha \tag{5.31}$$

$$\lambda_{AASS} = \frac{1}{2} \lambda_{H^+H^-SS} = \frac{\lambda_{S_1}}{4} \sin^2\beta + \frac{\lambda_{S_2}}{4} \cos^2\beta. \tag{5.32}$$

However, in order to recreate an effective H-h-χ-χ type interaction in a full model with S (as in Figure 17), there is a need for an h-H-S coupling. In the absence of the term multiplying mu_3, it is possible to generate this coupling if S acquires a VEV. In this study, it is necessary that S will indeed acquire a VEV and mix with h and H. The h-S mixing is also chosen to be small for simplicity. In the limit of zero mixing between h and S (as well as H and S), the expressions for various couplings have been shown above. Equation (5.28) indicates that the H-S-S coupling need not be small even in this limit, since α and β, which are the mixing angles from the doublet sector exclusively, are free parameters. If a mixing between S and the doublets is

turned on, Equation (5.28) will receive corrections through the additional mixing angle(s) introduced. However, in the case of small *h*-*S* mixing, the correction will also be small, and the *H*-*S*-*S* coupling will still remain sizeable. Therefore, it is assumed that the mixing of S with h is small enough (by interplay of various parameters in the potential) that it will not spoil any experimental bounds. The *h*-*H*-*S*-*S* interaction can be thought of as a source of the required *h*-*H*-*S* coupling if one replaces one S by its VEV in the *h*-*H*-*S*-*S* interaction.

5.4 PREDICTIVE PHENOMENOLOGY

This inclusion of S can open up various new possibilities in terms of search channels and phenomenology. With the mass range $m_h \lesssim m_S \lesssim m_H - m_h$ and $m_S > 2m_\chi$, one can find interesting physics in both $pp \to H \to Sh$ as well as $pp \to H \to hh$. There is an additional possibility to study a $H \to SS$ decay mode, which is more likely to be off-shell given the masses under consideration. Many interesting final states can be explored through these processes, keeping in mind that all decay modes of S (i.e. S decaying into jets, vector bosons, leptons, DM etc.) are possible.

Following an EFT approach, one can write down a Lagrangian for all the interactions necessary for phenomenological purposes in studying S. After EWSB, the Lagrangian for the singlet real scalar can be written as:

$$\mathcal{L}_S = \mathcal{L}_{\mathrm{K}} + \mathcal{L}_{SVV'} + \mathcal{L}_{Sf\bar{f}} + \mathcal{L}_{hHS} + \mathcal{L}_{S\chi}, \tag{5.33}$$

where each of the sectors is defined as follows:

$$\mathcal{L}_{\mathrm{K}} = \frac{1}{2}\partial_\mu S \partial^\mu S - \frac{1}{2} m_S^2 SS, \tag{5.34}$$

$$\begin{aligned}\mathcal{L}_{SVV'} = &\frac{1}{4}\kappa_{s_{gg}} \frac{\alpha_s}{12\pi v} S G^{a\mu\nu} G^a_{\mu\nu} + \frac{1}{4}\kappa_{s_{\gamma\gamma}} \frac{\alpha}{\pi v} S F^{\mu\nu} F_{\mu\nu} + \frac{1}{4}\kappa_{sZZ} \frac{\alpha}{\pi v} S Z^{\mu\nu} Z_{\mu\nu} \\ &+ \frac{1}{4}\kappa_{sZ_\gamma} \frac{\alpha}{\pi v} S Z^{\mu\nu} F_{\mu\nu} + \frac{1}{4}\kappa_{sWW} \frac{2\alpha}{\pi \sin^2\theta_w v} S W^{+\mu\nu} W^-_{\mu\nu},\end{aligned} \tag{5.35}$$

$$\mathcal{L}_{Sf\bar{f}} = -\sum_f \kappa_{sf} \frac{m_f}{v} S \bar{f} f, \tag{5.36}$$

$$\mathcal{L}_{HhS} = -\frac{1}{2}\, v \Big[\lambda_{hhS} hhS + \lambda_{hSS} hSS + \lambda_{HHS} HHS + \lambda_{HSS} HSS + \lambda_{HhS} HhS\Big], \tag{5.37}$$

$$\mathcal{L}_{S\chi} = -\frac{1}{2}\, v\, \lambda_{S\chi\chi} S\chi\chi - \frac{1}{2}\lambda_{SS\chi\chi} SS\chi\chi. \tag{5.38}$$

Here V refers to gluons, photons, and the weak vector bosons. The gauge field strength $W^{\pm}_{\mu\nu}$ is defined as $D_\mu W^{\pm}_\nu - D_\nu W^{\pm}_\mu$, where D_μ is the co-variant derivative $D_\mu W^{\pm}_\nu = \left[\partial_\mu \pm ieA_\mu\right] W^{\pm}_\nu$. Other possible self interaction terms for S are neglected here since they are not of any phenomenological interest in this study.

Using the schema described above, performing phenomenological studies becomes relatively simple, since the BRs of S to SM particles can be fixed to known values, and the BR to $\chi\chi$ is a free parameter that can be set. As an example, a simple study is that of revisiting the Higgs p_{T} spectrum after the introduction of the S mediator. Part of the rationale for designing a new production mechanism of the Higgs boson in association with DM (as described in Section 4.1) was to describe an apparent distortion in the intermediate range of the Higgs p_{T} spectrum. It is therefore instructive to understand how changing the structure of the production mechanism affects the Higgs p_{T}– that is, how the two different Feynman diagrams in Figure 17 predict the Higgs p_{T} spectrum. For a simple comparison, distributions were made using 50 000 events both generated and showered in `Pythia 8.2` [82]. These distributions can be seen in Figure 18. On the left, the SM Higgs p_{T} distribution from ggF can be seen compared to spectra predicted using the effective $H \to h\chi\chi$ decay process introduced in Section 4.1. Three mass points of H have been chosen, whereas the mass of χ has been fixed to 60 GeV $\sim m_h/2$. As in the study performed in Chapter 4, this suppresses the $h \to \chi\chi$ BR in accordance with experimental limits [77, 78]. On the right, the S mediator is introduced with $m_H = 300$ GeV and the mass of S is varied (the mass of χ has no effect on the p_{T} spectrum in this case). This is compared to a single mass point using the effective vertex approach. It was found that the kinematics for the effective vertex approach is similar to the full theory with a large width S at $m_S = m_H/2$ (in the

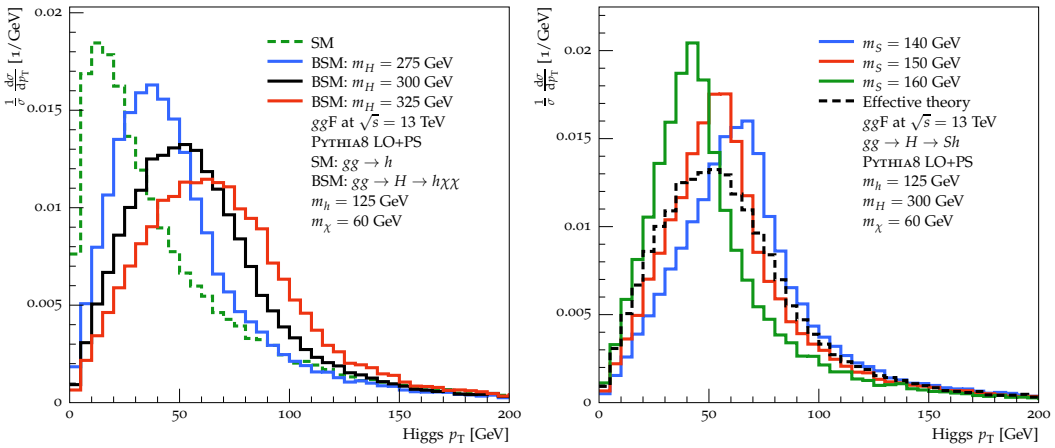


Figure 18: The Higgs p_T spectrum as predicted by various production mechanisms. On the left, through the effective decay process $H \to h\chi\chi$ with different mass hypotheses for H (solid lines), as well as through the SM ggF production mode (dashed line). On the right, one spectrum using the effective vertex (dashed line) is compared with three spectra as described by the S-mediated interaction (solid lines). In this case, m_H is fixed to 300 GeV and m_S is varied.

limit $m_\chi \to m_h/2$). Overall, using the S to mediate the decay process makes for a more resolved distribution compared with the effective vertex approach.

The most interesting decay mode to study, given that the S has a mass in the range $[m_h, m_H - m_h]$, is the $H \to Sh$ decay mode. The two couplings related to the H-S-S and H-h-S are of most importance here. As stated above, the origin for the consideration of the intermediate real scalar S demands that these two terms can explain a non-negligible BR for the $H \to h\chi\chi$ decay mode. In one sense, there is an equivalence of the couplings $\lambda_{Hh\chi\chi}$ (introduced in Section 4.1) compared with the cascade of λ_{HhS} and $\lambda_{S\chi\chi}$, so that the 3 body decay can be equated to a series of 2 body decays as shown in Figure 17. On the other hand, in order to minimise the number of free parameters in the theory, one can consider a ratio of couplings:

$$r = |\lambda_{HSS}|/|\lambda_{HhS}|. \tag{5.39}$$

This positive-definite ratio (i.e. so that there won't be any negative interference due to the choice of negative values) could be fixed in the range of theoretically allowed values, and then either one of the couplings λ_{HSS} or λ_{HhS} can be varied to control the relative rates of Sh and SS production through the decay of H.

Scalar	Decay modes
h	$b\bar{b}$, $\tau^+\tau^-$, $\mu^+\mu^-$, $s\bar{s}$, $c\bar{c}$, gg, $\gamma\gamma$, $Z\gamma$, W^+W^-, ZZ
H	hh, SS, Sh and all h decay modes
A	$t\bar{t}$, Zh, ZH, ZS, $W^\pm H^\mp$ and all h decay modes
$H^\pm$	$W^\pm h$, $W^\pm H$, $W^\pm S$
S	$\chi\chi$ and all h decay modes

Table 7: A list of all possible decay modes for the 2HDM scalars and S, based explicitly on the mass ranges considered in this work.

Below is a discussion on the various phenomenological aspects associated with this particle spectrum, applicable to collider signatures (in particular at the LHC). Given the mass ranges considered for each new scalar, their appropriate decay modes have been listed in Table 7. These are shown as a reference for purposes of discussing relevant experimental signatures. After the discussion, a list of specific experimental search channels for each of the new 2HDM scalars and S is shown in Table 8.

5.4.1 *Heavy scalar H*

In Section 4.1, the heavy scalar H was introduced in an effective theory, one of the goals of this was to explain an apparent distortion in the p_T spectrum of the Higgs boson. Moving on from the effective theory used there, one can now think of H as the heavier CP-even component of a 2HDM. Arguably, the motive now should be to fit parameters such as $\tan\beta$, α and the masses of A and $H^\pm$ in this specific model. However, the question arises as to whether one should think of a generalised 2HDM or a much more specifically defined type of 2HDM, as described in Reference [128]. On the other hand, experimental data from searches needs to be considered, which will constrain the available parameter space of this model. This was initially performed in Reference [4] and is also shown in Chapter 6 of this thesis. For the purposes of this study, it is explicitly chosen that the lighter CP-even component of the 2HDM is the experimentally observed Higgs-like boson scalar (i.e. that $m_h = 125$ GeV). With this fixed, the chosen mass range for H is

chosen to be in the range $2m_h < m_H < 2m_t$ for reasons which were alluded to in Section 4.2.1.

In the simplest case, the cross section for $gg \to H$ production (i.e. ggF, the dominant production mode) would be the same as a heavy Higgs boson – between 5 and 10 pb at $\sqrt{s} = 13$ TeV [33]. However, this number could be altered if one considers a re-scaling of the Yukawa coupling or the possibility of extra coloured particles running in the loop (as discussed above). In Section 4.1, the number β_g – which was assumed as a re-scaling of the Yukawa coupling – was estimated to be around 1.5. This implies that the $gg \to H$ production cross section could be enhanced by at least a factor of 2.

5.4.2 *CP-odd scalar A*

Whereas the statistical analysis performed in Section 4.1 was able to shed some light on the possible mass range for H, at this point the mass of A remains somewhat a free parameter in the model. Having said this, a separate and strongly linked study has been able to shed some light on the potential production of A in this model, where the mass of A finds compatibility with a mass of around 600 GeV [139]. There are a few generic points that should be noted when studying phenomenology related to A. In 2HDMs, the masses of the A and $H^\pm$ bosons are correlated. If one wishes to construct a 2HDM and fix it with a particular mass for A, then its compatibility with $m_{H^\pm}$ should also be taken seriously. Furthermore, known values for m_H and m_h (i.e. 125 GeV) should also necessarily be taken into account in tuning the parameters α and β.

In the case of a ggF production mode for A (through the effective g-g-A vertex), there will be a need for a scaling factor β_g^A similar to the treatment of H production, which scales with β_g. Consideration of the decay modes of A ($A \to \gamma\gamma$ in particular) also prompts the need for another scaling factor β_γ^A. This implies that one needs to control the $H \to \gamma\gamma$ decay rates via another parameter β_γ, since the form factors appearing in the calculation of $gg \to H, A$ and $H, A \to \gamma\gamma$ have a

different structure. They are also dependent on the masses of the particles under consideration (this is described in Reference [27]).

One should also consider the other possible decay modes of A which include pairs of W and Z bosons. These decays are possible only at loop level in 2HDMs, since the tree level A-V-V couplings are absent as a result of CP conservation issues. Depending on parameter choices, this model can predict an arbitrarily large cross section for $Z+$ jets $+E_{\mathrm{T}}^{\mathrm{miss}}$ production. It is important to think of the contribution of the decay mode of $A \to ZH$, where $H \to Sh$. This mode requires that $m_A > m_Z + m_H$. This decay mode also implies processes with multi-lepton final states, significantly in both same-sign (SS) and opposite-sign (OS) configurations, these processes are also produced in association with jets.

5.4.3 *Charged scalars $H^{\pm}$*

The 2HDM predicts the existence of charged scalars, $H^{\pm}$, which in principle can be produced at the LHC. Searches for these charged scalars are made based on assumptions of their mass. When $m_{H^{\pm}} > m_t$, the prominent decay modes of $H^{\pm}$ are $H^{\pm} \to tb$ and $H^{\pm} \to W^{\pm}h$. Noting that in this study $2m_h < m_H < 2m_t$, the decay mode $H^{\pm} \to W^{\pm}H$ could also be considered a prominent channel, assuming that $m_{H^{\pm}} \gg m_H$. The different predictions as a result of assumptions on the mass of the particles makes a discussion on the phenomenology slightly complicated. Therefore, it is assumed explicitly in this study that $m_{H^{\pm}} > m_t$, for simplicity.

In this case, there would be two main production mechanisms of $H^{\pm}$ at the LHC, both of which are top quark associated. These are as follows,

- $2 \to 2$: $gb \to tH^{\pm}$, and
- $2 \to 3$: $gg/qq \to tH^{\pm}b$.

Pair production of the charged scalars could also occur through a Drell-Yan like process (that is, $qq \to H^{+}H^{-}$). The associated production with W bosons, and pair production through ggF could also be studied, albeit with smaller cross sections.

Scalar	Production mode	Search channels
H	$gg \to H, Hjj$ (ggF and VBF)	Direct SM decays as in Table 7 $\to SS/Sh \to 4W \to 4\ell + E_{\mathrm{T}}^{\mathrm{miss}}$ $\to hh \to \gamma\gamma b\bar{b},\ b\bar{b}\tau\tau,\ 4b,\ \gamma\gamma WW$ etc. $\to Sh$ where $S \to \chi\chi \implies \gamma\gamma,\ b\bar{b},\ 4\ell + E_{\mathrm{T}}^{\mathrm{miss}}$
	$pp \to Z(W^{\pm})H\ (H \to SS/Sh)$	$\to 6(5)l + E_{\mathrm{T}}^{\mathrm{miss}}$ $\to 4(3)l + 2j + E_{\mathrm{T}}^{\mathrm{miss}}$ $\to 2(1)l + 4j + E_{\mathrm{T}}^{\mathrm{miss}}$
	$pp \to t\bar{t}H, (t+\bar{t})H\ (H \to SS/Sh)$	$\to 2W + 2Z + E_{\mathrm{T}}^{\mathrm{miss}}$ and b-jets $\to 6W \to 3$ same sign leptons + jets and $E_{\mathrm{T}}^{\mathrm{miss}}$
$H^{\pm}$	$pp \to tH^{\pm}\ (H^{\pm} \to W^{\pm}H)$	$\to 6W \to 3$ same sign leptons + jets and $E_{\mathrm{T}}^{\mathrm{miss}}$
	$pp \to tbH^{\pm}\ (H^{\pm} \to W^{\pm}H)$	Same as above with extra b-jet
	$pp \to H^{\pm}H^{\mp}\ (H^{\pm} \to HW^{\pm})$	$\to 6W \to 3$ same sign leptons + jets and $E_{\mathrm{T}}^{\mathrm{miss}}$
	$pp \to H^{\pm}W^{\pm}\ (H^{\pm} \to HW^{\pm})$	$\to 6W \to 3$ same sign leptons + jets and $E_{\mathrm{T}}^{\mathrm{miss}}$
A	$gg \to A$ (ggF)	$\to t\bar{t}$ $\to \gamma\gamma$
	$gg \to A \to ZH\ (H \to SS/Sh)$	Same as $pp \to ZH$ above, but with resonance structure over final state objects
	$gg \to A \to W^{\pm}H^{\mp}(H^{\mp} \to W^{\mp}H)$	$6W$ signature with resonance structure over final state objects
S	$gg \to S$ (ggF) *or* $H \to SS/Sh$ (associated production)	Resonantly through decays as in Table 7 ($\gamma\gamma$, $b\bar{b}$, $\tau\tau$, $ZZ \to 4\ell$) Non-resonantly through multilepton + $E_{\mathrm{T}}^{\mathrm{miss}}$ decays

Table 8: A list of potential search channels arising from the addition of the new scalars. This list is by no means complete, but contains clean search channels which could make for striking signatures at the LHC.

5.4.4 *The additional scalars S and χ*

The inclusion of S and χ in the model is especially significant in terms its phenomenology, since the signatures arising from the 2HDM scalars have mostly been addressed in other works already. With this in mind, the combination of the 2HDM with χ and S can lead to many interesting final states – lists of these can be seen in Table 7 and Table 8.

The dominant production mechanism of S is assumed to be through the decay processes $H \to SS$ and $H \to Sh$. The admixture of these decays is controlled by a ratio of BRs, defined as:

$$a_1 \equiv \frac{\mathrm{BR}(H \to SS)}{\mathrm{BR}(H \to Sh)}. \tag{5.40}$$

The parameter a_1 is related to the ratio r introduced in Equation (5.39), although uses the BRs instead of the couplings for simplicity. S is assumed to be similar to the SM Higgs boson, in the sense that it couples to SM particles with the same structure as h. These couplings are then dependent on m_S, and a choice of m_S therefore has implications on the final states that can be studied. Within the mass range considered (i.e. between m_h and $m_H - m_h$), S can be in one of two regions. The first is dominated by $S \to VV$, when $m_S \gtrsim 2m_W \sim 160$ GeV. The second is when $m_S \lesssim 2m_W$, and in this region S has non-negligible BRs to various decay products such as $b\bar{b}$, VV, gg, $\gamma\gamma$, $Z\gamma$ etc.

In this model, S is also assumed to be a portal to DM interactions through the decay mode $S \to \chi\chi$. With all other couplings to SM particles fixed, the BR to $\chi\chi$ is a free parameter in the theory. When adding this decay mode, all of the SM decay modes are scaled down by $1 - \mathrm{BR}(S \to \chi\chi)$, and the total width of S increases accordingly (although in practical studies, a narrow width approximation will suffice).

The SM Higgs boson has stringent experimental limits on its invisible BR. In this model, this is interpreted by the fact that the $h \to \chi\chi$ BR is suppressed by the choice of $m_\chi \sim m_h/2$. Therefore, S is an important component of the model since

it is useful to study events which can have an arbitrarily large amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ depending on m_H, m_S and $\mathrm{BR}(S \to \chi\chi)$.

5.5 ANALYSIS OF SELECTED LEPTONIC SIGNATURES

The model introduced above can produce a wide variety of interesting signatures for study, given the combination of production and decay modes discussed above (and shown in Table 8). As a case study, here the focus is on a few striking signatures driven by the production of multiple leptons. These signatures are also dependent on the production of a non-negligible amount of $E_{\mathrm{T}}^{\mathrm{miss}}$. Firstly, the production of 4 leptons in association with $E_{\mathrm{T}}^{\mathrm{miss}}$ via the production of 4 W bosons is analysed in Section 5.5.1. Thereafter, the unique signature of 6 W boson production in ttH is analysed with a final state of 3 SS leptons in Section 5.5.2. Finally, the $A \to ZH$ decay is analysed in Section 5.5.3 with some simple comparisons to limits in the LHC data. The analyses have been chosen such that the first two (Section 5.5.1 and Section 5.5.2) do not rely on the S's interaction with DM, whereas the third (Section 5.5.3) does. This is an example of how the "simplified model" approach is useful in that different searches can be used to constrain different parameters of the theory. The analyses treat the new scalars as being in the following mass ranges,

- Light Higgs: $m_h = 125$ GeV (this is the SM Higgs boson),
- Heavy Higgs: $2m_h < m_H < 2m_t$,
- CP-odd Higgs: $m_A > (m_H + m_Z)$,
- Charged Higgs: $(m_H + m_W) < m_{H^\pm} < m_A$,
- Additional scalars: $m_\chi < m_h/2$ and $m_h \lesssim m_S \lesssim (m_H - m_h)$.

In the following sections, some plots of key distributions are shown and discussed for the three different signatures. These plots were made from selecting `Pythia 8` [82] MC events using custom-made `Rivet` [83] routines. In all three cases,

500 000 events were generated and a selection efficiency was determined based on experimentally motivated cuts and criteria. These events are not passed through a detector simulation. The reason for this is that our intentions are not to model the profile of $E_{\mathrm{T}}^{\mathrm{miss}}$ with accuracy, but rather provide a signature of the general region in which $E_{\mathrm{T}}^{\mathrm{miss}}$ could be expected, given the parameter constraints.[5] For the first two analyses, leptons were defined as either electrons or muons with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.47$ (2.7) for electrons (muons). A crude lepton isolation is applied by vetoing any leptons which share a partner lepton within a cone of radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.2$ around it, and any leptons coming from a hadron decay are vetoed. At truth level, this is a good approximation of the traditional track or calorimeter based isolation procedures.

The mass points considered in these distributions are relatively close to the central points in the ranges considered here. The mass of S is fixed to 150 GeV, where it still enjoys a wide range of decay modes due to it's SM-like nature – at this mass the BRs to $b\bar{b}$ and VV are both non-negligible, allowing for sensitivity in di-jet and di-boson searches, while a lighter S runs the risk of being too close to the Higgs mass for a comfortable experimental resolution. The mass of H is considered at the two values 275 GeV and 300 GeV. A mass close to 275 GeV does have some motivation from the simple study done in Section 4.1, but is also interesting since the $H \to SS$ decay is then off-shell. The on-shell behaviour is probed by also selecting the point $m_H = 300$ GeV, and a_1 (Equation (5.40)) is chosen such that $\mathrm{BR}(H \to SS) = \mathrm{BR}(H \to Sh) = 0.5$ in order for both decay mechanisms to be explored evenly. The $S \to \chi\chi$ BR is chosen to be 0.5 to probe both the production of SM particles and DM particles through the decay of S.

5.5.1 $H \to 4W \to 4\ell + E_T^{miss}$

Assuming a large enough cross section for the single production of H, the decays $H \to SS, Sh$ can lead to a sizeable production of 4 Ws. The leptonic decays would

5 Having said this, the state of the art fast simulation package `Delphes 3`'s [140] predictions of detector effects in $E_{\mathrm{T}}^{\mathrm{miss}}$ are reasonable, but still not completely compatible with the full simulation packages used by ATLAS and CMS. Detector simulation could be studied in a future work.

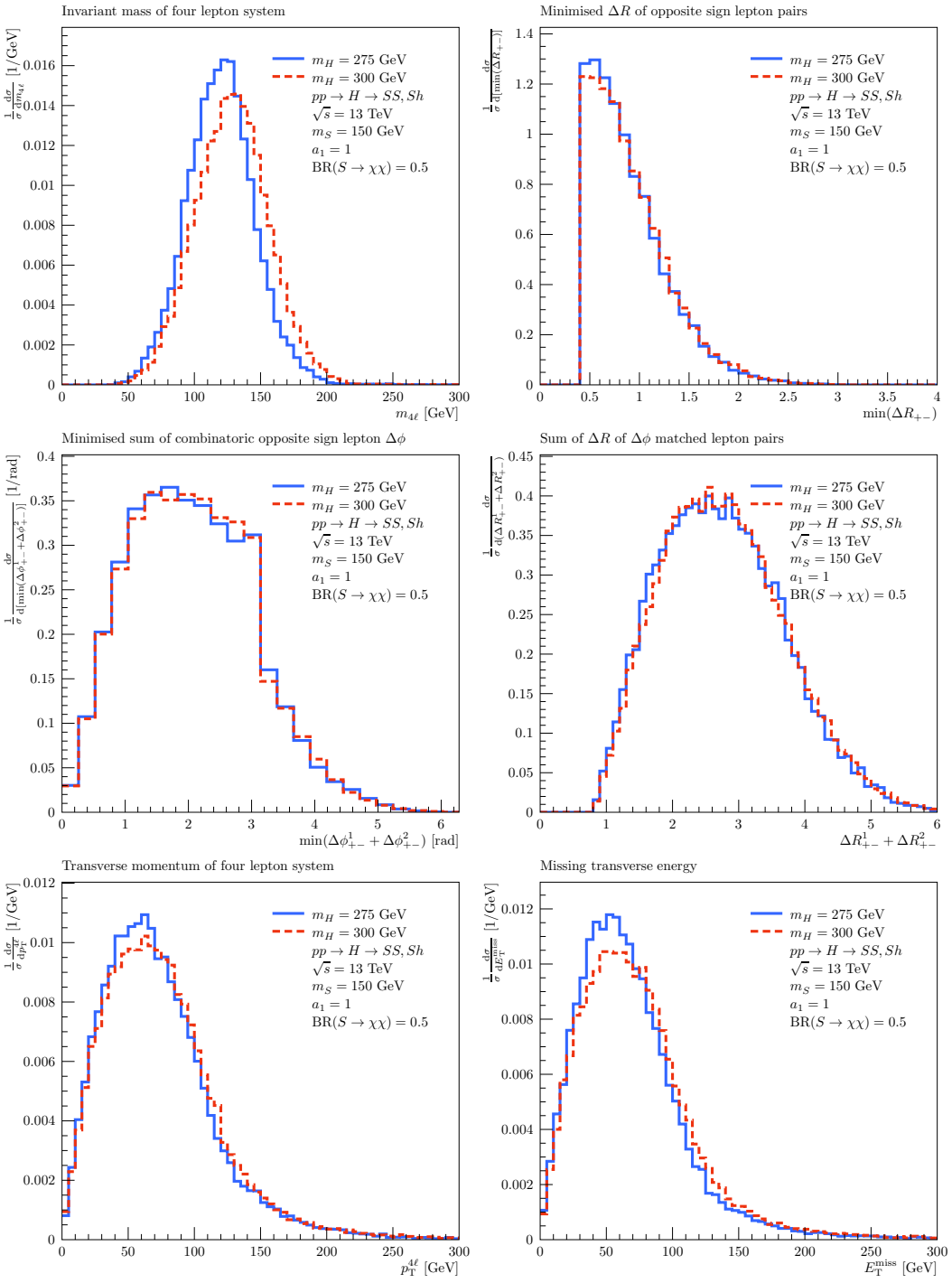


Figure 19: Various leptonic kinematic distributions (normalised to unity) pertaining to the process $H \to 4W \to 4\ell + E_{\mathrm{T}}^{\mathrm{miss}}$, as described in Section 5.5.1.

produce 4 charged leptons (electrons and muons) in conjunction with large $E_{\mathrm{T}}^{\mathrm{miss}}$. Due to the spin-0 nature of the S and h bosons, the leptons of the decay of each boson appear close together [141], leading to an even more striking signature.

Figure 19 displays the kinematics of the leptons for $m_H = 275, 300$ GeV and $m_S = 150$ GeV for a pp centre of mass energy of 13 TeV. Results are shown assuming $\mathrm{BR}(S \to \chi\chi) = 0.5$. In the event generation, both S and h are forced to decay to WW, and these Ws are forced to decay semi-leptonically (including $\tau\nu_\tau$ decays, since these can result in final states containing muons or electrons). Given the $gg \to H$ cross section range mentioned in Section 5.4.1, one could expect a cross section times BR of as much as about 50 fb for this process at the mass points considered here.

The upper left plot shows the four-lepton invariant mass ($m_{4\ell}$). In the mass range of interest here, the background is suppressed and it is dominated by the non-resonant production of Z boson pairs in which at least one is off-shell [142, 143]. The production of the SM Higgs boson would need to be taken into account as a background, although this could be suppressed by requiring a $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. The contribution from processes where at least one lepton arises from hadronic decays is sub-dominant to the production of $pp \to ZZ^* \to 4\ell$.

The upper right plot displays a distribution of the smallest ΔR between opposite sign leptons. This variable exploits the spin-0 nature of the S and h bosons.[6] The distribution features a cut-off due to the requirement that leptons be apart from each other by $\Delta R > 0.4$ due to isolation requirements. The plot in the middle-left displays the sum of the di-lepton azimuthal angle separation for the two opposite sign pairs ($\Delta\phi_{+-}$). Here the choice of lepton pairs is performed so as to minimise the sum of the di-lepton azimuthal angle separation. The corresponding sum of ΔR distances for this choice of lepton pairing is shown in the middle-right plot. The lower plots display the transverse momentum of the 4 lepton system and the $E_{\mathrm{T}}^{\mathrm{miss}}$. These distributions are significantly different from what one would expect from the residual backgrounds from $pp \to ZZ^* \to 4\ell$.

6 The kinematics of the decay depend on the tensor structure of the S-V-V coupling. This can be seen in Reference [4].

The production of $t\bar{t}Z$ is a source of four charged leptons [144]. This background can be suppressed by a combination of requirements, including placing a veto on the presence of jets and b-jets in the final state. The production $4W$s in the standard model is dominated by $t\bar{t}t\bar{t}$ [145, 81] and $t\bar{t}WW$ [81], which are significantly smaller and can be neglected. The production of $t\bar{t}t\bar{t}$ with other final states has been investigated and no significant excess in the data has been observed with respect to the SM prediction [146].

5.5.2 $t(t)H \to 6W \to \ell^{\pm}\ell^{\pm}\ell^{\pm} + X$

The production of double and single top quarks in association with the heavy scalar can produce up to 6 Ws in association with b-quarks. This leads to the possibility of producing three same-sign isolated charged leptons ($\ell^{\pm}\ell^{\pm}\ell^{\pm}$), a unique and striking signature at hadron colliders. The production of same-sign tri-leptons, including non-isolated leptons from heavy quark decays, was suggested in Reference [147] to tag top events. The production of isolated same-sign tri-leptons has been studied in the context of the search for new leptons [148] and in R-parity violating SUSY scenarios [149, 150]. Background studies performed in References [148, 150] indicate that the production of three same-sign isolated leptons is almost negligible at the LHC, with a cross section of less than 1×10^{-2} fb for a pp centre of mass of 13 TeV. The background would be dominated by the production of $t\bar{t}W$ with additional leptons from heavy flavour decays. This background is reducible by means of isolation, impact parameters and other requirements [142, 143]. With a reasonable choice of parameters a fiducial cross section of 0.5 fb can be predicted for 13 TeV centre of mass energy, rendering the search effectively background free.

It is relevant to study the kinematics of the final state here, as detailed in Figure 20. The event generation allowed for the decay of S and h into any channels involving a W, Z or τ. To ensure a clean signal, leptons were only selected if they did not come from a hadron decay – these processes contain many B-hadrons which can decay into leptons. Under these conditions, the efficiency to select at

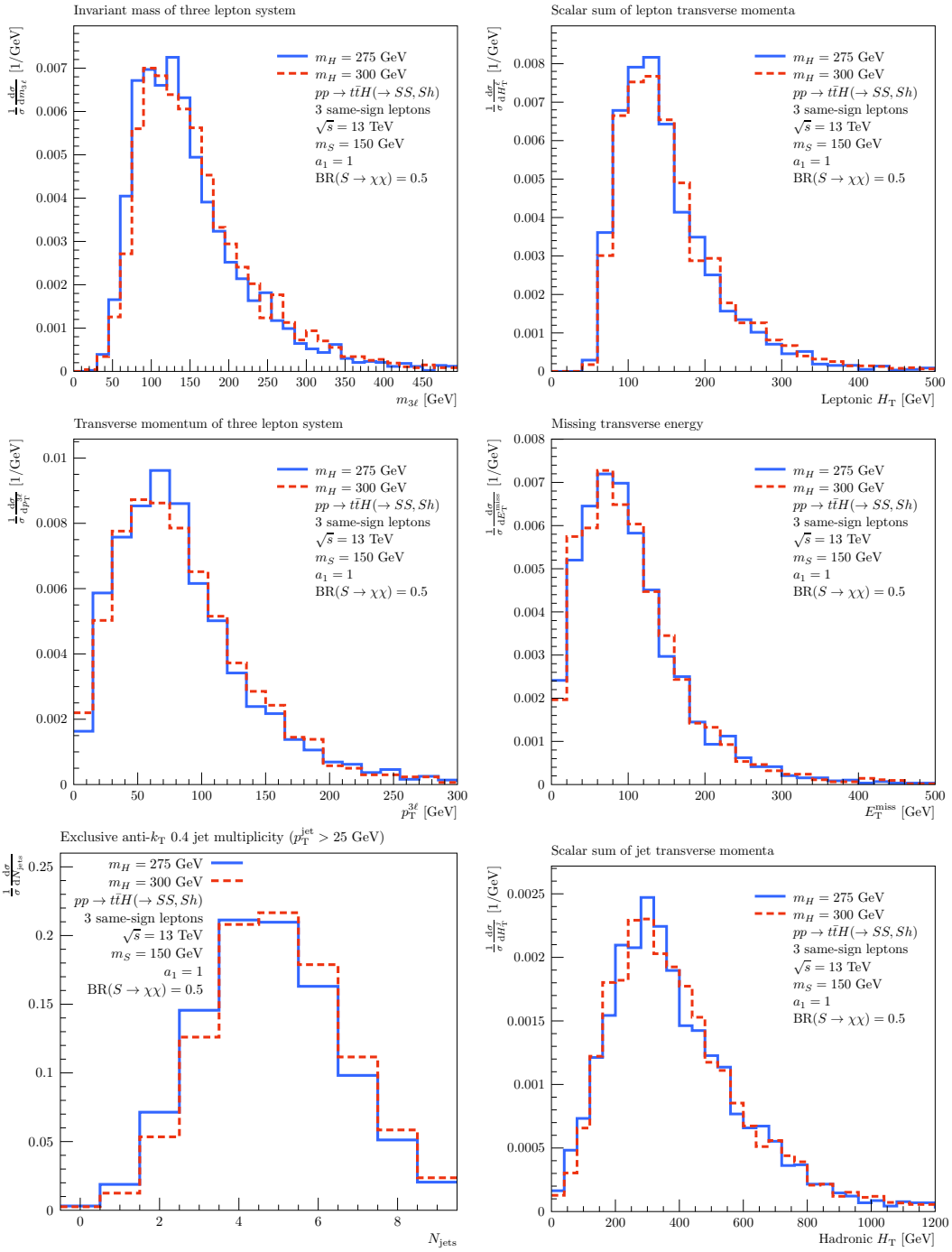


Figure 20: Various hadronic and leptonic kinematic distributions (normalised to unity) pertaining to the process $t\bar{t}H \to 6W \to \ell^{\pm}\ell^{\pm}\ell^{\pm} + X$, as described in Section 5.5.2.

least 3 leptons in an event was about 8%. Of these events, about 15% would contain a group of three same-sign leptons. The upper left and right plots display tri-lepton invariant mass and the scalar sum of the transverse momenta (H_{T}) of the leptons, respectively. The transverse momentum of the three leptons is shown in the middle-left plot. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is shown in the middle-right plot. The average $E_{\mathrm{T}}^{\mathrm{miss}}$ in these events is significant, and it adds to the uniqueness of the signature.

Since the production of three same-sign isolated leptons requires the presence of at least six weak bosons and/or τ leptons, a large number of jets is expected from those particles that do not decay leptonically. This makes the production of three same-sign isolated leptons even more striking. In the analysis done here, hadronic jets are defined using the anti-k_{T} clustering algorithm [151] with the parameter $R = 0.4$. Jets are required to have transverse momentum $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and to be central, that is, in the range $|\eta| < 2.5$. The jet multiplicity of jets is shown in the lower-left plot. The distribution peaks around 4-5 with a long tail stretching to 8 or more jets. The differences displayed by changing m_H are due to the fact that in the case of $m_H = 275$ GeV one of the S bosons in $H \to SS$ becomes off-shell, reducing the transverse momentum of the jets. The H_{T} constructed with jets is shown in the lower-right plot.

The production of H with single top is not suppressed with respect to the $t\bar{t}$ production, as it is in the production of the SM Higgs boson. The kinematic distributions shown in Figure 20 are similar to those displayed by the tH production with the exception of the jet multiplicity and the jet H_{T}, due to the reduced production of b-quarks. A similar discussion applies to the production of $H^{\pm} \to W^{\pm}H$. It is worth noting that the distributions shown in Figure 20 also apply to the combination of three leptons where the total charge is ± 1. There the SM backgrounds are significant, although the signal rate is about 6 times larger.

5.5.3 $A \to ZH \to Z + jets + E_T^{miss}$

In the literature regarding 2HDMs (the seminal example being Reference [27]), one can find the Lagrangians describing exactly how the SM gauge bosons interact with the 2HDM scalars. The sector of the Lagrangian describing the couplings between A, Z and the CP-even Higgs bosons are as follows:

$$\begin{aligned}\mathcal{L}_{V\phi\phi} \supset & \frac{M_W}{v\,\cos\theta_W}\sin(\beta-\alpha)Z_\mu\left(A\partial_\mu H - H\partial_\mu A\right) \\ & + \frac{M_W}{v\,\cos\theta_W}\cos(\beta-\alpha)Z_\mu(A\partial_\mu h - h\partial_\mu A).\end{aligned} \tag{5.41}$$

If one considers Equation (5.41), it can be noted that in the limit where $\cos(\beta - \alpha) \to 0$, the coupling strength in A-Z-H becomes large while the A-Z-H simultaneously becomes small. This limit applies in the case where H is SM-like. For this reason, in this model a prime search channel for A lies in the $A \to ZH$ decay, if m_A is large enough. If $H \to SS,\ Sh$, then there are two obvious LHC based searches which could already shed light on this decay mode. These are the typical SUSY $Z + E_{\mathrm{T}}^{\mathrm{miss}}$ [152, 153, 154] and the Zh (where $h \to b\bar{b},\ \tau\tau$) searches [155, 156, 157].

Using the model presented above, a `Rivet` analysis was designed to mimic the ATLAS Run 2 $Z + E_{\mathrm{T}}^{\mathrm{miss}}$ selection, and events were passed through this selection after being generated and showered at 13 TeV. The process which was generated is $gg \to A \to ZH$, and thereafter $Z \to \ell\ell$ (where $\ell = e, \mu$) and $H \to SS, Sh$. Both S and h are left open to decay, with S at 150 GeV and having SM-like BRs as well as $BR(S \to \chi\chi) = 0.5$. With $a_1 = 1$, the admixture of SS and Sh is considered to be equal. m_H was considered at 300 GeV, $m_\chi = 60$ GeV and m_A took on the values 600 and 800 GeV. With this choice of parameters, the process described here is well within current limits for mono-jet and $b\bar{b} + E_{\mathrm{T}}^{\mathrm{miss}}$ searches at the LHC, as shown in Table 9.

The results of this are shown in the first four plots in Figure 21. Comparing with the distributions in Reference [152], the shapes of the distributions seem consistent with those from data. The p_{T} of the di-lepton system is sensitive to the mass of A, and can be used as a discriminant for its search. The selection efficiencies for

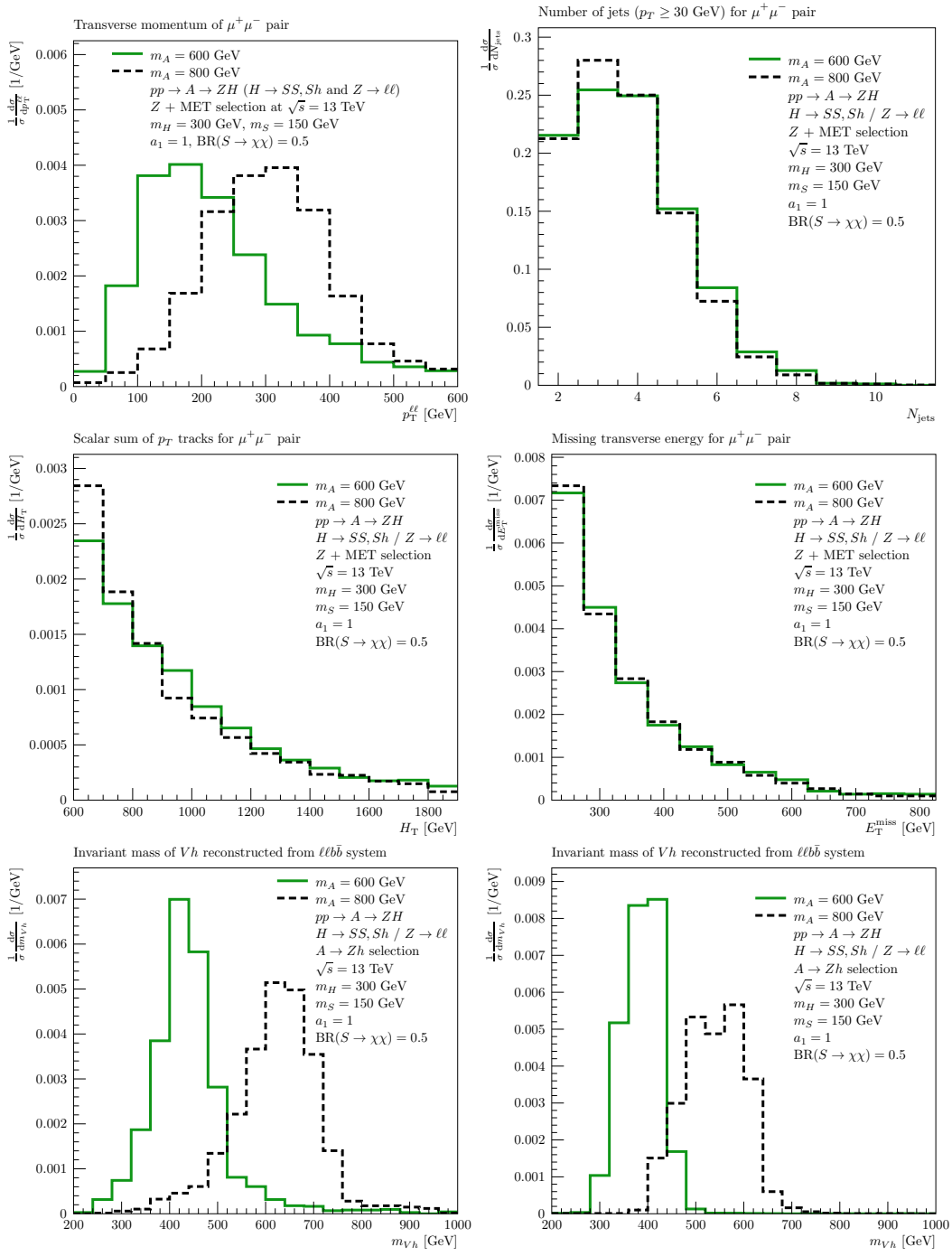


Figure 21: Kinematic distributions of the leptons in $A \to ZH$, where $H \to SS,\ Sh$. The top four pertain to the ATLAS Run 2 $Z + E_{\mathrm{T}}^{\mathrm{miss}}$ SR-Z selection. The bottom two figures pertain to the ATLAS Run 2 $A \to Zh$ $(h \to b\bar{b})$ selection, with the 1 b-tag category on the left and the 2 b-tag category on the right.

Channel/region	Prediction	Experimental limit
Mono-jet with $gg \to H \to SS \to 4\chi$ *at* $\sqrt{s} = 8$ *TeV*		
$E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV	15.1 ± 0.18 fb	229 fb
> 300 GeV	8.90 ± 0.063 fb	98.5 fb
> 350 GeV	5.42 ± 0.023 fb	48.8 fb
> 400 GeV	3.42 ± 0.0093 fb	20.2 fb
> 450 GeV	2.24 ± 0.0040 fb	7.82 fb
> 500 GeV	1.48 ± 0.0017 fb	6.09 fb
> 550 GeV	1.00 ± 0.00080 fb	7.21 fb
$b\bar{b} + E_T^{miss}$ *with* $gg \to H \to Sh \to b\bar{b}\chi\chi$ *at* $\sqrt{s} = 13$ *TeV*		
Signal region	0.10 ± 0.03 fb	1.38 fb
$\gamma\gamma + E_T^{miss}$ *with* $gg \to H \to Sh \to \gamma\gamma\chi\chi$ *at* $\sqrt{s} = 13$ *TeV*		
High $S_{E_{\mathrm{T}}^{\mathrm{miss}}}$, high $p_{\mathrm{T}}^{\gamma\gamma}$	0.265 ± 0.009 fb	12.1 fb
High $S_{E_{\mathrm{T}}^{\mathrm{miss}}}$, low $p_{\mathrm{T}}^{\gamma\gamma}$	0.675 ± 0.014 fb	12.1 fb
Intermediate $S_{E_{\mathrm{T}}^{\mathrm{miss}}}$	3.17 ± 0.03 fb	12.1 fb
Rest	2.80 ± 0.03 fb	12.1 fb

Table 9: Comparisons of the model's predictions for $gg \to H$ against (model-independent) visible cross section 95% CLs in the CMS Run 1 mono-jet [158], the ATLAS Run 2 $b\bar{b} + E_{\mathrm{T}}^{\mathrm{miss}}$ [159], and the ATLAS Run 2 $\gamma\gamma + E_{\mathrm{T}}^{\mathrm{miss}}$ [160] searches. For demonstration, the cross section of $gg \to H$ has been set to an optimistically high value of 10 (20) pb for $\sqrt{s} = 8$ (13) TeV, and yet the prediction is still well within the limits. The mass and parameter points considered here correspond to those chosen in Section 5.5.3. Binomial errors on selection efficiencies have been incorporated into the theoretical predictions. The $\gamma\gamma + E_{\mathrm{T}}^{\mathrm{miss}}$ experimental limit is not presented per category, so for each category the inclusive limit is shown.

the $m_A = 600$ and 800 GeV simulations are 0.68% and 1.86%, respectively. The ATLAS Run 2 excess of ~ 11 events at $L = 3.2$ fb^{-1} can therefore be explained by a $gg \to A \to ZH$ production cross section of the order of tens of picobarns.[7] However, contributions from $pp \to H \to SS, Sh$ production could also be a factor to account for, and in this case there would not only be contributions to the Z peak region (70 GeV $< m_{\ell\ell} <$ 110 GeV), but also in the regions where the di-lepton invariant mass ($m_{\ell\ell}$) is significantly smaller or larger than m_Z. This is due to the fact that in $H \to SS, Sh$, S can have a large BR to WW, and di-lepton pairs will come with $E_{\mathrm{T}}^{\mathrm{miss}}$ in the form of neutrinos for this decay, whereas jets could be found in the decay of the other S or h.

7 Note that since this study was performed, an updated result for the search has been published [161]. While the excess in the updated result is not as significant, it is still compatible with the statements made here.

The same events were passed through a selection mimicking the ATLAS Run 2 $A \to Zh$ (where $h \to b\bar{b}$) search [156]. While there has so far been no significant excess in this channel, it is interesting to understand how the kinematics look for $A \to ZH$. The discriminant of these searches is typically the mass of the vector boson and Higgs boson pair, as reconstructed through a di-lepton and $b\bar{b}$ system in the 2 lepton category (for the 0 lepton category, a transverse mass is calculated instead). The mass of the Zh system is shown by the last two plots in Figure 21. On the right is the 1 b-tag category and on the left is the 2 b-tag category. Both plots are shown in the categories with low p_{T} of the Z boson (the high p_{T} categories have a small selection efficiency). The selection efficiency is dominant in the 2 b-tag category with 2.2% and 1.8% for $m_A = 600$ and 800 GeV, respectively. The mass distributions do not peak at m_A because the final state is not just $\ell\ell b\bar{b}$ – more particles can come from the decay of $H \to SS,\ Sh$, making the final state more diverse. Note that there is also a mass dependence on the b-tag categorisation. This is due to the fact that the $b\bar{b}$ system four vector is scaled to the Higgs mass in the analysis, whereas in this case $S \to b\bar{b}$ could also occur, distorting the kinematics.

5.5.4 *Summary of the lepton studies*

The sections above have presented a set of interesting search channels that could be used to directly search for the model constructed up until this point. However, they pertain mostly to theoretical arguments, showing both that interesting search channels could be proposed and that the proposed model is well within current limits from the LHC. It is up to the experimental collaborations to perform searches that could constrain the parameter space of the model. So far, the ATLAS collaboration has performed a search for the $H \to SS$ decay mode for a few mass points [162], and has set limits in the production cross section times BR. The development of MC samples for this search is shown in Appendix A, along with several other implementations of the BSM model as part of different searches.

Without waiting for more experimental searches, the model can be constrained indirectly through measurements of SM multiple lepton production. There exists

a wealth of ATLAS and CMS results that report on the production of multiple leptons in a generic way. A set of these results is used in the next chapter (Chapter 6) to constrain the production cross section of H, thereby expanding on the studies done in this past chapter. This shall go part-way in testing the model against a more complete set of final states.

6

APPLICATION OF THE MODEL TO ANOMALOUS MULTIPLE LEPTON RESULTS

The iterative cycle of top-down and bottom-up studies discussed in the previous two chapters has ultimately led to the prediction of enhanced production rates of multiple leptons, both with and without additional b-jets. This progression was developed as follows. This early study done in 2015, discussed in Section 4.1, simply considered the possibility of a heavy scalar H being compatible with the set of results in Table 4. The result of this study had shown that with a single parameter β_g^2 (the scale factor for the production cross section of H) a set of ATLAS and CMS Higgs physics results could be fit with a significance of 3σ. The square root of the fit parameter was constrained to a value of $\beta_g = 1.5 \pm 0.6$, which pertained dominantly to the production of the SM Higgs boson in association with two dark matter particles χ from the decay process $H \to h\chi\chi$ through an effective vertex. The best fit mass of H was found to be at $m_H = 272^{+12}_{-9}$ GeV [3].

The next point of interest was to explore the possibility of introducing a scalar mediator S (instead of using effective vertices), such that H could decay to Sh, SS, and hh. This formalism was discussed in Chapter 5. The S was assumed to have globally re-scaled Higgs-like couplings, such that its BRs could be fixed. The possibility of embedding H into a Type-II 2HDM was also discussed [7]. More importantly, however, a predictive set of potential search channels for the new scalars was shown, as detailed in Table 8. A theoretical study on several final states with multiple leptons was shown in Section 5.5. Furthermore, the multiple lepton signatures predicted by the model were tested against the ATLAS and CMS data in Reference [4], where the β_g^2 parameter was constrained to the value of 1.38 ± 0.22 through a simultaneous fit to several independent data sets.

In a similar spirit, this chapter presents the results of an updated fit to the available relevant ATLAS and CMS data with final states containing multiple leptons. However, the focus of this study is shifted away from merely constraining the model. As shall be discussed in this chapter, several results for the production of multiple leptons have shown that the data are not well described by SM MC predictions. Here, it shall be shown that these anomalies in the data deviate significantly from the state-of-the-art MC predictions, and that the failure of SM tools to describe the data can be seen as indirect evidence for new physics processes present at the LHC.

The same simplified model described in Section 5.4 is used along with its assumptions, which acts as a source of multiple lepton production in association with b-jets. The salient points of the model are re-iterated and expanded upon below in Section 6.1. The masses of the new scalars are fixed and so only a single degree of freedom is necessary to be constrained (that is, β_g^2), which eliminates the need for a "look elsewhere effect". Using the simplified model, events are generated and analysed for a statistical comparison with the experimental results. The specific details of these procedures are outlined in Section 6.2. The ensemble of experimental results under consideration here is then discussed in Section 6.3, and thereafter a fit is made to each result with the BSM prediction of the simplified model considered. A combination fit for the entire ensemble of results is also shown. Finally in Section 6.4, the successes and failures of the introduced simplified model are discussed in light of the fit results.

6.1 THE SIMPLIFIED MODEL

The model used in this chapter is constructed as a result of the studies done in Chapters 4 and 5. It is used without appealing to any UV complete theory, and is an anomaly-free simplified model. Among the model's most important assumptions are the existence of two new scalar bosons, H and S. The masses of H and S are considered to be fixed for this study, and take on the values of $m_H = 270$ GeV and $m_S = 150$ GeV. These choices are based on the best-fit values obtained for

the masses from previous studies as performed in Section 4.1 and Reference [4]. This decision was made *a priori*, in order to avoid bias towards fitting the masses. This reduces the number of degrees of freedom in the fits, allowing for a more compelling statement to be made about the results. Therefore, it is not necessary to quantify the results in terms of a look elsewhere effect.

In terms of interactions, H is assumed to be linked to EWSB in that it has Yukawa couplings and tree-level couplings with the weak vector bosons V ($W^{\pm}$ and Z). After EWSB, the Lagrangian describing H is Higgs boson-like. Omitting the terms that are irrelevant in this Chapter, H interacts with the SM particles in the following way:

$$\mathcal{L}_{\text{int}} \supset -\beta_g \frac{m_t}{v} t\bar{t}H + \beta_V \frac{m_V^2}{v} g_{\mu\nu}\, V^\mu V^\nu H. \tag{6.1}$$

These are the Higgs-like couplings for H with the top quark (t) and the weak vector bosons, respectively, and is a more fundamental representation of the interactions of H compared with the effective Lagrangian in Section 4.1. As done in the previous chapters, the strength of each of the couplings is controlled by a free parameter: β_g for the H-t-t interaction and β_V for the H-V-V interaction. The vacuum expectation value v has a value of approximately 246 GeV. The omitted terms include the Yukawa couplings to the other SM fermions and self-interaction terms for H. It can be expected that the couplings to the other SM fermions would also differ by a factor like β_g, however the effect would not make a noticeable difference to the analysis considered in this article and therefore these terms are neglected. Such numbers could also not be deduced from the LHC data at its current reach, but could be considered with future searches for $H \to b\bar{b}$ and $\mu^+\mu^-$, for example.

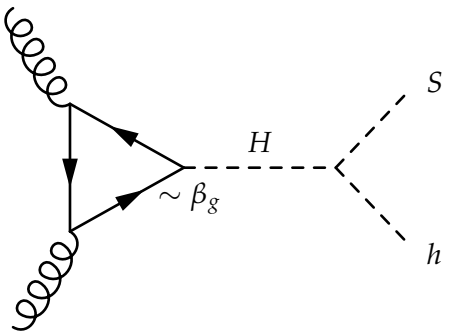
The first term in Equation (6.1) allows for the gluon fusion (ggF) production mode of H. As a baseline, β_g is set to unity such that the H is produced with a Higgs-like cross section. Due to the squaring of the matrix element in width calculations, production cross sections involving this Yukawa coupling are scaled by β_g^2. Therefore, the value of β_g^2 is used as a free parameter in fits to the data. We have set $\beta_V = 0$, such that the coupling of H to pairs of the weak vector bosons

is significantly small; the associated production of H with the weak vector bosons and VBF are negligible production modes.[8] The dominant production mode of H is therefore ggF, while both single (tH) and double (ttH) top associated production of H are also non-negligible. While single top associated production of a Higgs-like boson is usually suppressed due to interference, the implicit assumption of a significantly small H-V-V coupling allows for a sizeable tH production cross section [88]. It has been shown in previous studies [4, 3] (and in Chapter 4) that the tH cross section is enhanced to being approximately that of the ttH cross section. The representative Feynman diagrams for the production modes of H are shown in Figure 22, along with an indication of how the parameter β_g affects diagrams.

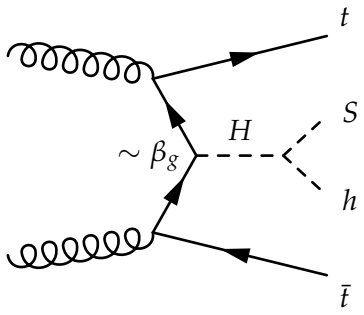
The S boson, on the other hand, is assumed not to be produced directly but rather through the decay of H. In principle, it is possible to include S as a singlet scalar that has interactions with H and the SM Higgs boson h. Doing this would allow the H to produce S bosons through the $H \to SS$ and Sh decay modes. Here it is assumed that the $H \to Sh$ decay mode to have a 100% BR (also shown in Figure 22). These assumptions are all achieved by introducing the effective interaction Lagrangians described in Section 5.4. S is given a VEV and couples to the scalar sector as in Equation (5.37). The couplings are fixed to ensure that the $H \to Sh$ BR is 100% (or equivalently that $a_1 = 0$ – see Equation (5.40)).

Secondly, S is given Higgs-like BRs by fixing the parameters in Equations (5.35) and (5.36). The couplings here are chosen to be globally re-scaled Higgs-like couplings. This is somewhat an arbitrary choice, although it has the dual advantage of fixing the BRs of S (which in turn reduces the number of free parameters in the model) and suppressing the direct production of S. The latter advantage is motivated by the LHC data, since there have been no observations of directly produced Higgs-like bosons near a mass of 150 GeV at the LHC as yet. It is possible to determine an upper limit on the value by which the couplings are re-scaled by considering the ATLAS and CMS searches for a Higgs boson in the $h \to ZZ^* \to 4\ell$ decay channel [163, 125], which provide event yields as a function of the Higgs

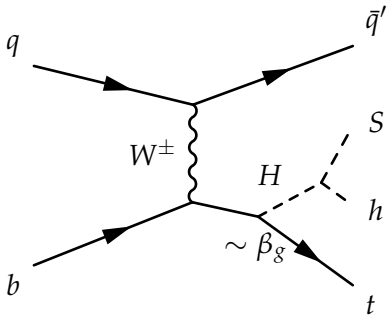
8 A study on the implications of including the VBF production mode is currently underway, but is for the moment left for future studies.



(a) Gluon fusion (ggF).



(b) Top pair associated production (ttH).



(c) Single top associated production (tH).

Figure 22: The representative Feynman diagrams for the leading order production modes of H and its subsequent decay to Sh.

boson mass. By considering the cross section and BR of a Higgs boson at a mass of 150 GeV, assuming that such a particle would have the same reconstruction efficiency and acceptance as the SM Higgs boson, and considering a systematic uncertainty of 10% on the SM prediction, the LHC data exclude a value larger than 0.23 at the 95% confidence level. Squaring this value gives 0.054, which is the limit on the relative rate for such a production compared with the un-scaled production rate. In the analysis done in this paper, the value does not have to be chosen explicitly since only the BRs are needed as inputs, and a direct production mechanism of S would not affect the results. A few of the shortcomings of making the arbitrary choice for the S couplings are discussed in Section 6.4.

6.2 ANALYSIS STRATEGY

6.2.1 *Process modelling*

Predictions of the shapes produced by the BSM processes described above were constructed from MC events as calculated by different event generators. For ggF, the built in matrix elements for BSM Higgs boson production from `Pythia 8` [82] were utilised for the hard scatter process. For the case of tH and ttH, the hard scatter process was simulated at LO using `aMC@NLO` [81]. In the case of tH, a five-flavour proton is used to construct the matrix element. The resonance decays, parton shower and hadronisation are all performed by `Pythia 8` for each BSM process. The masses of the relevant scalars were fixed to $m_H = 270$ GeV, $m_S = 150$ GeV and $m_h = 125.09$ GeV. The cross sections for each BSM process were scaled to the highest order SM Higgs-like production cross sections at a mass of 270 GeV, as taken from the LHC-HXSWG [33]. Such a cross section corresponds to $\beta_g^2 = 1$, and is used as the baseline before the fitting procedure. Following the discussion in Section 6.1, the tH cross section is set to be the same as the given ttH cross section.

The generated events are then passed through the `Delphes 3` [140] fast simulation package to model the response of the appropriate detector. This is done in conjunction with an event selection using the `CheckMATE 2` [164] analysis frame-

work (which, by design, uses the FastJet [165] method to reconstruct jets). Event selections were designed by hand to replicate the cuts and analysis techniques given in the experimental search results. The selection codes were validated by ensuring that the selection for a chosen process given in the experimental publications could be reproduced to within 10% of the value obtained by the custom designed selection codes used for the results discussed here. Given the known discrepancies between the Delphes 3 fast simulation and the full simulation done by the proprietary ATLAS and CMS software, the resulting distributions from the validation process performed very well when compared to the given distributions as shown by the experimental collaborations.

In order to maintain consistency with the experimental results, the SM background predictions and their associated systematic uncertainties were taken directly from the experimental publications. In several cases, additional systematic uncertainties were applied to the SM background predictions; these shall be explained in detail for each individual analysis in Section 6.3.

6.2.2 *Statistical tools*

All fits to the LHC data were performed using the HistFactory extension of the RooStats framework [166], which is a template-based method of performing fits based on maximising a profile likelihood ratio. The SM components of the fits are always taken directly from the published experimental distributions, along with their associated systematic uncertainties, which can be incorporated into the HistFactory schema as variations of a distribution's normalisation or shape. The BSM component is always constructed using a single mass point ($m_H = 270$ GeV and $m_S = 150$ GeV) so as to avoid potential bias and the look elsewhere effect, and therefore only one degree of freedom is considered under the assumptions stated in Section 6.1. The single degree of freedom is β_g^2, which maps directly to the normalisation of the BSM signal with respect to the SM Higgs-like production cross section of H.

The statistical likelihood function $L\left(\beta_g^2 \mid \theta\right)$ is constructed as the product of Poisson probabilities for each bin and in each considered measurement. Systematic uncertainties are incorporated as additional constraint factors in the likelihood, which vary according to their associated nuisance parameters (NPs) θ (Reference [166] contains a full description of how the likelihood is parameterised in terms of different kinds of systematic uncertainties). The general form of the profile likelihood ratio then takes the form:

$$\lambda\left(\beta_g^2\right) = \frac{L\left(\beta_g^2 \mid \hat{\hat{\theta}}\right)}{L\left(\hat{\beta}_g^2 \mid \hat{\theta}\right)}, \tag{6.2}$$

where $\hat{\hat{\theta}}$ is the set of NPs which maximise the likelihood function for a given value of β_g^2, and $\hat{\beta}_g^2$ and $\hat{\theta}$ are the values of β_g^2 and the set of NPs which maximise the likelihood function over the entire parameter space. The best-fit value of the parameter of interest (β_g^2) is identified as the minimum of $-2\log\lambda\left(\beta_g^2\right)$, where a deviation of one unit in this quantity is equivalent to a 1σ deviation from the best-fit point of the parameter of interest. Since the value $\beta_g^2 = 0$ corresponds to the SM-only hypothesis (the *null* hypothesis), the significance of each fit is calculated as:

$$Z = \sqrt{-2\log\lambda\left(0\right)}. \tag{6.3}$$

6.3 FITS TO LHC DATA

Experimental searches for final states containing multiple leptons in high energy proton-proton collisions at the LHC have been performed in a variety of contexts by the ATLAS and CMS experiments. The types of results include searches for the SM production of top quarks decaying to OS lepton pairs, searches for Higgs boson production in leptonic final states and BSM searches for the production of SS lepton pairs, to name a few. Many of these searches involve either a signal or dominant background component that contains top quarks in the final state. Therefore,

Data set	Reference	Selection
ATLAS Run 1	ATLAS-EXOT-2013-16 [167]	SS $\ell\ell$ and $\ell\ell\ell$ + b-jets
ATLAS Run 1	ATLAS-TOPQ-2015-02 [168]	OS $e\mu$ + b-jets
CMS Run 2	CMS-HIG-18-009 [169]	SS $e\mu$, $\mu\mu$ and $\ell\ell\ell$ + b-jets
CMS Run 2	CMS-TOP-17-018 [170]	OS $e\mu$
CMS Run 2	CMS-PAS-SMP-18-002 [171]	$\ell\ell\ell + E_{\mathrm{T}}^{\mathrm{miss}}$ (WZ)
ATLAS Run 2	ATLAS-EXOT-2016-16 [172]	SS $\ell\ell$ and $\ell\ell\ell$ + b-jets
ATLAS Run 2	ATLAS-CONF-2018-027 [173]	OS $e\mu$ + b-jets
ATLAS Run 2	ATLAS-CONF-2018-034 [174]	$\ell\ell\ell + E_{\mathrm{T}}^{\mathrm{miss}}$ (WZ)

Table 10: A list of the ATLAS and CMS experimental results pertaining to final states with multiple leptons that are considered here. For each result, a simple baseline selection is shown. The different kinematic cuts and categories are not shown here, but are described for each analysis below.

the results are often always dependent on the number of b-jets produced with the leptons. Note that any lepton ℓ refers to either an electron or a muon. Contributions from the production of τ-leptons are only relevant if they subsequently decay leptonically.

The ensemble of results considered is shown in Table 10. The majority of results come from the Run 2 data sets, due to the fact that the increased luminosity and cross sections of most of the processes implies a greater statistical precision in the data. The selection of charges for the leptons ensures that each data set is statistically independent, where any potential double counting could only arise through charge mis-identification, and is expected to be negligible. For each result in Table 10, a fit is made using the SM and BSM theoretical predictions discussed in Section 6.2.1 as inputs to the statistical method described in Section 6.2.2. The results of each fit are shown in the sections below. In Section 6.3.8, a combination of all the results is shown.

6.3.1 *ATLAS Run 1 search for SS leptons in association with b-jets*

The production of two SS leptons is a rare process in the SM. This makes it a striking signature for BSM theories that could predict SS lepton pairs via cascaded decays. The ATLAS Run 1 data set was used in a search for SS lepton pairs in

association with b-jets, with the goal of constraining BSM models that predict the existence of heavy vector-like quarks (VLQs) [167]. This kind of search is sensitive to the ttH and tH production modes of the simplified model considered in this study, since a SS lepton pair can be selected from the combination leptonic top quark decays and $S \to VV$ decays. The b-jets from the top quark and $h \to bb$ decays make for a high probability of reconstructing three b-jets in the final state; this is discussed in detail in Reference [139].

The data set for the search is statistically limited, and therefore the overall rates per signal region (SR) are used to fit the BSM prediction in this case (instead of the differential distributions). As a baseline selection, the analysis requires two or three leptons in the final state, with at least one SS lepton pair. The SRs are separated by b-jet multiplicity and different cuts on missing $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T}, which in this case is the scalar sum of the lepton and jet transverse momenta. These cuts are optimised to identify the signal from a model that predicts the production of VLQs, but are still sensitive to the simplified model used here due to relatively low cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$. The SRs are defined as follows:

SRVLQ0:	$N_{b\text{-jet}} = 1$;	$E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV;	$400 < H_{\mathrm{T}} < 700$ GeV,
SRVLQ1:	$N_{b\text{-jet}} = 2$;	$E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV;	$400 < H_{\mathrm{T}} < 700$ GeV,
SRVLQ2:	$N_{b\text{-jet}} \geq 3$;	$E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV;	$400 < H_{\mathrm{T}} < 700$ GeV,
SRVLQ3:	$N_{b\text{-jet}} = 1$;	$40 < E_{\mathrm{T}}^{\mathrm{miss}} < 100$ GeV;	$H_{\mathrm{T}} \geq 700$ GeV,
SRVLQ4:	$N_{b\text{-jet}} = 1$;	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 100$ GeV;	$H_{\mathrm{T}} \geq 700$ GeV,
SRVLQ5:	$N_{b\text{-jet}} = 2$;	$40 < E_{\mathrm{T}}^{\mathrm{miss}} < 100$ GeV;	$H_{\mathrm{T}} \geq 700$ GeV,
SRVLQ6:	$N_{b\text{-jet}} = 2$;	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 100$ GeV;	$H_{\mathrm{T}} \geq 700$ GeV,
SRVLQ7:	$N_{b\text{-jet}} \geq 3$;	$E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV;	$H_{\mathrm{T}} \geq 700$ GeV.

Given the categorisation shown above, a fit was made to the data using the SM+BSM hypothesis on the total rate per SR. An overall normalisation systematic uncertainty was applied to the SM prediction, ranging between 19% and 90% depending on the SR in question. The systematic uncertainty was assumed to be correlated over all of the SRs (i.e. the fit did not allow a bin-by-bin variation of the SM prediction). In the fit, the BSM fit parameter β_g^2 was best fit at a value of 6.51 ± 2.99. This result is relatively high compared to the other results calculated

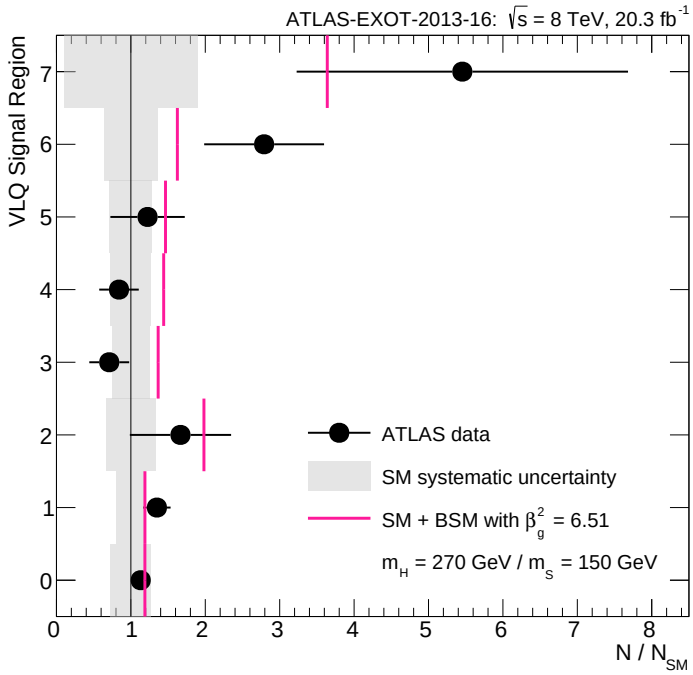


Figure 23: The SM+BSM fit result for the ATLAS Run 1 search for SS leptons in association with b-jets binned by the SRs defined in the text. The horizontal axis represents a production rate normalised to that of the SM. The BSM prediction is scaled to the best-fit value of β_g^2, while the SM and its systematic uncertainty do not incorporate the constraints imposed by the fitting procedure.

in this study. The implications of this are discussed in Section 6.4. In terms of the significance of this deviation from the SM-only hypothesis, this corresponds to $Z = 2.37\sigma$. The performance of the fit per SR can be seen in Figure 23.

Due to a lack of statistics and large systematic uncertainties in this measurement, the significance of the fit is not high enough for this measurement to be noteworthy on its own. However, the kinematic requirements for each SR shows the simplified model's strength in being able to describe excesses in the data with multiple leptons and at least three b-jets. The requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T} are relatively loose in this measurement, compared to other experimental searches that consider the production of heavy particles – SUSY for example. The simplified model is produced dominantly in the region of the phase space with low $E_{\mathrm{T}}^{\mathrm{miss}}$, making these searches of particular interest, as opposed to SUSY searches where the model does not produce a significant signal. The equivalent CMS and ATLAS Run 2 versions of this measurement are discussed in Section 6.3.4 and Section 6.3.5, respectively.

6.3.2 *ATLAS Run 1 di-lepton invariant mass spectrum*

Measurements related to the SM production of top quarks are not typically considered in the search for BSM physics. However, the very simple selection applied to the events considered in such measurements makes for a set of robust distributions against which new physics theories can be tested. From the ATLAS Run 1 data set, a set of differential distributions pertaining to the SM production of top quarks was reported [168]. The events selection in this measurement is a simple selection of an electron and a muon in association with at least one b-jet.

For this measurement, the di-lepton invariant mass is considered as a discriminating variable in the fit. This is because the simplified model used here produces a well defined narrow peak at a value of $m_{\ell\ell} \simeq 50$ GeV, and can be easily distinguished from the SM background peak at around 90 GeV. In addition to this, it is modelled relatively consistently with different event generators in the region where the BSM signal is concentrated. The slight variations of the SM prediction in this region can be covered easily with a systematic uncertainty that affects the SM normalisation.

A fit to this distribution was previously made using the same simplified model in Reference [4]. However, the previous fit was made only by maximising a likelihood for the entire distribution; the effects of systematic uncertainties were inferred after the fit result. Here a fit is presented that takes systematic uncertainties into account by minimising the profile likelihood ratio in Equation (6.2). The applied systematic uncertainty is determined as follows. With the assumption that no significant new physics signals appear in the tail of the distribution, the entire SM prediction is scaled to the data in the region where $m_{\ell\ell} > 110$ GeV; the scale factor was calculated to be 0.984. In doing this, many of the systematic uncertainties that affect the normalisation of the SM prediction become irrelevant. The uncertainties which are not affected by the scaling were then added up in quadrature and found to affect the normalisation of the SM by just under 2%. Therefore, a normalisation systematic uncertainty of 2%, correlated over all of the bins of the distribution is applied to the SM prediction in the fit. In order to avoid bias with

the scaling procedure, the fit is only performed on the bins where $m_{\ell\ell} \leq 110$ GeV. The 2% uncertainty on the normalisation includes the variation of scales for the SM prediction, which only varies by a normalisation factor in the fit region.

The result of the fit gives a best-fit value of β_g^2 at 4.09 ± 1.37, corresponding to a significance of $Z = 2.99\sigma$. The distribution overlaid with the SM+BSM fit is shown in the upper panel of Figure 24. The BSM prediction performs very well in its ability to explain the excess in the first few bins of the distribution. In the fit, the SM prediction is raised by slightly less than 1% in the profiling of the systematic uncertainty, such that the peak of the SM distribution is also fit well. It should be noted that the tail of the SM distribution, while not included in the fit, is still compatible with this constraint.

6.3.3 *CMS Run 2 di-lepton invariant mass spectrum*

Similar to the ATLAS Run 1 di-lepton invariant mass spectrum discussed above in Section 6.3.2, CMS have more recently published an $m_{\ell\ell}$ spectrum with a partial Run 2 data set [170]. The key difference between the ATLAS Run 1 result and the CMS Run 2 result is that CMS does not place any requirements on the number of b-jets in the final state. The CMS result, therefore, has a significant contribution from the SM Drell-Yan process, which enhances the SM background and shifts the inclusive peak to around 70 GeV. The fitting procedure for this measurement is similar to that of the ATLAS Run 1 results described in Section 6.3.2, with some differences that will be described below.

On the inspection of the $m_{\ell\ell}$ distribution in Reference [170], it is clear that there exists an excess of events at low invariant mass values, consistent with the results discussed in Section 6.3.2. However, it also becomes evident that the entire background SM prediction is poorly modelled with respect to the data, the tail of the distribution is underestimated by the theoretical prediction. This is most likely due to a discrepancy present in the nominal $t\bar{t}$ MC prediction used in the measurement (that is, the `POWHEG V2` [175] $t\bar{t}$ sample). This is justified by noting the $m_{\ell\ell}$ differential distribution in the measurement of $t\bar{t}$ fiducial cross sections [176],

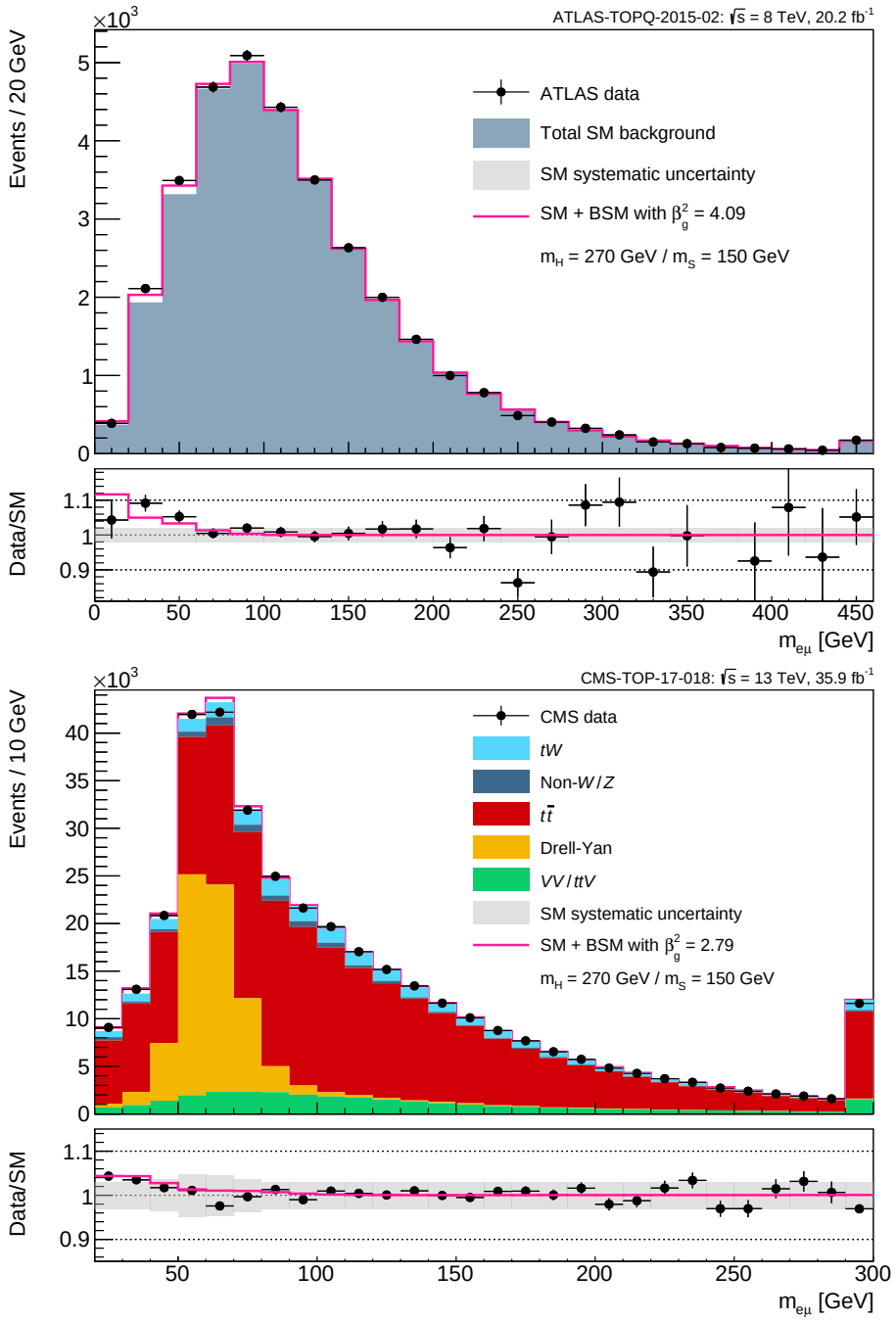


Figure 24: The SM+BSM fit result for the di-lepton invariant mass spectrum reported by ATLAS in Run 1 (above) and CMS in Run 2 (below). In each case, β_g^2 has been scaled to its best-fit value. The SM predictions and systematic uncertainties do not incorporate the constraints imposed by the fit.

which corresponds to a fiducial phase space enriched by $t\bar{t}$ events. One can note from this distribution that the same POWHEG V2 sample does not describe the data as well as the other two $t\bar{t}$ samples it is tested against. For this reason, the distribution in Reference [176] is used to re-weight the $t\bar{t}$ distribution in Reference [170]. In doing so, the large excess at low values of $m_{\ell\ell}$ is reduced and the tail is flattened, such that the entire distribution is able to describe the data far better after the re-weighting (up to a normalisation factor). The origin of this apparent discrepancy in the modelling of the $t\bar{t}$ process is uncertain.

The theoretical SM MC predictions are scaled to the data in the region where $m_{\ell\ell} > 110$ GeV, and the fit is done in the region where $m_{\ell\ell} < 110$ GeV (similarly to the ATLAS Run 1 $m_{\ell\ell}$ distribution). The Drell-Yan prediction, however, is not altered by this scaling. Instead, the Drell-Yan prediction is left at its nominal normalisation, and an exceptionally large normalisation systematic uncertainty of 6.86% is applied to it. All of the other SM components are given a normalisation systematic uncertainty of 3%. Similarly to Section 6.3.2, the scale uncertainties on the $t\bar{t}$ background only affect the normalisation of the distribution in the region where $m_{\ell\ell} < 110$ GeV, and so the 3% uncertainty incorporates the scale uncertainties. Note also that uncertainties related to the choice of event generator are small as in Section 6.3.2, and are therefore covered by the conservative normalisation uncertainty of 3%. The result of the re-weighting and scaling procedures is an $m_{\ell\ell}$ distribution that is very well described at the peak and in the tail, however still with a significant excess of events at low $m_{\ell\ell}$. This can be seen in the lower panel of Figure 24, along with the SM+BSM fit result.

The fitting process favours the SM+BSM hypothesis with a significance of $Z = 5.45\sigma$ (this fit has the highest significance for all of the individual fit results in the ensemble in Table 10). The corresponding best-fit value of β_g^2 is 2.79 ± 0.52. The statistical precision of the measurement is the main reason why the significance of the SM+BSM fit is so high. For one, the statistical uncertainty on the data is negligible compared with the systematic uncertainty on the SM prediction. In addition to this, the profiling of the systematic uncertainties places strong constraints on the best-fit normalisation of the SM background. Due to the statistical precision

of the data set, and the excess of events below $m_{\ell\ell} = 60$ GeV, there is a strong tension between the SM-only hypothesis and the data. Since the BSM prediction is distributed exactly where the excess of events is, the SM+BSM fit resolves this tension with a large significance. It should be noted that any significant variation of the normalisation or shape of the SM background in the fit would have negative consequences on the compatibility of the tail of the distribution with the data, which is so well described by the SM prediction. In the SM+BSM fit, however, the variation of the normalisation of the SM backgrounds is negligible (less than 0.1%).

With such a significant effect, it was deemed necessary to search for other measurements in which the excess might be localised. This was found in the $t\bar{t}/Wt$ control region (CR) of the ATLAS Run 2 Higgs production cross section measurement in the $WW^* \to e\nu\mu\nu$ decay channel [177]. The BSM model is sensitive to the selection criteria of this measurement since it requires one high p_T b-jet and exactly one central un-tagged jet. Since the discriminating variable of the search in this measurement is that of transverse mass (m_T), the excess may localise to a broad peak in $S \to WW^*$ decays. It was decided not to include this measurement in the fit, since its event selection overlaps significantly with that of the measurement discussed in Section 6.3.6, and including it in the combination would be double counting. In an attempt to reduce the experimental systematic uncertainties on the distribution, it was determined that the BSM model would only produce a signal in the region of $m_T < 200$ GeV, and therefore the distribution was scaled to match the integral of the data in the region above 200 GeV.

Doing this reveals a broad structure in the data, compared with the SM prediction, that peaks at around $m_T = 150$ GeV. This can be seen in Figure 25. Such a structure could be well described by a resonance decaying to a pair of W bosons in association with a b-jet; in this case the S boson (having a mass of 150 GeV) is a prime candidate. The BSM m_T distribution, however, does not have a peak where the data peaks. Due to the off-shell nature of the $H \to Sh$ decay, the distribution peaks below the Higgs mass.[9] The SM+BSM fit still improved on the SM-only hy-

9 This also implies that small changes in the masses of H and S would not drastically shift the peak. As long as the $H \to Sh$ decay is off-shell, the position of the peak is saturated towards a value below the Higgs mass.

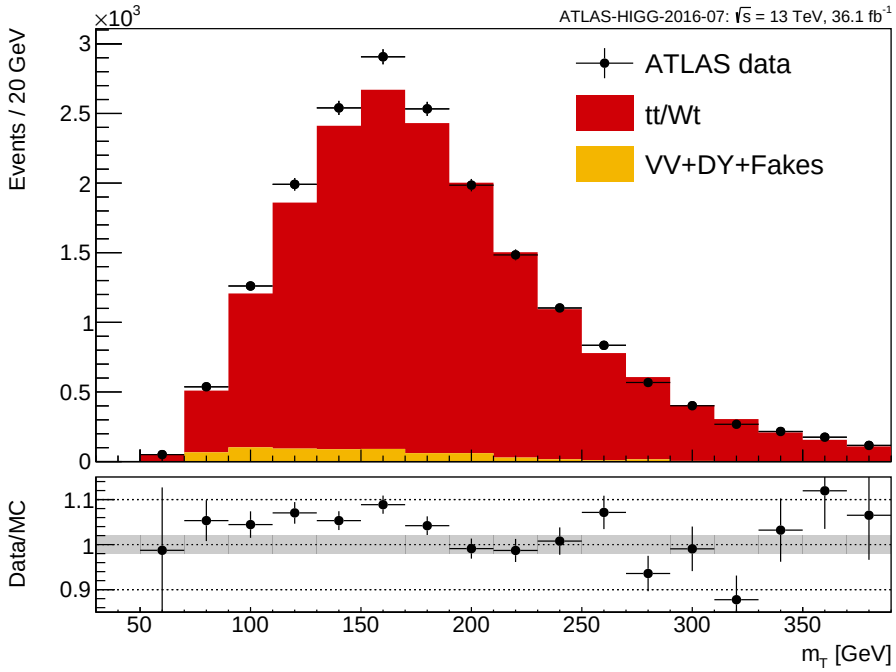


Figure 25: The transverse mass distribution in the ATLAS Run 2 $h \to WW^*$ top CR, scaled such that the integral in data and the SM MC match in the region of $m_{\mathrm{T}} > 200$ GeV. An estimated flat systematic uncertainty of 2% is drawn only as a baseline.

pothesis, however the improvement was weak (less than 2σ). This incompatibility with the simplified model described in this thesis is an exciting prospect, and such a measurement might aid in future developments of the model towards potentially more complicated new physics processes.

6.3.4 *CMS Run 2 search for SS leptons in association with b-jets*

The CMS Run 2 data set was used to make a search for the SM Higgs boson in association with a single top quark [169]. The event selection in this search is very similar to the ATLAS searches for SS leptons in association with b-jets (discussed in Section 6.3.1 for Run 1 and below in Section 6.3.5 for Run 2). The simple event pre-selection requires two or three leptons in the final state with at least one SS pair. In addition, at least one b-jet is required along with at least one additional un-tagged jet. The events are further categorised into three categories, $e\mu$, $\mu\mu$ and $\ell\ell\ell$, depending on the multiplicity and flavour of the leptons.

A very simple fit was made using this measurement in the results of Reference [4], where a total rate of events for $N_{\text{jet}} \geq 3$ with and without the BSM signal was compared for each category. Here, the fit is extended by combining all three categories for a greater statistical precision. The variable that is used to fit the BSM prediction is the highest pseudo-rapidity for high p_{T} jets (that is, jets with $p_{\text{T}} > 40$ GeV). This decision was made based on the fact that the systematic uncertainty on the SM is smaller for this variable than that of the other variables considered in the search (in addition to a reasonable signal to background discrimination). The applied systematic uncertainty is considered to be an overall normalisation variation, calculated as the sum of the total systematic variations from each category.

The BSM contamination in this signal region is non-trivial. Whilst one might expect that the dominant contribution to the event selection comes from the ttH production mode, it is actually the ggF production mode that is dominant. This is due to the much larger ggF cross section compared to that of ttH and tH, in addition to the fact that leptons from heavy-flavour decays in the ggF production mode contribute to the signal with a non-negligible probability. These effects are usually accounted for in the lepton isolation criteria used by the experiment. However, due to the ambiguity of these criteria described in this particular measurement, a crude estimate was made on the probabilities for both prompt and heavy-flavour decay leptons to be accepted or rejected in the selection. This estimate was validated using event yields for known processes given in Reference [169], and was found to perform well.

The result of the fitting process is shown on the left in Figure 26. Note that, although a bin-by-bin variation is shown as a systematic uncertainty in the plots in Figure 26, the fit makes use of an overall normalisation variation. The resulting best-fit value of β_g^2 is 1.41 ± 0.80. This corresponds to a significance of $Z = 1.75\sigma$. The overall agreement of the SM+BSM prediction compared with the data in the fit looks reasonable. The addition of the BSM prediction helps to explain the overall elevation of the data compared with the SM, although this is not a significant effect due to the fact that this elevation can be covered by the systematic uncertainty. A

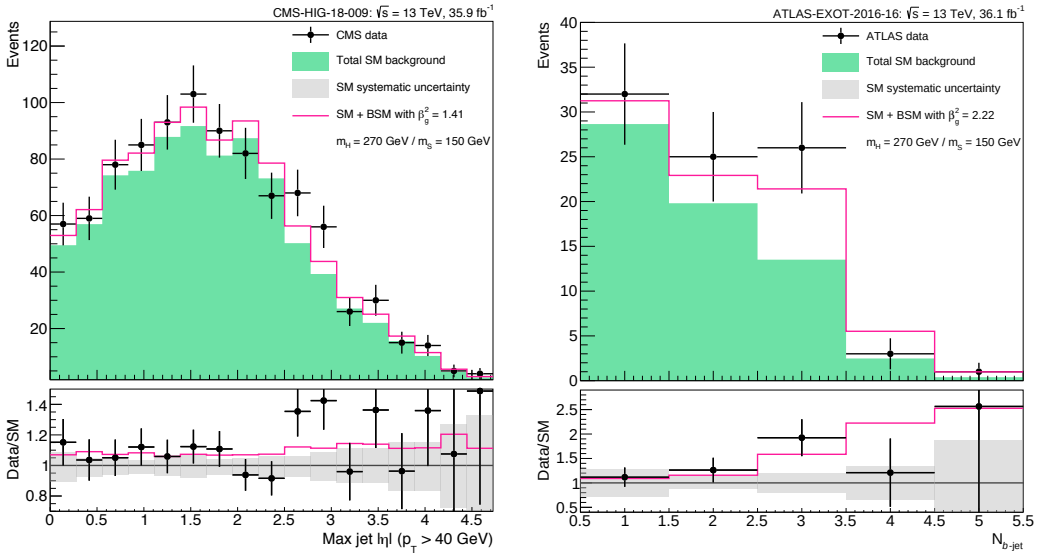


Figure 26: The SM+BSM fit results applied to searches for SS leptons in association with b-jets in the CMS (left) and ATLAS (right) Run 2 data sets. The discriminating variables in the fit are the highest value of pseudo-rapidity for high p_{T} jets in the CMS search and the b-jet multiplicity in the ATLAS search. In each case, β_g^2 has been scaled to its best-fit value, while the SM predictions are shown at their nominal values without any fit constraints. The systematic uncertainties are shown as bin-by-bin effects, as reported by the experimental collaborations. However, in the fit an overall normalisation systematic uncertainty is considered for each measurement.

more interesting feature is the ability of the BSM prediction to partially explain the greater elevation for events with forward jets (that is, $|\eta| > 2.5$). The signal contribution in this region is dominated by the tH and ttH production mechanisms. Due to the fact that the statistical uncertainty is still large for this measurement, it would be interesting to revisit this analysis with more data.

6.3.5 *ATLAS Run 2 search for SS leptons in association with b-jets*

The Run 2 version of the ATLAS search for SS leptons in association with b-jets [172] provides a more statistically precise and systematically constrained measurement than that of Run 1 (the result discussed in Section 6.3.1). The categorisation into SRs is slightly different, as in this case a selection of auxiliary plots show differential distributions combining all the SRs. Therefore, the details of the SRs are not important in this study; but it should be noted that the combination of all SRs have minimal cuts of $E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV and $H_{\mathrm{T}} > 500$ GeV.

Based on a study conducted on signal to background ratios (using the MC predictions), it was clear that the most sensitive variable to the BSM prediction is that of the b-jet multiplicity. In the ATLAS data, the distribution of b-jet multiplicity deviates from the SM in the bin with exactly 3 b-jets. The top associated production modes of the BSM scalar H come with several b-jets in the final state, due to the decays of both the top quarks and the intermediate SM Higgs boson in the process. The ggF BSM production mode has a relatively small acceptance into the event selection of this measurement, and therefore the overall BSM prediction does a good job of explaining the excess in the distribution of b-jet multiplicity. Similar to the CMS result in Section 6.3.4, a single normalisation systematic uncertainty was considered for this measurement, corresponding to the overall systematic uncertainty on the SM background.

The best-fit point for β_g^2 in the SM+BSM fit is found at 2.22 ± 1.19. This corresponds to a significance of $Z = 2.01\sigma$. The result of this fit is shown in the distribution of b-jet multiplicity on the right of Figure 26. While the significance of the fit is not particularly high, this is still an important result in this study. This is because Reference [172] provides the only distribution of b-jet multiplicity from ATLAS that can be related to the top associated production of the Higgs boson in multi-leptonic final states. As demonstrated in Reference [4], the ttH and tH BSM production modes discriminate most strongly against the SM prediction of b-jet multiplicity, and this effect can be seen in the data here. A similar (yet very strongly correlated) excess can be seen in the more recent ATLAS results in the search for four-top-quark production [146].

6.3.6 *ATLAS Run 2 "top spin correlations"*

Measurements of the azimuthal angle between OS leptons have historically been used to understand spin correlations in top quark decays. The most recent study of the distribution, as presented by the ATLAS collaboration using a partial Run 2 data set [173], has shown that a significant deviation from the SM exists. This result is interpreted by the experimental collaboration as being an indication of

mis-modelled top quark spin correlations in the SM $t\bar{t}$ MC predictions. Here it is argued that the deviation is not necessarily due to the mis-modelling of the SM $t\bar{t}$ production process, but that a contamination of the proposed BSM signal studied here can alleviate the discrepancy between the data and the SM prediction. Recent results on higher order corrections have claimed that the incompatibility with the data is alleviated in the NNLO description of the $t\bar{t}$ prediction [178]. However, as detailed in Reference [5], this is not holistically correct, since it does not take into account the correlation of higher order corrections with the shape of the $m_{\ell\ell}$ distribution.

It was decided to consider the "inclusive" distribution of $\Delta\phi(l^+, l^-)$, the difference in azimuthal angle between different-flavour OS (DFOS) di-leptons, in order to understand the BSM signal's effect on the measurement. This selection requires that events have at least one b-jet. The discrepancy can be seen in the detector level distribution, where in the data the lower values of $\Delta\phi(l^+, l^-)$ are underestimated by the SM prediction, and the higher values are overestimated. However, it is also clear that there exists some uncertainty in the different SM MC predictions at high values, whereas the different MC predictions agree relatively well at low values. Due to this discrepancy the `aMC@NLO + Pythia 8` prediction was chosen as a baseline, since it does the best job of describing the data in the region of high $\Delta\phi(l^+, l^-)$. Thereafter, a conservative bin-by-bin systematic uncertainty was applied to cover the variation of the other SM $t\bar{t}$ predictions, constructed using the difference of the MC predictions tested in the ATLAS result. This uncertainty varies from 0.3% at $\Delta\phi(l^+, l^-) = 0$ to around 2.5% at $\Delta\phi(l^+, l^-) = \pi$. In addition to this, all the appropriate experimental systematic uncertainties were applied to the SM predictions. The BSM contribution to the distribution is dominated by the ggF production mode. In general, the DFOS leptons come from the $S \to W^+W^- \to e^{\pm}\nu\mu^{\pm}\nu$ decays, whereas any extra b-jets come from the $h \to b\bar{b}$ decay mode. Since the di-lepton pair comes from a cascaded decay via the heavy scalar H, the $\Delta\phi(l^+, l^-)$ spectrum produced by the BSM ggF production process peaks at low values, which is opposite to that of SM $t\bar{t}$ production.

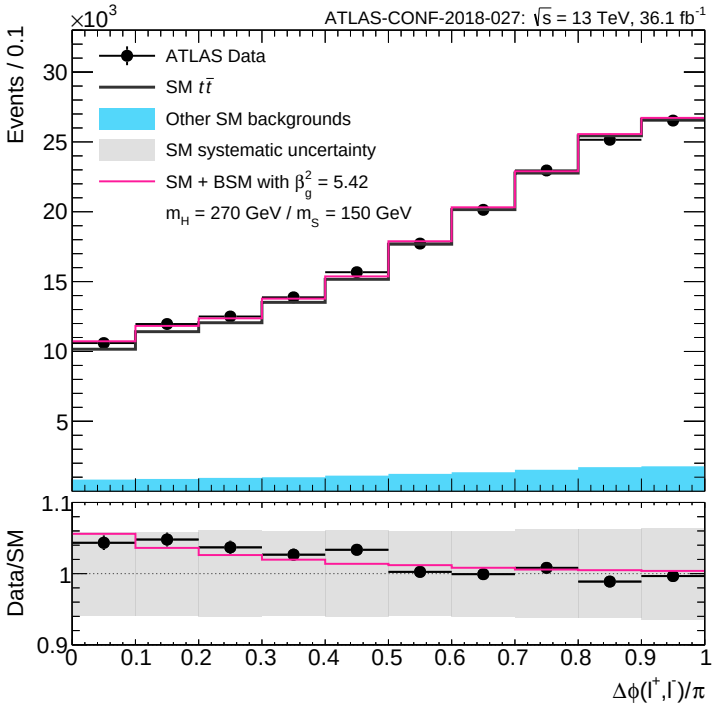


Figure 27: The SM+BSM fit result for the DFOS $\Delta\phi(l^+, l^-)$ distribution reported by ATLAS using a partial Run 2 data set. Here, β_g^2 has been scaled to its best-fit value. The SM predictions and systematic uncertainties do not incorporate the constraints imposed by the fit.

The SM+BSM fit does a remarkable job in describing the excess of data in the low end of the $\Delta\phi(l^+, l^-)$ spectrum. The best-fit value of β_g^2 is slightly higher than most of the other fit results, and is fit at 5.42 ± 1.28, corresponding to a significance of 4.06σ. The result of this fit can be seen in Figure 27. Note that, like all of the other data comparisons, the systematic uncertainties and SM predictions are shown at their nominal and un-scaled values. It is only the BSM prediction that has been scaled to its best-fit normalisation. The inability of the SM to describe the data (even within systematic uncertainties) is due to the fact that the dominant systematic uncertainties only affect the overall normalisation of the SM, whereas the excess in the data clearly has a shape dependence. The BSM prediction matches this shape dependence very well, and hence the SM+BSM fit has a high significance.

6.3.7 *ATLAS and CMS Run 2 measurements on WZ production*

Up until this point, this study has been concerned mostly with measurements that have dominant components relating to top quark production. It is therefore fair to assume that the fit results up until this point might be biased towards potential mis-modelling of SM top production processes. However, it can be shown that multi-lepton excesses exist also in measurements that are dominantly sensitive to EW processes.

To demonstrate this, one can consider the SM measurements of WZ production as presented by the ATLAS [174] and CMS [171] experiments in their Run 2 data sets. These measurements both select events with exactly three leptons, two of which must be a same-flavour opposite-sign (SFOS) pair with a mass close to the Z boson mass. A cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ is also made to select events containing a leptonically decaying W boson. The main difference between the ATLAS and CMS event selections is that the CMS selection vetoes events containing b-jets, whereas the ATLAS selection does not apply such a constraint. The event selections applied to these searches are almost completely orthogonal to the other measurements considered here.

The only common distribution shown in the SR for both ATLAS and CMS is that of the Z boson p_{T} (that is, the p_{T} of the SFOS di-lepton system with a mass closest to the Z boson mass). Since this variable relies only on the performance of reconstructing the momentum of light leptons, it is therefore relatively robust and not likely to suffer from theoretical mis-modelling. For this reason, it was chosen to be the discriminating variable in the SM+BSM fit. The SM prediction of the Z boson p_{T} in the WZ production process was calculated at NLO in terms of QCD corrections in both the ATLAS and CMS measurements, but not with the NLO EW corrections. A study on the recent literature in SM WZ production at the LHC has shown that the current predictions are relatively robust, with the overall NLO EW corrections having only a small effect on the Z p_{T} spectrum [179, 180].

For these measurements, the ggF BSM production mode again dominates over the top associated modes in terms of contamination into the SRs. An MC study

showed that the BSM prediction studied here seldom produces a SFOS lepton pair close to the Z mass, and therefore the acceptance is still relatively low. However, due to the fact that the Z boson is most often produced through a cascaded off-shell decay (through the $h \to ZZ$ or $S \to ZZ$ decay mode), it has a very low p_{T} on average. Therefore, it described the mild excesses seen in the ATLAS and CMS data at low Z p_{T} relatively well.

It was decided that the Z boson p_{T} would not be scaled to match the integral of the data in the tail of the distribution (which was done in the $m_{\ell\ell}$ distributions discussed in Section 6.3.2 and Section 6.3.3). Therefore, the full set of experimental systematic uncertainties was applied to the SM distributions in the fitting procedure. It was also evident that a comprehensive study on the theoretical uncertainties in the region of low Z p_{T} for WZ production must be understood so that the significance of the excess in the data could be characterised more accurately. This was done in a few different respects.

Firstly, a scale uncertainty as a function of the Z p_{T} was derived by varying the dynamical renormalisation and factorisation scales (μ_{R} and μ_{F}, respectively) at NLO using `aMC@NLO` and `Pythia 8`. The central dynamic scale is calculated on an event-by-event basis, and is defined as half of the scalar sum of transverse momenta for the final state particles ($H_{\mathrm{T}}/2$). For the check, events were generated and passed through the `Delphes 3` fast simulation package. The CMS Run 2 WZ event selection was used for the purposes of this check. The effect of varying μ_{R} and μ_{F} by a factor of 0.5, 1, and 2 was determined by finding the maximum and minimum deviations from the nominal case in terms of the differential cross section as a function of the Z p_{T}, as shown in the top of Figure 29. As can be seen in the spectrum, the associated scale uncertainty is of the order of 5% in the region of $p_{\mathrm{T}} < 100$ GeV, and grows to around 10% at high p_{T}. This check was also performed with the `POWHEG` event generator and the `Herwig 7` parton shower, the results of which are compatible with the result above.

Secondly, the effect of changing the matrix element and parton shower of the event generation process was studied by comparing the Z p_{T} for all the combinations of events generated with `POWHEG` and `aMC@NLO`, and showered with `Pythia 8`

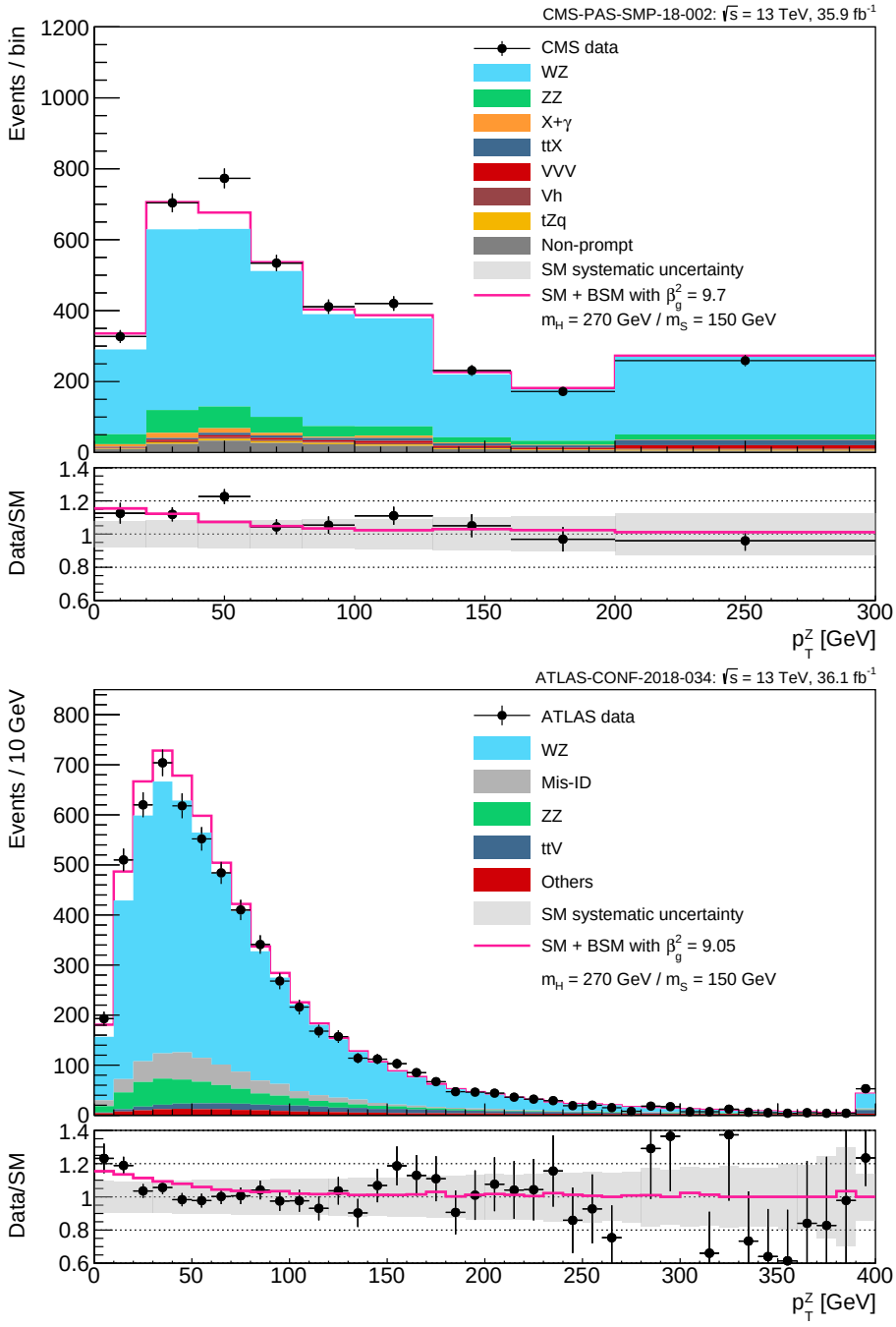


Figure 28: The SM+BSM fit result for the CMS (top) and ATLAS (bottom) measurements of SM WZ production. Only the BSM prediction has been scaled to its best-fit value. The SM predictions and their associated uncertainties have been left at their nominal values.

and Herwig 7 [181]. The normalised Z p_{T} spectrum for the four different combinations can be seen in the middle of Figure 29. Here, the nominal case can be thought of as POWHEG + Pythia 8, since both ATLAS and CMS use this for their measurements. The biggest deviation from the nominal is therefore aMC@NLO + Pythia 8, the reason of which is unclear. However, it can be said that changing from POWHEG to aMC@NLO would have accentuated the apparent excess at low Z p_{T} by an unrealistic amount. Therefore, it was decided that POWHEG + Pythia 8 models the SM WZ production the best, and so this prediction was used in the fits.

Next, the effects of changing PDF sets was studied by producing the Z p_{T} spectrum with the CT14 [182], MMHTnlo [183] and PDF4LHCnlo [184] PDF sets. This effect was seen to only alter the normalisation of the prediction slightly, and therefore was not considered a significant systematic uncertainty.

Finally, due to the fact that the MC event generators which are used in the measurements do not account for higher order EW corrections, an estimate on the effect of the real radiative corrections to the Z p_{T} was made by altering the Pythia 8 parton shower. SM WZ events were generated at LO in QCD using Pythia 8, in order to separate the EW corrections from the QCD corrections. The Z p_{T} spectrum with and without photonic emissions in the shower can be seen in the bottom of Figure 29. It can be seen that the most significant correction to the p_{T} in terms of real EW corrections comes at high p_{T}, which is far from the region in which the BSM model discussed here predicts a signal. In this region, the current statistical imprecision of the data dominates over the potential for the Z p_{T} spectrum to be significantly altered by EW corrections. Of course, virtual EW corrections should also play a role, however these will be much smaller than the real corrections.

In both the ATLAS and CMS measurements the SM+BSM fit favoured the best-fit value of β_g^2 to be rather high, at 9.05 ± 3.35 for the ATLAS measurement and 9.70 ± 3.88 for the CMS measurement. This corresponds to significance values of 2.52σ and 2.36σ, respectively. The distributions with the BSM prediction scaled to the best-fit value can be seen in Figure 28.

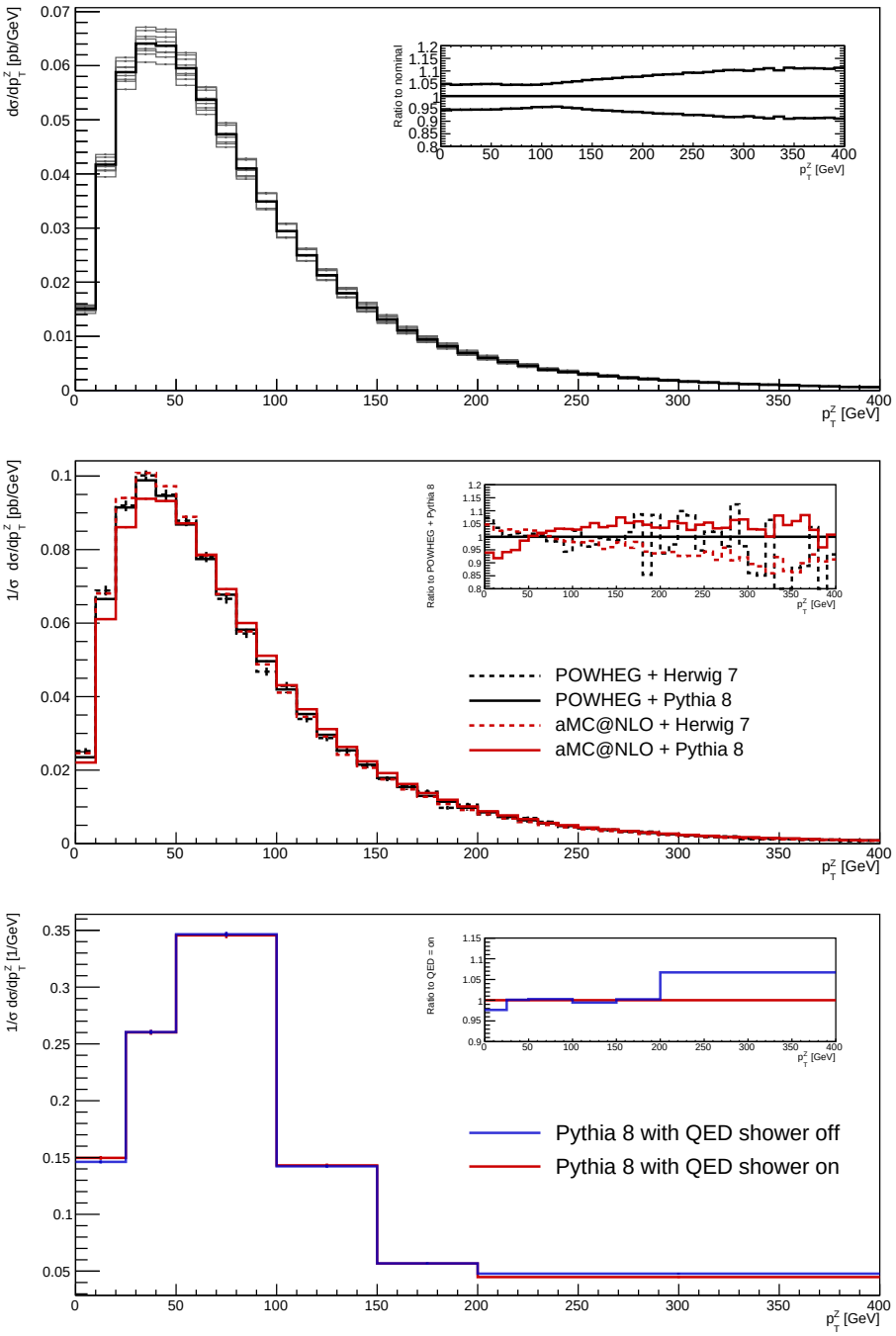


Figure 29: Theoretical studies on how different effects alter the differential cross section and shape of the SM WZ process as a function of the Z p_{T}. On the top, scale uncertainties as determined by `aMC@NLO` and `Pythia 8`. The thick black line represents the spectrum at the nominal scale, and each grey line is a variation of the scale. The insert shows the maximum and minimum relative deviations for all scale variations. In the middle the effect of changing the event generator and parton shower program, normalised to unity. On the bottom, the approximate effect of EW corrections determined by switching photonic emissions on and off in `Pythia 8`, normalised to unity.

Selection	Best-fit β_g^2	Significance
ATLAS Run 1 SS leptons + b-jets	6.51 ± 2.99	2.37σ
ATLAS Run 1 DFOS di-lepton + b-jets	4.09 ± 1.37	2.99σ
ATLAS Run 2 SS leptons + b-jets	2.22 ± 1.19	2.01σ
CMS Run 2 SS leptons + b-jets	1.41 ± 0.80	1.75σ
CMS Run 2 DFOS di-lepton	2.79 ± 0.52	5.45σ
ATLAS Run 2 DFOS di-lepton + b-jets	5.42 ± 1.28	4.06σ
CMS Run 2 tri-lepton + $E_{\mathrm{T}}^{\mathrm{miss}}$	9.70 ± 3.88	2.36σ
ATLAS Run 2 tri-lepton + $E_{\mathrm{T}}^{\mathrm{miss}}$	9.05 ± 3.35	2.52σ
Combination	2.92 ± 0.35	8.04σ

Table 11: A summary of the SM+BSM fit results for each measurement, along with the result of their combination.

6.3.8 *Combination*

Each of the results studied here makes use of a profile likelihood ratio to constrain the single fit parameter β_g^2 under the SM+BSM hypothesis. With these profile likelihood ratios constructed as a function of β_g^2, it is relatively straightforward to perform a simultaneous fit on all of the results considered, and therefore make a combination of the independent data sets under the SM+BSM hypothesis. The combined profile likelihood is constructed by multiplying the profile likelihood ratios for each individual measurement. Then, the best-fit value of β_g^2 and significance can be calculated similarly to the individual results (i.e. by minimising Equation (6.2) and using Equation (6.3)). Doing so constrains the parameter β_g^2 to the value 2.92 ± 0.35, which corresponds to a significance of $Z = 8.04\sigma$ in favour of the SM+BSM hypothesis over the SM-only hypothesis. A summary of all the individual fit results, as well as the combination, can be seen in Table 11. In addition to this, each of the individual profile likelihood ratios are shown in Figure 30, with the combined case shown in black.

In terms of the combination that we have performed, the calculated significance can only be treated as an estimate of the "true" value, due to the fact that we lack much of the necessary information to formally combine the results. For one, statistical correlations for the effects of systematic uncertainties have not been accounted

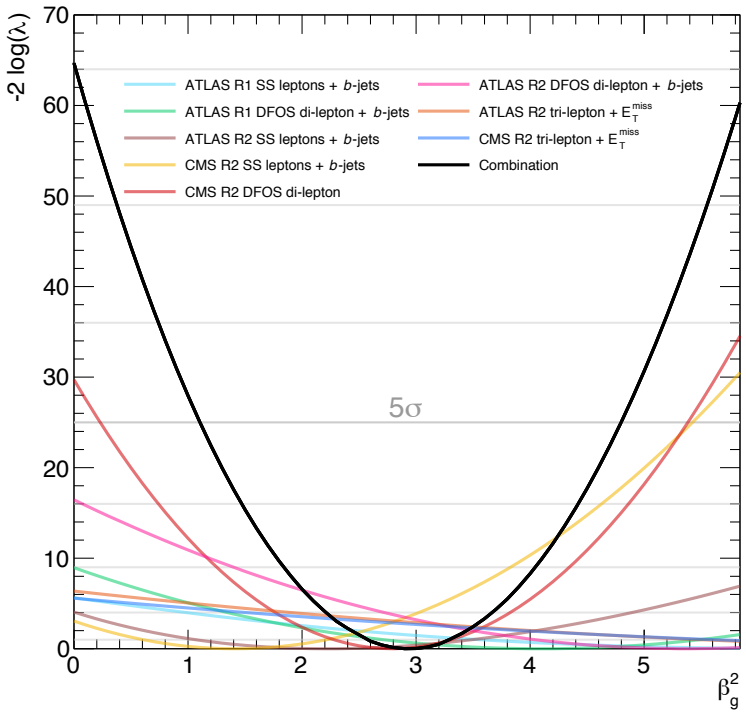


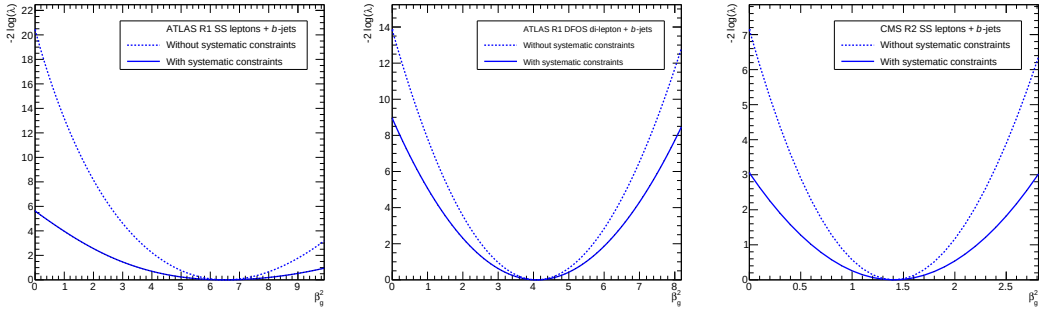
Figure 30: The (negative 2 times the logarithm of the) profile likelihood ratios for each of the individual fit results, overlaid with that of their combination. The significance of a result is calculated as the square root of the point which intersects the y-axis (that is, the SM-only hypothesis where $\beta_g^2 = 0$).

for. Having said this, since there exists a diverse set of measurements in the combination, it is not obvious that incorporating such correlations would have a big effect on the final fit result. Another necessary bit of information that is lacking is the exact way in which the experimental systematic uncertainties affect the shapes of the SM distributions that have been fit. The best possible approximation that could have been made was to incorporate bin-by-bin variations for those that do have understandable effects. However, it stands to reason that a more rigorous and insightful application of systematic uncertainties to the distributions could change the final fit results.

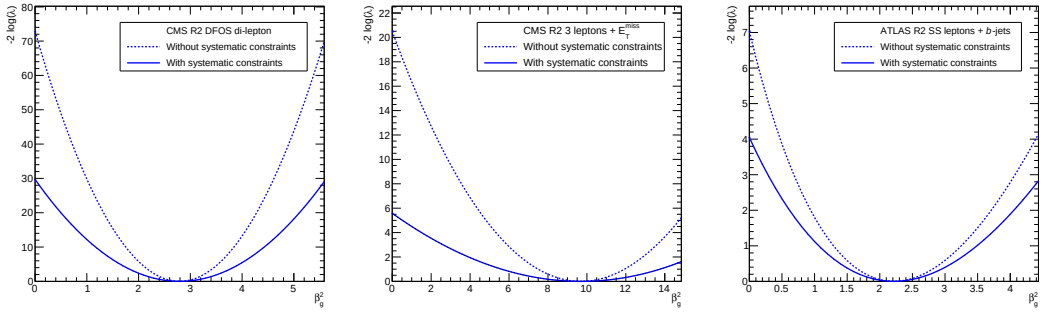
Due to the high statistical precision of many of the results studied here, it is important to understand the effect of systematic uncertainties on each of the fit results. In Figure 31, for each measurement considered, the profile likelihood ratio is shown overlaid with the corresponding negative log likelihood (NLL), as a function of β_g^2. Since the NLL does not contain information about the overall constraints of the systematic uncertainties, the comparison of the two is a good measure by which one can understand the effects of systematic uncertainties on the fit results.

6.4 DISCUSSION

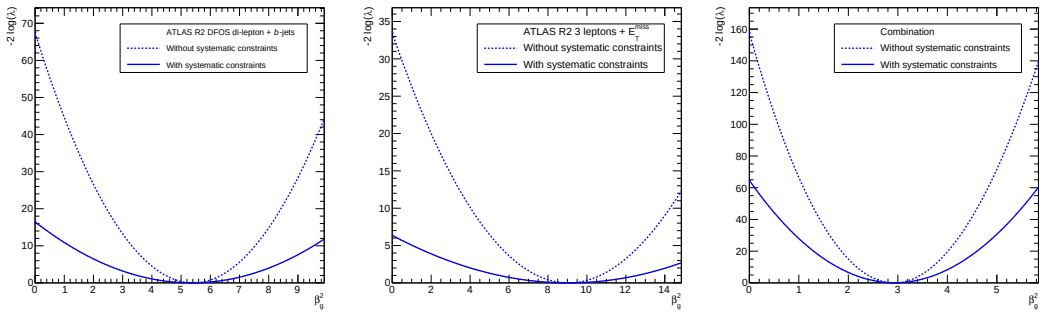
It goes without saying that the large combined significance discussed above should be subject to some criticism. From a statistics standpoint, a common issue raised in such circumstances is that of a look elsewhere effect. Traditionally, a look elsewhere effect will suppress the significance of a fit result that was performed on an unexpected (or un-predicted) deviation from the null hypothesis. This has the advantage of reducing cognitive bias in terms of model building, such that it is a safeguard against tuning a model's parameters in order to describe a fluctuation. In terms of the fits performed here, it is strongly believed that a look elsewhere effect is not appropriate. The measurements against which the model was fit were decided based on the signatures of the model as described in previous studies, both in Chapter 4 and in Reference [4]. In addition to this, the mass points were decided to be fixed *a priori*, and therefore the model was not tuned to explain the



(a) ATLAS-EXOT-2013-16 (b) ATLAS-TOPQ-2015-02 (c) CMS-HIG-18-009



(d) CMS-TOP-17-018 (e) CMS-PAS-SMP-18-002 (f) ATLAS-EXOT-2016-16



(g) ATLAS-CONF-2018-027 (h) ATLAS-CONF-2018-034 (i) Combination

Figure 31: The profile likelihood ratio overlaid with the NLL for each of the fit results performed. These figures provide a good reference for how the systematic uncertainties affect the significance of the obtained fit results.

data in terms of its parameters. In truth, a mass scan would be interesting to study, but this is left for a future work.

Since the initial time of documenting the work in this chapter, an additional anomaly has been identified in the ATLAS Run 2 data which can be explained similarly to those mentioned in Section 6.3. This anomaly is seen in results reported by the ATLAS collaboration in di-lepton final states with a full hadronic jet veto [185], where the dominant SM process is the non-resonant production of W boson pairs. Figure 32 displays the di-lepton invariant mass in $e\mu$ events with a full hadronic jet veto after the application of the aforementioned corrections. Here, the QCD NNLO corrections to $q\bar{q} \to W^+W^-$ production [186, 187, 188, 189], QCD NLO corrections to non-resonant $gg \to W^+W^-$ [190] and EW NLO corrections [191] have been applied. A study of this distribution where the SM MC has been normalised to the data in the tail is shown in Reference [8]. The discrepancy re-emerges here with $m_{\ell\ell} < 100$ GeV, as has already been predicted and studied in this chapter and Chapter 5, showing similar features compared to the discrepancies in di-lepton final states with b-jets. The deviation seen in Figure 32 is not included in the result shown in Figure 30, and neither were those already identified with Run 1 data, as discussed in Chapter 4.

One might ask why such a prominent excess should show up in SM measurements instead of those that aim to discover new physics, particularly in multi-lepton final states (with or without b-jets). The best example is in searches for SUSY, where multiple leptons are often produced in cascade decays of charginos and neutralinos. The reason that the model discussed in this thesis is not sensitive to such searches is that a key signature in SUSY searches is the requirement of large values of $E_{\mathrm{T}}^{\mathrm{miss}}$, H_{T}, or the invariant mass of various multi-lepton systems. Due to the fact that the model discussed here produces its final state via off-shell cascaded decays, the produced leptons and b-jets tend to be far softer than those produced in SUSY processes, and therefore the predicted phase space tends to be quite different than what the typical SUSY searches require.

There are several issues that deserve some attention with regards to the BSM model considered here. With the single degree of freedom β_g^2, the BSM model is

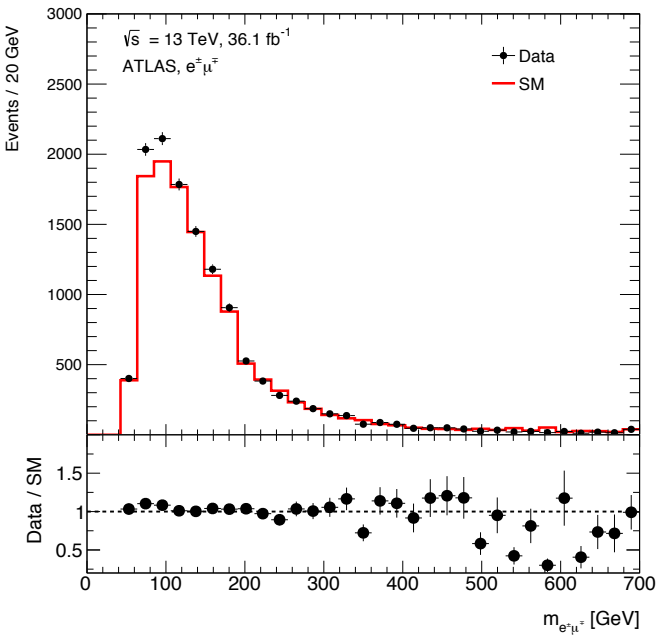


Figure 32: The di-lepton invariant mass spectrum in events with a full hadronic jet veto after the application of NNLO QCD and NLO EW corrections, as published in Reference [185] and analysed in Reference [8].

able to shed light on the magnitude of the excesses in the data. However, it is clear from Table 11 that in the fits there exists a tension between different final states. In particular, the fit results for the WZ measurements seem to exclusively require more BSM signal events than any other measurement. Interestingly enough, the ATLAS and CMS WZ results are consistent with one another. However, it cannot be said that the results are consistent with the rest of the ensemble, where for the WZ results β_g^2 is over a factor of 3 larger than for most of the other results. Apart from the WZ results, the rest of the ensemble exhibits a non-negligible spread around the combined best-fit mean value. Most notably, β_g^2 for the ATLAS Run 2 OS (spin correlation) measurement is almost a factor of 2 larger than the best-fit value. If one is to believe that the excesses in data truly are the result of new physics processes at the LHC, then what can be said is that the BSM model used here does not predict the correct relative mixture of events in terms of lepton and b-jet multiplicity. The simplified assumption of only one degree of freedom appears to be incapable of constructing a coherent prediction that is able to concurrently explain all of the excesses discussed. As mentioned above, one interesting avenue

to explore is multiple mass points of the BSM mode. Due to the sensitivity of the mass dependent BRs for the Higgs-like scalar S, a wide range of b-jet and lepton multiplicities could be explored.

Having said this, it is also possible that the simplified model as part of a 2HDM+S, as detailed in Chapter 5, predicts additional production mechanisms that can make for significant contributions to the results considered in this chapter. This has the potential to alleviate the tension in the β_g^2 values described above. In particular, the $A \to ZH$ process mode would certainly have a high acceptance in the WZ measurements considered in Section 6.3.7. The degree of contamination in the other measurements is yet to be determined. Studies have already been made on this process in Reference [139], and further developments can be anticipated in time.

It is clear that, if new physics is indeed responsible for the excesses discussed here, the simplified BSM model presented in this work is not an ideal candidate to explain them consistently. This statement is made stronger by the inability for the simplified model to describe the localised excess seen in the ATLAS Run 2 m_{T} distribution (Figure 25). In addition to a mass scan for S and H, it would be worthwhile looking into different assumptions for the decays of S. After all, allowing the S to have Higgs-like BRs was a convenient assumption to begin with, only because it assists in reducing the number of degrees of freedom in the model. An interesting alternative could be to study the heavy neutrino model introduced in Reference [4], since many of the kinematic distributions are similar to the model used here, and the extra degrees of freedom may make for a better fit (albeit probably not as statistically significant). Having said this, the BSM model used here does a remarkable job given that it requires only one degree of freedom.

Another interesting prospect would be to consider a VBF production mode for H. It was stated in Section 6.1 that a small effective H-W-W coupling, which has the advantage of enhancing the single top associated production mode, would suppress VBF. Throughout the study, however, it was determined that the single top associated production mode of H does not have a significant impact on the fit results for all of the considered measurements. The only non-negligible contributions of the tH production mechanism for the measurements considered here

are those that search for SS leptons in association with b-jets. Removing the tH contribution to the signal could have as much as a 20% effect on the best-fit values for β_g^2 in these measurements. However, they are not very sensitive measurements, and tH has a negligible impact on the most sensitive measurements studied in this study. Therefore, the assumption on the H-W-W coupling could indeed be relaxed, which would open up new possibilities of measurements that could be probed. Still more interesting would be to explore different models of the H and S bosons themselves, in terms of different possible spins and decay modes. This is left for future studies.

Should the 8.04σ significance of the combined fit stand the test of time and scientific criticism, it will present a challenge to our current understanding of physics at the LHC. Theoretically, it appears that the current set of tools used to describe SM processes is failing to do so, even in measurements of quantities as simple as the momenta of leptons. Whether or not this failure is due to BSM physics at the LHC remains to be seen. Any contamination of BSM physics relating to the Higgs sector would have profound impacts on the measurements of its mass and couplings. In any event, it is necessary to try and understand the data with as little bias as possible as we strive to solidify our understanding of the SM and beyond.

7

THE SEARCH FOR A HIGH MASS SCALAR DECAYING TO A Z BOSON PAIR IN THE FOUR LEPTON FINAL STATE

This chapter presents the search results for a heavy resonance using the ATLAS detector. The particular signature in which the search is made is for the hypothetical heavy resonance decaying to a Z boson pair with a 4ℓ final state. The search makes use of pp data collected by the ATLAS detector at $\sqrt{s} = 13$ TeV from 2015 until 2017. The total integrated luminosity used in this search is 79.8 fb^{-1}.

The goal in this search is to look for an excess in the $m_{4\ell}$ spectrum. If such an excess exists, its significance can be quantified in order to determine whether or not such a heavy resonance is produced at the LHC. The signal hypothesis used is that of a narrow-width Higgs-like boson, H, produced through ggF – much like the hypothetical heavy scalar discussed earlier in this thesis. The mass range considered for H is between 200 and 1400 GeV.

The specific search considered in this chapter is motivated by the 2016 result for the same search [96], which revealed two mild excesses in the $m_{4\ell}$ spectrum. Each excess had a local significance of 3.6 σ. The peaks were located at around 240 GeV and 700 GeV. The peak at 700 GeV was not present in the $2\ell 2\nu$ channel, and was therefore considered to be excluded at 95% CL. The $2\ell 2\nu$ result, however, had no effect on the 240 GeV peak due to the fact that the search is not sensitive below masses of 300 GeV. The primary goal in this chapter is to check whether the excesses diminish or re-appear when one includes the 2017 data. Note that this will only be done in the context of ggF-like signatures (that is, in the $4e$, 4μ and $2e2\mu$ channels), in which the excesses primarily existed in 2016.

7.1 DATA AND SIMULATED SAMPLES

The data considered for the search in this chapter includes pp collision data collected in the ATLAS detector from 2015 to 2017. Each year contained different pile-up conditions. The $\langle\mu\rangle$ changed from 13.6 in 2015 to 24.9 in 2016, and then to 37.8 in 2017. In terms of integrated luminosity, the total raw data collected over the years was 86.76 fb^{-1}, corresponding to 3.86 fb^{-1} in 2015, 36.0 fb^{-1} in 2016 and 46.9 fb^{-1} in 2017. However, the data are subjected to quality control requirements which ensure that all components of the detector were fully operational at the time of recording. After applying a "good-runs list" that ensures the quality of the data, the resulting data used in the search amount to 36.2 fb^{-1} for 2015 and 2016, and 43.6 fb^{-1} in 2017. The entire data set is processed and reconstructed in terms of physics objects by proprietary ATLAS analysis software, version 21.2. This is in contrast to the 2016 result [96], which made use of version 20.7, and therefore a number of improvements can be expected compared to the older results.

Two different methods are used for reconstructing MC simulations in this search, namely `mc16a` and `mc16d`. The former method was applied to the 2015 and 2016 data sets, whereas the latter was applied to the 2017 data set. The main difference between the two methods is the way in which they deal with simulating pile-up. In both cases, reconstruction was performed using the ATLAS analysis software, version 21.2. The proprietary ATLAS detector simulation was also used to simulate the response of the ATLAS detector. Depending on the sample in question, either a full simulation (FS) or fast simulation (also know as AF2) was used [192]. A list of all the MC samples used in this study can be seen in Table 12.

7.2 RECONSTRUCTION OF PHYSICS OBJECTS

Since the search in this chapter is concerned only with the detection of four leptons through a direct production mechanism, the most important objects necessary to be reconstructed are electrons and muons. This section briefly discusses how these

Process	Modelling	Generator(s)	PDF	Simulation
		Backgrounds		
$qq \to ZZ$	SM	Sherpa	NNPDF30NNLO	FS
$gg \to ZZ$	SM (without Higgs)	Sherpa	NNPDF30NNLO	AF2
$pp \to VVV$	SM	Sherpa	NNPDF30NNLO	FS
$pp \to ttV$	SM	aMC@NLO + Pythia 8	NNPDF23LO	FS
$qq \to Z$+jets	SM	Sherpa	NNPDF30NNLO	FS
$pp \to tt$	SM	POWHEG + Pythia 8	NNPDF23LO	FS
		Signals		
$gg \to H \to 4\ell$	m_H=200 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2
	m_H=300 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	FS+AF2
	m_H=400 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2
	m_H=500 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2
	m_H=600 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2
	m_H=700 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	FS+AF2
	m_H=800 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2
	m_H=900 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2
	m_H=1000 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2
	m_H=1200 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2
	m_H=1400 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2
	m_H=1600 GeV, NWA	aMC@NLO + Pythia 8	NNPDF23LO	AF2

Table 12: A list of the MC samples needed for the search for a high mass four-lepton resonance produced through ggF. Both the signals and backgrounds are shown here. NWA refers to signals generated with the narrow-width approximation.

leptons are reconstructed in the ATLAS detector, which was described in Chapter 3. In addition to this, some key improvements that were introduced to version 21.2 of the ATLAS analysis software are discussed. Note that in the absence of a search for VBF-like signatures, jets are not considered for reconstruction in the analysis presented here.

Electrons are reconstructed from clusters of calorimeter cells using a dynamic topological clustering technique. This technique improves the energy resolution of the detected electron, especially for electrons that radiate a photon via brehmsstrahlung processes [193]. The electrons are constructed as "super-clusters" that are associated with tracks from the ID. The complete matching of the clusters and tracks is done with a Gaussian-sum filter, which accounts for potential radiation. In addition to this, extra information coming from both the EM shower shape and track hit multiplicity is used to reject backgrounds (i.e. fake electrons). With these additional degrees of freedom, a series of working points are defined for the reconstruction of electrons. In this analysis the "loose" working point is used, which has a reconstruction efficiency of 90% at an electron p_{T} of 20 GeV, and 96% for electrons with a p_{T} greater than 60 GeV.

Different types of muons are reconstructed using information from the ID and the MS [194]. If a matching track is found in both the ID and the MS, this is classified as a *combined* muon. For the central region of the detector where the MS has limited coverage (that is, for $|\eta| < 0.1$) only the ID is used for the determination of the muon momentum; these are called *segment-tagged* muons. In this region, muons that can be correlated with calorimetric energy consistent with a minimum ionising particle are known as *calorimeter-tagged* muons. In forward regions where the ID is limited (that is, $2.5 < |\eta| < 2.7$) the momentum determination is made only using MS information. These are known as *stand-alone* muons. The analysis in this chapter considers all of these different types of muons. Improvements in version 21.2 of the reconstruction software helps in reducing reconstructing muons falsely from fake tracks, due to improving the requirements on the number of hits in the different trackers. Similarly for electrons, different working points are

defined for muons. The analysis presented here uses the "loose" working point, which has a reconstruction efficiency of over 98%.

7.3 EVENT SELECTION

In this analysis, events are classified into three different channels based on the flavours of the selected leptons. These channels are 4μ, $4e$ and $2\mu 2e$, and are assigned based on which triggers are activated in the event. The trigger thresholds varied between 2015 and 2017, but overall their efficiency is about 98%. For each event, there should also be at least one vertex with two associated tracks having $p_{\mathrm{T}} > 500$ MeV. Previously, in version 20.7 of the ATLAS analysis software, this number was 400 MeV. Of the reconstructed vertices, the *primary* vertex is identified as the one with the highest sum of track p_{T}^2.

Leptons are then sorted into quadruplets based on all the available SFOS lepton pairs in the event. Leptons can only be assigned to a quadruplet if they have sufficient p_{T} and fall within the geometric acceptance range. For electrons, the p_{T} should be greater than 7 GeV and they should lie within $|\eta| < 2.47$. Muons should have a p_{T} greater than 5 GeV and be within $|\eta| < 2.7$. In each quadruplet, the three highest p_{T} leptons should have at least 20 GeV, 15 GeV and 10 GeV of p_{T}, respectively. Each quadruplet can have at maximum only one segment-tagged, calorimeter-tagged or stand-alone muon. If more than one quadruplet exists per channel, only one is selected by virtue that the SFOS pairs in the quadruplet are the two that have the closest mass to the Z boson mass. The masses of these SFOS systems are labelled as m_{12} and m_{34} for the closest and second closest pairs, respectively. The selected quadruplet is required to satisfy 50 GeV $< m_{12} <$ 106 GeV and $m_{34} <$ 115 GeV. The lower bound on m_{34} is 12 GeV for $m_{4\ell} <$ 140 GeV and 50 GeV for $m_{4\ell} >$ 190 GeV. For any value of $m_{4\ell}$ in between 140 and 150 GeV, the bound grows according to a linear slope.

Several additional cuts are placed on the leptons to suppress fake or non-prompt backgrounds. Firstly, the leptons are all required to be separated from each other by at least $\Delta R = 0.1$ if they are same-flavour, and $\Delta R = 0.2$ if they are different-flavour.

For the $4e$ and 4μ channels, quadruplets are removed if any two SFOS leptons have a mass of less than 5 GeV. This suppresses the production of leptons through the decay of the J/ψ meson. If after all these requirements, multiple quadruplets still exist, then the selected quadruplet is the one with the highest expected signal rate, i. e. in the following order: 4μ, $2\mu 2e$ and $4e$.

Backgrounds which involve non-resonant and/or the associated production of four leptons, such as the $Z+$ jets and $t\bar{t}$ processes, are suppressed using impact parameter and isolation requirements. The transverse impact parameter significance is defined as $|d_0|/\sigma_{d_0}$, where the denominator is the standard deviation for measurements of the transverse impact parameter d_0. This significance is required to be less than 3 for muons and 5 for electrons. Additionally, all leptons are required to be associated with the same originating vertex. These requirements help ensure that the leptons originate from a single object in the hard interaction.

Leptons are required to be isolated using both track-based and calorimeter-based discriminants. The track-based discriminant takes into account the scalar p_{T} sum of all tracks in a cone of width $\Delta R = 0.3$ for muons and 0.2 for electrons (excluding the lepton itself). The ratio of this sum to the lepton p_{T} should be less than 0.15. The tracks that enter the sum are also required to originate from the primary vertex in order to mitigate pile-up contributions. If the leptons are considered to be high in p_{T} (that is, at least 33 GeV for muons and 50 GeV for electrons), the cone size is reduced to 10 GeV divided by the p_{T} of the lepton in question. The calorimeter-based discriminant similarly adds up the cluster E_{T} values used to reconstruct jets within a cone of width $\Delta R = 0.2$ around the barycentre of the candidate lepton. This E_{T} sum divided by the p_{T} of the lepton is required to be less than 0.3 for muons and 0.2 for electrons. Any known contributions to the E_{T} sum that come from the candidate lepton are excluded, and pile-up contributions are subtracted on an event-by-event basis.

Once the four lepton quadruplets are selected, the leptons' four momenta can be corrected for final state radiation (FSR) using information from the calorimeters. FSR photons are searched for both colinearly and non-colinearly (the former only applying to muons) with a maximum of one FSR photon per event. In the event

that such a photon is found, the lepton's four momentum is re-computed by making a fit to the Z mass peak line-shape, which is modelled as a Breit-Wigner peak with a Gaussian curve assigned to each lepton to account for detector resolution effects.

The event selection described above is designed to maximise the acceptance for ggF-like signals while suppressing the four lepton backgrounds. For the ggF MC samples introduced in Section 7.1, the acceptance grows with the mass of the hypothetical scalar. The acceptances in each channel are as follows:

- 4μ channel: from 52% at $m_H = 200$ GeV to 62% at $m_H = 1600$ GeV,
- $2\mu 2e$ channel: from 41% at $m_H = 200$ GeV to 58% at $m_H = 1600$ GeV,
- $4e$ channel: from 34% at $m_H = 200$ GeV to 55% at $m_H = 1600$ GeV.

7.4 BACKGROUND ESTIMATION

For the ggF-enriched categories considered in this chapter, the dominant backgrounds come from the non-resonant production of Z boson pairs. The dominant background, comprising of 86% of the total background composition, is the quark initiated $q\bar{q} \to ZZ$ process. The next most prominent background is the gluon initiated equivalent $gg \to ZZ$ process, which makes up 10% of the total background composition. The sub-dominant backgrounds each contribute to the total composition on the level of approximately 1% each, these being Z+jets, $t\bar{t}$ and the EW production of two Z bosons. The dominant non-resonant backgrounds are modelled using MC simulations, whereas the others are estimated using a combination of data-driven techniques and MC simulations.

The properties of the background processes are significantly sensitive to the flavour of the sub-leading SFOS lepton pair in the quadruplet. Therefore, the analysis on the backgrounds is performed separately for $\ell\ell\mu\mu$ and $\ell\ell ee$ channels. The main reason for this is that background muons tend to come from heavy-flavour jet decays in association with Z bosons, or the decays of top quarks. Background

electrons, on the other hand, tend to be the result of the mis-identification of light-flavour jets.

For the $\ell\ell\mu\mu$ channels, four CRs are defined in the data in order to fix the normalisation of the Z+jets and $t\bar{t}$ backgrounds. All of these CRs relax the requirement that the four selected leptons should originate from a common vertex. Then the different CRs are constructed by inverting and/or relaxing the isolation and impact parameter requirements. Additional CRs are also made by forcing the sub-leading pair of muons to be same-sign, or by requiring a different-flavour leading pair of leptons (this is targeted at constraining the $t\bar{t}$ background). The normalisation factors are fixed by fitting the leading lepton pair's invariant mass simultaneously in the CRs. Thereafter, the normalisation of the sample in the SR is fixed by an extrapolation of the normalisation in the CRs, using the acceptance determined in the CRs. This is known as a transfer factor, and is determined using MC simulations.

Only one CR is used for the $\ell\ell ee$ channels. This CR, which is referred to as the $3\ell + X$ CR, requires the sub-leading electron pair to be same-sign. The lowest E_{T} electron also has its identification and isolation criteria relaxed, such that it can be a mis-identified light-flavour jet or photon, or an electron coming from a heavy-flavour decay. The heavy-flavour component of this CR comes from MC simulations, whereas the others come from using the ${}_s\mathcal{P}lot$ method [195] on information from the ID. Similar to the $\ell\ell\mu\mu$ channels, transfer factors are determined from MC simulations.

The other sub-dominant backgrounds, including tri-boson and $t\bar{t}V$ processes, are estimated from MC simulations, along with the WZ process in the $\ell\ell\mu\mu$ channels. However, in the $\ell\ell ee$ channels, the WZ process is included as a component in the data-driven calculations.

7.5 SIGNAL AND BACKGROUND PARAMETERISTAION

The search results are obtained by performing fits to the data using signal and background models. These fits are performed as a function of the four-lepton invariant mass. The MC simulated samples and their accompanying estimates from

the data (described above) are used to construct analytic functions for the fitting process. These functions and their parameterisations are discussed below for the signal and background hypotheses.

7.5.1 *Signal model*

In this chapter only a narrow-width scalar is assumed for the search, and therefore the width in the signal $m_{4\ell}$ distribution is completely due to detector resolution effects. These effects are typically modelled using the sum of a Crystal Ball (CB) function [196] and a Gaussian. The corresponding probability density function can be written down as:

$$P_s(m_{4\ell}) = f_C \times C\left(m_{4\ell};\ \mu,\ \sigma_C,\ \alpha_C,\ n_C\right) + (1 - f_C) \times G\left(m_{4\ell};\ \mu,\ \sigma_G\right), \qquad (7.1)$$

where C denotes the function and variables relevant to the CB, and G denotes those relevant to the Gaussian. Both functions share the same mean value μ, but differ in their widths (σ_C and σ_G). The additional parameters in the CB function (α_C and n_C) control the shape of the exponential tail. The relative normalisation of the two functions is controlled by the parameter f_C.

All the parameters in Equation (7.1) (with the exception of n_C) are fit to the MC simulations for a narrow-width heavy scalar at the various mass points listed in Table 12. The parameter n_C is chosen to be constant for all the mass points, to improve the stability of the fit. As a function of the resonance mass, the best-fit parameter points are then fit with polynomials, such that the signal model can be determined for an arbitrary value of m_H in between those that have been simulated. Using the analytic function induces a bias when extracting signal yields from the data; this bias was determined to be less than 1.5%. In addition, the use of polynomial fits for the interpolation introduces a bias of about 1%.

Examples of the fits made for the signal parameterisation are shown in Figure 33. The mass point chosen for illustration is $m_H = 300$ GeV. The different components have been split up and displayed, as well as the total fit function. In all mass

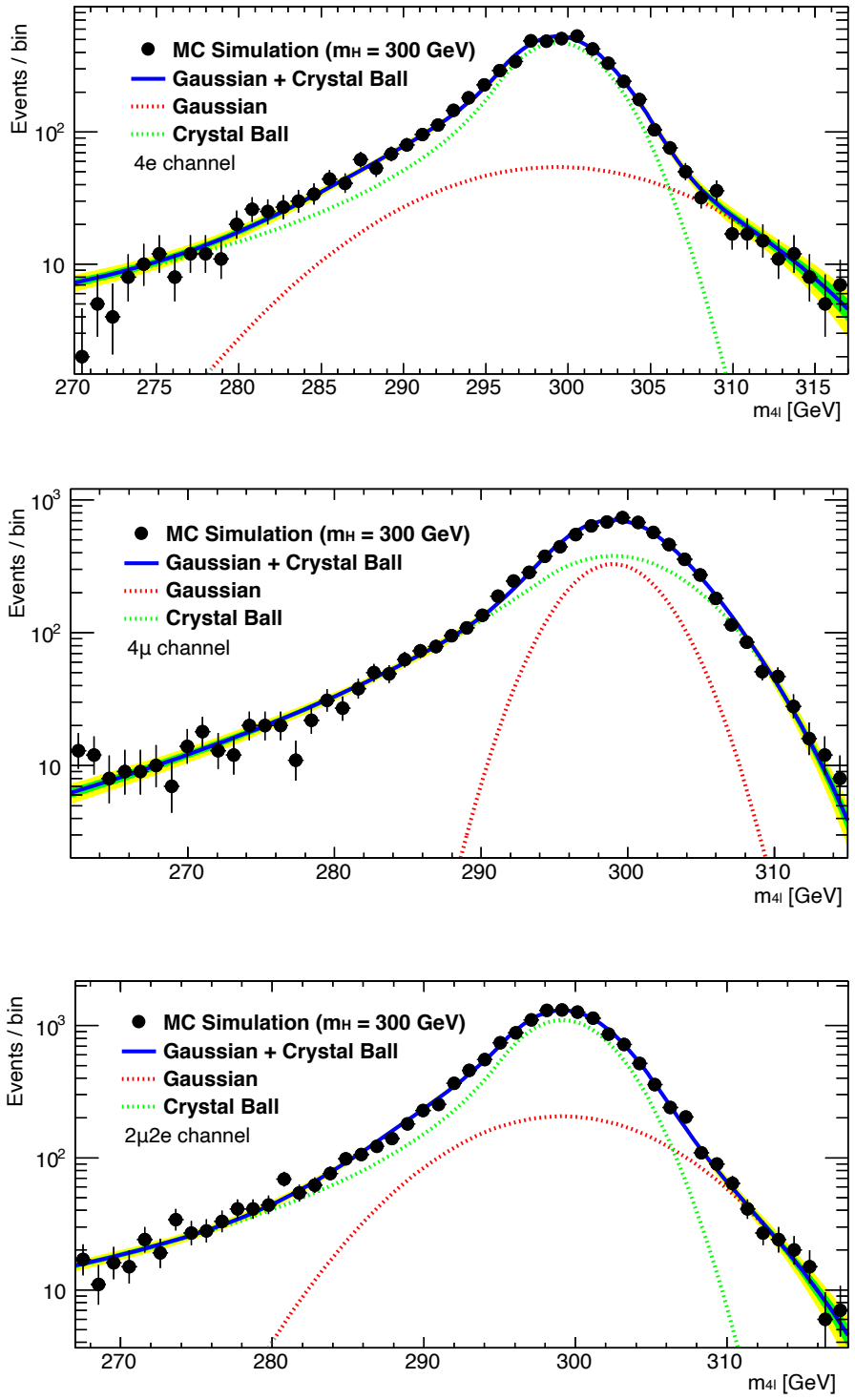


Figure 33: A demonstration of the performance of the analytic fit function for the signal in Equation (7.1). The overall fit is shown compared with the MC simulation, and the different components of the fit. An uncertainty band on the overall fit is shown, with the green band representing a 1σ deviation and the yellow representing 2σ. All three different channels are shown, with $4e$ on the top, 4μ in the middle and $2\mu 2e$ on the bottom.

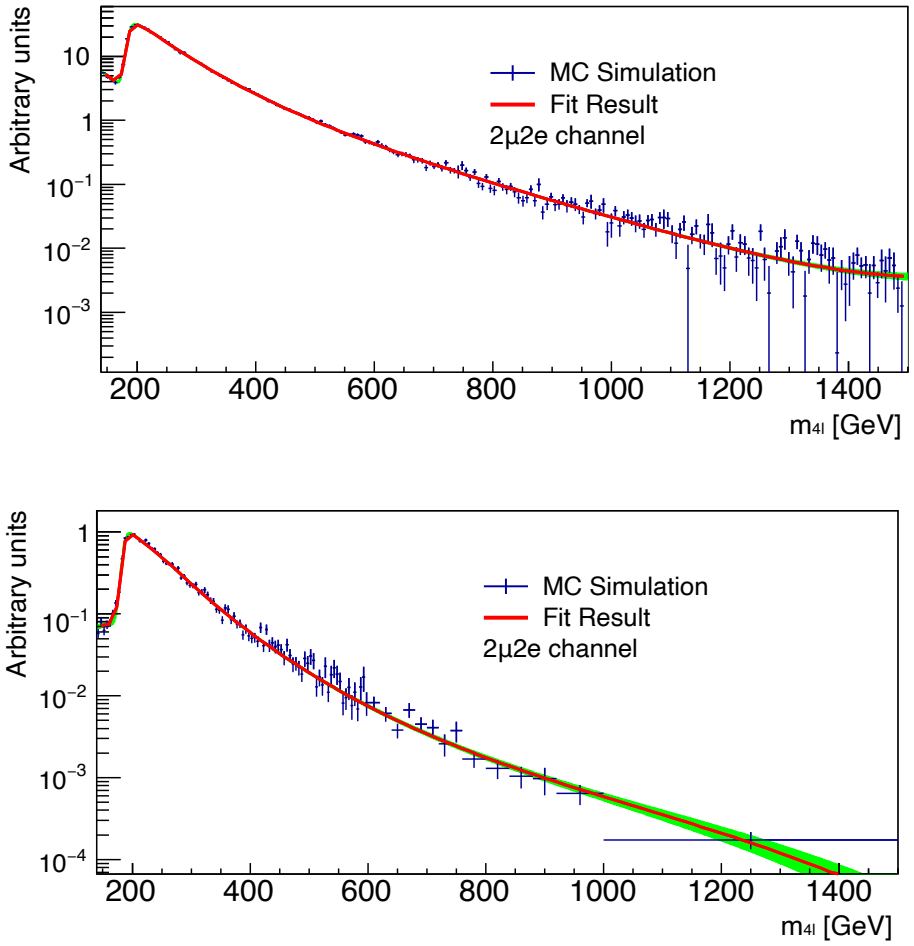


Figure 34: An illustration of the analytic functions used to fit the continuum backgrounds in the high-mass $H \to ZZ$ search. Here the $qq \to ZZ$ (top) and $gg \to ZZ$ (bottom) processes are shown in the $2\mu 2e$ channel, as an example. The green band represents the uncertainty in the fit parameters.

points, the tail of the distribution for the 4μ channel is accurately described by the CB, whereas for the channels that include electrons, the Gaussian assists in describing the tails. For all mass points, it was found that the analytic function in Equation (7.1) describes the MC simulation well, and the interpolation polynomial fits, which give values for the fit parameters as a function of the signal mass, vary smoothly between the mass points.

7.5.2 *Background model*

The dominant continuum ZZ background predictions (i.e. $qq \to ZZ$ and $gg \to ZZ$) are estimated with MC simulations, on which an analytic function is fit. The analytic function used is a piece-wise function of $m_{4\ell}$, and can be written as:

$$f^p(m_{4\ell}) = \begin{cases} [f_1(m_{4\ell}) + f_2(m_{4\ell})] \times C_0^p & m_{4\ell} \leq m_0^p \\ f_3^p(m_{4\ell}) & m_{4\ell} > m_0^p, \end{cases} \tag{7.2}$$

where $p \in \{qq, gg\}$ denotes whether the function applies to the quark or gluon initiated background shape, respectively. The points at which the functions change is controlled by m_0^p, which are chosen by ensuring that the two functions are smooth at the transition. The values are approximately $m_0^{qq} = 240$ GeV and $m_0^{gg} = 265$ GeV. The normalisation parameters, C_0^p, ensure that the functions are also continuous at m_0^p, and are defined as:

$$C_0^p = \frac{f_3^p(m_0^p)}{f_1(m_0^p) + f_2(m_0^p)}. \tag{7.3}$$

The different parts of Equation (7.2) each explain different underlying physical processes in the SM production of Z boson pairs. The equations describing their shapes can be written in terms of their parameters as follows:

$$f_1(m_{4\ell};\ a_1,\ a_2) = \mathrm{Exp}\left(a_1 + a_2 \cdot m_{4\ell}\right), \tag{7.4}$$

$$f_2(m_{4\ell};\ b_1,\ b_2,\ b_3) = \frac{1 + \mathrm{Erf}\left(\frac{m_{4\ell}-b_1}{b_2}\right)}{2} \times \frac{1}{1 + \mathrm{Exp}\left(\frac{m_{4\ell}-b_1}{b_3}\right)}, \tag{7.5}$$

$$f_3^{qq}(m_{4\ell};\ c_1,\ldots,c_5) = \mathrm{Exp}\left(c_1 + c_2 \cdot m_{4\ell} + c_3 \cdot m_{4\ell}^2 + c_4 \cdot m_{4\ell}^3 + c_5 \cdot m_{4\ell}^4\right), \tag{7.6}$$

$$f_3^{gg}(m_{4\ell};\ c_1,\ldots,\ c_4) = \mathrm{Exp}\left(c_1 + c_2 \cdot m_{4\ell} + c_3 \cdot m_{4\ell}^2 + c_4 \cdot m_{4\ell}^{2.7}\right). \tag{7.7}$$

The functions covering the low mass part of the spectrum are designed to explain the $m_{4\ell}$ shape below the kinematic threshold at a mass of $2m_Z$. The first part, f_1, provides an exponentially decaying shape relevant when one of the Z bosons is

off-shell. The second part, f_2, becomes relevant at the kinematic threshold and describes the sharp increase in the spectrum. The relatively generic exponential f_3^p functions describe the tail of the distributions. The parameters a_i, b_i and c_i are used to fit the functions to the MC simulations for the processes. An example of this can be seen in Figure 34, where the $2\mu 2e$ channel has been chosen for illustration.

7.6 SYSTEMATIC UNCERTAINTIES

7.6.1 *Experimental systematic uncertainties*

The signal and background predictions used in this work are affected by a variety of sources of experimental systematic uncertainty. These systematic uncertainties are treated as NPs and are profiled in the statistical machinery used for the final results. The dominant experimental systematic uncertainties arise from the energy/-momentum scales and reconstruction and identification efficiencies of the leptons and jets. Several sub-dominant systematic uncertainties are considered and will be discussed below.

The calculation of the experimental systematic uncertainties is done as follows. First, the $m_{4\ell}$ distribution is considered using the standard weights and a nominal configuration of the NPs – the *nominal distribution*. Then, for each NP, the $m_{4\ell}$ distribution is made using a (1σ) variation of the NP in question – the *modified distribution*. The variation is considered for the "up" and "down" configuration of the NP ($\pm 1\sigma$). For NPs that only affect the normalisation of the distribution, the relative change of the total event yield from the modified distribution is compared to that of the nominal distribution. For NPs that also affect the shape of the $m_{4\ell}$ distribution, the relative change to the mean and RMS of the modified distribution, with respect to the nominal distribution, is also calculated. These shape variations are used to construct an analytical variation to the distributions in the fitting procedure. The logical flow of the procedure in calculating the effects of systematic uncertainties is shown in Figure 35.

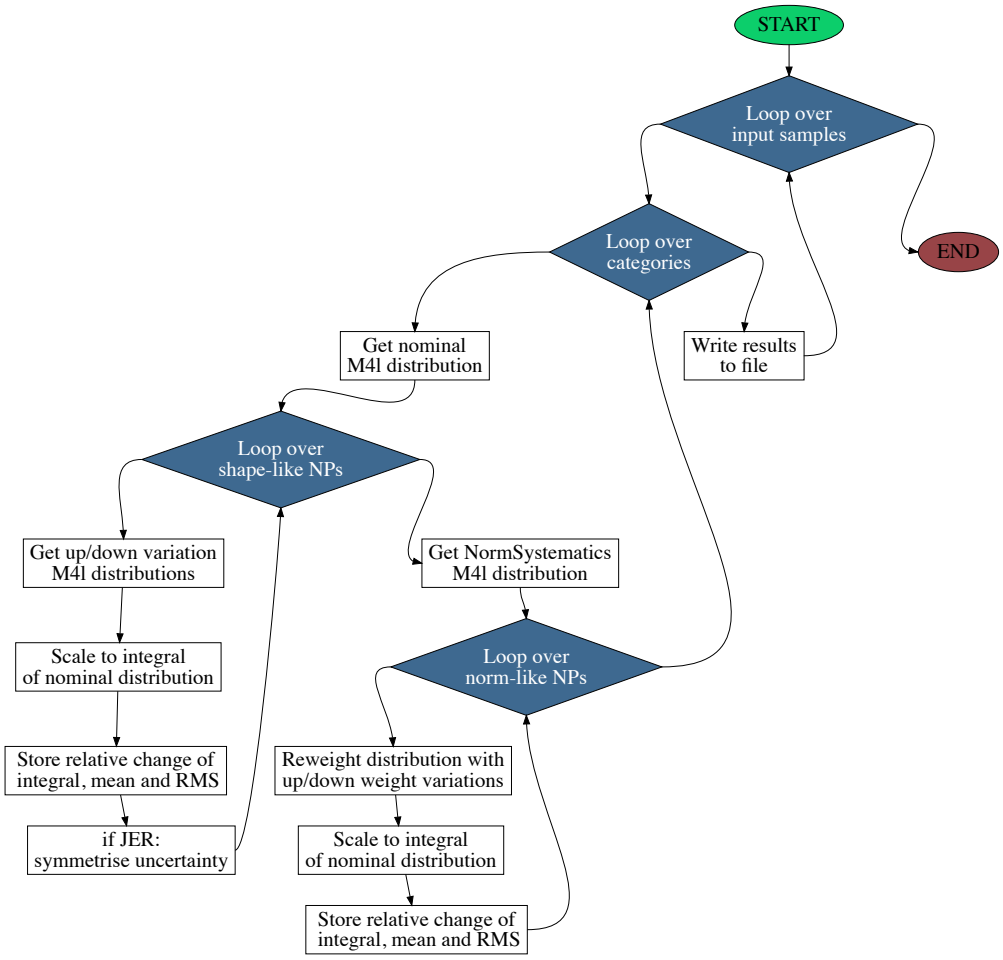


Figure 35: A simple flowchart of the logic followed when calculating the effect of systematic variations of NPs on the $m_{4\ell}$ distribution. For technical reasons, the calculation is performed separately for NPs that only affect the normalisation of the $m_{4\ell}$ spectrum ("norm-like NPs") and those that can also affect the shape ("shape-like NPs").

The systematic uncertainties are calculated for each of the three categories considered in this chapter: $4e$, 4μ and $2e2\mu$. In the calculation, the leptons in the quadruplet are treated as completely correlated, so as to obtain conservative systematic uncertainties. That is, if the nominal and modified weights of the i-th lepton are w_i and $w_i + \delta w_i$, respectively, then the nominal and modified weights of the lepton quadruplet are $w_1 \cdot w_2 \cdot w_3 \cdot w_4$ and $(w_1 + \delta w_1) \cdot (w_2 + \delta w_2) \cdot (w_3 + \delta w_3) \cdot (w_4 + \delta w_4)$, respectively. The systematic uncertainties are calculated using the recommendations from the ATLAS Combined Performance groups. The standard set of systematics is made of 96 NPs, each with an "up" and "down" variation, with the exception of the jet energy resolution NP, which is assumed to have a perfectly symmetric systematic uncertainty. A list of all the NPs used in the analysis is shown in Table 13.

The different effects that the systematic uncertainties address are:

- **Electron reconstruction and identification**: for reconstruction, a single NP is used, `EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR`. The identification efficiency uncertainty is estimated using `EL_EFF_ID_CorrUncertaintyNP[0-15]` and `EL_EFF_ID_SIMPLIFIED_UncorrUncertaintyNP[0-17]`.

- **Electron energy scale and resolution**: effects on the electron energy scale arising from correlated effects are evaluated with `EG_SCALE_ALLCORR`, whereas those arising from detector components are evaluated with `EG_SCALE_E4SCINTILLATOR`, `EG_SCALE_LARTEMPERATURE_EXTRA2015PRE`, `EG_SCALE_LARTEMPERATURE_EXTRA2016PRE` and `EG_SCALE_LARCALIB_EXTRA2015PRE`.

- **Muon reconstruction and identification**: for muons, the reconstruction and identification systematics are different for high and low p_{T} muons, that is, muons with greater or smaller p_{T} than 15 GeV. The reconstruction systematic is calculated with `MUON_EFF_RECO_STAT`, `MUON_EFF_RECO_STAT_LOWPT`, `MUON_EFF_RECO_SYS` and `MUON_EFF_RECO_SYS_LOWPT`.

- **Muon momentum scale and resolution**: the momentum scale and resolution are determined from tracking properties that are different in the different

Normalisation NPs	Shape NPs
Electrons	
EL_EFF_ID_CorrUncertaintyNP[0-15]	EG_RESOLUTION_ALL
EL_EFF_ID_SIMPLIFIED_UncorrUncertaintyNP[0-17]	EG_SCALE_ALLCORR
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	EG_SCALE_E4SCINTILLATOR
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	EG_SCALE_LARCALIB_EXTRA2015PRE
	EG_SCALE_LARTEMPERATURE_EXTRA2015PRE
	EG_SCALE_LARTEMPERATURE_EXTRA2016PRE
Muons	
MUON_EFF_ISO_STAT	MUON_ID
MUON_EFF_ISO_SYS	MUON_MS
MUON_EFF_RECO_STAT	MUON_SAGITTA_RESBIAS
MUON_EFF_RECO_STAT_LOWPT	MUON_SAGITTA_RHO
MUON_EFF_RECO_SYS	MUON_SCALE
MUON_EFF_RECO_SYS_LOWPT	
MUON_EFF_TTVA_STAT	
MUON_EFF_TTVA_SYS	
Jets	
	JET_BJES_Response
	JET_EffectiveNP_[1-7]
	JET_EffectiveNP_8restTerm
	JET_EtaIntercalibration_Modelling
	JET_EtaIntercalibration_NonClosure_highE
	JET_EtaIntercalibration_NonClosure_negEta
	JET_EtaIntercalibration_NonClosure_posEta
	JET_EtaIntercalibration_TotalStat
	JET_Flavor_Composition
	JET_Flavor_Response
	JET_JER_SINGLE_NP
	JET_JvtEfficiency
	JET_Pileup_OffsetMu
	JET_Pileup_OffsetNPV
	JET_Pileup_PtTerm
	JET_Pileup_RhoTopology
	JET_PunchThrough_MC16
	JET_SingleParticle_HighPt
Flavour tagging	
FT_EFF_Eigen_B_[0-2]	
FT_EFF_Eigen_C_[0-2]	
FT_EFF_Eigen_Light_[0-4]	
FT_EFF_extrapolation	
FT_EFF_extrapolation_from_charm	
Other	
PRW_DATASF	

Table 13: A list of the NPs used for systematic uncertainties in this analysis. The NPs have been separated by whether they only affect the normalisation (left), or if they affect the shape (right) of the $m_{4\ell}$ distribution. They are further subdivided into the reconstructed objects that they affect.

tracking systems. Overall, the scale is evaluated with MUON_SCALE. The ID systematic comes from MUON_ID and the MS part comes from MUON_MS. The tracking NPs MUON_SAGITTA_RESBIAS and MUON_SAGITTA_RHO are also used.

- **Muon track-to-vertex-association (TTVA)**: there exists an independent uncertainty on the performance of matching a muon track to a vertex. This uncertainty is evaluated with MUON_EFF_TTVA_STAT and MUON_EFF_TTVA_SYS.

- **Isolation**: electron isolation systematics are evaluated with the NP EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR, whereas muons use the NPs MUON_EFF_ISO_STAT and MUON_EFF_ISO_SYS.

- **Jet energy scale and resolution**: the jet energy resolution is evaluated with one NP, JET_JER_SINGLE_NP. All other systematic effects from jet reconstruction and scaling are evaluated with the large number of NPs listed in Table 13.

- **Flavour tagging**: these sub-dominant effects don't have a significant impact on the analysis, yet they are still used in the evaluation of the systematic uncertainties. The effects of tagging b-, c- and light jets are studied using FT_EFF_Eigen_B_[0-2], FT_EFF_Eigen_C_[0-2] and FT_EFF_Eigen_Light_[0-4], respectively. Extrapolation uncertainties are propagated with FT_EFF_extrapolation and FT_EFF_extrapolation_from_charm.

- **Pile-up**: the effects of pile-up have some effect on the reconstruction of jets, but the dominant uncertainty comes from PRW_DATASF. This comes from the scale factor that re-scales the $\langle\mu\rangle$ value from its nominal re-scaling, a number determined from data to MC comparisons. The uncertainty is often one sided, since the overall scaling is never considered to be bigger than 1.0.

For the continuum background samples and one signal sample, the systematic variations as a function of the NPs considered can be seen in Figure 36, Figure 37 and Figure 38, respectively. These figures show how the different NPs affect the normalisation of the $m_{4\ell}$ distribution, and these values are used as systematic uncertainties for the final fit results.

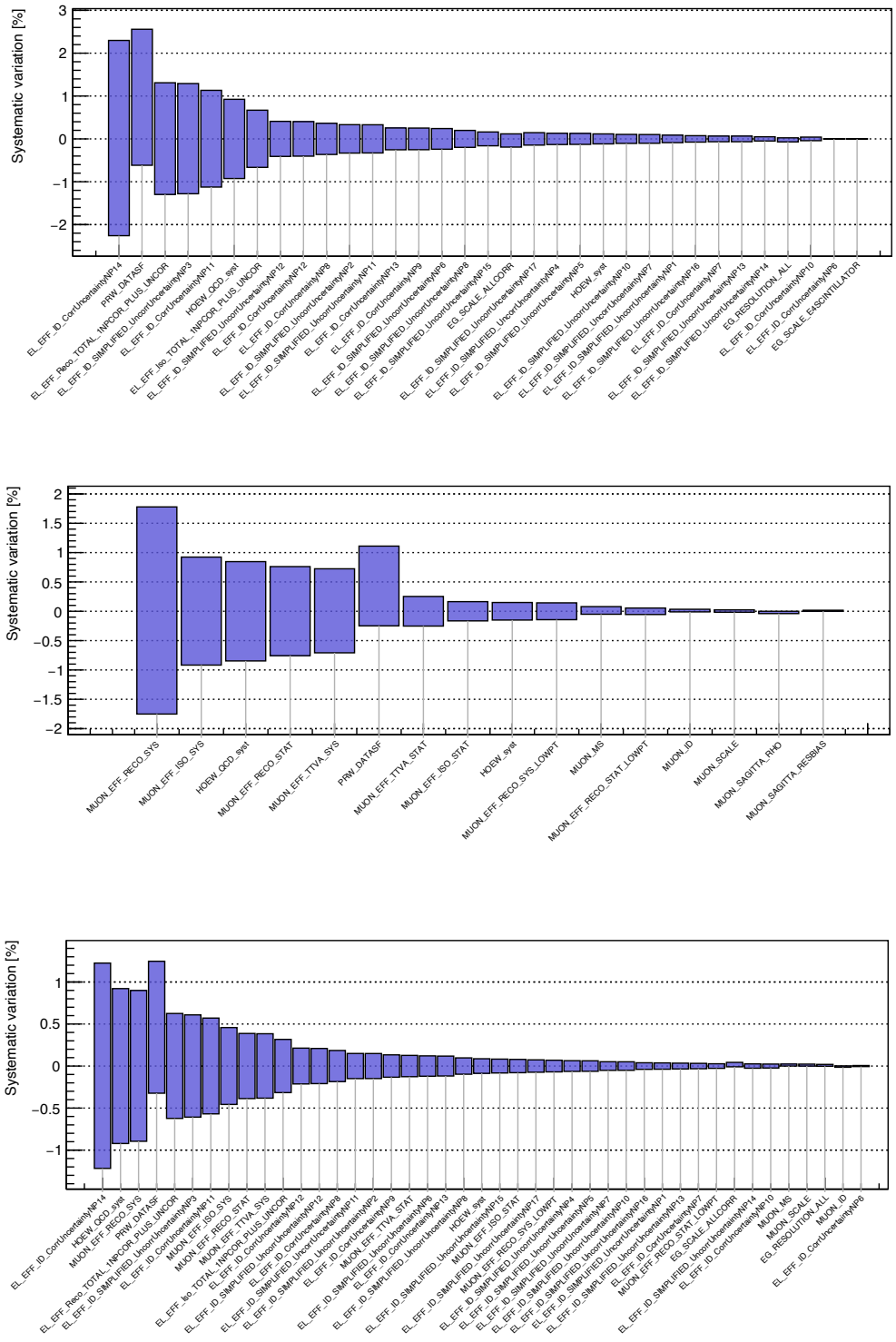


Figure 36: Experimental systematic uncertainties in the normalisation of the $m_{4\ell}$ spectrum for the $qq \to ZZ$ continuum background sample, as a function of the NPs considered. Each of the different categories are shown with $4e$ on the top, 4μ in the middle and $2\mu 2e$ on the bottom. Any NPs with an effect of less than 0.01% have been removed.

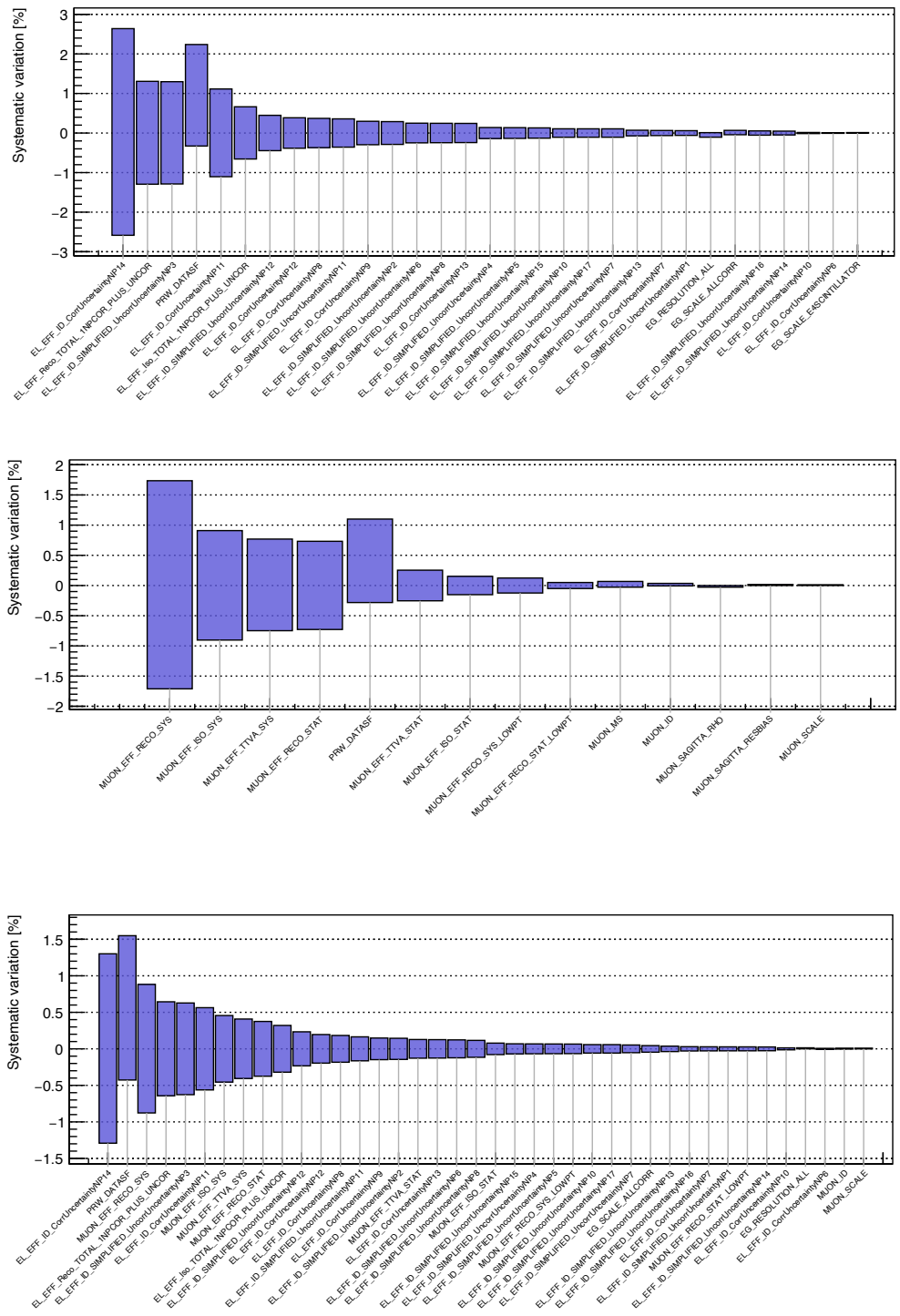


Figure 37: Experimental systematic uncertainties in the normalisation of the $m_{4\ell}$ spectrum for the $gg \to ZZ$ continuum background sample, as a function of the NPs considered. Each of the different categories are shown with $4e$ on the top, 4μ in the middle and $2\mu 2e$ on the bottom. Any NPs with an effect of less than 0.01% have been removed.

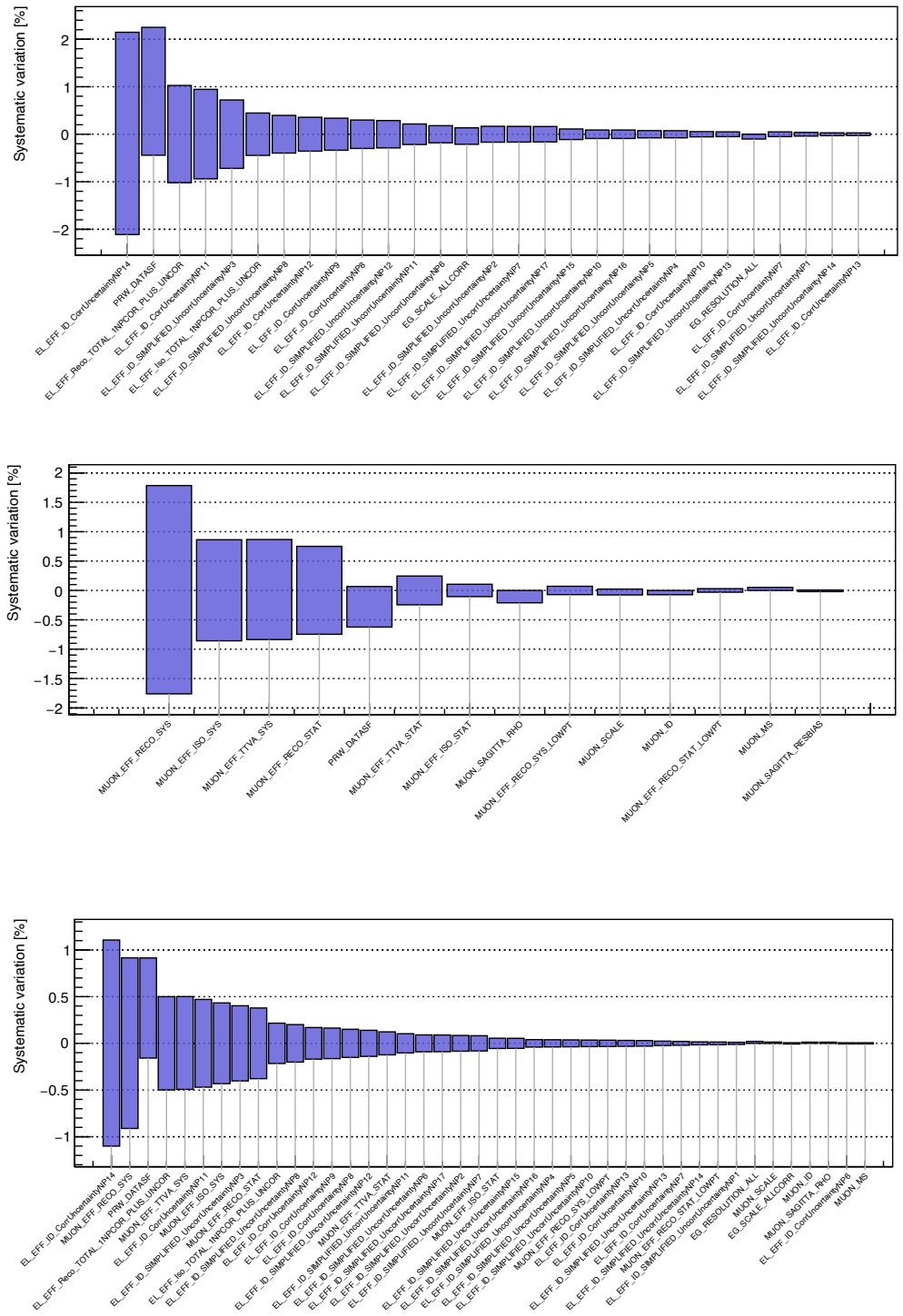


Figure 38: Experimental systematic uncertainties in the normalisation of the $m_{4\ell}$ spectrum for the ggF signal sample with $m_H = 300$ GeV, as a function of the NPs considered. Each of the different categories are shown with $4e$ on the top, 4μ in the middle and $2\mu 2e$ on the bottom. Any NPs with an effect of less than 0.01% have been removed.

In addition to the systematic uncertainties introduced above, there also exist systematic uncertainties related to the total integrated luminosity and trigger inefficiencies. The uncertainty in the combined 2015-2017 integrated luminosity is 2.0%. A systematic uncertainty of 0.5% is considered for the muon trigger efficiency differences in the data and the simulation.

7.6.2 *Theoretical systematic uncertainties*

In this analysis, theoretical systematic uncertainties are considered for the dominant continuum ZZ backgrounds and the signal. The sources of these uncertainties include variations of the central renormalisation and factorisation scales, PDF sets in association with their associated eigenvectors, and configurations of the parton shower modelling. In all cases, the resulting systematic uncertainty is calculated as the change in the selection acceptance while varying certain parameters related to the source of the uncertainty. In this section, the methods of estimating these uncertainties are discussed along with a summary of their results.

The dominant $qq \to ZZ$ background simulated by `Sherpa 2.2.2` [197] was used with PDF and scale variations to estimate the theoretical uncertainty on the prediction of the shape and normalisation of the $m_{4\ell}$ spectrum. Variations on the PDFs used considered the eigenvectors of the `CT10nlo`, `MSTW2008nlo` and `NNPDF30NNLO` PDF sets. The variation (treated as the theoretical systematic uncertainty) is of the order of 4-5% over the $m_{4\ell}$ spectrum. QCD scale uncertainties are determined by varying μ_R and μ_F up and down by a factor of 2 from the central scale choice of $\mu_R = \mu_F = m_{4\ell}$, and is found to vary between 2 and 4% over the $m_{4\ell}$ spectrum.

The $gg \to ZZ$ background, having its LO production being a loop process, has significantly larger QCD uncertainties than those for $qq \to ZZ$ mentioned above. The uncertainties are determined in the same way as for the $qq \to ZZ$ process, although in this case the central scale is set to $\mu_R = \mu_F = 2m_Z$. The QCD scale uncertainty was calculated to be of the order of 30%. On the other hand, the PDF uncertainty was only around 4%, and was counted to be negligible compared to the scale uncertainty.

The signal samples are also passed through the same prescription for PDF and scale uncertainties, the latter using the central scale of $\mu_R = \mu_F = m_{4\ell}$. Each of these effects were determined to be less than a 1% effect. In addition to this, the signals gain an additional theory uncertainty from variations of certain parton shower parameters in `Pythia 8`. The prescription for these variations arise from the method by which `Pythia 8` was tuned by ATLAS (the `A14` tune [198]). These variations were also less than 1%.

7.7 RESULTS

Comparisons between the data and the background MC samples, normalised to their appropriate cross sections and luminosity values, were thereafter analysed as a function of $m_{4\ell}$. The figures cannot be shown due to issues of data propriety, however they will be made public by the ATLAS collaboration in due course. All of the predictions and data analysed were selected using the physics objects described in Section 7.2 and the selection procedure in Section 7.3.

The overall agreement between the MC and the data was fair and within systematic and statistical uncertainties, apart from the two excesses mentioned at the beginning of this chapter. On an inspection of the 2015-2016 data, the mild excesses at 240 GeV and 700 GeV were visible without having to perform any further analysis. In order to quantify these effects, a statistical study was needed, and this is performed in the next section. Unfortunately, due to data propriety issues, the 2017 results cannot be discussed in this thesis.

7.8 STATISTICAL INTERPRETATION

The statistical treatment of the results makes use of a test statistic, which is used for hypothesis testing and limit setting. The test statistic used is a profile likelihood ratio, defined in a similar way to Equation (6.2):

$$\lambda\left(\mu_{ggF}\right) = \frac{L\left(\mu_{ggF}, \hat{\hat{\theta}}\right)}{L\left(\hat{\mu}_{ggF}, \hat{\theta}\right)}. \tag{7.8}$$

Given a value of $m_{4\ell}$, the parameter of interest here is μ_{ggF}, the signal strength. A value of $\mu_{ggF} = 0$ corresponds to no signal, whereas a value of $\mu_{ggF} = 1$ corresponds to a ggF signal with a Higgs-like production cross section that depends on the mass of H, taken from Reference [33]. However, since the BR to ZZ is not known, what is ultimately constrained by the statistical procedure is the cross section times BR: $\sigma_{ggF} \times \text{BR}(H \rightarrow ZZ)$. The NPs, θ, represent the estimates of the systematic uncertainties, and are each modelled as a Gaussian variation from their nominal values. Each likelihood function L is a product of Poisson probabilities and systematic constraint functions. The overall maximum likelihood is given by the denominator of Equation (7.8), $L\left(\hat{\mu}_{ggF}, \hat{\theta}\right)$, and therefore the profile likelihood ratio will have a value between 0 and 1. The values of the NPs that maximise the likelihood function for an input value of $\mu_{ggF} = 0$ is $\hat{\hat{\theta}}$.

For each category, likelihood fits were performed with different hypotheses, both involving the signal and background predictions. These are made using the parameterised functions described in Section 7.5. For the signal hypothesis, a scan was made on the values of $m_{4\ell}$, ranging between 200 GeV and 1400 GeV. The first test calculated a p_0 value, which is the probability that a deviation in the spectrum was produced by a fluctuation of the background hypothesis (i. e. the SM). The results of this cannot be shown, but will appear in future ATLAS publications of the data.

From the p_0 study, the two excesses at 240 GeV and 700 GeV could be seen with their relative significance values using the 2015-2016 data. The 700 GeV excess appeared in a region where the background is small, and so it is possible that the value of the significance could be artificially enhanced to a large extent by any

form of mis-modelling in this region. The excess at 240 GeV, however, had a more coherent structure. In the case that this may be due to a real signal, it is likely that the signal has a large width, and therefore the significance is limited by the fact that this analysis only considers narrow-width signals. Large-width signals are slightly more complicated to implement, due to their interference with the $gg \to ZZ$ background; this will be explored in a future work. Large width studies without the interference have been considered in Reference [123].

Although the two mild excesses could be seen in the p_0 study, they were not large enough to be considered significant excesses. Because of this, limits were set on the cross section times BR for the production of H through ggF. These limits were set at the 95% CL and calculated as a function of m_H using the profile likelihood ratio in Equation (7.8) as a test statistic. Again, the result cannot be shown but will appear in a future ATLAS publication. Since the combined data set is roughly double the size of the two smaller data sets, the expected limit improved by roughly a factor of $1/\sqrt{2}$. As anticipated, the excesses at 245 GeV and 700 GeV were visible as deviations of the observed limit in excess of the expected limit. In both cases, the observed limits exceeded the 2σ uncertainty band, corresponding with the results from the p_0 study.

Interpreting the data as being produced by a heavy scalar decaying to a pair of Z bosons is directly comparable to the BSM physics model discussed in Chapters 4 to 6. Therefore, it is worthwhile considering the possibility that the excess at 245 GeV is related to the excess at 270 GeV found in earlier parts of this thesis, even though there exists a slight difference between the two best-fit masses. A likely possibility is that the heavy scalar has a large width, and therefore big fluctuations in the mass value are possible with this limited data set. An additional possibility is that there exist two separate structures, which may or may not be due to the same initiating resonance. This hypothesis is corroborated by the results of the study in Reference [123], which considers a combination of the Run 2 ATLAS and CMS $H \to 4\ell$ results [96, 125] and the ATLAS $H \to 2\mu$ result [199]. The p_0 plot for these different components using a narrow-width scalar is shown in Figure 39. In the plot, there appear to be two different structures for the ensemble of results

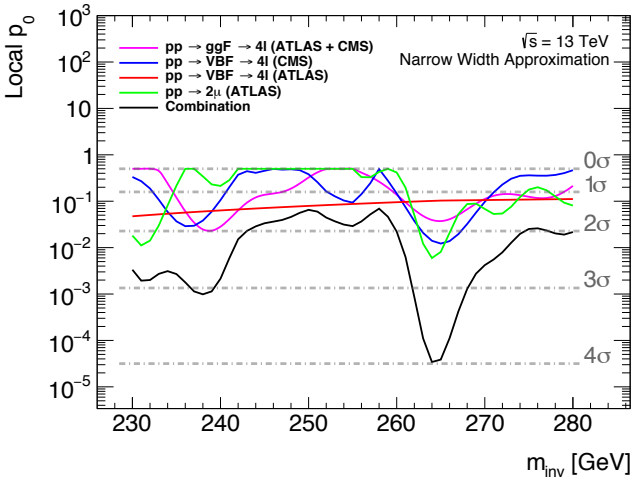


Figure 39: The local p_0 values for the Run 2 searches for $H \to 4\ell$ by ATLAS and CMS, and the ATLAS search for $H \to 2\mu$, as performed in the study in Reference [123]. On the x-axis is the invariant mass of the final state system; here it can be interpreted as m_H.

considered: one at a mass of around 240 GeV and the other at a higher mass of around 265 GeV. Knowing this, a different question arises if one considers that both peaks are due to new physics: is the double-peak structure due to the decay of one large-width object, or two different objects with smaller widths? This question can be addressed in future studies and with more sophisticated tools, such as machine learning techniques geared towards a BSM interpretation of the data.

8

CONCLUSIONS

The underlying purpose of this thesis has been to understand whether undetected new physics might exist in the limited LHC data set considered thus far. In particular, a study has been made on searching for a hypothetical heavy scalar boson with a mass around the EW scale. This study has been done using a the pragmatic methodology described in Section 1.1, where each step of the model's development can be interpreted holistically.

8.1 RESEARCH SUMMARY

The starting point of the discussion considered a statistically limited set of ATLAS and CMS Higgs-related results that seemed to mildly deviate from the SM expectation. The most accessible such result was that of the Higgs boson p_{T} spectrum, which at first seemed to indicate an enhancement in the region between 20 GeV and 100 GeV. From this point a simplified model was designed to try and explain this enhancement through the decay of a heavy scalar H to an as-yet unspecified associated Higgs production mode ($H \to h + X$). The model assumptions were simple: the H was produced dominantly via ggF and the overall rate of the production of H was controlled by a single parameter β_g^2, which was used to make a fit to the data. In assuming a ggF production mode for H, the production of H in association with top quarks is also necessary, and so this mode was also included in the study. While fitting this model to the Higgs p_{T} spectrum, it was also decided to include limits on heavy resonances and searches for top associated Higgs production, which further constrained the fit.

The result from this early study was that the limited data set appeared to be compatible with the hypothesis of a heavy scalar with a mass of around 270 GeV.

The compatibility was further quantified as a 3σ effect [3, 6]. The model that was chosen was to study the three-body decay process $H \to h\chi\chi$, with χ being a DM particle with a mass of around $m_h/2$. As newer data sets were published by the experimental collaborations, the discrepancies in the Higgs p_{T} spectrum, di-boson resonance searches, and top associated Higgs production remained compatible with this early result. These anomalies remain to this day.

This was an interesting result that warranted further investigation. The first issue to address was the apparent "unnaturalness" of the large BR of the three-body decay process, where one would expect a two-body decay to dominate. This problem was addressed by introducing a new scalar S, which acts as a DM mediator through the $S \to \chi\chi$ decay mode [7]. The three-body decay process $H \to h\chi\chi$ was therefore replaced by the $H \to Sh$, $S \to \chi\chi$ cascaded decay process. In doing this, a new question arose: what else could S potentially couple to? This was answered with a simple assumption, that S would have globally re-scaled Higgs-like couplings to the SM particles. This had the dual effect of both reducing the number of free parameters in the theory, as well as expanding the search prospects for S through a host of new decay modes. The mass of S was considered to be slightly higher than the mass of the Higgs boson. With $m_H = 270$ GeV, the distortion of the Higgs p_{T} spectrum discussed above was maintained with an S mass of around 150 GeV. This choice of masses was also supported by a separate set of Run 1 results relating to the multiple lepton sector [4].

With $m_S = 150$ GeV, one of the key search prospects was the signature of the production of multiple leptons in association with b-jets in the $H \to Sh$ decay. The S boson decays mostly to W boson pairs at this mass, which have a relatively high BR to leptons, whereas the Higgs boson decays mostly to b-quarks. With the specific set of parameters described above, the model could produce a relatively significant signal in a number of LHC related searches. An ensemble of these results was compiled and tested for compatibility with the simplified BSM model [4, 5, 8]. The ensemble consisted of a variety of ATLAS and CMS results from Run 1 and 2 of the LHC. As a result of testing the background-only and signal plus background hypotheses with the ensemble of results, it was found that the simplified

model improves on the background-only hypothesis at the level of 8σ. However exceptional this result might sound, it is made even more remarkable when one considers that it is obtained using a simplified model with only one fit parameter. The fit parameter (β_g^2) found a best-fit value of 2.92 ± 0.35. It has become clear that the SM on its own cannot explain the anomalies present in the LHC data which led to this result. Therefore, it has been demonstrated that the result can be seen as the first indirect evidence of new physics beyond the SM in laboratory conditions.

This result can be compared with direct searches for heavy scalars. In particular, the ATLAS search for a heavy Higgs-like boson that decays to a pair of Z bosons with a four lepton final state was shown in Chapter 7. A mild excess was found with 79.8 fb^{-1} of data at a mass of 245 GeV, hinting that this excess may be correlated with the 270 GeV excess described above. The conclusion of this study leads to the possibility of a large width heavy scalar being produced, which can result in the double-peak structure. This structure was also identified in a separate combination study of $H \to 4\ell$ and $H \to \mu\mu$ studies [123].

8.2 DISCUSSION AND PROSPECTS

The initial research question posed in Chapter 1 was:

> *Given the current reach of physics results from high energy physics experiments, to what extent can anomalous features of the data be explained and constrained by a simplified scalar extension to the Standard Model?*

In order to answer this question appropriately, it is helpful for one to formulate an opinion about the data collected at the LHC (the high energy physics experiment in question). By and large, the results in this thesis tend to inspire an optimistic opinion of the search for new physics at the LHC.

Consider the simplified model designed in Chapter 4 and extended in Chapter 5. Starting with an extremely limited data set (compared to what is available currently), this model was only able to describe a small set of anomalous features of the data to a significance of 3σ. Making only a few assumptions about the

simplified model, and following some logical conclusions about its shortcomings, resulted in a study of a more recent data set, presented in Chapter 6. With only one degree of freedom, the overall improvement of the data compared with the SM reached 8σ of significance. Even though it is acknowledged that the model in this thesis has its shortcomings, the result makes for a very strong argument that there is physics present at the LHC that is not understood.

The answer to the research question is therefore quite concrete. A simplified extension to the SM has been able to explain anomalous features of the LHC data to a large extent. With this in mind, a new question could be posed for the future of this work. The new question would read something akin to the following:

> *What is the precise cause of the large discrepancies we see in the multi-lepton data?*

The answer to this question requires that we appeal to a more comprehensive approach of analysing the data. One cannot always rely on traditional methods and inclusive analyses of the data. Recent advancements in the fields of machine learning and classification techniques are promising avenues for untangling the potentially complicated physics processes that remain to be understood from the data.

With Run 2 of the LHC now officially complete, the popular opinion of the particle physics community has started to show signs of negativity towards future research (see for instance Reference [200]). While this sentiment is still far from being the opinion of the majority, the consistent reliance on traditional methods has led to frustration at no new discoveries at the expense of years of hard work. In this thesis, it has been demonstrated that if new physics processes are being produced at the LHC, it is not so clear that the underlying cause should present itself to the community in an obvious way.

A

MODELLING BSM PROCESSES FOR EXPERIMENTAL SEARCHES

When one is concerned with experimental searches and measurements of physics in the SM and beyond, theoretical predictions are almost always made using MC predictions. This appendix describes the technical details for producing a set of MC samples for BSM processes related to the production of a heavy scalar H, such that they can be used in physics searches in the ATLAS collaboration.

In order to standardise the MC production procedure in the ATLAS collaboration, a centralised software package called `Athena` is used to generate events. As an input, `Athena` reads in a job options (JO) file, which specifies details of the specific event generator(s) to be used. `Athena` can then be used to generate events first at *TRUTH* level[10] into what is called an `EVGEN` file. In this appendix, the necessary inputs are described to generate `EVGEN` files (i. e. without detector simulation) and distributions are shown at TRUTH level for the purpose of validating the inputs and extracting interesting information about the modelled processes themselves.

A.1 SEARCHES FOR THE HIGGS BOSON IN ASSOCIATION WITH MISSING ENERGY

As discussed in Chapters 4 and 5, a starting point for the BSM model discussed in this thesis is the production of the Higgs boson in association with $E_{\mathrm{T}}^{\mathrm{miss}}$ in the form of DM. This is a commonly searched for signature in the experiments at the LHC, and so it was decided that this could be used as a search with the

10 TRUTH level refers to a fully evolved event without any detector simulation. Note that the name should not necessarily imply that the event is modelled *truly*. It is up to the discretion of the duly appointed authorities in the ATLAS collaboration to decide whether or not produced MC samples are modelled correctly.

ATLAS detector to constrain the parameters of the model using the $h \to \gamma\gamma$ decay mode. In this section, the first part describes how the MC prediction was made for the effective H-h-χ-χ vertex introduced in Chapter 4, and the second part uses the intermediate scalar S, as introduced in Chapter 5. The former prediction was used in a search that was published, and provided an important constraint on the parameter β_g^2.

A.1.1 *With an effective H-h-χ-χ vertex*

The theoretical study in Chapter 4 made use of a custom-designed UFO [80] model for use in `MadGraph5_aMCNLO` [81] to generate the $gg \to H \to h\chi\chi$ process at LO. The same approach was followed for the MC production process in the ATLAS collaboration with one key modification. For the ATLAS software to correctly identify the χ particles as DM candidates, it was decided to change their identification code to that of a SUSY neutralino. This has no effect on the physics of the process, but is merely a technicality to make the event generation more compatible with `Athena`.

The commands used to generate the process in `MadGraph5_aMCNLO` are therefore as follows:

```
import model sm
define p = g u c d s u~ c~ d~ s~
define j = g u c d s u~ c~ d~ s~
import model HHDM_intMET
generate p p > h2, h2 > h n1 n1
```

Here the name of the modified UFO model is `HHDM_intMET`, `h2` refers to the heavy scalar H, `h` refers to the SM Higgs boson and `n1` refers to the DM candidate χ. The masses of the particles could be specified in the corresponding parameter card. In this analysis a wide range of samples were produced with masses for H between 260 GeV and 350 GeV, and the mass for χ was either 50 GeV or 60 GeV. The BR of $H \to h\chi\chi$ was assumed to be 100% for the analysis, and therefore a limit on

the production cross section times BR can be interpreted as a direct limit on the $H \to h\chi\chi$ BR.

The decays, parton shower and hadronisation were all handled by `Pythia 8` [82]. In ATLAS, the default prescription for the `Pythia 8` parton shower is to use the `A14` tune as a configuration [198]. `Athena` also deals with τ-lepton decays and photons, respectively, with `TAUOLA` and `PHOTOS` [201]. The decays of B-hadrons are handled by the `EvtGen` software package [202].

A few interesting distributions from the production process are shown for four different mass points in Figure 40. The chosen mass points are the 4 combinations of $m_H = \{270, 330\}$ GeV and $m_\chi = \{50, 60\}$ GeV. The first plot is the transverse momentum of the di-photon system, which in this case is extremely similar to the $E_{\mathrm{T}}^{\mathrm{miss}}$. The peak shifts towards higher values when the difference between m_H and $m_h + 2m_\chi$ becomes larger. The second plot is the transverse momentum of the $\gamma\gamma\chi\chi$ system, which is a reconstruction of H. This is the typical spectrum for ggF at LO, where without the parton shower the p_{T} would always be zero, and the tail is determined by the parton shower. Thirdly and finally, is an interesting observable that could not be measured experimentally, the invariant mass of the $\chi\chi$ system. The distribution is still interesting, as it can provide an approximated mapping between the simple effective H-h-χ-χ vertex and any similar processes that predict an intermediate particle decaying to $\chi\chi$ instead, such as the process discussed in Chapter 5 and the next section.

As mentioned above, these MC samples were used in a Run 2 ATLAS analysis that placed limits on the production cross section times BR of the $gg \to H \to h\chi\chi$ hypothesis [160]. The resulting 95% CL as a function of m_H can be seen in Figure 41. In this figure the theoretical cross section times BR for the process (with its associated uncertainty) is shown as a blue band. This theoretical prediction includes a β_g^2 scaling of the order of approximately two. Since the explicit assumption was made that $\mathrm{BR}(H \to h\chi\chi) = 100\%$, this limit should be interpreted as a limit on the BR if one assumes that the production cross section is fixed. Therefore, it can be interpreted that at 95% CL, the $\mathrm{BR}(H \to h\chi\chi)$ is excluded for values larger than 50% for the lower masses of H, and 10% for the higher masses.

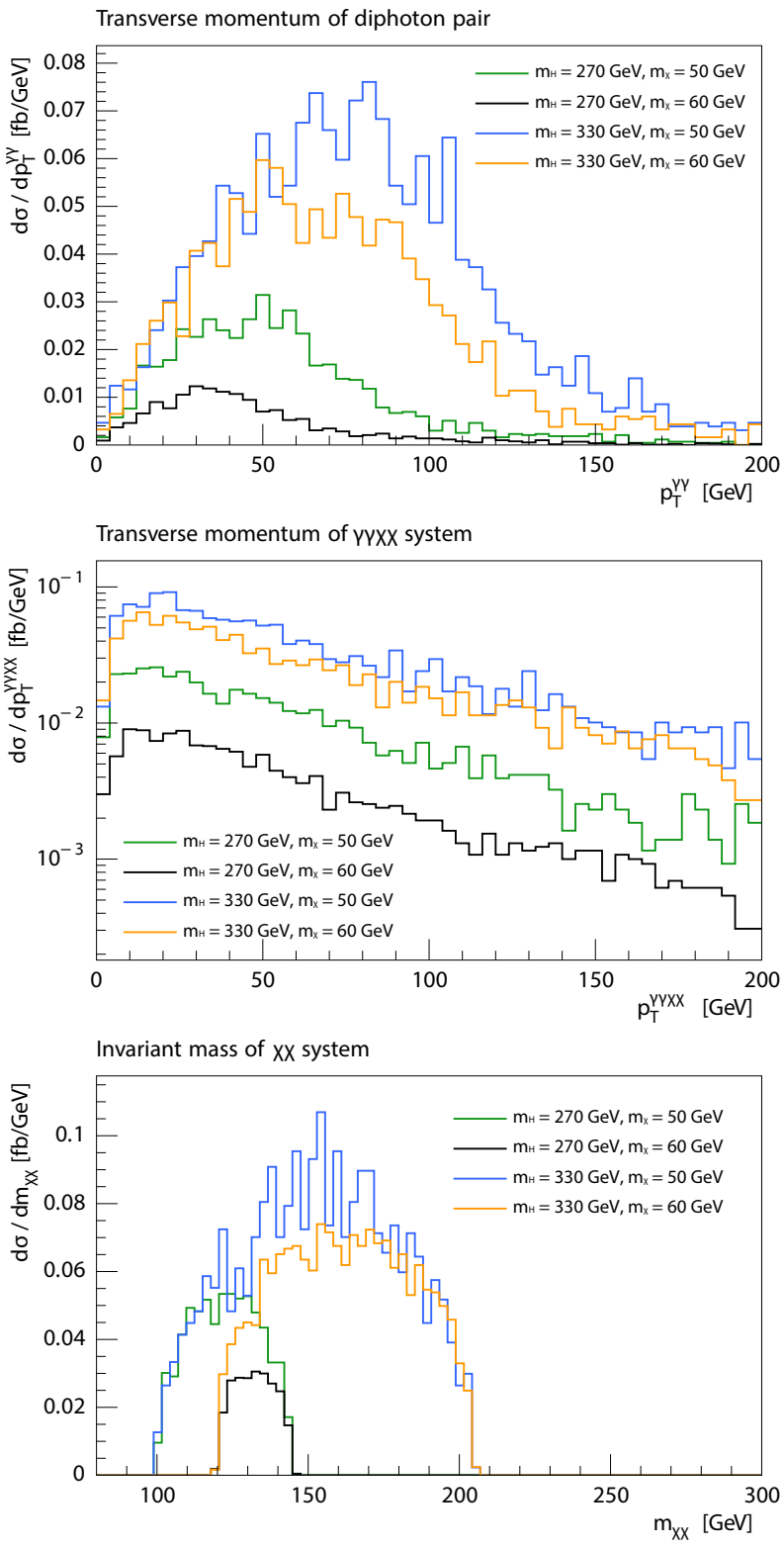


Figure 40: A series of TRUTH level validation plots for the MC production of $gg \to H \to h\chi\chi$ ($h \to \gamma\gamma$) using `MadGraph5_aMCNLO` and `Pythia 8` in `Athena` for a set of mass points. From the top to the bottom are distributions of the transverse momentum of the di-photon pair, the transverse momentum of H and the invariant mass of the $\chi\chi$ system.

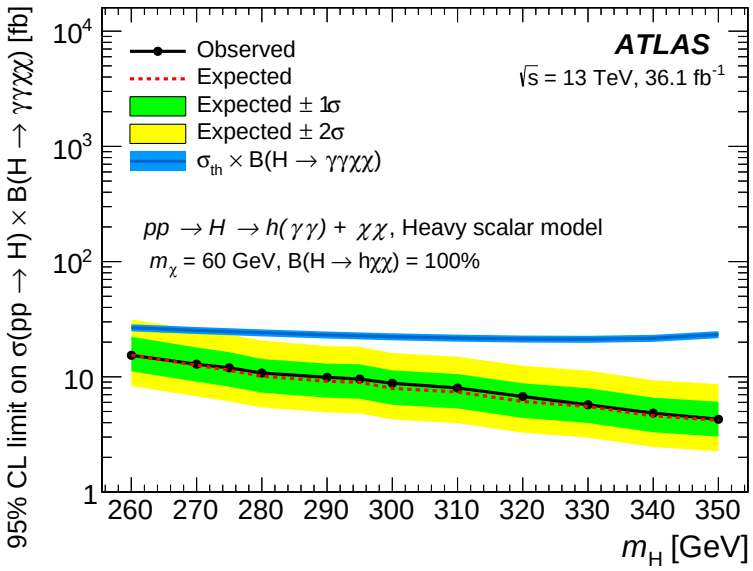


Figure 41: The 95% CL as a function of m_H in the ATLAS Run 2 search for a Higgs boson decaying to two photons in association with $E_{\mathrm{T}}^{\mathrm{miss}}$ [160].

A.1.2 *With the S mediator*

If one considers the production of the Higgs boson in association with DM and/or other particles via a mediator S (such as what was done in Chapter 5 with the $H \to Sh$ decay mode), new degrees of freedom are introduced into the theory. The new degrees of freedom will at least include the mass and width of S is introduced. However, if one considers this model to be used in an experimental search, the narrow-width approximation can be invoked to fix the width of S, and the mass therefore becomes the only free parameter. In addition to this, the degrees of freedom related to χ fall away, since the fully reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ will only depend on the parameters of S in the $S \to \chi\chi$ decay mode. Therefore, using S as a mediator does not necessarily complicate the process described above, but rather offers a simple way to explore the structure of $E_{\mathrm{T}}^{\mathrm{miss}}$ production with different masses for H and S. This is briefly discussed in this section.

The production of H through ggF, and its subsequent decay to Sh or SS, can be modelled easily in `Pythia`, since all of the involved particles are Higgs-like bosons (the $H \to Sh$ and SS decays are isotropic). As an example, shown below are the required input commands for the JOs to generate the $gg \to H \to Sh$ process, where

the Higgs boson decays to $\gamma\gamma$ and the S boson decays to all other possible final states:

```
Higgs:useBSM = on
ParticleDecays:mSafety = 0.0
HiggsBSM:gg2A3 = on
Higgs:clipWings = off
36:mWidth = 0.01
36:doForceWidth = on
36:addChannel = 1 1 100 35 25
36:onMode = off
36:onIfMatch = 35 25
36:mayDecay = on
35:mMin = 50.0
25:mMin = 50.0
25:onMode = off
25:onIfMatch = 22 22
SLHA:readFrom = 2
SLHA:file = H300_S140_Sh.slha
SLHA:useDecayTable = on
SLHA:allowUserOverride = on
```

Many of the details here are of a technical nature, but the important points are as follows. The H is modelled as a particle with a particle identification (PID) code of 36.[11] The production of H is therefore specified on line 3. The PID codes of the Higgs and S bosons are 25 and 35, respectively. Line 7 therefore adds the $H \to Sh$ decay mode, and it is activated as the only decay mode with lines 8-10. The $h \to \gamma\gamma$ decay mode is specified on lines 13-14 (the photon has a PID code of 22).

Finally, the mass, width and BRs of the S boson were specified with a file following the SUSY Les-Houches Accord (SLHA). `Pythia` reads the SLHA file in with lines 15-18 above. The format of an SLHA file specifies different blocks for masses and decays. In this case, the S boson is assumed to have SM Higgs boson-like de-

11 Strictly speaking, the PID code of 36 is reserved for a pseudo-scalar. However, the kinematics of the production for a pseudo-scalar by ggF is identical to that of a scalar (it only differs up to an overall cross section).

cays with an additional $S \to \chi\chi$ BR of 50%. The appropriate SLHA file needed to achieve this is as follows:

```
BLOCK MASS
36          300.0
35          140.0
1000022     60.0

DECAY   35 0.01
        0.5         2   1000022 1000022
        0.0298      2   21      21
        0.000965    2   22      22
        0.001225    2   21      23
        0.2505      2   24      -24
        0.03435     2   23      23
        0.1575      2   5       -5
        0.0176      2   15      -15
        6.1e-05     2   13      -13
        0.00795     2   3       -3
        6.7e-05     2   4       -4
        0.0         2   6       -6
```

Lines 1-4 above are the `MASS` block, and simply specify the masses of the different particles. In order, they specify that $m_H = 300$ GeV, $m_S = 140$ GeV and $m_\chi = 60$ GeV. The other lines (in the `DECAY` block) deal with the decays of S. The total width is set to a narrow value of 0.01 GeV. The BRs to the SM particles are all SM-like divided by 2, and the BR to $\chi\chi$ is fixed at 50% in this case.

Similar JOs to the one shown above were made for the $H \to SS$ where one S decays to $\gamma\gamma$, and $H \to Sh$ where the S decays to $\gamma\gamma$ and the h decays openly. One can see the effect on the di-photon p_{T} for the different production modes in Figure 42, as well as the total truth $E_{\mathrm{T}}^{\mathrm{miss}}$. From these figures it appears that there is some mild discrimination ability between the different decay modes of H, if one attempts to reconstruct the transverse momenta of the different bosons. Note also

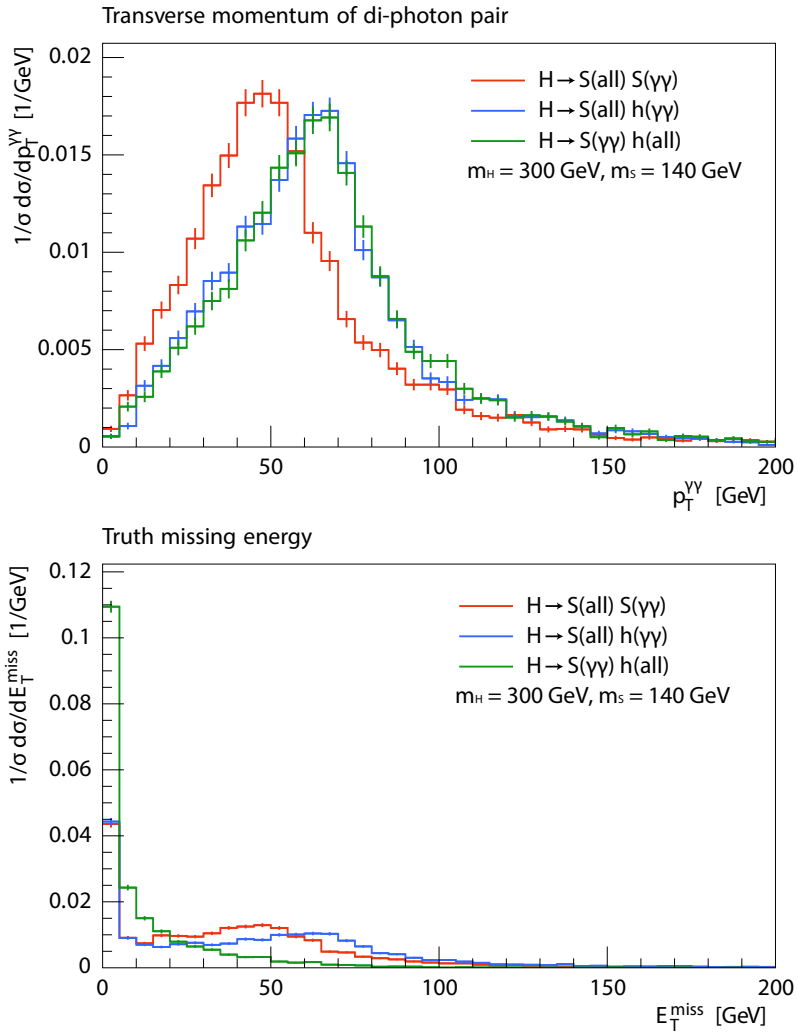


Figure 42: Comparisons of the $H \to SS$ and $H \to Sh$ decay modes for a mass point of $m_H = 300$ GeV and $m_S = 140$ GeV. The parentheses indicate the subsequent decays of the scalars, where "all" refers to an open decay.

that the green curve doesn't involve any production of DM, and this is why its $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution peaks so low.

In a search for these production modes, however, it makes more sense to look at jet-related variables for a better degree of discrimination. Several of these are shown in Figure 43, namely the jet multiplicity, the b-jet multiplicity and the p_{T} of the first jet. For instance, if one were to reconstruct the S boson via the $S \to \gamma\gamma$ decay mode, the jet and b-jet multiplicities already provide a big separation between the SS and Sh decays. Additionally, the discrimination power of the leading jet p_{T} can be used.

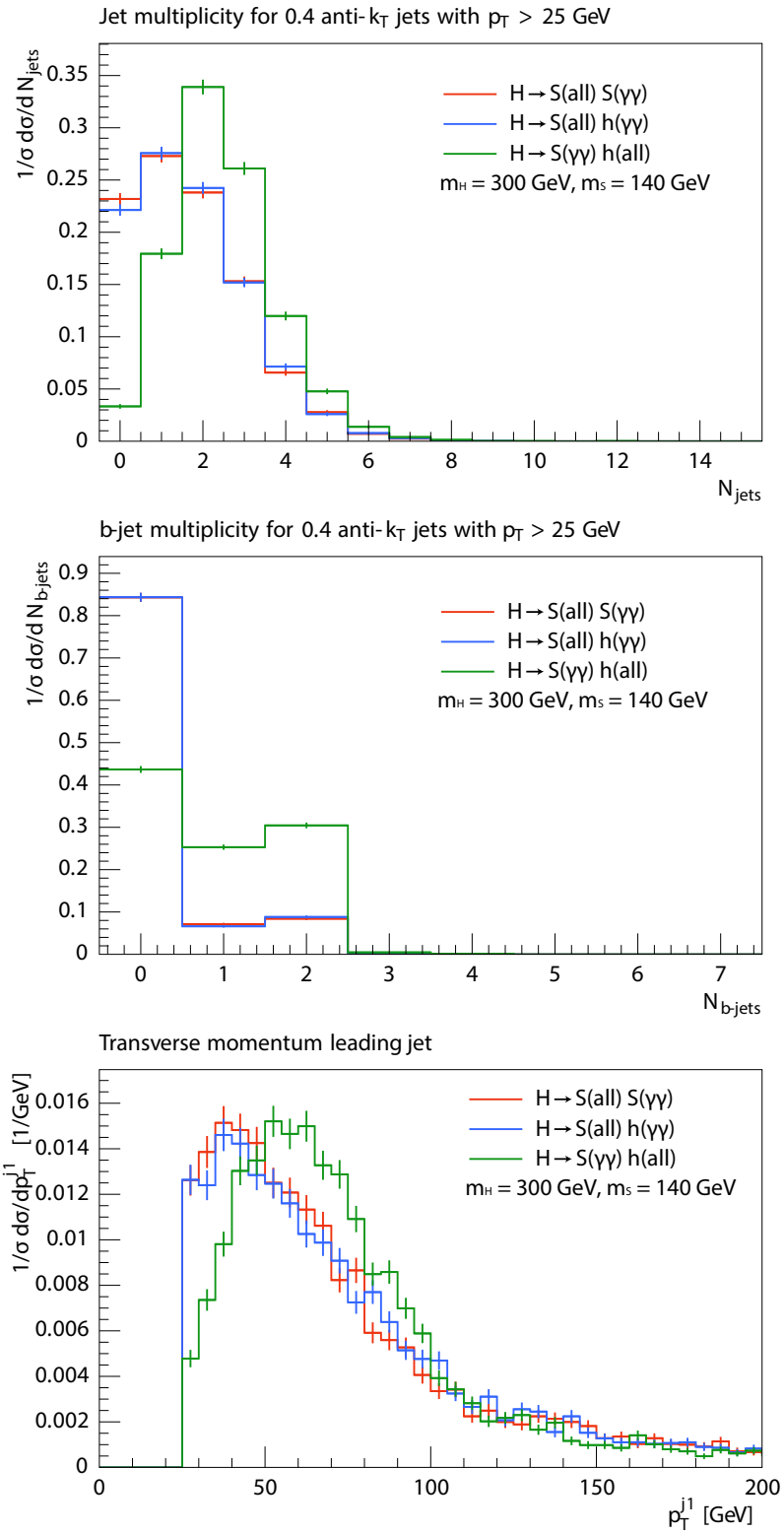


Figure 43: Comparisons of the $H \to SS$ and $H \to Sh$ decay modes for a mass point of $m_H = 300$ GeV and $m_S = 140$ GeV, using variables related to the production of additional jets. The parentheses indicate the subsequent decays of the scalars, where "all" refers to an open decay.

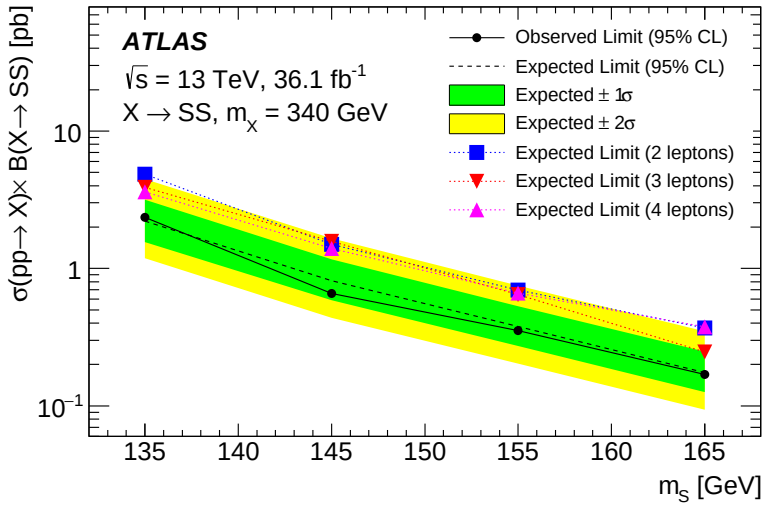
The signal shown above might appear promising, since the SM background is expected to be low for final states with a pair of photons and additional jets and/or b-jets. However, the low BR of the h and $S \to \gamma\gamma$ decay mode means that a signal would not be striking with the current reach of the LHC. In order to maximise the discovery potential of the $H \to SS$ and Sh decay modes, it makes more sense to look at a decay mode with a higher BR. As the mass of S becomes closer to $2m_W$, the $S \to WW$ BR becomes dominant (see Figure 3), and therefore the $SS/Sh \to 4W$ search becomes optimal. The ATLAS collaboration did indeed perform a search for this final state [162] using a similar method for MC production described above.

The final results are in the form of 95% CL plots as a function of the resonance masses, shown in Figure 44. No significant excess was found in the search, although it should be noted that the ATLAS search only considered the $H \to SS$ decay mode at specific masses above the kinematic threshold (i. e. the off-shell production of an S boson was not considered, which is an important part of the parameter space for the model). However, the limits are still useful as a gauge of the potential cross sections which might be at play for the considered BSM scenario. In future, other potential searches could be proposed to explore these regions of the phase space.

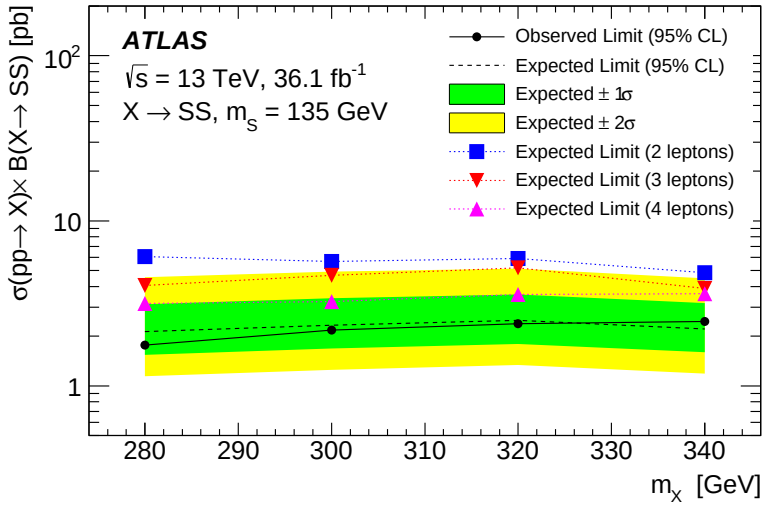
A.2 SEARCHES FOR ASSOCIATED H PRODUCTION

Up until this point in the thesis, the primary goal has been to study the potential existence of a heavy scalar H through its associated decay to the SM Higgs boson. However, as the reach of the LHC grows (and will continue to grow in the years to come), one might start to consider if the heavy scalar itself might be constructed directly via its decay to SM particles. In this section, a simple model is designed to anticipate this eventuality.

The experimental collaborations already make searches for heavy bosons, in most cases under the interpretation of a 2HDM. Many of these searches were discussed in Section 4.2. These searches will always consider the direct production of H, and at maximum might also consider a VBF-like production (that is, H with

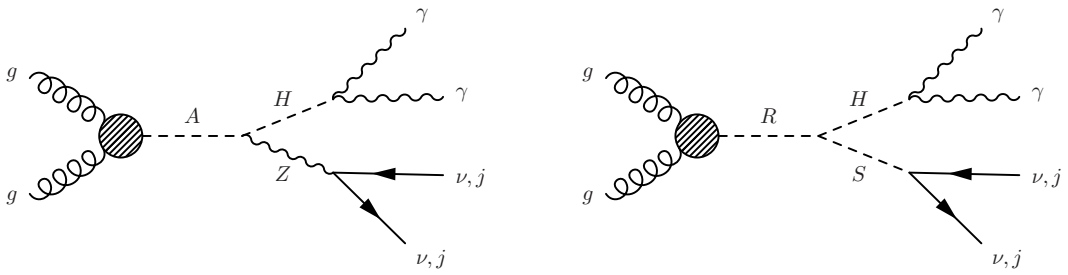


(a) With fixed $m_H = 340$ GeV.



(b) With fixed $m_S = 135$ GeV.

Figure 44: The results of the ATLAS search for the $H \to SS \to 4W$ decay process, in the form of 95% CL plots. In (a) the mass of H (labelled as X by ATLAS) is fixed to 340 GeV and the mass of S is scanned over, and in (b) the mass of S is fixed to 135 GeV and the mass of H is scanned over. The different coloured lines represent the limits from the 2, 3, and 4-lepton categories in the analysis. The black line shows the combined limit for all the categories.



(a) AZH model. (b) RSH model.

Figure 45: Representative Feynman diagrams for the topologies studied using the AZH and RSH models.

two additional jets). Whereas a direct production is the simplest case, one might also consider the associated production of H with extra degrees of freedom that can control the associated final state objects. This is possible by including new particles and interactions to the model that has been described so far, and was attempted here through the design of the so-called *AZH* and *RSH* models.

The AZH model is designed to replicate the topology of the production of a pseudo-scalar (A) production through ggF, which thereafter decays to ZH. This topology is shown as a Feynman diagram in Figure 45a. The Z boson can then decay to neutrinos to produce $E_{\mathrm{T}}^{\mathrm{miss}}$, leptons or jets. The H boson can be reconstructed most simply through any normal Higgs-like decay, $H \to \gamma\gamma$ for example. Since the mass, width and BRs of the Z boson are fixed, one can control the kinematics of the topology with at least 2 free parameters: the masses of A and H. The widths of A and H can also be changed, although one can invoke the narrow-width approximation and assume that their width is dominated by detector resolution effects.

The RSH model, on the other hand, has a similar topology to the AZH model, but introduces a new heavy scalar R (with $m_R > m_H + m_S$). Therefore, the RSH model has at least 3 free parameters that can be used to control the kinematics of the final state objects: the masses of R, H and S. Additionally their widths can be controlled, or fixed to a narrow value; differences in width will be studied in Appendix A.2.2. The topology can be seen in Figure 45b. Again, the decays of S can be chosen to be neutrinos if one wants to study $E_{\mathrm{T}}^{\mathrm{miss}}$, jets, vector bosons or leptons as an associated production mechanism.

Importantly, it should be noted that the AZH and RSH models are not intended to be theoretically robust models. They merely serve the purpose of providing a tool to explore the associated production of H in potential searches at the LHC.

The processes shown in Figure 45 can be produced using `Pythia` with relative ease. Below, it will be shown how this is done for the $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ decays modes. Some comparative distributions will also be shown in order to understand the effects of changing the models' free parameters.

A.2.1 *Modelling $H \to \gamma\gamma + X$*

The JOs needed to produce H in association with other final sate objects largely follow the same logic as the production process described in Appendix A.1.2, although without any need of an SLHA file. Firstly, for the AZH model, the production can be specified with the following `Pythia` commands:

```
Higgs:useBSM = on
HiggsBSM:gg2A3 = on
36:m0 = 400.0
36:mWidth = 0.01
36:doForceWidth = yes
36:onMode = off
36:onIfMatch = 25 23
25:m0 = 250.0
25:doForceWidth = yes
25:mWidth = 0.01
25:onMode = off
25:onIfMatch = 22 22
23:onMode = off
```

In this case, the A boson has a PID code of 36 and its mass and width are specified on lines 3 and 4, respectively. The $A \to ZH$ decay is activated on line 7, where the PID codes for Z and H are 23 and 25, respectively. The mass of H is set on line 8, and in lines 11 and 12 it is specified that the H should decay to two photons.

To specify the decay of Z, one can do the following to force it to decay to neutrinos:

```
23:onIfAny = 12 14 16
```

In the case of the $Z \to jj$ decay, one can write:

```
23:onIfAny = 1 2 3 4 5
```

And finally, for $Z \to \ell\ell$ decays (where ℓ is an electron or muon), one can write:

```
23:onIfAny = 11 13
```

The RSH model, on the other hand, follows a similar logic but requires a few extra lines:

```
Higgs:useBSM = on
HiggsBSM:gg2A3 = on
36:m0 = 450.0
36:mWidth = 0.01
36:doForceWidth = yes
36:addChannel = 1 1 103 35 25
36:onMode = off
36:onIfMatch = 35 25
25:m0 = 250.0
25:doForceWidth = yes
25:mWidth = 0.01
25:onMode = off
25:onIfMatch = 22 22
35:m0 = 160.0
35:mWidth = 0.01
35:doForceWidth = on
35:onMode = off
```

In this case, the R assumes the PID code of 36. The important additional lines are lines 6-8, which first define the $R \to SH$ decay and then activate it. Again in this case, the H has a PID code of 25, but the S has a PID code of 35. Therefore, the mass of the S is set on line 14, and its width is set on line 15.

The decay modes of S can then be chosen. For the $S \to jj$ decay mode, one can write:

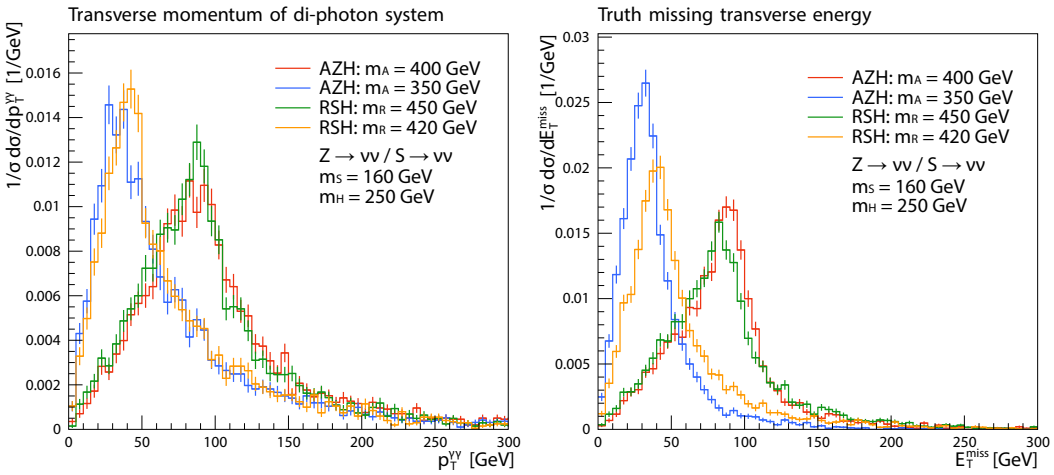
```
35:onIfAny = 21 21
```

Specifically, this activates the $S \to gg$ decay mode. The quarks can also be used as jets by specifying their PID codes (1-5). For the $S \to \nu\nu$ decay, the mode first has to be created, since the vertex does not exist *a priori* for a Higgs-like boson (recall that this model is in no way theoretically motivated). The decay mode can be created and activated by using the following commands:

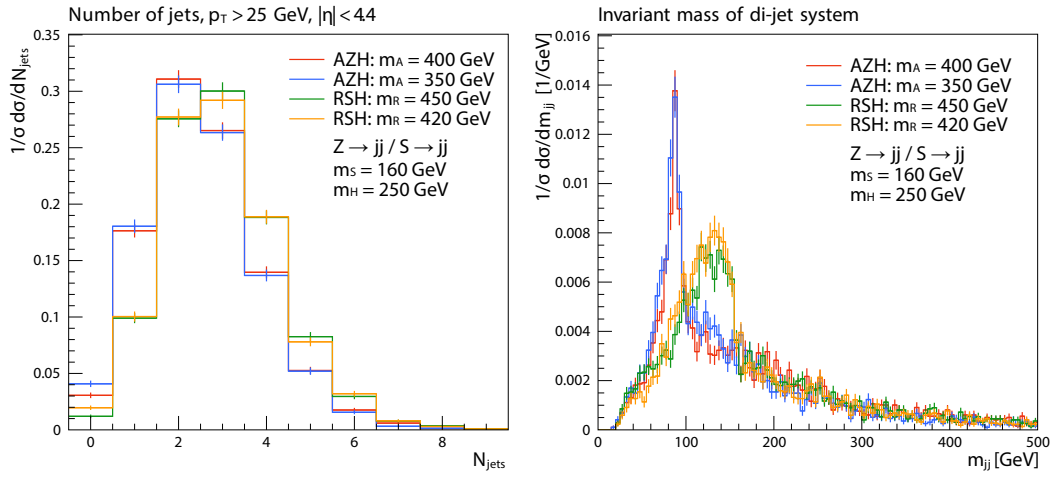
```
35:addChannel = 1 1 103 12 -12
35:onIfMatch = 12 -12
```

Using the commands above, a set of mass points was chosen and events were generated for comparative purposes. Distributions of a few important variables can be seen in Figure 46 for both the $S/Z \to \nu\nu$ and jj decay choices. For the RSH model, the mass of S is fixed to 160 GeV. For both models, different masses of A and R are considered, but the mass of H is fixed to 250 GeV. The chosen decay mode for H is to a pair of photons, although the plots are relevant to any decay mode of H (as long as it can be reconstructed confidently).

Some interesting points arise from the plots in Figure 46. Firstly as expected, the transverse momentum of the $\gamma\gamma$ system (i. e. the H boson) changes according to the difference in mass between the parent particle and the sibling particle. For example, the larger mass difference between R and S (when $m_R = 450$ GeV) boosts the H to a higher transverse momentum than when the difference is smaller (when $m_R = 420$ GeV). This also affects the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum. In the case where S and Z decay to jets, the jet multiplicity becomes dependent on the mass of the heaviest resonance in the process, since more energetic gluons are needed to produce a heavy resonance, and these energetic gluons tend to radiate more. This can be used to discriminate between mass hypotheses. This may be helpful, since it becomes difficult to reconstruct the mass of S or Z from the di-jet mass spectrum, as can be seen in the plots. Even at truth level (i. e. without detector resolution effects), the m_{jj} spectrum becomes wide due to energy losses during the formation of the jet.



(a) For the $S/Z \to \nu\nu$ decay mode.



(b) For the $S/Z \to jj$ decay mode.

Figure 46: A number of interesting observables for comparison of the AZH and RSH models for (a) the $S/Z \to \nu\nu$ decay choice and (b) the $S/Z \to jj$ decay choice. In all plots, $m_S = 160$ GeV and $m_H = 250$ GeV.

A.2.2 *Width effects with $H \to ZZ \to 4\ell + X$*

Generating events with the AZH or RSH model that have a different H decay mode requires only a small modification to the JOs shown above in Appendix A.2.1. In this brief section, the effect of changing the H width is determined using the RSH model with the $H \to ZZ \to 4\ell$ and $S \to \nu\nu$ decay modes. The necessary `Pythia` commands are as follows:

```
Higgs:useBSM = on
HiggsBSM:gg2A3 = on
36:m0 = 450.0
36:mWidth = 0.01
36:doForceWidth = yes
36:addChannel = 1 1 103 35 25
36:onMode = off
36:onIfMatch = 25 35
25:m0 = 250.0
25:doForceWidth = yes
25:onMode = off
25:onIfMatch = 23 23
23:onMode = off
23:onIfAny = 11 13
35:m0 = 160.0
35:mWidth = 0.01
35:doForceWidth = on
35:addChannel = 1 1 103 12 -12
35:onMode = off
35:onIfMatch = 12 -12
```

The $H \to ZZ \to 4\ell$ decay mode is specified in lines 11-14.

Thereafter, changing the width of H is trivial. Here, the width is considered at 1%, 3% and 5% of the nominal mass of H. If the mass of H is fixed to 250 GeV here, changing the width to 1% of the nominal mass requires the following:

```
25:mWidth = 2.5
```

Changing the width to 3% is done as follows:

```
25:mWidth = 7.5
```

And finally, changing the width to 5% is done with:

```
25:mWidth = 12.5
```

Using the JOs shown above, a few mass points were selected to understand the effect of the H boson width on the kinematics of its decay products. The selected mass points were:

- $m_R = 390$ GeV and $m_H = 220$ GeV,
- $m_R = 450$ GeV and $m_H = 250$ GeV, and
- $m_R = 800$ GeV and $m_H = 500$ GeV.

By and large, the kinematics behave similarly to the study shown in Appendix A.2.1. The biggest effects are on the mass and transverse momentum of the 4ℓ system, as shown in Figure 47.

The effect of the width on the mass of the 4ℓ system (i. e. the H boson) is predictable, since the width will change the structure of the Breit-Wigner factor in the propagator. Larger values of the width lead to a wider mass spectrum, by definition. The effect on the p_{T} spectrum is significantly more mild. But it is evident from Figure 47 that a larger width also widens the p_{T} spectrum. This is because the mass of H will fluctuate more with a larger width, and so the total available energy for the decay process will also fluctuate, leading to variations in the p_{T}. Since the S boson decays entirely to $E_{\mathrm{T}}^{\mathrm{miss}}$, the same effect can be seen in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution.

A.2.3 *Prospects for the AZH and RSH models*

It should be noted again that the AZH and RSH models are not theoretically motivated, but are useful for data-driven searches for heavy scalars. The freedom of

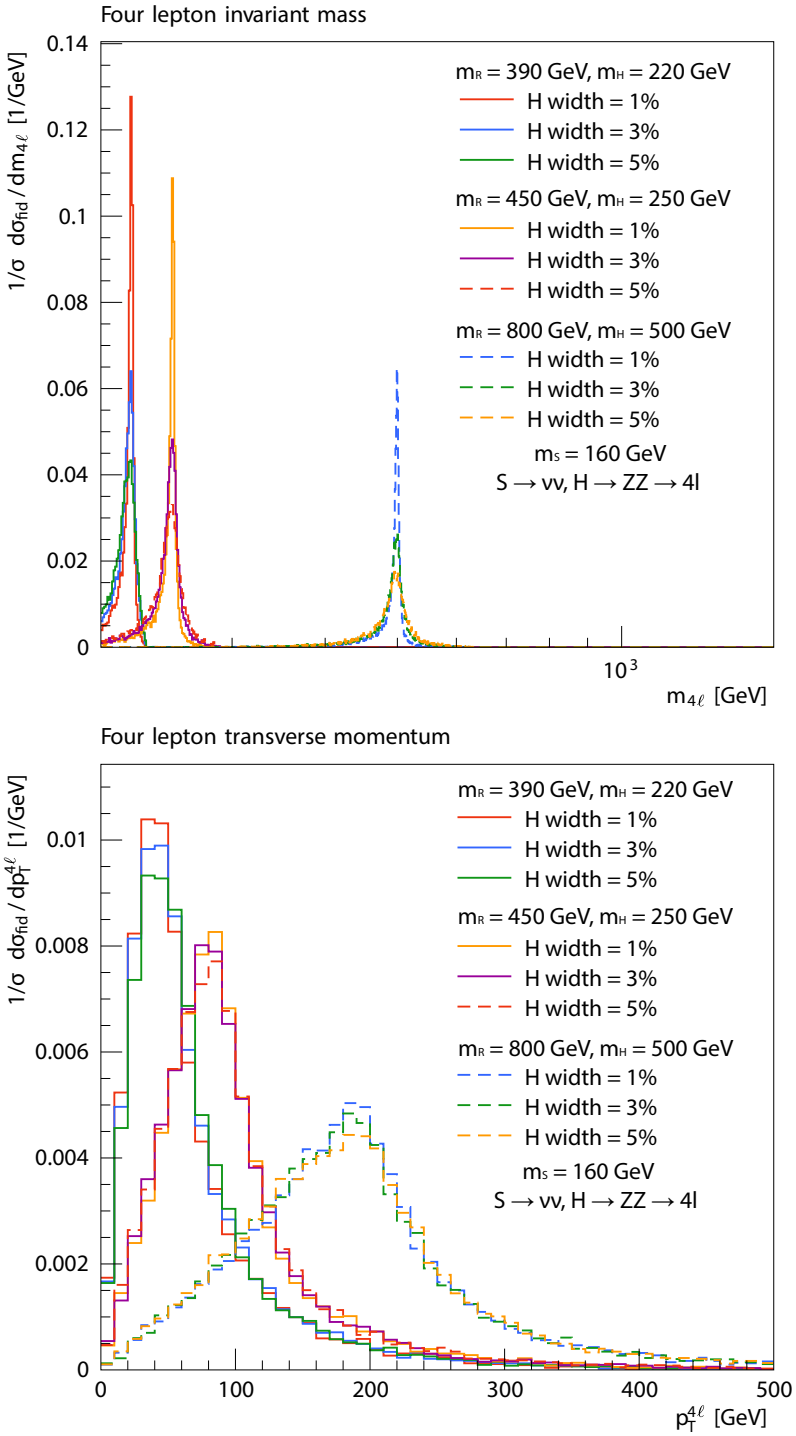


Figure 47: A comparison of different mass points in the RSH model for the $H \to ZZ \to 4\ell$ and $S \to \nu\nu$ decay modes. Different widths of H are also considered, at 1%, 3% and 5% of the nominal H mass.

choice in the parameters allows one to do a simple MC study useful for optimisation criteria and the expected magnitude of interesting variables.

Up until this point, the AZH and RSH models have not been used in a published experimental result. They have been useful in determining important variables and their cuts for some searches in ATLAS, particularly in the $H + E_{\mathrm{T}}^{\mathrm{miss}}$ channel. But these searches are still underway and may not be published in the near future.

BIBLIOGRAPHY

[1] **ATLAS** Collaboration, G. 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** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[2] **CMS** Collaboration, S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett.* **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[3] S. von Buddenbrock, N. Chakrabarty, A. S. Cornell, D. Kar, M. Kumar, T. Mandal, B. Mellado, B. Mukhopadhyaya, and R. G. Reed, "The compatibility of LHC Run 1 data with a heavy scalar of mass around 270 GeV," arXiv:1506.00612 [hep-ph].

[4] S. von Buddenbrock, A. S. Cornell, A. Fadol, M. Kumar, B. Mellado, and X. Ruan, "Multi-lepton signatures of additional scalar bosons beyond the Standard Model at the LHC," *J. Phys.* **G45** no. 11, (2018) 115003, arXiv:1711.07874 [hep-ph].

[5] S. Buddenbrock, A. S. Cornell, Y. Fang, A. Fadol Mohammed, M. Kumar, B. Mellado, and K. G. Tomiwa, "The emergence of multi-lepton anomalies at the LHC and their compatibility with new physics at the EW scale," *JHEP* **10** (2019) 157, arXiv:1901.05300 [hep-ph].

[6] S. von Buddenbrock *et al.*, "Constraining hypothetical extensions to the Higgs sector at the LHC," in *The Proceedings of SAIP2016, the 61st Annual Conference of the South African Institute of Physics*, S. Peterson and S. Yacoob, eds., pp. 158–163. ISBN: 978-0-620-77094-1.

[7] S. von Buddenbrock, N. Chakrabarty, A. S. Cornell, D. Kar, M. Kumar, T. Mandal, B. Mellado, B. Mukhopadhyaya, R. G. Reed, and X. Ruan, "Phenomenological signatures of additional scalar bosons at the LHC," *Eur. Phys. J.* **C76** no. 10, (2016) 580, arXiv:1606.01674 [hep-ph].

[8] S. von Buddenbrock and B. Mellado, "Anomalies in the production of multiple leptons at the LHC," arXiv:1909.04370 [hep-ph].

[9] **Particle Data Group** Collaboration, M. Tanabashi *et al.*, "Review of Particle Physics," *Phys. Rev.* **D98** no. 3, (2018) 030001.

[10] M. Gell-Mann, "The interpretation of the new particles as displaced charge multiplets," *Nuovo Cim.* **4** no. S2, (1956) 848–866.

[11] M. Goldhaber, L. Grodzins, and A. W. Sunyar, "Helicity of neutrinos," *Phys. Rev.* **109** (Feb, 1958) 1015–1017.

[12] S. L. Glashow, "Partial Symmetries of Weak Interactions," *Nucl. Phys.* **22** (1961) 579–588.

[13] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[14] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc.* **C680519** (1968) 367–377.

[15] **ATLAS, CMS** Collaboration, G. Aad *et al.*, "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments," *Phys. Rev. Lett.* **114** (2015) 191803, arXiv:1503.07589 [hep-ex].

[16] C.-N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance," *Phys. Rev.* **96** (1954) 191–195.

[17] T. Nakano and K. Nishijima, "Charge Independence for V-particles," *Prog. Theor. Phys.* **10** (1953) 581–582.

[18] E. Fermi, "Tentativo di una teoria dei raggi β," *Il Nuovo Cimento (1924-1942)* **11** no. 1, (Sep, 2008) 1.

[19] E. Fermi, "Versuch einer theorie der β-strahlen. i," *Zeitschrift für Physik* **88** no. 3, (Mar, 1934) 161–177.

[20] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.* **13** (1964) 508–509.

[21] P. W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons," *Phys. Rev.* **145** (1966) 1156–1163.

[22] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.* **13** (1964) 321–323.

[23] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," *Phys. Rev. Lett.* **13** (1964) 585–587. [,162(1964)].

[24] D. Kirzhnits and A. Linde, "Symmetry behavior in gauge theories," *Annals of Physics* **101** no. 1, (1976) 195 – 238.

[25] Y. Nambu, "Quasi-particles and gauge invariance in the theory of superconductivity," *Phys. Rev.* **117** (Feb, 1960) 648–663.

[26] J. Goldstone, "Field theories with superconductor solutions," *Il Nuovo Cimento (1955-1965)* **19** no. 1, (Jan, 1961) 154–164.

[27] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, "The Higgs Hunter's Guide," *Front. Phys.* **80** (2000) 1–404.

[28] B. Mistlberger, "Higgs boson production at hadron colliders at N^3LO in QCD," *JHEP* **05** (2018) 028, arXiv:1802.00833 [hep-ph].

[29] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, "Higgs Boson Gluon-Fusion Production in QCD at Three Loops," *Phys. Rev. Lett.* **114** (2015) 212001, arXiv:1503.06056 [hep-ph].

[30] F. A. Dreyer and A. Karlberg, "Vector-Boson Fusion Higgs Production at Three Loops in QCD," *Phys. Rev. Lett.* **117** no. 7, (2016) 072001, arXiv:1606.00840 [hep-ph].

[31] R. V. Harlander, J. Klappert, S. Liebler, and L. Simon, "vh@nnlo-v2: New physics in Higgs Strahlung," *JHEP* **05** (2018) 089, arXiv:1802.04817 [hep-ph].

[32] A. Broggio, A. Ferroglia, R. Frederix, D. Pagani, B. D. Pecjak, and I. Tsinikos, "Top-quark pair hadroproduction in association with a heavy boson at NLO+NNLL including EW corrections," *JHEP* **08** (2019) 039, arXiv:1907.04343 [hep-ph].

[33] **LHC Higgs Cross Section Working Group** Collaboration, D. de Florian *et al.*, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector," arXiv:1610.07922 [hep-ph].

[34] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[35] **ATLAS** Collaboration, A. Airapetian *et al.*, *ATLAS detector and physics performance: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1999.

[36] **ATLAS** Collaboration, *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1997.

[37] **ATLAS** Collaboration, S. Haywood, L. Rossi, R. Nickerson, and A. Romaniouk, *ATLAS inner detector: Technical Design Report, 2*. Technical Design Report ATLAS. CERN, Geneva, 1997.

[38] **ATLAS** Collaboration, *ATLAS central solenoid: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997.

[39] **ATLAS** Collaboration, N. Wermes and G. Hallewel, *ATLAS pixel detector: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998.

[40] **ATLAS** Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.

[41] **ATLAS** Collaboration, *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.

[42] **ATLAS** Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997.

[43] A. Ruiz-Martinez and ATLAS Collaboration, "The Run-2 ATLAS Trigger System," Tech. Rep. ATL-DAQ-PROC-2016-003, CERN, Geneva, Feb, 2016.

[44] **ATLAS** Collaboration, *ATLAS level-1 trigger: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998.

[45] **ATLAS** Collaboration, P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2003.

[46] ATLAS Collaboration, "LuminosityPublicResultsRun2," 2019. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. Last accessed: 16/01/2020.

[47] **ATLAS** Collaboration, M. Aaboud *et al.*, "Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector," *Eur. Phys. J.* **C77** no. 9, (2017) 580, arXiv:1705.02211 [hep-ex]. [Erratum: Eur. Phys. J.C77,no.10,712(2017)].

[48] **ATLAS** Collaboration, "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC," Tech. Rep. ATLAS-CONF-2019-021, 2019.

[49] **CMS** Collaboration, G. L. Bayatian *et al.*, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006.

[50] **ALICE** Collaboration, *ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC*. LHC Tech. Proposal. CERN, Geneva, 1995.

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

[52] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π decay of the k_2^0 meson," *Phys. Rev. Lett.* **13** (Jul, 1964) 138–140.

[53] **LHCb** Collaboration, R. Aaij *et al.*, "Measurement of the ratio of branching fractions $\mathcal{B}(\overline{b}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{b}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$," *Phys. Rev. Lett.* **115** (Sep, 2015) 111803.

[54] W.-L. Guo and Y.-L. Wu, "The Real singlet scalar dark matter model," *JHEP* **10** (2010) 083, arXiv:1006.2518 [hep-ph].

[55] Y. Cai, X.-G. He, and B. Ren, "Low Mass Dark Matter and Invisible Higgs Width In Darkon Models," *Phys. Rev.* **D83** (2011) 083524, arXiv:1102.1522 [hep-ph].

[56] M. Gonderinger, H. Lim, and M. J. Ramsey-Musolf, "Complex Scalar Singlet Dark Matter: Vacuum Stability and Phenomenology," *Phys. Rev.* **D86** (2012) 043511, arXiv:1202.1316 [hep-ph].

[57] A. Djouadi, A. Falkowski, Y. Mambrini, and J. Quevillon, "Direct Detection of Higgs-Portal Dark Matter at the LHC," *Eur. Phys. J.* **C73** no. 6, (2013) 2455, arXiv:1205.3169 [hep-ph].

[58] **ATLAS** Collaboration, G. Aad *et al.*, "Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8$ TeV with ATLAS," *JHEP* **09** (2014) 112, arXiv:1407.4222 [hep-ex].

[59] **ATLAS** Collaboration, G. Aad *et al.*, "Fiducial and differential cross sections of Higgs boson production measured in the four-lepton decay channel in pp collisions at $\sqrt{s}$=8 TeV with the ATLAS detector," *Phys. Lett.* **B738** (2014) 234–253, arXiv:1408.3226 [hep-ex].

[60] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 8$ TeV," *Eur. Phys. J.* **C76** no. 1, (2016) 13, arXiv:1508.07819 [hep-ex].

[61] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of differential and integrated fiducial cross sections for Higgs boson production in the four-lepton decay channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV," *JHEP* **04** (2016) 005, arXiv:1512.08377 [hep-ex].

[62] **ATLAS** Collaboration, G. Aad *et al.*, "Searches for Higgs boson pair production in the $hh \to bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector," *Phys. Rev.* **D92** (2015) 092004, arXiv:1509.04670 [hep-ex].

[63] **CMS** Collaboration, "Search for the resonant production of two Higgs bosons in the final state with two photons and two bottom quarks,". CMS-PAS-HIG-13-032.

[64] **CMS** Collaboration, V. Khachatryan *et al.*, "Searches for a heavy scalar boson H decaying to a pair of 125 GeV Higgs bosons hh or for a heavy pseudoscalar boson A decaying to Zh, in the final states with $h \to \tau\tau$," *Phys. Lett.* **B755** (2016) 217–244, arXiv:1510.01181 [hep-ex].

[65] **CMS** Collaboration, V. Khachatryan *et al.*, "Searches for heavy Higgs bosons in two-Higgs-doublet models and for $t \to ch$ decay using multilepton and diphoton final states in pp collisions at 8 TeV," *Phys. Rev.* **D90** (2014) 112013, arXiv:1410.2751 [hep-ex].

[66] **ATLAS** Collaboration, G. Aad *et al.*, "Search for $H \to \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector," *Phys. Lett.* **B740** (2015) 222–242, arXiv:1409.3122 [hep-ex].

[67] **ATLAS** Collaboration, G. Aad *et al.*, "Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector," *Phys. Lett.* **B749** (2015) 519–541, arXiv:1506.05988 [hep-ex].

[68] **ATLAS** Collaboration, G. Aad *et al.*, "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector," *Eur. Phys. J.* **C75** no. 7, (2015) 349, arXiv:1503.05066 [hep-ex].

[69] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for the associated production of the Higgs boson with a top-quark pair," *JHEP* **09** (2014) 087, arXiv:1408.1682 [hep-ex]. [Erratum: JHEP10,106(2014)].

[70] **ATLAS** Collaboration, G. Aad *et al.*, "Search for a high-mass Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *JHEP* **01** (2016) 032, arXiv:1509.00389 [hep-ex].

[71] **ATLAS** Collaboration, G. Aad *et al.*, "Search for an additional, heavy Higgs boson in the $H \to ZZ$ decay channel at $\sqrt{s} = 8$ TeV in pp collision data with the ATLAS detector," *Eur. Phys. J.* **C76** no. 1, (2016) 45, arXiv:1507.05930 [hep-ex].

[72] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for a Higgs boson in the mass range from 145 to 1000 GeV decaying to a pair of W or Z bosons," *JHEP* **10** (2015) 144, arXiv:1504.00936 [hep-ex].

[73] **Planck** Collaboration, P. A. R. Ade *et al.*, "Planck 2013 results. XVI. Cosmological parameters," *Astron. Astrophys.* **571** (2014) A16, arXiv:1303.5076 [astro-ph.CO].

[74] **LUX** Collaboration, D. S. Akerib *et al.*, "First results from the LUX dark matter experiment at the Sanford Underground Research Facility," *Phys. Rev. Lett.* **112** (2014) 091303, arXiv:1310.8214 [astro-ph.CO].

[75] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov, "micrOMEGAs_3: A program for calculating dark matter observables," *Comput. Phys. Commun.* **185** (2014) 960–985, arXiv:1305.0237 [hep-ph].

[76] J. M. Cline, K. Kainulainen, P. Scott, and C. Weniger, "Update on scalar singlet dark matter," *Phys. Rev.* **D88** (2013) 055025, arXiv:1306.4710 [hep-ph]. [Erratum: Phys. Rev.D92,no.3,039906(2015)].

[77] **ATLAS** Collaboration, M. Aaboud *et al.*, "Combination of searches for invisible Higgs boson decays with the ATLAS experiment," *Phys. Rev. Lett.* **122** no. 23, (2019) 231801, arXiv:1904.05105 [hep-ex].

[78] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Search for invisible decays of a Higgs boson produced through vector boson fusion in proton-proton collisions at $\sqrt{s} = 13$ TeV," *Phys. Lett.* **B793** (2019) 520–551, arXiv:1809.05937 [hep-ex].

[79] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, "FeynRules 2.0 - A complete toolbox for tree-level phenomenology," *Comput. Phys. Commun.* **185** (2014) 2250–2300, arXiv:1310.1921 [hep-ph].

[80] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, "UFO - The Universal FeynRules Output," *Comput. Phys. Commun.* **183** (2012) 1201–1214, arXiv:1108.2040 [hep-ph].

[81] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP* **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[82] T. Sjostrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An Introduction to PYTHIA 8.2," *Comput. Phys. Commun.* **191** (2015) 159–177, arXiv:1410.3012 [hep-ph].

[83] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz, and F. Siegert, "Rivet user manual," *Comput. Phys. Commun.* **184** (2013) 2803–2819, arXiv:1003.0694 [hep-ph].

[84] A. van Hameren, "OneLOop: For the evaluation of one-loop scalar functions," *Comput. Phys. Commun.* **182** (2011) 2427–2438, arXiv:1007.4716 [hep-ph].

[85] K. Hamilton, P. Nason, and G. Zanderighi, "MINLO: Multi-Scale Improved NLO," *JHEP* **10** (2012) 155, arXiv:1206.3572 [hep-ph].

[86] R. Boughezal, F. Caola, K. Melnikov, F. Petriello, and M. Schulze, "Higgs boson production in association with a jet at next-to-next-to-leading order," *Phys. Rev. Lett.* **115** no. 8, (2015) 082003, arXiv:1504.07922 [hep-ph].

[87] R. Boughezal, C. Focke, W. Giele, X. Liu, and F. Petriello, "Higgs boson production in association with a jet at NNLO using jettiness subtraction," *Phys. Lett.* **B748** (2015) 5–8, arXiv:1505.03893 [hep-ph].

[88] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, and A. Thamm, "Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson," *JHEP* **05** (2013) 022, arXiv:1211.3736 [hep-ph].

[89] K. Pearson F.R.S., "X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **50** no. 302, (1900) 157–175.

[90] **ATLAS** Collaboration, G. Aad *et al.*, "Search for Scalar Diphoton Resonances in the Mass Range $65 - 600$ GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8$ *TeV*," *Phys. Rev. Lett.* **113** no. 17, (2014) 171801, arXiv:1407.6583 [hep-ex].

[91] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for diphoton resonances in the mass range from 150 to 850 GeV in pp collisions at $\sqrt{s} = 8$ TeV," *Phys. Lett.* **B750** (2015) 494–519, arXiv:1506.02301 [hep-ex].

[92] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for neutral MSSM Higgs bosons decaying to $\mu^+\mu^-$ in pp collisions at $\sqrt{s} = 7$ and 8 TeV," *Phys. Lett.* **B752** (2016) 221–246, arXiv:1508.01437 [hep-ex].

[93] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Search for Higgs boson pair production in the $bb\tau\tau$ final state in proton-proton collisions at $\sqrt{(s)} = 8$ TeV," *Phys. Rev.* **D96** no. 7, (2017) 072004, arXiv:1707.00350 [hep-ex].

[94] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for two Higgs bosons in final states containing two photons and two bottom quarks in proton-proton collisions at 8 TeV," *Phys. Rev.* **D94** no. 5, (2016) 052012, arXiv:1603.06896 [hep-ex].

[95] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs

in proton-proton collisions at 8 TeV," *Phys. Lett.* **B749** (2015) 560–582, arXiv:1503.04114 [hep-ex].

[96] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for heavy ZZ resonances in the $\ell^+\ell^-\ell^+\ell^-$ and $\ell^+\ell^-\nu\bar{\nu}$ final states using proton-proton collisions at $\sqrt{s} = 13$TeV with the ATLAS detector," *Eur. Phys. J.* **C78** no. 4, (2018) 293, arXiv:1712.06386 [hep-ex].

[97] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for heavy resonances decaying into WW in the $e\nu\mu\nu$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Eur. Phys. J.* **C78** no. 1, (2018) 24, arXiv:1710.01123 [hep-ex].

[98] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for new phenomena in high-mass diphoton final states using 37 fb^{-1} of proton–proton collisions collected at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Phys. Lett.* **B775** (2017) 105–125, arXiv:1707.04147 [hep-ex].

[99] **ATLAS** Collaboration, M. Aaboud *et al.*, "Searches for the $Z\gamma$ decay mode of the Higgs boson and for new high-mass resonances in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP* **10** (2017) 112, arXiv:1708.00212 [hep-ex].

[100] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in 36 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP* **01** (2018) 055, arXiv:1709.07242 [hep-ex].

[101] **ATLAS** Collaboration, "Combination of searches for Higgs boson pairs in pp collisions at 13 TeV with the ATLAS experiment.," Tech. Rep. ATLAS-CONF-2018-043, CERN, Geneva, Sep, 2018.

[102] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for Higgs boson pair production in the $\gamma\gamma WW^*$ channel using pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Eur. Phys. J.* **C78** no. 12, (2018) 1007, arXiv:1807.08567 [hep-ex].

[103] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Search for a new scalar resonance decaying to a pair of Z bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV," *JHEP* **06** (2018) 127, arXiv:1804.01939 [hep-ex].

[104] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV," arXiv:1803.06553 [hep-ex].

[105] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Search for beyond the standard model Higgs bosons decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV," arXiv:1805.12191 [hep-ex].

[106] **CMS Collaboration** Collaboration, "Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV," Tech. Rep. CMS-PAS-HIG-17-030, CERN, Geneva, 2018.

[107] A. Azatov, R. Contino, and J. Galloway, "Model-Independent Bounds on a Light Higgs," *JHEP* **04** (2012) 127, arXiv:1202.3415 [hep-ph]. [Erratum: JHEP04,140(2013)].

[108] S. von Buddenbrock, "Di-Higgs production and the Madala Hypothesis," in *Proceedings of the HEPP Workshop 2018*, A. Cornell and B. Mellado, eds. https://indico.cern.ch/event/682259/. Accepted for publication.

[109] K. Hamilton, P. Nason, and G. Zanderighi, "Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator," *JHEP* **05** (2015) 140, arXiv:1501.04637 [hep-ph].

[110] **ATLAS** Collaboration, M. Aaboud *et al.*, "Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Phys. Rev.* **D98** (2018) 052005, arXiv:1802.04146 [hep-ex].

[111] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of inclusive and differential Higgs boson production cross sections in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV," arXiv:1807.03825 [hep-ex].

[112] **ATLAS** Collaboration, M. Aaboud *et al.*, "Measurement of inclusive and differential cross sections in the $H \to ZZ^* \to 4\ell$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP* **10** (2017) 132, arXiv:1708.02810 [hep-ex].

[113] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV," *JHEP* **11** (2017) 047, arXiv:1706.09936 [hep-ex].

[114] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of fiducial differential cross sections of gluon-fusion production of Higgs bosons decaying to $WW^* \to e\nu\mu\nu$ with the ATLAS detector at $\sqrt{s} = 8$ TeV," *JHEP* **08** (2016) 104, arXiv:1604.02997 [hep-ex].

[115] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of the transverse momentum spectrum of the Higgs boson produced in pp collisions at $\sqrt{s} = 8$ TeV using $H \to WW$ decays," *JHEP* **03** (2017) 032, arXiv:1606.01522 [hep-ex].

[116] **ATLAS** Collaboration, M. Aaboud *et al.*, "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector," *Phys. Lett.* **B784** (2018) 173–191, arXiv:1806.00425 [hep-ex].

[117] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Observation of $t\bar{t}H$ production," *Phys. Rev. Lett.* **120** no. 23, (2018) 231801, arXiv:1804.02610 [hep-ex].

[118] S. von Buddenbrock, "Exploring LHC Run 1 and 2 data using the Madala hypothesis," *J. Phys. Conf. Ser.* **878** no. 1, (2017) 012030, arXiv:1706.02477 [hep-ph].

[119] S. von Buddenbrock, "Production of the Madala boson in association with top quarks," in *The Proceedings of SAIP2017, the 62nd Annual Conference of the*

South African Institute of Physics, J. Engelbrecht, ed., pp. 187–192. ISBN: 978-0-620-82077-6.

[120] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for Higgs boson production in association with top quarks in multilepton final states at $\sqrt{s} = 13$ TeV,". CMS-PAS-HIG-17-004.

[121] **ATLAS** Collaboration, G. Aad *et al.*, "Search for the Associated Production of a Higgs Boson and a Top Quark Pair in Multilepton Final States with the ATLAS Detector,". ATLAS-CONF-2016-058.

[122] **ATLAS** Collaboration, G. Aad *et al.*, "Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector," arXiv:1906.02025 [hep-ex].

[123] P. Rapheeha, X. Ruan, and B. Mellado, "Compatibility of a simplified BSM model with the observed excesses in multi-lepton production at the LHC," in *The Proceedings of SAIP2019, the 64th Annual Conference of the South African Institute of Physics*. Submitted for publication.

[124] **ATLAS** Collaboration, "Combined measurement of the total and differential cross sections in the $H \to \gamma\gamma$ and the $H \to ZZ^* \to 4\ell$ decay channels at $\sqrt{s}$ =13 TeV with the ATLAS detector,". ATLAS-CONF-2019-032.

[125] **CMS** Collaboration, "Measurements of properties of the Higgs boson in the four-lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV,". CMS-PAS-HIG-19-001.

[126] S. von Buddenbrock, A. S. Cornell, M. Kumar, and B. Mellado, "The Madala hypothesis with Run 1 and 2 data at the LHC," *J. Phys. Conf. Ser.* **889** no. 1, (2017) 012020, arXiv:1709.09419 [hep-ph].

[127] **ATLAS** Collaboration, G. Aad *et al.*, "Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment," *Phys. Rev.* **D101** no. 1, (2020) 012002, arXiv:1909.02845 [hep-ex].

[128] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, "Theory and phenomenology of two-Higgs-doublet models," *Phys. Rept.* **516** (2012) 1–102, arXiv:1106.0034 [hep-ph].

[129] M. E. Peskin and T. Takeuchi, "Estimation of oblique electroweak corrections," *Phys. Rev.* **D46** (1992) 381–409.

[130] J. Erler and M. Schott, "Electroweak Precision Tests of the Standard Model after the Discovery of the Higgs Boson," *Prog. Part. Nucl. Phys.* **106** (2019) 68–119, arXiv:1902.05142 [hep-ph].

[131] **CLEO** Collaboration, S. Anderson *et al.*, "Improved upper limits on the flavor-changing neutral current decays $B \to K\ell^+\ell^-$ and $B \to k^*(892)\ell^+\ell^-$," *Phys. Rev. Lett.* **87** (Oct, 2001) 181803.

[132] B. Coleppa, F. Kling, and S. Su, "Charged Higgs search via $AW^\pm/HW^\pm$ channel," *JHEP* **12** (2014) 148, arXiv:1408.4119 [hep-ph].

[133] B. Coleppa, F. Kling, and S. Su, "Constraining Type II 2HDM in Light of LHC Higgs Searches," *JHEP* **01** (2014) 161, arXiv:1305.0002 [hep-ph].

[134] B. Coleppa, F. Kling, and S. Su, "Exotic Higgs Decay via AZ/HZ Channel: a Snowmass Whitepaper," arXiv:1308.6201 [hep-ph].

[135] A. Drozd, B. Grzadkowski, J. F. Gunion, and Y. Jiang, "Extending two-Higgs-doublet models by a singlet scalar field - the Case for Dark Matter," *JHEP* **11** (2014) 105, arXiv:1408.2106 [hep-ph].

[136] D. Sabatta, A. S. Cornell, A. Goyal, M. Kumar, B. Mellado, and X. Ruan, "Connecting the Muon Anomalous Magnetic Moment and the Multi-lepton Anomalies at the LHC," *Chin. Phys. C* (2020) , arXiv:1909.03969 [hep-ph].

[137] C.-Y. Chen, M. Freid, and M. Sher, "Next-to-minimal two Higgs doublet model," *Phys. Rev.* **D89** no. 7, (2014) 075009, arXiv:1312.3949 [hep-ph].

[138] V. Barger, P. Langacker, M. McCaskey, M. Ramsey-Musolf, and G. Shaughnessy, "Complex Singlet Extension of the Standard Model," *Phys. Rev.* **D79** (2009) 015018, arXiv:0811.0393 [hep-ph].

[139] S. von Buddenbrock, A. S. Cornell, E. D. R. Iarilala, M. Kumar, B. Mellado, X. Ruan, and E. M. Shrif, "Constraints on a 2HDM with a singlet scalar and implications in the search for heavy bosons at the LHC," *J. Phys.* **G46** no. 11, (2019) , arXiv:1809.06344 [hep-ph].

[140] **DELPHES 3** Collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaitre, A. Mertens, and M. Selvaggi, "DELPHES 3, A modular framework for fast simulation of a generic collider experiment," *JHEP* **02** (2014) 057, arXiv:1307.6346 [hep-ex].

[141] M. Dittmar and H. K. Dreiner, "How to find a Higgs boson with a mass between 155-GeV - 180-GeV at the LHC," *Phys. Rev.* **D55** (1997) 167–172, arXiv:hep-ph/9608317 [hep-ph].

[142] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the properties of a Higgs boson in the four-lepton final state," *Phys. Rev.* **D89** no. 9, (2014) 092007, arXiv:1312.5353 [hep-ex].

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

[144] A. Lazopoulos, T. McElmurry, K. Melnikov, and F. Petriello, "Next-to-leading order QCD corrections to $t\bar{t}Z$ production at the LHC," *Phys. Lett.* **B666** (2008) 62–65, arXiv:0804.2220 [hep-ph].

[145] G. Bevilacqua and M. Worek, "Constraining BSM Physics at the LHC: Four top final states with NLO accuracy in perturbative QCD," *JHEP* **07** (2012) 111, arXiv:1206.3064 [hep-ph].

[146] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp

collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *Phys. Rev.* **D99** no. 5, (2019) 052009, arXiv:1811.02305 [hep-ex].

[147] V. D. Barger and R. J. N. Phillips, "Trileptons From the Top," *Phys. Rev.* **D30** (1984) 1890.

[148] V. E. Ozcan, S. Sultansoy, and G. Unel, "Possible Discovery Channel for New Charged Leptons at the LHC," *J. Phys.* **G36** (2009) 095002, arXiv:0903.3177 [hep-ex]. [Erratum: J. Phys.G37,059801(2010)].

[149] B. Mukhopadhyaya and S. Mukhopadhyay, "Same-sign trileptons and four-leptons as signatures of new physics at the CERN Large Hadron Collider," *Phys. Rev.* **D82** (2010) 031501, arXiv:1005.3051 [hep-ph].

[150] S. Mukhopadhyay and B. Mukhopadhyaya, "Same-sign trileptons at the LHC: A Window to lepton-number violating supersymmetry," *Phys. Rev.* **D84** (2011) 095001, arXiv:1108.4921 [hep-ph].

[151] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," *JHEP* **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[152] **ATLAS** Collaboration, "A search for Supersymmetry in events containing a leptonically decaying Z boson, jets and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector," Tech. Rep. ATLAS-CONF-2015-082, CERN, Geneva, Dec, 2015.

[153] **ATLAS** Collaboration, G. Aad *et al.*, "Search for supersymmetry in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector," *Eur. Phys. J.* **C75** no. 7, (2015) 318, arXiv:1503.03290 [hep-ex]. [Erratum: Eur. Phys. J.C75,no.10,463(2015)].

[154] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for Physics Beyond the Standard Model in Events with Two Leptons, Jets, and Missing Transverse Momentum in pp Collisions at sqrt(s) = 8 TeV," *JHEP* **04** (2015) 124, arXiv:1502.06031 [hep-ex].

[155] **ATLAS** Collaboration, G. Aad *et al.*, "Search for a CP-odd Higgs boson decaying to Zh in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Phys. Lett.* **B744** (2015) 163–183, arXiv:1502.04478 [hep-ex].

[156] **ATLAS** Collaboration, "Search for a CP-odd Higgs boson decaying to Zh in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATLAS-CONF-2016-015, CERN, Geneva, Mar, 2016.

[157] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for a pseudoscalar boson decaying into a Z boson and the 125 GeV Higgs boson in $\ell^+\ell^- b\bar{b}$ final states," *Phys. Lett.* **B748** (2015) 221–243, arXiv:1504.04710 [hep-ex].

[158] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for dark matter, extra dimensions, and unparticles in monojet events in proton-proton collisions at $\sqrt{s} = 8$ TeV," *Eur. Phys. J.* **C75** no. 5, (2015) 235, arXiv:1408.3583 [hep-ex].

[159] **ATLAS** Collaboration, "Search for Dark Matter production associated with bottom quarks with 13.3 fb-1 of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC,".

[160] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for dark matter in association with a Higgs boson decaying to two photons at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *Phys. Rev.* **D96** no. 11, (2017) 112004, arXiv:1706.03948 [hep-ex].

[161] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for new phenomena in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in $\sqrt{s} = 13$ pp collisions with the ATLAS detector," *Eur. Phys. J.* **C77** no. 3, (2017) 144, arXiv:1611.05791 [hep-ex].

[162] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for Higgs boson pair production in the $WW^{(*)}WW^{(*)}$ decay channel using ATLAS data recorded at $\sqrt{s} = 13$ TeV," *JHEP* **05** (2019) 124, arXiv:1811.11028 [hep-ex].

[163] **ATLAS** Collaboration, "Measurements of the Higgs boson inclusive, differential and production cross sections in the 4ℓ decay channel at $\sqrt{s}$ = 13 TeV with the ATLAS detector,". ATLAS-CONF-2019-025.

[164] D. Dercks, N. Desai, J. S. Kim, K. Rolbiecki, J. Tattersall, and T. Weber, "CheckMATE 2: From the model to the limit," *Comput. Phys. Commun.* **221** (2017) 383–418, arXiv:1611.09856 [hep-ph].

[165] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual," *Eur. Phys. J.* **C72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[166] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, "HistFactory: A tool for creating statistical models for use with RooFit and RooStats," Tech. Rep. CERN-OPEN-2012-016, New York U., New York, Jan, 2012.

[167] **ATLAS** Collaboration, G. Aad *et al.*, "Analysis of events with b-jets and a pair of leptons of the same charge in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *JHEP* **10** (2015) 150, arXiv:1504.04605 [hep-ex].

[168] **ATLAS** Collaboration, M. Aaboud *et al.*, "Measurement of lepton differential distributions and the top quark mass in $t\bar{t}$ production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Eur. Phys. J.* **C77** no. 11, (2017) 804, arXiv:1709.09407 [hep-ex].

[169] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Search for associated production of a Higgs boson and a single top quark in proton-proton collisions at $\sqrt{s} =$ TeV," *Phys. Rev.* **D99** no. 9, (2019) 092005, arXiv:1811.09696 [hep-ex].

[170] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV," *JHEP* **10** (2018) 117, arXiv:1805.07399 [hep-ex].

[171] **CMS** Collaboration, "Measurements of the pp $\rightarrow$ WZ inclusive and differential production cross section and constraints on charged anomalous triple gauge couplings at $\sqrt{s} = 13$ TeV.," Tech. Rep. CMS-PAS-SMP-18-002, CERN, Geneva, 2018.

[172] **ATLAS** Collaboration, M. Aaboud *et al.*, "Search for new phenomena in events with same-charge leptons and b-jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP* **12** (2018) 039, arXiv:1807.11883 [hep-ex].

[173] **ATLAS** Collaboration, "Measurements of top-quark pair spin correlations in the $e\mu$ channel at $\sqrt{s} = 13$ TeV using pp collisions in the ATLAS detector," Tech. Rep. ATLAS-CONF-2018-027, CERN, Geneva, Jul, 2018.

[174] **ATLAS** Collaboration, "Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATLAS-CONF-2018-034, CERN, Geneva, 2018.

[175] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP* **06** (2010) 043, arXiv:1002.2581 [hep-ph].

[176] **CMS** Collaboration, "Measurements of differential cross sections for $\mathrm{t\bar{t}}$ production in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons," Tech. Rep. CMS-PAS-TOP-17-014, CERN, Geneva, 2018. Figure 35.

[177] **ATLAS** Collaboration, M. Aaboud *et al.*, "Measurements of gluon-gluon fusion and vector-boson fusion Higgs boson production cross-sections in the $H \to WW^* \to e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," arXiv:1808.09054 [hep-ex].

[178] A. Behring, M. Czakon, A. Mitov, R. Poncelet, and A. S. Papanastasiou, "Higher order corrections to spin correlations in top quark pair production at the lhc," *Phys. Rev. Lett.* **123** (Aug, 2019) 082001. https://link.aps.org/doi/10.1103/PhysRevLett.123.082001.

[179] B. Biedermann, A. Denner, and L. Hofer, "Next-to-leading-order electroweak corrections to the production of three charged leptons plus missing energy at the LHC," *JHEP* **10** (2017) 043, arXiv:1708.06938 [hep-ph].

[180] J. Baglio and N. Le Duc, "Fiducial polarization observables in hadronic WZ production: A next-to-leading order QCD+EW study," *JHEP* **04** (2019) 065, arXiv:1810.11034 [hep-ph].

[181] J. Bellm *et al.*, "Herwig 7.1 Release Note," arXiv:1705.06919 [hep-ph].

[182] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, "New parton distribution functions from a global analysis of quantum chromodynamics," *Phys. Rev.* **D93** no. 3, (2016) 033006, arXiv:1506.07443 [hep-ph].

[183] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs," *Eur. Phys. J.* **C75** no. 5, (2015) 204, arXiv:1412.3989 [hep-ph].

[184] J. Butterworth *et al.*, "PDF4LHC recommendations for LHC Run II," *J. Phys.* **G43** (2016) 023001, arXiv:1510.03865 [hep-ph].

[185] **ATLAS** Collaboration, M. Aaboud *et al.*, "Measurement of fiducial and differential W^+W^- production cross-sections at $\sqrt{s}$ =13 TeV with the ATLAS detector," arXiv:1905.04242 [hep-ex].

[186] T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhoefer, A. von Manteuffel, S. Pozzorini, D. Rathlev, and L. Tancredi, "W^+W^- Production at Hadron Colliders in Next to Next to Leading Order QCD," *Phys. Rev. Lett.* **113** no. 21, (2014) 212001, arXiv:1408.5243 [hep-ph].

[187] M. Grazzini, S. Kallweit, S. Pozzorini, D. Rathlev, and M. Wiesemann, "W^+W^- production at the LHC: fiducial cross sections and distributions in NNLO QCD," *JHEP* **08** (2016) 140, arXiv:1605.02716 [hep-ph].

[188] K. Hamilton, T. Melia, P. F. Monni, E. Re, and G. Zanderighi, "Merging WW and WW+jet with MINLO," *JHEP* **09** (2016) 057, arXiv:1606.07062 [hep-ph].

[189] E. Re, M. Wiesemann, and G. Zanderighi, "NNLOPS accurate predictions for W^+W^- production," *JHEP* **12** (2018) 121, arXiv:1805.09857 [hep-ph].

[190] F. Caola, K. Melnikov, R. Roentsch, and L. Tancredi, "QCD corrections to W^+W^- production through gluon fusion," *Phys. Lett.* **B754** (2016) 275–280, arXiv:1511.08617 [hep-ph].

[191] B. Biedermann, M. Billoni, A. Denner, S. Dittmaier, L. Hofer, B. Jaeger, and L. Salfelder, "Next-to-leading-order electroweak corrections to $pp \to W^+W^- \to 4$ leptons at the LHC," *JHEP* **06** (2016) 065, arXiv:1605.03419 [hep-ph].

[192] **ATLAS** Collaboration, G. Aad *et al.*, "The ATLAS Simulation Infrastructure," *Eur. Phys. J.* **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[193] **ATLAS** Collaboration, "Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach," Tech. Rep. ATL-PHYS-PUB-2017-022, CERN, Geneva, Dec, 2017.

[194] **ATLAS** Collaboration, G. Aad *et al.*, "Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV," *Eur. Phys. J.* **C76** no. 5, (2016) 292, arXiv:1603.05598 [hep-ex].

[195] M. Pivk and F. R. Le Diberder, "SPlot: A Statistical tool to unfold data distributions," *Nucl. Instrum. Meth.* **A555** (2005) 356–369, arXiv:physics/0402083 [physics.data-an].

[196] M. Oreglia *et al.*, "Study of the reaction $\psi' \to \gamma\gamma\frac{J}{\psi}$," *Phys. Rev. D* **25** (May, 1982) 2259–2277.

[197] F. Cascioli, S. Hoeche, F. Krauss, P. Maierhoefer, S. Pozzorini, and F. Siegert, "Precise Higgs-background predictions: merging NLO QCD and squared quark-loop corrections to four-lepton + 0,1 jet production," *JHEP* **01** (2014) 046, arXiv:1309.0500 [hep-ph].

[198] **ATLAS** Collaboration, "ATLAS Run 1 Pythia8 tunes," Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov, 2014.

[199] **ATLAS** Collaboration, G. Aad *et al.*, "Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector," *Phys. Lett.* **B796** (2019) 68–87, arXiv:1903.06248 [hep-ex].

[200] S. Hossenfelder, "How the LHC may spell the end of particle physics," 2018. http://backreaction.blogspot.com/2018/12/how-lhc-may-spell-end-of-particle.html. Last accessed: 20/09/2019.

[201] P. Golonka, B. Kersevan, T. Pierzchala, E. Richter-Was, Z. Was, and M. Worek, "The Tauola photos F environment for the TAUOLA and PHOTOS packages: Release. 2.," *Comput. Phys. Commun.* **174** (2006) 818–835, arXiv:hep-ph/0312240 [hep-ph].

[202] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

[203] S. H. Abidi *et al.*, "Search for heavy ZZ resonances in the $\ell\ell\ell\ell$ final state using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-ANA-HIGG-2018-09, CERN, Geneva, Sep, 2018.

[204] X. Ruan *et al.*, "Search for a Higgs boson with the di-photon 2 decay in association with missing energy in p p collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-1732, CERN, Geneva, Nov, 2016.

[205] S. von Buddenbrock *et al.*, "Search for resonant pair production in $H \to SS, Sh, hh$ channels decaying to 4W with a four lepton final state at $\sqrt{s} = 13$ TeV in the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-1622, CERN, Geneva, Feb, 2017.

[206] L. Cerda Alberich *et al.*, "Supporting note: Selection and performance for the $h \to \gamma\gamma$ and $h \to Z\gamma$ analyses," Tech. Rep. ATL-COM-PHYS-2016-862, CERN, Geneva, Jul, 2016.

[207] S. von Buddenbrock *et al.*, "Supporting note: Search for a generic heavy boson using $\gamma\gamma + X$ in p p collisions at$\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-473, CERN, Geneva, May, 2016.

[208] S. von Buddenbrock *et al.*, "Supporting note: Search for dark matter in events with missing transverse momentum and a Higgs boson decaying to two photons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-708, CERN, Geneva, Jun, 2016.

[209] E. Petit *et al.*, "Supporting note: Selection and performance for the $h \to \gamma\gamma$ and $h \to Z\gamma$ analyses," Tech. Rep. ATL-COM-PHYS-2016-109, CERN, Geneva, Feb, 2016.

[210] X. Ruan *et al.*, "Search for a Higgs boson with the di-photon decay in association with missing energy in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2015-1589, CERN, Geneva, Dec, 2015.

[211] J. Adelman *et al.*, "Search for resonant and enhanced non-resonant dihiggs production in the yybb channel with XXX fb-1 of data at 13 TeV," Tech. Rep. ATL-COM-PHYS-2015-1073, CERN, Geneva, Aug, 2015.

[212] S. von Buddenbrock *et al.*, "Supporting note: Search for dark matter in events with missing transverse momentum and a Higgs boson decaying to two photons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector using data in full years of 2015 and 2016," Tech. Rep. ATL-COM-PHYS-2016-1735, CERN, Geneva, Nov, 2016.

[213] J. Adelman *et al.*, "Supporting note: Selection and performance for the $h \to \gamma\gamma$ and $h \to Z\gamma$ analyses, Spring 2017," Tech. Rep. ATL-COM-PHYS-2017-357, CERN, Geneva, Apr, 2017.

[214] Y. Fang *et al.*, "Search for Higgs boson pair production in the final state of $\gamma\gamma WW^*(\to lvjj)$ using 36.1 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2016-1406, CERN, Geneva, Sep, 2016.