Search for Higgs boson decays to beyond-the-Standard-Model light bosons in four-lepton events with the ATLAS detector at the LHC

by

Justin Chiu
B.Sc., University of Victoria, 2014
M.Sc., University of Victoria, 2016

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Physics and Astronomy

Search for Higgs boson decays to beyond-the-Standard-Model light bosons in four-lepton events with the ATLAS detector at the LHC

by

Justin Chiu
B.Sc., University of Victoria, 2014
M.Sc., University of Victoria, 2016

Supervisory Committee

Dr. Michel Lefebvre, Supervisor
(Department of Physics and Astronomy)

Dr. Robert McPherson, Departmental Member
(Department of Physics and Astronomy)

Dr. Ian Putnam, Additional Member
(Department of Mathematics and Statistics)

Supervisory Committee

Dr. Michel Lefebvre, Supervisor
(Department of Physics and Astronomy)

Dr. Robert McPherson, Departmental Member
(Department of Physics and Astronomy)

Dr. Ian Putnam, Additional Member
(Department of Mathematics and Statistics)

ABSTRACT

This thesis presents the search for the dark sector process $h \to Z_d Z_d \to 4\ell$ in events collected by the ATLAS detector at the Large Hadron Collider in 2015–2018. In this theorized process, the Standard Model Higgs boson (h) decays to four leptons (4ℓ) via two intermediate Beyond-the-Standard-Model particles each called Z_d. This process arises from interactions of the Standard Model with a *dark sector*. A dark sector consists of one or more new particles that have limited or zero interaction with the Standard Model, such as the new vector boson Z_d (*dark photon*). It could have a rich and interesting phenomenology like the visible sector (the Standard Model) and could naturally address many outstanding problems in particle physics. For example, it could contain a particle candidate for dark matter. In particular, Higgs decays to Beyond-the-Standard-Model particles are well-motivated theoretically and are not tightly constrained; current measurements of Standard Model Higgs properties permit the fraction of such decays to be as high as approximately 30%. The results of this search do not show evidence for the existence of the $h \to Z_d Z_d \to 4\ell$ process and are therefore interpreted in terms of upper limits on the branching ratio $\mathcal{B}(h \to Z_d Z_d)$ and the effective Higgs mixing parameter κ'.

Contents

Supervisory Committee **ii**

Abstract **iii**

Table of Contents **iv**

List of Tables **viii**

List of Figures **ix**

List of Abbreviations **xii**

Declaration **xiv**

Acknowledgements **xv**

Dedication **xvii**

1 Introduction **1**

2 Motivation **3**

3 Theory **8**

4 LHC and ATLAS **18**

4.1 Large Hadron Collider 18
4.2 ATLAS detector 21
4.2.1 Overview 21
4.2.2 Inner Detector 22
4.2.3 Calorimeter 24

4.2.4 Muon Spectrometer 29
4.2.5 Trigger and data acquisition system 33

5 $h \to Z_d Z_d \to 4\ell$ search strategy 35
5.1 Channels 37
5.2 Search strategy 37
5.3 Event pre-selection 40
5.4 Selection of electrons and muons 41
5.4.1 Electrons 41
5.4.1.1 Electron reconstruction 42
5.4.1.2 Electron calibration 44
5.4.1.3 Electron identification 45
5.4.1.4 Electron isolation 46
5.4.1.5 Kinematic requirements on electrons 48
5.4.2 Muons 49
5.4.2.1 Muon reconstruction 49
5.4.2.2 Muon calibration 51
5.4.2.3 Muon identification 52
5.4.2.4 Muon isolation 53
5.4.2.5 Kinematic requirements on muons 53
5.5 Overlap removal 55
5.6 Ranking and selection of quadruplets 55
5.7 Event selection 57

6 Data and Monte Carlo-based signal and background estimation 63
6.1 Data 63
6.2 Monte Carlo-based signal and background estimation 65
6.2.1 Signal simulation 67
6.2.2 Background simulation 71
6.2.2.1 $h \to ZZ^* \to 4\ell$ 71
6.2.2.2 $ZZ^* \to 4\ell$ 72
6.2.2.3 Tri-boson, higher-order electroweak, and $t\bar{t} + Z$ . . . 72
6.2.2.4 Reducible backgrounds 72
6.2.2.5 Pileup treatment 76
6.2.3 Estimate of statistical and systematic uncertainties 76

6.2.3.1 Experimental systematic uncertainties 77
6.2.3.2 Theoretical systematic uncertainties 78

7 Data-driven fake background estimation **83**
7.1 Introduction 83
7.2 Methodology 85
7.3 Fake factors 89
7.4 Fake estimates 93
7.4.1 Results 93
7.4.2 Systematic uncertainty 96
7.4.3 Closure test 97

8 Results **98**
8.1 Validation of background estimates 98
8.1.1 Overview and yields 98
8.1.2 Validation region 1 101
8.1.3 Validation region 2 101
8.1.4 Validation region 3 105
8.1.5 Validation region 4 105
8.1.6 Validation region 5 109
8.2 Unblinding of signal region 113

9 Interpretation **120**
9.1 Methodology 120
9.1.1 Hypothesis testing and p-value 120
9.1.2 Likelihood construction 122
9.1.3 Test statistic for discovery 123
9.1.4 Test statistic for upper limit on μ 126
9.1.5 $\mathrm{CL_s}$ method for upper limits 127
9.1.6 Toy-based construction of test statistic PDFs 128
9.2 Implementation 129
9.3 Transformation of signal strength limit 131
9.4 Fitting results 132
9.5 p_0 and significance scan 136
9.6 Upper limits 137

10 Conclusions **141**

A 2016–2018 CMS result **143**

Bibliography **147**

List of Tables

4.1 Triggers used in the $h \to Z_d Z_d \to 4\ell$ search with 2015–2018 data and simulation . . . 34

5.1 Selection criteria of the $h \to Z_d Z_d \to 4\ell$ search . . . 39

6.1 Signal region cutflow for $m_{Z_d} = 25$ GeV signal in 2015–2018 simulation 70
6.2 $h \to ZZ^* \to 4\ell$ simulation samples . . . 73
6.3 Non-resonant $ZZ^* \to 4\ell$ simulation samples . . . 73
6.4 Tri-boson, higher-order electroweak, and $t\bar{t} + Z$ simulation samples . 73
6.5 Reducible background simulation samples . . . 73
6.6 Signal region cutflow for $h \to ZZ^* \to 4\ell$ background in 2015–2018 simulation . . . 74
6.7 Signal region cutflow for $ZZ^* \to 4\ell$ background in 2015–2018 simulation 75
6.8 Relative experimental (theoretical) systematic uncertainties on signal region yield (acceptance) for $h \to Z_d Z_d \to 4\ell$ signal . . . 81
6.9 Relative systematic uncertainties on signal region yield for $h \to ZZ^* \to 4\ell$ and $ZZ^* \to 4\ell$ backgrounds . . . 82

7.1 Z+jets estimates with the fake factor method . . . 95

8.1 Comparison of validation region selection criteria . . . 100
8.2 Expected background yields and observed yields in the validation regions per channel . . . 100
8.3 Expected background yield *predicted from simulation* and observed yield in the signal region per data-taking campaign . . . 114
8.4 Expected background yield per process and observed yield in the signal region in 2015–2018 . . . 114
8.5 Signal region cutflow for 2015–2018 data . . . 117

List of Figures

2.1 Distributions of dilepton mass in searches by ATLAS with 2010–2012 and 2015–2016 data and simulation 6

3.1 Feynman diagrams of $h \to ZZ_d \to 4\ell$ and $h \to Z_dZ_d \to 4\ell$ 9
3.2 Comparison of upper limits set on kinetic mixing parameter ϵ versus dark photon mass 12
3.3 Branching ratio of $Z_d \to \ell\ell$ predicted from theory 14
3.4 Branching ratios of $h \to Z_{(d)}Z_d \to 4\ell$ predicted from theory 15
3.5 Upper limit on the effective Higgs mixing parameter κ' set by ATLAS with 2015–2016 data 16
3.6 Expected upper limits on the effective Higgs mixing parameter κ' from experiments at the LHC 16
3.7 Upper limit on the branching ratio $\mathcal{B}(h \to Z_dZ_d)$ set by ATLAS with 2015–2016 data 17
3.8 Expected upper limits on the branching ratio $\mathcal{B}(h \to Z_dZ_d)$ from experiments at the LHC 17

4.1 Distribution of μ during 2015–2018 pp collision data-taking 20
4.2 Cut-away view of the ATLAS detector 22
4.3 Structure of the LAr electromagnetic barrel calorimeter 27
4.4 Schematic of a portion of the LAr electromagnetic barrel calorimeter 27
4.5 Schematic of a portion of the tile hadronic calorimeter 28
4.6 Layout of the Muon Spectrometer 30
4.7 Cross-section of a monitored drift tube 32
4.8 Monitored drift tube 32
4.9 Cathode strip chamber 32

5.1 Electron identification efficiencies in $Z \to ee$ events in 2015–2017 data and simulation . . . 47
5.2 m_{ee} distribution in $Z \to ee$ events in 2015–2017 data and simulation . 47
5.3 Muon identification efficiencies in $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events in early 2015 data and simulation . . . 54
5.4 $m_{\mu\mu}$ distribution in $Z \to \mu\mu$ events with combined muons in early 2015 data and simulation . . . 54
5.5 Expected discovery significance for $m_{Z_d} = 60$ GeV signal . . . 59
5.6 Original medium signal region and addition from re-optimization . . . 62
5.7 Modulating function $F(m_{12})$ of re-optimized signal region and exponential-Gaussian background model $F(\langle m_{\ell\ell} \rangle)$ . . . 62

6.1 Total integrated luminosity delivered to and recorded and certified by ATLAS in 2015–2018 . . . 64
6.2 Generator-level distributions of $h \to Z_d Z_d \to 4\ell$ signal . . . 69

7.1 Categorization of leptons into regions of the fake factor method . . . 87
7.2 Electron fake factors and fake rates per data-taking campaign with contamination correction . . . 90
7.3 Muon fake factors and fake rates per data-taking campaign with contamination correction . . . 91
7.4 Electron and muon fake factors combined across data-taking campaigns with and without contamination correction . . . 92

8.1 Distributions of $\langle m_{\ell\ell} \rangle$ in validation region 1 in 2015–2018 data and simulation . . . 102
8.2 Distributions of $(m_{14} + m_{32})/2$ in validation region 1 in 2015–2018 data and simulation . . . 103
8.3 Distributions of the mass of the dilepton pair with mass closest to m_Z in validation region 1 in 2015–2018 data and simulation . . . 104
8.4 Distributions of $\langle m_{\ell\ell} \rangle$ in validation region 2 in 2015–2018 data and simulation . . . 106
8.5 Distributions of $m_{4\ell}$ in validation region 3 in 2015–2018 data and simulation . . . 107
8.6 Distributions of $\langle m_{\ell\ell} \rangle$ in validation region 3 in 2015–2018 data and simulation . . . 108

8.7 Distributions of $\langle m_{\ell\ell} \rangle$ in validation region 4 in 2015–2018 data and simulation . . . 110
8.8 Distributions of $\langle m_{\ell\ell} \rangle$ in validation region 5 in 2015–2018 data and simulation . . . 111
8.9 Distributions of $m_{4\ell}$ in validation region 5 in 2015–2018 data and simulation . . . 112
8.10 Distributions of $\langle m_{\ell\ell} \rangle$ in the signal region in 2015–2018 data and simulation with 5 GeV–wide bins . . . 115
8.11 Distribution of $\langle m_{\ell\ell} \rangle$ in the signal region in 2015–2018 data and simulation with all channels combined with 2.5 GeV–wide bins . . . 116
8.12 m_{34} vs. m_{12} distribution of events observed in the signal region . . . 118
8.13 m_{34} vs. m_{12} distribution of events observed passing requirements up to and including the Higgs window requirement . . . 119

9.1 Pre- and post-fit $\langle m_{\ell\ell} \rangle$ distributions for the $m_{Z_d} = 20$ GeV signal hypothesis . . . 134
9.2 Pull and ranking plot for the $m_{Z_d} = 35$ GeV signal hypothesis . . . 135
9.3 Local p_0 value vs. m_{Z_d} . . . 136
9.4 p-value vs. μ for signal hypotheses $m_{Z_d} = 18, 21, 23$ GeV . . . 138
9.5 Upper limit on the branching ratio $\mathcal{B}(h \to Z_d Z_d)$ . . . 139
9.6 Acceptance times efficiency factor vs. m_{Z_d} . . . 140
9.7 Upper limit on the effective Higgs mixing parameter κ' . . . 140

A.1 Distributions of $\langle m_{\ell\ell} \rangle$ in the signal region of the CMS 137 fb^{-1} search 145
A.2 Limits on the branching ratio and Higgs mixing parameter set by the CMS 137 fb^{-1} search . . . 146

List of Abbreviations

ATLAS A Toroidal LHC ApparatuS

BSM Beyond-the-Standard-Model

CMS Compact Muon Solenoid

CP Charge-Parity

CSC cathode strip chamber

EM electromagnetic

GRL Good Runs List

HAHM Hidden Abelian Higgs Model

HLT High-Level Trigger

ID Inner Detector

L1 Level-1

LAr liquid argon

LHC Large Hadron Collider

LO leading order

MC Monte Carlo

MDT monitored drift tube

MLE maximum likelihood estimate

MPI multiple parton interaction

MS Muon Spectrometer

NLO next-to-leading order

NNLO next-to-next-to-leading order

PDF probability distribution function

PDF parton distribution function

QCD quantum chromodynamics

QED quantum electrodynamics

RPC resistive plate chamber

SCT Semiconductor Tracker

SM Standard Model

TDAQ trigger and data acquisition

TGC thin gap chamber

TRT Transition Radiation Tracker

DECLARATION

The results of the ATLAS Experiment are made possible by the collective efforts of the ATLAS Collaboration. The results described in this thesis are no exception. This Declaration characterizes the contributions of the author to the analysis described in this thesis and his role within the analysis group and Collaboration.

The author is the lead analyst for the $h \to Z_d Z_d \to 4\ell$ channel (*high-mass*). This channel is the focus of this thesis. The author developed from the ground up the software needed to perform the analysis of this channel on the common datasets. This software was later adapted for the ZX channel. For the high-mass channel, one of the chief contributions of the author was the improvement of the fake background estimation with a data-driven technique (described in Chapter 7). All of the results described in this thesis—mostly contained within Chapters 7, 8, and 9—were produced by the author, except if noted otherwise. These results include the distributions of the validation and signal regions, the estimates of the systematic uncertainties, and the interpretation of the signal region results in terms of limits.

The author served as co-analysis contact for the $h \to XX/ZX \to 4\ell$ analysis group between October 2018 and May 2020. This role involved leading the analysis group of twenty-seven scientists and five analysis channels and facilitating the interaction between the group and the wider Collaboration. This analysis group belongs to the Higgs and Light Resonances Searches sub-group within the Higgs and Diboson Searches group. The software used to produce the common datasets for the five channels was originally developed by Dr. Will Buttinger; after his departure from the analysis group, the software was further developed and maintained by the author. The common datasets were produced and maintaned by the author. The author led the effort to internally document the five channels for review by the Collaboration and contributed the common and high-mass portions. The author organized the group's responses to questions raised by the Collaboration. The author has contributed the high-mass results to a three-channel journal publication that is currently in progress and the author is assisting with its ongoing review.

The author provided comments and software assistance for the other four channels, but did not contribute significantly. The high-mass channel—like the other channels and other ATLAS analyses—used data and simulation datasets, calibrations, and physics prescriptions (e.g. standard event cleaning cuts, object cuts) produced and defined by other members of the Collaboration.

ACKNOWLEDGEMENTS

I would not have achieved this and would not be the person I am today without the support and guidance of my colleagues and friends. Thank you, from the bottom of my heart.

First, I want to thank my supervisor, Prof. Michel Lefebvre, for his direction, expertise, and confidence. You have provided me with countless opportunities and have taught me more than physics. I have learned leadership, patience, and perseverance. I will carry what I've learned from you forward into the future.

Along the way, Dr. Kenji Hamano has been incredibly helpful and generous with his time. Thank you for sharing your analysis and publication expertise, your assistance with the documentation, and for continuing to help field questions from me and the Collaboration.

My work builds on the brilliance of Dr. Will Buttinger, who entrusted me with this analysis in 2017 and gave me the opportunity to take a leading role. His software, physics, and statistics expertise has been invaluable, and he has always found time to help, even after his departure from the analysis team. Thank you.

It has been a pleasure to work with the analysis team. I want to thank Diallo Boye for his help as co-analysis contact, and the others on the team for sharing their knowledge of the analysis and providing comments on my work.

My time at the university has been an incredible transformative journey, in no small part due to my friends. They have given me unflagging support, provided me with warm and precious company, and helped me evolve:

Kayla McLean and Nicholas Lange: thank you both for your help and support, going back to my undergraduate years. I can't believe it's been ten years! I remember us on the fourth floor and in classes in Elliot, and hurrying from class to class in the quad. Thanks for being my *original* friends and showing me what that is all about.

Tony Kwan: thank you for your wisdom, patience, and understanding. Your counsel has been invaluable. I wonder today how, years down the road, we will look back and how things will have changed. There will be ups and downs but hopefully we will find what fulfills us, in due course. I really hope.

Danika MacDonell: thank you for your delightful company, friendship, and thoughtfulness, and for introducing the outdoors to me! I've really enjoyed our time

running, brunching, hiking, traveling, climbing...and cycling soon! I look back fondly at memories of the time we've shared here and abroad.

Savino Longo: thank you for the good times in our shared office, and for sharing your passion for music, your jam sessions, and *"The Boss"*.

Maheyer Shroff: thank you for being my running "coach" and for your never-ending lightheartedness and humour!

Alexandre Beaubien and Charlie Chen: thank you both for the opportunity to share with you some of what I've learned, and for the celebrations we've had at your place.

I also want to give a special shout-out to Traviss Ram. It's been a blast spending late nights with you at Forge and VBB. *"Thanks for everything!"*

Finally, I want to thank my parents for their love and support. Both of you have gone to great lengths to support me and have given me so much opportunity. With what you've provided me, I can take flight.

DEDICATION

To my Mom, for “loving me to a crisp”,

and

To my Dad, for educating me about the world,

and

To my friends, for their unwavering kindness and camaraderie.

Chapter 1

Introduction

The Standard Model (SM) of particle physics is the best theoretical description we have to date of fundamental particles and their interactions. It has succesfully predicted the existence of the W, Z, and h (Higgs) bosons[1] and has been tested to remarkable precision.[2] It is, however, an incomplete theory. For example, it does not explain the presence of *dark matter* in our universe. Nor does it explain the asymmetry in the amounts of matter and anti-matter. It also does not incorporate gravity. These are some of the deficiencies and experimental observations that motivate theories that *extend* the Standard Model. The following chapter gives an overview of some of these challenges.

An interesting extension to the SM proposes the decay of the SM Higgs boson to two *dark sector* particles each called Z_d. This decay occurs in models where the Higgs interacts with a dark sector which could have rich and interesting phenomenology like the SM. The focus of this thesis is the search for the dark sector process $h \to Z_d Z_d \to 4\ell$ with data collected by the ATLAS detector at the Large Hadron Collider (LHC) in 2015–2018. This search builds on an existing search conducted with data collected in

---

[1] The W and Z bosons were predicted in the 1960s by Glashow, Weinberg, and Salam in their theory unifying the electromagnetic and weak interactions. The bosons were discovered in 1983 by the UA1 and UA2 experiments at CERN. The Higgs boson and the Higgs mechanism (responsible for the masses of the W, Z, and the Higgs boson itself) were theorized around the same time by Brout and Englert, Higgs, and Guralnik, Hagen, and Kibble. The Higgs boson was discovered in 2012 by the ATLAS and Compact Muon Solenoid (CMS) experiments.

[2] For example, the mass of the Z boson determined from fitting the SM with other observables is 91.1882 ± 0.0020 GeV, and from experimental measurements is 91.1876 ± 0.0021 GeV [1]. This difference is only -0.3σ. *Electroweak precision tests* not only test the SM but are also interpreted as experimental constraints on new theories.

2015–2016 by quadrupling the dataset size and improving the background estimation and validation and signal regions.[3]

This thesis is organized as follows:

Following a summary of experimental and theoretical motivations for new physics and dark sector models in Chapter 2, the theory of the model proposing the $h \to Z_d Z_d \to 4\ell$ process is introduced in Chapter 3. The interactions, parameters, and resulting phenomenology of the model and existing constraints are discussed. The Large Hadron Collider and the ATLAS detector are described in Chapter 4, with particular focus on the reconstruction and identification of the leptons that are searched for: electron and muons. Chapter 5 details the selection applied to collision events used to search for the process in the data collected by ATLAS. This selection is applied to the data and simulation-based signal and background predictions described in Chapter 6. The data-driven Z+jets background estimate—a significant improvement for this iteration of the search with the 2015–2018 dataset—is detailed in Chapter 7. Finally, the results of the search—the unblinded validation and signal regions—are presented in Chapter 8. The interpretation methodology and the interpretation of the results are presented in Chapter 9 and are followed by conclusions.

---

[3] Two additional channels ($h \to aa \to 4\tau$ and $S \to XX \to 4\ell$) were also added. An overview of the four related channels is given in Section 5.1. The focus of this thesis is on the "high-mass" channel, $h \to Z_d Z_d \to 4\ell$, where each Z_d is hypothesized to decay to two electrons or two muons, and are hypothesized to have similar masses between 15 GeV and 60 GeV.

Chapter 2

Motivation

There are strong experimental reasons to believe that the Standard Model (SM) is part of a larger theoretical structure. While the SM has been successful in its predictions—culminating with the discovery of the Higgs in 2012—it does not explain some well-established experimental observations and measurements. These include:

- Dark matter: a form of matter making up approximately 84% of the mass in the universe [1] that has been observed to interact gravitationally, but not through any other interaction. Its existence has been inferred but its exact makeup is unknown. Some of the strongest evidence for its existence comes from galactic rotation dynamics. Most of the visible matter of a galaxy lies near its centre. This suggests that objects rotating inside the galaxy should have velocities inversely proportional to the square of their radial distance $1/\sqrt{r}$ due to weakening gravitional force. This is inconsistent with the observation that velocities are approximately constant as radial distance increases. This observation suggests that there must be additional *non-visible* (dark) matter. The feebleness of its interaction through other means has made it exceptionally difficult to detect *directly* and to produce. To date, the nature of dark matter has not been discovered.[1] The SM does not incorporate dark matter.

---

[1] There are three main approaches to detecting dark matter other than through its gravitational effects: direct detection, indirect detection, and collider production. *Direct detection* seeks to observe dark matter scattering off SM particles (e.g. recoiling against nuclei inside a liquid medium such as xenon). *Indirect detection* seeks to observe dark matter annihilating into SM particles (e.g. annihilation inside the sun into highly energetic neutrinos that pass through the earth). Finally, dark matter could be produced in particle collisions (for example, in pp collisions at the LHC, with the dark matter particles decaying into SM particles or escaping detection and leaving a *missing energy or momentum* signature).

- Baryon asymmetry: there is an unexplained excess of baryons over anti-baryons. In the early universe, quarks and anti-quarks were pair-produced and annihilated ($\gamma + \gamma \leftrightarrow q + \bar{q}$) at equal rates (*thermal equilibrium*). As the universe cooled, quarks and anti-quarks annihilated into photons at a rate greater than pair production, but an unexplained excess of quarks over anti-quarks remained. These quarks formed in groups of three into baryons.[2] This process of generating the asymmetry is known as *baryogenesis*. The relative excess of baryons over anti-baryons is $\mathcal{O}(10^{-10})$, as determined from the observed abundances of light elements (*big bang nucleosynthesis*) and from the observed fluctuations in the cosmic microwave background. In order for baryogenesis to occur, Sakharov proposed three conditions: i) baryon number violation, ii) *Charge* (C) and *Charge-Parity* (CP) violation, and iii) departure from thermal equilibrium. The SM exhibits all three conditions, but not to the levels necessary to explain the observed asymmetry. C and CP symmetries are violated by the weak interaction and the universe departed from thermal equilibrium (cooled) as it expanded. However, the level of CP violation and the value of the Higgs mass make it impossible for baryogenesis to occur within the SM.

- Neutrino mass: the SM assumes all neutrinos are left-handed[3] and therefore all neutrinos are massless.[4] However, observations of neutrinos oscillating from one flavour to another imply that neutrinos have mass [1].

From an aesthetic perspective, there are concerns about the extraordinary fine-tuning required to make the predictions of the SM consistent with observations. For example, the SM predicts that the leading correction to the Higgs mass-squared is

$$\delta m_h^2 = \frac{3G_F}{4\sqrt{2}\pi^2}\left(4m_t^2 - 2m_W^2 - m_Z^2 - m_h^2\right)\Lambda^2 \tag{2.1}$$

---

[2] Mesons can be formed from one quark and one anti-quark, but are short-lived (unlike the proton and the neutron).

[3] The *handedness* refers to the helicity or chirality of the particle. If a particle has right-handed (left-handed) helicity, then the direction of its spin is the same as (opposite to) its momentum. The helicity can be changed under a boost of reference frame only if the particle is moving at $v < c$; this is only possible if the particle is massive. Chirality, however, is a related but intrinsic property of *fermions* that is invariant under boost (same in all frames). It is defined for fermions as the eigenstates of the γ^5 matrix. In the massless limit, helicity and chirality are the same.

[4] Similarly, the SM assumes all anti-neutrinos are right-handed. The absence of chiral partners for neutrinos and anti-neutrinos suggested that they need not be included in the SM; by construction, the SM does not include right-handed neutrinos and therefore, no mass terms are generated for neutrinos. However, nothing in principle forbids the addition of right-handed neutrinos to the SM.

where m_T, m_W, m_Z, m_h are the masses of the top quark and the W, Z, and Higgs bosons, G_F is Fermi coupling constant, and Λ is the maximum energy scale of the theory [2]. The observed values of the masses imply that $\Lambda \lesssim 1$ TeV. This contradicts the expectation that a complete theory should reach the scale at which quantum gravity is expected to appear: the Planck scale $\Lambda \sim 10^{16}$ TeV. If this expectation is to be maintained—along with the observation that $m_h \approx 125$ GeV—then there must be counter-terms to m_h that combine to $\mathcal{O}(10^{-16})$ TeV. This is an extraordinary level of fine-tuning.[5]

The problems described above are a few of the reasons to expect physics beyond the SM. A well-motivated possibility is the existence of a *dark sector* that may interact in a limited manner with the *visible sector* (i.e. what forms the SM as we know it). As with the SM, the dark sector could consist of one or more particles with its own set of interactions. What we observe and identify as dark matter could be one or more particles of a dark sector. The absence of evidence for a dark sector to date suggests that, if it exists, it must interact with the SM at a level that is substantially weaker than the interactions within the SM. The theory of the dark sector model that is searched for in this thesis is described in Chapter 3.

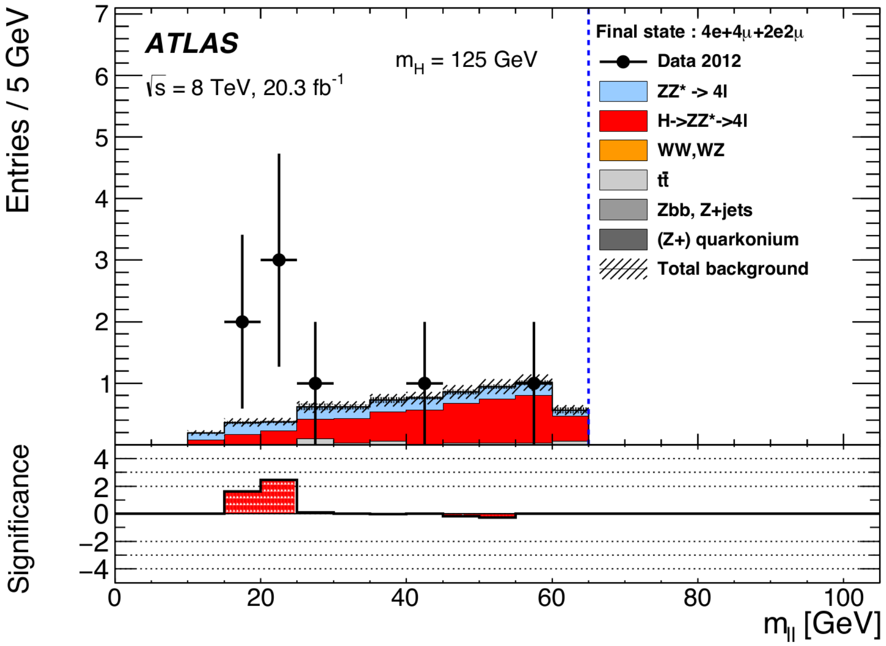
Besides possibly addressing the nature of dark matter, a dark sector could also address other open questions. For example, CP violation could be transferred to the visible sector with an interaction between a dark sector particle and SM particles [3]. This could allow baryogenesis to occur. A dark sector could also easily increase the muon's predicted anomalous magnetic moment and therefore resolve the disagreement between its observed and predicted values [4].

The work of this thesis builds on searches by ATLAS for the dark sector process $h \to Z_d Z_d \to 4\ell$ conducted with data collected in 2010–2012 [5] and 2015–2016 [6]. The main results of those searches are shown in Figure 2.1. These distributions compare, in a *signal region*, the number of lepton pairs (*dileptons*) observed with the number of lepton pairs predicted to come from background processes. This region is a kinematic region that is designed to select leptons from Z_d decay with high purity while rejecting background. A statistically significant excess in the observed yield over the predicted background yield could indicate the presence of the signal process $h \to Z_d Z_d \to 4\ell$.

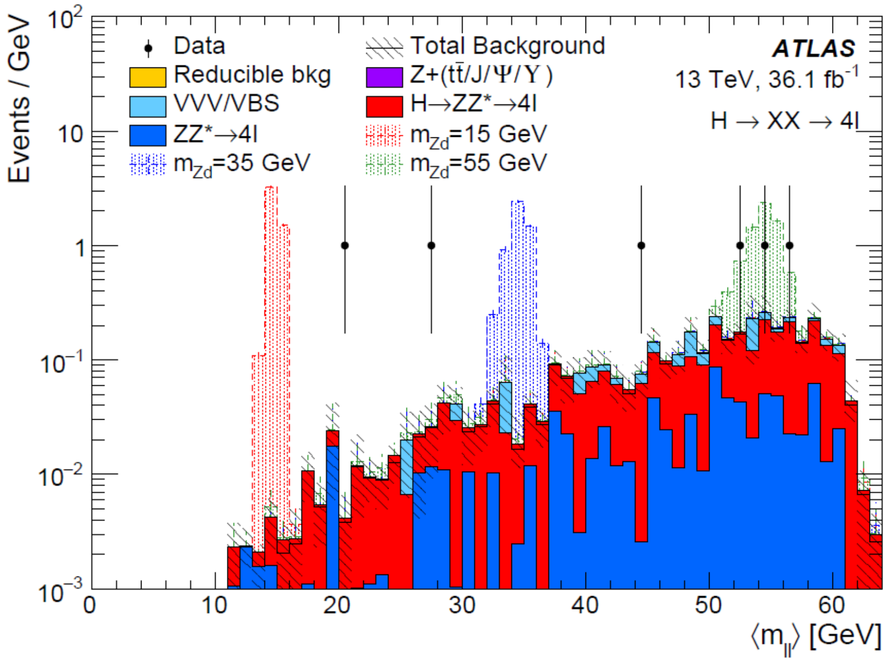
One event with $m_{Z_d} \approx 20$ GeV corresponding to 1.9σ global significance (3.2σ local significance) was found in the ATLAS 2015–2016 dataset search [6]. Two events

---

[5] Also known as the *hierarchy problem*. Why is the Higgs so much lighter than the Planck scale?



(a) Dilepton mass $m_{\ell\ell} \equiv m_{12,34} \sim m_{Z_d}$ in 2010–2012 data and simulation [5]. Two entries are plotted per event.



(b) Average dilepton mass $\langle m_{\ell\ell}\rangle \equiv (m_{12} + m_{34})/2 \sim m_{Z_d}$ in 2015–2016 data and simulation [6]. One entry is plotted per event.

Figure 2.1: Distributions of dilepton mass in searches by ATLAS with 2010–2012 and 2015–2016 data and simulation.

in a similar mass range were also found in the ATLAS 2010–2012 dataset search [5] and in the CMS 2010–2012 dataset search [7]. The excesses could be due to statistical fluctuation and/or background under-estimation. The search described in this thesis quadruples the amount of data used and improves the background estimation. By increasing the size of the dataset, the uncertainty on the observed yields due to finite statistics (the *statistical uncertainty*) is reduced. Notably, the aforementioned excess of the 2015–2016 search is due to the observation of one event in a region with negligible background. This excess could also be due to an under-estimate of background from jets being misidentified or mis-reconstructed as leptons. While this background is expected to be negligible, it is difficult to estimate accurately because of its *fake* nature. The estimate of this background can be improved by replacing the Monte Carlo-based simulation used in the 2015–2016 search with a more accurate *data-driven* method.

Chapter 3

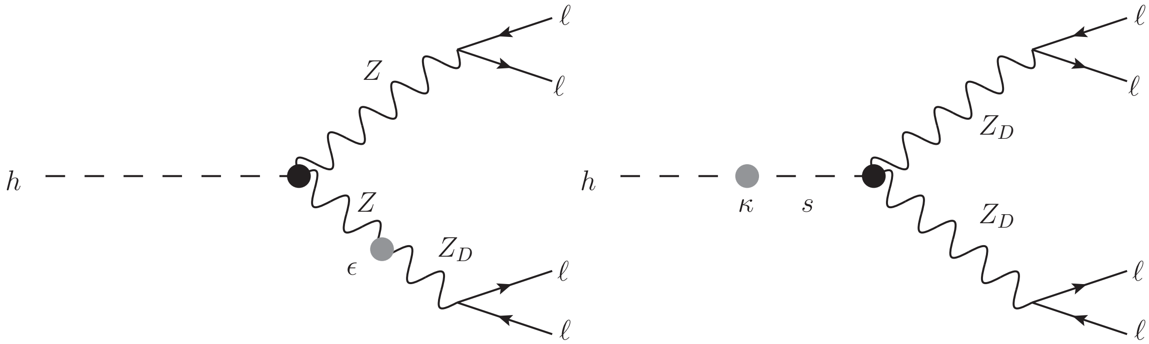
Theory

A natural way for a dark sector to interact with the Standard Model (SM) is through coupling with the SM Higgs boson. The small width of the Higgs means that even a small coupling to a Beyond-the-Standard-Model (BSM) particle could result in sizeable new decay modes. If the BSM contribution to the width of the Higgs is Γ_{BSM}, then the branching ratio of the Higgs to a pair of BSM particles is proportional to $\Gamma_{\mathrm{BSM}}/(\Gamma_{\mathrm{BSM}}+\Gamma_{\mathrm{SM}})$, where Γ_{SM} is the width of the SM Higgs. Since Γ_{SM} is relatively small ($\Gamma_h/m_h \sim \mathcal{O}(10^{-5})$ while $\Gamma_Z/m_Z \sim \mathcal{O}(10^{-2})$, for instance), the branching ratio to BSM particles would be non-negligible. This branching ratio is not well-constrained from experimental measurements: ATLAS and CMS currently constrain the branching ratio of SM Higgs to BSM particles to be less than approximately 30% [8], and future projections for the LHC physics program suggest that this constraint could tighten to at most 10% [9]. Furthermore, the addition of a BSM particle that couples to the SM Higgs can be done with minimal modification to the SM. The dark sector model that is searched for in this thesis is an example of such a minimal extension.

One possibility is for the SM Higgs to decay to one or two new gauge (vector) bosons Z_d. In the minimal scenario, the *dark photon* Z_d interacts with the SM through kinetic mixing between $U(1)_d$ and hypercharge $U(1)_\gamma$. $U(1)_d$ is an additional "dark" gauge symmetry analogous to the $U(1)_\gamma$ hypercharge gauge symmetry of the SM. The new term in the Lagrangian responsible for kinetic mixing is [9, 10]:

$$\mathcal{L} \supset \frac{1}{2}\frac{\epsilon}{\cos\theta_w} B_{\mu\nu} Z_d^{\mu\nu} \tag{3.1}$$

where ϵ is the new kinetic mixing parameter, θ_w is the weak mixing angle (Weinberg angle), and $B_{\mu\nu}$ and $Z_d^{\mu\nu}$ are the $U(1)_\gamma$ and $U(1)_d$ field strengths, respectively. This



(a) Hypercharge portal

(b) Higgs portal

Figure 3.1: Feynman diagrams of $h \to ZZ_d \to 4\ell$ and $h \to Z_dZ_d \to 4\ell$ via the hypercharge and Higgs portals. ϵ is the kinetic mixing parameter of the hypercharge portal interaction and κ is the Higgs mixing parameter of the Higgs portal interaction. S is the dark Higgs.

interaction between the SM and dark sector is referred to as the *hypercharge portal* and allows for $h \to ZZ_d$, $h \to Z_dZ_d$, and $Z_d \to f\bar{f}$ decays. The decay $h \to ZZ_d \to 4\ell$ through the hypercharge portal is visualized in Figure 3.1(a).

To leading order in $m^2_{Z_d}/m^2_Z$, the partial width for this decay mode is [9]:

$$\Gamma\left(h \to ZZ_d\right) = \frac{\epsilon^2 \tan^2\theta_W}{16\pi} \frac{m^2_{Z_d}\left(m^2_h - m^2_Z\right)^3}{m^3_h m^2_Z v^2} \tag{3.2}$$

where v is the vacuum expectation value of the SM Higgs ($v = (\sqrt{2}G_F)^{-1/2} \approx 246$ GeV, where the Fermi coupling G_F is well-determined from experiment [1]). Notably, the width is proportional to ϵ^2. The addition of kinetic mixing alone leads to $h \to ZZ_d$ as the dominant Higgs decay mode to Z_d. If only kinetic mixing is present, then the decay to two Z_d particles is still possible (via the hypercharge portal) but it is highly suppressed by a factor of ϵ^4 since both Z bosons in the decay would need to mix with Z_d.

If a *dark Higgs mechanism* that spontaneously breaks the $U(1)_d$ symmetry is added—much like an Abelian Higgs mechanism—then mixing between the SM Higgs and dark Higgs is introduced and appears in the potential as:

$$V(H,S) \supset \kappa|S|^2|H|^2 \tag{3.3}$$

where H is the SM Higgs doublet, S is the dark Higgs singlet that breaks the $U(1)_d$ symmetry, and κ is the new Higgs mixing parameter. This interaction is referred to as the *Higgs portal*. The decay to two Z_d particles can occur via the Higgs portal; with the addition of Higgs portal, the branching ratio of this decay can be much larger. The decay via the Higgs portal is visualized in Figure 3.1(b). To leading order in κ, the partial width for the decay mode $h \to Z_d Z_d$ with the addition of the Higgs portal is:

$$\Gamma\left(h \to Z_d Z_d\right) = \kappa'^2 \frac{1}{32\pi} \frac{v^2}{m_h} \sqrt{1 - \frac{4m_{Z_d}^2}{m_h^2}} \frac{\left(m_h^2 + 2m_{Z_d}^2\right)^2 - 8\left(m_h^2 - m_{Z_d}^2\right) m_{Z_d}^2}{m_h^4} \tag{3.4}$$

where the *effective Higgs mixing parameter* κ' is:

$$\kappa' = \kappa \frac{m_h^2}{|m_h^2 - m_s^2|} \tag{3.5}$$

The term in the Lagrangian responsible for coupling of Z_d to SM fermions (electrons, muons, taus, neutrinos, and quarks) is:

$$\mathcal{L} \supset g_{Z_d f \bar{f}} Z_d^{\mu} \bar{f} \gamma_\mu f \tag{3.6}$$

where the couplings $g_{Z_d f \bar{f}}$ are proportional to ϵ and $1/m_{Z_d}^2$. These couplings lead to decays of Z_d to SM fermions. If Z (Z_d) and Z_d are produced, then the signature will be four fermions, with two fermions from each decay. If the fermions are leptons, then each pair of leptons will create a dilepton mass resonance, and the four leptons together will create a resonance at the Higgs mass. This phenomenology motivates the identification of lepton pairs and quadruplets, and the selections used in this search. To leading order, the partial widths for decays of Z_d to pairs of fermions are:

$$\Gamma\left(Z_d \to \bar{f} f\right) = \frac{N_{\text{colours}}}{12\pi m_{Z_d}} \sqrt{1 - \frac{4m_f^2}{m_{Z_d}^2}} \left(m_{Z_d}^2 g_{Z_d f \bar{f}}^2 + 2m_f^2 g_{Z_d f \bar{f}}^2\right) \tag{3.7}$$

where N_{colours} is 1 for leptons and 3 for quarks.

The branching ratio to leptons, shown in Figure 3.3, is quite sizable.[1] To one flavour of leptons, it is approximately 15% for $m_{Z_d} \sim 1$ GeV and decreases to approximately 9% for $m_{Z_d} \gtrsim 60$ GeV. While $\Gamma(h \to Z_d Z_d)$ does not depend on ϵ, $\Gamma(h \to Z Z_d)$ and $\Gamma(Z_d \to f\bar{f})$ do. The dependence of the latter on ϵ means that observation of $h \to Z_d Z_d \to 4\ell$ decay implies a non-zero value for ϵ due to the leptonic portion of the

---

[1] The branching ratio for decay mode i is $\Gamma_i/\Gamma_{\text{total}}$, where the numerator is the partial decay width and the denominator is the total decay width $\Gamma_{\text{total}} = \sum_i \Gamma_i$.

decay. The dependence of the former on ϵ means that observation of $h \to ZZ_d \to 4\ell$ decay also has the same implication, but phenomenological studies [10] show that the experimental sensitivity to ϵ via this decay is limited to $\epsilon \gtrsim 10^{-3}$, even with the next generation of collider experiments.[2] The only requirement on ϵ for the $h \to Z_d Z_d \to 4\ell$ decay to occur is that it is large enough to enable the $Z_d \to \ell\ell$ decay. Following from $\Gamma(Z_d \to \ell\ell)$ and the fact that a particle's lifetime is inversely proprotional to its decay width, *prompt* $Z_d \to \ell\ell$ decay occurs for $\epsilon \gtrsim 10^{-5}$. Therefore, observation of prompt $h \to Z_d Z_d \to 4\ell$ decay would provide sensitivity to this level of ϵ. If $\Gamma(Z_d \to f\bar{f})$ is extremely small (e.g. $\epsilon \lesssim 10^{-5}$), then Z_d would have a long lifetime and decay at a displaced location in the detector. Further sensitivity could be gained by searching for these displaced decays, with the level of sensitivity dependent on the displacement $c\tau$ (speed of light multiplied by mean lifetime) and the detector's acceptance and efficiency at such displacements.

Multiplying the branching ratio to leptons with the branching ratio of the Higgs to one or two Z_d particles yields the total branching ratios $\mathcal{B}(h \to Z_{(d)} Z_d \to 4\ell)$ shown in Figure 3.4.

Searches for Z_d for $m_{Z_d} \sim \mathcal{O}(1\text{ GeV})$ have been conducted at electron-positron experiments and for $m_{Z_d} \sim \mathcal{O}(10\text{ GeV})$ at proton-proton experiments. BaBar and BESIII have searched for $e^+e^- \to \gamma A'$, where the dark photon A' can decay leptonically ($A' \to e^+e^-$ or $\mu^+\mu^-$, $0.02 < m_{A'} < 10.2$ GeV for BaBar and $1.5 < m_{A'} < 3.4$ GeV for BESIII) [11, 12] or invisibly to a pair of lighter dark sector particles ($A' \to \chi\bar{\chi}$, $m_{A'} < 8$ GeV for BaBar) [13].

In the searches with leptonic final states, the detector signature is two charged tracks and a single photon. The signal could appear as a resonance in the dilepton distribution. In the search for invisible decays, the detector signature is a mono-energetic initial-state radiation photon. This photon has energy $E^\star_\gamma = (s - m^2_{A'})/(2\sqrt{s})$, where s is the centre-of-mass energy; therefore, the signal could appear as a resonance in the $m^2_{A'} = s - 2E^\star_\gamma\sqrt{s}$ distribution. This search requires a single-photon trigger, which was only implemented for the final running period of BaBar (effective analysis threshold of $E^\star_\gamma \geq 1.8$ GeV). The dominant SM backgrounds are $e^+e^- \to \gamma\gamma(\gamma)$ with only one of the photons detected; the others are not detected and/or are out of detector acceptance. This search will be performed by Belle-II, which has a dedicated single-photon trigger, improved detector hermiticity, and greater luminosity. In these mass ranges, BaBar and BESIII have set upper limits of $\epsilon \lesssim 10^{-3}$.

---

[2] Largely due to the SM $h \to ZZ^* \to 4\ell$ background.

LHCb has searched for $A' \to \mu^+\mu^-$ in proton-proton collisions [14, 15]. For prompt A' decay, the searches considered $m_{A'}$ from the dimuon threshold up to 70 GeV, and for non-prompt A' decay, the searches considered $214 < m_{A'} < 350$ MeV. The A' couplings are inherited from the SM couplings, therefore $\gamma^\star \to \mu\mu$ has the same production and decay kinematics as $A' \to \mu\mu$. The two signatures are experimentally indistinguishable and therefore have the same detector efficiencies. This is exploited with a data-driven approach: the observed $A'(\gamma^\star) \to \mu\mu$ are fitted and then normalized to expectation derived from simulation in order to obtain limits on ϵ. This results in a cancellation of experimental systematics. The latter LHCb search has set upper limits of $\epsilon \lesssim 10^{-3}$. Limits on ϵ are compared in Figure 3.2.

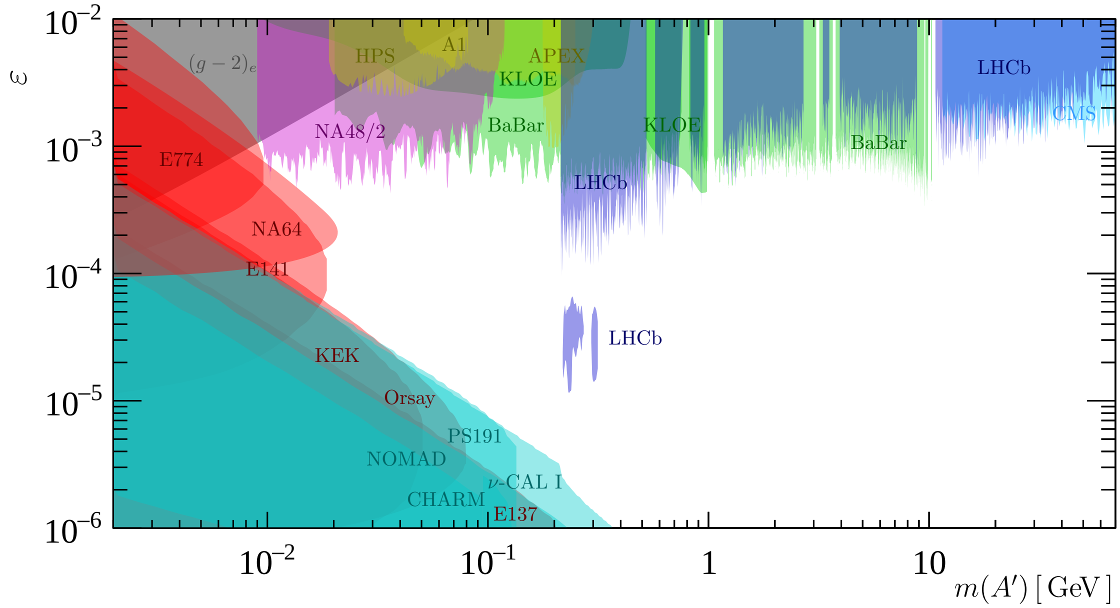


Figure 3.2: Comparison of upper limits set on the kinetic mixing parameter ϵ versus dark photon mass $m_{A'}$ by various experiments (90% confidence level). From [15] and generated with the recasting approach described in [16]. Details of the searches can be found in [16].

Below the mass of the Z boson ($10 < m_{Z_d} < 80$ GeV), re-interpretation of Drell-Yan ($pp \to \gamma^*/Z \to \ell\ell$) cross-section measurements sets limits of $\epsilon \lesssim 10^{-2}$ [17, 18]. In the immediate vicinity of $m_Z \approx 91.2$ GeV, fits to electroweak precision observables strengthen the limit to $\epsilon \lesssim 3 \times 10^{-3}$ [10]. The search for $h \to ZZ_d \to 4\ell$ by ATLAS with data collected in 2015–2016 set a limit of $\epsilon \lesssim 5 \times 10^{-2}$ in the range $15 < m_{Z_d} < 55$ GeV [6].

The strongest upper limits on κ' have been set by searches for $h \to Z_d Z_d \to 4\ell$ by ATLAS and CMS.[3] The search by ATLAS with the 2015–2016 dataset set a limit of $\kappa' \lesssim 10^{-3}$ [6], as shown in Figure 3.5. This thesis describes a search for this process with approximately four times the amount of data (36 fb^{-1} to 139 fb^{-1}). In the absence of discovery, the upper limit should strengthen to $\epsilon \lesssim 2 \times 10^{-4}$, as predicted in Figure 3.6. The upper limit on $\mathcal{B}(h \to Z_d Z_d)$ set by the same ATLAS search is $\lesssim 10^{-4}$, as shown in Figure 3.7. Similarly, this limit should strengthen to $\lesssim 3 \times 10^{-5}$, as predicted in Figure 3.8. The methodology used to set limits on the branching ratio and κ' is described in Chapter 9.

---

[3] In May 2020, CMS published preliminary results of a search similar to the one described in this thesis, conducted on the CMS 2016–2018 (137fb^{-1}) dataset. This search is described and compared against the ATLAS search in Appendix A. Previous CMS searches for exotic Higgs decays with leptonic final states include a search for $h \to aa + X \to 4\mu + X$ and $h \to n_1 n_1 \to \gamma_D \gamma_D n_D n_D \to \mu\mu n_D n_D$ [19] and a search for $h \to aa \to 4\tau$ $(2\mu 2b, 2\mu 2\tau)$ [7].

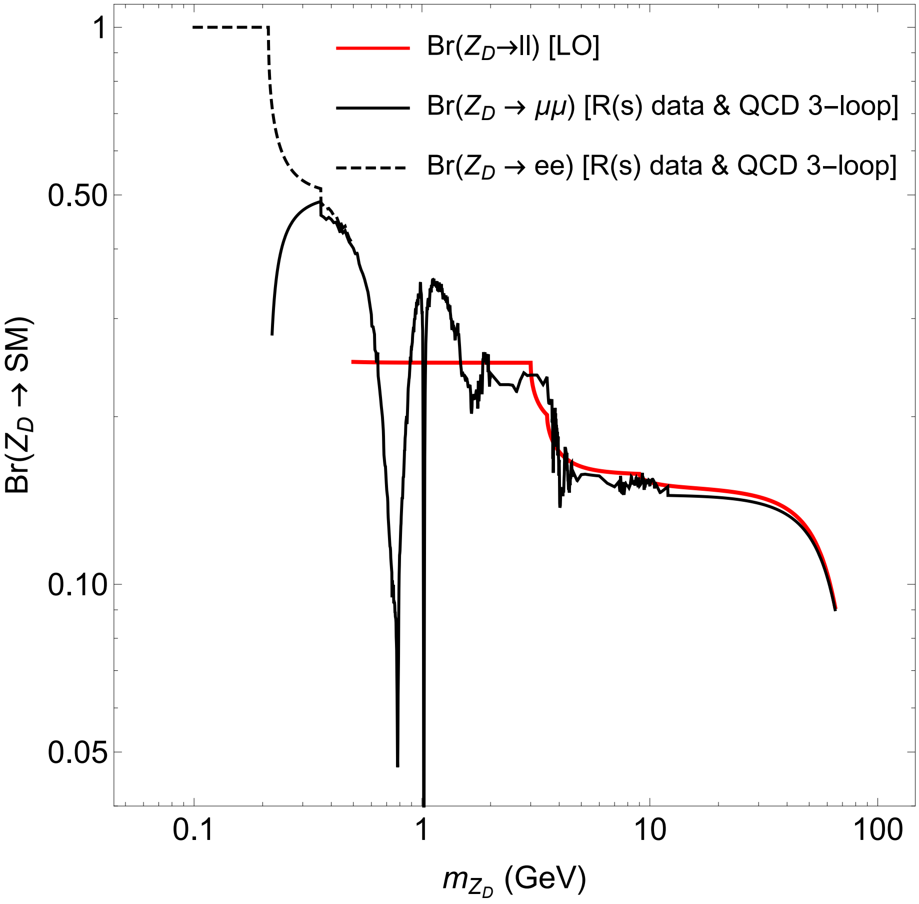
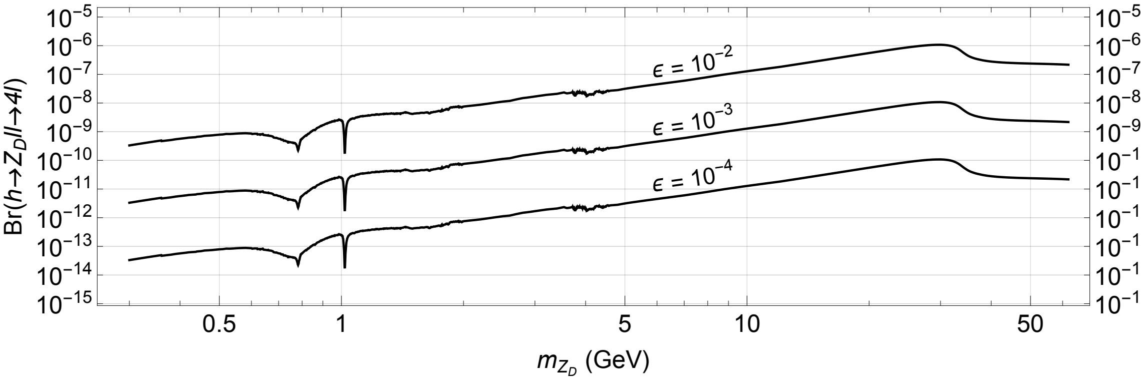
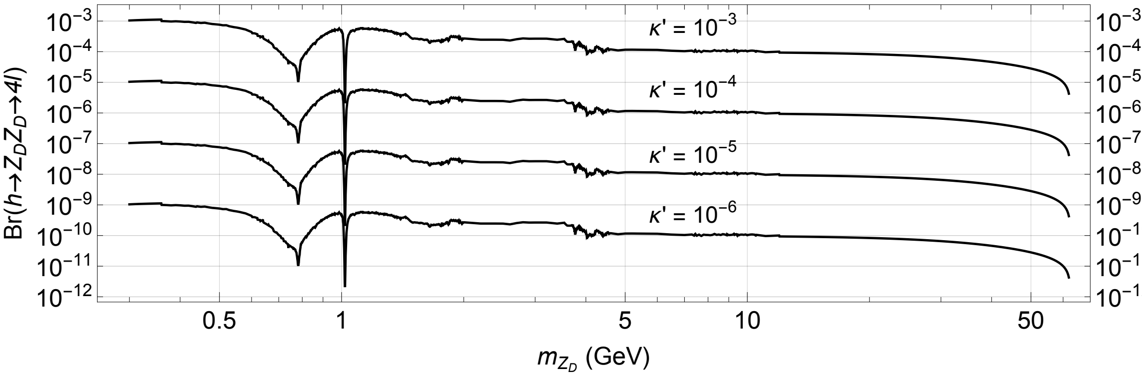


Figure 3.3: Branching ratio of $Z_d \to \ell\ell$ ($\ell = e$ or μ) predicted from theory [10]. The branching ratio calculated at leading order (red line) is a good approximation of the branching ratios calculated with higher-order QCD corrections (black lines). Higher-order calculations of the branching ratio require knowledge of $R_{Z_d} \equiv \Gamma(Z_d \to \text{hadrons})/\Gamma(Z_d \to \mu^+\mu^-)$. For $m_{Z_d} < 12$ GeV, measurements of the hadronic to muonic cross-section in e^+e^- collisions $R(s) \equiv \sigma(e^+e^- \to \text{hadrons})/\sigma(e^+e^- \to \mu^+\mu^-)$ can be used to fully determine R_{Z_d}. For $m_{Z_d} > 12$ GeV, R_{Z_d} can be calculated with existing higher-order predictions for $R(s)$ by replacing SM couplings with the couplings $g_{Z_d f\bar{f}}$.



(a) $\mathcal{B}(h \to ZZ_d \to 4\ell)$



(b) $\mathcal{B}(h \to Z_d Z_d \to 4\ell)$

Figure 3.4: Branching ratios of $h \to Z_{(d)} Z_d \to 4\ell$ predicted from theory for different values of kinetic mixing parameter ϵ and effective Higgs mixing parameter κ' [10].

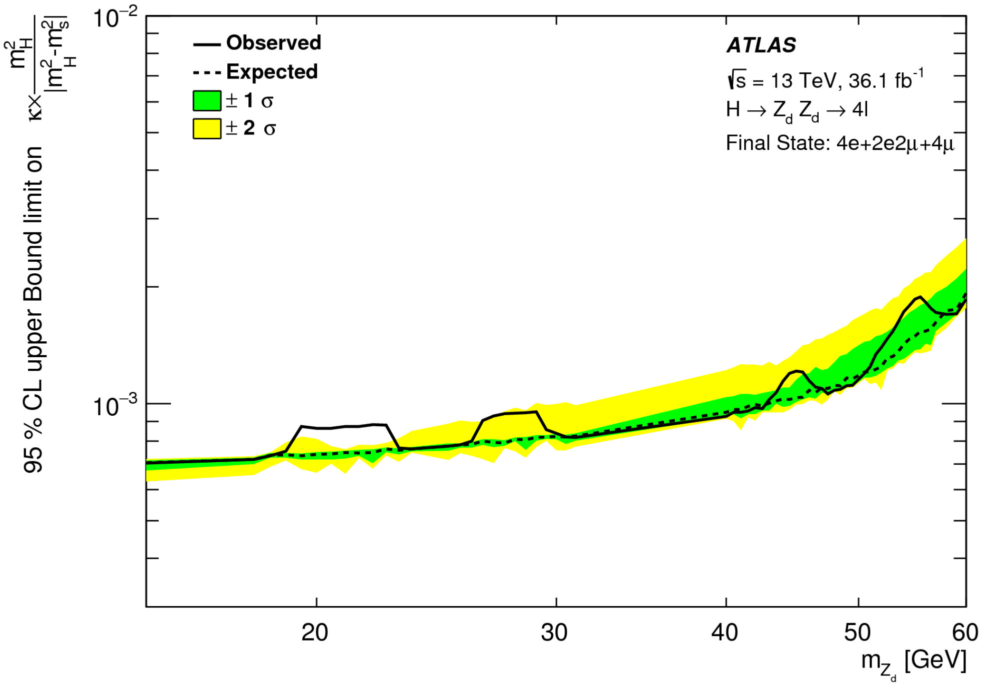


Figure 3.5: Upper limit on the effective Higgs mixing parameter κ' set by ATLAS with 2015–2016 data [6].

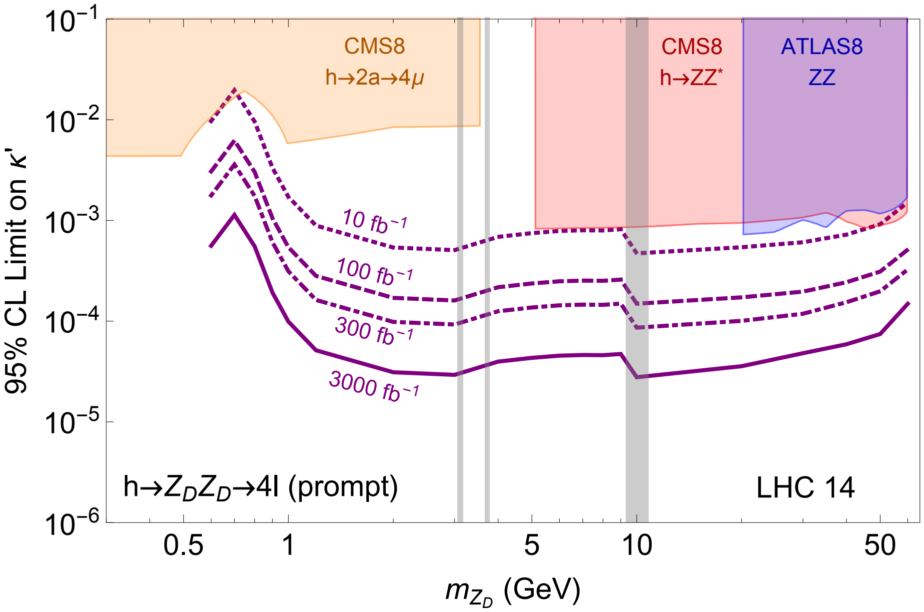


Figure 3.6: Expected upper limits on the effective Higgs mixing parameter κ' from experiments at the LHC at centre-of-mass energy $\sqrt{s} = 14$ TeV [10]. The upper limit set by ATLAS with 2015–2016 data (36 fb^{-1}) is shown in Figure 3.5.

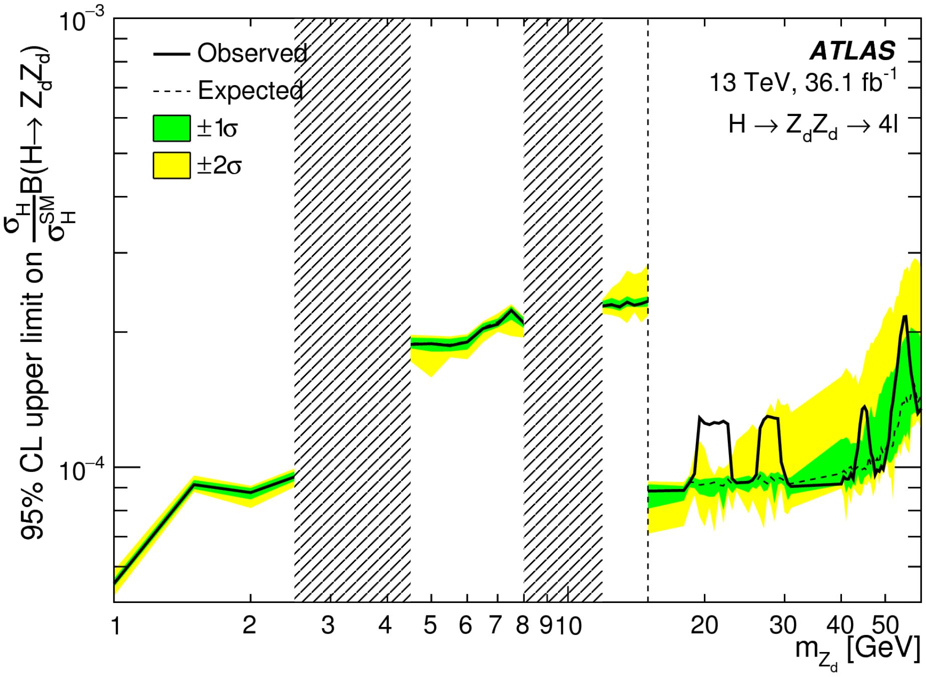


Figure 3.7: Upper limit on the branching ratio $\mathcal{B}(h \to Z_dZ_d)$ set by ATLAS with 2015–2016 data [6].

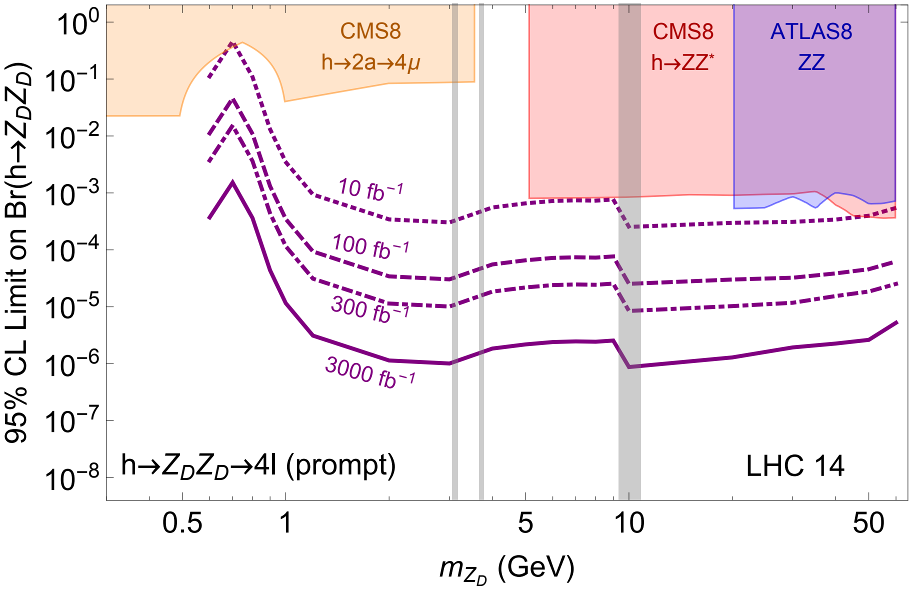


Figure 3.8: Expected upper limits on the branching ratio $\mathcal{B}(h \to Z_dZ_d)$ from experiments at the LHC at centre-of-mass energy $\sqrt{s} = 14$ TeV [10]. The upper limit set by ATLAS with 2015–2016 data (36 fb^{-1}) is shown in Figure 3.7.

Chapter 4

LHC and ATLAS

This chapter describes the ATLAS detector and the Large Hadron Collider (LHC). The protons produced and accelerated by the LHC are collided inside ATLAS and three other interaction points. The ATLAS detector consists of multiple sub-detectors designed to accurately reconstruct and identify particles produced in the collisions. Particle reconstruction and identification entails measuring their momentum, energy, and other properties. The $h \to Z_d Z_d \to 4\ell$ search requires four well-reconstructed and identified electrons and/or muons and thus capitalizes on the exceptional performance of the ATLAS detector—in particular, the calorimeter and Muon Spectrometer for electron and muon reconstruction and identification, respectively.

4.1 Large Hadron Collider

The Large Hadron Collider (LHC) is currently the world's most powerful particle accelerator. It is situated, on average, 100 m underground and has a circumference of approximately 27 km. Protons are accelerated by a series of smaller accelerators before they are injected into the LHC, where they are boosted to a final energy of 6.5 TeV per proton beam. Sixteen radiofrequency cavities accelerate the particles and 1232 dipole and 392 quadrupole magnets guide and focus the beams [20].

The two counter-rotating beams are focused and collided at four interaction points along the circumference of the LHC. At each interaction point, a detector observes the resulting *pp* or *heavy ion* collisions.[1] Two of the detectors—ATLAS and CMS—are

---

[1] The LHC can also accelerate xenon and lead nuclei. Besides *pp* collisions, Xe-Xe, Pb-Pb, and *p*-Pb collisions have also be studied by the ATLAS detector.

general purpose detectors that were designed to be sensitive to a wide range of physics phenomena at the TeV scale.

Each 6.5 TeV proton beam contains approximately 2232 proton bunches, with each bunch containing approximately 1.1×10^{11} protons.[2] Every 25 ns, bunches are *crossed* at the four interaction points along the circumference of the LHC in order to produce collisions inside each detector. This time interval is commonly referred to as the *bunch spacing*.

The instantaneous luminosity quantifies the ability of a collider to produce interactions and is defined as the ratio of the number of events per second to the cross-section for that type of event, as given in Equation 4.1. In practice, luminosity is expressed in terms of beam parameters since neither the cross-section nor the number of events per second can be directly manipulated. Equation 4.2 expresses luminosity in terms of the revolution frequency f, the number of bunches N_b, the number of protons per bunch N_p, and the x and y root-mean-square beam widths σ_x and σ_y.

$$L = \frac{1}{\sigma} \frac{dN_{\text{events}}}{dt} \tag{4.1}$$

$$L = \frac{1}{4\pi} \frac{f N_b N_{\text{p}}^2}{\sigma_x \sigma_y} \tag{4.2}$$

The mean number of interactions per bunch crossing μ varies proportionally to the square of the number of protons per bunch. This variation produces a spread in μ, as shown in Figure 4.1.

---

[2] Initial parameters for *Run-2* (2015 and onwards) operation.

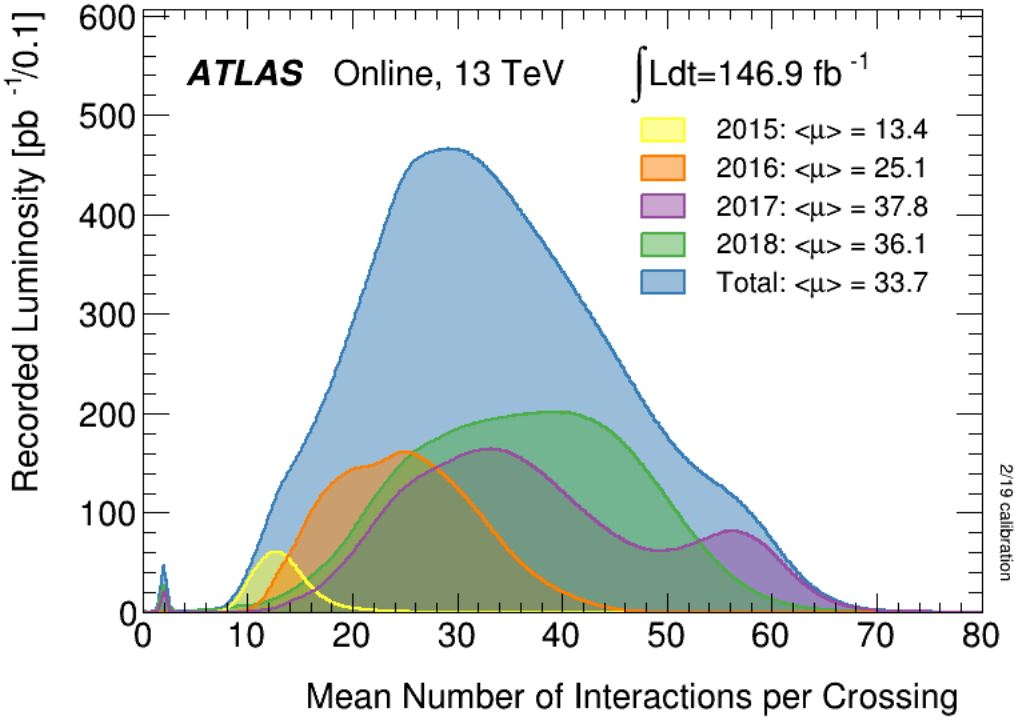


Figure 4.1: Distribution of μ during each year of 2015–2018 pp collision data-taking [21]. The shape of the distribution changed in 2017 due to a change in the LHC filling scheme for improved beam stability.

4.2 ATLAS detector

4.2.1 Overview

The ATLAS (A Toroidal LHC ApparatuS) detector was designed to exploit the high energy and luminosity of the LHC in order to explore a wide range of physics ranging from precision measurements of known processes, to discovery of theorized and even unexpected physics.

The main components of the ATLAS detector (shown in Figure 4.2) beginning from the innermost to outermost are:

- The **Inner Detector (ID)**, which determines the originating vertices of interactions, identifies charged particles, and measures the trajectories of charged particles via energy deposition in semiconductors and ionization of xenon and argon gas.

- The **electromagnetic (EM) and hadronic calorimeters**, which determine the energy of particles that interact electromagnetically (i.e. electrons or photons) and hadronically (i.e. jets) by absorbing them and measuring the energy deposited.

- The **Muon Spectrometer (MS)**, which identifies muons and measures their trajectories as they pass through and ionize gases in the MS drift chambers.

- The **magnets**, consisting of a solenoid surrounding the Inner Detector and three toroid magnets (one barrel toroid and two end-cap toroids). The bending of the trajectories of charged particles by the magnets enables the Inner Detector and Muon Spectrometer to measure their momenta.

- The **trigger and data acquisition (TDAQ) system**, which receives and buffers the output of the detector, selects events by performing a fast preliminary reconstruction, and writes selected events to permanent storage.

Signals from the detector are used to *reconstruct* physics objects such as electrons, muons, and jets. Candidates from reconstruction are then *calibrated* to account for detector effects such as energy response and *identified* for quality and origin.

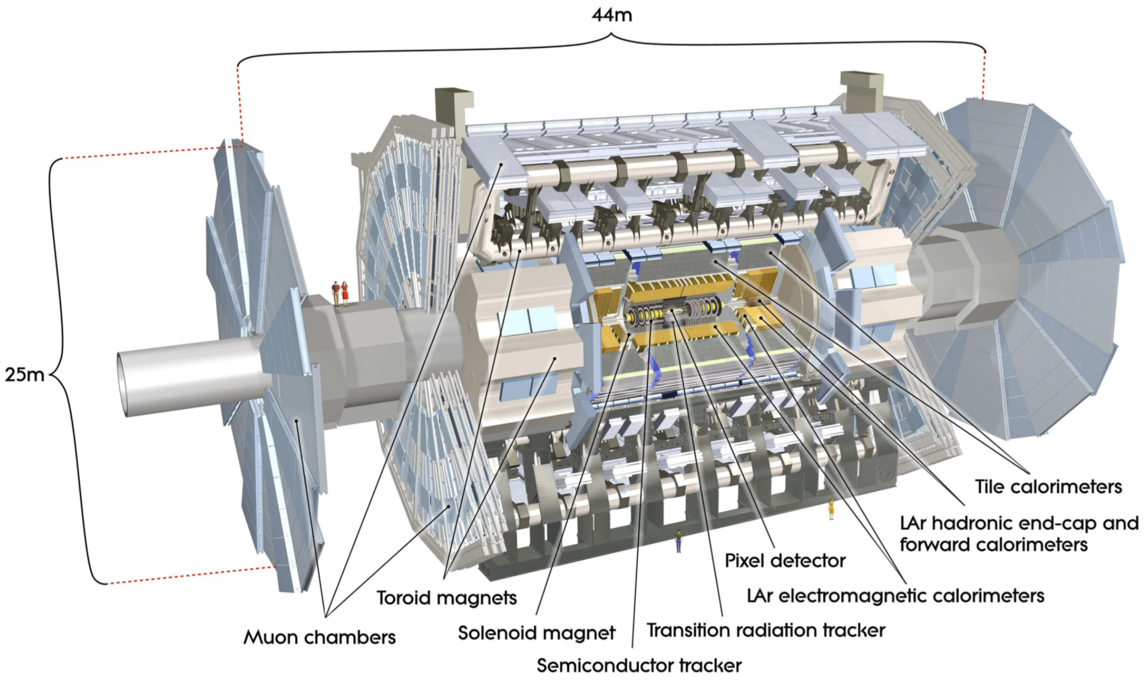


Figure 4.2: Cut-away view of the ATLAS detector [22]. The detector possesses cylindrical symmetry and was designed to be as hermetic as possible. It has close to 4π coverage in solid angle.[3]

4.2.2 Inner Detector

The ATLAS Inner Detector (ID) consists of three sub-detectors: the pixel detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). The entire ID is immersed in a 2 T magnetic field that deflects the trajectories of charged particles. The ID was designed to reconstruct these trajectories (*tracks*) in the region of the interaction point within a range of $|\eta| < 2.5$. Interactions happen at an extended distance along the beam pipe due to the spatial spread of the bunches. The ID establishes the origins of the hard scatter and other pp interactions (*vertices*). The tracking performance requirements and the proximity of the ID to the interaction point and beam pipe impose requirements on the electronics and material thickness: the electronics of the ID must maintain low occupancy amidst high hit rates, have efficient readout and be radiation hard, and the amount of material must be minimized to

---

[3] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.

reduce the number of unwanted photon conversions and other interactions producing secondary particles. As a charged particle passes through the ID, it will register hits in the semiconductors of the pixel detector and SCT, and ionize the gas of the TRT through primary ionization and transition radiation.

The $h \rightarrow Z_d Z_d \rightarrow 4\ell$ search considers electrons and muons in the final state. Electrons are reconstructed by matching tracks in the ID with energy depositions in the calorimeters and muons are reconstructed by matching tracks in the ID with tracks in the Muon Spectrometer. The reconstruction and identification of electrons and muons are discussed in depth in Sections 5.4.1 and 5.4.2, respectively.

The pixel detector—the innermost layer of the ID—has four barrel layers and two end-caps with three disks each for a total of 92 million readout channels [23, 24]. The pixel detector has a silicon-covered surface area of approximately 2 m^2. Layer-0 of the pixel detector—the Insertable B-Layer—is new for *Run-2* and improves tracking by providing measurements closer to the interaction point. The reduced distance improves the tracking resolution, which in turn improves the vertexing and b-tagging performance. Furthermore, it provides the pixel detector a margin of robustness against high hit rates by continuing to provide hits even when the readout electronics of the other layers are saturated. The intrinsic measurement accuracy of each sensor is approximately 10 μm in R–ϕ and 115 μm in z (for the barrel layers) or R (for the end-caps) [22].

The Semiconductor Tracker (SCT)—the middle layer of the ID—has four barrel layers and two end-caps with nine disks [25]. It has a silicon-covered surface area of approximately 60 m^2. It has less granularity than the pixel detector (6.3 million readout channels) and consequently has less accuracy. The intrinsic measurement accuracy of each sensor is approximately 17 μm in R–ϕ, and 580 μm in z (for the barrel layers) or R (for the end-caps) [22].

The Transition Radiation Tracker (TRT)—the outermost layer of the ID—is a drift tube system consisting of approximately $350,000$ drift tubes (also referred to as straw tubes) filled with xenon-based and argon-based gases.[4] The structure of each

---

[4] The TRT initially operated entirely with a xenon-based gas, but due to large and irreparable leaks in *Run-1* and the high cost of xenon, parts of the TRT operated with an argon-based gas instead of a xenon-based gas in *Run-2*. Gases with a higher-Z (xenon) are preferred over gases with lower-Z (argon) for transition radiation production because the photoelectric absorption cross-section is proportional to Z^n $(4 < n < 5)$ [1]. For this reason, discrimination between electrons and pions was negatively impacted by the change in gas. This impact has been partially compensated by the implementation of a likelihood-based discriminant [26]. Conversely, tracking performance has improved slightly because the pulse shapes with argon gas allow for a lower effective threshold and consequently increased sensitivity to low amounts of ionization [27].

drift tube acts as a cathode while the tungsten wire strung along the centre of each tube acts as an anode. When a charged particle passes through, it ionizes the gas and produces ionization charges (electrons and ions). The charges are accelerated by the electric field towards the anode wire and produce further ionization. This additional ionization is commonly referred to as an *ionization cascade*. The charges are then collected on the anode wire, which registers a signal that is proportional to the amount of ionization. In addition to the charges of the primary ionization and subsequent ionization cascade, additional charges are produced by the absorption of *transition radiation* by the gas.[5]

4.2.3 Calorimeter

The ATLAS calorimeter is comprised of electromagnetic (EM) and hadronic calorimeters that cover an overall range of $|\eta| < 4.9$. It is of a *sampling* design; each calorimeter consists of passive absorbing material and an active measurement medium. When a particle is absorbed, it generates a cascade of secondary particles (*shower*) that ionizes the active medium. In the electromagnetic calorimeters, the hadronic end-cap calorimeters, and the forward calorimeters, particles are absorbed mainly in dense material and the ionization of the liquid argon (LAr) is measured by charge collection on electrodes. LAr was chosen as the active medium because of its linear response and radiation hardness. In the tile calorimeter, particles are absorbed by steel plates and ultraviolet light from the ionization of scintillating polystyrene tiles is collected by photomultiplier tubes. The ATLAS calorimeter is also *non-compensating*; it possesses a greater response to electromagnetic showers than to hadronic showers due to invisible and escaped energy in hadronic interactions that is not immediately compensated for. For measurements of missing transverse momentum and jets, it is crucial that this difference in response is properly accounted for through calibration. The calorimeter

---

[5] Transition radiation consists of X-ray photons that are produced when highly-relativistic charged particles pass through boundaries. In the TRT, straw tubes are interleaved with fibres and foils in order to induce transition radiation. This effect creates a much larger signal than ionization [1, 28] and is dependent on how relativistic the particle is moving $\left(\gamma = \sqrt{1 + \left(\frac{p}{m_0}\right)^2}\right)$. The energy deposition from transition radiation produced by electrons ranges from 8 keV to 10 keV, while pions produce negligible transition radiation and deposit approximately 2 keV [29]. This is because electrons are significantly lighter than pions (approximately 278 times lighter than a $\pi^{\pm}$). This difference is exploited by the TRT in order to distinguish electrons from other minimally-ionizing particles such as pions.

has a total of $187,652$ cells,[6] with increased granularity in the EM calorimeter and near the interaction point.

The electromagnetic calorimeter consists of a barrel calorimeter covering up to $|\eta| < 1.475$ and two *electromagnetic* end-cap calorimeters covering up to $|\eta| < 3.2$. Its high granularity enables it to precisely measure electrons and photons in conjunction with the ID. The absorbing media (lead plates) and active media (LAr) are organized in an accordion-like geometry as shown in Figure 4.3. This geometry provides full coverage in ϕ without any cracks and provides enough depth to contain most showers—at least 22 radiation lengths in the barrel and at least 24 radiation lengths in the EM end-caps. Additional depth to contain hadronic showers is provided by the surrounding hadronic calorimetry.

The solenoid surrounding the ID bends charged particles into helicoidal paths (circles in the transverse plane (xy) traveling along the beam axis). This results in a spread in the transverse plane but not in the longitudinal (rz) plane.[7] For this reason, the first layer of the barrel is more finely segmented in η than in ϕ. The second layer usually collects the most energy from an EM shower out of all layers, and the third layer usually collects only the tail. The segmentation of the EM barrel calorimeter is shown in Figure 4.4.

A presampler covering up to $|\eta| < 1.8$ consisting of one layer of LAr accounts for the energy lost in front of the EM calorimeter (i.e. in the innermost region of the detector).

The hadronic calorimeter consists of two *hadronic* end-caps calorimeters, a tile calorimeter, and two forward calorimeters. The hadronic end-caps, covering $1.5 < |\eta| < 3.24$, use LAr as the active medium due to the increased radiation close to the beam pipe and are interleaved with copper plates. The tile calorimeter, covering up to $|\eta| < 1.7$, uses scintillating tiles and photomultiplier tubes and is interleaved with steel plates as shown in Figure 4.5. The geometry of the tile calorimeter enables full coverage in ϕ. The forward calorimeters, covering $3.1 < |\eta| < 4.9$, each consist of three modules: an electromagnetic module that is closest to the interaction point, and

---

[6] $182,468$ cells (excluding presamplers) in the liquid argon calorimeters (9344 presampler cells in addition) and 5184 cells in the tile calorimeters [30].

[7] In other words, the separation between an electron and the photons it radiates through bremsstrahlung is mostly in ϕ. For a given material, the scale in the transverse direction of an EM shower is the *Molière radius* and is defined as the radius of a cylinder that on average encompasses 90% of the energy deposition of a shower. For liquid argon, the Molière radius is approximately 9 cm [1].

behind it, two hadronic modules. For the absorbing media, copper is used in the EM module and tungsten is used in the hadronic modules.

The $h \rightarrow Z_d Z_d \rightarrow 4\ell$ search considers electrons (and muons) in the final state. Electrons are reconstructed by identifying clusters of energy deposition in the electromagnetic calorimeter and matching them with tracks in the ID. The reconstruction, calibration, and categorization of these candidates are described in Section 5.4.1.

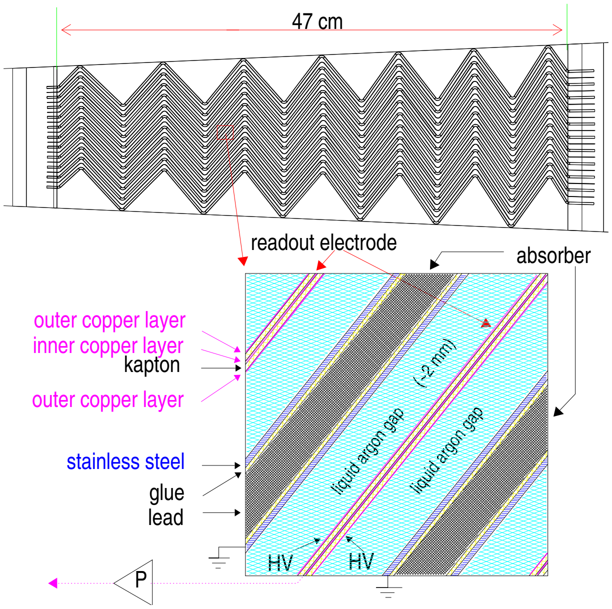


Figure 4.3: Accordion-like structure of the LAr electromagnetic barrel calorimeter [31]. The absorbing medium and active medium are interleaved.

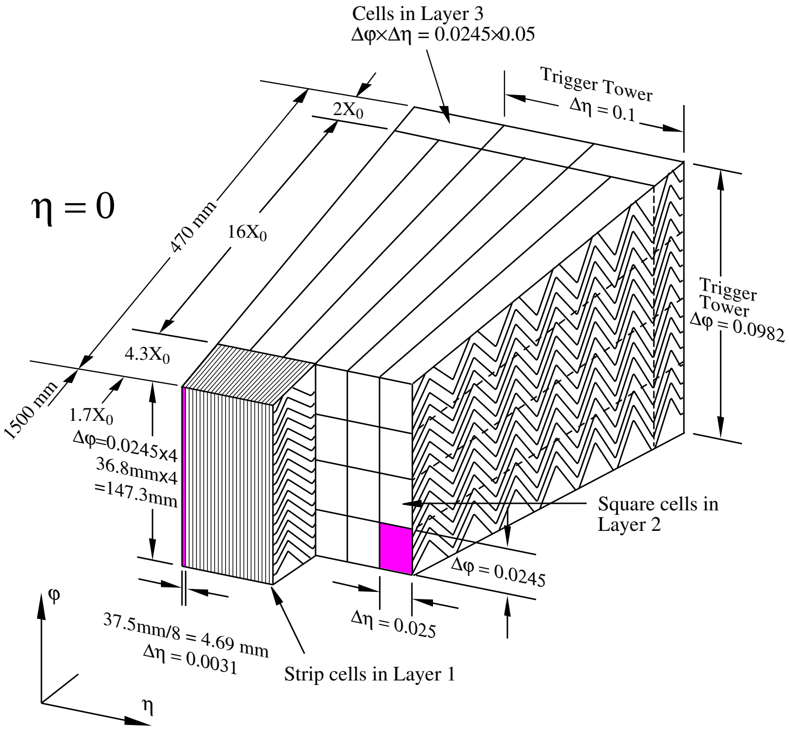


Figure 4.4: Schematic of a portion of the LAr electromagnetic barrel calorimeter [22].

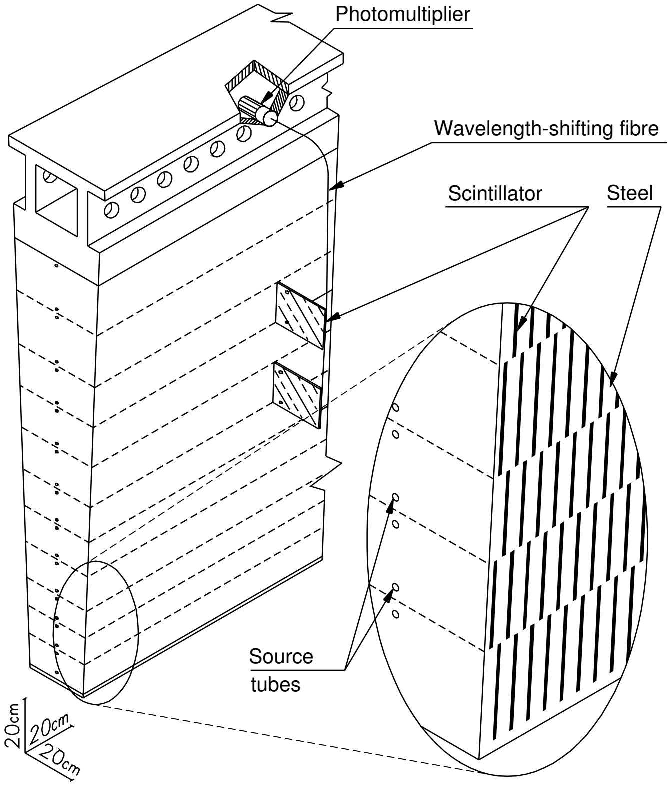


Figure 4.5: Schematic of a portion of the tile hadronic calorimeter [22]. The scintillating tiles and steel plates are oriented radially. Each of the three barrels (one central and two extended) consists of 64 modules.

4.2.4 Muon Spectrometer

The Muon Spectrometer (MS) is the outermost sub-detector of the ATLAS detector [32]. It was designed to measure muons with transverse momenta (p_T) ranging from 3 GeV to 3 TeV within a range of $|\eta| < 2.7$. The MS operates within a magnetic field produced by one barrel toroid covering $|\eta| < 1.4$ and two end-cap toroids covering $1.6 < |\eta| < 2.7$. The magnetically-deflected muon trajectories are measured by a combination of drift chambers: monitored drift tubes (MDTs), cathode strip chambers (CSCs), resistive plate chambers (RPCs), and thin gap chambers (TGCs). The first two types are primarily used for precision tracking and the last two types are primarily used for triggering.

The chambers operate on the principle of gaseous ionization (the production of electron-ion pairs). Each chamber is filled with a gas: Ar/CO_2 for MDTs and CSCs, tetrafluoroethane for RPCs, and CO_2/n-pentane for TGCs. A muon traveling through the gas contained in each chamber will cause *primary ionization*. The electrons from primary ionization will cause additional ionization, with the amount dependent on the type of gas chosen. In the presence of an electric field, electrons and ions will drift in opposite directions; the structure of each chamber (tube) acts as a cathode (attracts positive charge) while the wire strung along the centre of each chamber (tube) acts as an anode (attracts negative charge). As shown in Figure 4.7, the electrons will drift towards the anode, at a *drift velocity* that is dependent primarily on the gas used. Additionally, the acceleration of the electrons by the electric field causes collisions with the gas molecules and will result in further ionization (*Townsend avalanche*). By measuring the signal pulse on the anode precisely, the drift time and distance can be determined. For MDTs, the drift time is about 700 ns.

Monitored drift tubes are used for precision tracking in the region $|\eta| < 2.7$ (except in the innermost layer where they are replaced with cathode strip chambers in the forward region $2 < |\eta| < 2.7$). As shown in Figure 4.8, each MDT chamber consists of layers of drift tubes with the tubes oriented along the magnetic field lines (i.e. along ϕ). In the middle and outer layers, each chamber has two *multi-layers* of three tube layers each. In the innermost layer, each multi-layer has four tube layers each instead. A muon that passes through all six layers of a chamber yields six coordinates r_{min} in the plane perpendicular to the tubes (the plane corresponding to the cross-section view of each tube). The z resolution of a chamber is approximately 35 μm and of a tube is 80 μm.

As a muon travels through the MS, the toroids bend its trajectory in the rz plane. The layout of the MS in this plane is shown in Figure 4.6. The bending causes the muon's trajectory to vary in η.[8]

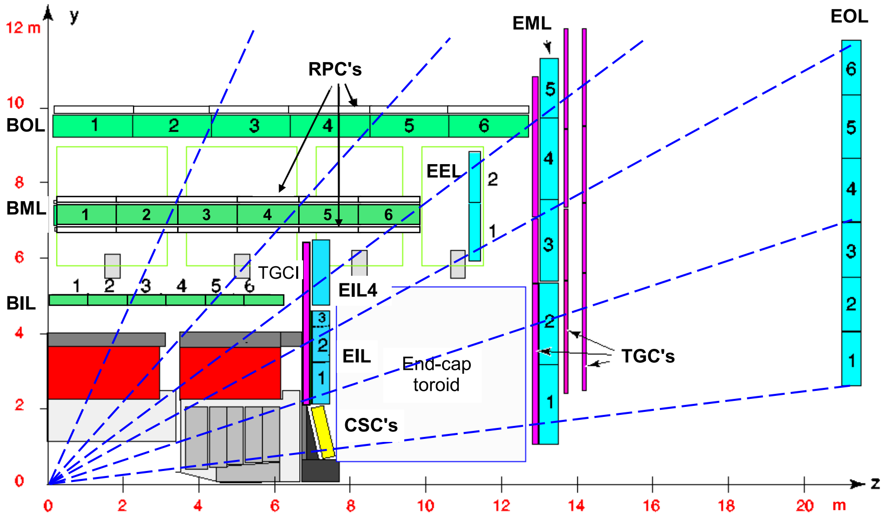


Figure 4.6: Layout of the Muon Spectrometer in the rz plane [22]. The MDTs are in green and teal, the CSCs in yellow, the TGCs in purple, and the RPCs in white rectangles enclosed by black lines.

A trajectory constructed from these measurements yields the *sagitta*. This is the distance from the midpoint of the trajectory's arc to a straight line connecting the first and last hits (chord). The sagitta is measured with a resolution of $\lesssim 50\ \mu$m. From the sagitta, the component of the momentum in the rz plane (radial component) can be determined. This component is the dominant component of the muon's momentum.[9]

In the forward region of the innermost layer where the hit rate exceeds the capacity of MDTs (> 150 Hz/cm^2), cathode strip chambers (CSCs) are used. They operate on the same principle of gaseous ionization as the MDTs. Instead of one wire per tube (acting as the anode and cathode respectively), multiple wires are strung in the radial direction inside a gap sandwiched between two cathodes and filled with gas. The strips of one cathode are oriented perpendicular to the wires and the strips of the

---

[8] The solenoidal field enveloping the ID causes the trajectory to become helicoidal along the beam axis. Consequently, as a particle moves through the ID, its trajectory will vary in ϕ. This can be visualized as a curve in the transverse plane. Once it exits the ID, it adopts a straight trajectory in the transverse plane. Since this straight trajectory does not trace back to the origin, it will vary slightly in ϕ.

[9] Ignoring the effect of the ID mentioned above and energy losses in the calorimeter.

other cathode are oriented parallel to the wires (with the former segmented more finely than the latter). This layout, shown in Figure 4.9, enables the determination of both η and ϕ. The strips perpendicular to the wires have a R resolution of approximately 40 μm and the strips parallel to the wires have a ϕ resolution of approximately 5 mm. Four CSCs are layered in order to provide four measurements of η and ϕ.

For triggering, the MDTs are sandwiched by resistive plate chambers (RPCs) in the barrel and by thin gap chambers (TGCs) in the end-caps, as shown in Figure 4.6. Each package of RPCs and MDTs form a *trigger station* (RPC/MDT/RPC in the middle layer and RPC/MDT in the outer layer). A resistive plate chamber (RPC) consists of two parallel resistive plates made of bakelite with an electric field across a gap filled with gas. Readout strips measure the ionization occuring within the gap via capacitive coupling of the resistive plates [33] and are laid out orthogonal to each other (η-strips and ϕ-strips). This layout enables the determination of both η and ϕ in the same manner as CSCs (but unlike drift tubes). Two layers of RPCs are layered in order to provide two measurements of η and ϕ; a muon that travels through all three stations will produce six measurements of η and ϕ, with precision in η provided by the MDTs. RPCs have a resolution of approximately 10 mm in z and ϕ within $|\eta| < 1.05$. Thin gap chambers (TGCs) are multi-wire proportional chambers like CSCs, with one triplet and two doublet layers in the end-caps ($1.05 < |\eta| < 1.92$) and one doublet layer only in the inner (forward) region ($1.92 < |\eta| < 2.4$). TGCs have a resolution of 2 to 6 mm in R and 3 to 7 mm in ϕ within $|\eta| < 2.4$.

A challenge in muon triggering is maintaining resolution in p_{T} over $|\eta|$. Muon momentum increases rapidly with $|\eta|$, while the integrated bending power increases only modestly. In the forward region, hit rates from background are greater and the chambers covering this region (TGCs) are outside of the magnetic field. These challenges necessitate increased granularity in the end-cap region, provided by the TGCs.

Reconstruction of muons can be done with the MS alone or in combination with the ID and/or calorimeter by extrapolation and association of tracks between sub-detectors [34]. The $h \to Z_d Z_d \to 4\ell$ search primarily uses muons first identified independently in the ID and MS and then fitted with a track incoporating both ID and MS hits (*combined muons*). The types of muons and their calibration and identification are discussed in Section 5.4.2.

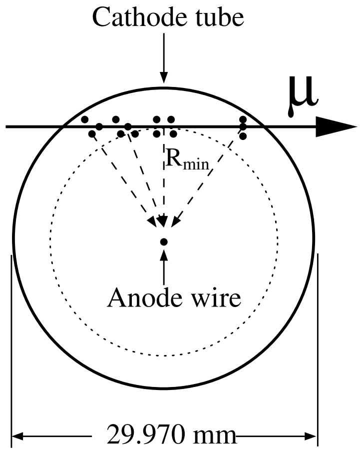


Figure 4.7: Cross-section of a monitored drift tube [22]. R_{min} is the track coordinate and μ is the trajectory of a muon.

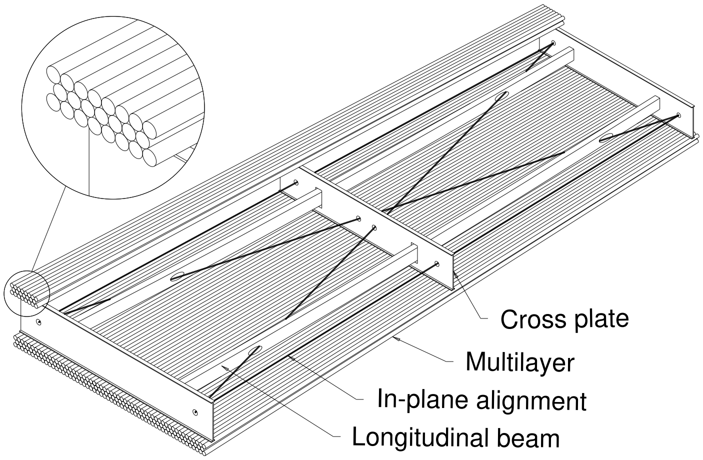


Figure 4.8: Monitored drift tube consisting of six layers of tubes [32].

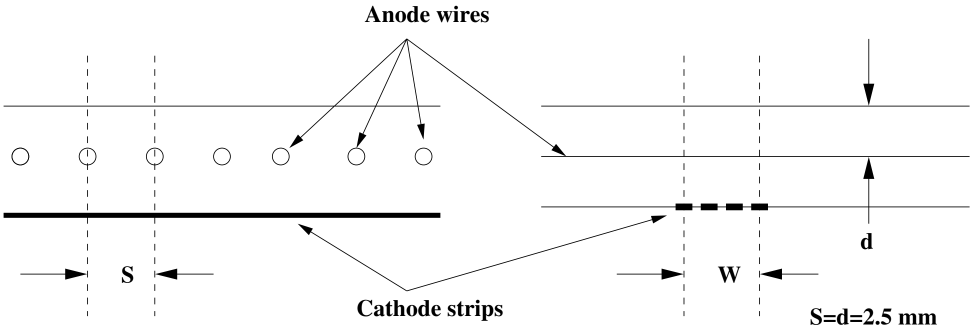


Figure 4.9: Cathode strip chamber [22]. On the left, the anode wires are going into the page. The CSCs are *multi-wire proportional chambers*.

4.2.5 Trigger and data acquisition system

The ATLAS detector features a multi-level trigger and data acquisition (TDAQ) system that reduces the raw event rate provided by the LHC (40 MHz) to a more manageable 1.2 kHz [22, 35]. The system receives the output of the detectors and within 2.5 μs, identifies object candidates and makes a decision to accept the event for further processing or not. This decision—made by the Level-1 (L1) trigger using coarse detector output—reduces the rate from 40 MHz to maximally 100 kHz. If the event is accepted by the L1 trigger, the full detector output and regions-of-interest(s) identified by the L1 trigger are passed onto the High-Level Trigger (HLT). The HLT, within 0.2 s, makes a final decision using more complex algorithms and access to the full granularity of the detector. If the event is accepted by the HLT, the full detector output (along with any objects reconstructed by the trigger) is written to permanent storage for offline reconstruction and analysis.

The L1 trigger and HLT are configured in *trigger menus*. These menus are generally static for a run period (duration of days to months) and change only when operating conditions change (e.g. significant changes in instantaneous luminosity/pileup). Each trigger menu consists of combinations of L1 trigger configurations and HLT configurations chained together into *trigger chains*.

For example, the chain `HLT_2mu14_L12MU10` translates to a requirement of at least two muons with $p_T \geq 10$ GeV identified by the L1 trigger, followed by a requirement of at least two muons with $p_T \geq 14$ GeV identified by the HLT.[10] Additional requirements can be imposed such as cut-based identification (e.g. `HLT_e60_medium`), likelihood-based identification (e.g. `HLT_e60_lhmedium`), and isolation requirements (e.g. `HLT_mu20_iloose_L1MU15`). Due to limitations in available bandwidth, a trigger may have its rate reduced by allowing only a pre-defined fraction to pass (*trigger prescale*). The $h \to Z_d Z_d \to 4\ell$ search uses the un-prescaled triggers with the lowest p_{T} thresholds. These triggers are listed in Table 4.1.

---

[10] The HLT algorithms examine the region-of-interest(s) identified by the L1 trigger algorithms. Consequently, the objects that fire the L1 trigger are usually the same as the ones that fire the HLT.

Leptons	Trigger name	2015	2016	2017	2018
e	HLT_e24_lhmedium_L1EM20VH	✓			
	HLT_e24_lhmedium_iloose_L1EM18VH	✓			
	HLT_e24_lhtight_nod0_ivarloose		✓		
	HLT_e26_lhtight_nod0				✓
	HLT_e26_lhtight_nod0_ivarloose		✓	✓	✓
	HLT_e60_lhmedium	✓	✓		
	HLT_e60_lhmedium_nod0		✓	✓	✓
	HLT_e60_medium		✓		
ee	HLT_2e12_lhloose_L12EM10VH	✓			
	HLT_2e15_lhvloose_nod0_L12EM13VH		✓		
	HLT_2e17_lhvloose_nod0		✓		
	HLT_2e17_lhvloose_nod0_L12EM15VHI			✓	✓
	HLT_2e24_lhvloose_nod0			✓	✓
μ	HLT_mu20_iloose_L1MU15	✓			
	HLT_mu26_ivarmedium		✓	✓	✓
	HLT_mu24_iloose (_L1MU15 in MC)		✓		
	HLT_mu24_imedium		✓		
	HLT_mu24_ivarloose (_L1MU15 in MC)		✓		
	HLT_mu24_ivarmedium		✓		
	HLT_mu40	✓	✓		
	HLT_mu50	✓	✓	✓	✓
	HLT_mu60_0eta105_msonly	✓	✓	✓	✓
$\mu\mu$	HLT_mu18_mu8noL1	✓			
	HLT_mu20_mu8noL1		✓		
	HLT_mu22_mu8noL1		✓	✓	✓
	HLT_2mu10	✓	✓		
	HLT_2mu14		✓	✓	✓
$e\mu$	HLT_e17_lhloose_mu14	✓			
	HLT_e17_lhloose_nod0_mu14		✓	✓	✓
	HLT_e24_lhmedium_nod0_L1EM20VHI_mu8noL1		✓		
	HLT_e26_lhmedium_nod0_mu8noL1			✓	✓
	HLT_e26_lhmedium_nod0_L1EM22VHI_mu8noL1		✓		
	HLT_e7_lhmedium_mu24	✓			
	HLT_e7_lhmedium_nod0_mu24		✓	✓	✓

Table 4.1: Triggers used in the $h \to Z_d Z_d \to 4\ell$ search with 2015–2018 data and Monte Carlo (MC)-based simulation. Triggers are only used when they are un-prescaled. A separate Good Runs List (Section 5.3) is maintained for each trigger that records which luminosity blocks the trigger was un-prescaled. Prescaled triggers are not included because of their low p_T thresholds. The events that pass only prescaled triggers would not pass the minimum lepton p_T requirements used in this search. Those requirements are summarized in Table 5.1 and described in Sections 5.4 and 5.6.

Chapter 5

$h \to Z_d Z_d \to 4\ell$ search strategy

This chapter describes the $h \to Z_d Z_d \to 4\ell$ search strategy. The goal of this search is to look for the $h \to Z_d Z_d \to 4\ell$ process in proton-proton collision events collected by the ATLAS detector. The strategy consists of a series of selection criteria that form a *signal region*. The selection criteria are optimized to accept events with the $h \to Z_d Z_d \to 4\ell$ process (*signal events*) and reject events with other processes (*background events*). A statistically significant excess in observed events over estimated background events in the signal region would be evidence of the existence of the process.[1]

The probability of a process occuring is quantified by its *production cross-section* σ. It is expressed in units of area, and in its most basic form, it is the number of events with this process N divided by the instantaneous luminosity (events per time per area) integrated over time:

$$\sigma \equiv \frac{N}{\int \mathcal{L} \, \mathrm{d}t} \tag{5.1}$$

After applying the selection criteria, the number of expected background events can be compared to the number of events observed. A discovery would be quantified primarily by its *significance* and would be followed by measurements of properties such as the production cross-section of $h \to Z_d Z_d \to 4\ell$. If there is no statistically significant excess, then an upper limit on the cross-section is set instead.

The development of the search is conducted *blind* in order to avoid bias. Data is not included in the signal region until the strategy is fully developed and validated:

---

[1] Typically, a global significance (Chapter 9) of at least 3σ is required to suggest *evidence* and a global significance of at least 5σ is required to declare *discovery*, where σ is the standard deviation of a standard Gaussian (Section 9.1.1).

1. A selection is designed, guided by the phenomenology of $h \to Z_d Z_d \to 4\ell$ and intuition of the detector signature it would leave. The design of the selection is also guided by prior knowledge of processes that produce similar detector signatures (backgrounds). The selection requirements[2] are optimized for signal efficiency and suppression of background through simulation-based studies of the signal region. This step is the focus of this chapter.

2. The background estimates in the signal region are validated by comparing against data in *validation regions*. These regions share similarities in kinematic variables with the signal region but do not overlap with it.

3. Uncertainties on the signal and background estimates are assessed. Three kinds of uncertainties are considered: statistical, experimental, and theoretical. *Statistical* uncertainties arise from the finite amount of signal and background simulation available due to computational constraints. *Experimental* systematic uncertainties account for uncertainties in the simulation of detector effects such as detector resolution. *Theoretical* systematic uncertainties account for modelling effects such as uncertainties in the knowledge of how momenta is distributed amongst incoming partons.

4. A statistical analysis for the unblinded signal region is designed. If there is no discovery (no significant excess), then upper limits on the $h \to Z_d Z_d \to 4\ell$ production cross-section will be set. This procedure is exercised while the analysis is blind by computing upper limits with the expectation of no excess (*expected upper limits*). In addition to validating the statistical analysis, these limits can be used to optimize the selection such that maximum exclusion is achieved if there is no discovery. They can also serve as a benchmark for comparing the search against previous and competing searches.

5. The signal region is *unblinded*. If there is an excess of data over background estimates, the discovery significance is calculated in order to determine if the excess is statistically significant. If it is not or if there is no excess at all, then the upper limits are computed with the observed data.

---

[2] Also referred to as selection *cuts*, which form a selection *cutflow*.

5.1 Channels

The $h \to Z_d Z_d \to 4\ell$ search that is the focus of this thesis shares similarities with four other searches conducted by ATLAS. These five searches all look for Higgs (or scalar) decays to Beyond-the-Standard-Model (BSM) particles in four lepton final states. They differ in their progenitor particle (Standard Model (SM) Higgs or BSM scalar S), intermediate particle X ($Z_d Z_d$, ZZ_d, or BSM pseudoscalars aa), and final state (combinations of electrons and muons, all muons, or all taus). In conceptual terms, they are similarly motivated and look for similar intermediate particles and final states in comparable mass ranges. For these reasons, they share similarities in search strategy, signal and background modeling, and statistical interpretation. They are also referred to as *search channels* and are as follows:

1. High-mass (*the focus of this thesis*): $h \to Z_d Z_d \to 4\ell$, where $\ell = e, \mu$ and m_{Z_d} is between 15 GeV and 60 GeV,
2. ZZ_d: $h \to ZZ_d \to 4\ell$, where $\ell = e, \mu$ and m_{Z_d} is between 15 GeV and 55 GeV,[3]
3. Low-mass (4μ): $h \to aa \to 4\mu$, where m_a is between 1 GeV and 15 GeV,[4]
4. 4τ: $h \to aa \to 4\tau$, where m_a is between 15 GeV and 60 GeV,
5. Additional scalar: $S \to XX \to 4\ell$ where $\ell = e, \mu$, m_S is up to 750 GeV, and m_X is up to 250 GeV.[5]

The focus of this thesis is on the high-mass channel. $h \to Z_d Z_d \to 4\ell$ and *high-mass* can be used interchangeably. Aspects that pertain only to a particular search channel will mention the search channel explicitly.

5.2 Search strategy

If the $h \to Z_d Z_d \to 4\ell$ process is produced in the ATLAS detector, it would leave a detector signature consisting of four leptons.[6] Each Z_d would decay to two leptons

---

[3] The mass ranges for the ZX and high-mass channels are different because the mass distributions of the X and Z bosons start to overlap significantly at approximately 55 GeV ($m_Z = 91.2$ GeV).

[4] The low-mass channel only considers the 4μ final state because muon triggers and reconstruction are able to reach lower p_{T}. Also, heavy flavour backgrounds become significant at low p_{T}.

[5] The additional scalar channel has two signal regions. The first searches for S with mass below the SM Higgs (30 GeV $< m_S <$ 115 GeV and 15 GeV $< m_X <$ 50 GeV) and the second searches for S with mass above the SM Higgs (130 GeV $< m_S <$ 750 GeV and 20 GeV $< m_X <$ 250 GeV).

[6] Only prompt decays of Z_d are considered.

with the same flavour (two electrons or two muons) and opposite electric charge. The goal of the search strategy is to reconstruct these final state leptons with acceptable efficiency and suppression of background.

The search starts with an event pre-selection consisting of basic requirements such as passing data quality checks and the firing of a trigger. Events that pass pre-selection are then required to have at least four leptons (electrons and/or muons) that pass the *baseline* selection criteria. Leptons are then paired together into same-flavour opposite-sign (SFOS) *dileptons* and dileptons are paired into *quadruplets*. The quadruplets are then ranked and filtered through a *quadruplet* selection. The highest-ranked surviving quadruplet is the *selected quadruplet* of the event. Finally, the selected quadruplet is subject to the event selection. The events that survive the event selection are the events of the signal region.[7] The selections are summarized in Table 5.1.

---

[7] Equivalently, the quadruplets that survive the event selection are the quadruplets of the signal region.

Category	Name	Requirement
Event pre-selection	Good Runs Lists	Event is from a luminosity block w/o detector problems
	Primary vertex	Event has a primary vertex (vertex w/ highest Σp_T^2 of its tracks, must have at least two tracks)
	Trigger	Event fired at least one lowest un-prescaled (multi-)lepton trigger
	Detector flags	No LAr calorimeter noise burst and/or data corruption No tile calorimeter data corruption No single-event upsets in SCT No events with missing detector information
Baseline electrons	p_T	$p_T > 7$ GeV
	η	$\lvert\eta\rvert < 2.47$ (including crack regions)
	Identification	VeryVeryLoose with B-layer cut
	Object quality	Not built with a bad calorimeter cluster
Baseline muons	p_T	$p_T > 5$ GeV (15 GeV if calorimeter-tagged)
	η	$\lvert\eta\rvert < 2.7$
	Identification	Loose
	Impact parameter	Transverse: $d_0 < 1$ mm (stand-alone muons exempt)
Quadruplet formation	Quadruplet	At least one quadruplet of leptons consisting of two same-flavour opposite-sign leptons
Quadrulet selection	p_T	Three leading-p_T leptons satisfying $p_T > 20$ GeV, 15 GeV, 10 GeV
	Muon type	Number of calorimeter-tagged muons plus number of stand-alone muons not greater than one
	Overlap removal	No overlap-removed e or μ
	Trigger matching	At least one lepton in quadruplet responsible for firing at least one trigger In the case of multi-lepton triggers, all leptons of the trigger must match to leptons in the quadruplet
	Lepton separation	$\Delta R(\ell, \ell') > 0.10\,(0.20)$ for same-flavour (different-flavour) leptons in the quadruplet
Quadruplet ranking	Minimal $\Delta m_{\ell\ell}$	Select quadruplet with smallest $\Delta m_{\ell\ell} \equiv \lvert m_{12} - m_{34}\rvert$
Event selection	Isolation	`FCLoose` for electrons, `FCLoose_FixedRad` for muons
	Electron identification	Loose with B-layer cut
	Electron impact parameters	Longitudinal: $\lvert z_0 \sin\theta\rvert < 0.5$ mm Transverse significance: $d_0/\sigma_{d_0} < 5$
	Muon impact parameters	Longitudinal: $\lvert z_0 \sin\theta\rvert < 0.5$ mm (stand-alone muons exempt) Transverse significance: $d_0/\sigma_{d_0} < 3$
	Quarkonia (heavy flavour) veto	Reject event if: $(m_{J/\Psi} - 0.25\text{ GeV}) < m_{12,34,14,32} < (m_{\Psi(2S)} + 0.30\text{ GeV})$ and/or $(m_{\Upsilon(1S)} - 0.70\text{ GeV}) < m_{12,34,14,32} < (m_{\Upsilon(3S)} + 0.75\text{ GeV})$
	Higgs window	$115\text{ GeV} < m_{4\ell} < 130\text{ GeV}$
	Z veto	$10\text{ GeV} < m_{12,34} < 64\text{ GeV}$ $4e$ and 4μ channels: $5\text{ GeV} < m_{14,32} < 75\text{ GeV}$
	Low mass veto	$m_{12,34,14,32} > 5$ GeV
	Loose signal region	$m_{12,34} > 10$ GeV
	Medium signal region	$m_{34} > 0.85 m_{12} - 0.1125 f(m_{12}) m_{12}$

Table 5.1: Selection criteria of the $h \to Z_d Z_d \to 4\ell$ search. The *event pre-selection* selects events that are suitable for analysis and is described in Section 5.3. The *baseline* electron and muon selections select events with electrons and/or muons passing basic quality and kinematic requirements and are described in Section 5.4. These leptons are combined into dileptons and quadruplets, and one quadruplet is selected per event, as described in Section 5.6. Finally, the *event selection* applies kinematic requirements that reject quadruplets from background processes and select quadruplets compatible with final state of the $h \to Z_d Z_d \to 4\ell$ process. This selection is described in Section 5.7.

If the leptons from $h \to Z_d Z_d \to 4\ell$ are observed, then the invariant mass[8] of each dilepton would correspond to the mass of each Z_d in the intermediate state. If the intermediate Z_d are hypothesized to have similar mass, then the dilepton masses would also be similar. The average of the two dilepton masses, $\langle m_{\ell\ell} \rangle \equiv (m_{12} + m_{34})/2$, would also correspond to the mass of the Z_d. Therefore, in the signal region, the chosen kinematic variable of interest is $\langle m_{\ell\ell} \rangle$. Any excess observed in data over expected background in the average dilepton mass distribution may indicate the existence of Z_d with m_{Z_d} corresponding to the excess in $\langle m_{\ell\ell} \rangle$. Shown in Figure 2.1(b) of Chapter 2 is the distribution of $\langle m_{\ell\ell} \rangle$ in the signal region with 2015–2016 data and expected background from simulation. This distribution is from a previous iteration of this search and was published in 2018 [6].

5.3 Event pre-selection

The event pre-selection removes events that fail basic requirements. In rare cases, the detector may not be fully operational or operating correctly. *Event cleaning* vetoes these events by checking Good Runs Lists (GRLs) and detector error flags. These vetoes are also known as data quality checks and are only applied to events in data. After event cleaning, two other basic requirements are enforced: the existence of a primary vertex and the firing of at least one of the selected triggers. In detail, event pre-selection consists of the following:

Good Runs Lists The event must be from luminosity blocks white-listed in the Good Runs Lists. Luminosity blocks are time periods of data-taking with approximately constant beam and detector conditions (e.g. stable instantaneous luminosity and trigger pre-scales and menu). They are approximately one to two minutes long. If problems with the detector within a luminosity block make the data collected unusable for physics analysis, the block is excluded from the GRLs. Effectively, the events within that block are vetoed and the integrated luminosity in that block is not counted.

Primary vertex The event must have a primary vertex. The primary vertex is the vertex of the event with the highest Σp_T^2 of its associated tracks [36, 37]. The tracks considered for primary vertex construction are a subset of all reconstructed

---

[8] $m_{12}^2 = (E_1 + E_2)^2 - |\mathbf{p}_1 + \mathbf{p}_2|^2 = m_1^2 + m_2^2 + 2\,(E_1 E_2 - \mathbf{p}_1 \cdot \mathbf{p}_2)$

tracks. They must pass impact parameter requirements and meet minimum numbers of hits in various layers of the Inner Detector. Candidate vertices are reconstructed by repeatedly fitting these tracks with a seed vertex position, weighing the tracks by compatibility with the position, and re-computing the vertex. In addition to being the vertex with the highest Σp_T^2, the primary vertex must have at least two associated tracks.

Trigger At least one of the triggers listed in Table 4.1 must have fired in the event.

Detector errors The following flags must be checked on an event-by-event basis:

- Liquid argon calorimeters must have no noise bursts and no data corruption,
- Tile calorimeters must have no data corruption,
- No single-event upsets in the Semiconductor Tracker,[9]
- No events with missing detector information (if the timing-trigger-control systems are restarted to recover from a busy detector condition, the subsequent events in the luminosity block may be incomplete).

5.4 Selection of electrons and muons

The events that pass pre-selection are checked for the existence of electrons and/or muons. The final states of $h \to Z_d Z_d \to 4\ell$ considered in this search are four muons, four electrons, and two electrons and two muons. Any electrons and muons that are found are then subject to a selection criteria. At least four leptons passing the following criteria are required to pass this selection stage.

5.4.1 Electrons

Electrons are reconstructed by identifying and matching together clusters of energy deposits in the electromagnetic calorimeters and tracks in the Inner Detector (ID).

---

[9] Data corruption that is caused by a single ionizing particle interacting with a semiconductor (e.g. detector readout electronics).

Candidates are then calibrated to correct for detector effects and then categorized with a likelihood-based method.[10]

5.4.1.1 Electron reconstruction

Topocluster reconstruction Calorimeter cells with physically significant energy deposits are clustered together into *topoclusters*. The topological clustering algorithm builds topoclusters by selecting seed cells with energy significantly above their expected noise threshold and then iteratively collecting neighbouring cells if they also exceed their noise thresholds. Topoclusters offer a few notable advantages over the fixed-size clusters formed by sliding-window algorithm used in the past:[11]

- Topoclusters are able to reconstruct showers down to $\mathcal{O}(100\ \mathrm{MeV})$ in energy. This enables the recovery of low energy photons from bremsstrahlung in the ID. Topoclusters built from these low energy deposits (*satellite topoclusters*) are later matched with the primary topocluster associated with the candidate (*seed topocluster*) in order to form a *supercluster*. Candidates built from topoclusters collect more energy than those built from fixed-sized clusters and therefore require less correction through calibration post-reconstruction.

- Topoclusters are also able to adapt better to the varying and irregular shapes of showers than fixed-window clusters. Shower sizes vary significantly with the particle's energy and momentum, and showers can deposit energy across several layers of the calorimeter. The dynamic nature of topoclustering makes it possible to closely follow the development of a shower across multiple layers and only collect cells with significant energy. This results in superior resolution and a linear energy response across a wide range of energies.

Track reconstruction Independently of topocluster reconstruction, tracks in the ID are reconstructed [41]. First, pixels and strips with energy deposits are clustered together if they are adjacent [42]. In the pixel detector, one cluster forms one

---

[10] Photons are reconstructed in a similar way. Photons primarily lose energy through i) the photoelectric effect at energies less than approximately 1 MeV, ii) Compton scattering at intermediate energies up to approximately 1 GeV, and iii) pair production at higher energies [1]. Given that photons in ATLAS are in the higher range of energy, energy losses through pair production are the primary concern. For photons, clusters are matched to the vertex at which the photon converts to a e^+e^- pair.

[11] More details on topoclustering can be found in [30, 38, 39]. Details on the fixed-sized clusters and sliding-window algorithm used in 2015–2016 operation can be found in [40].

space-point, and in the Semiconductor Tracker, two clusters form two space-points. *Track seeds* are formed from combining three space-points and are filtered to remove duplicates and low-quality seeds according to momentum and impact parameter. Each filtered track seed, along with other space-points compatible with a preliminary trajectory, are fed into a Kalman filter to build a track candidate. This track building expects the energy loss at successive space-points to follow that of a pion. If it is not possible to build a track with this hypothesis, a second attempt that allows for up to 30% energy loss due to bremsstrahlung at each material intersection is made. Track candidates are then fitted and given a track score that is a function of the χ^2 of the fit and weights of the clusters associated with the track. The scores are used to resolve ambiguities from candidates sharing clusters.

Electrons are much lighter than muons (0.5 MeV vs. 106 MeV). For this reason, energy loss due to bremsstrahlung is more significant for electrons. As an electron loses energy, its p_T decreases, which in turn increases the curvature of its trajectory in the presence of a magnetic field. This effect is accounted for by re-fitting tracks that loosely match fixed-window calorimeter clusters in η and ϕ and have at least four silicon hits [38] with a Gaussian-sum filter [43, 44]. This filter approximates the Bethe-Heitler model for bremsstrahlung as a sum of Gaussians. The re-fit improves the track parameters and consequently the electron identification efficiencies.

Track-to-topocluster matching Tracks are then extrapolated to the calorimeter and matched to topoclusters by requiring $|\Delta\eta| < 0.05$ and $-0.10 < q \cdot (\phi_{\text{track}} - \phi_{\text{clus}}) < 0.05$, where q is the charge of the track.[12,13] If multiple tracks are matched to a cluster, the tracks are ranked by the number of hits, the hit locations, and ΔR, and the highest-ranked track is chosen.

Supercluster reconstruction In order to recover low energy photons from bremsstrahlung, *satellite topoclusters* are combined with the track-matched topoclusters from the previous step. Topoclusters that match to a track with at least four silicon

---

[12] Matching is attempted with the track's momentum and also with the track's momentum rescaled to match the topocluster's energy. This is to allow for electrons that have lost significant energy through bremsstrahlung in the ID. Note that the pseudorapidity can be expressed as $\eta = \frac{1}{2}\ln\left(\frac{|\mathbf{p}|+p_z}{|\mathbf{p}|-p_z}\right) = \operatorname{arctanh}\left(\frac{p_z}{|\mathbf{p}|}\right)$, where the polar angle θ is the angle between particle's three-momentum $\mathbf{p}$ and the positive direction of the beam axis.

[13] The second requirement is asymmetric because extrapolated tracks can miss some of the energy lost through bremsstrahlung.

hits and have $E_T > 1$ GeV[14] are considered as *seed topoclusters*. For each seed topocluster, topoclusters which are i) within a $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ cone[15] or ii) within an extended window of 0.125×0.300 and have the same track match as the seed are considered satellites.

Track-to-supercluster matching Finally, superclusters are matched to tracks in the same way that topoclusters are matched to tracks previously. Those with a successful match are deemed electron candidates.

5.4.1.2 Electron calibration

Electron candidates are then calibrated to correct differences in energy response across the layers of the calorimeter, differences between data and simulation, and residual effects [38, 45, 46]:

1. The difference in energy response between the first and second layers of the EM calorimeter is corrected in data. This correction is derived by comparing, between data and simulation, the energy deposits of muons in $Z \to \mu\mu$ events in the two layers. Muons are used because they lose minimal energy in the material in front of the calorimeter.

2. The energies of the constituent topoclusters in both data and simulation are calibrated with corrections derived only from simulation. These corrections, binned in p_T and η, are derived with a multivariate analysis of simulated electrons[16] that takes into account ratios of topocluster energy deposition between different layers of the calorimeter.

3. Residual effects are corrected in data, such as i) energy shifts due to pileup (these shifts vary due to bunch-to-bunch luminosity variations), ii) geometric

---

[14] Electron and photon candidates are built from calorimeter cells and clusters with energy measurements, while muon candidates are built from tracks with momentum measurements. Therefore, it is customary to state transverse energy (E_T) for electrons and photons and transverse momentum (p_T) for muons. This distinction is not significant because photons are massless and both types of leptons have small masses relative to the energy scale of ATLAS ($\mathcal{O}$(GeV) and greater).

[15] As shown in Figure 4.4, each cell in the second layer of the electromagnetic (EM) calorimeter has dimensions $\Delta\eta = 0.025$ and $\Delta\phi = 0.0245$. Therefore, this region corresponds to three (five) units of cells in η (ϕ). As described in Section 4.2.3, the window is asymmetric because the magnetic field of the ID causes energy depositions from interactions with the ID to be spread in the ϕ-direction.

[16] Also applies to photons.

effects (variations in gaps between absorbers), and iii) differences in high-voltage settings.

4. The energies of electrons in data are calibrated with corrections derived from $Z \rightarrow ee$ events in data and simulation. The events in both are split into regions corresponding to the pseudorapidities of the two electrons (η_i, η_j) and distributions of the dielectron mass are created for each region. The average dielectron mass in data is parametrized as $m_{ij}^{\text{data}} = m_{ij}^{\text{sim}}(1 + \alpha_{ij})$, where α_{ij} is the shift in the average mass needed to correct the distribution of simulated events to the distribution of observed events (or vice versa). The corrections α_{ij} are determined by χ^2 fitting of the distributions of simulation to data. The energy resolution in simulation is calibrated with corrections derived similarly, with the parametrization $(\sigma_m/m)_{ij}^{\text{data}} = (\sigma_m/m)_{ij}^{\text{sim}} \oplus c_{ij}$.

5.4.1.3 Electron identification

Following calibration, the candidates are identified with a likelihood approach that incorporates shower shape, ratios of energy deposition between different layers of the calorimeter, track hits, and differences in kinematic quantities between the track and clusters [38, 40]. Multi-dimensional probability distribution functions (PDFs) derived from simulation are computed for signal electrons and background consisting of jets faking electrons, electrons from photon conversion due to interactions with the beam pipe, and non-prompt electrons from heavy flavour decays. A discriminant that is the ratio of the signal PDF to the sum of the signal PDF and background PDF is transformed into a sigmoid distribution and is evaluated for each electron candidate. Values of this transformed discriminant are chosen as electron identification working points. These working points are choices in electron identification efficiency and background rejection. This likelihood-based approach differs from the cut-based approach used for muon identification and has two notable advantages: no single discriminant alone can cause an electron to fail identification, and discriminants that share similar distributions and would cause significant losses in efficiency if used together can be combined in a likelihood without such a penalty.

Electrons considered in the $h \rightarrow Z_d Z_d \rightarrow 4\ell$ search are required to pass the *Loose* working point.[17] This working point requires at least two hits in the pixel detector,

---

[17] This requirement is imposed at the end of the $h \rightarrow Z_d Z_d \rightarrow 4\ell$ selection in the *event selection* in order to avoid double-counting an event in both the signal region and the regions of the data-driven fake background estimate. A detailed explanation is given in Section 5.7.

seven hits total in the pixel and silicon-strip detectors, and one hit in the inner-most pixel layer (B-Layer). The last requirement reduces the background from photon conversions. As shown in Figure 5.1, the efficiency of this working point varies from approximately 85% at $E_T = 20$ GeV to 96% at $E_T = 100$ GeV compared to 55% to 90% for the *Tight* working point. The increased efficiency comes at the cost of decreased background rejection—the rejection factor for jets mis-identified as electrons is approximately five times less for *Loose* compared to *Tight*. As shown in Figure 5.2, the dielectron mass resolution is approximately 2 GeV in $Z \to ee$ events.

5.4.1.4 Electron isolation

Electrons are also required to be well-isolated in the detector. Particles that come from signal processes are characterized by signatures in the detector that have little activity around them. These *isolated* signatures are distinct from the less well-isolated signatures left by background processes such as non-prompt decays (e.g. heavy flavour decays),[18] jets misidentified as leptons, and photons converting to e^+e^- pairs. In order to help differentiate between signal particles and background, requirements on the maximum amount of activity around the particle in the calorimeter and around its tracks in the ID are applied. Electrons in the $h \to Z_d Z_d \to 4\ell$ search are required to pass the `FCLoose` isolation working point.[19] For this working point, the calorimeter and track isolation requirements for electrons are:

- Calorimeter isolation (`topoetcone20`/$p_T < 0.2$):
    1. The E_T of clusters within a cone of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.2$ around the electron are summed.
    2. In order to remove the energy deposits of the electron itself, the E_T of clusters in a $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ cone around the particle are subtracted from the sum (core subtraction). A *leakage* correction derived from simulation to compensate for deposition of energy outside of this cone is also subtracted from the sum.

---

[18] For example, a b-quark decaying to a muon and a c-quark. The c-quark could also decay to a muon and other light hadrons. This is a *double semi-leptonic* decay.

[19] As for electron identification, electron isolation is imposed in the event selection in order to avoid double-counting an event in both the signal region and the regions of the data-driven fake background estimate. A detailed explanation is given in Section 5.7.

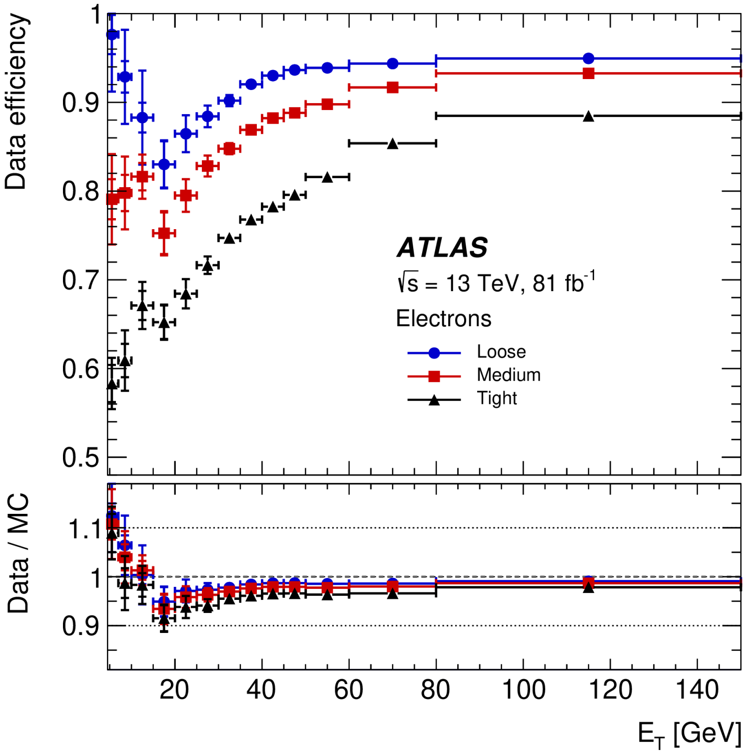


Figure 5.1: Electron identification efficiencies in $Z \to ee$ events in 2015–2017 data and simulation for the *Loose*, *Medium*, and *Tight* likelihood-based working points as a function of E_T [38].

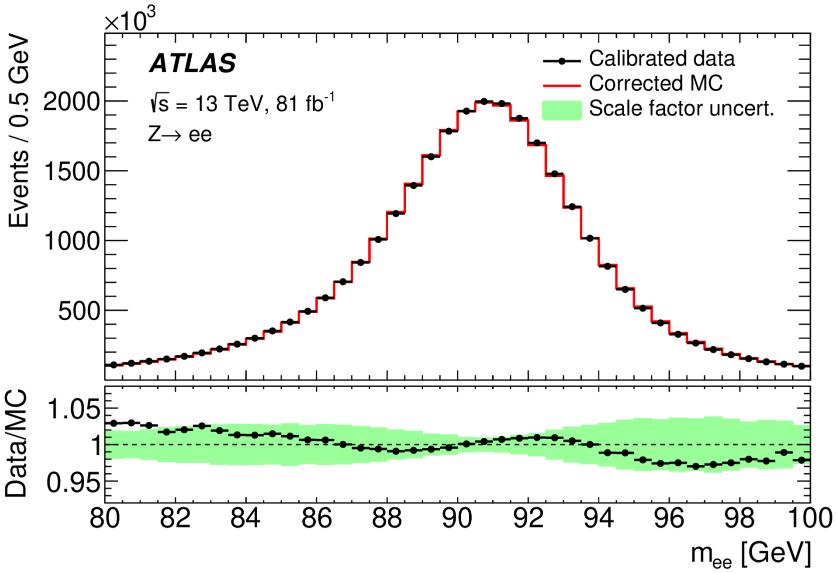


Figure 5.2: m_{ee} distribution in $Z \to ee$ events in 2015–2017 data and simulation [38].

3. A correction for energy deposited by pileup and the underlying event is subtracted from the sum. This correction is derived from the energy density measured across the entire calorimeter [47].
4. The resulting sum `topoetcone20` is divided by the p_T of the electron and the final value is required to be less than 0.2. The abbreviation `FC` in the name of the working point refers to the fixed cut of 0.2.

- Track isolation (`ptvarcone20_TightTTVA_pt1000`/$p_T < 0.15$):

1. The p_T of tracks that satisfy $p_T > 1$ GeV, $|\eta| < 2.5$, $|z_0 \sin\theta| < 0.5$ mm, and track quality requirements (minimum number of hits and maximum number of holes in various layers) within a cone around the electron are summed. The cone size is a function of the electron's p_T and becomes smaller as p_T increases.[20] The abbreviation `TightTTVA` in the variable name refers to the track-to-vertex association requirement $|z_0 \sin\theta| < 0.5$ mm.
2. The p_T of the electron's track is subtracted from the sum.
3. A correction for bremsstrahlung radiation is subtracted from the sum. This correction is calculated by extrapolating all tracks into the second layer of the EM calorimeter and summing the p_T of extrapolated tracks within a $\Delta\eta \times \Delta\phi = 0.05 \times 0.1$ cone around the electron.
4. The resulting sum `ptvarcone20_TightTTVA_pt1000` is divided by the p_T of the electron and the final value is required to be less than 0.15.

5.4.1.5 Kinematic requirements on electrons

In addition to passing identification and isolation, electrons are also required to pass the following:

$p_{\mathrm{T}} > 7$ GeV The transverse momentum must be greater than 7 GeV. In *Run-1* operation (2010–2012), this was the minimum p_{T} supported by electron identification. In *Run-2* operation (2015 and onwards), this was lowered to 4.5 GeV, but the $h \to Z_d Z_d \to 4\ell$ search did not change accordingly; as shown in Figure 5.1, efficiency decreases steeply as p_{T} decreases due to increases in pileup, multi-jet background, and jets being misidentified as electrons.

---

[20] Cone sizes smaller than the $\Delta R = 0.2$ size used for calorimeter isolation can be used for track isolation due to the higher granularity of the ID. For track isolation, $\Delta R = \min\left(\frac{10\ \text{GeV}}{p_T}, R_{\max}\right)$.

$|\eta| < 2.47$ The absolute pseudorapidity of the calorimeter cluster forming the electron candidate must be less than 2.47. For this cut, the pseudorapidity is calculated from the barycentre of the cluster in the second calorimeter sampling layer and the centre of the ATLAS detector, instead of from the corresponding track and the primary vertex. Electrons that are in the calorimeter crack region $1.37 < |\eta| < 1.52$ (the transition region between the barrel and endcap electromagnetic calorimeters) are included.

Impact parameter $|z_0 \sin\theta| < 0.5$ mm The longitudinal impact parameter z_0 multiplied by $\sin\theta$ (polar angle) of the track must be less than 0.5 mm. Impact parameter requirements select leptons that originate from the hard scatter of the event. The longitudinal impact parameter is calculated in the xy plane and relative to the beam spot.[21]

No bad calorimeter cluster A calorimeter cluster is considered bad if it is in a region with a non-operational calorimeter front-end board or without high-voltage supply, or if it contains a calorimeter cell that has been masked off (usually due to noise bursts).

5.4.2 Muons

Muons are reconstructed by forming tracks from hits in the Inner Detector (ID) and Muon Spectrometer (MS) independently, and then combined. After reconstruction, muon candidates are calibrated to account for detector effects and categorized using a cut-based method.

5.4.2.1 Muon reconstruction

Muon reconstruction in the ID The tracks used in the first stage of track reconstruction for electrons are also used for muon reconstruction. At the energies relevant to ATLAS, muons behave like minimally-ionizing particles with minimal energy losses due to bremsstrahlung.

---

[21] In *Run-2*, it was recommended to calculate the impact parameter relative to the beam spot instead of relative to the primary vertex (as was done in *Run-1*). This is because the expected size of the beam spot is much smaller than the primary vertex resolution in the transverse direction obtained from track information only.

Muon reconstruction in the MS In each station (RPC/MDT or RPC/MDT/RPC), a Hough transform [48] is used to search for hits that are aligned in a straight line in the non-bending plane (xy). In this plane, trajectories are relatively straight. In the bending plane (rz) however, muons follow a curved trajectory. This trajectory is reconstructed from multiple straight-line *track segments* starting with the hits found by the Hough transform. MDT track segments are found by fitting straight lines across the hits found in each multi-layer (three layers of tubes). CSC track segments are found by performing a combinatorial search in η and ϕ. Track candidates are then built using a segment-seeded combinatorial search. This search is seeded by segments found in the middle layer where more hits are registered. Each segment in this layer is extrapolated into the outer and inner layers—while accounting for the magnetic field—in order to search for matching segments. At least two matching segments are required to build a candidate[22] and candidates are allowed to share segments initially (duplicates are removed after ID-MS matching and muon identification). The candidates are then fitted with a χ^2 fit. Hits that provide a large contribution to the χ^2 are removed and the fit is repeated. A hit recovery procedure is also performed in order to recover hits that are consistent with the trajectory. This procedure looks for hits in the vicinity of chambers that are traversed by the track but do not have any hits. Hits that are found are added to a re-fit. Finally, track candidates are accepted according to lowest χ^2.

Muon reconstruction in the ID and MS combined As described in Section 4.2.4, muon reconstruction also combines track candidates from the ID and MS as well as energy deposits in the calorimeter. Four types of muons are defined [34]:

- Combined: a track is formed from a re-fit across hits found in both the ID and MS. Most muons of this type are found by first starting with tracks reconstructed in the MS and then extrapolating inward and matching to a track reconstructed in the ID.

- Segment-tagged: a track reconstructed in the ID is extrapolated into the MS and is matched to at least one segment in the MDTs or CSCs. This type of muon recovers acceptance for low p_{T} muons ($p_{\mathrm{T}} \lesssim 8$ GeV) that do not cross

---

[22] Except for in the barrel-endcap transition region $1.3 < |\eta| < 1.65$, where only a single high-quality segment is required to form a track candidate. In this region, the tracks are nearly straight because of inhomogeneities in the magnetic field. The overlapping of the barrel and end-cap toroid fields leads to a reduction in the integrated bending power [22].

more than one layer of the MS and in regions that are not fully instrumented (calorimeter and ID cabling and other services). Segment-tagged muons are considered only in $|\eta| < 0.1$.

- Calorimeter-tagged: a track reconstructed in the ID is matched with an energy deposit in the calorimeter that is compatible with a minimally-ionizing particle. This type of muon recovers acceptance in regions of the MS that are not fully instrumented. Calorimeter-tagged muons have the lowest purity of all types and are considered only in $|\eta| < 0.1$.
- Extrapolated (*stand-alone*): a track reconstructed in the MS is compatible with originating from the interaction point. This extrapolation takes into account energy loss in the calorimeters (mainly via ionization [1]). This type of muon recovers acceptance in the region $2.5 < |\eta| < 2.7$ that is not instrumented in the ID.

If muons of different types share the same ID track, preference is given in the above order. Extrapolated muons that share MS tracks are resolved by selecting the muon candidate that has the best track fit and highest number of hits.

5.4.2.2 Muon calibration

Muon candidates in simulation are then calibrated to correct differences in momentum scale and resolution between data and simulation [34]. These differences arise from inaccuracies in the simulation of the magnetic field description, the geometric description of the MS, and material effects (mainly energy loss in the calorimeter). In an approach similar to the approach used to derive electron calibration corrections with $Z \to ee$ events, muon momentum calibrations are derived with $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events containing only combined muons in data and simulation. Both populations are split into regions of η and ϕ, and the transverse momentum is parametrized in terms of corrections for each of the differences described previously. The corrections are derived using maximum-likelihood fits[23] of the dimuon mass distributions of simulation to data.

---

[23] These fits are similar to the fits performed in the interpretation of the search results (Chapter 9).

5.4.2.3 Muon identification

Following calibration, muons are identified with a cut-based approach [34]. The main backgrounds to prompt muons are decays of hadrons: charged pion decays ($\pi^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu})$) and kaon decays (e.g. $\mathrm{K}^{+} \rightarrow \mu^{+} + \nu_{\mu}$ and kaons decaying to charged pions). The displacement of these decays from the interaction point (*non-prompt decays*) causes the reconstructed track to have a kink. These tracks can be distinguished from the tracks of prompt muons by their lower fit quality and incompatibility in momentum between the constituent ID and MS tracks. For combined muons, three variables are considered:

- q/p significance: the absolute value of the difference between the ratios of charge to momentum measured in the ID and MS, normalized by the sum in quadrature of their uncertainties
- ρ': the absolute value of the difference between the p_T measured in the ID and MS, normalized by the p_T of the combined track
- χ^2 of the combined track fit

Requirements on the minimum number of hits and maximum number of holes[24] in the ID and MS are also imposed.

Muon identification is separated into four working points: *Loose*, *Medium*, *Tight*, and *High-*p_{T}. The working points impose requirements on the muon type in addition to those mentioned above. Muons considered in the $h \rightarrow Z_d Z_d \rightarrow 4\ell$ search are required to pass the *Loose* working point. This working point considers all muon types:

- Combined muons must have at least three hits in at least two MDT multi-layers, except for tracks in the $|\eta| < 0.1$ region, where tracks with at least one MDT multi-layer but no more than one MDT multi-layer with holes and no hits are alowed. In addition, the q/p significance must be at least seven.
- Extrapolated muons (those with a MS track compatible with the interaction point) must have at least three hits in at least three MDT and/or CSC layers and are considered only in the region without ID acceptance ($2.5 < |\eta| < 2.7$).
- Calorimeter-tagged muons only in the region in which they are defined ($|\eta| < 0.1$)

---

[24] An active sensor that is traversed but not hit.

- Segment-tagged muons, similarly ($|\eta| < 0.1$)

The *Loose* working point was specifically optimized for reconstructing 4ℓ final states originating from Higgs decays. In the region $|\eta| < 2.5$, approximately 97.5% of *Loose* muons are combined, 1.5% are calorimeter-tagged, and 1% are segment-tagged. As shown in Figure 5.3, the efficiency of this working point is over 98% for most muons (4 GeV $< p_T < 100$ GeV). As shown in Figure 5.4, the dimuon mass resolution is less than 2 GeV in $Z \rightarrow \mu\mu$ events.

5.4.2.4 Muon isolation

As with electrons, muons are also required to be well-isolated in the detector. Muons in the $h \rightarrow Z_d Z_d \rightarrow 4\ell$ search are required to pass the `FCLoose_FixedRad` isolation working point.[25] The calorimeter and track isolation requirements applied to muons are similar to those applied to electrons:

- Calorimeter isolation (`topoetcone20`$/p_T < 0.3$):
    1. The muon's track is extrapolated into the EM calorimeter and the E_T of topoclusters within a cone of $\Delta R < 0.2$ around the track are summed.
    2. In order to remove the energy deposits of the muon itself, the E_T of topoclusters within $\Delta R < 0.05$ are subtracted from the sum.
    3. A correction for energy deposited by pileup and the underlying event is subtracted from the sum (estimated in a similar way as for electron calorimeter isolation).
- Track isolation (`ptvarcone30_TightTTVA_pt1000`$/p_T < 0.15$ for $p_T < 50$ GeV, `ptcone20_TightTTVA_pt1000`$/p_T < 0.15$ otherwise): in the same manner as electron track isolation. For improved background rejection at high transverse momentum ($p_T > 50$ GeV), a fixed-size cone is used instead of a cone whose size decreases as p_T increases.

5.4.2.5 Kinematic requirements on muons

In addition to passing identification and isolation, muons are also required to pass the following:

---

[25] As for electron identification and isolation, muon isolation is applied in the event selection in order to avoid double-counting an event in both the signal region and regions of the data-driven fake background estimate. A detailed explanation is given in Section 5.7.

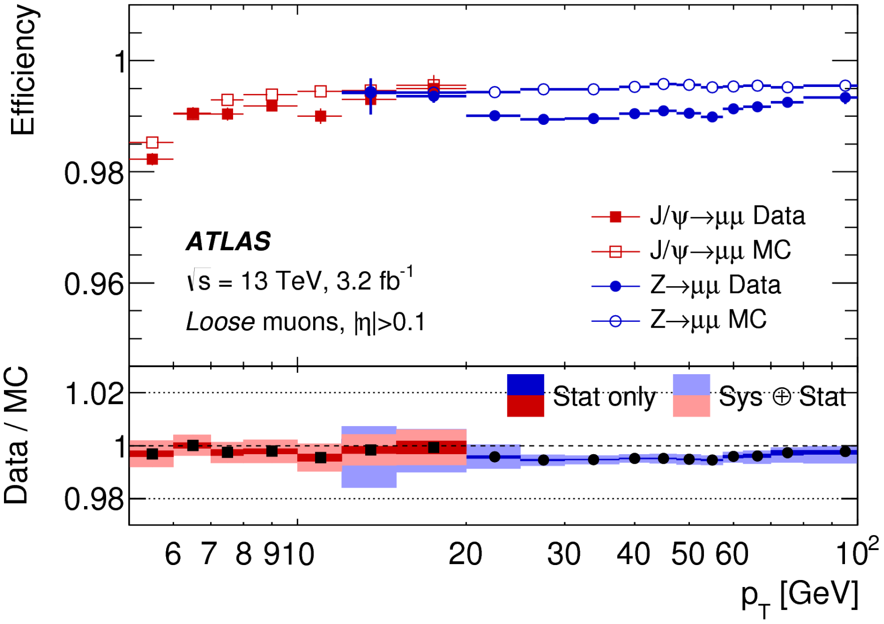


Figure 5.3: Muon identification efficiencies in $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events in early 2015 data and simulation for the *Loose* cut-based working point as a function of p_T [34].

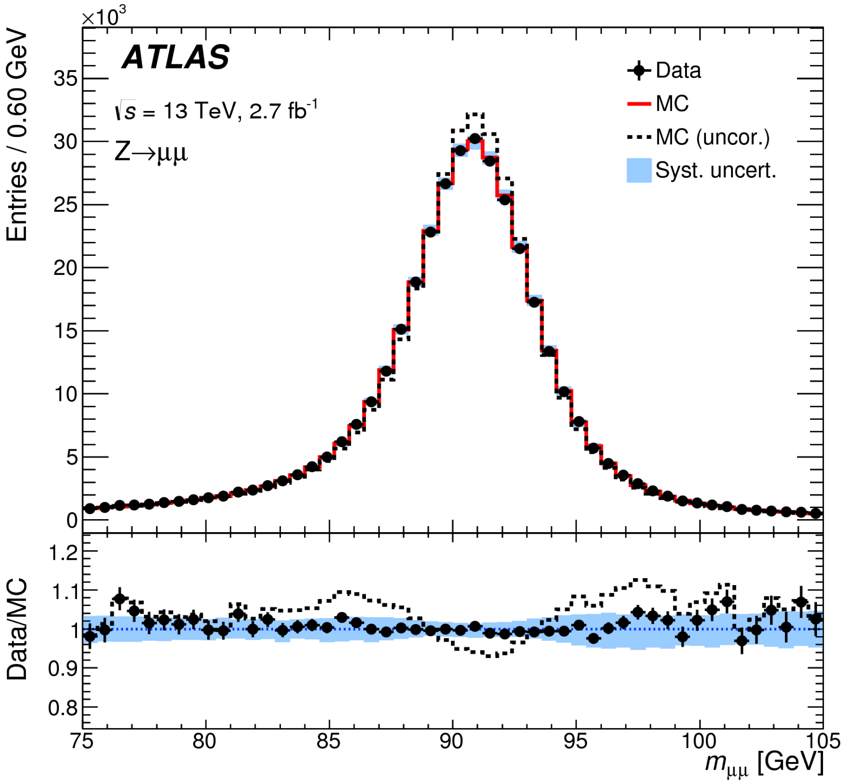


Figure 5.4: $m_{\mu\mu}$ distribution in $Z \to \mu\mu$ events with combined muons in early 2015 data and simulation [34].

$\boldsymbol{p_\mathrm{T} > 5}$ **GeV (15 GeV *if calorimeter-tagged*)** Below approximately 5 GeV, muon reconstruction efficiency decreases and fake backgrounds increase. Calorimeter-tagged muons are optimized for 15 GeV $< p_T <$ 100 GeV.

$\boldsymbol{|\eta| < 2.7}$ The region $2.5 < |\eta| < 2.7$ is not covered by the ID. Muons extrapolated from the MS recover acceptance in this region.

Impact parameters $\boldsymbol{|z_0 \sin\theta| < 0.5}$ **mm,** $\boldsymbol{|d_0| < 1}$ **mm** These requirements are similar to those applied to electrons. Muons extrapolated from the MS are exempt from both requirements because they do not have ID tracks. The second requirement rejects muons from cosmic rays.

5.5 Overlap removal

Overlap removal is applied in order to avoid reconstructing the same detector signature as multiple electrons or muons. The leptons passing the requirements above are examined for overlaps in the following order:[26]

1. Electron-electron: if two electrons share a track and/or if the cluster associated to each electron are within $\Delta\eta = 3 \times 0.025$ and $\Delta\phi = 5 \times 0.025$ of each other, then the two electrons are considered as overlapping. To resolve the overlap: if one of the electrons is considered *ambiguous*,[27] then the other electron is accepted, followed by accepting the electron with the higher p_T.
2. Muon-electron: if an electron and a muon share a track, they are considered as overlapping. To resolve the overlap: if the muon is a calorimeter-tagged muon, then electron is accepted, otherwise the muon is accepted.

5.6 Ranking and selection of quadruplets

All possible same-flavour opposite-sign dileptons are formed from the selected leptons. These dileptons can be (e^+e^-) and $(\mu^+\mu^-)$. These dileptons are paired to form

---

[26] As described in Section 5.4.2.1, muon-muon overlaps are resolved during muon reconstruction.

[27] Superclusters for electrons and photons are built independently, so it is possible for an electron and a photon to share the same seed topocluster. If it is not possible to identify whether the object is an electron (a cluster with a track and no photon conversion vertex) or a photon (cluster alone), then the electron is labelled as ambiguous.

quadruplets $(e^+e^-)(e^+e^-)$, $(e^+e^-)(\mu^+\mu^-)$, and $(\mu^+\mu^-)(\mu^+\mu^-)$. The dilepton with mass closest to the mass of the SM Z boson $m_Z \equiv 91.1876$ GeV [1] is labelled ℓ_{12} with mass m_{12} and the other pair is labelled ℓ_{34} with mass m_{34} with the labels ℓ_1 and ℓ_3 corresponding to the positively charged lepton of each pair. The first pair is referred to as the *primary* or *leading* pair and the second pair is referred to as the *secondary* or *sub-leading* pair. In case of observation of Z_d decaying to two leptons, the dilepton that is reconstructed is intended to correspond to the two leptons from the Z_d decay. Therefore, the masses of the dileptons (m_{12} and m_{34}) are intended to correspond to the masses of the observed Z_dZ_d.

Alternatively, the positively charged lepton from the primary (secondary) pair can be paired with the negatively charged lepton of the secondary (primary) pair to create the alternative pairing ℓ_{14} (ℓ_{32}).

The quadruplets are ranked by the absolute difference in mass between the primary and secondary pairs $\Delta m_{\ell\ell} \equiv |m_{12} - m_{34}|$. The smaller the difference, the higher the ranking of the quadruplet. This follows from the hypothesis that the intermediate particles in the decay, Z_dZ_d, have similar mass. Quadruplets with smaller mass difference are preferred over those with larger mass difference.

After formation and ranking, the *quadruplet selection* applies kinematic, reconstruction, and trigger requirements on each quadruplet:

- The three leading-p_T (E_T) leptons must have p_T (E_T) > 20 GeV, 15 GeV, 10 GeV.[28]
- In case of four muons, at least three must be combined muons (reconstructed by combining ID and MS).[29]
- The leptons of the quadruplet must have been responsible for firing at least one trigger (for multi-lepton triggers, all leptons of the trigger must match to leptons of the quadruplet).[30]

---

[28] The p_T requirement is applied to muons and the E_T requirement is applied to electrons.

[29] In other words, the number of extrapolated (stand-alone) muons plus calorimeter-tagged muons cannot be greater than one.

[30] Put another way: at least one of the triggers must have been fired by the lepton(s) of the quadruplet (and not in combination with leptons not belonging to the quadruplet). For example, if one lepton from the quadruplet fired a single lepton trigger, the quadruplet is qualified as trigger matched. If two leptons from the quadruplet together fired a dilepton trigger, the quadruplet is also qualified as trigger matched. If one lepton from the quadruplet in combination with another lepton not belonging to the quadruplet together fired a dilepton trigger, the quadruplet is not qualified as trigger matched by this trigger. If only one lepton of a dilepton trigger is matched, the quadruplet is also not qualified as trigger matched by this trigger.

- $\Delta R(\ell, \ell') > 0.10$ (0.20) for same-flavour (different-flavour) leptons. This is to ensure that the leptons are well-separated.

The highest ranking quadruplet (smallest $\Delta m_{\ell\ell}$) that survives the selection is chosen as the quadruplet of the event. If no quadruplet is selected, then the event is discarded. At this stage, events can be classified into three flavour channels: 4μ, $4e$, and $2e2\mu$.

5.7 Event selection

The event selection imposes further requirements on the selected quadruplet of an event that are more specific to the $h \to Z_d Z_d \to 4\ell$ search channel.[31] The events that pass the event selection are the events of the signal region. The first set of requirements are on the leptons of the selected quadruplet, and the second set of requirements are on the kinematics of the quadruplet:

- All electrons of the quadruplet must pass the *Loose* identification working point. This requirement is imposed in the event selection instead of earlier in order to avoid double-counting of events between the signal region and regions of the data-driven fake background estimate (Chapter 7). Electron identification is reversed as part of the data-driven fake background estimation. If this requirement was imposed earlier instead of at event-level, it would be possible for an event to have one quadruplet that passes the requirement and enters the signal region, and a different quadruplet *in the same event* that fails the requirement and enters one of the regions used in the estimate. This event would therefore be double-counted. For muons, the analogous identification requirement is imposed earlier in the lepton selection because it is not reversed.

- All leptons of the quadruplet must pass isolation requirements: `FCLoose` for electrons and `FCLoose_FixedRad` for muons.

- All leptons of the quadruplet must pass impact parameter significance requirements: $|d_0|/\sigma(d_0) < 5$ for electrons and $|d_0|/\sigma(d_0) < 3$ for muons. These impact parameter requirements are calculated in the rz plane and relative to the beam spot.[32]

---

[31] As distinct from the other channels listed in Section 5.1.

[32] Impact parameter significance is required in the event selection instead of earlier for the same reason as electron identification.

- Higgs boson mass window: 115 GeV $< m_{4\ell} <$ 130 GeV (applied to the invariant four-lepton mass $m_{4\ell}$)
- Z boson veto: 10 GeV $< m_{12,34} <$ 64 GeV and 5 GeV $< m_{14,32} <$ 75 GeV
- Quarkonia veto: $(m_{J/\Psi} - 0.25\text{ GeV}) < m_{12,34,14,32} < (m_{\Psi(2S)} + 0.30\text{ GeV})$ and $(m_{\Upsilon(1S)} - 0.70\text{ GeV}) < m_{12,34,14,32} < (m_{\Upsilon(3S)} + 0.75\text{ GeV})$
- Low-mass veto: $m_{12,34,14,32} > 5$ GeV
- Medium signal region: $m_{34} > 0.85 m_{12} - 0.1125 f(m_{12}) m_{12}$

Higgs boson mass window This requirement on the quadruplet invariant mass follows from the hypothesis that the progenitor of $Z_d Z_d$ is a SM Higgs with mass $m_h \approx$ 125 GeV. This window is asymmetric around m_h in order to cover bremsstrahlung losses in final states with electrons and is also applied in SM $h \to ZZ^* \to 4\ell$ measurements [49, 50].

***Z* boson veto** This requirement on the dilepton masses vetoes processes that produce Z bosons. The two dominant backgrounds in the $h \to Z_d Z_d \to 4\ell$ signal region are $h \to ZZ^* \to 4\ell$ and $ZZ^* \to 4\ell$.[33] This veto is carried over from the 2015–2016 $h \to Z_d Z_d \to 4\ell$ search [6]. It was optimized by choosing the values of the veto that maximize the expected[34] discovery significance [1, 51]:

$$\text{med}[Z_0|\mu = 1] = \sqrt{2\left((s+b)\ln\left(1+\frac{s}{b}\right) - s\right)} \tag{5.2}$$

Any of the dilepton pairings $\ell_{12}, \ell_{34}, \ell_{14}$, and ℓ_{32} can reconstruct a Z boson. In the $4e$ and 4μ channels, it is possible to mis-pair the leptons coming from the intermediate particles (Z or Z_d). In other words, a lepton from one Z_d decay may be incorrectly paired with a lepton coming from the other Z_d decay, instead of with the other lepton coming from the same Z_d decay. The rate of mis-pairing is estimated to be

---

[33] The latter is also known as non-resonant/continuum ZZ^* production.

[34] More precisely, the *median* discovery significance. This approach characterizes the *experimental sensitivity*. The significance Z_μ is a function of the observed data and is therefore a random variable with a distribution. This distribution corresponds to a value of signal strength μ (typically $\mu = 0$ for the background-only hypothesis and $\mu = 1$ for the nominal signal hypothesis). Therefore, one can consider the distribution of Z_μ as a function of μ, $f(Z_\mu|\mu')$, and identify its median value as the expected significance. The median is preferred over the mean because the former is always invariant under monotonic transformations such as between the p-value and significance. This is not always true for the latter.

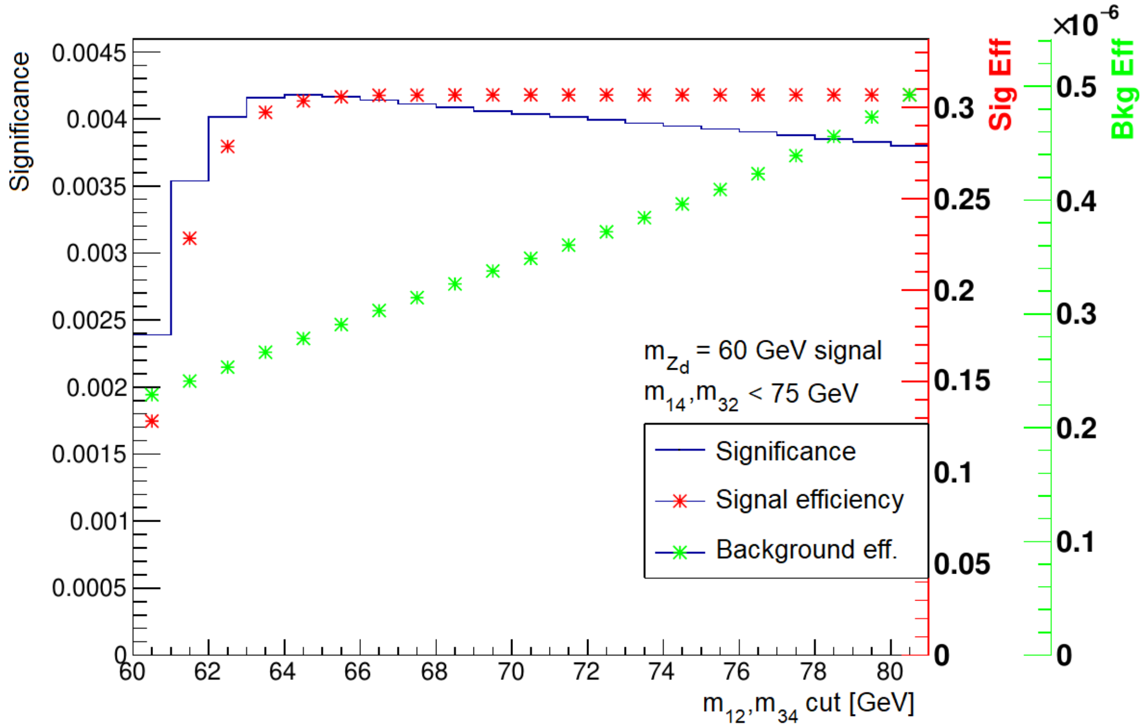


Figure 5.5: Expected discovery significance for $m_{Z_d} = 60$ GeV signal as a function of m_{12}, m_{34} requirement, with $m_{14}, m_{32} < 75$ GeV imposed [52]. The signal (background) efficiency is the signal (background) yield after imposing the requirement on m_{12}, m_{34} divided by the yield before.

approximately 8% (5%) in the $4e$ (4μ) channels for $m_{Z_d} = 60$ GeV signal, falling to 2% (1%) for $m_{Z_d} = 15$ GeV. Therefore, requirements on m_{14}, m_{32} are applied in addition to requirements on m_{12}, m_{34} in order to veto background events in which the leptons from ZZ^* decay are mis-paired.

For the highest m_{Z_d} signal hypothesis considered, $m_{Z_d} = 60$ GeV, the dilepton masses m_{12} and m_{34} do not exceed 65 GeV. Therefore, no requirement is needed to recover events with m_{12}, m_{34} above $m_Z \approx 91.2$ GeV. From the same signal, there is no contribution to m_{14} and m_{32} above m_Z, so no requirement on m_{14}, m_{32} is needed to recover events with m_{14}, m_{32} above m_Z, either. If the requirement $m_{14}, m_{32} < 75$ GeV is imposed, choosing $m_{12}, m_{34} < 64$ GeV maximizes the significance for $m_{Z_d} = 60$ GeV signal, as shown in Figure 5.5. If the latter requirement is imposed, choosing $m_{14}, m_{32} < 80$ (70) GeV maximizes the significance for $m_{Z_d} = 15$ (60) GeV signal. In order to keep the requirement consistent for all m_{Z_d} considered, $m_{14}, m_{32} < 75$ GeV was chosen, along with $m_{12}, m_{34} < 64$ GeV.

Quarkonia veto This requirement on the dilepton masses vetoes quarkonia decay processes (e.g. $J/\Psi \to e^+e^-$ $(\mu^+\mu^-)$). The masses of J/Ψ, $\Upsilon(2S)$, $\Upsilon(1S)$, and $\Upsilon(3S)$ are taken from [1].

Low-mass veto This requirement on the dilepton masses ensures that the signal region does not extend into a kinematic region that is not fully covered by the simulation-based $ZZ^* \to 4\ell$ background estimate. This background was simulated with a generator-level $m_{\ell\ell} > 4$ GeV requirement. In other words, the kinematic region with dilepton mass below 4 GeV was not simulated. The low-mass veto limits the kinematic region to $m_{\ell\ell} > 5$ GeV. Without it, discrepancies between data and expected background are observed. This veto renders the J/ψ component of the quarkonia veto redundant, but for clarity and harmonization with the $h \to Z_dZ_d \to 4\mu$ search (*low-mass* search), the quarkonia veto continues to include a J/ψ component.[35]

Medium signal region This is a requirement on the compatibility of m_{12} and m_{34}. It follows from the hypothesis that the two Z_d particles are of similar mass. In the 2015–2016 iteration of this search, this requirement was simply $m_{34}/m_{12} > 0.85$, but it is re-optimized for the 2015–2018 iteration to allow for Z_d with larger width. In the (m_{12}, m_{34}) plane depicted in Figure 5.6, the former signal region is in red and the addition to it from re-optimization is in green.

$m_{34}/m_{12} > 0.85$ This requirement follows from the hypothesis that the Z_d width is narrow and that the dilepton mass resolution is approximately 15%. Initial theoretical models hypothesized that the Z_d width is relatively narrow ($\Gamma_{Z_d} \sim \mathcal{O}(10^{-9})$ GeV [10, 53] compared to $\Gamma_h \sim 4 \times 10^{-3}$ GeV [1, 54]). Therefore, the observed width would be dominated by detector resolution. With the assumption that the lepton energy resolutions are $\Delta E_E \approx 3.5\%$ and $\Delta E_\mu \approx 3.8\%$,[36] it follows that a $\pm 2\sigma$ window around the dilepton mass is approximately 15%.

---

[35] The low-mass channel searches for Z_d with mass between 1 GeV and 15 GeV and therefore does not apply a low-mass veto. In this mass range, heavy flavour background becomes significant. This background is not well-modelled by simulation. For this reason, a data-driven technique is used to estimate the $ZZ^* \to 4\ell$ background instead of simulation.

[36] More recent references on electron [38] and muon [34] reconstruction performance place the energy and momentum resolutions between 2% to 3%, depending on E_T (p_T), η, and the process producing electrons and muons.

$m_{34} > 0.85m_{12} - 0.1125f(m_{12})m_{12}$ The aformentioned requirement is re-optimized to allow for Z_d with larger width. A $\pm 3.5\sigma$ window increases the signal efficiency for $m_{Z_d} = 15$ GeV signal with minimal increase in background, but increases the background significantly for $m_{Z_d} = 60$ GeV. Instead of a fixed cut on m_{34} and m_{12}, a function that decreases the window width as m_{Z_d} and background increases is used. The function applies a $\pm 3.5\sigma$ window starting at $m_{Z_d} = 10$ GeV, gradually decreases the window width to $\pm 2\sigma$ at $m_{Z_d} = 50$ GeV, and applies the original $\pm 2\sigma$ window ($m_{34}/m_{12} > 0.85$) onwards. The change in window width is controlled by a modulating function (Figure 5.7) that varies from 100% to 0% as the background—modelled by a exponential tail matched to a Gaussian—varies from minimum to maximum.

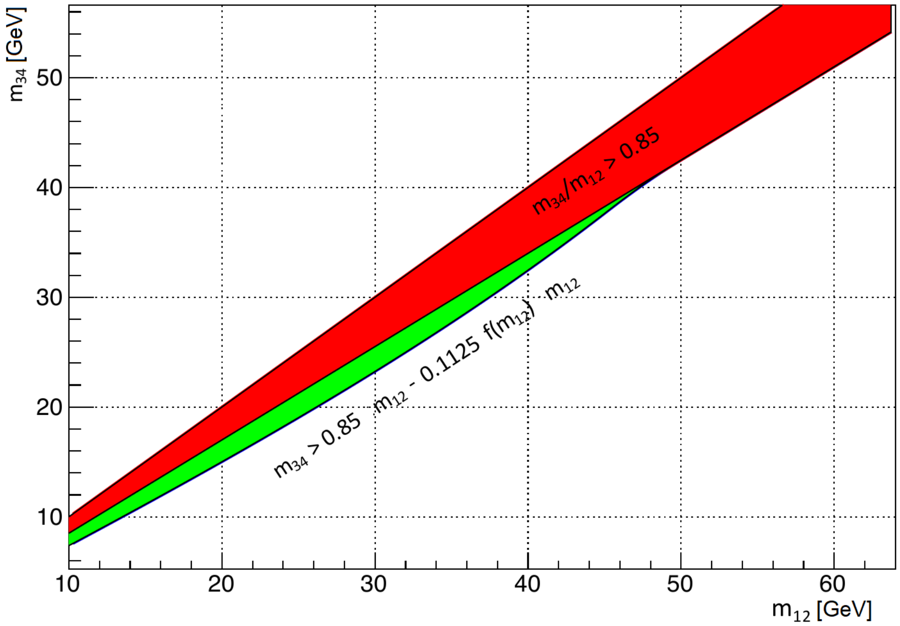


Figure 5.6: Original medium signal region defined by $m_{34}/m_{12} > 0.85$ (in red) and the addition to it from re-optimization (in green) [55]. The re-optimized region encompasses the original region and is defined by $m_{34} > 0.85m_{12} - 0.1125f(m_{12})m_{12}$.

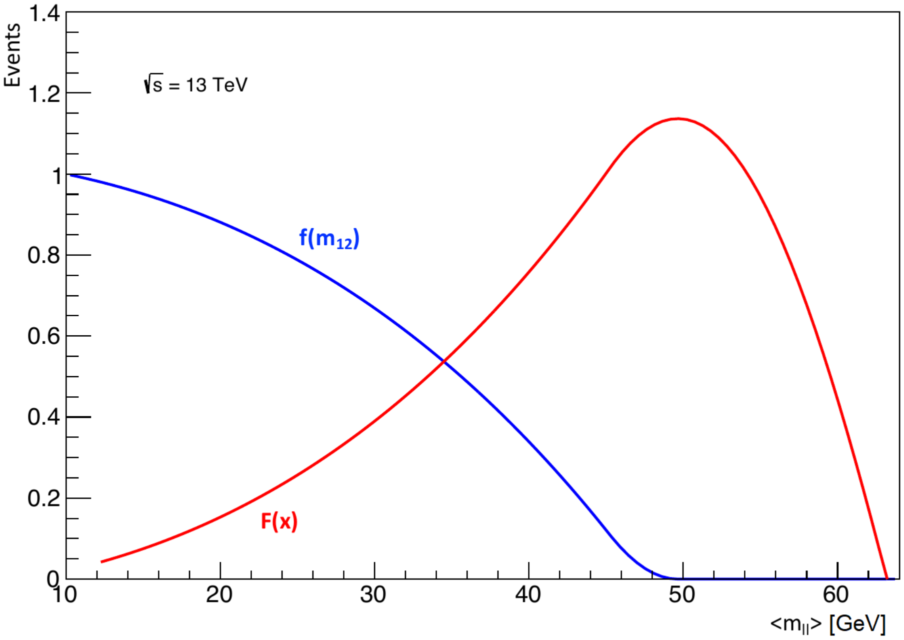


Figure 5.7: Modulating function $f(m_{12})$ of re-optimized signal region and exponential-Gaussian background model $F(\langle m_{\ell\ell}\rangle)$ [55].

Chapter 6

Data and Monte Carlo-based signal and background estimation

In essence, the $h \to Z_d Z_d \to 4\ell$ search compares distributions of events observed in data with distributions of background and signal events. With the exception of Z+jets background (Chapter 7), the signal and background events used in this search were generated with Monte Carlo (MC)-based simulation. This chapter describes the data and MC-based simulation used.

6.1 Data

The ATLAS detector records collisions of protons (or heavy ions) produced and accelerated by the Large Hadron Collider (LHC). Beginning in May 2015, the LHC began delivering proton beams of 6.5 TeV energy each, for a centre-of-mass energy of 13 TeV. In previous operation, the LHC delivered beams at lower centre-of-mass energy (7 TeV in 2010 and 2011, and 8 TeV in 2012 and 2013). In addition to delivering higher energy beams than before, the LHC also delivered higher instantaneous luminosity. The peak luminosity delivered to ATLAS in 2018 was 21.0×10^{33} $\text{cm}^{-2}\,\text{s}^{-1}$, compared to no more than 8×10^{33} $\text{cm}^{-2}\,\text{s}^{-1}$ in 2012. The increases in energy and luminosity have a direct impact on many analyses, including precision measurements of the Standard Model, measurements of the recently discovered SM Higgs boson, and searches for new physics.

After a commissioning period, ATLAS began taking data for physics analysis in August 2015. The search described in this thesis is conducted with data collected from

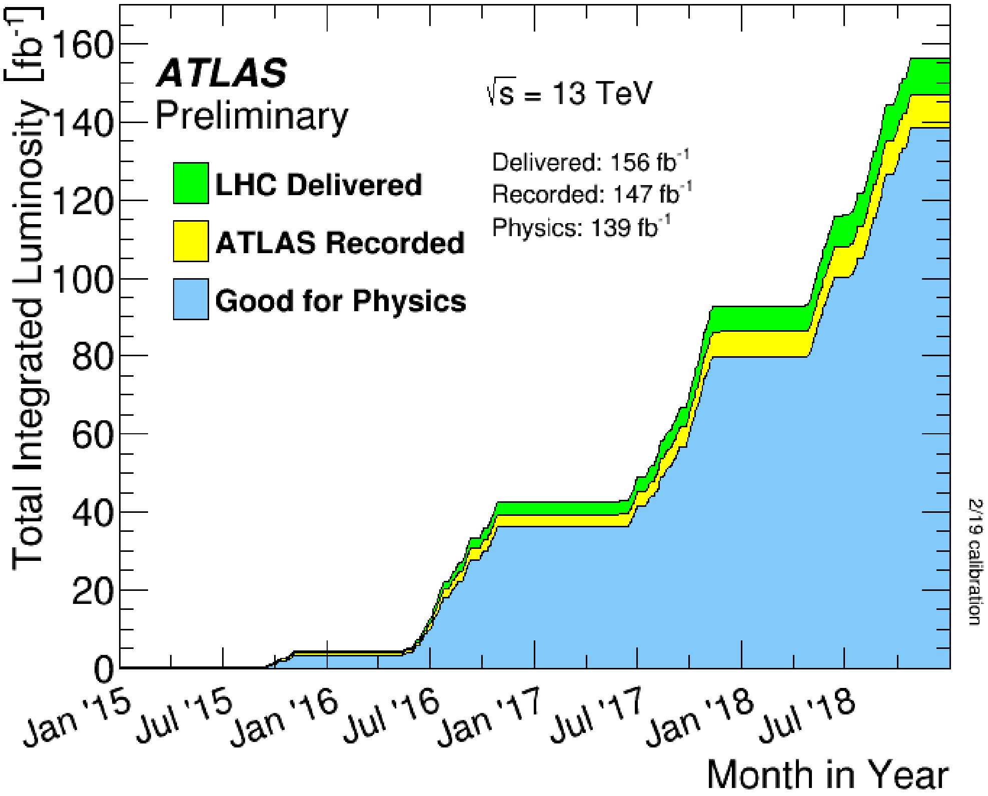


Figure 6.1: Total integrated luminosity delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be Good for Physics (blue) in 2015–2018 [21]. The delivered luminosity is the luminosity delivered from the start of stable beams until the LHC requests ATLAS to put the detector in a safe standby mode to allow a beam dump or beam studies. The recorded luminosity reflects the data acquisition inefficiency, as well as the inefficiency of the "warm start": when the stable beam flag is raised, the high-voltage supply for the tracking detectors ramp-up and the preamplifiers of the pixel system turn on.

August 2015 to October 2018. The LHC delivered 156 fb^{-1} of integrated luminosity to ATLAS in this period, of which 147 fb^{-1} were recorded and 139 fb^{-1} were usable for physics analysis [56].[1] The uncertainty in the combined 2015–2018 integrated luminosity is ±2.4 fb^{-1} (±1.7%) [56], obtained using the LUCID-2 detector [57] for the primary luminosity measurements.

---

[1] Corresponds to the integrated luminosity of the luminosity blocks passing the Good Runs Lists.

6.2 Monte Carlo-based signal and background estimation

Distributions of events observed in data are compared to distributions of simulated events containing signal and background processes (signal and background *samples*). Each process is simulated individually with a multi-step approach:

1. The four-momenta of the collision products and subsequent decay products are *generated*.
2. The passage of the decay products through the detector is *simulated*.
3. The response of the detector to the simulated energy deposits of the decay products and particles created from material interactions is *digitized*.
4. Finally, the electronic signals are *reconstructed* into analyzable physics objects.

First, the events containing the process are randomly generated with a Monte Carlo (MC)-based approach. This approach can be understood as numerical integration of the probability for the process to occur. Each point in the integrated momentum space is an event and the integral corresponds to the production cross-section.[2] The probability is given by the quantum mechanical matrix element-squared. The matrix elements are calculated with perturbation theory; due to computational constraints, usually only the first few orders in pertubation theory can be calculated. In most cases, calculations are leading order (LO) and provide sufficient accuracy.[3] In rarer cases, calculations can be done at next-to-leading order (NLO) or higher orders.[4] In addition to the probability of a process occuring, the probability distributions of the incoming quarks and gluons (*partons*) in the pp collision (*hard scatter*) also need to be taken into account. These are given by parton distribution functions (PDFs). PDFs define the probability to find a parton with a given momentum fraction x as a function of

---

[2] This space is multi-dimensional. Three momentum components per particle and energy-momentum conservation contribute $3n - 4$ dimensions, and additional dimensions and constraints come from the flavours and spins of the particles.

[3] Usually, leading order corresponds to the *tree-level* Feynman diagrams of the process and ignore higher-order diagrams (these may have loops and additional legs).

[4] For example, next-to-next-to-leading order (NNLO) in QCD is possible for $gg \to h$ with the POWHEG NNLOPS generator [58]. At leading order, the shapes of distributions are usually reliable, but not the normalization of distributions. If the cross-section can be computed at NLO, the ratio of the NLO to LO cross-sections (K-factor) can be applied to a LO distribution in order correct its normalization.

the squared energy scale Q^2.[5] The Monte Carlo-based evaluation of these probabilities produces events containing the four-momenta of the decay products. This step is known as *event generation*. Many different event generators exist, with some providing increased accuracy for particular processes. For example, the $h \to Z_d Z_d \to 4\ell$ signal samples used in this search were generated at leading order in perturbation theory with the MadGraph5_aMC@NLO event generator [59] and the NNPDF23LO parton distribution function set [60].

Following generation of the decay products, the passage of the decay products through the detector is simulated. Interaction with the detector's materials will result in direct energy deposition (e.g. bremsstrahlung of an electron) and/or indirect energy deposition through the creation of new particles (e.g. pair production of e^+e^- or ionization of an atom, followed by bremsstrahlung of the electrons). Simulation requires detailed modeling of electromagnetic and hadronic interactions with materials and of the detector's materials and geometry. Most of the samples used in this search are simulated with GEANT4 [61]. Like many simulators, GEANT4 is Monte Carlo-based; it divides the passage of each particle into steps, with the behaviour at each step determined by randomly sampling probability distributions that model the material interactions. For computational efficiency, a few background samples used in this search were simulated with parametrizations instead of with MC-based simulation.[6] These parametrizations are energy deposit distributions for a variety of particle types and kinematics and are pre-computed with single-particle GEANT4 simulation [62].

In both data and simulation, events are always assumed to be statistically independent from one another. For this reason, events are never counted more than once within an analysis. Samples are generated and simulated for each data-taking campaign independently in order to maintain statistical independence and also to take into account differences in beam and detector conditions. These differences include changes in the pileup distribution between campaigns (as shown in Figure 4.1) and improvements in detector modelling (for instance, the modeling of additional material due to the Insertable B-Layer of the Inner Detector).

The simulated energy deposits are then converted into electronic signals in the

---

[5] PDFs at various Q^2 and x are determined from measurement and are evolved by theory to other Q^2 and x (Section 6.2.3.2).

[6] Gluon-initiated ZZ^* production and higher-order electroweak production were simulated with parametrizations. All other processes were simulated with GEANT4. All background and signal processes considered in this search are not sensitive to the method of simulation.

digitization step. This requires modeling of the detector readout electronics. These models include details of how charge is collected and converted, as well as details about electronic noise, pulse shaping, and high occupancy behaviour (the behaviour in high pileup conditions).

Finally, the electronic signals are reconstructed into physics objects such as electrons and muons, as described in Sections 5.4.1 and 5.4.2 respectively.

6.2.1 Signal simulation

$h \to Z_d Z_d \to 4\ell$ signal samples were generated according to the Hidden Abelian Higgs Model (HAHM) [9, 10] with the MadGraph5_aMC@NLO event generator and the NNPDF23_lo_as_0130_qed leading order parton distribution function set. This model follows the theory described in Chapter 3 and is accurate to leading order.[7] The configurable parameters of this model are:

- Kinetic mixing parameter, ϵ: fixed to 10^{-4} for all m_{Z_d}
- Higgs mixing parameter, κ: fixed to 10^{-4} for all m_{Z_d}[8]
- Mass of dark Higgs, m_S: fixed to 200 GeV for all m_{Z_d}
- Mass of Z_d, m_{Z_d}: from 15 GeV to 60 GeV inclusive, in steps of 5 GeV

This model produces SM Higgs bosons only through the gluon-gluon fusion production mode (ggF). ggF is the dominant production mode at the LHC and has a cross-section that is roughly ten times larger than the cross-section of the other production modes combined [1, 54].[9]

After the hard scatter events $gg \to h \to Z_d Z_d \to 4\ell$ were generated, the parton shower, hadronization, and underlying event were modelled by PYTHIA8 with the A14 parameter set. Finally, all decay products were fully simulated in the ATLAS detector with GEANT4.

---

[7] The partial widths for $h \to Z_d Z_d$ and $Z_d \to \ell\ell$ evaluated with this model agree with the leading order analytical expressions given in Chapter 3. The Hidden Abelian Higgs Model does not include $\mathcal{O}(1)$ QCD corrections for *quark* final states.

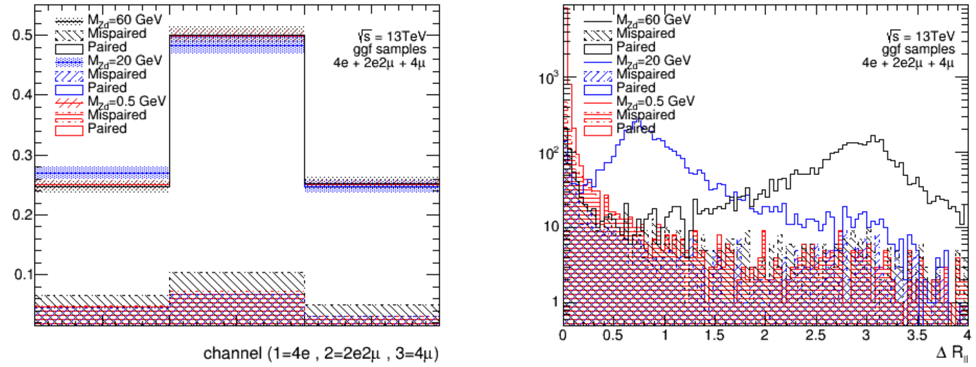
[8] In the limit $\epsilon \gg \kappa$, kinetic mixing dominates and only $h \to Z Z_d$ decays are produced. In the limit $\kappa \gg \epsilon$, Higgs mixing dominates and only $h \to Z_d Z_d$ decays are produced. When generating the samples for this search, MadGraph was configured to produce only $h \to Z_d Z_d$ decays.

[9] In order of decreasing cross-section, the other main production modes are: vector-boson fusion (VBF), quark-initiated Higgs-strahlung (associated production with a gauge boson, or $q\bar{q} \to VH$), and $t\bar{t}H$ (associated production with a pair of top quarks).

Shown in Figure 6.2 are distributions of the decay products at generator-level. The objects in the events forming these distributions were later simulated and reconstructed for analysis.[10] The reduction in the number of simulated signal events from imposing the signal region selections described in Chapter 5 is shown in Table 6.1.

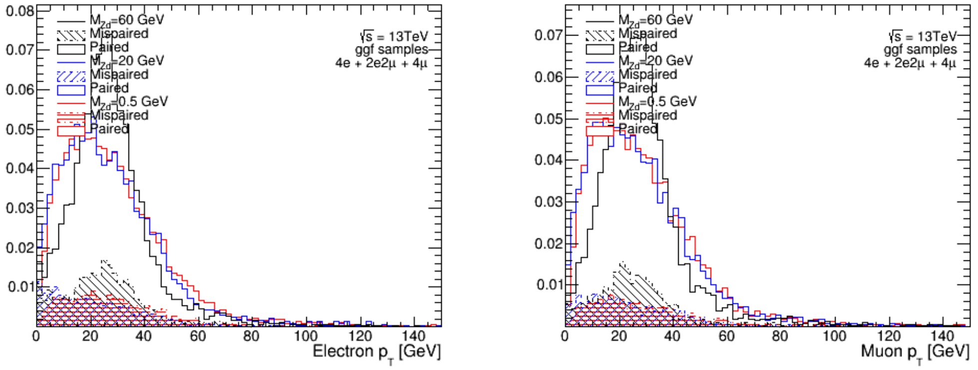
---

[10] Generator-level distributions are also referred to as *truth-level* distributions.



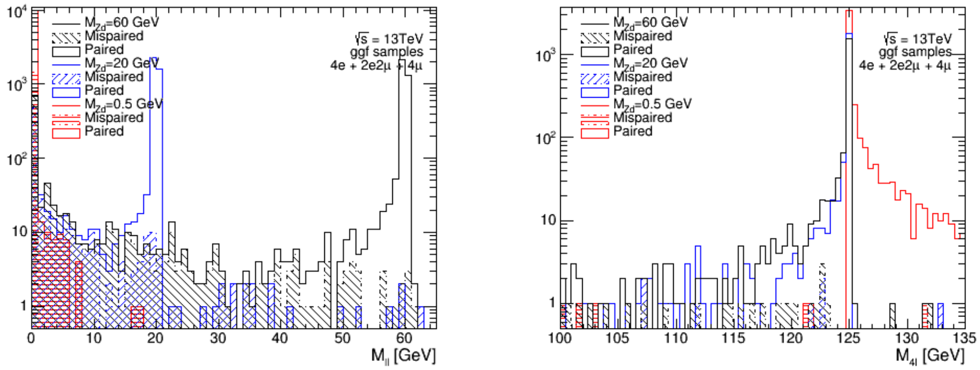
(a) Fraction of generated events in each flavour channel

(b) ΔR between leptons of each Z_d candidate



(c) Electron p_T normalized to unit integral

(d) Muon p_T normalized to unit integral



(e) Dilepton mass (mass of Z_d candidate)

(f) Four-lepton mass (mass of Higgs candidate)

Figure 6.2: Generator-level distributions of $h \to Z_d Z_d \to 4\ell$ signal [63]. In these events, QED final state radiation is recovered by adding the four-momenta of photons within a $\Delta R = 0.1$ cone around each lepton to the lepton's four-momentum [64].

Requirement	Number of events							
Event cleaning and at least one primary vertex	837.1							
At least one lowest-unprescaled trigger fired	708.9							
	$4e$		$2e2\mu$		4μ		All	
	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}
Same-flavour opposite sign quadruplet	99.79	100 %	208.8	100 %	106.4	100 %	415.0	100 %
Non-overlapping quadruplet leptons	99.75	100 %	208.7	100 %	106.5	100 %	414.9	100 %
$p_T(\ell_1) > 20$, $p_T(\ell_2) > 15$, $p_T(\ell_3) > 10$ GeV	95.91	96.1 %	201.0	96.2 %	102.7	96.5 %	399.6	96.3 %
Leptons in quadruplet fired at least one trigger	95.43	95.6 %	201.0	96.3 %	102.7	96.5 %	399.1	96.2 %
$\Delta R > 0.1\,(0.2)$ for same (opposite) flavour leptons	94.01	94.2 %	198.7	95.2 %	102.8	96.6 %	395.5	95.3 %
Not more than one calo-tagged or stand-alone muon	94.03	94.2 %	198.6	95.1 %	102.6	96.4 %	395.3	95.2 %
Isolation (Electrons: FCLoose) (Muons: FCLoose_FixedRad)	58.41	58.5 %	143.4	68.7 %	84.91	79.8 %	286.7	69.1 %
Electron ID (all quadruplet leptons Loose or better)	40.54	40.6 %	119.0	57.0 %	84.91	79.8 %	244.4	58.9 %
Impact parameter (Electrons: $z_0 \sin\theta < 0.5$ mm) (Muons: $z_0 \sin\theta < 0.5$ mm, $d_0 < 1$ mm, d_0 significance ≤ 3)	40.08	40.2 %	115.9	55.5 %	81.72	76.8 %	237.7	57.3 %
Quarkonia veto	40.05	40.1 %	115.9	55.5 %	81.56	76.6 %	237.5	57.2 %
Low mass veto (require $m_{\ell\ell} > 5$ GeV)	40.03	40.1 %	115.9	55.5 %	81.54	76.6 %	237.5	57.2 %
Higgs window ($115 < m_{4\ell} < 130$ GeV)	37.70	37.8 %	110.6	53.0 %	78.81	74.0 %	227.1	54.7 %
Z veto (require $m_{12,34} < 64$ and $m_{32,14} < 75$ GeV)	25.29	25.3 %	110.6	53.0 %	51.90	48.8 %	187.8	45.3 %
Loose signal region ($m_{12,34} > 10$ GeV)	25.29	25.3 %	110.6	53.0 %	51.90	48.8 %	187.8	45.3 %
Medium signal region (re-optimized for wider width) ($m_{34} > 0.85 m_{12} - 0.1125 f(m_{12}) m_{12}$)	25.24	25.3 %	110.3	52.8 %	51.83	48.7 %	187.4	45.1 %

Table 6.1: Signal region cutflow for $m_{Z_d} = 25$ GeV signal in 2015–2018 simulation. The yields from simulation of each data-taking campaign are normalized to the integrated luminosity of the campaign (36.2 fb^{-1} for 2015–2016, 44.3 fb^{-1} for 2017, and 59.9 fb^{-1} for 2018) and then combined (139 fb^{-1} total), with signal estimates corresponding to a signal strength of 1 (equivalent to the SM $gg \to h \to ZZ^* \to 4\ell$ cross-section). Per-channel, the quadruplet cutflow can increase because events can move from one channel to another during the quadruplet selection if a higher-ranked quadruplet fails a cut that a lower-ranked one can pass (however, the cutflow summed across all channels can only decrease).

6.2.2 Background simulation

All backgrounds in the $h \to Z_d Z_d \to 4\ell$ search are estimated from simulation.[11] The two main backgrounds in this search are $h \to ZZ^* \to 4\ell$ and $q\bar{q} \to ZZ^* \to 4\ell$ and account for approximately 72% and 24% of the background in the signal region, respectively. In order to make distributions of simulated backgrounds comparable to distributions of observed data, the former must be properly normalized by taking into account the integrated luminosity of the data, the number of events generated, and the cross-section of the simulated process:

$$\frac{\left(\int \mathcal{L}\ \mathrm{d}t\right)_{\text{data}}}{\left(\int \mathcal{L}\ \mathrm{d}t\right)_{\text{MC}}} = \left(\int \mathcal{L}\ \mathrm{d}t\right)_{\text{data}} \frac{\sigma_{\text{gen,MC}}}{N_{\text{gen,MC}}} = \left(\int \mathcal{L}\ \mathrm{d}t\right)_{\text{data}} \frac{\sigma_{\text{MC}}\epsilon_{\text{filter}}}{N_{\text{gen,MC}}} \tag{6.1}$$

After event generation, a filter may be applied so that only events in the relevant kinematic region are simulated. The fraction of events passing the filter is the generator filter efficiency, ϵ_{filter}. If the cross-section of the process is known at a higher order, the normalization of the lower order distribution can be corrected to higher order with the ratio of the two cross-sections (K-factor). This improves the precision of the distribution without having to generate events at higher order.

6.2.2.1 $h \to ZZ^* \to 4\ell$

Decays of the SM Higgs to two Z bosons (one on-shell, the other off-shell) which finally decay to four leptons (electrons and/or muons) form the dominant background in the signal region. Table 6.2 lists the $h \to ZZ^* \to 4\ell$ samples used, in order of largest Higgs production cross-section to smallest.

Events with τ leptons can contribute to the signal region because τ leptons can decay to $e\nu\nu$ or $\mu\nu\nu$. The additional yield to the validation and signal regions was assessed with a sample containing τ decays and was found to be approximately 1% of the total background prediction. This is less than the uncertainty (stat. $\oplus$ syst.) on the total background prediction ($\pm 3\%$). Since the sample without τ decays has a greater number of generated events than the sample with τ decays, the sample without τ decays is used instead.

---

[11] As discussed in Chapter 7, fake backgrounds from Z+jets processes were cross-checked with a data-driven method.

The reduction in the number of simulated $h \rightarrow ZZ^* \rightarrow 4\ell$ events from imposing the signal region selections is shown in Table 6.6.

6.2.2.2 $ZZ^* \rightarrow 4\ell$

Non-resonant (*continuum*) production of ZZ^* followed by decay to four leptons forms the second-largest background in the signal region. Table 6.3 lists the samples used to simulate quark-initiated and gluon-initiated ZZ^* production.

The reduction in the number of simulated $ZZ^* \rightarrow 4\ell$ events from imposing the signal region selections is shown in Table 6.7.

6.2.2.3 Tri-boson, higher-order electroweak, and $t\bar{t} + Z$

Vector boson fusion and vector boson scattering processes involving three bosons (ZZZ, WZZ, WWZ) that decay to four or more leptons with neutrinos, and higher-order electroweak processes (couplings of α^6) contribute negligibly ($< 0.1\%$ of total background) to the signal region. Production of a top quark pair in association with a Z boson with the Z decaying to two leptons ($t\bar{t} + (Z \rightarrow \ell\ell)$) also contributes negligibly ($< 1\%$). The samples used to simulate these processes are listed in Table 6.4.

6.2.2.4 Reducible backgrounds

Z+jets, $t\bar{t}$, and WZ processes produce less than four leptons but can contribute to the signal region if jets are misidentified or misreconstructed as additional leptons. In the case of Z+jets, the Z boson decays to two electrons (or muons) and at most four jets are simulated. If two leptons are mistakenly identified or reconstructed from the jets, then the event can enter the signal region. The negligible estimate from Monte Carlo-based simulation is cross-checked with a data-driven estimate (described in Chapter 7). Dileptonic decays of top quark pairs ($t\bar{t} \rightarrow WbWb \rightarrow \ell\ell\nu\nu bb$) produce two leptons and two bottom quarks and can enter the signal region in a similar manner. If the bottom quark(s) hadronize (i.e. into two b-tagged jets), they can also be misidentified or misreconstructed as leptons. WZ decaying to three leptons, a neutrino, and jets can also contribute to the signal region. Dileptonic $t\bar{t}$ decay and WZ decay to three leptons contribute $\sim 3\%$ and $< 1\%$ to the total background in the signal region, respectively. The samples used to simulate reducible processes are listed in Table 6.5.

Higgs production process	Generator	Accuracy	PDF set	Parton showering, hadronization, MPI (Tune & PDF)	K-factor, event re-weighting
Gluon-gluon fusion (ggF)	POWHEG [65–67] NNLOPS [58]	NNLO	PDF4LHC15 NLO [68]	PYTHIA8 w/ AZNLO [69] & CTEQ6L1 [70]	1.1 (N3LO [71])
Vector boson fusion (VBF)	POWHEG	NLO	CT10 NLO [72]	PYTHIA8 w/ AZNLO & CTEQ6L1	0.9786 (NLO EW corrections [73, 74]), 1 (NNLO [75])
Quark-initiated Higgs-strahlung ($q\bar{q} \rightarrow W(Z), VH$)	PYTHIA8 [76]	LO	NNPDF23 LO [77]	PYTHIA8 w/ A14 [78] & NNPDF23 LO	None
Higgs production with two top quarks ($t\bar{t}H$)	POWHEG	NLO	NNPDF30 NLO [79]	PYTHIA8 w/ A14 & NNPDF23 LO	None
Higgs production with two bottom quarks (bbH)	MadGraph5_aMC@NLO [59, 80]	NLO	CT10 NLO	PYTHIA8 w/ A14 & NNPDF23 LO	None
Loop-induced Higgs and Z production ($gg \rightarrow ZH$)	POWHEG	NLO	PDF4LHC15 NLO	PYTHIA8 w/ AZNLO & CTEQ6L1	None

Table 6.2: $h \rightarrow ZZ^* \rightarrow 4\ell$ simulation samples. All-hadronic ($t(\bar{t}) \rightarrow W(\rightarrow q\bar{q})b$), dileptonic ($t(\bar{t}) \rightarrow W(\rightarrow \ell\nu)b$), and semi-leptonic (one top decaying hadronically and the other top decaying leptonically) decays of the top quark pair were simulated. MPI stands for multiple parton interaction.

ZZ^* production process	Generator	Accuracy	PDF set	Parton showering, hadronization, MPI (Tune & PDF)	K-factor, event re-weighting
Quark-initiated ($q\bar{q} \rightarrow ZZ^*$)	Sherpa [81–83] v2.2.2	NLO	NNPDF30 NNLO [79]	Sherpa v2.2.2 w/ NNPDF30 NNLO PDF & tune	NLO EW re-weighting [84, 85]
Gluon-initiated ($gg \rightarrow ZZ^*$)	Sherpa v2.2.2	LO	NNPDF30 NNLO	Sherpa v2.2.2 w/ NNPDF30 NNLO PDF & tune	1.7 (NLO [86])

Table 6.3: Non-resonant (continuum) $ZZ^* \rightarrow 4\ell$ simulation samples. MPI stands for multiple parton interaction.

Process	Generator	Accuracy	PDF set	Parton showering, hadronization, MPI (Tune & PDF)	K-factor, event re-weighting
$ZZZ \rightarrow 4\ell 2\nu / 6\ell 0\nu$	Sherpa v2.2.2	NLO	NNPDF30 NNLO	Sherpa v2.2.2 w/ NNPDF30 NNLO PDF & tune	None
$WZZ \rightarrow 5\ell 1\nu$	Sherpa v2.2.2	NLO	NNPDF30 NNLO	Sherpa v2.2.2 w/ NNPDF30 NNLO PDF & tune	None
$WWZ \rightarrow 4\ell 2\nu$	Sherpa v2.2.2	NLO	NNPDF30 NNLO	Sherpa v2.2.2 w/ NNPDF30 NNLO PDF & tune	None
Higher-order electroweak $\rightarrow 4\ell 2j$	Sherpa v2.2.2	LO	NNPDF30 NNLO	Sherpa v2.2.2 w/ NNPDF30 NNLO PDF & tune	None
$t\bar{t} + (Z \rightarrow \ell\ell)$	Sherpa v2.2.0	LO	NNPDF30 NNLO	Sherpa v2.2.0 w/ NNPDF30 NNLO PDF & tune	1.09 (NNLO [87])

Table 6.4: Tri-boson (vector boson fusion/scattering), higher-order electroweak, and $t\bar{t} + Z$ simulation samples. MPI stands for multiple parton interaction.

Process	Generator	Accuracy	PDF set	Parton showering, hadronization, MPI (Tune & PDF)	K-factor, event re-weighting
$Z(\rightarrow \ell\ell)$+jets	Sherpa v2.2.1	$0, 1, 2j$@NLO, $3, 4j$@LO	NNPDF30 NNLO	Sherpa v2.2.1 w/ NNPDF30 NNLO PDF & tune	0.9751 (NNLO [88])
$t\bar{t}$ dileptonic	POWHEG	NLO	NNPDF30 NLO	PYTHIA8 w/ A14 & NNPDF23 LO	1.1398 (NNLO [89])
$WZ \rightarrow \ell\nu\ell\ell$	POWHEG	NLO	CT10 NLO	PYTHIA8 w/ AZNLO & CTEQ6L1	1.14 (NNLO [90])

Table 6.5: Reducible background simulation samples. MPI stands for multiple parton interaction.

Requirement	Number of events							
Event cleaning and at least one primary vertex	624.6							
At least one lowest-unprescaled trigger fired	589.1							
	$4e$		$2e2\mu$		4μ		All	
	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}
Same-flavour opposite sign quadruplet	134.3	100 %	261.5	100 %	116.7	100 %	512.5	100 %
Non-overlapping quadruplet leptons	134.2	99.92 %	261.4	99.95 %	116.7	100.04 %	512.3	99.96 %
$p_T(\ell_1) > 20$, $p_T(\ell_2) > 15$, $p_T(\ell_3) > 10$ GeV	126.0	93.79 %	246.1	94.11 %	113.6	97.39 %	485.7	94.77 %
Leptons in quadruplet fired at least one trigger	124.2	92.49 %	247.0	94.44 %	113.8	97.51 %	485.0	94.63 %
$\Delta R > 0.1\,(0.2)$ for same (opposite) flavour leptons	120.4	89.61 %	237.6	90.84 %	113.9	97.60 %	471.8	92.05 %
Not more than one calo-tagged or stand-alone muon	120.4	89.61 %	237.6	90.83 %	113.7	97.50 %	471.7	92.03 %
Isolation (Electrons: FCLoose) (Muons: FCLoose_FixedRad)	49.89	37.14 %	117.6	44.96 %	89.03	76.32 %	256.5	50.05 %
Electron ID (all quadruplet leptons Loose or better)	31.61	23.53 %	92.32	35.30 %	89.03	76.32 %	213.0	41.55 %
Impact parameter (Electrons: $z_0 \sin\theta < 0.5$ mm) (Muons: $z_0 \sin\theta < 0.5$ mm, $d_0 < 1$ mm, d_0 significance ≤ 3)	31.19	23.22 %	89.41	34.19 %	85.18	73.02 %	205.8	40.15 %
Quarkonia veto	29.26	21.78 %	86.58	33.10 %	79.51	68.15 %	195.4	38.12 %
Low mass veto (require $m_{\ell\ell} > 5$ GeV)	28.86	21.49 %	85.92	32.85 %	78.28	67.09 %	193.1	37.67 %
Higgs window ($115 < m_{4\ell} < 130$ GeV)	26.63	19.83 %	80.48	30.77 %	74.40	63.78 %	181.5	35.42 %
Z veto (require $m_{12,34} < 64$ and $m_{32,14} < 75$ GeV)	4.380	3.26 %	10.73	4.10 %	10.65	9.13 %	25.77	5.03 %
Loose signal region ($m_{12,34} > 10$ GeV)	4.348	3.24 %	10.54	4.03 %	10.57	9.06 %	25.47	4.97 %
Medium signal region (re-optimized for wider width) ($m_{34} > 0.85m_{12} - 0.1125f(m_{12})m_{12}$)	2.230	1.66 %	3.944	1.51 %	4.943	4.24 %	11.12	2.17 %

Table 6.6: Signal region cutflow for $h \to ZZ^* \to 4\ell$ background in 2015–2018 simulation with all production modes combined. The yields from simulation of each data-taking campaign are normalized to the integrated luminosity of the campaign (36.2 fb^{-1} for 2015–2016, 44.3 fb^{-1} for 2017, and 59.9 fb^{-1} for 2018) and then combined (139 fb^{-1} total). Per-channel, the quadruplet cutflow can increase because events can move from one channel to another during the quadruplet selection if a higher-ranked quadruplet fails a cut that a lower-ranked one can pass (however, the cutflow summed across all channels can only decrease).

Requirement	Number of events							
Event cleaning and at least one primary vertex	$9.617 \cdot 10^4$							
At least one lowest-unprescaled trigger fired	$6.960 \cdot 10^4$							
	$4e$		$2e2\mu$		4μ		All	
	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}
Same-flavour opposite sign quadruplet	4730	100 %	8799	100 %	3139	100 %	$1.667 \cdot 10^4$	100 %
Non-overlapping quadruplet leptons	4726	99.91 %	8787	99.87 %	3141	100.05 %	$1.665 \cdot 10^4$	99.92 %
$p_T(\ell_1) > 20$, $p_T(\ell_2) > 15$, $p_T(\ell_3) > 10$ GeV	4257	90.00 %	7646	86.89 %	2682	85.42 %	$1.458 \cdot 10^4$	87.50 %
Leptons in quadruplet fired at least one trigger	4227	89.38 %	7646	86.90 %	2684	85.50 %	$1.456 \cdot 10^4$	87.34 %
$\Delta R > 0.1\,(0.2)$ for same (opposite) flavour leptons	4016	84.91 %	7150	81.27 %	2687	85.60 %	$1.385 \cdot 10^4$	83.12 %
Not more than one calo-tagged or stand-alone muon	4016	84.91 %	7149	81.25 %	2683	85.46 %	$1.385 \cdot 10^4$	83.08 %
Isolation (Electrons: FCLoose) (Muons: FCLoose_FixedRad)	1500	31.71 %	3558	40.44 %	2237	71.26 %	7295	43.77 %
Electron ID (all quadruplet leptons Loose or better)	926.7	19.59 %	2779	31.58 %	2237	71.26 %	5943	35.65 %
Impact parameter (Electrons: $z_0 \sin\theta < 0.5$ mm) (Muons: $z_0 \sin\theta < 0.5$ mm, $d_0 < 1$ mm, d_0 significance ≤ 3)	909.9	19.24 %	2685	30.52 %	2126	67.74 %	5721	34.32 %
Quarkonia veto	860.4	18.19 %	2569	29.19 %	1948	62.05 %	5377	32.26 %
Low mass veto (require $m_{\ell\ell} > 5$ GeV)	832.0	17.59 %	2455	27.90 %	1822	58.02 %	5108	30.65 %
Higgs window ($115 < m_{4\ell} < 130$ GeV)	15.72	0.33 %	56.35	0.64 %	50.79	1.62 %	122.9	0.74 %
Z veto (require $m_{12,34} < 64$ and $m_{32,14} < 75$ GeV)	1.704	0.04 %	6.288	0.07 %	4.551	0.14 %	12.54	0.08 %
Loose signal region ($m_{12,34} > 10$ GeV)	1.559	0.03 %	4.293	0.05 %	4.113	0.13 %	9.966	0.06 %
Medium signal region (re-optimized for wider width) ($m_{34} > 0.85 m_{12} - 0.1125 f(m_{12}) m_{12}$)	0.6389	0.01 %	1.229	0.01 %	1.512	0.05 %	3.380	0.02 %

Table 6.7: Signal region cutflow for $ZZ^* \to 4\ell$ background in 2015–2018 simulation with all production modes combined. The yields from simulation of each data-taking campaign are normalized to the integrated luminosity of the campaign (36.2 fb^{-1} for 2015–2016, 44.3 fb^{-1} for 2017, and 59.9 fb^{-1} for 2018) and then combined (139 fb^{-1} total). Per-channel, the quadruplet cutflow can increase because events can move from one channel to another during the quadruplet selection if a higher-ranked quadruplet fails a cut that a lower-ranked one can pass (however, the cutflow summed across all channels can only decrease).

6.2.2.5 Pileup treatment

The generation of the simulated samples includes the effect of multiple proton-proton interactions in the same and nearby bunch crossings (*pileup*). This was simulated at leading order with Pythia8 using the MSTW 2008 parton distribution function set [91] and the A2 parameter set [92].

6.2.3 Estimate of statistical and systematic uncertainties

Three kinds of uncertainties need to be assessed:

- Statistical: from random fluctuations in the measurement. In this search, statistical uncertainties are fluctuations in the yields of the background predictions and observed data in the signal region. This is a Poisson-distributed process in the asymptotic limit, therefore the statistical uncertainty is simply $\sqrt{n}$.[12]
- Experimental systematics: from the nature of the detector, such as uncertainties associated with electron and muon identification and uncertainties in momentum resolution and energy scale. These uncertainties apply to simulation-based background and signal estimates. The estimation of these uncertainities depends on the kinematic variable and type of object.
- Theoretical systematics: from uncertainties on parameters and modelling used for signal simulation, e.g. quantum chromodynamics (QCD) scale, parton distribution functions, and parton showering.

The effect of these uncertainties on simulated signal and background yields in the signal region are tabulated in Tables 6.8 and 6.9, respectively. For signal (background), systematic uncertainties combine in quadrature to $\lesssim 12\%$ ($\lesssim 14\%$), with the largest effect in the $4e$ channel (primarily due to electron identification efficiency uncertainty).

---

[12] Since the number of bunch crossings and number of particles per bunch is finite, the number of events passing some criteria is technically a binomial process. In the limit $n \to \infty$, the process is Poisson-distributed. The estimate of the standard deviation is therefore $\sqrt{n}$. However, this is misleading in the case of small n (e.g. 0 ± 0 and 1 ± 1). This is the case for the signal region. For small n, an alternative is to define confidence intervals (a prescription is given in [1]). For large n, the uncertainty derived from defining a 68.3% central confidence interval is equivalent to $\sqrt{n}$.

6.2.3.1 Experimental systematic uncertainties

Luminosity As described in Section 6.1, the uncertainty on the integrated luminosity is ±1.7%.

Pileup re-weighting As described in Section 6.2.2.5, simulated pileup is overlaid on the simulated process. The amount of pileup in each simulated event is determined by drawing from a reference distribution of the mean number of interactions per bunch crossing (Figure 4.1). This distribution may not match the distribution that is measured. This is corrected by re-weighting simulated events with *scale factors* such that the agreement is improved. In order to estimate the uncertainty associated with pileup re-weighting, events are also re-weighted with $\pm 1\sigma$ variations of the nominal scale factor instead. The difference in the event yield between re-weighting with the nominal and the $+(-)1\sigma$ variation is taken as the $+(-)$ uncertainty on the yield.

Electrons The resolution, energy scale, and identification/isolation/reconstruction efficiencies of electrons in simulation are corrected in order to improve agreement with data. Electron resolution and energy scale corrections are applied to each electron (Section 5.4.1), and identification/isolation/reconstruction efficiency corrections are applied through event re-weighting. As with the uncertainty associated with pileup re-weighting, uncertainties associated with electron corrections are evaluated by applying a $+(-)1\sigma$ variation of each correction independently and then re-determining the event yield.

EG_RESOLUTION_ALL: uncertainty associated with smearing of electron energies in simulation in order to improve resolution agreement between data and simulation.

EG_SCALE_ALL(AF2): uncertainty associated with calibration of electron energy scale in simulation. A special set of calibrations and uncertainties are applied for samples simulated with the Atlfast-II (AF2) parametrization instead of with GEANT4.

EL_EFF_ID(Iso,Reco)_TOTAL_1NPCOR_PLUS_UNCOR: uncertainties associated with re-weighting of simulated events such that identification, isolation, and reconstruction efficiencies in simulation agree with those in data. Identification efficiency uncertainty is approximately ±10% on the signal region yield and is the dominant systematic uncertainty.

Muons The resolution, momentum scale, and isolation/reconstruction/track-to-vertex association efficiencies of muons in simulation are also corrected. Muon resolution and momentum scale corrections are applied to each muon, and isolation/ reconstruction/track-to-vertex association efficiency corrections are applied through event re-weighting. Uncertainties associated with these corrections are evaluated in the same way as they are evaluated for electrons.

MUON_ID: uncertainty associated with charge-agnostic smearing of simulated muon p_T in the Inner Detector (ID) in order to improve muon ID p_T resolution agreement between data and simulation.

MUON_MS: uncertainty associated with charge-agnostic smearing of simulated muon p_T in the Muon Spectrometer (MS) in order to improve muon MS p_T resolution agreement between data and simulation.

MUON_SCALE: uncertainty associated with calibration of the muon momentum scale in simulation.

MUON_SAGITTA_RHO: uncertainty associated with correction of muon momenta for charge-dependent sagitta biases in simulation. Geometric deformations of the detectors affect the sagitta measurement and consquently the momentum.

MUON_SAGITTA_RESBIAS: uncertainty associated with correction of muon momenta for residual sagitta bias in simulation.

MUON_EFF_ISO(RECO)(TTVA)_STAT(SYS)_LOWPT: similar to the efficiency uncertainties for electrons but broken up into statistical and systematic errors on the weights.

6.2.3.2 Theoretical systematic uncertainties

Generation of events requires knowledge of the momentum fraction x carried by each parton and a choice for the QCD renormalization and factorization scales. Parton distribution functions (PDFs) are determined experimentally and QCD normalization and factorization scales are chosen as appropriate. The associated systematic uncertainties are assessed by varying the PDFs and the QCD scales.

Parton distribution functions are extracted from experimental data with modelling assumptions and differential equations that govern how the PDFs evolve with the squared energy scale Q^2. These are the *DGLAP equations* (named after authors Dokshizer, Gribov, Lipatov, Altarelli, and Parisi). Modelling assumptions include the choice of parametrization of the input distributions to be fitted and the choice of the strong coupling constant $\alpha_S(Q^2)$. The coupling constant may be fixed to a certain value or left floating. The DGLAP equations are expressed in terms of kernels that are powers of the strong coupling constant and are only known analytically to the first few orders. Choices of input parametrization and truncation of the DGLAP equations are only two examples of uncertainties in PDF construction.

Different PDFs use different modelling and theoretical assumptions. Therefore, the uncertainties related to these assumptions can be assessed by varying the choice of PDF. Each PDF set has multiple PDFs, with the *nominal* PDF corresponding to the "best fit" parameters and each of the other PDFs corresponding to a systematic variation on a fitted parameter. Additionally, the nominal PDFs from multiple PDF sets can be considered, in order to take into account intrinsic differences in PDF set construction between sets.

Signal uncertainties For signal, the effects of theoretical systematic uncertainties are assessed on the signal acceptance and do not enter the statistical interpretation of the results. The signal acceptance is the number of events in the signal region divided by the number of events before selection (i.e. the number of generated events).[13] Consequently, variations on PDFs and QCD scales affect both the numerator and the denominator. The following uncertainties are assessed:

- **Intra-PDF set uncertainties**: The alternative PDFs of the nominal PDF set NNPDF23_LO_AS_0130_QED (LHAPDF ID 247000) were used to generate alternative event weights for each event. Each alternative PDF (one hundred in total) is used to generate an alternative event weight (one hundred alternative weights per event). The alternative weights are used in place of the nominal weight in a generator-level implementation of the signal region selection. The signal region yield is calculated with each replacement of the weight, resulting in one hundred alternative yields y_j each corresponding to a PDF within the set. The *asymmetric* systematic uncertainty on the acceptance is assessed as the 15.87% and 84.14% quantiles of the distribution of normalized yields $y_j^N = \frac{s_0}{s_j} y_j$, where s_0 and s_j are the sums of event weights before selection for the nominal and alternative PDF, respectively.

---

[13] More precisely: the sum of *event weights*. Events in Monte Carlo-based simulation are weighted, unlike events observed in data.

- **Inter-PDF set uncertainties**: The nominal PDFs of the alternative PDF sets CT14lo (LHAPDF ID 13200) and MMHT2014lo68cl (LHAPDF ID 25000) are used to generate alternative event weights for each event. The alternative weights (two per event) are processed in the same way as for intra-PDF set uncertainties. The asymmetric systematic uncertainty on the acceptance is assessed as the largest deviations of the normalized yields $y_j^N = \max(|y_j^N - y_0|)$ above and below the nominal yield (envelope).

- **QCD scale uncertainties**: Six pairwise variations of the renormalization μ_r and factorization μ_f scales are considered: $(\mu_r, \mu_f) = (0.5, 0.5)$, $(0.5, 1.0)$, $(1.0, 0.5)$, $(1.0, 2.0)$, $(2.0, 1.0)$, $(2.0, 2.0)$. Alternative signal region yields are calculated with the approach described for PDF uncertainties. The asymmetric systematic uncertainty on the acceptance is assessed as the envelope, as described for inter-PDF set uncertainties.

The PDF uncertainties *on the signal region acceptance* combine in quadrature to $\lesssim 3\%$ and the QCD uncertainties are $\lesssim 1\%$. The PDF uncertainties are relatively small because the generation is at leading order only, and the QCD uncertainties are negligible because of the absence of hadronization in the final state.

Background uncertainties For background, the effects of theoretical systematic uncertainties are assessed in the same way as experimental (detector) systematic uncertainties. Unlike the signal uncertainties, they are assessed on the signal region yields and enter the statistical interpretation of the results. For the two dominant backgrounds $gg \to h \to ZZ^* \to 4\ell$ and $ZZ^* \to 4\ell$, the PDF and QCD scale uncertainties are taken from prior calculations following the same procedure described above ([93–96] and [97] for each background, respectively).

Systematic	$m_{Z_d} = 20$ GeV			$m_{Z_d} = 35$ GeV			$m_{Z_d} = 50$ GeV		
	$4e$	$2e2\mu$	4μ	$4e$	$2e2\mu$	4μ	$4e$	$2e2\mu$	4μ
EG_RESOLUTION_ALL	$^{-0.44}_{+0.29}$	$^{-0.10}_{+0.05}$	$^{+0.00}_{-0.06}$	$^{-1.10}_{+0.64}$	$^{-0.20}_{+0.09}$	$^{+0.08}_{-0.00}$	$^{-0.78}_{+0.38}$	$^{-0.27}_{+0.11}$	$^{+0.00}_{-0.01}$
EG_SCALE_AF2	–	–	–	–	–	–	–	–	–
EG_SCALE_ALL	$^{-0.59}_{+0.74}$	$^{+0.05}_{-0.15}$	$^{+0.00}_{-0.04}$	$^{-0.99}_{+0.39}$	$^{+0.01}_{-0.06}$	$^{+0.07}_{-0.00}$	$^{-0.44}_{+0.31}$	$^{+0.00}_{-0.08}$	$^{+0.00}_{-0.05}$
EIGENTUNESIGNAL	±0.50	±0.50	±0.50	±0.50	±0.50	±0.50	±0.50	±0.50	±0.50
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	$^{+10.96}_{-10.04}$	$^{+5.40}_{-5.26}$	–	$^{+12.09}_{-10.96}$	$^{+5.69}_{-5.54}$	–	$^{+11.93}_{-10.86}$	$^{+5.78}_{-5.64}$	–
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	±0.84	±0.48	–	±0.91	±0.50	–	±0.80	$^{+0.43}_{-0.42}$	–
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	$^{+2.63}_{-2.59}$	±1.40	–	$^{+2.95}_{-2.90}$	$^{+1.50}_{-1.49}$	–	$^{+2.85}_{-2.80}$	±1.45	–
INTERPDFSIGNAL	$^{+1.70}_{-0.40}$	$^{+1.70}_{-0.40}$	$^{+1.70}_{-0.40}$	$^{+1.70}_{-0.40}$	$^{+1.70}_{-0.40}$	$^{+1.70}_{-0.40}$	$^{+1.70}_{-0.40}$	$^{+1.70}_{-0.40}$	$^{+1.70}_{-0.40}$
INTRAPDFSIGNAL	$^{+1.50}_{-1.80}$	$^{+1.50}_{-1.80}$	$^{+1.50}_{-1.80}$	$^{+1.50}_{-1.80}$	$^{+1.50}_{-1.80}$	$^{+1.50}_{-1.80}$	$^{+1.50}_{-1.80}$	$^{+1.50}_{-1.80}$	$^{+1.50}_{-1.80}$
LUMI	±1.70	±1.70	±1.70	±1.70	±1.70	±1.70	±1.70	±1.70	±1.70
MUONS_ID	–	$^{-0.04}_{+0.07}$	$^{-0.44}_{+0.05}$	–	$^{-0.08}_{+0.10}$	$^{-0.30}_{+0.19}$	–	$^{-0.11}_{+0.03}$	$^{+0.00}_{-0.19}$
MUONS_MS	–	$^{-0.26}_{+0.03}$	$^{-0.54}_{+0.05}$	–	$^{-0.21}_{+0.11}$	$^{-0.42}_{+0.07}$	–	∓0.18	$^{-0.25}_{+0.18}$
MUONS_SAGITTA_RESBIAS	–	$^{+0.16}_{-0.00}$	$^{+0.15}_{-0.00}$	–	$^{+0.14}_{-0.00}$	$^{+0.12}_{-0.00}$	–	$^{+0.20}_{-0.00}$	$^{+0.06}_{-0.00}$
MUONS_SAGITTA_RHO	–	$^{+0.16}_{-0.00}$	$^{+0.08}_{-0.00}$	–	$^{+0.11}_{-0.00}$	$^{+0.11}_{-0.00}$	–	$^{+0.17}_{-0.00}$	$^{+0.00}_{-0.02}$
MUONS_SCALE	–	$^{-0.06}_{+0.02}$	$^{+0.27}_{-0.56}$	–	$^{+0.00}_{-0.02}$	$^{+0.15}_{-0.35}$	–	$^{+0.00}_{-0.04}$	$^{+0.00}_{-0.15}$
MUON_EFF_ISO_STAT	–	±0.15	±0.25	–	±0.16	±0.27	–	±0.12	±0.22
MUON_EFF_ISO_SYS	–	±0.80	$^{+1.52}_{-1.50}$	–	$^{+0.82}_{-0.81}$	$^{+1.57}_{-1.55}$	–	±0.79	$^{+1.54}_{-1.52}$
MUON_EFF_RECO_STAT	–	±0.14	±0.29	–	±0.14	±0.28	–	±0.15	±0.30
MUON_EFF_RECO_STAT_LOWPT	–	±0.09	±0.15	–	±0.10	±0.16	–	±0.08	±0.15
MUON_EFF_RECO_SYS	–	$^{+0.65}_{-0.63}$	$^{+1.35}_{-1.32}$	–	$^{+0.66}_{-0.65}$	$^{+1.35}_{-1.31}$	–	$^{+0.69}_{-0.68}$	$^{+1.41}_{-1.38}$
MUON_EFF_RECO_SYS_LOWPT	–	±0.30	±0.50	–	±0.31	±0.53	–	±0.25	±0.45
MUON_EFF_TTVA_STAT	–	$^{+0.13}_{-0.14}$	$^{+0.22}_{-0.24}$	–	$^{+0.13}_{-0.14}$	$^{+0.23}_{-0.25}$	–	$^{+0.11}_{-0.12}$	$^{+0.20}_{-0.21}$
MUON_EFF_TTVA_SYS	–	±0.02	±0.03	–	±0.02	$^{+0.04}_{-0.03}$	–	$^{+0.02}_{-0.01}$	±0.03
PileupWeight	$^{-1.39}_{+1.23}$	$^{-0.77}_{+0.74}$	$^{-0.38}_{+0.28}$	$^{-1.43}_{+1.37}$	$^{-0.82}_{+0.68}$	$^{-0.79}_{+0.61}$	$^{-1.42}_{+1.31}$	$^{-0.95}_{+0.80}$	∓0.35
QCDSCALESIGNAL	$^{+0.80}_{+0.00}$	$^{+0.80}_{+0.00}$	$^{+0.80}_{+0.00}$	$^{+0.80}_{+0.00}$	$^{+0.80}_{+0.00}$	$^{+0.80}_{+0.00}$	$^{+0.80}_{+0.00}$	$^{+0.80}_{+0.00}$	$^{+0.80}_{+0.00}$
Total	$^{+11.78}_{-10.83}$	$^{+6.49}_{-6.19}$	$^{+3.70}_{-3.46}$	$^{+12.93}_{-11.84}$	$^{+6.75}_{-6.46}$	$^{+3.77}_{-3.49}$	$^{+12.73}_{-11.65}$	$^{+6.83}_{-6.55}$	$^{+3.72}_{-3.38}$

Table 6.8: Relative experimental (theoretical) systematic uncertainties in percent on signal region yield (acceptance) for $h \to Z_d Z_d \to 4\ell$ signal in 2015–2018 simulation. In cases where the UP and DOWN variations result in uncertainties with different signs, the sign has been preserved, with the uncertainty in the upper (lower) position corresponding to the UP (DOWN) variation. If they are equal in magnitude, one number is shown, with the upper (lower) sign corresponding to the UP (DOWN) variation. In cases where both variations result in uncertainties that are both positive (negative), the larger (smaller) of the two is shown with the sign preserved in the upper (lower) position. The other position is set to zero. In cases where both variations result in zero uncertainties, – is shown. Systematics that affect the event weight usually increase/decrease the yield with UP/DOWN variations, thus resulting in a positive/negative relative uncertainty. Systematics that affect the kinematics may not be consistent in this manner. The totals are computed by summing in quadrature the most positive (negative) uncertainties.

Systematic	$h \rightarrow ZZ^* \rightarrow 4\ell$			$ZZ^* \rightarrow 4\ell$		
	$4e$	$2e2\mu$	4μ	$4e$	$2e2\mu$	4μ
EG_RESOLUTION_ALL	$^{-0.33}_{+0.02}$	$^{-0.22}_{+0.17}$	$^{-0.00}_{-0.02}$	$^{+0.00}_{-1.33}$	$^{+0.57}_{-0.00}$	$^{+0.07}_{-0.00}$
EG_SCALE_AF2	–	–	–	–	$^{+0.00}_{+0.23}$	–
EG_SCALE_ALL	$^{-0.97}_{+0.84}$	$^{-0.31}_{+0.17}$	$^{-0.08}_{+0.07}$	$^{+0.00}_{-0.93}$	$^{+0.49}_{-0.00}$	$^{-0.22}_{+0.03}$
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	$^{+11.87}_{-10.80}$	$^{+5.68}_{-5.54}$	–	$^{+13.18}_{-11.83}$	$^{+6.26}_{-6.09}$	–
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	$^{+0.79}_{-0.78}$	±0.39	–	±1.26	±0.62	–
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	$^{+2.78}_{-2.73}$	$^{+1.38}_{-1.37}$	–	$^{+3.63}_{-3.54}$	$^{+1.75}_{-1.74}$	–
LUMI	±1.70	±1.70	±1.70	±1.70	±1.70	±1.70
MUONS_ID	–	$^{+0.06}_{-0.00}$	$^{+0.00}_{-0.09}$	–	$^{+0.79}_{-0.41}$	$^{+0.52}_{-0.05}$
MUONS_MS	–	$^{-0.24}_{+0.19}$	$^{-0.34}_{+0.17}$	–	$^{+0.17}_{-0.54}$	$^{+0.27}_{-0.00}$
MUONS_SAGITTA_RESBIAS	–	$^{+0.06}_{-0.00}$	$^{+0.00}_{-0.01}$	–	$^{+0.65}_{-0.00}$	$^{+0.34}_{-0.00}$
MUONS_SAGITTA_RHO	–	$^{+0.04}_{-0.00}$	$^{+0.00}_{-0.09}$	–	$^{+0.00}_{-0.17}$	$^{+0.06}_{-0.00}$
MUONS_SCALE	–	$^{+0.07}_{-0.18}$	$^{+0.48}_{-0.66}$	–	$^{+0.25}_{-0.52}$	$^{-0.77}_{+0.09}$
MUON_EFF_ISO_STAT	–	±0.12	±0.22	–	±0.19	±0.39
MUON_EFF_ISO_SYS	–	±0.76	$^{+1.50}_{-1.48}$	–	±0.92	$^{+1.79}_{-1.76}$
MUON_EFF_RECO_STAT	–	±0.15	±0.31	–	±0.15	±0.30
MUON_EFF_RECO_STAT_LOWPT	–	±0.07	±0.14	–	±0.11	±0.23
MUON_EFF_RECO_SYS	–	$^{+0.70}_{-0.68}$	$^{+1.42}_{-1.39}$	–	$^{+0.67}_{-0.66}$	$^{+1.32}_{-1.29}$
MUON_EFF_RECO_SYS_LOWPT	–	±0.23	±0.42	–	±0.36	±0.76
MUON_EFF_TTVA_STAT	–	$^{+0.11}_{-0.12}$	$^{+0.20}_{-0.21}$	–	$^{+0.15}_{-0.16}$	$^{+0.32}_{-0.34}$
MUON_EFF_TTVA_SYS	–	$^{+0.02}_{-0.01}$	±0.03	–	±0.02	$^{+0.05}_{-0.04}$
PDFALPHASggHZZ	±2.82	±2.80	±2.79	–	–	–
PDFZZ	–	–	–	$^{+2.50}_{-1.58}$	$^{+2.41}_{-1.52}$	$^{+2.53}_{-1.59}$
PileupWeight	$^{-1.65}_{+1.57}$	$^{-0.93}_{+0.90}$	$^{-0.54}_{+0.48}$	$^{-0.26}_{+0.25}$	$^{-0.83}_{+0.56}$	$^{-0.67}_{+0.28}$
QCDSCALEZZ	–	–	–	$^{+4.82}_{-4.36}$	$^{+4.64}_{-4.20}$	$^{+4.86}_{-4.40}$
QCDSCALEggHZZ	$^{+6.91}_{-7.36}$	$^{+6.88}_{-7.33}$	$^{+6.85}_{-7.30}$	–	–	–
Total	$^{+14.53}_{-13.91}$	$^{+9.72}_{-9.97}$	$^{+7.92}_{-8.33}$	$^{+14.86}_{-13.46}$	$^{+8.74}_{-8.14}$	$^{+6.28}_{-5.62}$

Table 6.9: Relative systematic uncertainties in percent on signal region yield for $h \rightarrow ZZ^* \rightarrow 4\ell$ and $ZZ^* \rightarrow 4\ell$ backgrounds in 2015–2018 simulation. In cases where the UP and DOWN variations result in uncertainties with different signs, the sign has been preserved, with the uncertainty in the upper (lower) position corresponding to the UP (DOWN) variation. If they are equal in magnitude, one number is shown, with the upper (lower) sign corresponding to the UP (DOWN) variation. In cases where both variations result in uncertainties that are both positive (negative), the larger (smaller) of the two is shown with the sign preserved in the upper (lower) position. The other position is set to zero. In cases where both variations result in zero uncertainties, – is shown. Systematics that affect the event weight usually increase/decrease the yield with UP/DOWN variations, thus resulting in a positive/negative relative uncertainty. Systematics that affect the kinematics may not be consistent in this manner. The totals are computed by summing in quadrature the most positive (negative) uncertainties.

Chapter 7

Data-driven fake background estimation

7.1 Introduction

Signal regions are optimized to accept signal and reject background. For a Beyond-the-Standard-Model (BSM) search such as $h \to Z_d Z_d \to 4\ell$, Standard Model (SM) processes form the background. These backgrounds can be divided into two categories: reducible backgrounds and irreducible backgrounds:

- Irreducible backgrounds: SM processes that produce the same final state as the BSM process (e.g. $h \to ZZ^* \to 4\ell$ and $ZZ^* \to 4\ell$).
- Reducible (fake) backgrounds: SM processes that do not have the same final state as the BSM process, but can "fake" the same final state (e.g. jets from the Z+jets process being misidentified or misreconstructed as electrons). These are also known as *fake* backgrounds.

Simulation usually provides both *accurate* and *precise* estimates of irreducible backgrounds. Event generation of these processes and simulation of true leptons through the ATLAS detector and their reconstruction are generally well-established and considered reliable. These backgrounds are simulated in quantities that are large enough to enable precise predictions (i.e. with statistical uncertainties that are small relative to the estimated yields themselves). Reducible (fake) backgrounds, however, are difficult to quantify. For example, pions contained inside a jet will decay to photons, electrons, and muons. These leptons can be identified or reconstructed independently from the jet and are considered *fake leptons* not from the hard scatter interaction.[1] *Punch-through jets* that are not fully contained within the calorimetry and enter the Muon Spectrometer can result in tracks that are reconstructed into fake muons. The modelling of misidentification and misreconstruction that leads to fake leptons is not accurate, unlike the modelling of true leptons. Additionally, misidentification and misreconstruction are thought to be rare phenomena that occur at rates that are negligible compared to the rates of irreducible processes. Therefore, simulations of events containing leptons from misidentification and misreconstruction are also imprecise due to lack of sufficient statistics.

The detector signature of this search—four prompt leptons—is rare in itself. The application of kinematic and quality requirements on the leptons results in a further winnowing of events, to the extent that only a handful of events with four leptons from irreducible processes survive to enter the signal region. If a lepton from misidentification or misreconstruction is an uncommon occurence compared to the occurence of a true lepton, then the occurence of an event with four leptons—where one or more are fake—is *extraordinarily* rare. Indeed, estimates of Z+jets from simulation predict many bins with negative (unphysical) yields,[2] with statistical uncertainties on the overall yield (yield summed over bins) of the order of the yield itself.

An alternative to estimating fake backgrounds from simulation is to estimate them from data. In general, *data-driven* (*in-situ*) techniques measure the rate of fake leptons occurring in kinematic regions that are rich in fake leptons and extrapolate the rate to the signal region in order to obtain an estimated yield. The benefit of these techniques is that they rely minimally on the modelling of misidentification and misreconstruction and also minimally on the amount of simulated events.

In the 2015–2016 iteration of this search, an attempt was made to cross-check simulation-based fake background estimates with a data-driven estimate but due to insufficient statistics in data, the data-driven estimate was not usable. Similar to how

---

[1] Some of these backgrounds not exactly fake, in the sense that they are true electrons and muons that were created in *non-prompt* decays. The mechanism of jet formation (the hadronization of quarks and gluons) produces neutral and charged pions. Neutral pions decay to photons and electrons and charged pions decay to electrons and muons. Charged pions in particular have a relatively long lifetime (10^{-8} s compared to 10^{-17} s for neutral pions). This phenomenon is distinct from prompt electrons being identified from energy deposits in the electromagnetic calorimeter, for example.

[2] Events from Monte Carlo-based event generation are sometimes negatively weighted. This is due to how the phase space is integrated. Events generated this way are meant to be considered in aggregate (unlike events in data, which are always considered independently). If the number of generated events is limited, then it is possible for the sum of negatively-weighted events to exceed the sum of positively-weighted events. This results in yields that are negative.

insufficient Monte Carlo statistics and negatively-weighted Monte Carlo events can conspire to give negative yields, insufficient statistics *in data* in particular kinematic regions and negative transfer factors can also result in negative yields. Furthermore, systematic uncertainties on the data-driven estimate were not assessed. In the 2015–2018 iteration, the data-driven estimate was improved by quadrupling the dataset size ($139\ \text{fb}^{-1}$ vs. $36\ \text{fb}^{-1}$) and by assessing systematic uncertainties.

7.2 Methodology

The data-driven Z+jets background estimate uses (i) inverted signal regions that are rich in the fake background to be estimated (i.e. with one or two fake leptons) in order to map the fake factors to the signal region, and (ii) regions that are non-overlapping with the signal region and also enriched in the fake background in order to derive the fake factors.

First, inverted lepton selection cuts need to be defined. These cuts are not the exact opposite of the nominal selection. In the nominal lepton selection (used for the signal region, as described in Chapter 5 and in Table 5.1):

- Electrons are required to pass the *Loose* likelihood identification working point and pass the `FCLoose` isolation working point, and
- Muons are required to pass the `FCLoose_FixedRad` isolation working point and have transverse impact parameter significance $|d_0/\sigma_{d_0}| < 3$.[3]

The inverted lepton selection cuts are defined as:

- Electrons fail either the identification working point or the isolation working point (but not both), and
- Muons fail either (or both) the isolation working point and the transverse impact parameter significance $|d_0/\sigma_{d_0}|$.

The regions used to map the fake factors to the signal region are defined as:

- Signal region A: the region in which the fake background is to be estimated.

---

[3] This is in addition to *Loose* identification for muons and other kinematic requirements.

- Inverted signal regions B1, B1′, and B2: same as region A except one or two leptons in the secondary pair ℓ_{34} pass the inverted lepton selection (from here, ℓ_1 and ℓ_2 are used as labels for the two leptons of the ℓ_{34} pair):[4]
  - Region B1: ℓ_1 passes the nominal selection ("good") and ℓ_2 passes the inverted selection ("bad").
  - Region B1′: ℓ_1 passes the inverted selection ("bad") and ℓ_2 passes the nominal selection ("good").
  - Region B2: both ℓ_1 and ℓ_2 pass both the inverted selection (both "bad").

The regions used for deriving the fake factors are defined as:

- Regions C and D: Enriched in Z+jets events. Events must have two good leptons that reconstruct a Z boson and at least one other lepton. The Z boson candidate with the smallest $|m_{\ell\ell} - m_Z|$ is chosen, with the requirement that $|m_{\ell\ell} - m_Z| < 10$ GeV.
  - Region C: extra lepton passes the nominal selection ("good").
  - Region D: extra lepton passes the inverted selection ("bad").

If regions C and D are highly enriched in Z+jets events, then every extra lepton—aside from the leptons forming the Z boson candidate—is likely to be fake. The number of leptons entering regions C and D are first measured in data as a function of lepton flavour, p_T, and η. Every extra lepton is counted as an entry into a region. In order to correct for real (*true*) leptons contamininating regions C and D, the same measurements are also made in simulation of processes with multiple real leptons and are subtracted from the measurements in data.

After correction through subtraction, *fake factors* are calculated by dividing the number of leptons entering region C by the number of leptons entering region D. They are calculated separately for muons and electrons, and binned in p_{T} and η.[5] The fake factor is the ratio of good fake leptons to bad fake leptons and counts the number of

---

[4] The dilepton closest to m_Z is labelled ℓ_{12} and the other dilepton is labelled ℓ_{34}. Since ℓ_{34} is farther away from m_Z and has less stringent kinematic requirements than ℓ_{12}, it is more likely for ℓ_{34} to have one or two fake leptons than ℓ_{12}.

[5] The binning in η of the electron fake factors takes into account the geometry of the calorimeter. The *crack region* $1.37 < |\eta| < 1.52$ is the transition region between the barrel and end-cap calorimeters, and contains services for the Inner Detector. The additional material in this region affects resolution, therefore separate bins are assigned to this region.

leptons, not the number of events (i.e. if an event has multiple extra leptons, each extra lepton contributes to the numerator or denominator).

$$f = N_C/N_D \tag{7.1}$$

The fake factors are used to calculate the fake background contributions in region A (signal region) by *transferring* events measured in regions B1, B1′, and B2, as shown in Figure 7.1:

$$N_{\text{good,fake}} = f \cdot N_{\text{bad,fake}} \tag{7.2}$$

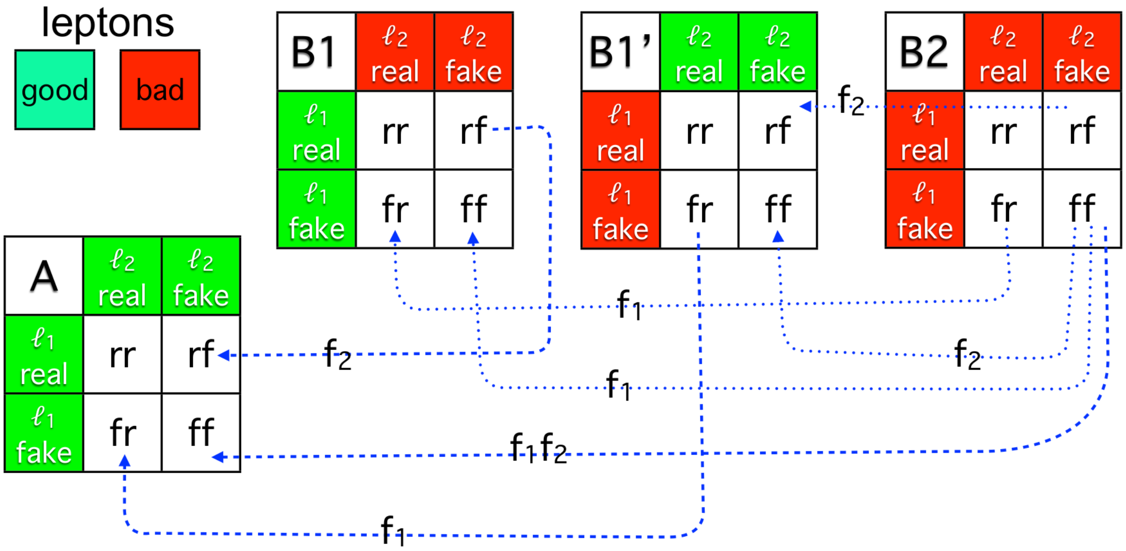


Figure 7.1: Categorization of leptons into regions of the fake factor method [98]. Region B2—where both leptons are “bad”—is used to extrapolate the yields in regions B1 and B1′ where one lepton is “bad” and the other is “good”. The white-coloured squares represent the four possible combinations of ℓ_1 and ℓ_2 being real or fake. In the signal region (A), ℓ_1 and ℓ_2 are identified as “good” (even if they may be fake). To estimate the number of events with fakes in the signal region, the fake factor is evaluated for ℓ_1 (f_1) or ℓ_2 (f_2). Equation 7.2 is then used to calculate the estimate.

The number of events in each region is the sum of events with two real leptons, two fake leptons, and one real and one fake lepton:

$$\begin{aligned} N_A &= N_A^{rr} + N_A^{rf} + N_A^{fr} + N_A^{ff} \equiv N_A^{rr} + N_A^{\text{fake}} \\ N_{B1} &= N_{B1}^{rr} + N_{B1}^{rf} + N_{B1}^{fr} + N_{B1}^{ff} \\ N_{B1'} &= N_{B1'}^{rr} + N_{B1'}^{rf} + N_{B1'}^{fr} + N_{B1'}^{ff} \\ N_{B2} &= N_{B2}^{rr} + N_{B2}^{rf} + N_{B2}^{fr} + N_{B2}^{ff} \end{aligned} \tag{7.3}$$

The terms corresponding to events with fake lepton(s) is estimated with the fake factors using Equation 7.2:

$$\begin{aligned} &N_A^{ff} = f_1 f_2 N_{B2}^{ff} \quad && N_{B1}^{ff} = f_1 N_{B2}^{ff} \quad && N_{B1'}^{ff} = f_2 N_{B2}^{ff} \\ &N_A^{rf} = f_2 N_{B1}^{rf} && N_{B1}^{fr} = f_1 N_{B2}^{fr} && N_{B1'}^{rf} = f_2 N_{B2}^{rf} \\ &N_A^{fr} = f_1 N_{B1'}^{fr} \end{aligned} \tag{7.4}$$

The final fake background estimate is given by:

$$\begin{aligned} N_A^{\text{fake}} &\equiv N_A^{rf} + N_A^{fr} + N_A^{ff} \\ &= f_2 N_{B1}^{rf} + f_1 N_{B1'}^{fr} + f_1 f_2 N_{B2}^{ff} \\ &= f_2 \left(N_{B1} - N_{B1}^{rr} - N_{B1}^{fr} - N_{B1}^{ff} \right) + f_1 \left(N_{B1'} - N_{B1'}^{rr} - N_{B1'}^{rf} - N_{B1'}^{ff} \right) \\ &\quad + f_1 f_2 N_{B2}^{ff} \\ &= f_2 \left(N_{B1} - N_{B1}^{rr} - f_1 N_{B2}^{fr} - f_1 N_{B2}^{ff} \right) + f_1 \left(N_{B1'} - N_{B1'}^{rr} - f_2 N_{B2}^{rf} - f_2 N_{B2}^{ff} \right) \\ &\quad + f_1 f_2 N_{B2}^{ff} \\ &= f_2 \left(N_{B1} - N_{B1}^{rr} \right) + f_1 \left(N_{B1'} - N_{B1'}^{rr} \right) - f_1 f_2 \left(N_{B2}^{rf} + N_{B2}^{fr} + N_{B2}^{ff} \right) \\ &= f_2 \left(N_{B1} - N_{B1}^{rr} \right) + f_1 \left(N_{B1'} - N_{B1'}^{rr} \right) - f_1 f_2 \left(N_{B2} - N_{B2}^{rr} \right) \\ &= \underbrace{f_2 N_{B1} + f_1 N_{B1'} - f_1 f_2 N_{B2}}_{\text{estimated from data}} - \underbrace{f_2 N_{B1}^{rr} + f_1 N_{B1'}^{rr} - f_1 f_2 N_{B2}^{rr}}_{\text{estimated from } 4\ell \text{ irreducible sim.}} \end{aligned} \tag{7.5}$$

The first term is estimated purely from data. The second term is the fake contribution from events with four real leptons. This term is estimated from the irreducible background simulation (e.g. $h \to ZZ^* \to 4\ell$ and $ZZ^* \to 4\ell$), with a *truth matching* requirement. This requirement ensures that each reconstructed lepton in simulation originates from a generated (real) W or Z boson.[6]

---

[6] In simulation, the origin of each lepton is known and can be traced.

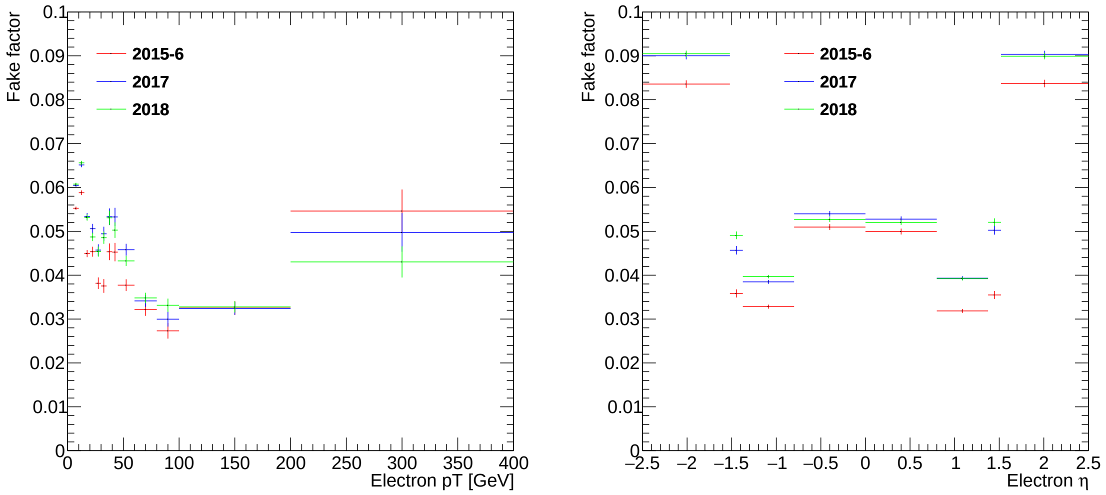
7.3 Fake factors

The electron and muon fake factors are parametrized in p_T and η and are applied in the fake estimates as a function of p_T and η. Shown in Figures 7.2 and 7.3 are the fake factors and fake rates per data-taking campaign projected onto one variable only and corrected for true lepton contamination with simulation. Also shown is the *fake rate* defined as $f/(1+f) = N_{\text{good}}/(N_{\text{good}} + N_{\text{bad}})$. It is a measure of the number of fakes passing the "good" criteria divided by *all fakes*.[7] The fake factors and fake rates can vary for each data-taking campaign due to differences in operating conditions (for example, increased pileup and multi-jet background resulting in increased rates of misidentification and misreconstruction). The differences are small enough that the fake factors are combined across data-taking campaigns in order to reduce the statistical uncertainty.

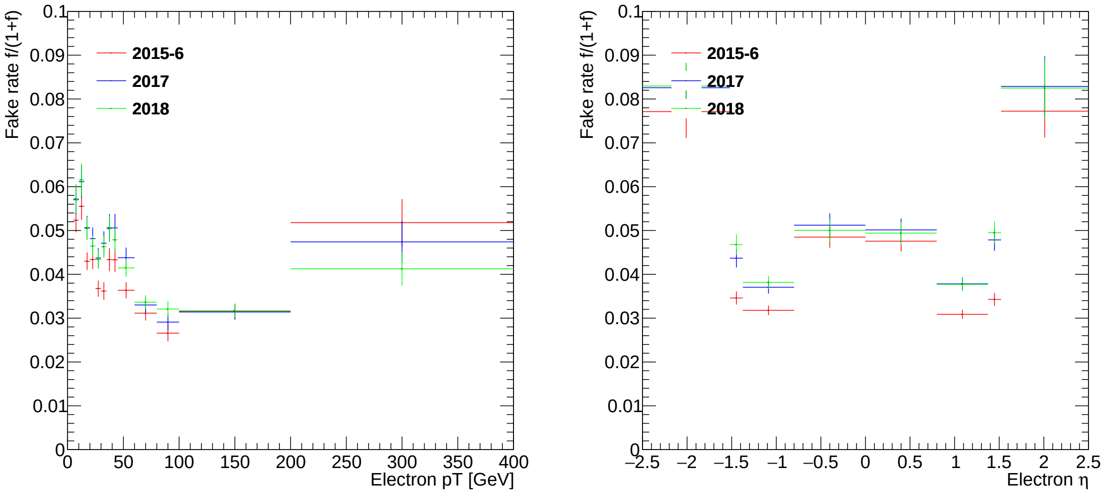
Shown in Figure 7.4 are the electron and muon fake factors combined across data-taking campaigns with and without the correction for true lepton contamination. The contamination effect is significant, especially from the $WZ \rightarrow \ell\nu\ell\ell$ process. In region C of the muon fake factors, the contamination from this process is 12% ($WZ \rightarrow \ell\nu\ell\ell$ events/data). Specifically in the highest p_T bin of this region, it is 40%. The contamination from $h \rightarrow ZZ^* \rightarrow 4\ell$ and $ZZ^* \rightarrow 4\ell$ processes is 2% across all bins. With this correction, the fake factors become smaller, and consequently, the fake estimates also become smaller.

---

[7] It is not possible to measure the *true* fake rate: number of fakes divided by all leptons.

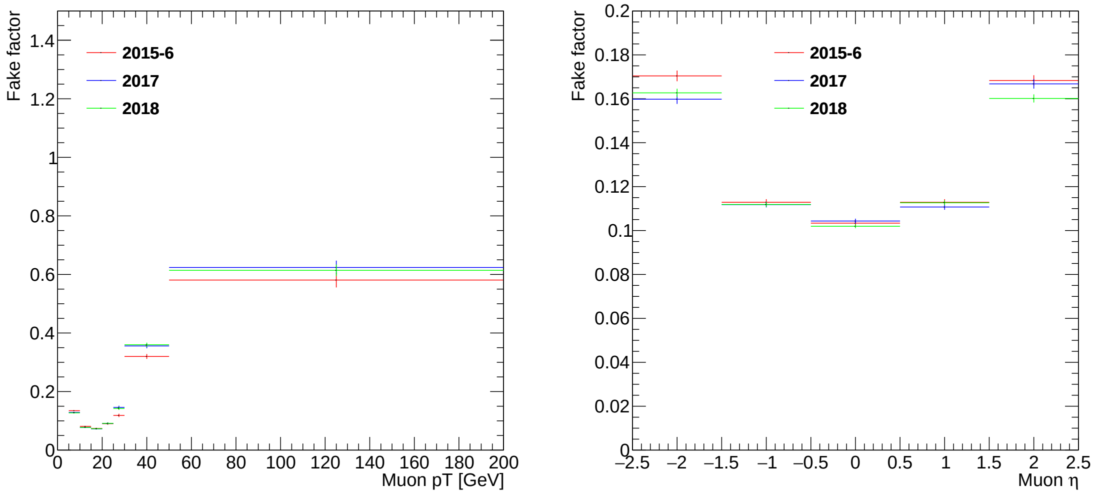


(a) Electron fake factors in p_T (b) Electron fake factors in η

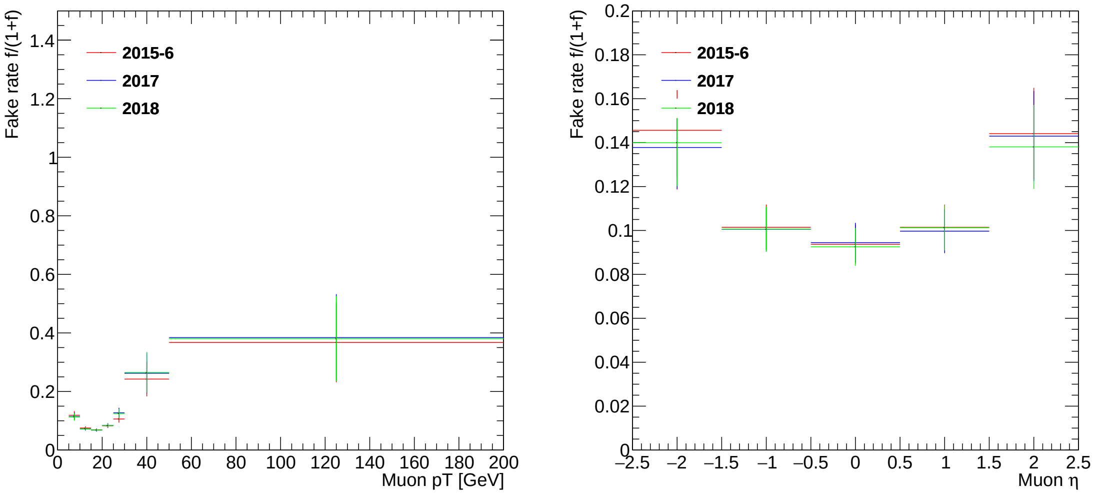


(c) Electron fake rate in p_T (d) Electron fake rate in η

Figure 7.2: Electron fake factors and fake rates derived from 2015–2016, 2017, and 2018 data projected onto one variable only and corrected for true lepton contamination with simulation-based estimates.

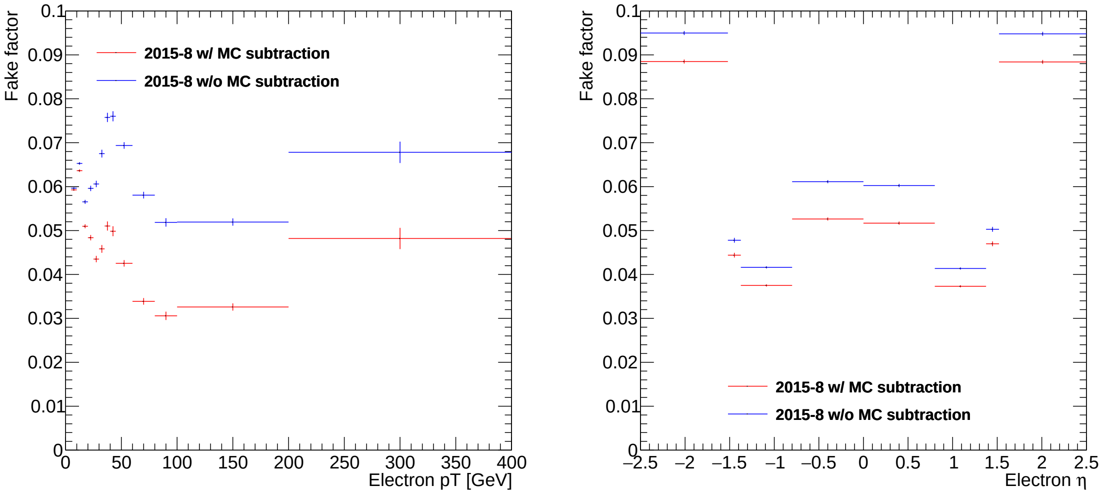


(a) Muon fake factors in p_T (b) Muon fake factors in η

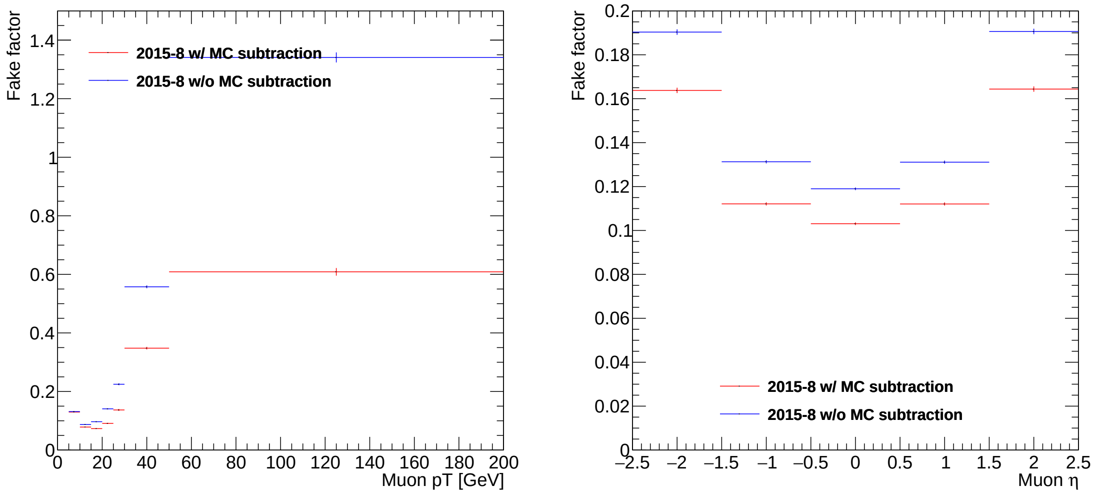


(c) Muon fake rate in p_T (d) Muon fake rate in η

Figure 7.3: Muon fake factors and fake rates derived from 2015–2016, 2017, and 2018 data projected onto one variable only and corrected for true lepton contamination with simulation-based estimates.



(a) Electron fake factors in p_T (b) Electron fake factors in η



(c) Muon fake factors in p_T (d) Muon fake factors in η

Figure 7.4: Electron and muon fake factors combined across data-taking campaigns (2015–2018) projected onto one variable only, with and without correction for true lepton contamination.

7.4 Fake estimates

7.4.1 Results

The fake estimates are shown in Table 7.1. The estimate of Z+jets with the fake factor method, combined across all channels and data-taking campaigns, is 0.43 with a systematic uncertainty of $^{+0.17}_{-0.01}$ and statistical uncertainty of ±0.39. The estimate is compatible with zero within approximately 1σ of the statistical uncertainty.

Fake estimates are derived from data (four leftmost columns) and 4ℓ Monte Carlo (MC)-based simulation (four middle columns). The second is subtracted from the first in order to produce the final fake estimate (four rightmost columns).[8] The fake estimates per channel are summed together to produce the fake estimate across all channels.

Some of the estimates are negative due to the simulation subtraction step and statistical fluctuations. Requiring truth matching reduces the estimates from 4ℓ simulation by approximately 20%.

The *statistical uncertainty on the estimate* arises from the finiteness of the numbers of events entering regions B1, B1′, and B2. The statistical uncertainty on the final estimate is ±0.39.

To estimate the systematic uncertainty due to the *statistical uncertainty on the fake factors*, statistical variations of the factors are made by varying each (p_{T}, η) bin within its statistical uncertainty. These variations are drawn from a Gaussian centered at the nominal fake factor with standard deviation equal to the statistical uncertainty on the nominal fake factor. Only positive-valued variations are accepted. Ten variations are generated for the data-derived fake factors and the simulation-derived fake factors. The fake estimate is then computed with each set of fake factors and the envelope of the fake estimates derived from statistical variation is taken as the systematic uncertainty associated with the fake factors. This systematic uncertainty on the final estimate is ±0.01.

The estimation of the systematic uncertainties *on the method* is described in detail in the following section. This systematic uncertainty is $^{+0.17}_{-0.00}$ and is combined with

---

[8] The second estimate is the fake estimate from events with four real leptons and is assessed from 4ℓ simulation. Since the first term is the fake estimate from data and thus includes fakes from four real leptons, the second term must be subtracted from the first in order to produce a estimate that includes only fake leptons.

the aforementioned systematic uncertainty ±0.01 in quadrature to yield the final systematic uncertainty $^{+0.17}_{-0.01}$.

	Yield ± Stat (statistical uncertainty on fake estimate yield)											
Systematic definition	Fake estimate from data				Fake estimate from 4ℓ simulation				**Final fake estimate (data - simulation)**			
	$4e$	4μ	$2e2\mu$	All	$4e$	4μ	$2e2\mu$	All	$4e$	4μ	$2e2\mu$	All
Nominal	0.15 ± 0.13	-0.29 ± 0.16	0.73 ± 0.33	0.59 ± 0.39	0.03 ± 0.00	0.07 ± 0.00	0.06 ± 0.00	0.16 ± 0.00	0.13 ± 0.13	-0.36 ± 0.16	0.67 ± 0.33	0.43 ± 0.39
Systematic uncertainty from FF statistical uncertainty: max(|yield w/ nominal FF - yield w/ Gaussian-varied nominal FF|)												± 0.01
Looser	0.24 ± 0.03	0.03 ± 0.12	0.40 ± 0.15	0.67 ± 0.20	0.00 ± 0.00	0.05 ± 0.00	0.03 ± 0.00	0.08 ± 0.00	0.23 ± 0.03	-0.02 ± 0.12	0.38 ± 0.15	0.60 ± 0.20
LooserElKH	0.28 ± 0.04	-0.29 ± 0.16	0.61 ± 0.28	0.60 ± 0.33	0.00 ± 0.00	0.07 ± 0.00	0.04 ± 0.00	0.11 ± 0.00	0.28 ± 0.04	-0.36 ± 0.16	0.57 ± 0.28	0.49 ± 0.33
d0sigElKH	0.28 ± 0.04	-0.29 ± 0.16	0.63 ± 0.28	0.62 ± 0.33	0.00 ± 0.00	0.07 ± 0.00	0.04 ± 0.00	0.11 ± 0.00	0.27 ± 0.04	-0.36 ± 0.16	0.59 ± 0.28	0.50 ± 0.33
Systematic uncertainty from definition: max(|yield w/ nominal def. - yield w/ varied def.|)												${}^{+0.17}_{-0}$
Yield from Z+jets simulation-based closure test (FF and estimation from Z+jets simulation)												0.09 ± 0.11
Final fake estimate ± stat. ± syst.												$0.43 \pm 0.39 \, {}^{+0.17}_{-0.01}$

Table 7.1: Z+jets estimates with the fake factor method. The estimates from 4ℓ simulation are subtracted from the estimates from data in order to produce the final fake estimates. The estimates are determined per channel and then summed together. The systematic uncertainty arising from the statistical uncertainty on the fake factors is assessed by varying the fake factors within their statistical uncertainties and re-calculating the estimates. The systematic uncertainties arising from the fake factor method itself is assessed by re-calculating the fake factors and fake estimates with variations of the definitions. For both uncertainties, the envelope of the fake estimates is taken as the final uncertainty.

7.4.2 Systematic uncertainty

To estimate the systematic uncertainty on the fake factor method, three variations on the lepton "good" and "bad" definitions are considered:

Nominal definition, as described in Section 7.2:

- Good electron: required to pass the *Loose* likelihood identification working point and pass the `FCLoose` isolation working point.
- Bad electron: fails either *Loose* identification or `FCLoose` isolation, but not both (electrons failing both are considered to be too removed from the good definition and therefore are not considered).
- Good muon: required to pass the `FCLoose_FixedRad` isolation working point and have transverse impact parameter significance $|d_0/\sigma_{d_0}| < 3$.
- Bad muon: fails either `FCLoose_FixedRad` isolation and/or $|d_0/\sigma_{d_0}| < 3$ (an inclusive OR is used here for increased statistics).

Looser definition: remove $z_0 \sin\theta < 0.5$ requirement from the nominal cut flow and use it for inversion in the good/bad lepton definitions, and use inclusive OR for bad electrons:

- Good electron: passes *Loose* identification and `FCLoose` isolation and $z_0 \sin\theta < 0.5$.
- Bad electron: fails any of the above.
- Good muon: passes $|d_0/\sigma_{d_0}| < 3$ and `FCLoose_FixedRad` isolation and $z_0 \sin\theta < 0.5$.
- Bad muon: fails any of the above.

LooserElKH definition: variation only in electron good/bad definition by using an inclusive OR for bad electrons:

- Good electron: passes *Loose* identification and `FCLoose` isolation.
- Bad electron: fails any of the above.
- Good muon and bad muon: same as nominal definition.

d0sigElKH definition: variation only in electron good/bad definition by additionally inverting $|d_0/\sigma_{d_0}| < 5$ cut (< 5 for electrons and < 3 for muons):

- Good electron: passes *Loose* identification and `FCLoose` isolation and $|d_0/\sigma_{d_0}| < 5$.
- Bad electron: fails any of the above.
- Good muon and bad muon: same as nominal definition.

The systematic uncertainty associated with the fake factor method is $^{+0.17}_{-0.00}$. This is assessed by taking the envelope of the fake estimates derived with the varied fake factor lepton definitions.

7.4.3 Closure test

To test the method for closure, fake factors are derived from Z+jets MC-based simulation and used to produce a fake estimate also from Z+jets simulation. This test estimates a yield of 0.09 ± 0.11. This yield is compatible with the yield of Z+jets simulation in the signal region (0.02 ± 0.02) and therefore no systematic uncertainty is assigned to incomplete closure.

Chapter 8

Results

This chapter presents the results of the $h \to Z_d Z_d \to 4\ell$ search with data. Prior to unblinding the signal region, the background estimates are studied in validation regions. An overview of the validation regions is given in the following section, and the kinematic distributions of data and background predictions are compared in each validation region in the subsequent sections. Distributions of the signal region with data are shown in the final section and followed with a discussion of the results. The results of the signal region are interpreted with a statistical analysis in Chapter 9.

8.1 Validation of background estimates

8.1.1 Overview and yields

Validation regions are used to study the agreement of background predictions with data prior to unblinding the signal region. By construction, they should be:

- non-overlapping with the signal region, so that signal cannot appear in it,
- share similarities in the relevant kinematic variables with the signal region,
- should contain sufficient background statistics so that a meaningful comparison between the prediction and observation can be made.

One or more event selection requirements are inverted in order to define validation regions that are non-overlapping with the signal regions. To increase the number of events entering the region, certain requirements are also removed. The predictions of the two dominant background processes $h \to ZZ^* \to 4\ell$ and $ZZ^* \to 4\ell$ are validated in five regions, described below and compared in Table 8.1:

> **Validation region 1 (VR 1)**: Reversal of the $m_{14,32} < 75$ GeV requirement (part of the Z veto) to $m_{14,32} \geq 75$ GeV. Additionally, removal of the final signal region requirement $m_{34} > 0.85m_{12} - 0.1125f(m_{12})m_{12}$. This produces an enriched selection of Higgs and non-resonant ZZ^* backgrounds. Only $4e$ and 4μ quadruplets can be formed in this validation region because of the requirement that lepton pairs consist of same-flavour opposite-sign leptons.[1]
>
> **Validation region 2 (VR 2)**: All four dilepton mass requirements are replaced with one dilepton mass requirement $m_{12} \geq 64$ GeV. Additionally, the final signal region requirement is removed. This also produces an enriched selection of Higgs and non-resonant ZZ^* backgrounds. Since there are no cuts on the alternative pairings, this validation region can contain $2e2\mu$ quadruplets.
>
> **Validation region 3 (VR 3)**: Inversion of the final signal region requirement and inversion of the Higgs window requirement 115 GeV $< m_{4\ell} <$ 130 GeV. This validation region is dominated by the $Z \to 4\ell$ process.
>
> **Validation region 4 (VR 4)**: Inversion of the final signal region requirement and replacement of all four dilepton mass requirements with $m_{\ell\ell} < 55$ GeV. The dilepton mass requirements are lowered from $m_{12,34} < 64$ GeV and $m_{14,32} <$ 75 GeV to $m_{12,34,14,32} < 55$ GeV. This validation region is dominated by the $h \to ZZ^* \to 4\ell$ process.
>
> **Validation region 5 (VR 5)**: Inversion of the Higgs window requirement and require 100 GeV $< m_{4\ell} <$ 115 GeV or 130 GeV $< m_{4\ell} <$ 170 GeV. Additionally, remove the Z veto and final signal region requirements. This validation region examines the side-bands of the ZZ^* prediction around the Higgs mass ($m_h \approx 125$ GeV).

The overall yields of the $\langle m_{\ell\ell} \rangle$ distributions in each validation region are given in Table 8.2.

---

[1] Consider dileptons $\ell_1\ell_2 \Leftrightarrow (e, e)$ and $\ell_3\ell_4 \Leftrightarrow (\mu, \mu)$. This forms a $2e2\mu$ quadruplet. The alternative pairings are $\ell_1\ell_4 \Leftrightarrow (e, \mu)$ and $\ell_3\ell_2 \Leftrightarrow (\mu, e)$. These are not same-flavour opposite-sign pairings and are thus not permitted. Therefore, only $4e$ and 4μ quadruplets can be formed in VR 1.

Category	Name	Signal region requirement	VR 1	VR 2	VR 3	VR 4	VR 5
Event selection	Quarkonia veto	Reject event if: $(m_{J/\Psi} - 0.25\ \text{GeV}) < m_{12,34,14,23} < (m_{\Psi(2S)} + 0.30\ \text{GeV})$ and/or $(m_{\Upsilon(1S)} - 0.70\ \text{GeV}) < m_{12,34,14,23} < (m_{\Upsilon(3S)} + 0.75\ \text{GeV})$	Same as SR	Same as SR	Same as SR	Same as SR	Same as SR
	Higgs window	$115\ \text{GeV} < m_{4\ell} < 130\ \text{GeV}$	Same as SR	Same as SR	Inverted	Same as SR	$100\ \text{GeV} < m_{4\ell} < 115\ \text{GeV}$ or $130\ \text{GeV} < m_{4\ell} < 170\ \text{GeV}$
	Z veto	$10\ \text{GeV} < m_{12,34} < 64\ \text{GeV}$ $4e$ and 4μ channels: $5\ \text{GeV} < m_{14,32} < 75\ \text{GeV}$	$m_{14,32} > 75\ \text{GeV}$	$m_{12} > 64\ \text{GeV}$	Same as SR	$m_{12,34,14,32} < 55\ \text{GeV}$	Removed
	Signal region	$m_{34} > 0.85 m_{12} - 0.1125 f(m_{12}) m_{12}$	Removed	Removed	Inverted	Inverted	Removed

Table 8.1: Comparison of validation region selection criteria.

Channel	VR 1		VR 2		VR 3	
	Expected	Observed	Expected	Observed	Expected	Observed
$4e$	$31.24 \pm 0.20 \, ^{+1.66}_{-1.69}$	32	$4.25 \pm 0.06 \, ^{+0.23}_{-0.23}$	5	$65.23 \pm 3.23 \, ^{+3.62}_{-3.14}$	70
4μ	$95.97 \pm 1.98 \, ^{+4.79}_{-4.84}$	85	$14.84 \pm 0.27 \, ^{+0.76}_{-0.77}$	18	$235.43 \pm 1.66 \, ^{+14.08}_{-12.20}$	237
$2e2\mu$	–	–	$125.28 \pm 4.26 \, ^{+6.40}_{-6.48}$	135	$159.18 \pm 1.22 \, ^{+9.17}_{-7.95}$	165
Combined	$127.21 \pm 1.77 \, ^{+6.44}_{-6.53}$	117	$142.23 \pm 2.36 \, ^{+7.37}_{-7.47}$	158	$459.84 \pm 3.83 \, ^{+26.87}_{-23.28}$	472

Channel	VR 4		VR 5	
	Expected	Observed	Expected	Observed
$4e$	$0.22 \pm 0.04 \, ^{+0.01}_{-0.01}$	0	$68.68 \pm 2.72 \, ^{+3.29}_{-2.87}$	90
4μ	$0.50 \pm 0.01 \, ^{+0.02}_{-0.02}$	1	$169.83 \pm 3.69 \, ^{+9.20}_{-7.99}$	201
$2e2\mu$	$5.96 \pm 0.37 \, ^{+0.26}_{-0.26}$	5	$172.18 \pm 5.94 \, ^{+8.34}_{-7.28}$	195
Combined	$6.68 \pm 0.38 \, ^{+0.29}_{-0.30}$	6	$410.69 \pm 7.56 \, ^{+20.82}_{-18.14}$	486

Table 8.2: Expected background yields and observed yields in the validation regions per channel. The uncertainties listed are statistical and systematic, respectively. The yields exclude any underflow and overflow in the $\langle m_{\ell\ell} \rangle$ distributions.

8.1.2 Validation region 1

In validation region 1, the alternative lepton pairings ℓ_{14} and ℓ_{32} fail the Z veto. This is imposed by reversing the veto: m_{14} or $m_{32} > 75$ GeV. This allows the admission of on- and off-shell Z bosons reconstructed by ℓ_{14} and ℓ_{32}. To increase the number of events, the final signal requirement is removed. Distributions of $\langle m_{\ell\ell} \rangle$ in this region are shown in Figure 8.1. In this distribution and following distributions:

- The significance is calculated using a formula derived from the asymptotic approximations of profile likelihood statistics [99].
- The *asymmetric* error bars on data and the tabulated observed yields are calculated with 68.3% frequentist central confidence intervals (following the prescription given in [1]).
- The contribution of Z+jets is estimated from Monte Carlo-based simulation.[2]

In terms of overall yield, 127 ± 7 (stat. $\oplus$ syst.) events are predicted and 117 events are observed. The prediction is within 1σ of the statistical uncertainty of the observation. Therefore, the yields are considered to be compatible. In terms of the $\langle m_{\ell\ell} \rangle$ distribution however, two bins around $\langle m_{\ell\ell} \rangle \approx 35$ GeV have disagreement at the level of $\lesssim 2\sigma$. The events entering these bins were scrutinized and no abnormalities were found. The disagreement is likely due to statistical fluctuation.

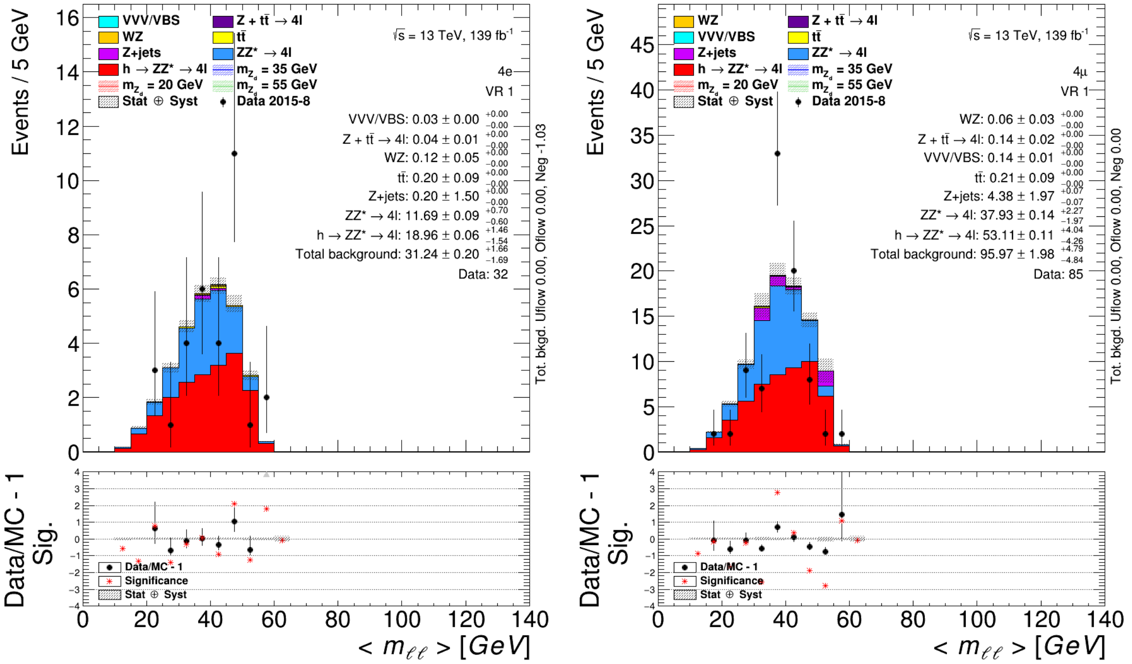
Unlike ℓ_{14} and ℓ_{32}, ℓ_{12} and ℓ_{34} do not correspond to any particles. Therefore, it is not meaningful to study $\langle m_{\ell\ell} \rangle = (m_{12} + m_{34})/2$ in this validation region. No significant excesses or deficits are observed in other kinematic variables such as $(m_{14} + m_{32})/2$ (Figure 8.2) and the mass of the alternative dilepton pair (ℓ_{14} or ℓ_{32}) closest to m_Z (Figure 8.3). The latter reconstructs m_Z acceptably.

8.1.3 Validation region 2

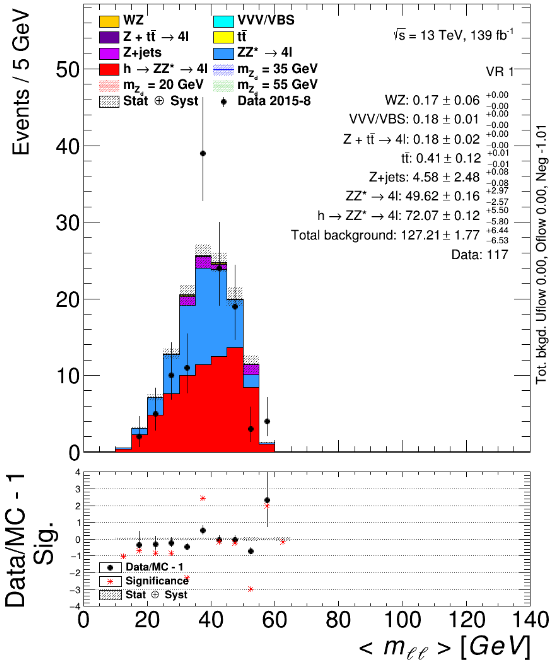
In validation region 2, all components of the Z veto are removed and replaced with the requirement $m_{12} > 64$ GeV. On- and off-shell Z bosons can be reconstructed by any lepton pairing. To increase the number of events, the final signal requirement is removed. Distributions of $\langle m_{\ell\ell} \rangle$ in this region are shown in Figure 8.4.

---

[2] As described in Section 7.1, events from Monte Carlo-based simulation are sometimes negatively weighted and are meant to be taken in aggregate with positively-weighted events. If the number of events entering a region is severely limited, the negatively-weighted events can outweigh the positively-weighted events and result in a negative (unphysical) event yield.

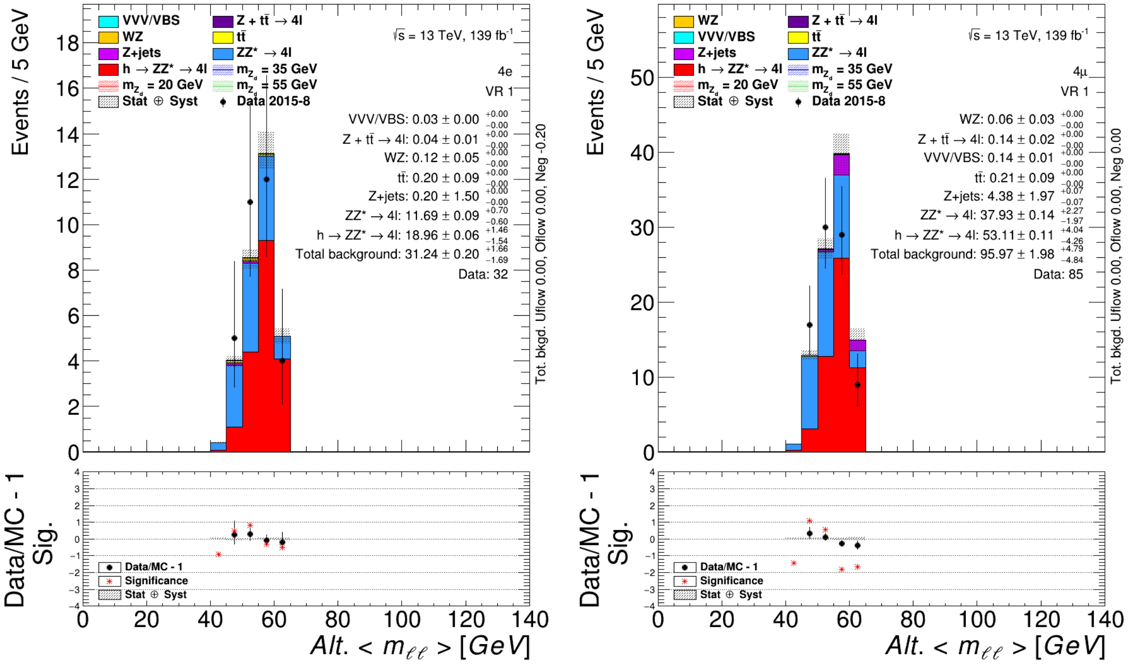


(a) $4e$ (b) 4μ

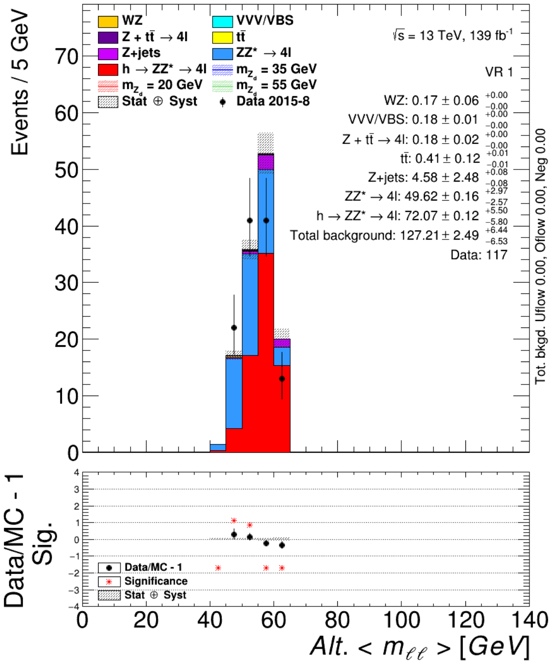


(c) All channels

Figure 8.1: Distributions of $\langle m_{\ell\ell} \rangle$ in validation region 1 in 2015–2018 data and simulation for the $4e$ and 4μ channels, and all channels combined.

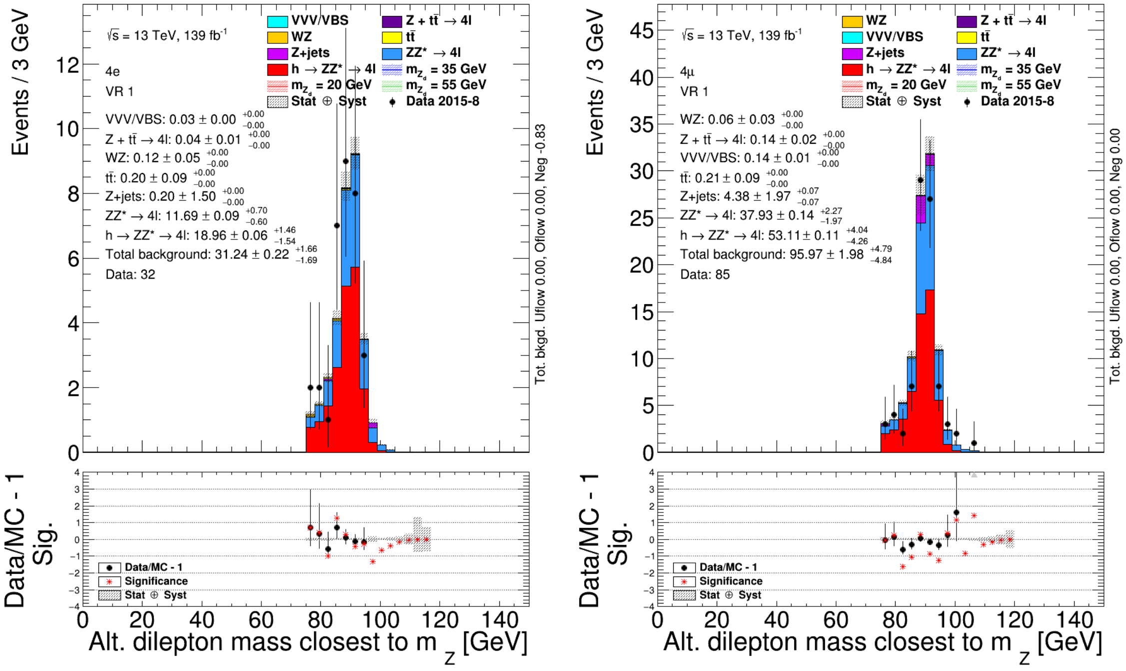


(a) $4e$ (b) 4μ



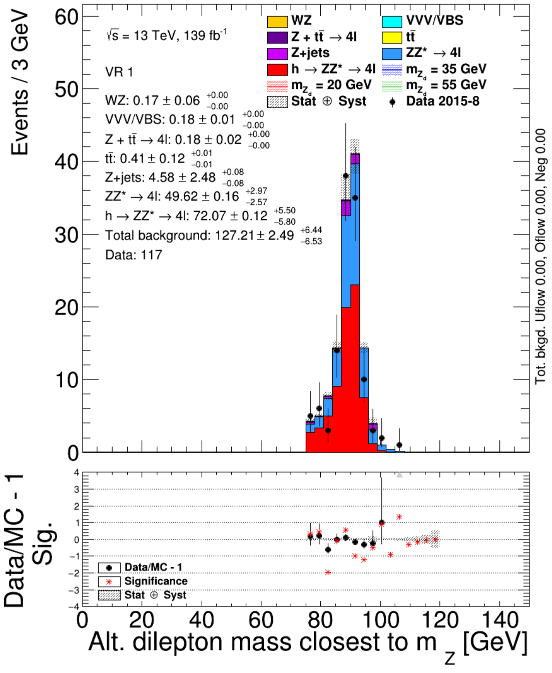
(c) All channels

Figure 8.2: Distributions of $(m_{14} + m_{32})/2$ in validation region 1 in 2015–2018 data and simulation for the $4e$ and 4μ channels, and all channels combined.



(a) $4e$

(b) 4μ



(c) All channels

Figure 8.3: Distributions of the mass of the dilepton pair (ℓ_{14} or ℓ_{32}) with mass closest to m_Z in validation region 1 in 2015–2018 data and simulation for the $4e$ and 4μ channels, and all channels combined.

In terms of overall yield, 142 ± 8 (stat. $\oplus$ syst.) events are predicted and 158 events are observed. In terms of $\langle m_{\ell\ell} \rangle$ distribution, all bins have prediction within 1σ. Both the overall yield and shape show acceptable agreement.

8.1.4 Validation region 3

In validation region 3, the Higgs window and final signal region requirements are inverted. The inversion of the Higgs window excludes the contribution of $h \rightarrow ZZ^* \rightarrow 4\ell$ processes. The Z boson of the dominant process $Z \rightarrow 4\ell$ is reconstructed by the invariant four-lepton mass $m_{4\ell}$ as shown in Figure 8.5. Distributions of $\langle m_{\ell\ell} \rangle$ are shown in Figure 8.6.

In terms of overall yield, 460 ± 25 (stat. $\oplus$ syst.) events are predicted and 472 events are observed. In terms of $\langle m_{\ell\ell} \rangle$ distribution, all bins have prediction within 1σ. Both the overall yield and shape show acceptable agreement.

8.1.5 Validation region 4

Validation regions 4 and 5 reduce the overlap between the validation and signal regions.[3] Validation region 4 inverts the final signal region requirement and replaces all four dilepton mass requirements with $m_{12,34,14,32} < 55$ GeV. By lowering the dilepton mass requirements, the overlap with the ZZ_d signal region is reduced from 14 GeV to 5 GeV in terms of $m_{12,34}$.[4]

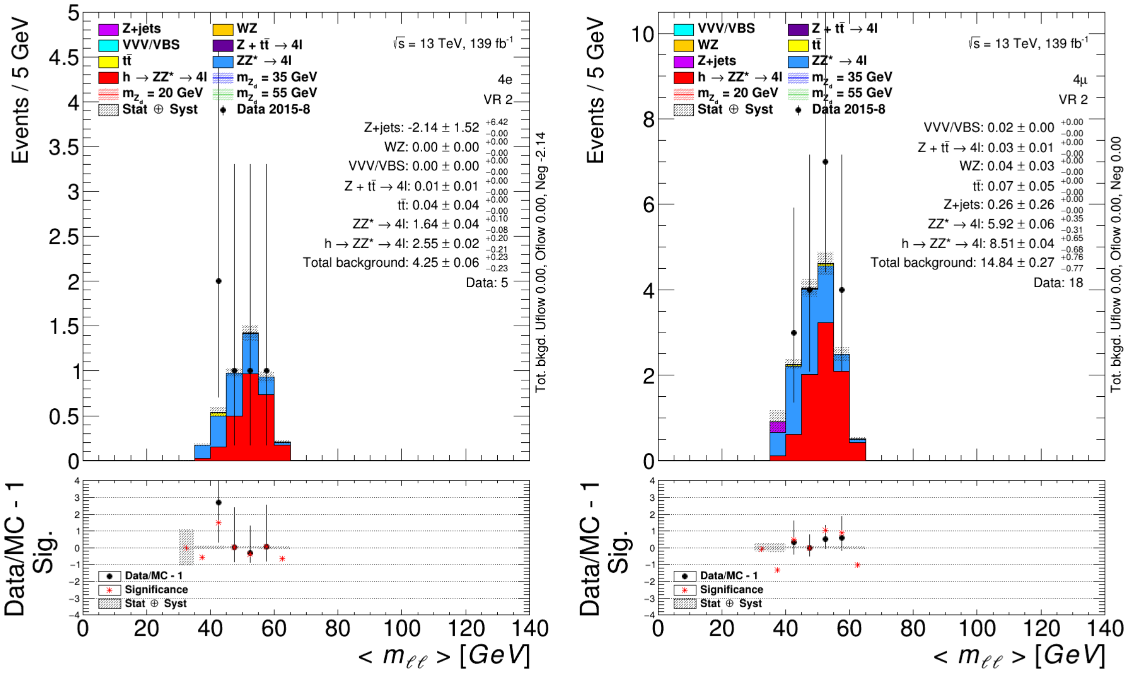
Validation region 4 is closer to $h \rightarrow Z_d Z_d \rightarrow 4\ell$ kinematic region than validation regions 1 and 2. Validation regions 1 and 2 examine the background prediction around m_Z, while validation region 4 examines the background prediction below m_Z and in the mass range of the Z_d hypothesis. In this respect, it is a more relevant validation region for the signal region of interest than validation regions 1 and 2.

Figure 8.7 shows distributions of $\langle m_{\ell\ell} \rangle$ in this region. Due to insufficient statistics, no statement can be made on compatibility in terms of distribution shape. The

---

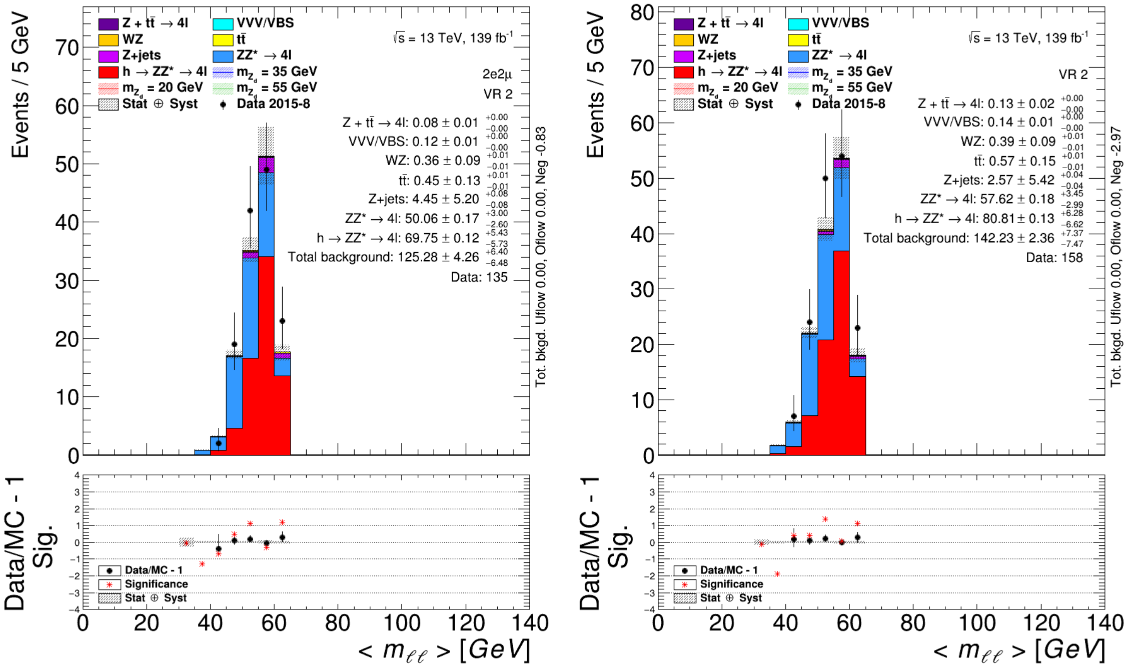
[3] The signal region of $h \rightarrow ZZ_d \rightarrow 4\ell$ overlaps with validation regions 1 and 2, and the signal region of $S \rightarrow XX \rightarrow 4\ell$ overlaps with validation region 3. In order to admit a Z boson, the ZZ_d signal region admits m_{12} between 50 GeV and 106 GeV and m_{34} between 12 GeV and 115 GeV. Validation region 1 can admit an on-shell Z boson reconstructed as $\ell_{14,32}$, and validation region 2 can also admit an on-shell Z boson reconstructed as ℓ_{12}. The other pairings can admit Z_d. Consequently, it is possible for ZZ_d signal to leak into validation regions 1 and 2. In order to admit $S > m_h$, the Higgs window is removed. Validation region 3 reverses the Higgs window and can therefore admit S signal.

[4] The leakage of ZZ_d signal into validation region 4 was assessed with signal simulation and found to be negligible.



(a) $4e$

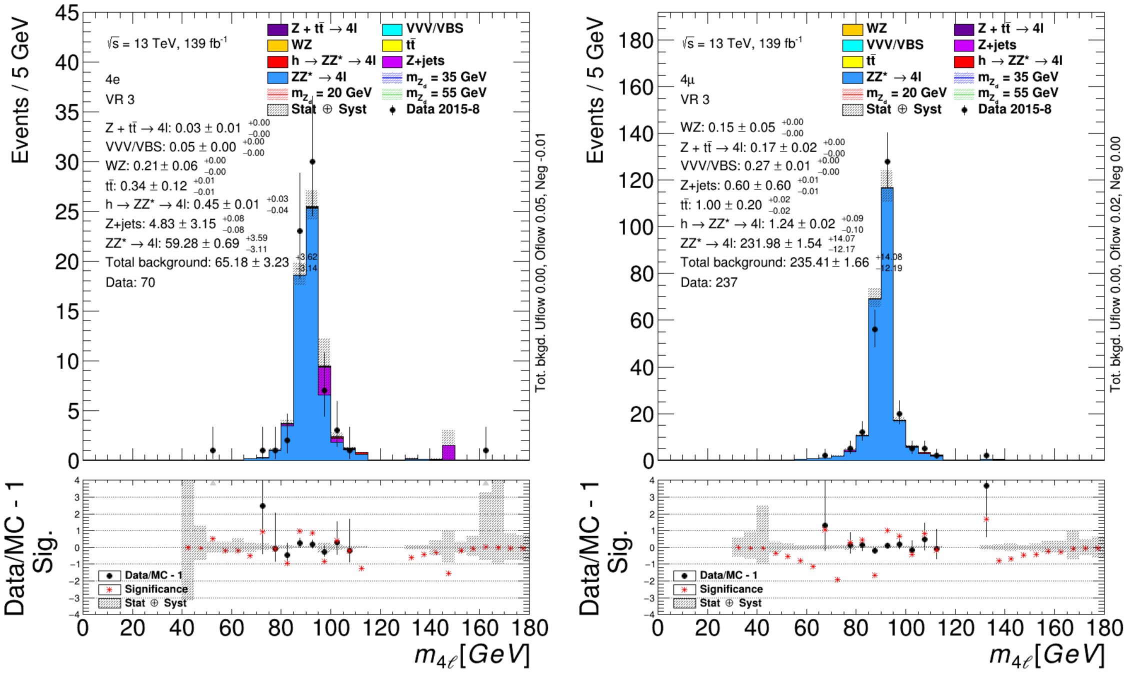
(b) 4μ



(c) $2e2\mu$

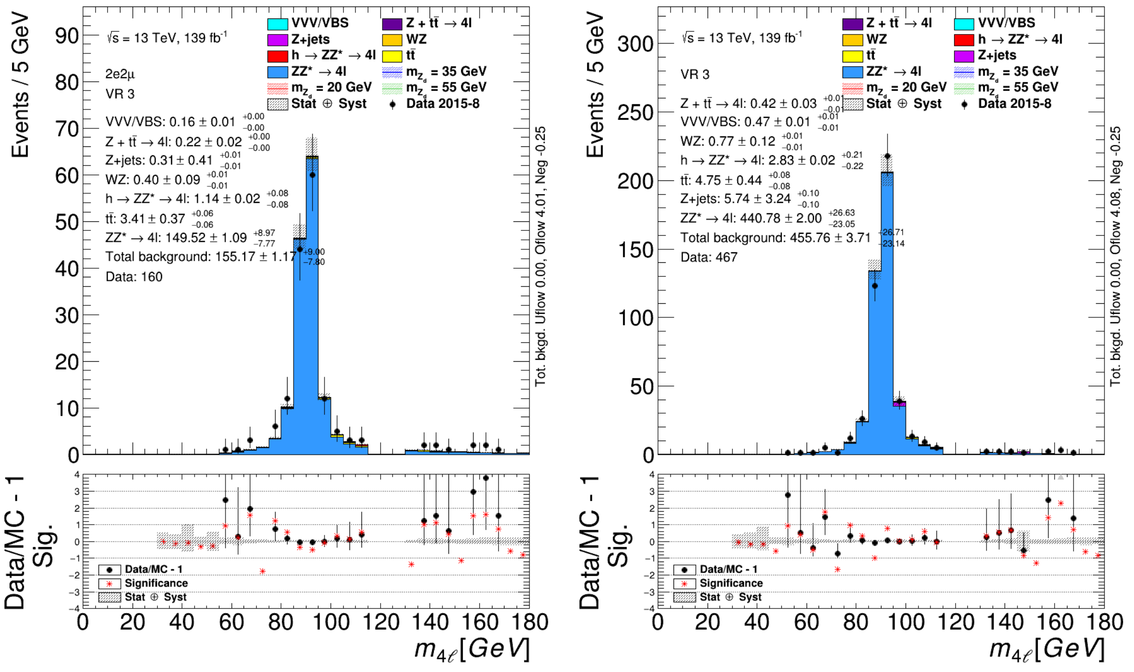
(d) All channels

Figure 8.4: Distributions of $\langle m_{\ell\ell} \rangle$ in validation region 2 in 2015–2018 data and simulation for the $4e$, 4μ, and $2e2\mu$ channels, and all channels combined.



(a) $4e$

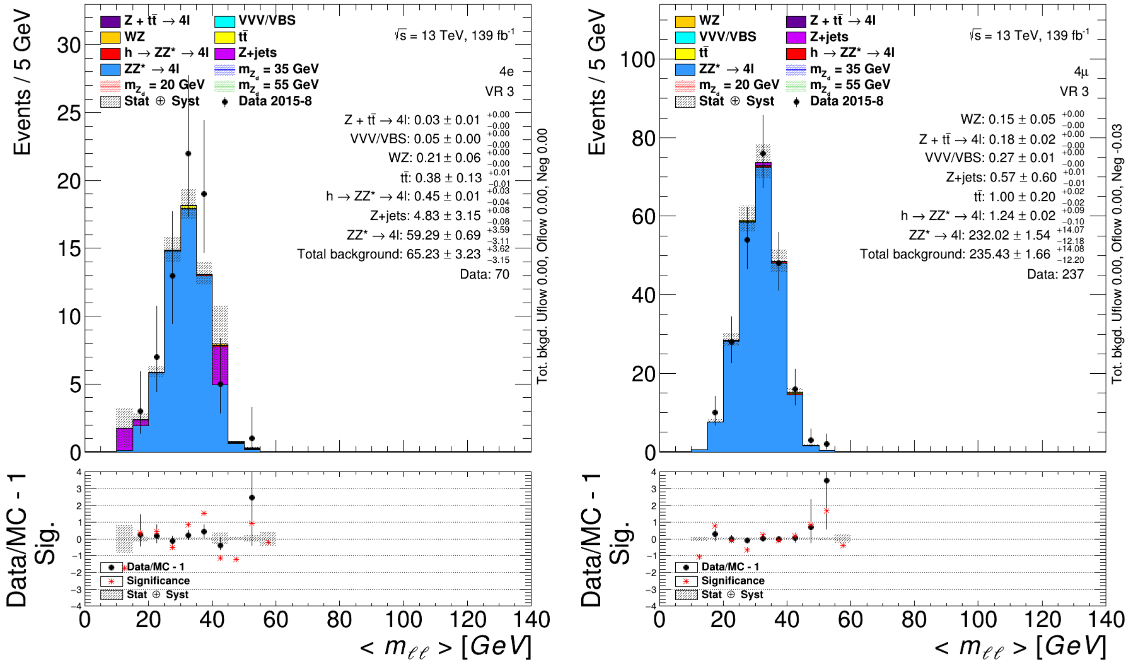
(b) 4μ



(c) $2e2\mu$

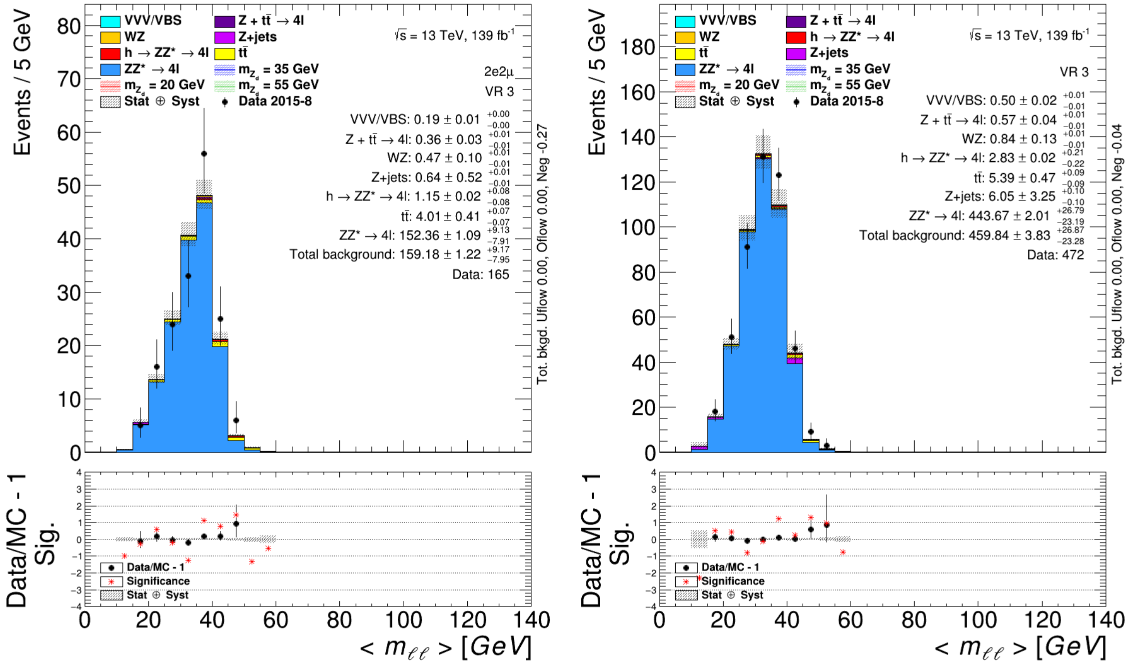
(d) All channels

Figure 8.5: Distributions of $m_{4\ell}$ in validation region 3 in 2015–2018 data and simulation for the $4e$, 4μ, and $2e2\mu$ channels, and all channels combined.



(a) $4e$

(b) 4μ



(c) $2e2\mu$

(d) All channels

Figure 8.6: Distributions of $\langle m_{\ell\ell} \rangle$ in validation region 3 in 2015–2018 data and simulation for the $4e$, 4μ, and $2e2\mu$ channels, and all channels combined.

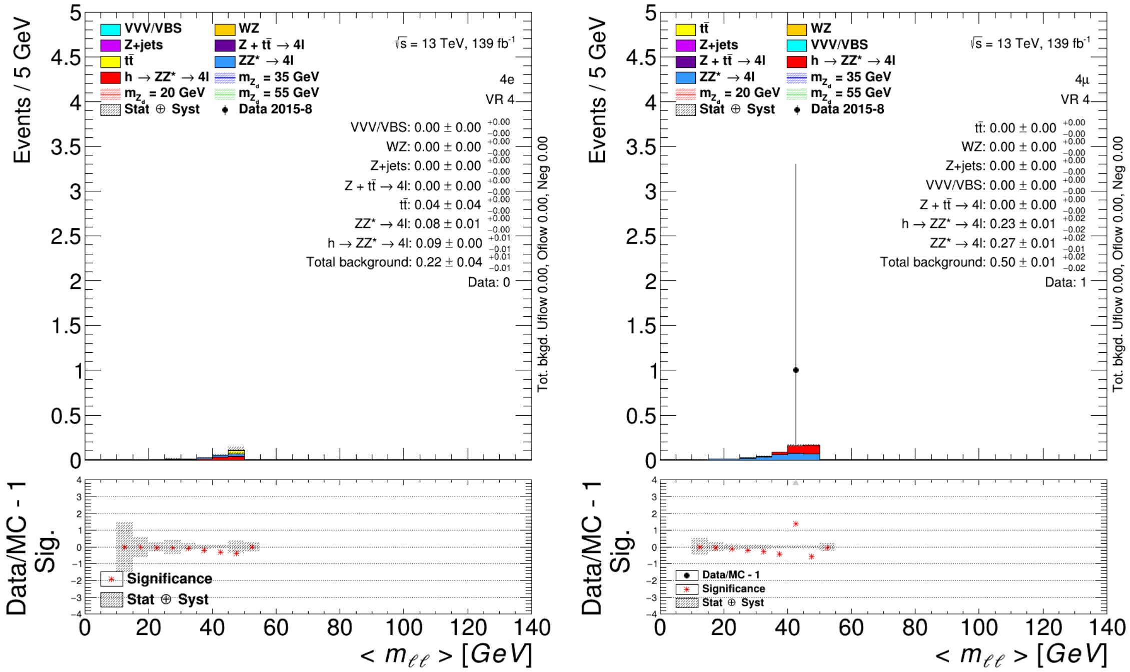
compatibility in terms of overall yield is acceptable: 6.7 ± 0.5 (stat. $\oplus$ syst.) events are predicted and 6 events are observed.

8.1.6 Validation region 5

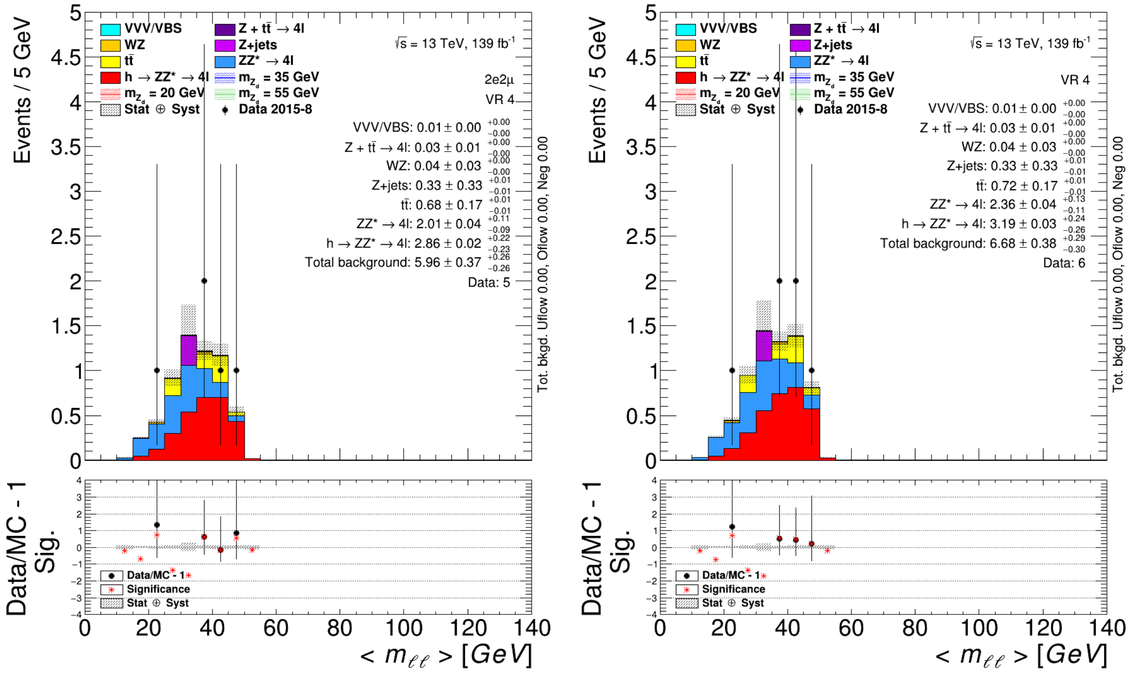
In this validation region, the side-bands of the ZZ^* prediction around the Higgs mass are examined by inverting the Higgs window requirement. The kinematic region is also restricted to 100 GeV $< m_{4\ell} <$ 170 GeV and the dilepton mass requirements (Z veto) and final signal region requirements are removed. The removal of the Z veto increases the amount of Z+jets background considerably. Monte Carlo-based simulation predicted Z+jets contributing approximately 8% of the total background in this region. The agreement in terms of overall yield is marginal: 411 ± 21 (stat. $\oplus$ syst.) events are predicted and 486 are observed. Since validation region 3 demonstrates acceptable agreement of the ZZ^* prediction with data, the marginal agreement in validation region 5 is attributed to the simulation-based Z+jets estimate.[5] Distributions of $\langle m_{\ell\ell} \rangle$ and $m_{4\ell}$ in this region are shown in Figures 8.8 and 8.9, respectively.

---

[5] The data-driven estimate is available only for the signal region and is unbinned.

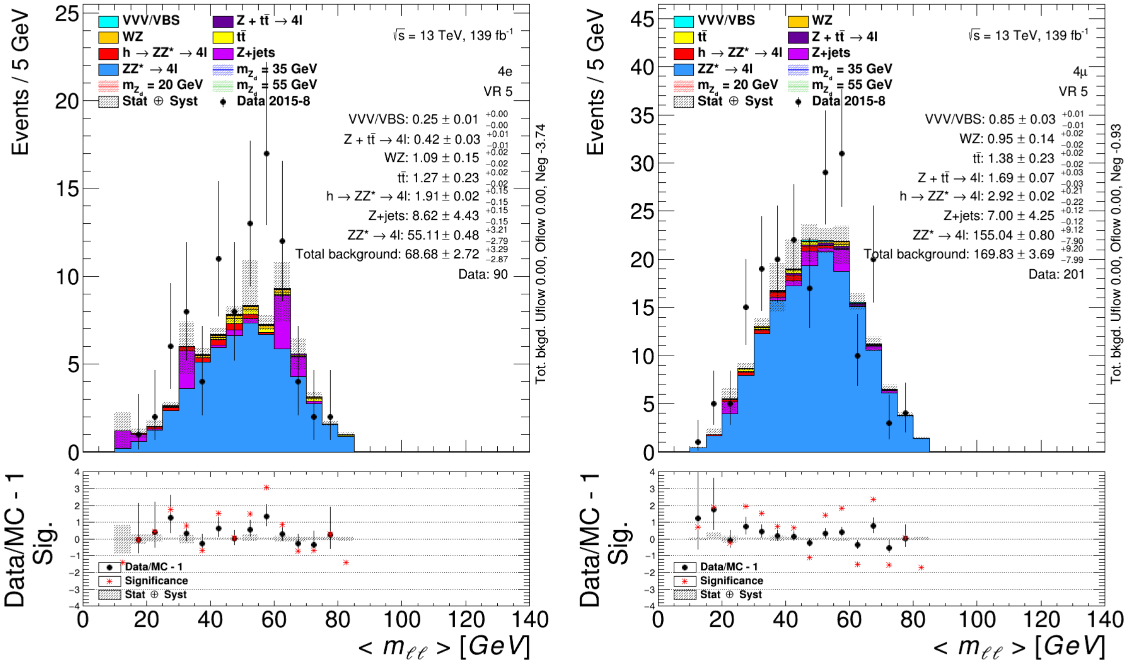


(a) $4e$ (b) 4μ



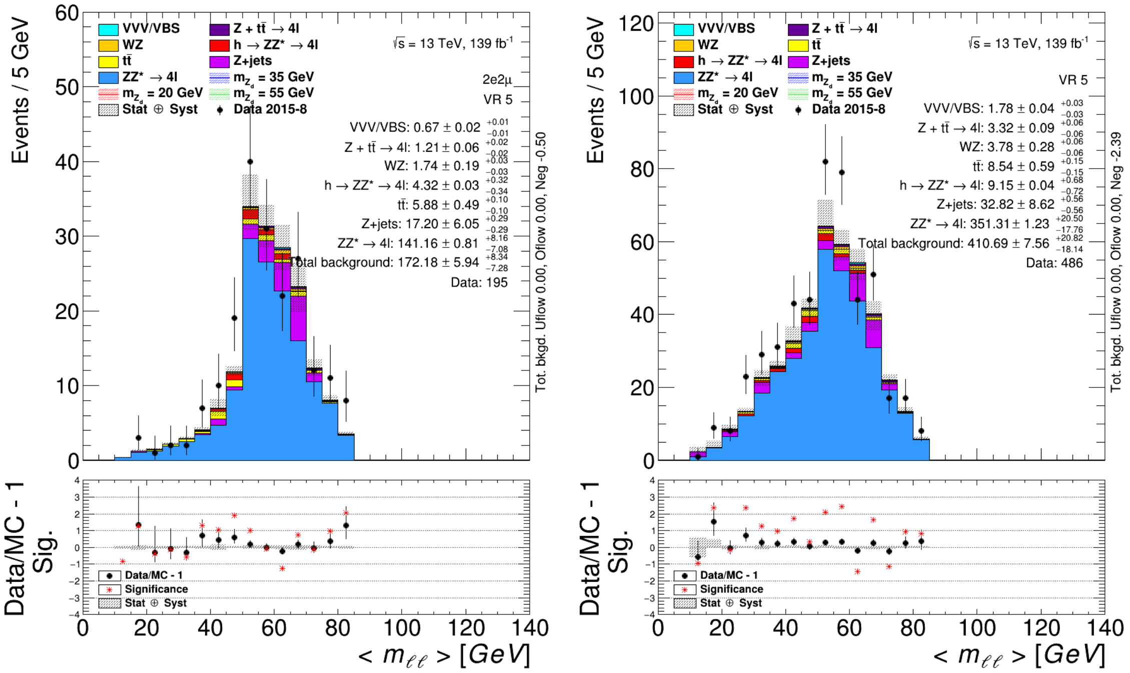
(c) $2e2\mu$ (d) All channels

Figure 8.7: Distributions of $\langle m_{\ell\ell}\rangle$ in validation region 4 in 2015–2018 data and simulation for the $4e$, 4μ, and $2e2\mu$ channels, and all channels combined.



(a) 4e

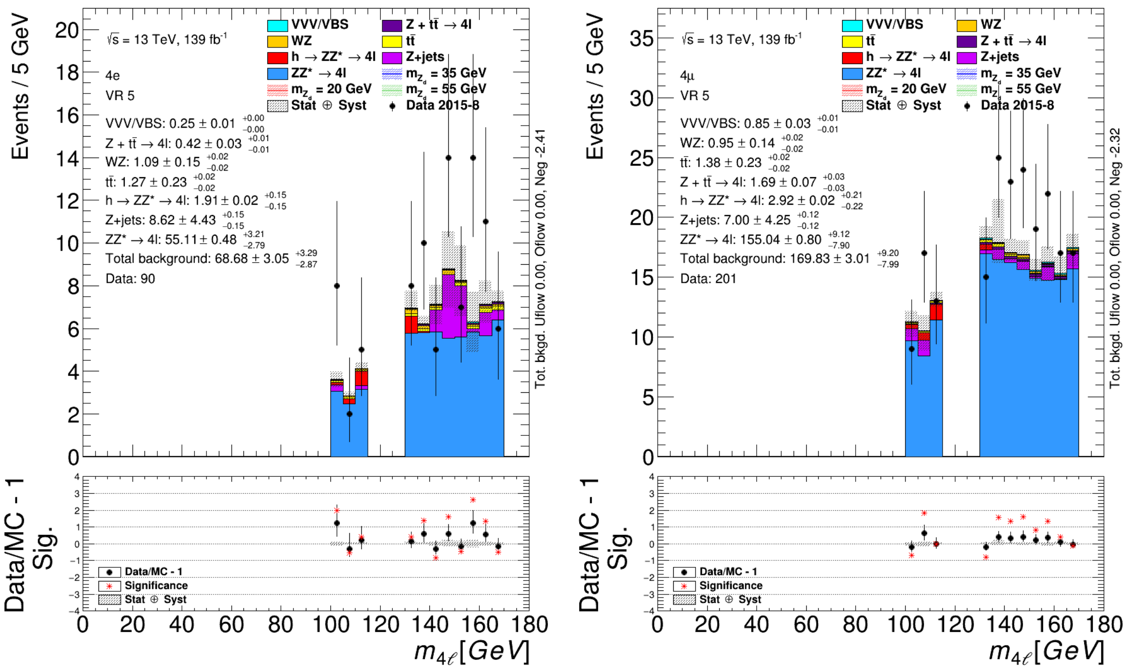
(b) 4μ



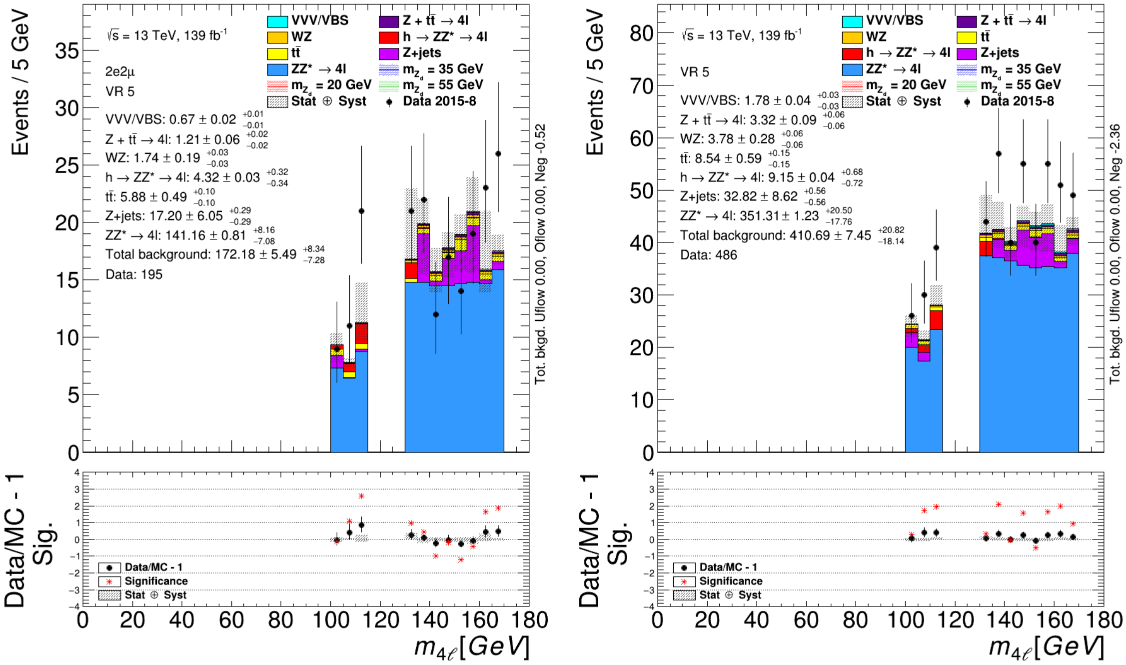
(c) $2e2\mu$

(d) All channels

Figure 8.8: Distributions of $\langle m_{\ell\ell} \rangle$ in validation region 5 in 2015–2018 data and simulation for the $4e$, 4μ, and $2e2\mu$ channels, and all channels combined.



(a) $4e$ (b) 4μ



(c) $2e2\mu$ (d) All channels

Figure 8.9: Distributions of $m_{4\ell}$ in validation region 5 in 2015–2018 data and simulation for the $4e$, 4μ, and $2e2\mu$ channels, and all channels combined.

8.2 Unblinding of signal region

In the signal region, 20 events are observed and $15.6 \pm 0.4 \pm 1.2$ are predicted in background. In terms of overall yield, the observed excess is not statistically significant. The yields per channel and data-taking campaign, and per background are summarized in Table 8.3 and Table 8.4 respectively, and distributions of $\langle m_{\ell\ell} \rangle$ in the signal region are shown in Figures 8.10 and 8.11.[6] Of the twenty events observed, three events are outside the range considered for setting limits (15 GeV $< \langle m_{\ell\ell} \rangle < 60$ GeV): one event in the $2e2\mu$ channel with $\langle m_{\ell\ell} \rangle \approx 13.5$ GeV and two events in the 4μ channel with $\langle m_{\ell\ell} \rangle \approx 60.2$ GeV and $\langle m_{\ell\ell} \rangle \approx 60.04$ GeV. They are included in the distributions shown in Figures 8.10 and 8.11.

The reduction in the number of observed events from imposing the signal region selections described in Chapter 5 is shown in Table 8.5.

The results do not show evidence for the $h \to Z_d Z_d \to 4\ell$ process.

The m_{34} vs. m_{12} distribution of observed events passing all selection requirements is shown in Figure 8.12. Each marker corresponds to an event observed in the signal region and the shaded area represents the kinematic region accepted by the final signal region requirement on m_{34} and m_{12}. Figure 8.13 shows the same events as well as events failing the Z veto (crossed-through markers) and the final signal region requirement (markers outside the shaded area).[7] In the latter figure, the shaded area is split into two: the green area corresponds to the kinematic region accepted by the *original* final signal region requirement $m_{34}/m_{12} > 0.85$ and the red area corresponds to the additional kinematic region accepted by the *re-optimized* final signal requirement. As evident in the signal (Table 6.1) and background (Tables 6.6 and 6.7) cutflow tables, the Z veto has a significant impact on signal efficiency, but it is applied in order to suppress contributions from background processes with Z bosons.

---

[6] The events observed in the search with 2015–2016 data only are recovered.

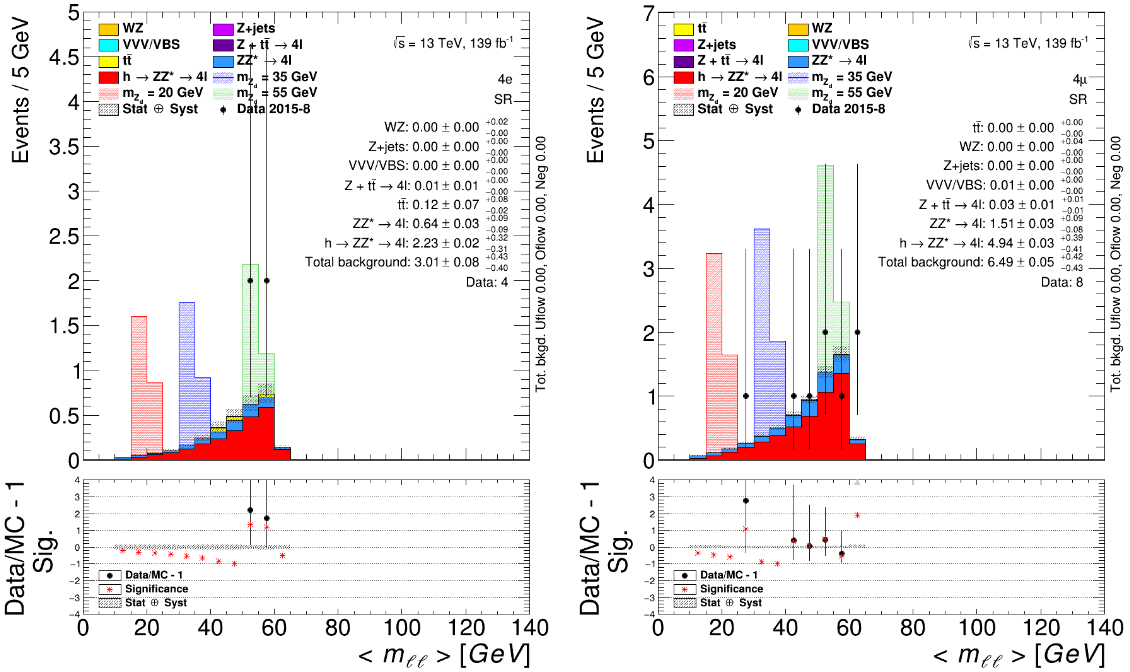
[7] All events in both distributions passed requirements up to and including the Higgs window requirement, as tabulated in Table 8.5. 336 events are observed up to this point, of which 170 are in the m_{12} and m_{34} ranges plotted in Figure 8.13.

Channel	Signal region in 2015–2016 (36.2 fb^{-1})		Signal region in 2017 (44.3 fb^{-1})		Signal region in 2018 (58.4 fb^{-1})		Signal region in 2015–2018 (139 fb^{-1})	
	Expected background (predicted from simulation)	Observed	Expected background (predicted from simulation)	Observed	Expected background (predicted from simulation)	Observed	Expected background (predicted from simulation)	Observed
$4e$	$0.88 \pm 0.06\ ^{+0.20}_{-0.18}$	1	$0.88 \pm 0.02\ ^{+0.09}_{-0.09}$	0	$1.25 \pm 0.05\ ^{+0.15}_{-0.14}$	3	$3.01 \pm 0.08\ ^{+0.43}_{-0.40}$	4
4μ	$1.76 \pm 0.02\ ^{+0.12}_{-0.12}$	2	$2.05 \pm 0.02\ ^{+0.14}_{-0.14}$	3	$2.68 \pm 0.03\ ^{+0.17}_{-0.18}$	2	$6.49 \pm 0.05\ ^{+0.42}_{-0.43}$	8
$2e2\mu$	$1.60 \pm 0.08\ ^{+0.19}_{-0.19}$	3	$1.75 \pm 0.07\ ^{+0.12}_{-0.13}$	2	$2.29 \pm 0.07\ ^{+0.17}_{-0.17}$	3	$5.64 \pm 0.13\ ^{+0.47}_{-0.48}$	8
Combined	$4.24 \pm 0.10\ ^{+0.44}_{-0.43}$	6	$4.68 \pm 0.08\ ^{+0.32}_{-0.33}$	5	$6.22 \pm 0.09\ ^{+0.44}_{-0.45}$	8	$15.13 \pm 0.16\ ^{+1.17}_{-1.18}$	20

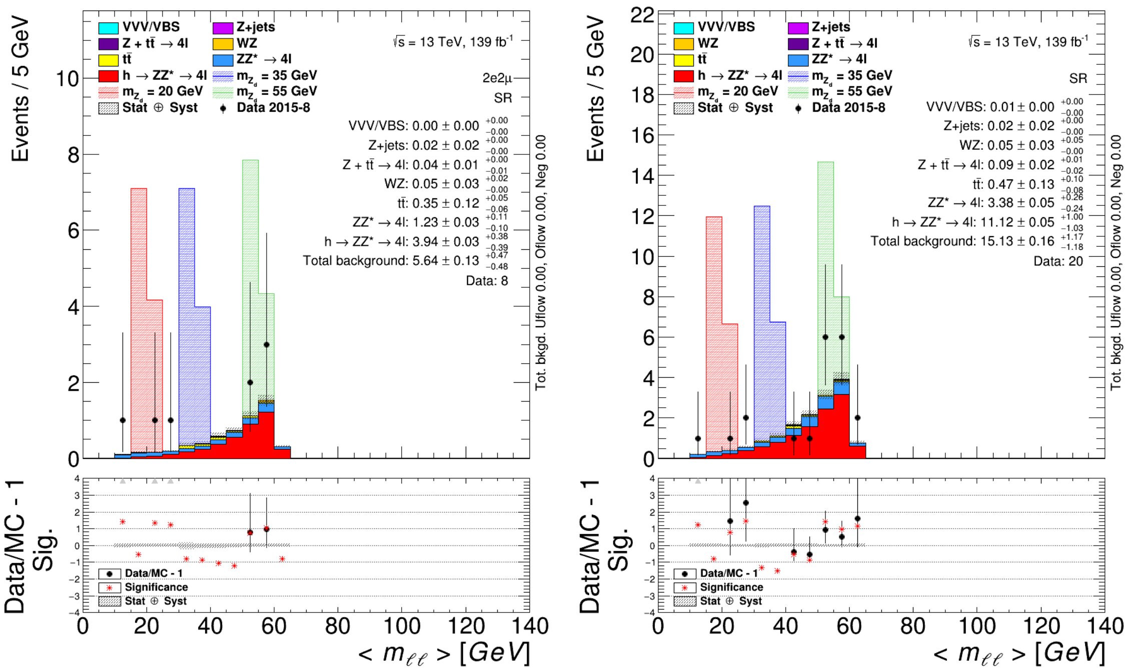
Table 8.3: Expected background yield *predicted from simulation* and observed yield in the signal region per data-taking campaign. For the Z+jets process, estimates from simulation are used for the total expected background tabulated above.

Background process	Prediction in signal region
$h \to ZZ^* \to 4\ell$	$11.12 \pm 0.05\ ^{+1.00}_{-1.03}$
$ZZ^* \to 4\ell$	$3.38 \pm 0.05\ ^{+0.26}_{-0.24}$
$t\bar{t}$	$0.47 \pm 0.13\ ^{+0.10}_{-0.08}$
Z+jets	$0.43 \pm 0.39\ ^{+0.17}_{-0.01}$ (from data-driven method)
$Z + t\bar{t} \to 4\ell$	$0.09 \pm 0.02\ ^{+0.01}_{-0.02}$
WZ	$0.05 \pm 0.03\ ^{+0.05}_{-0.00}$
VVV/VBS	$0.01 \pm 0.00\ ^{+0.00}_{-0.00}$
Combined	$15.6 \pm 0.4 \pm 1.2$
Observed	20

Table 8.4: Expected background yield per process and observed yield in the signal region in 2015–2018 (139 fb^{-1}). The uncertainties listed are statistical and systematic, respectively. For the Z+jets background process, the estimate from simulation is $0.02 \pm 0.02 \pm 0.00$ and the estimate from the data-driven method (Chapter 7) is $0.43 \pm 0.39^{+0.17}_{-0.01}$. The estimates for all other processes are from simulation.



(a) $4e$ (b) 4μ



(c) $2e2\mu$ (d) All channels

Figure 8.10: Distributions of $\langle m_{\ell\ell} \rangle$ in the signal region in 2015–2018 data and simulation for the $4e$, 4μ, and $2e2\mu$ channels, and all channels combined. The bins of the distributions are 5 GeV wide. The example signal distributions correspond to the expected yield normalized with $\sigma(gg \to h \to Z_d Z_d \to 4\ell) = \frac{1}{10}\sigma_{\text{SM}}(gg \to h \to ZZ^* \to 4\ell)$. The estimate from Z+jets shown in the distribution and written in the inset is from simulation. The data-driven Z+jets estimate is $0.43 \pm 0.39^{+0.17}_{-0.01}$ (detailed in Chapter 7). With the data-driven estimate, the total background prediction is $15.6 \pm 0.4 \pm 0.5$.

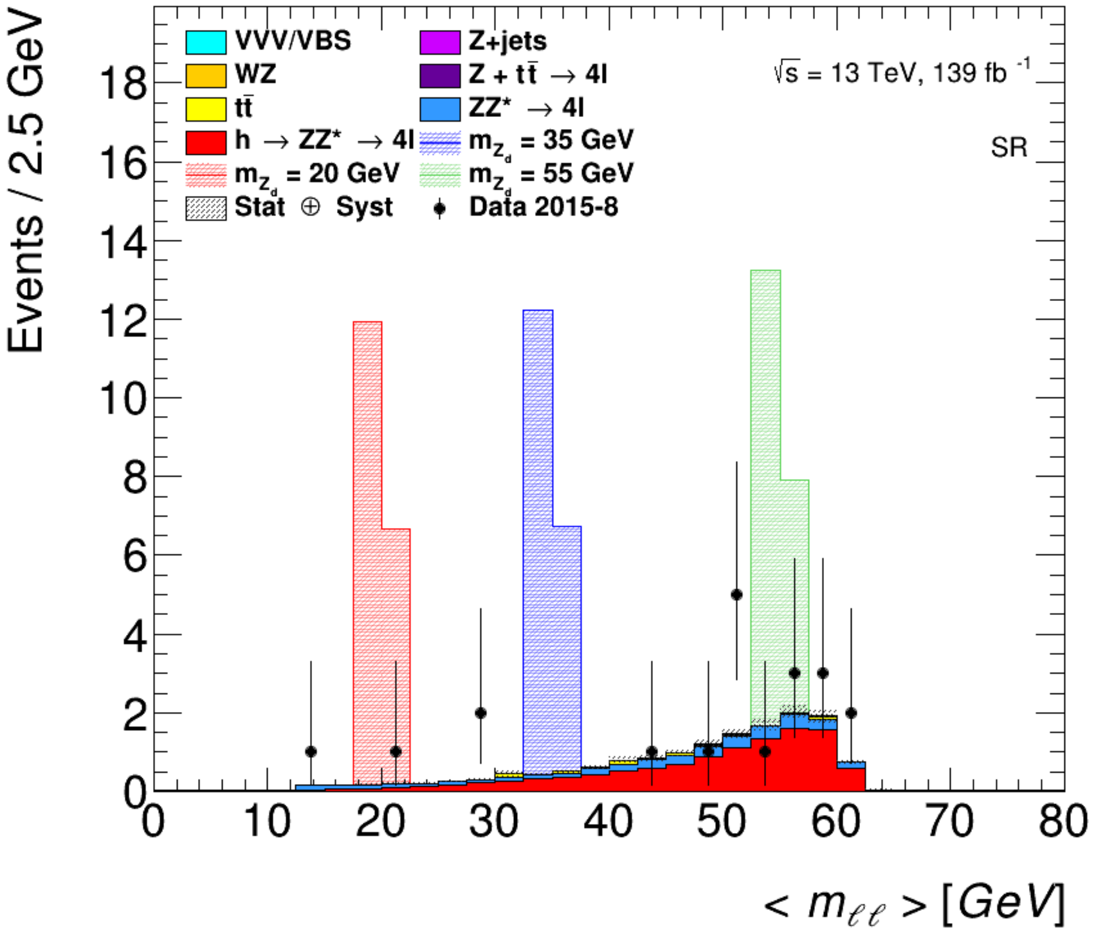


Figure 8.11: Distribution of $\langle m_{\ell\ell}\rangle$ in the signal region in 2015–2018 data and simulation with all channels combined. The bins of the distribution are 2.5 GeV wide. The example signal distributions correspond to the expected yield normalized with $\sigma(gg \rightarrow h \rightarrow Z_d Z_d \rightarrow 4\ell) = \frac{1}{10}\sigma_{\text{SM}}(gg \rightarrow h \rightarrow ZZ^* \rightarrow 4\ell)$. The estimate from Z+jets shown in the distribution is from simulation.

Requirement	Number of events							
Number of events in reprocessed AOD (Analysis Object Data)	$\sim 1.8 \cdot 10^{10}$							
Number of events in HIGG2D1 DAOD (Derived AOD, dilepton filter)	$\sim 3.4 \cdot 10^{9}$							
Event cleaning (GRL, primary vertex, detector flags) & At least one lowest-unprescaled trigger fired	$\sim 1.5 \cdot 10^{8}$							
	$4e$		$2e2\mu$		4μ		All	
	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}	N_{SFOS}	N/N_{SFOS}
Same-flavour opposite sign quadruplet	$8.505 \cdot 10^{6}$	100 %	$7.685 \cdot 10^{7}$	100 %	$2.485 \cdot 10^{5}$	100 %	$8.560 \cdot 10^{7}$	100 %
Non-overlapping quadruplet leptons	$8.491 \cdot 10^{6}$	99.8 %	$7.679 \cdot 10^{7}$	99.9 %	$2.486 \cdot 10^{5}$	100 %	$8.553 \cdot 10^{7}$	99.9 %
$p_T(\ell_1) > 20$, $p_T(\ell_2) > 15$, $p_T(\ell_3) > 10$ GeV	$6.404 \cdot 10^{6}$	75.3 %	$5.366 \cdot 10^{7}$	69.8 %	$8.970 \cdot 10^{4}$	36.1 %	$6.015 \cdot 10^{7}$	70.3 %
Leptons in quadruplet fired at least one trigger	$6.334 \cdot 10^{6}$	74.5 %	$5.323 \cdot 10^{7}$	69.3 %	$9.143 \cdot 10^{4}$	36.8 %	$5.965 \cdot 10^{7}$	69.7 %
$\Delta R > 0.1\,(0.2)$ for same (opposite) flavour leptons	$4.918 \cdot 10^{6}$	57.8 %	$2.060 \cdot 10^{7}$	26.8 %	$5.415 \cdot 10^{4}$	21.8 %	$2.557 \cdot 10^{7}$	29.9 %
Not more than one calo-tagged or stand-alone muon	$4.918 \cdot 10^{6}$	57.8 %	$2.060 \cdot 10^{7}$	26.8 %	$5.406 \cdot 10^{4}$	21.8 %	$2.557 \cdot 10^{7}$	29.9 %
Isolation (Electrons: FCLoose) (Muons: FCLoose_FixedRad)	$1.329 \cdot 10^{5}$	1.6 %	$1.818 \cdot 10^{5}$	$2.4 \cdot 10^{-1}$ %	4682	1.9 %	$3.193 \cdot 10^{5}$	$3.7 \cdot 10^{-1}$ %
Electron ID (all quadruplet leptons Loose or better)	1516	$1.8 \cdot 10^{-2}$ %	4954	$6.4 \cdot 10^{-3}$ %	4682	1.9 %	$1.115 \cdot 10^{4}$	$1.3 \cdot 10^{-2}$ %
Impact parameter (Electrons: $z_0 \sin\theta < 0.5$ mm) (Muons: $z_0 \sin\theta < 0.5$ mm, $d_0 < 1$ mm, d_0 significance ≤ 3)	1308	$1.5 \cdot 10^{-2}$ %	3810	$5.0 \cdot 10^{-3}$ %	3239	1.3 %	8357	$9.8 \cdot 10^{-3}$ %
Quarkonia veto	1163	$1.4 \cdot 10^{-2}$ %	3453	$4.5 \cdot 10^{-3}$ %	2648	1.1 %	7264	$8.5 \cdot 10^{-3}$ %
Low mass veto (require $m_{\ell\ell} > 5$ GeV)	1046	$1.2 \cdot 10^{-2}$ %	2974	$3.9 \cdot 10^{-3}$ %	2135	$8.6 \cdot 10^{-1}$ %	6155	$7.2 \cdot 10^{-3}$ %
Higgs window ($115 < m_{4\ell} < 130$ GeV)	48	$5.6 \cdot 10^{-4}$ %	153	$2.0 \cdot 10^{-4}$ %	135	$5.4 \cdot 10^{-2}$ %	336	$3.9 \cdot 10^{-4}$ %
Z veto (require $m_{12,34} < 64$ and $m_{32,14} < 75$ GeV)	10	$1.2 \cdot 10^{-4}$ %	18	$2.3 \cdot 10^{-5}$ %	25	$1.0 \cdot 10^{-2}$ %	53	$6.2 \cdot 10^{-5}$ %
Loose signal region ($m_{12,34} > 10$ GeV)	10	$1.2 \cdot 10^{-4}$ %	16	$2.1 \cdot 10^{-5}$ %	23	$9.3 \cdot 10^{-3}$ %	49	$5.7 \cdot 10^{-5}$ %
Medium signal region (re-optimized for wider width) ($m_{34} > 0.85 m_{12} - 0.1125 f(m_{12}) m_{12}$)	4	$4.7 \cdot 10^{-5}$ %	8	$1.0 \cdot 10^{-5}$ %	8	$3.2 \cdot 10^{-3}$ %	20	$2.3 \cdot 10^{-5}$ %

Table 8.5: Signal region cutflow for 2015–2018 data combined (36.2 fb^{-1} for 2015–2016, 44.3 fb^{-1} for 2017, and 59.9 fb^{-1} for 2018). Per-channel, the quadruplet cutflow can increase because events can move from one channel to another during the quadruplet selection if a higher-ranked quadruplet fails a cut that a lower-ranked one can pass (however, the cutflow summed across all channels can only decrease).

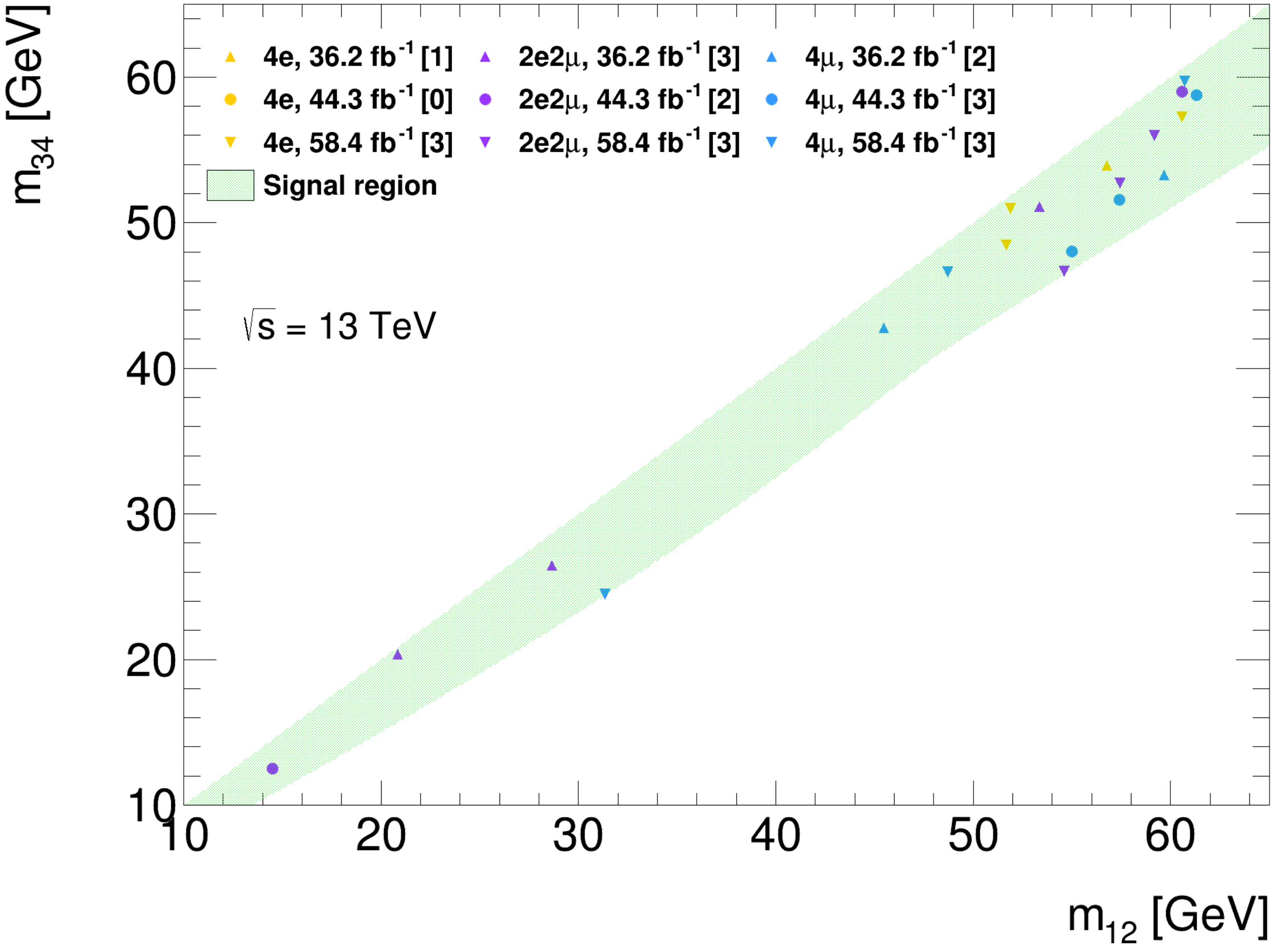


Figure 8.12: m_{34} vs. m_{12} distribution of events observed in the signal region. Each marker corresponds to an event that passed all selection requirements. The markers and events are differentiated by channel and by data-taking campaign (2015–2016, 2017, and 2018), with the number of events in the signal region per channel and data-taking campaign written within the square brackets.

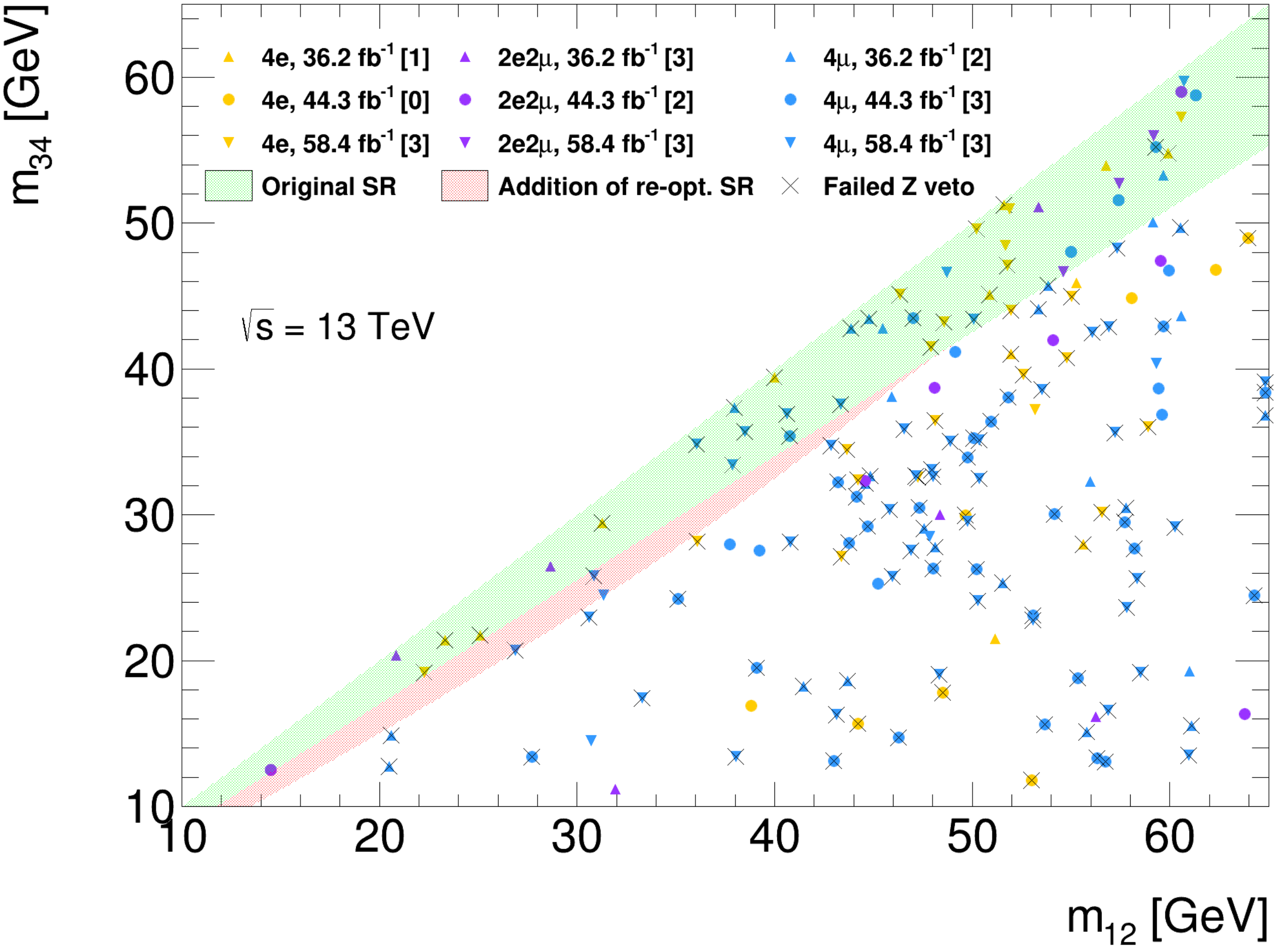


Figure 8.13: m_{34} vs. m_{12} distribution of events observed passing requirements up to and including the Higgs window requirement. The crossed-through markers correspond to events that fail the Z veto. The events inside (outside) the shaded area are the events that pass (fail) the final signal region requirement. The markers and events are differentiated by channel and by data-taking campaign (2015–2016, 2017, and 2018), with the number of events in the signal region per channel and data-taking campaign written within the square brackets.

Chapter 9

Interpretation

The results shown in the previous section (Section 8.2) do not suggest that the $h \to Z_d Z_d \to 4\ell$ process exists. In terms of overall yield, the excess in observed data over background prediction has significance of less than 1σ. In terms of the $\langle m_{\ell\ell} \rangle$ distribution, the bin with the largest excess has approximately 2σ significance.[1] Neither excesses are significant enough to declare evidence or discovery. A more accurate calculation of the p-values and corresponding significances per bin in $\langle m_{\ell\ell} \rangle$ is presented in this chapter. In the absence of significant excesses, the results are interpreted in terms of upper limits on the branching ratio and mixing parameter of the model described in Chapter 3 and simulated with the approach described in Section 6.2.1.

The first portion of this chapter describes the statistical model and methods used in this search. The second portion describes the implementation and presents the interpretation of the results.

9.1 Methodology

9.1.1 Hypothesis testing and p-value

In the *frequentist* approach, probabilities are defined as the frequency of a certain outcome out of an infinite number of identical trials. Parameters are thought to have true (fixed) values that are unknown and cannot be treated as random variables

---

[1] Local significance estimated with a formula that is only a function of observed yield, expected yield, and uncertainty on expected yield. This chapter will present an approach to calculating the p-values and corresponding significances that incorporates the complete statistical model.

that can be modelled; therefore, probabilities are not assigned to parameters having certain values. This differs from the *Bayesian* approach, where *prior probabilities* for parameters are defined before an experiment (e.g. as a probability distribution function) and updated with experimental data to give *posterior probabilities.* The posterior probability distribution function (PDF) can then be used to determine the range of parameter values that have a certain probability to contain the true value of the parameter. Notably, the Bayesian approach does not concern itself with multiple trials. In the frequentist approach, a measurement is thought to be repeated an infinite number of trials, with each trial yielding a different *frequentist interval.* The boundaries of each interval are determined from each observation of the parameter with the *Neyman* procedure. The fraction of intervals that encompass the true value of the parameter is the frequentist probability. One advantage of this approach is that no prior probability is required. The interpretation of the search results presented in this chapter takes the frequentist approach.

In statistical terms, the primary objective of this search is to *test* for the absence of signal. Two hypotheses are considered: the *background-only* hypothesis and the *signal-plus-background* hypothesis. The former is taken as the *null hypothesis* H_0 we wish to test and hopefully reject. If the null hypothesis is rejected, then the alternative (signal-plus-background) hypothesis H_1 is accepted. The hypothesis PDFs $f(\mathbf{x} \mid H)$ depend intrinsically on the measurement values $\mathbf{x}$ and are usually obtained using Monte Carlo methods. The null hypothesis is tested by checking whether or not the observed data belongs to a *critical region* in $\mathbf{x}$ that corresponds to a small probability under $f(\mathbf{x} \mid H_0)$ and a non-negligible probability under $f(\mathbf{x} \mid H_1)$. If the observed data enters the critical region, then the null hypothesis is rejected; otherwise, it belongs to the bulk of the null hypothesis PDF and therefore the null hypothesis is accepted. The critical region should be chosen such that the probability of correctly accepting the true hypothesis is maximized and the probabilities for accepting the null hypothesis when the alternative hypothesis is true (*type-II error*) and vice versa (*type-I error*) are minimized. The probability under the null hypothesis that corresponds to the critical region, α, is known as the *significance level* of the test:

$$\alpha = \int_{t_{\mathrm{crit}}}^{\infty} f(t \mid H_0) \tag{9.1}$$

where the measured values $\mathbf{x}$ have been transformed into a scalar function $t(\mathbf{x})$ known as the *test statistic.* Once the significance level α has been chosen, then the test is to compare α against the probability corresponding to $t(\mathbf{x}_{\text{obs}})$ under the null hypothesis:

$$p = \int_{t_{\text{obs}}}^{\infty} f(t \mid H_0) \tag{9.2}$$

Typically, the significance level required for discovery is $p \leq 2.87 \times 10^{-7}$. In terms of a standard Gaussian, this p-value corresponds to the probability that is found in one tail that is five standard deviations from the mean ("5σ"). The *significance* in units of standard Gaussian standard deviation is $Z(p) = \Phi^{-1}(1-p)$, where Φ^{-1} is the inverse cumulative distribution function of a standard Gaussian.

9.1.2 Likelihood construction

Typically, the hypotheses are modelled as likelihoods with Poisson distributions representing the signal and background processes (commonly referred to as an *extended likelihood*). The likelihoods are functions of the parameter of interest (e.g. signal strength μ) and nuisance parameters $\boldsymbol{\theta}$. The nuisance parameters are additional parameters that have an effect on the parameter of interest. They are needed to fully define the experimental model and are measured in *auxiliary measurements*. For example, knowledge of electron identification efficiency is a required part of the model and is represented in the likelihood by a nuisance parameter. The efficiencies are measured in dedicated tag-and-probe studies of $Z \to ee$ events (the auxiliary measurements) and are propagated to the electrons measured in the signal region through a nominal efficiency correction and $\pm 1\sigma$ systematic variations of the correction.[2] The nuisance parameters themselves are assumed to follow a Gaussian distribution, and are therefore incorporated into the likelihood by Gaussian *constraint terms*. The integrated luminosity is another example of a systematic uncertainty that is parametrized by a nuisance parameter.

A *binned extended likelihood* has the general form:

$$L(\mu, \boldsymbol{\theta}) = \prod_{i}^{n_{\text{bins}}} e^{-(\mu s_i + b_i)} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} \prod_{j}^{n_{\text{systs}}} C(\theta_{i,j}) \tag{9.3}$$

- Signal strength μ (parameter of interest): $\mu = 0$ corresponds to the background-

---

[2] Resolution and energy scale are corrected to their nominal values and adjusted to their systematic variations by changing the lepton kinematics. Efficiencies are corrected and adjusted by re-weighting the events with *scale factors*.

only hypothesis and $\mu > 0$ corresponds to the signal-plus-background hypothesis,[3]

- Nuisance parameters $\boldsymbol{\theta}$: additional parameters of the model, principally used to parametrize systematic uncertainties,
- Per-bin expected signal and background yields s_i and b_i,[4][5]
- Nuisance parameter constraints $C(\theta_{i,j})$: typically Gaussians with width proportional to the effect of the systematic variation on the yield.[6] Systematics that are correlated between samples are assigned one nuisance parameter for all samples and systematics that are uncorrelated between samples are assigned one nuisance parameter per sample.

9.1.3 Test statistic for discovery

The *Neyman-Pearson* lemma states that the test with minimal type-II error[7] is the likelihood ratio $\lambda(\mathbf{x}) = f(\mathbf{x} \mid H_1)/f(\mathbf{x} \mid H_0)$.[8] The tests used in this search are formed from the *profile likelihood ratio*:[9]

$$\lambda(\mu) = \frac{L\left(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu)\right)}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{9.4}$$

Specific values of the nuisance parameters are chosen through *profiling* of the likelihood.[10] The values of μ and $\boldsymbol{\theta}$ that simultaneously maximize the likelihood are the *unconditional maximum likelihood estimates* and are denoted $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$, respectively. The values of $\boldsymbol{\theta}$ that maximize the likelihood for a given value of μ are the *conditional maximum likelihood estimates* and are denoted $\hat{\hat{\boldsymbol{\theta}}}(\mu)$. $\hat{\mu}$ can be interpreted as the

---

[3] $\langle N \rangle = \mu s + b$, where $\mu = 1$ corresponds to the hypothesized *nominal* signal strength

[4] Signal distributions are obtained from simulation. Background distributions can be obtained from simulation or measured (i.e. in-situ).

[5] Per-bin yields provided *per background* complete the statistical model by providing information about the shape of each background.

[6] These widths are assumed to be known (provided in auxiliary measurements and propagated to the analysis objects). In practice, the nuisance parameter and thus the observable are scaled such that the Gaussian has unit variance, e.g. for nuisance parameter $\theta_j \in \boldsymbol{\theta}$ and yield n of an observable, $G(n \mid \theta_j, 1)$. This is the HistFactory convention [100].

[7] The complement of the probability for type-II error, $1 - \beta$, is the *power* of the test.

[8] This test aims to reject the hypothesis of the numerator H_1 in favour of hypothesis of the denominator H_0.

[9] The likelihood ratio is a function of $\mathbf{x}$ and is therefore also referred to as a test statistic.

[10] Equivalently, the nuisance parameters are *marginalized* through profiling.

observed value of μ. It is the value of μ that is determined from maximizing the likelihood with the observed data.

If $\hat{\mu}$ is assumed to follow a Gaussian distribution, then the ratio can be re-written as [101]:

$$t_\mu = -2\ln\lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}) \tag{9.5}$$

where σ is the Gaussian standard deviation and N is the sample size.[11] In the limit $N \to \infty$, t_μ approaches a *noncentral* χ^2 distribution.[12] The asymptotic distribution of t_μ and monotonic nature of the *log-likelihood* as a function of μ suggest a test statistic of this form.

The p-value of this test statistic quantifies the level of incompatibility between a hypothesized value of μ and the value that is found from profiling the data, $\hat{\mu}$:

$$p_\mu = \int_{t_{\mu,\mu=\hat{\mu}}}^{\infty} f(t_\mu \mid \mu) dt_\mu \tag{9.6}$$

where $t_{\mu,\mu=\hat{\mu}}$ is the test statistic evaluated at $\mu = \hat{\mu}$ and $f(t_\mu \mid \mu)$ is the PDF of the test statistic assuming the hypothesized value of μ.

For the signal-plus-background model where μ is hypothesized to be greater than or equal to 0, the test statistic should respond only to excesses in observed yield over predicted background yield (upward fluctuations found in the data, $\hat{\mu} > 0$). This requirement defines the alternative test statistics $\tilde{t}_\mu$ and $\tilde{\lambda}$:

$$\tilde{t}_\mu = -2\ln\tilde{\lambda}(\mu) = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu} \geq 0 \\ -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \end{cases} \tag{9.7}$$

In the context of the Neyman-Pearson lemma, this test aims to reject the hypothesized value of μ in favour of the value of μ that best corresponds to the data, $\hat{\mu}$. The p-value can be approximated by exploiting the asymptotic limit of the distribution of $\tilde{t}_\mu$, $F(\tilde{t}_\mu \mid \mu)$ [51]:

---

[11] The standard deviation is obtained from the covariance matrix of the parameters μ and $\boldsymbol{\theta}$.

[12] If additionally $\mu = \hat{\mu}$, then the distribution approaches a standard χ^2 distribution [102].

$$\begin{aligned} p_\mu &= 1 - F(\tilde{t}_\mu \mid \mu) \\ F\left(\tilde{t}_\mu \mid \mu\right) &= \begin{cases} 2\Phi\left(\sqrt{\tilde{t}_\mu}\right) - 1 & \tilde{t}_\mu \leq \mu^2/\sigma^2 \\ \Phi\left(\sqrt{\tilde{t}_\mu}\right) + \Phi\left(\frac{\tilde{t}_\mu + \mu^2/\sigma^2}{2\mu/\sigma}\right) - 1 & \tilde{t}_\mu > \mu^2/\sigma^2 \end{cases} \end{aligned} \tag{9.8}$$

where $F(\tilde{t}_\mu \mid \mu)$ is the cumulative distribution of the test statistic PDF $f(\tilde{t}_\mu \mid \mu)$ and Φ is the cumulative distribution of a standard Gaussian.

In order to test for discovery, the test statistic should test the background-only hypothesis, $\mu = 0$ against $\hat{\mu}$. This requirement defines the test statistic q_0:

$$q_0 = \begin{cases} -2\ln\lambda(0) = -2\ln\left(\frac{L\left(0,\hat{\hat{\boldsymbol{\theta}}}(0)\right)}{L(\hat{\mu},\hat{\boldsymbol{\theta}})}\right) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{9.9}$$

with approximate p-value for $\mu = 0$ given by:

$$\begin{aligned} p_0 &= 1 - F(q_0 \mid \mu = 0) \\ F(q_0 \mid \mu = 0) &= \Phi(\sqrt{q_0}) \\ q_0 &= \begin{cases} \hat{\mu}^2/\sigma^2 & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \end{aligned} \tag{9.10}$$

Similarly, this test aims to reject the background-only hypothesis $\mu = 0$ in favour of $\hat{\mu}$. The smaller the value of p_0, the greater the *significance* Z of the excess in observed yield over predicted background yield.

For a Gaussian-distributed measurement, the probability that a measurement is outside of $\pm 1\sigma$ from the mean (i.e. into the positive/negative tail in the case of upwards/downwards fluctuation) is approximately 32%. If each bin of the observable distribution is assumed to be independent, then the probability to observe a $\pm 1\sigma$ excess (deficit) is approximately 32% (one excess or deficit every three bins). Similarly, the probability to observe a $\pm 2\sigma$ excess (deficit) is approximately 4.6% (one excess or deficit every twenty-two bins). Therefore, a single bin in the $\langle m_{\ell\ell} \rangle$ distribution (with over twenty populated bins) having an excess of approximately 2σ is not statistically significant.

In many cases, there is no prior knowledge of where an excess may occur in a distribution. The probability to find a single bin with an excess *anywhere in a distribution* is higher than the probability to find a similar excess when only a particular bin is considered. This is known as the *look elsewhere effect.* The p_0 and significance values reported in this section do not take into account this effect and are hence referred to as *local* values.[13]

9.1.4 Test statistic for upper limit on μ

In order to set an upper limit on μ, the test statistic should test the value of μ obtained from the data, $\hat{\mu}$, against hypothesized values of μ. Specifically for an *upper* limit, when $\hat{\mu} > \mu$, the data should not be regarded as incompatible and hence the test statistic q_μ is set to 0 (in other words, q_μ is 0 if the data fluctuates upwards, unlike q_0, which is 0 if the data fluctuates downwards):

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{9.11}$$

The upper limit on μ is found by determining the value of μ such that p_μ corresponds to the desired *confidence level* $1 - \alpha$:

$$\alpha = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q_\mu \mid \mu) dq_\mu \tag{9.12}$$

The test statistic can be expressed in terms of the alternate test statistic $\tilde{\lambda}$:

$$\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \\ -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{9.13}$$

In numerical tests, the alternate form $\tilde{q}_\mu$ is similar to q_μ. With the approximations described previously, both yield the same upper limit. For both q_μ and $\tilde{q}_\mu$, the approximate p-value is analytic and therefore the upper limit on μ can be calculated analytically:

---

[13] Global values can be calculated with the prescription given in [1]. Global significance is of interest when the local significance is especially large ($\gtrsim$ 3–4σ).

$$\begin{aligned} F\left(\tilde{q}_{\mu} \mid \mu\right) &= \begin{cases} \Phi(\sqrt{\tilde{q}_{\mu}}) & 0 < \tilde{q}_{\mu} \leq \mu^2/\sigma^2 \\ \Phi\left(\frac{\tilde{q}_{\mu}+\mu^2/\sigma^2}{2\mu/\sigma}\right) & \tilde{q}_{\mu} > \mu^2/\sigma^2 \end{cases} \\ p_{\mu} &= 1 - F\left(\tilde{q}_{\mu} \mid \mu\right) \\ \mu_{\text{up}} &= \hat{\mu} + \sigma\Phi^{-1}(1-\alpha) \end{aligned} \tag{9.14}$$

For a limit at $1-\alpha = 95\%$ confidence level, $\Phi^{-1}(0.95) = 1.64$. With σ determined from the covariance matrix of the parameters μ and $\boldsymbol{\theta}$, the above asymptotic approximations can be used to obtain an estimate of the upper limit without computationally expensive toy-based construction of the test statistic PDFs *for large N*. Tests of the asymptotic approximations suggest that they produce accurate results when the number of background events per bin is at least ten. All of the bins of the observable ($\langle m_{\ell\ell}\rangle$) in the signal region of $h \to Z_d Z_d \to 4\ell$ have less than ten background events, and therefore asymptotic approximations have questionable validity. While the results with approximations are not fully accurate, approximations are still helpful in developing the interpretation software and studying the effects of systematics and other changes. In Section 9.1.6, toy-based construction of the test statistic PDFs is described, and in Section 9.6, upper limits produced using both approaches are compared.

9.1.5 CL$_\text{s}$ method for upper limits

The upper limit of a parameter is usually interpreted as the largest value that is not excluded. In order to avoid exclusion in the case of marginal sensitivity, this search uses the *CL$_s$ method*. This method is recommended by the ATLAS Collaboration and is also used by other experiments.

In the frequentist interpretation, the interval between 0 and the upper limit at 95% confidence level covers the true value in at least 95% of cases. However, the remainder of the cases correspond to excluded hypotheses to which the experiment has no sensitivity or at best, marginal sensitivity. For example, a hypothesis that expects $n_{\text{signal}} \ll n_{\text{background}}$ would be excluded if $n_{\text{obs}} \lesssim n_{\text{background}}$ (downward fluctuation), with a probability less than or equal to α. In other words, it is possible for exclusions to be set in cases of limited sensitivity. In terms of the hypothesis PDFs, this scenario occurs when the alternative (signal-plus-background) and null (background-only) hypothesis PDFs overlap significantly.

According to the test described in Section 9.1.1, a signal model is excluded if $p_b < \alpha$. In order to protect against exclusion in case of little or no sensitivity, the $\mathrm{CL_s}$ method [103] defines a signal as excluded if:[14]

$$\frac{p_{s+b}}{1 - p_b} < \alpha \tag{9.15}$$

where $p_{s+b} = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q_\mu \mid \mu) dq_\mu$ is the p-value of the signal-plus-background hypothesis and $p_b = \int_{-\infty}^{q_{\mu,\mathrm{obs}}} f(q_\mu \mid 0) dq_\mu$ is the p-value of the background-only hypothesis.[15]

If the PDFs do not overlap significantly, then the denominator $1 - p_b$ is close to 1 and the numerator p_{s+b} is marginally penalized. In cases of poor sensitivity, the probability under the background-only hypothesis PDF is large and p_{s+b} is penalized.

9.1.6 Toy-based construction of test statistic PDFs

The calculation of p-values and upper limits requires knowledge of the test statistic PDF $f(q_0 \mid 0)$ for testing discovery, and $f(\tilde{q}_\mu \mid \mu)$ for determining upper limits. The approximations given in Section 9.1.3 and Section 9.1.4 were derived by neglecting the $\mathcal{O}(1/\sqrt{N})$ terms of $t_\mu = -2 \ln \lambda(\mu)$ (the large N limit). To achieve accuracy for small N—the case for the $h \to Z_d Z_d \to 4\ell$ signal region—the PDFs must be generated with toy Monte Carlo. The observed upper limit on μ is determined with the following procedure:

1. Calculate the *observed test statistic* $\tilde{q}_{\mu,\mathrm{obs}}$ by evaluating $\tilde{q}_\mu$ for a range of values of μ with the unconditional and conditional maximum likelihood estimates (MLEs) determined from profiling the observed data.

2. For each value of μ:

    (a) Determine $f(\tilde{q}_\mu \mid \mu)$ for the fixed value of $\mu = \mu'$ by generating toy data:

        i. With values for the signal and background yields drawn from a Poisson distribution, generate a toy signal-and-background dataset $\mathbf{x}$ according to the signal-plus-background PDF $f(\mathbf{x} \mid \mu, \boldsymbol{\theta})$, with the MLEs determined from profiling the observed data and μ'.

---

[14] Declaring exclusion if $p_b < \alpha$ or $p_{s+b} < \alpha$ are referred to as the $\mathrm{CL_b}$ and $\mathrm{CL_{s+b}}$ methods, respectively.

[15] The alternate form $\tilde{q}_\mu$ can be used interchangeably with q_μ.

        ii. Calculate $\tilde{q}_\mu$ for $\mu = \mu'$. This requires profiling the toy dataset and will therefore result in a set of MLEs for the toy dataset.
        iii. The above two steps are repeated until the probability distribution of the test statistic, $f(\tilde{q}_\mu \mid \mu')$, is well-populated.
    (b) Evaluate $p_\mu = \int_{\tilde{q}_{\mu,\mathrm{obs}}}^{\infty} f(\tilde{q}_\mu \mid \mu)$ for $\mu = \mu'$ with the generated distribution and $\tilde{q}_{\mu,\mathrm{obs}}(\mu = \mu')$ (the latter having been determined in the first step).
3. Determine the value of μ such that $p_\mu = 0.05$ from a plot of p_μ vs. μ. This is the observed upper limit on μ.

The expected upper limit on μ is determined by repeating the above with multiple toy datasets generated with $f(\mathbf{x} \mid \mu = 0, \hat{\hat{\boldsymbol{\theta}}}(0, \mathrm{obs}))$:

1. Determine the conditional MLEs $\hat{\hat{\boldsymbol{\theta}}}(0, \mathrm{obs})$ from profiling the observed data.
2. With a value for the background yield drawn from a Poisson distribution, generate a toy background-only dataset $\mathbf{x}'$ according to the signal-plus-background PDF $f(\mathbf{x} \mid \mu, \hat{\hat{\boldsymbol{\theta}}}(0, \mathrm{obs}))$.
3. Determine the upper limit using the procedure *for the observed upper limit* described above using the toy background-only dataset as input (i.e. $\mathbf{x} \to \mathbf{x}'$).
4. Repeat, by generating another toy background-only dataset $\mathbf{x}'$ and re-determining the observed upper limit μ^{up}.
5. From the distribution of the upper limits $f(\mu^{\mathrm{up}} \mid \mu = 0, \hat{\hat{\boldsymbol{\theta}}}(0, \mathrm{obs}))$:[16]
    - Median: expected limit
    - 68% and 95% bands: $\pm 1\sigma$ and $\pm 2\sigma$ of the expected limit, respectively

9.2 Implementation

The binned extended likelihood is built from 1 GeV-binned distributions of the observable ($\langle m_{\ell\ell} \rangle$) in the signal region with the three channels combined. The data is represented by a single distribution, while the background and signal predictions are represented by distributions *per process* and *per systematic variation*. Distributions of

---

[16] Usually follows a Gaussian distribution.

each prediction are produced with the nominal experimental parameters (i.e. nominal lepton corrections, integrated luminosity, and background modeling), and with $\pm 1\sigma$ variations of each parameter for a total of $2N_{\text{syst}} + 1$ distributions per prediction. As described in Section 6.2.3.1, the variations represent the $\pm 1\sigma$ systematic uncertainties on each parameter. Each parameter and its uncertainties are estimated in external studies (auxiliary measurements) and are propagated to the analysis objects by prescribed changes to object kinematics (e.g. nominal and $\pm 1\sigma$ values of a lepton's p_T), event weights (weight for nominal correction and weights for $\pm 1\sigma$ of the nominal correction), or prediction normalization in the case of integrated luminosity and its uncertainties. In the likelihood, each parameter with systematic uncertainties is assigned a nuisance parameter α. The observable is *vertically morphed* [100] between its nominal value and $\pm 1\sigma$ variations, with the morphing controlled by varying α between -1 and 1. Nuisance parameters are also assigned for the statistical uncertainty in each bin of each prediction. All nuisance parameters are modelled in the likelihood with Gaussian constraints.

As described in Section 6.2.1, events with the $h \rightarrow Z_d Z_d \rightarrow 4\ell$ signal process were simulated with the Hidden Abelian Higgs Model at steps of $\Delta m_{Z_d} = 5$ GeV across the range 15 GeV $\leq m_{Z_d} \leq$ 60 GeV. Distributions of $\langle m_{\ell\ell} \rangle$ in the signal region of these events are produced at every step in m_{Z_d}. Figure 8.10 shows distributions for $m_{Z_d} = 20,\ 35,\ 55$ GeV. At intermediate steps of $\Delta m_{Z_d} = 0.5$ GeV, the signal shape is generated by moment morphing [104] with the simulated shapes at every $\Delta m_{Z_d} = 5$ GeV as templates. The signal normalization for the morphed distributions is determined from linear interpolation of the signal yields. As with the background predictions, the distributions of each m_{Z_d} signal prediction are produced with nominal and $\pm 1\sigma$ values of the experimental parameters.

In order to incorporate the data-driven Z+jets estimate in the likelihood, the nominal estimate (unbinned) is divided equally among bins of $\langle m_{\ell\ell} \rangle$, with the relative (percent) statistical and systematic uncertainties in each bin equal to the relative uncertainties on the unbinned estimate.

The interpretation is implemented with the HistFitter software framework [105]. This framework interfaces with HistFactory [100], RooStats [106], and RooFit [107]. The distributions, nuisance parameters, constraint types, and parameter of interest are specified in a HistFitter configuration file. HistFitter utilizes HistFactory to create a RooFit workspace containing the likelihood, variables representing the parameters, the distributions, and a statistical model relating the contents. Profiling ("fitting") of the likelihood is implemented in RooFit and statistical tests are implemented in RooStats. HistFitter provides interfaces to both operations.

9.3 Transformation of signal strength limit

The upper limit on the signal strength μ is transformed into a limit on the branching ratio $\mathcal{B}(h \to Z_d Z_d)$ with the following transformation:

$$\sigma_{h\to Z_d Z_d \to 4\ell} = \frac{\mu N_{\text{signal,SR}}}{\mathcal{L}} \tag{9.16}$$

$$\mathcal{B}(h \to Z_d Z_d \to 4\ell) = \frac{\sigma_{h\to Z_d Z_d \to 4\ell}}{\sigma_h} \tag{9.17}$$

$$\mathcal{B}(h \to Z_d Z_d) = \frac{\mathcal{B}(h \to Z_d Z_d \to 4\ell)}{4[\mathcal{B}(Z_d \to \ell\ell)]^2} \tag{9.18}$$

where:

- $N_{\text{signal,SR}}$ is the predicted signal yield in the signal region,
- σ_h is the Standard Model (SM) Higgs production cross-section limit (48.58 pb at N^{3}LO from [54, 93–95]),[17]
- $\mathcal{B}(h \to Z_d Z_d \to 4\ell)$ is the branching ratio summed over all three possible 4ℓ channels ($4e$, $2e2\mu$, and 4μ),
- $\mathcal{B}(Z_d \to \ell\ell)$ is the branching ratio to one flavour of $\ell\ell$. Partial fractions of 0.25:0.25:0.25:0.25 for the $4e$:$2e2\mu$:4μ:$2\mu 2e$ channels are assumed, with $\mathcal{B}(Z_d \to \mu\mu) + \mathcal{B}(Z_d \to ee)$ taken from theory [10] for different m_{Z_d}.

This limit is divided by acceptance and efficiency in order to produce limits that are independent of the detector. This factor is defined as the ratio of $N_{\text{signal,SR}}$ to the generated number of signal events.

In addition to setting a limit on the $h \to Z_d Z_d$ branching ratio, a limit on the effective Higgs mixing parameter $\kappa' = \kappa \cdot m_h^2/|m_h^2 - m_S^2|$ is also set with the transformation described in [9, 10]. This limit is shown in Figure 9.7. The limit on

---

[17] This assumes that the presence of Beyond-the-Standard-Model decays of the Higgs boson does not signicantly alter the Higgs boson production cross-section from the SM prediction.

κ' is calculated with the branching ratio limit unfolded for detector effects with the following transformation:

$$\kappa'^2 = \frac{\Gamma_{\mathrm{SM}}}{f(m_{Z_d})} \frac{\mathcal{B}(h \to Z_d Z_d)}{1 - \mathcal{B}(h \to Z_d Z_d)} \tag{9.19}$$

where

$$f(m_{Z_d}) = \frac{v^2}{32\pi m_h} \sqrt{1 - \frac{4m_{Z_d}^2}{m_h^2}} \frac{\left(m_h^2 + 2m_{Z_d}^2\right)^2 - 8\left(m_h^2 - m_{Z_d}^2\right) m_{Z_d}^2}{m_h^4} \tag{9.20}$$

The branching ratio $\mathcal{B}(Z_d \to \ell\ell)$ is the ratio of partial width $\Gamma(Z_d \to \ell\ell)$ to the sum of widths across all possible decays. To leading order, the ratio of widths results in a cancellation of dependence on the kinetic mixing parameter ϵ, and therefore this branching ratio is independent of ϵ and no limits are set on ϵ. The branching ratio $\mathcal{B}(h \to Z_d Z_d)$ depends only on the Higgs mixing parameter κ. The limit is set on the *effective* Higgs mixing parameter κ' because the dark Higgs mass m_S is not known.

9.4 Fitting results

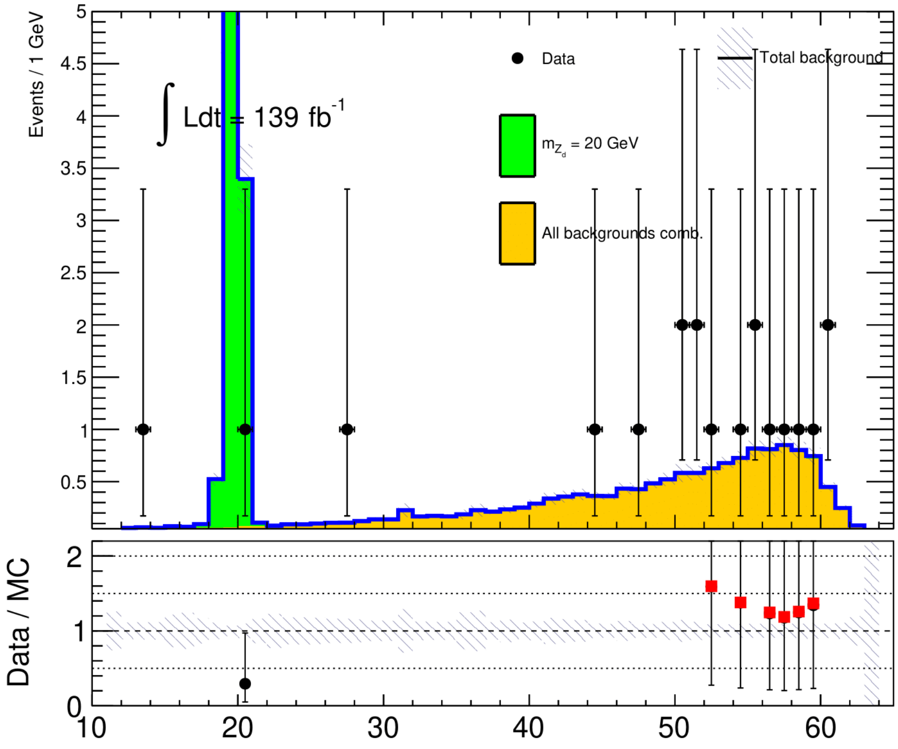
The "best fit" values for the parameters μ and $\boldsymbol{\theta}$ (unconditional MLEs) are determined by fitting the model (likelihood with predicted signal and background yields) with the observed data. The maximum likelihood fit adjusts the parameters in order to achieve the best possible agreement with the observed data. For example, Figure 9.1 shows the $\langle m_{\ell\ell} \rangle$ distribution before and after fitting with the $m_{Z_d} = 20$ GeV signal hypothesis.

In the vicinity of the signal hypothesis, one event is observed with $\langle m_{\ell\ell} \rangle \approx 21$ GeV. The fit adjusts the signal strength such that the signal hypothesis more closely matches the observed yield; however, it is not adjusted to match the observed yield exactly because of the absence of events in adjacent bins and the width of the hypothesis (greater than 1 GeV).

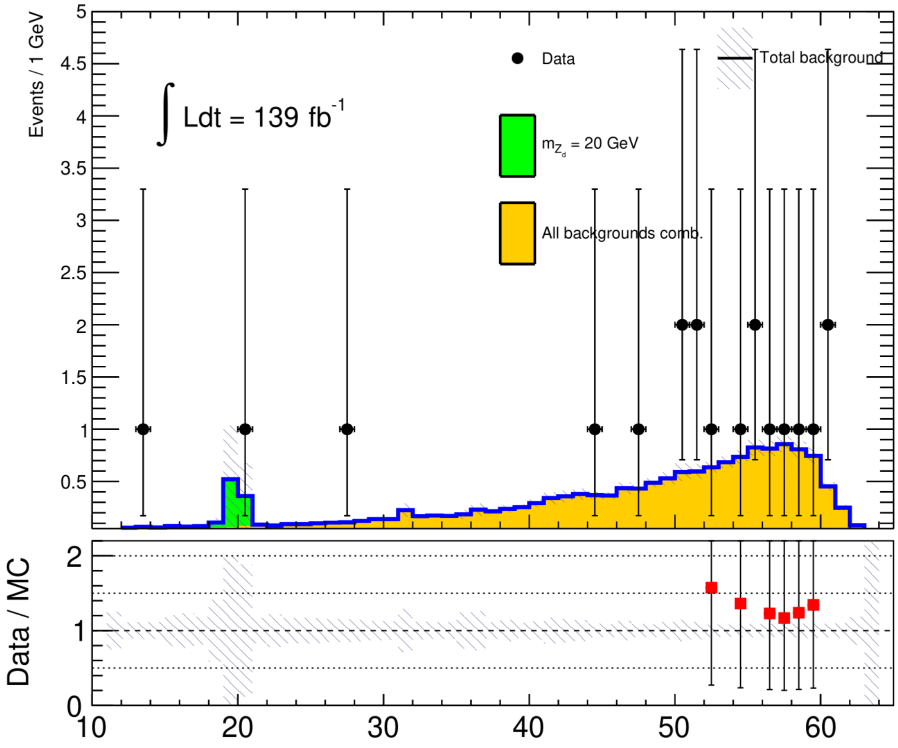
The behaviour of the nuisance parameters in the fit can be visualized in a *ranking and pull plot*. Figure 9.2 shows the ranking and pull plot for the $m_{Z_d} = 35$ GeV signal hypothesis and corresponding fit. The *pull* of a nuisance parameter is defined as the difference between the post-fit and pre-fit value of the nuisance parameter divided by the uncertainty on the parameter, $(\hat{\theta} - \theta_0)/\Delta\theta$. The plot shows that none of the central values of the nuisance parameters are pulled by the fit (i.e. the central values were accurately estimated pre-fit). The black band corresponds to the $\pm 1\sigma$ values of the nuisance parameter (uncertainty). If the black band is smaller (larger) than $\pm 1\sigma$ (red band), then the uncertainty on the nuisance parameter is reduced (increased) by the fit—i.e. it was over-estimated (under-estimated) pre-fit.[18] None of the nuisance parameter uncertainties are significantly adjusted by the fit. The nuisance parameters are ranked from highest to least impact on the signal strength. The impact, $\Delta\mu$, is the difference between the nominal best fit value of μ (determined with all parameters floating in the fit) and the value of μ determined with the nuisance parameter fixed to its best fit value *shifted by its pre- or post-fit uncertainty* $\pm\Delta\theta$ or $\pm\Delta\hat{\theta}$. As expected, the nuisance parameters corresponding to larger experimental uncertainties (e.g. electron identification efficiency) have greater impact than those corresponding to smaller uncertainties. The pre- and post-fit impact bands are aligned; none of the nuisance parameters exhibit a change in impact on the signal strength after fitting. This indicates that the nuisance parameters and their uncertainties were accurately modelled pre-fit.

---

[18] To reduce the uncertainty is to *constrain* the nuisance parameter.



(a) Pre-fit



(b) Post-fit

Figure 9.1: Pre- and post-fit $\langle m_{\ell\ell} \rangle$ distributions for the $m_{Z_d} = 20$ GeV signal hypothesis.

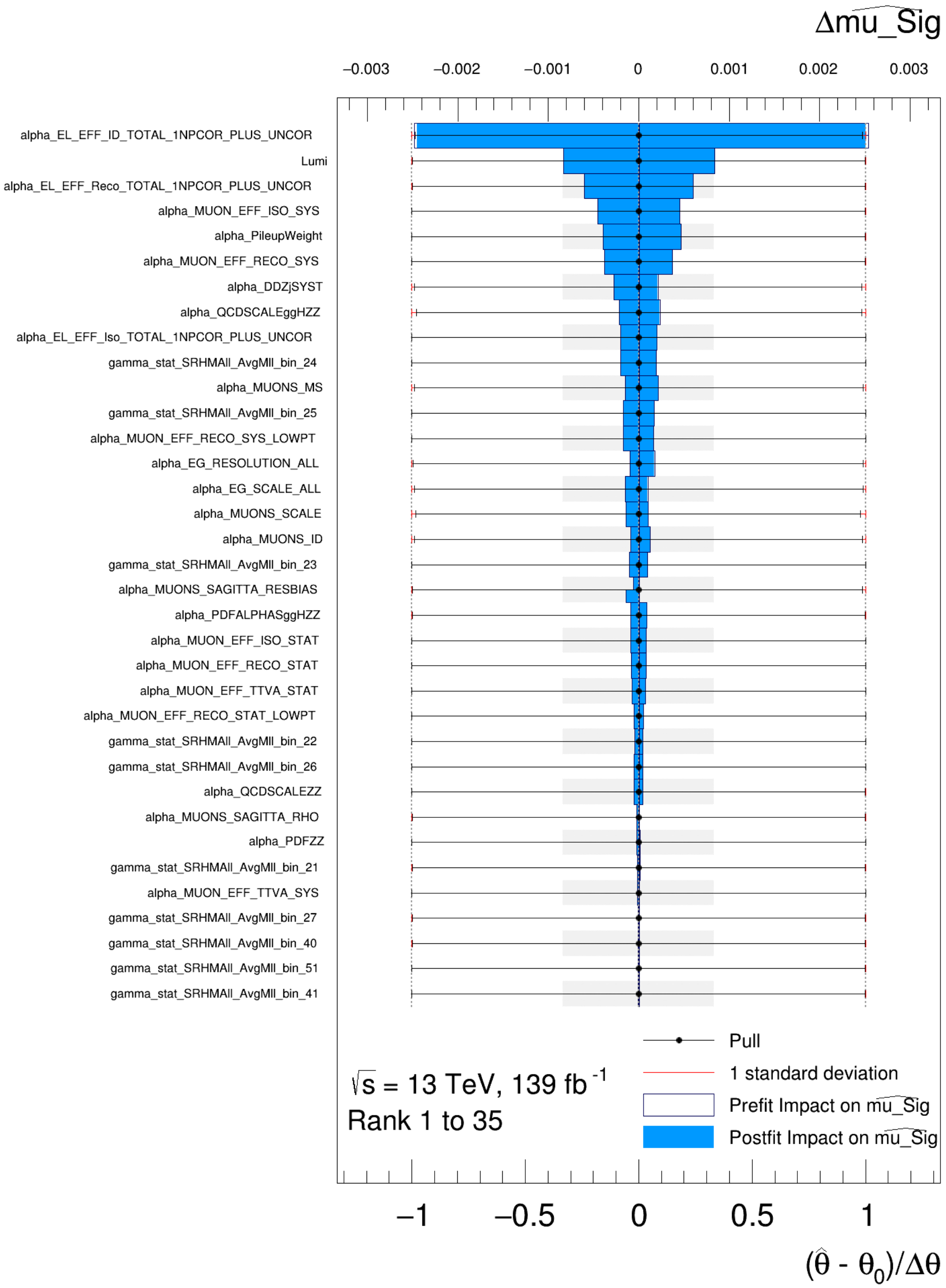


Figure 9.2: Pull and ranking plot for the $m_{Z_d} = 35$ GeV signal hypothesis. The nuisance parameters are ordered from highest (top) to lowest (bottom) impact. The interpretation of the pull and impact is described in the main text. `gamma_stat_SRHMAll_AvgMll_bin_num` are nuisance parameters corresponding to the statistical uncertainties of the predictions in each bin.

9.5 p_0 and significance scan

Shown in Figure 9.3 are the local p_0-values across the range of m_{Z_d} hypotheses (p_0 *scan*). The largest excess in the $\langle m_{\ell\ell}\rangle$ distribution is from the observation of two events with $\langle m_{\ell\ell}\rangle \approx 27.6$ GeV and $\langle m_{\ell\ell}\rangle \approx 27.9$ GeV. The local significance of this excess determined by the toy-based p_0 scan is approximately 2.5σ. This result is larger than the result calculated with asymptotic formulae ($\approx 2\sigma$) shown in the sub-plot of Figure 8.10. This is because the toy-based p_0 scan correctly takes into account detector resolution and the Z_d width—as long as the binning is narrow enough—while the formulae give the significance *for a given bin* and hence depend on the bin width.

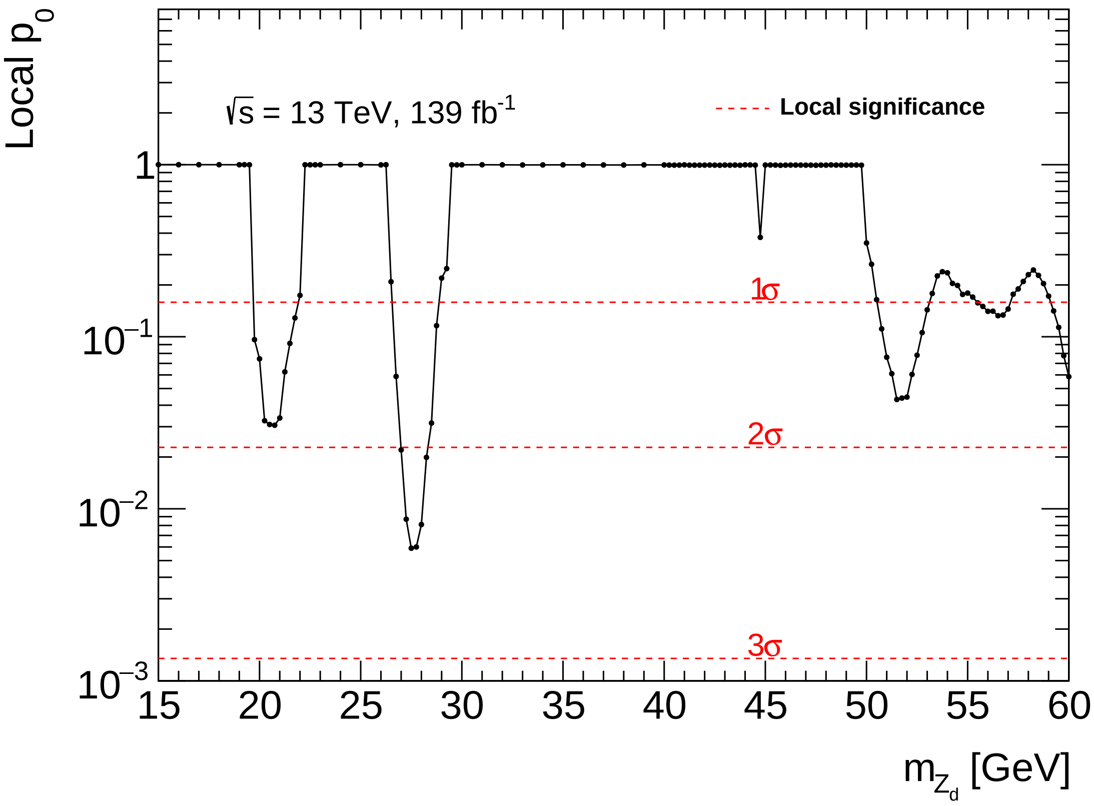


Figure 9.3: Local p_0 value vs. m_{Z_d} calculated with toy-based generation of the test statistic PDF $f(q_0 \mid 0)$. The local significance of the excess at $m_{Z_d} \approx 27$ GeV is approximately 2.5σ. Ten thousand toys are used per signal mass point. In the vicinity of observed data, the local p_0 is calculated at steps of $\Delta m_{Z_d} = 0.25$ GeV with signal distributions generated by moment morphing.

9.6 Upper limits

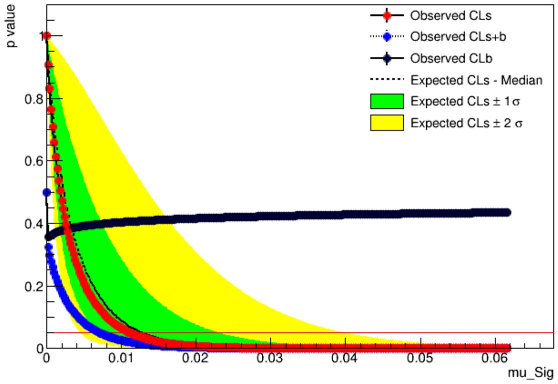
The upper limit at 95% confidence level on the signal strength μ is determined by plotting p_μ versus μ and identifying the value of $\mu = \mu^{\text{obs}}$ such that $p_\mu = 0.05$. Shown in Figure 9.4 are the p-values calculated with the CL_s, $\text{CL}_\text{s+b}$, and CL_b methods and asymptotic formulae plotted against μ for signal hypotheses $m_{Z_d} = 18, 21, 23$ GeV. These signal hypotheses are in the vicinity of one observed event with $\langle m_{\ell\ell} \rangle \approx 21$ GeV (shown in Figures 9.1 and 8.10). The excursion in the upper limit for the $m_{Z_d} = 21$ GeV hypothesis relative to the other two hypotheses is due to the observed event.

The upper limit on the signal strength is transformed into a *detector-independent* upper limit on the branching ratio with the formulae given in Section 9.3. Shown in Figure 9.5 is the upper limit on the branching ratio $\mathcal{B}(h \to Z_d Z_d)$ calculated with the asymptotic formulae described in Section 9.1.4 (top) and with toy-based generation of the test statistic PDF $f(\tilde{q}_\mu \mid \mu)$ described in Section 9.1.6 (bottom). The excursions in the upper limit on μ in the presence of non-zero observed yield evident in Figure 9.4 are also visible in the upper limit on the branching ratio. The asymptotic formulae give a good approximation for the observed and expected limits, but not for the expected $\pm 1\sigma$ and $\pm 2\sigma$ bands. Over the range of m_{Z_d} hypotheses, the average upper limit on the branching ratio is approximately 4×10^{-5}.

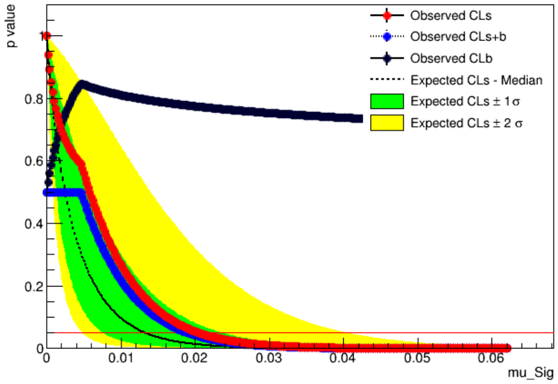
Shown in Figure 9.6 is the acceptance times efficiency factor used to derive the detector-independent limit. For lighter m_{Z_d}, the final state leptons are softer (lower momentum) and consequently have lower reconstruction efficiencies.

Finally, the upper limit on the branching ratio is transformed into the upper limit on the effective Higgs mixing parameter $\kappa' = \kappa \times m_h^2 / |m_h^2 - m_S^2|$ (shown in Figure 9.7) using the transformation given by Equations 9.19 and 9.20. Over the range of m_{Z_d} hypotheses, the average upper limit on κ' is approximately 3×10^{-4}.

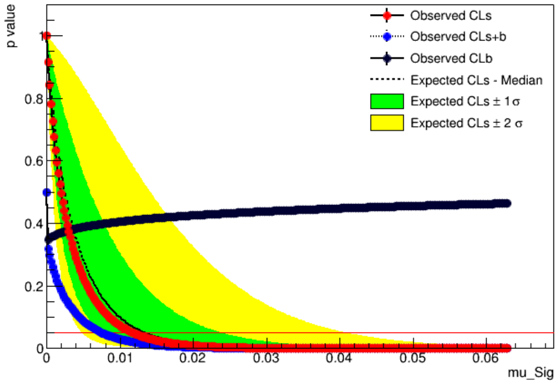
Both limits improve on the limits of the previous published search (using data collected in 2015–2016 *only*) by a factor of four, and are the strongest limits to date for m_{Z_d} in this mass range. The limits are also in line with the expectations from the phenomenological studies discussed in Chapter 3.



(a) $m_{Z_d} = 18$ GeV

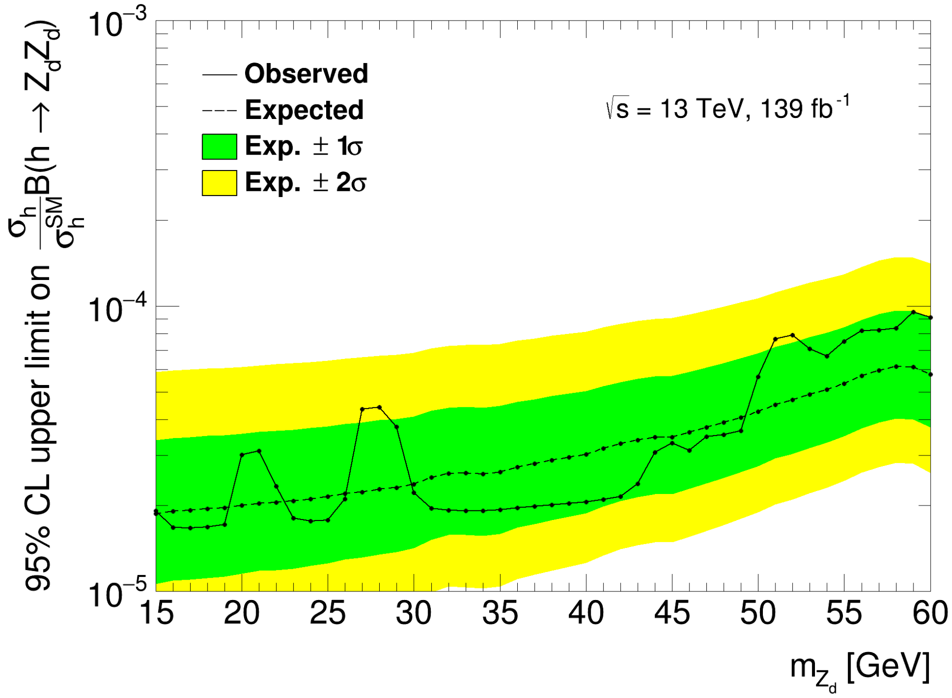


(b) $m_{Z_d} = 21$ GeV

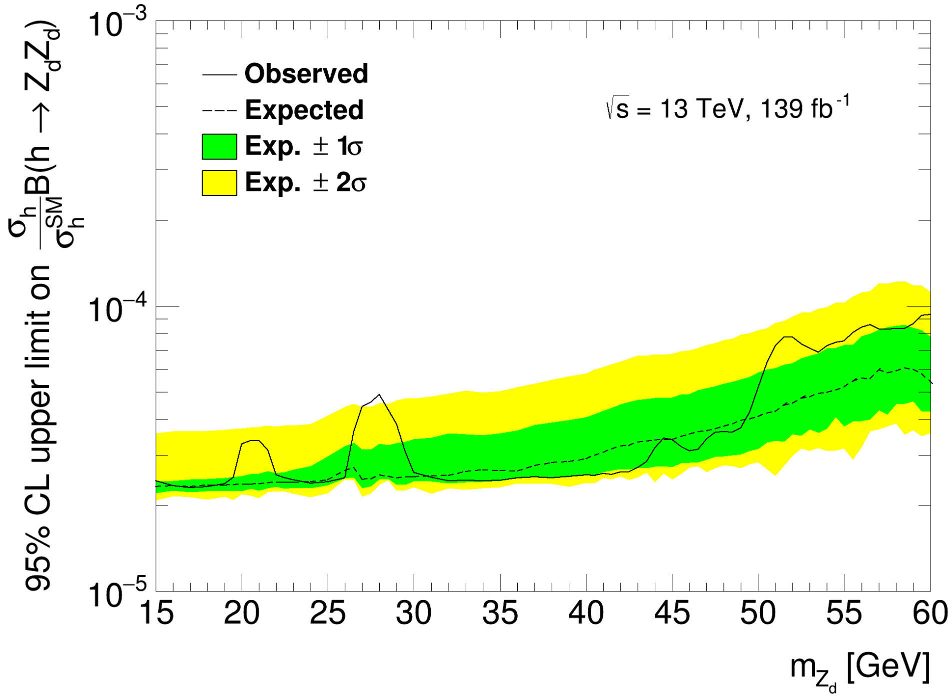


(c) $m_{Z_d} = 23$ GeV

Figure 9.4: p-value vs. μ for signal hypotheses $m_{Z_d} = 18, 21, 23$ GeV calculated with asymptotic formulae and with the $\mathrm{CL_s}$, $\mathrm{CL_{s+b}}$, and $\mathrm{CL_b}$ methods. The value of μ at which the horizontal red line $p = 0.05$ intersects with the calculated p-values is the observed upper limit on μ at 95% confidence level. The excursion in the upper limit for the $m_{Z_d} = 21$ GeV hypothesis relative to the other two hypotheses is due to the observation of one event with $\langle m_{\ell\ell} \rangle \approx 21$ GeV.



(a) Calculated with asymptotic formulae



(b) Calculated with toys

Figure 9.5: Upper limit on the branching ratio $\mathcal{B}(h \to Z_dZ_d)$ at 95% confidence level. For the limit set with toys: in the vicinity of observed data, the limit is set every $\Delta m_{Z_d} = 0.5$ GeV with 150000 toys. The limit is set every 1 GeV with 50000 toys elsewhere.

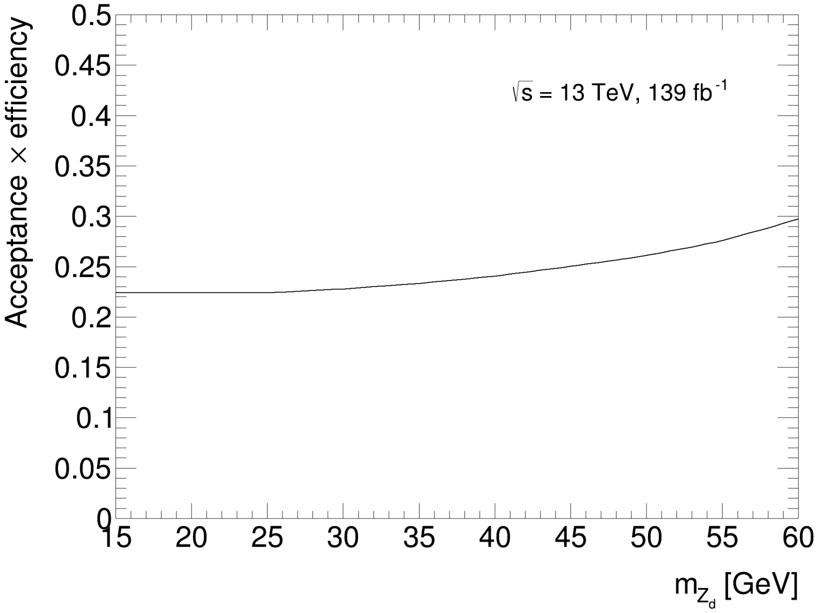


Figure 9.6: Acceptance times efficiency factor vs. m_{Z_d}.

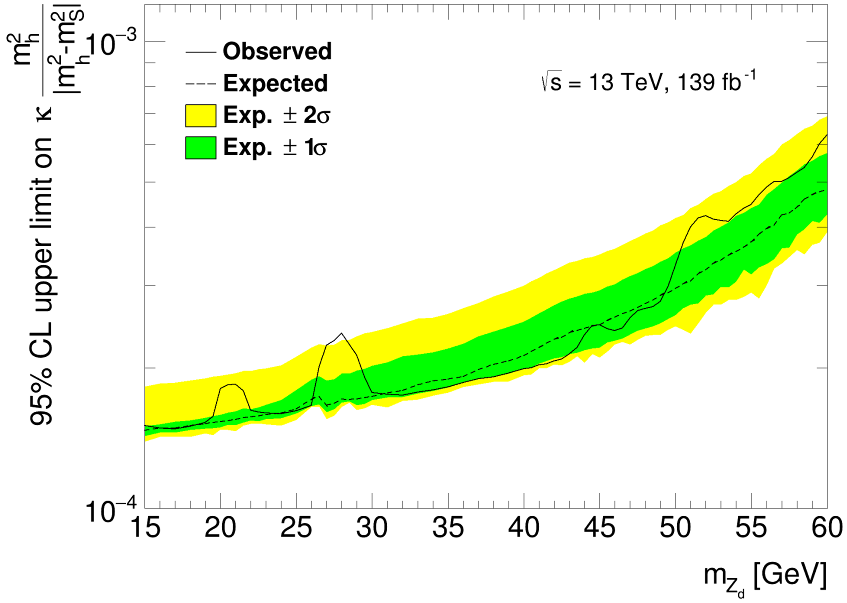


Figure 9.7: Upper limit on the effective Higgs mixing parameter κ' at 95% confidence level. Calculated using the toy-based branching ratio limit shown in Figure 9.5.

Chapter 10

Conclusions

This thesis presents the search for the dark sector process $h \rightarrow Z_d Z_d \rightarrow 4\ell$ with data collected by the ATLAS detector at the Large Hadron Collider in 2015–2018. In this theorized process, the Standard Model (SM) Higgs decays to two bosons belonging to a dark sector, each of which decays to two electrons or muons. As described in Chapter 3, the addition of a $U(1)$ dark sector via kinetic mixing to the SM gives rise to this process. The production of $Z_d Z_d$ from SM Higgs ($h \rightarrow Z_d Z_d$) is enhanced significantly by the addition of mixing between the SM Higgs and *dark Higgs*. This process is searched for in data corresponding to an integrated luminosity of 139 fb^{-1} by reconstructing and identifying leptons compatible with $Z_d Z_d$ production from a SM Higgs and compatible with decay from Z_d. The strategy and analysis objects used for this search are detailed in Chapter 5.

The backgrounds from SM processes that share the same decay signature are estimated with Monte Carlo (MC)-based simulation and a data-driven technique. One of the chief improvements of this iteration of the search is the introduction of a more accurate Z+jets background estimate. The result of the previous iteration observed a large excess[1] which could be attributed to an under-estimate from simulation of jets from the Z+jets process *faking* leptons. In order to more accurately assess this background, a *fake factor* method is implemented (Chapter 7). This method estimates the Z+jets background in the signal region by measuring transfer factors in kinematic regions enriched in Z+jets background and applying the transfer factors in kinematic regions related to the signal region. The estimate from this method is compatible with zero. The contribution of other processes to the signal region are estimated with

---

[1] 3.2σ local significance and 1.9σ global significance.

simulation. These simulation-based estimates (Chapter 6.2) are validated against data in specialized kinematic regions and are shown to have acceptable agreement (Section 8.1).

The results presented in Chapter 8 show no evidence for the existence of the $h \to Z_d Z_d \to 4\ell$ process. The data in the signal region are statistically compatible with SM background estimates. A local significance scan incorporating the complete statistical model is performed in order to accurately assess the excesses of observed events over estimated background events in the $\langle m_{\ell\ell} \rangle$ distribution (Section 9.5). The largest excess is from two events observed with $\langle m_{\ell\ell} \rangle \approx 27.6$ GeV and $\langle m_{\ell\ell} \rangle \approx 27.9$ GeV, with a local significance of approximately 2.5σ. In the absence of significant excesses, the results are interpreted by setting upper limits on the branching ratio $\mathcal{B}(h \to Z_d Z_d)$ and effective Higgs mixing parameter κ' (Section 9.6). The observed upper limits are $\mathcal{B}(h \to Z_d Z_d) \lesssim 4 \times 10^{-5}$ and $\kappa' \lesssim 3 \times 10^{-4}$ over the mass range of Z_d hypotheses considered (15 GeV $< m_{Z_d} <$ 60 GeV) and are a factor of four smaller than the limits of the previous search. These limits help to constrain specifically the existence of the $U(1)$ *Higgs portal*, and are the strongest limits to date for m_{Z_d} in this mass range.

Appendix A

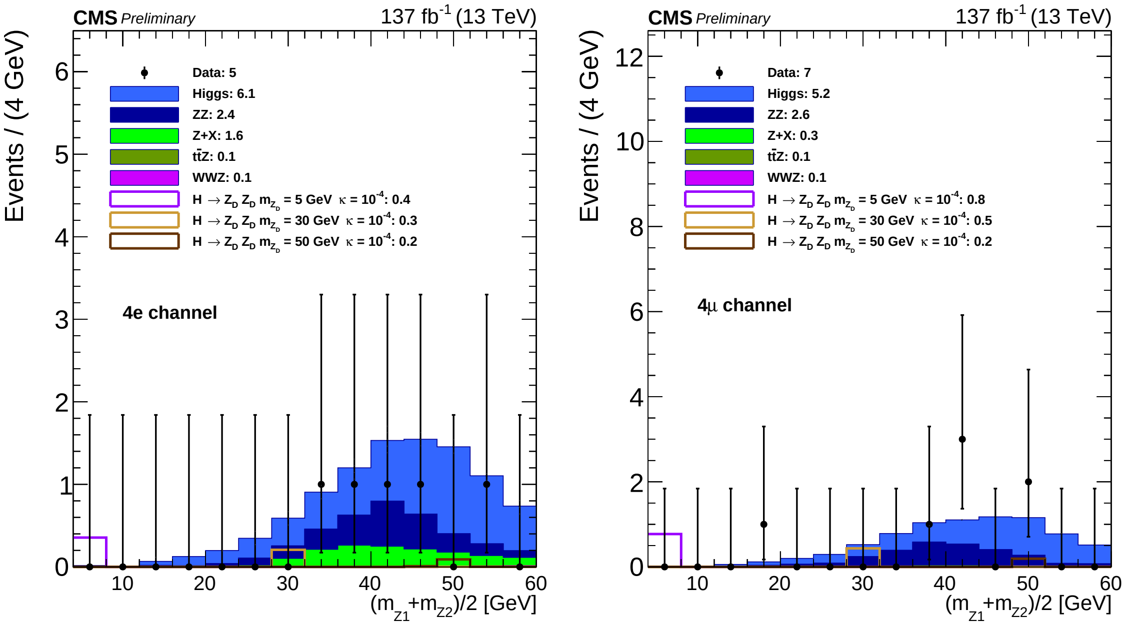
2016–2018 CMS result

In May 2020, the CMS experiment published preliminary results of its *search for low-mass dilepton resonances in Higgs boson decays to four-lepton final states* with the CMS 2016–2018 dataset (137 fb^{-1}) [108]. This analysis searched for the decay of the SM Higgs to one or two BSM particles with four leptons (electrons and/or muons) in the final state ($h \to ZX \to 4\ell$ or $h \to XX \to 4\ell$). Two models were considered: one where the BSM particle X is the dark photon Z_d arising from the addition of a $U(1)_d$ gauge symmetry (as in the ATLAS search described in this thesis) and the other where X is an axion-like particle (pseudoscalar a). The mass range considered was 4 GeV $< m_X <$ 62.5 GeV.

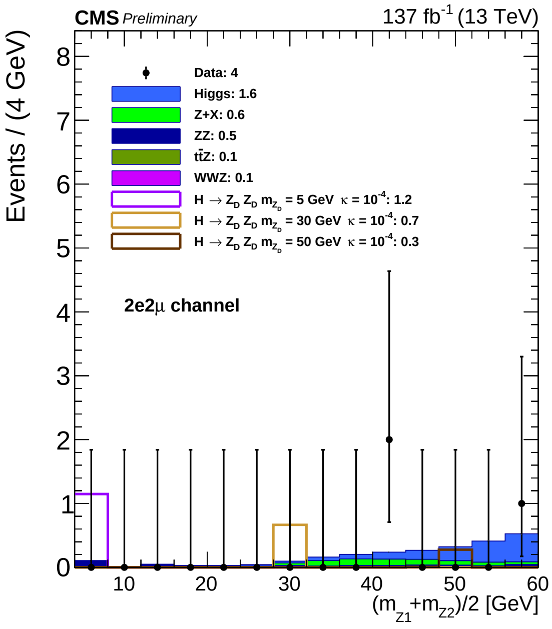
The CMS analysis is similar to the ATLAS analysis. Both analyses impose p_T and $\Delta R(\ell, \ell')$ requirements on leptons and a Higgs mass window on the quadruplet invariant mass. Unlike the ATLAS analysis, the CMS analysis does not apply a Z veto, selects the quadruplet with smallest $(m_{12} - m_{34})/(m_{12} + m_{34})$, and does not apply any requirement on the compatibility of m_{12} and m_{34}. Irreducible backgrounds in both analyses are estimated with Monte Carlo-simulated events and Z+jets background (reducible) in both analyses are estimated with similar data-driven techniques. The technique described in the CMS preliminary result derives fake factors in Z+jets-enriched regions and applies the factors in regions with one or two leptons passing inverted identification criteria. The inverted criteria are not defined in the publication. Contributions from $ZZ^* \to 4\ell$ events are estimated with simulation.

No evidence for the $h \to Z_d Z_d \to 4\ell$ process was found in the CMS search. Shown in Figure A.1 are the distributions of $(m_{Z1} + m_{Z2})/2 \Leftrightarrow \langle m_{\ell\ell} \rangle$ in the signal region of the CMS search. These distributions can be compared against the $\langle m_{\ell\ell} \rangle$ distributions of the ATLAS search shown in Figure 8.10. Across the three channels, 16 (20) events are observed and 21.5 ($15.6 \pm 0.4 \pm 1.2$) events are predicted by CMS (ATLAS). The results of the two experiments are compatible. For Z+jets background, CMS predicts 2.5 events with a systematic uncertainty varying from 48% to 100% and ATLAS predicts $0.43 \pm 0.39\,^{+0.17}_{-0.01}$. While the mass range considered by CMS extends to lower mass (4 GeV instead of 15 GeV), no events with $\langle m_{\ell\ell} \rangle < 16$ GeV are observed by CMS.

The results of both searches are also interpreted similarly. Shown in Figure A.2 are limits on the branching ratio and Higgs mixing parameter set by the CMS search. The branching ratio limit set by CMS is stronger by approximately one order of magnitude than the limit set by ATLAS (Figure 9.5). The Higgs mixing parameter limits are similar (Figure 9.7).

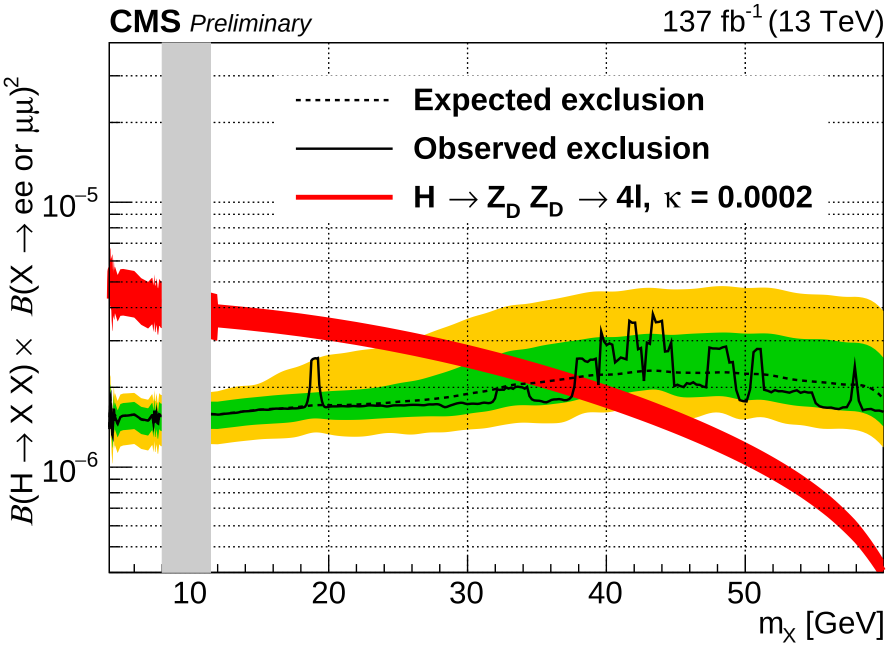


(a) $4e$ (b) 4μ

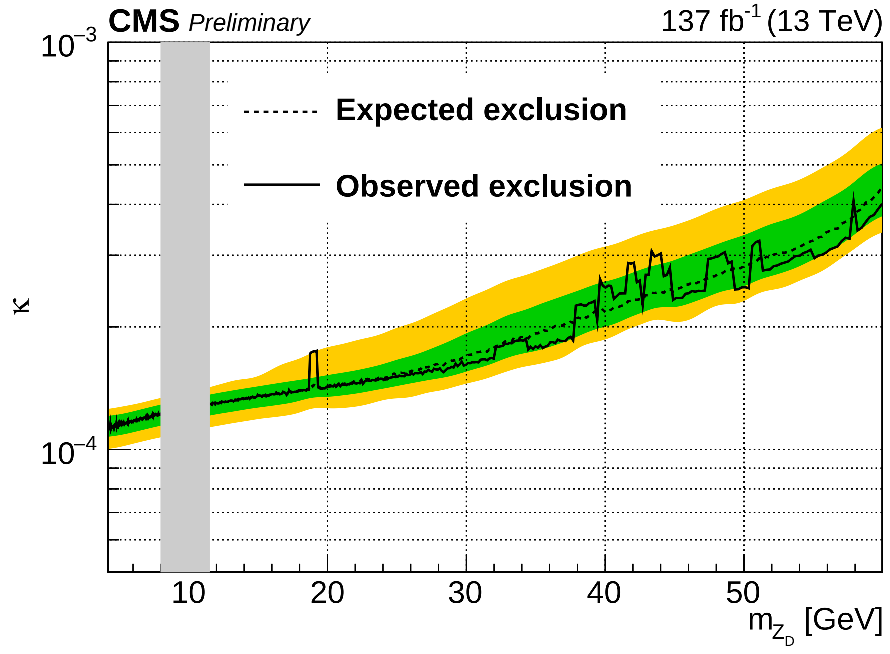


(c) $2e2\mu$

Figure A.1: Distributions of $(m_{Z1} + m_{Z2})/2 \Leftrightarrow \langle m_{\ell\ell} \rangle$ in the signal region of the CMS 137 fb^{-1} search for the $4e$, 4μ, and $2e2\mu$ channels [108].



(a) Limit on branching ratio



(b) Limit on Higgs mixing parameter

Figure A.2: 95% CL limits on the branching ratio $\mathcal{B}(H \to XX) \times \mathcal{B}(X \to ee \text{ or } \mu\mu)^2$ and Higgs mixing parameter κ [108].

Bibliography

[1] M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001 (cit. on pp. 1, 3, 4, 9, 23–25, 42, 51, 56, 58, 60, 67, 76, 101, 126).

[2] G. F. Giudice, *Naturally Speaking: The Naturalness Criterion and Physics at the LHC*, (2008) 155, ed. by G. Kane and A. Pierce, arXiv: `0801.2562 [hep-ph]` (cit. on p. 5).

[3] J. M. Cline, K. Kainulainen, and D. Tucker-Smith, *Electroweak baryogenesis from a dark sector*, Phys. Rev. D **95** (2017) 115006, arXiv: `1702.08909 [hep-ph]` (cit. on p. 5).

[4] M. Pospelov, *Secluded U(1) below the weak scale*, Phys. Rev. **D80** (2009) 095002, arXiv: `0811.1030 [hep-ph]` (cit. on p. 5).

[5] ATLAS Collaboration, *Search for new light gauge bosons in Higgs boson decays to four-lepton final states in pp collisions at* $\sqrt{s} = 8\,TeV$ *with the ATLAS detector at the LHC*, Phys. Rev. D **92** (2015) 092001, arXiv: `1505.07645 [hep-ex]` (cit. on pp. 5–7).

[6] ATLAS Collaboration, *Search for Higgs boson decays to beyond-the-Standard-Model light bosons in four-lepton events with the ATLAS detector at* $\sqrt{s} = 13\,TeV$, JHEP **06** (2018) 166, arXiv: `1802.03388 [hep-ex]` (cit. on pp. 5, 6, 12, 13, 16, 17, 40, 58).

[7] CMS Collaboration, *Search for light bosons in decays of the* $125\,GeV$ *Higgs boson in proton–proton collisions at* $\sqrt{s} = 8\,TeV$, JHEP **10** (2017) 076, arXiv: `1701.02032 [hep-ex]` (cit. on pp. 7, 13).

[8] ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at* $\sqrt{s} = 7$ *and* $8\,TeV$, JHEP **08** (2016) 045, arXiv: `1606.02266 [hep-ex]` (cit. on p. 8).

[9] D. Curtin et al., *Exotic decays of the 125 GeV Higgs boson*, Phys. Rev. **D90** (2014) 075004, arXiv: `1312.4992 [hep-ph]` (cit. on pp. 8, 9, 67, 131).

[10] D. Curtin, R. Essig, S. Gori, and J. Shelton, *Illuminating Dark Photons with High-Energy Colliders*, JHEP **02** (2015) 157, arXiv: `1412.0018 [hep-ph]` (cit. on pp. 8, 11, 12, 14–17, 60, 67, 131).

[11] J. Lees et al., *Search for a Dark Photon in e^+e^- Collisions at BaBar*, Phys. Rev. Lett. **113** (2014) 201801, arXiv: `1406.2980 [hep-ex]` (cit. on p. 11).

[12] M. Ablikim et al., *Dark Photon Search in the Mass Range Between 1.5 and 3.4 GeV/c^2*, Phys. Lett. B **774** (2017) 252, arXiv: `1705.04265 [hep-ex]` (cit. on p. 11).

[13] J. Lees et al., *Search for Invisible Decays of a Dark Photon Produced in e^+e^- Collisions at BaBar*, Phys. Rev. Lett. **119** (2017) 131804, arXiv: `1702.03327 [hep-ex]` (cit. on p. 11).

[14] R. Aaij et al., *Search for Dark Photons Produced in 13 TeV pp Collisions*, Phys. Rev. Lett. **120** (2018) 061801, arXiv: `1710.02867 [hep-ex]` (cit. on p. 12).

[15] R. Aaij et al., *Search for $A' \to \mu^+\mu^-$ Decays*, Phys. Rev. Lett. **124** (2020) 041801, arXiv: `1910.06926 [hep-ex]` (cit. on p. 12).

[16] P. Ilten, Y. Soreq, M. Williams, and W. Xue, *Serendipity in dark photon searches*, JHEP **06** (2018) 004, arXiv: `1801.04847 [hep-ph]` (cit. on p. 12).

[17] I. Hoenig, G. Samach, and D. Tucker-Smith, *Searching for dilepton resonances below the Z mass at the LHC*, Phys. Rev. D **90** (2014) 075016, arXiv: `1408.1075 [hep-ph]` (cit. on p. 12).

[18] J. M. Cline, G. Dupuis, Z. Liu, and W. Xue, *The windows for kinetically mixed Z'-mediated dark matter and the galactic center gamma ray excess*, JHEP **08** (2014) 131, arXiv: `1405.7691 [hep-ph]` (cit. on p. 12).

[19] CMS Collaboration, *A search for pair production of new light bosons decaying into muons*, Phys. Lett. B **752** (2016) 146, arXiv: `1506.00424 [hep-ex]` (cit. on p. 13).

[20] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001 (cit. on p. 18).

[21] ATLAS Collaboration, *Run-2 luminosity public results*, (2020), URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2` (cit. on pp. 20, 64).

[22] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 22, 23, 27, 28, 30, 32, 33, 50).

[23] ATLAS Collaboration, *ATLAS Inner Detector: Technical Design Report, 2*, ATLAS-TDR-5, 1997, URL: `https://cds.cern.ch/record/331064` (cit. on p. 23).

[24] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19, 2010, URL: `https://cds.cern.ch/record/1291633` (cit. on p. 23), Addendum: ATLAS-TDR-19-ADD-1, 2012, URL: `https://cds.cern.ch/record/1451888`.

[25] ATLAS Collaboration, *Operation and performance of the ATLAS semiconductor tracker*, JINST **9** (2014) P08009, arXiv: `1404.7473 [hep-ex]` (cit. on p. 23).

[26] B. Mindur, *ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider*, Nucl. Instrum. Meth. A **845** (2017) 257, ed. by G. Badurek et al. (cit. on p. 23).

[27] ATLAS Collaboration, *Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties*, JINST **12** (2017) P05002, arXiv: `1702.06473 [hep-ex]` (cit. on p. 23).

[28] A. Andronic and J. Wessels, *Transition radiation detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **666** (2012) 130, Advanced Instrumentation, ISSN: 0168-9002, URL: `http://www.sciencedirect.com/science/article/pii/S0168900211018134` (cit. on p. 24).

[29] V. A. Mitsou, "The ATLAS transition radiation tracker", *8th International Conference on Advanced Technology and Particle Physics (ICATPP 2003): Astroparticle, Particle, Space Physics, Detectors and Medical Physics Applications*, 2003 497, arXiv: `hep-ex/0311058` (cit. on p. 24).

[30] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: `1603.02934 [hep-ex]` (cit. on pp. 25, 42).

[31] ATLAS Collaboration, *Drift Time Measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using Cosmic Muons*, Eur. Phys. J. C **70** (2010) 755, arXiv: `1002.4189 [hep-ex]` (cit. on p. 27).

[32] ATLAS Collaboration, *ATLAS Muon Spectrometer: Technical Design Report*, ATLAS-TDR-10, 1997, URL: `https://cds.cern.ch/record/331068` (cit. on pp. 29, 32).

[33] G. Cattani, *The Resistive Plate Chambers of the ATLAS experiment: performance studies*, Journal of Physics: Conference Series **280** (2011) 012001, URL: `https://doi.org/10.1088%2F1742-6596%2F280%2F1%2F012001` (cit. on p. 31).

[34] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: `1603.05598 [hep-ex]` (cit. on pp. 31, 50–52, 54, 60).

[35] ATLAS Collaboration, *Trigger Menu in 2018*, ATL-DAQ-PUB-2019-001, 2019, URL: `https://cds.cern.ch/record/2693402` (cit. on p. 33).

[36] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, Eur. Phys. J. C **77** (2017) 332, arXiv: `1611.10235 [hep-ex]` (cit. on p. 40).

[37] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-026, 2015, URL: `https://cds.cern.ch/record/2037717` (cit. on p. 40).

[38] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: `1908.00005 [hep-ex]` (cit. on pp. 42–45, 47, 60).

[39] ATLAS Collaboration, *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*, ATL-PHYS-PUB-2017-022, 2017, URL: `https://cds.cern.ch/record/2298955` (cit. on p. 42).

[40] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: `1902.04655 [hep-ex]` (cit. on pp. 42, 45).

[41] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: `1704.07983 [hep-ex]` (cit. on p. 42).

[42] A. Rosenfeld and J. L. Pfaltz, *Sequential Operations in Digital Picture Processing*, J. ACM **13** (1966) 471, ISSN: 0004-5411, URL: `https://doi.org/10.1145/321356.321357` (cit. on p. 42).

[43] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, URL: `https://cds.cern.ch/record/1449796` (cit. on p. 43).

[44] R. Frühwirth, *A Gaussian-mixture approximation of the Bethe–Heitler model of electron energy loss by bremsstrahlung*, Computer Physics Communications **154** (2003) 131, ISSN: 0010-4655, URL: `http://www.sciencedirect.com/science/article/pii/S0010465503002923` (cit. on p. 43).

[45] G. Aad et al., *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv: `1407.5063 [hep-ex]` (cit. on p. 44).

[46] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data*, JINST **14** (2019) P03017, arXiv: `1812.03848 [hep-ex]` (cit. on p. 44).

[47] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. B **659** (2008) 119, arXiv: `0707.1378 [hep-ph]` (cit. on p. 48).

[48] J. Illingworth and J. Kittler, *A survey of the hough transform*, Computer Vision, Graphics, and Image Processing **44** (1988) 87, ISSN: 0734-189X, URL: `http://www.sciencedirect.com/science/article/pii/S0734189X88800331` (cit. on p. 50).

[49] ATLAS Collaboration, *Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **03** (2018) 095, arXiv: `1712.02304 [hep-ex]` (cit. on p. 58).

[50] ATLAS Collaboration, *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]` (cit. on p. 58).

[51] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: `1007.1727 [physics.data-an]` (cit. on pp. 58, 124), Erratum: Eur. Phys. J. C **73** (2013) 2501.

[52] R. Wang, Private communication, 2016 (cit. on p. 59).

[53] S. Gopalakrishna, S. Jung, and J. D. Wells, *Higgs boson decays to four fermions through an abelian hidden sector*, Phys. Rev. D **78** (2008) 055002, arXiv: `0801.3456 [hep-ph]` (cit. on p. 60).

[54] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: `1610.07922 [hep-ph]` (cit. on pp. 60, 67, 131).

[55] D. Boye, Private communication, 2019 (cit. on p. 62).

[56] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: `https://cds.cern.ch/record/2677054` (cit. on p. 64).

[57] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017 (cit. on p. 64).

[58] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, *NNLOPS simulation of Higgs boson production*, JHEP **10** (2013) 222, arXiv: `1309.0017 [hep-ph]` (cit. on pp. 65, 73).

[59] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: `1405.0301 [hep-ph]` (cit. on pp. 66, 73).

[60] R. D. Ball et al., *Parton distributions with QED corrections*, Nucl. Phys. B **877** (2013) 290, arXiv: `1308.0598 [hep-ph]` (cit. on p. 66).

[61] S. Agostinelli et al., GEANT4 – *a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 66).

[62] ATLAS Collaboration, *The new Fast Calorimeter Simulation in ATLAS*, ATL-SOFT-PUB-2018-002, 2018, URL: `https://cds.cern.ch/record/2630434` (cit. on p. 66).

[63] W. Buttinger, Private communication, 2019 (cit. on p. 69).

[64] ATLAS Collaboration, *Proposal for particle-level object and observable definitions for use in physics measurements at the LHC*, ATL-PHYS-PUB-2015-013, 2015, URL: `https://cds.cern.ch/record/2022743` (cit. on p. 69).

[65] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: `hep-ph/0409146` (cit. on p. 73).

[66] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: `0709.2092 [hep-ph]` (cit. on p. 73).

[67] 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]` (cit. on p. 73).

[68] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: `1510.03865 [hep-ph]` (cit. on p. 73).

[69] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09** (2014) 145, arXiv: `1406.3660 [hep-ex]` (cit. on p. 73).

[70] J. Pumplin et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **07** (2002) 012, arXiv: `hep-ph/0201195` (cit. on p. 73).

[71] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, JHEP **05** (2016) 058, arXiv: `1602.00695 [hep-ph]` (cit. on p. 73).

[72] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv: `1007.2241 [hep-ph]` (cit. on p. 73).

[73] M. Ciccolini, A. Denner, and S. Dittmaier, *Strong and Electroweak Corrections to the Production of Higgs + 2 Jets via Weak Interactions at the Large Hadron Collider*, Phys. Rev. Lett. **99** (2007) 161803, arXiv: `0707.0381 [hep-ph]` (cit. on p. 73).

[74] M. Ciccolini, A. Denner, and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC*, Phys. Rev. D **77** (2008) 013002, arXiv: `0710.4749 [hep-ph]` (cit. on p. 73).

[75] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, *Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD*, Phys. Rev. Lett. **105** (2010) 011801, arXiv: `1003.4451 [hep-ph]` (cit. on p. 73).

[76] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: `0710.3820 [hep-ph]` (cit. on p. 73).

[77] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: `1207.1303 [hep-ph]` (cit. on p. 73).

[78] ATLAS Collaboration, *ATLAS Pythia 8 tunes to* 7 *TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: `https://cds.cern.ch/record/1966419` (cit. on p. 73).

[79] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: `1410.8849 [hep-ph]` (cit. on p. 73).

[80] M. Wiesemann et al., *Higgs production in association with bottom quarks*, JHEP **02** (2015) 132, arXiv: `1409.5301 [hep-ph]` (cit. on p. 73).

[81] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: `0811.4622 [hep-ph]` (cit. on p. 73).

[82] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: `0808.3674 [hep-ph]` (cit. on p. 73).

[83] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: `1111.5206 [hep-ph]` (cit. on p. 73).

[84] B. Biedermann, A. Denner, S. Dittmaier, L. Hofer, and B. Jäger, *Electroweak corrections to* $pp \to \mu^+\mu^-e^+e^- + X$ *at the LHC: a Higgs background study*, Phys. Rev. Lett. **116** (2016) 161803, arXiv: `1601.07787 [hep-ph]` (cit. on p. 73).

[85] B. Biedermann, A. Denner, S. Dittmaier, L. Hofer, and B. Jager, *Next-to-leading-order electroweak corrections to the production of four charged leptons at the LHC*, JHEP **01** (2017) 033, arXiv: `1611.05338 [hep-ph]` (cit. on p. 73).

[86] F. Caola, K. Melnikov, R. Röntsch, and L. Tancredi, *QCD corrections to ZZ production in gluon fusion at the LHC*, Phys. Rev. D **92** (2015) 094028, arXiv: `1509.06734 [hep-ph]` (cit. on p. 73).

[87] S. Frixione, V. Hirschi, D. Pagani, H. .-.-S. Shao, and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, JHEP **06** (2015) 184, arXiv: `1504.03446 [hep-ph]` (cit. on p. 73).

[88] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. D **69** (2004) 094008, arXiv: `hep-ph/0312266` (cit. on p. 73).

[89] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through* $O(\alpha_S^4)$, Phys. Rev. Lett. **110** (2013) 252004, arXiv: `1303.6254 [hep-ph]` (cit. on p. 73).

[90] ATLAS Collaboration, *Multi-boson simulation for* 13 *TeV ATLAS analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: `https://cds.cern.ch/record/2119986` (cit. on p. 73).

[91] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189, arXiv: `0901.0002 [hep-ph]` (cit. on p. 76).

[92] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012, URL: `https://cds.cern.ch/record/1474107` (cit. on p. 76).

[93] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, CERN-2011-002 (CERN, Geneva, 2011), arXiv: `1101.0593 [hep-ph]` (cit. on pp. 80, 131).

[94] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, CERN-2012-002 (CERN, Geneva, 2012), arXiv: `1201.3084 [hep-ph]` (cit. on pp. 80, 131).

[95] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, CERN-2013-004 (CERN, Geneva, 2013), arXiv: `1307.1347 [hep-ph]` (cit. on pp. 80, 131).

[96] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, CERN Yellow Reports: Monographs, 869 pages, 295 figures, 248 tables and 1645 citations. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG, 2016, URL: `https://cds.cern.ch/record/2227475` (cit. on p. 80).

[97] ATLAS Collaboration, *$ZZ \to \ell^+\ell^-\ell'^+\ell'^-$ cross-section measurements and search for anomalous triple gauge couplings in 13 TeV pp collisions with the ATLAS detector*, Phys. Rev. D **97** (2018) 032005, arXiv: `1709.07703 [hep-ex]` (cit. on p. 80).

[98] M. Lefebvre, Private communication, 2019 (cit. on p. 87).

[99] *Formulae for Estimating Significance*, tech. rep. ATL-PHYS-PUB-2020-025, CERN, 2020, URL: `https://cds.cern.ch/record/2736148` (cit. on p. 101).

[100] 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., 2012, URL: `https://cds.cern.ch/record/1456844` (cit. on pp. 123, 130).

[101] A. Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, Transactions of the American Mathematical Society **54** (1943) 426, ISSN: 00029947, URL: `http://www.jstor.org/stable/1990256` (cit. on p. 124).

[102] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Ann. Math. Statist. **9** (1938) 60, URL: `https://doi.org/10.1214/aoms/1177732360` (cit. on p. 124).

[103] A. L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693 (cit. on p. 128).

[104] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke, *Interpolation between multi-dimensional histograms using a new non-linear moment morphing method*, Nucl. Instrum. Meth. A **771** (2015) 39, arXiv: `1410.7388 [physics.data-an]` (cit. on p. 130).

[105] M. Baak et al., *HistFitter software framework for statistical data analysis*, Eur. Phys. J. C **75** (2015) 153, arXiv: `1410.1280 [hep-ex]` (cit. on p. 130).

[106] L. Moneta et al., *The RooStats Project*, 2010, arXiv: `1009.1003 [physics.data-an]` (cit. on p. 130).

[107] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: `physics/0306116 [physics.data-an]` (cit. on p. 130).

[108] *Search for a low-mass dilepton resonance in Higgs boson decays to four-lepton final states at* $\sqrt{s} = 13$ TeV, tech. rep. CMS-PAS-HIG-19-007, CERN, 2020, URL: `https://cds.cern.ch/record/2718976` (cit. on pp. 143, 145, 146).