DIRECT SEARCH FOR THE HIGGS BOSON TO CHARM COUPLING WITH THE ATLAS EXPERIMENT

António Manuel Mendes Jacques da Costa

Thesis submitted for the degree of
Doctor of Philosophy

UNIVERSITY OF BIRMINGHAM

Particle Physics Group,
School of Physics and Astronomy,
University of Birmingham

February 2023

ABSTRACT

A direct search for the Standard Model Higgs boson decays to a pair of charm quarks is presented in this thesis, probing the Higgs boson Yukawa couplings to the second generation of fermions. This analysis makes use of the full LHC Run 2 dataset collected with the ATLAS detector, corresponding to an integrated luminosity of 139 fb^{-1} of proton-proton collisions at a centre-of-mass energy of 13 TeV. An extrapolation of this search to the High-Luminosity LHC conditions - a centre-of-mass energy of 14 TeV, an integrated luminosity of 3000 fb^{-1} - is also presented. The production of the Higgs boson in association with a W or Z boson is targeted, where only leptonic decays of the W or Z boson are considered. Both charm and bottom jet tagging algorithms are used to identify the signature of the Higgs boson decay to charm quarks, while reducing contamination from Higgs boson decays to bottom quarks.

The full Run 2 search improves the constraint on the cross-section times branching fraction for a Higgs boson with a mass of 125 GeV previously presented by ATLAS, using an integrated luminosity of 36 fb^{-1} of proton-proton collisions at the same centre-of-mass energy. A direct limit on the charm Yukawa coupling modifier from Higgs boson to charm quark decays is also set. The HL-LHC extrapolation of this search has the potential to further improve the constraints.

The ATLAS detector is being upgraded for the next years of operation, with the new Level 1 calorimeter sub-trigger system presented in this thesis. The development of the bitwise simulation, which reproduces the behaviour of the trigger's firmware, and the monitoring framework are reported in this context, focusing in particular in the e/γ trigger candidates.

DECLARATION OF AUTHOR CONTRIBUTIONS

The author contributed to the validation and development of the bitwise simulation (mimicking the behaviour of the firmware in the trigger electronic boards) framework of the ATLAS Run 3 upgraded Level 1 calorimeter sub-trigger system, with particular emphasis on the e/γ Feature Extractor algorithm. This work included the implementation of the energy encoding scheme, and the study and implementation of the seed finder algorithm and e/γ isolation variables and associated working points setting (with the exception of the w_{stot} variable which was studied and implemented by a collaborator). The author contributed to the development of the ATLAS monitoring framework for the upgraded Level 1 calorimeter system, implementing the plotting of the e/γ and τ-lepton histograms resulting of the Feature Extractor algorithm. In addition the author performed Level 1 calorimeter on-call shifts at CERN, ensuring the correct functioning of the trigger test facility.

The author was part of the 0-lepton channel team in the $VH(H \to c\bar{c})$ ATLAS full Run 2 analysis, contributing to the analysis framework development and background modelling. The latter work was particularly focused on the W/Z+jets treatment, with the derivation of the modelling uncertainties (with the exceptions of the channel and signal region/control region extrapolations; the author contributed only to the initial derivation stage of the shape uncertainties). The author undertook additionally studies on the flavour splitting of the associated jets of the Z+jets background. The author contributed to the 0-lepton event selection optimisation, in particular to the ΔR requirements, and developed and implemented the 0 c-tag control region in the 2-lepton channel. The author worked also on the statistical analysis, implementing, understanding and improving the fit model and associated behaviour, with focus on the 0-lepton individual and combined lepton channels fits. The author performed also fit studies for the extrapolation of this analysis to the HL-LHC scenario.

ACKNOWLEDGEMENTS

This PhD was a journey during troubled times in the world, but in which I grew as a person, formed new friendships and embarked on new physics projects.

I must start by thanking my supervisors, Alan Watson and Paul Thompson, who were always available to help me throughout the years. I'll remember fondly the stories and lessons from the conversations with Alan; Paul made me see the brighter side of more frustrating occasions, and always gave me much appreciated help.

I want to thank my viva examiners, Prof. David Evans and Prof. Pedro Teixeira-Dias, for the discussions, as well as for the comments on the thesis. I also thank Evgueni Goudzovski for the smooth chairing of the viva.

I am thankful to all the people in the Higgs group at Birmingham during these years, for all their support and help. Kostas Nikolopoulos helped me grow as both a physicist and a person, and provided me useful life advices in addition to guidance in the physics path. I am very grateful to Andy Chisholm, which contributed to the supervision of my analysis work, offering valuable help and support, and for all the discussions we had, particularly in the post-pandemic period. I want to thank Tom Neep, who was always available to thoroughly discuss my questions, and Panagiotis Bellos, for our interesting conversations in a stressful period of the PhD.

The people at the particle physics group in Birmingham provided me with the best of the environments, with a particular thanks to Dave Charlton, Paul Newman and Maria O'Connor for their managing of the group, and to all the people that contributed with their experience and suggestions to the occasional questions. Robert Ward, Tom Bache, Kalun Bedingfield, Naomi Cooke and Alexis Stabekis gave me the best welcome I could hope for, and I am happy for the the friendship we developed across the years and how we helped each other. In particular without you the pandemic years would have been much harder to bear.

The students in and around the Birmingham ATLAS office changed throughout the years, but we always managed to have a joyful environment. I would like to thank in particular Elliot Reynolds and James Kendrick for their initial welcome and support, and Adrien Auriol, Júlia Silva and Mihaela Marinescu for all the

enjoyable moments we shared together. I am also thankful to Patrick Knights and Enrico Andriolo for their support.

I remember with happiness the coffee breaks, journeys and discussions I had with Paulo Gomes, for which I thank him. I want to thank as well Ana Évora, Carlos Silva and Inês Barroso for the great times.

One of the key outputs of this PhD was the $VH(H \to c\bar{c})$ analysis, which was the result of the amazing work and dedication of all the team I had the privilege of being a part of. For this I am thankful to Hannah Arnold, Tristan du Pree, Guy Koren, Hadar Cohen, Maria Mironova, Marko Stamenkovic, Jonathan Shlomi, Zijun Xu and Tao Wang. Hannah in particular often provided me advice and support, especially in the fit studies, helping me navigate a new topic.

My time in Switzerland was made much happier by the shared moments with João Ferreira, Ana Luísa Carvalho and Carolina Amoedo, and by Harry Cooke, Gareth Bird and many other LTA students, which made for some very enjoyable six months for which I am very grateful.

I am thankful to Rhys Owen and Silvia Franchino, which in my time at CERN offered me their knowledge and availability to accompany them in technical tasks from which I benefited immensely, and to Damir Rassloff, for all the help during the L1Calo on-call shifts.

My gratitude goes also to all of the L1Calo offline software team, with particular thanks to Alison Elliot and Jacob Kempster for all the discussions and support, which made me very much enjoy working with them. I am also thankful to the whole L1Calo group for all these years of collaboration, in particular to Hans-Christian Schultz-Coulon, Juraj Bracinik and Steve Hillier for their support. Francesco Gonnella helped me not only in the L1Calo work but also in providing me with much needed accommodation in the last months of the PhD, becoming during the PhD a friend for which I am grateful.

I would like to thank also Duarte Drumond, Pedro Gomes, Pedro Pinto and Ricardo Gonçalo, for all the times in which they had the patience to listen to me and have a chat, despite the busy lives we all lead.

I am fortunate to have had Margarida Cordeiro by my side during these years, and I could not be happier and prouder of the lives we are building together and how we make each other grow. A realidade é que sem ela não há paz, não há beleza.

I finish with a profound gratitude and love for my parents and sister, which have always been supporting me either at home or when I am away. Um muito obrigado do fundo do coração.

Contents

1 Introduction **1**

2 High Energy Physics **5**

2.1 Standard Model of Particle Physics **6**

2.2 Proton-proton collisions and hadronic products **14**

2.3 Higgs boson Production and Couplings **20**

2.4 Kappa Framework **23**

2.5 Searches for the Higgs-charm coupling **27**

2.5.1 Exclusive $H \to J/\psi\,\gamma$ decays **28**

2.5.2 Charm quark initiated Higgs boson production **29**

2.5.3 Inclusive $H \to c\bar{c}$ decays **30**

2.6 Beyond the Standard Model **33**

3 The ATLAS Experiment **36**

3.1 The Large Hadron Collider at CERN **37**

3.2 The ATLAS Detector **40**

3.2.1 Inner Detector **43**

3.2.2 Calorimeters **45**

3.2.3 Muon Spectrometer **50**

3.3 ATLAS Run 3 Level-1 Calorimeter Trigger Upgrade **52**

3.3.1 Bitwise Simulation **57**

3.3.2 Monitoring **67**

3.4 The ATLAS detector at the High-Luminosity Large Hadron Collider **71**

4 Physics Objects **73**

4.1 Tracks and Interaction vertices . . . **73**

4.2 Electrons . . . **74**

4.3 Muons . . . **76**

4.4 Jets . . . **77**

4.5 Missing Transverse Energy . . . **78**

4.6 Overlap Removal . . . **80**

4.7 Heavy Flavour Hadron Tagging . . . **81**

5 Search for Higgs Boson Decays to Charm Quarks **87**

5.1 Introduction . . . **88**

5.2 Data and Simulation Samples . . . **92**

5.3 Object Selection . . . **96**

5.4 Event Selection . . . **98**

5.4.1 0-lepton channel Signal Region Event Selection . . . **101**

5.4.2 1-lepton channel Signal Region Event Selection . . . **101**

5.4.3 2-lepton channel Signal Region Event Selection . . . **102**

5.4.4 Control Regions . . . **104**

5.5 Background Composition . . . **105**

5.6 Systematic Uncertainties . . . **108**

5.6.1 Experimental Uncertainties . . . **109**

5.6.2 Signal and Background Modelling . . . **112**

5.7 Statistical Analysis . . . **121**

5.7.1 Likelihood Definition . . . **121**

5.7.2 Fit Model . . . **124**

5.8 Results . . . **128**

5.8.1 Signal Strengths and Significances . . . **129**

5.8.2 Limit on $\mu_{VH(H\to c\bar{c})}$ . . . **130**

5.8.3 Post-fit $m_{c\bar{c}}$ distributions **132**

5.8.4 Breakdown of uncertainties **135**

5.8.5 Kappa Interpretation **138**

5.9 Combination of $VH(H \to c\bar{c})$ and $VH(H \to b\bar{b})$ analyses **140**

5.10 ATLAS sensitivity to $H \to b\bar{b}$ and $H \to c\bar{c}$ decays in VH production at the HL-LHC **144**

6 Conclusion **150**

References **153**

A Post-fit m_{cc} distributions of analysis signal regions **163**

List of Tables

2.1 Isospin I_3, hypercharge Y and electric charge Q for fermions. 9

2.2 SM cross-sections for main production modes of a Higgs boson with $m_H = 125$ GeV in pp collisions [1]. 22

2.3 SM branching ratios for a Higgs boson with $m_H = 125$ GeV [19, 20]. 23

2.4 Fit results for κ_Z, κ_W, κ_b, κ_t, κ_τ and κ_μ [29]. 26

5.1 Thresholds for each data taking period of the triggers used in the analysis. 93

5.2 Signal and background processes and their corresponding MC generators used in the analysis. NNLO and ((N)NLL stand for next-to-next-to-leading order and ((next-to-)next-to-)leading log, respectively [57]. 94

5.3 Summary of event categorisation of signal sensitive regions. 100

5.4 Summary of the signal region event selection in the 0-, 1- and 2-lepton channels. Jet1 and jet2 refer to the two *signal* jets and H refers to the jet1–jet2 system [57]. 103

5.5 Summary of the experimental systematic uncertainties applied in the $VH(H \to c\bar{c})$ analysis. 109

5.6 Systematic uncertainties on the $VH(H \to c\bar{c})$ signal processes. . . . 115

5.7 Summary of the background modelling systematic uncertainties considered. CR and SR stand for control region and signal region [57]. 120

5.8 Summary of the signal and control regions used in the analysis. ΔR_{SR}^{cut} is the upper cut of $\Delta R(\text{jet1}, \text{jet2})$ based on the signal optimisation described in Section 5.4. 124

5.9 Expected and observed significances for the $VW(W \to cq)$ and $VZ(Z \to c\bar{c})$ processes. 130

5.10 Comparison between 36.1 fb^{-1} and 139 fb^{-1} ATLAS analyses. . . . 131

5.11 Breakdown of absolute contributions to the uncertainty in the fitted values of $\mu_{VH(H\to c\bar{c})}$, $\mu_{VW(W\to cq)}$ and $\mu_{VZ(Z\to c\bar{c})}$. The sum in quadrature of uncertainties from different sources may differ from the total due to correlations [57]. 137

5.12 Main differences between the $VH(H \to c\bar{c})$ and $VH(H \to b\bar{b})$ analyses. 141

5.13 Centre-of-mass scale factors from 13 to 14 TeV for signal and background expected yields [65]. 145

5.14 Scale factors applied to systematic uncertainties to reflect expected improvements in the HL-LHC scenario. A scale factor of 1 represents no improvement foreseen [65]. 146

5.15 Breakdown of contributions to the uncertainty in the fitted values of $\mu_{VH(H\to c\bar{c})}$ from a fit to a simulated dataset scaled to $\sqrt{s} = 14$ TeV and an integrated luminosity of 3000 fb^{-1}. The sum in quadrature of different sources of uncertainty may differ from the total due to correlations [65]. 148

List of Figures

1.1 Elementary particles. 1

1.2 Range of observed fermion masses. 2

2.1 The Standard Model particles [4]. 7

2.2 Schematic of a proton-proton collision. The hard process is represented in red, the parton shower in blue, hadron production in light green blobs and their decay products in dark green. Gluon radiation in the (red) gluons leading to the hard-scatter process correspond to initial state radiation, while gluon emission in the (red) hard-scattering process represent final state radiation. The underlying event is shown in purple [16]. 17

2.3 Schematic diagram of heavy flavour hadron decay [17] 18

2.4 Anatomy of a b-jet. 19

2.5 Anatomy of a c-jet. 20

2.6 Anatomy of a light flavour jet. 20

2.7 Feynman diagrams for the main Higgs boson production modes. . 21

2.8 SM cross-sections for Higgs boson production modes as a function of center-of-mass energy, with $m_H = 125$ GeV [19]. 22

2.9 SM branching ratios for Higgs decays as a function of its mass [19, 20]. 23

2.10 Example of Higgs coupling modifiers in a $H \to c\bar{c}$ process. 24

2.11 Best-fit values and uncertainties for Higgs boson coupling modifiers per particle type with effective photon, $Z\gamma$ and gluon couplings and either $B_{\text{i.}} = B_{\text{u.}} = 0$ (left), or $B_{\text{i.}}$ and $B_{\text{u.}}$ included as free parameters in the model (right) [29]. 27
2.12 $H \to J/\psi\,\gamma$ decay. Q is a vector bound state of quarks and the hatched circle represents a set of one-loop diagrams. 28
2.13 Higgs boson production involving Higgs-charm coupling. 29
2.14 $H \to c\bar{c}$ decay in VH production. 30
2.15 Representative V+jets processes. 32
2.16 Representative top processes. 32
2.17 Representative diboson processes. 33

3.1 Schematic view of the CERN accelerator complex [2]. 37
3.2 The ATLAS detector [2]. 40
3.3 The ATLAS inner detector barrel region [45]. 44
3.4 The ATLAS Inner Detector end-cap regions [2] before the IBL installation. 44
3.5 The ATLAS calorimeters [2]. 47
3.6 A module of the LAr electromagnetic calorimeter in the barrel region [2]. 48
3.7 The ATLAS muon system [2]. 51
3.8 Overview of the Level-1 trigger system planned for Run 3 [47]. . . . 53
3.9 The trigger granularity from each 0.1×0.1 trigger tower after the upgrade of the electromagnetic calorimeter electronics, resulting in ten supercells. [47]. 54
3.10 3×3 (in η and ϕ) tower window in layer-2 for e/γ algorithm. Core (environment) supercells represented in orange (green). 55

3.11 Seed finder conditions for each of the four supercells in the central tower. Seed candidate represented in blue. 55
3.12 3×2 (in η and ϕ) supercell cluster in layer-2. Core supercells represented in orange and environment supercells in green. Seed candidate represented in blue. 56
3.13 Total TOB transverse energy. Core supercells represented in orange and environment supercells in green. Seed candidate represented in blue. 56
3.14 eFEX energy encoding scheme. 59
3.15 The eFEX R_η cluster (yellow) and environment (blue) in layer-2, where the seed cell **A** is the biggest local maximum in η and the one above **B** is the biggest neighbor in ϕ (up or down) of the seed **A**. 60
3.16 R_η distribution for a simulated signal sample of a Z boson decaying to a pair of electrons. 61
3.17 R_η distribution for a simulated background sample of multi jets. . . 61
3.18 $1 - R_\eta$ ratio as a function of C_η. 62
3.19 Cell selection for the R_{had} hadronic condition. 63
3.20 R_{had} distribution for a simulated signal sample of a Z boson decaying to a pair of electrons. 63
3.21 R_{had} distribution for a simulated background sample of multi jets. . 63
3.22 D_{had} parameter as a function of the R_{had} parameter. 64
3.23 The w_{stot} isolation condition. 66
3.24 Selection of e/γ TOB monitored variables. From left to right, and top to bottom: E_T, η, ϕ, and strictest working point passed for R_η, R_{had} and w_{stot}. 69
3.25 Selection of τ-lepton TOB monitored variables. From left to right, and top to bottom: E_T, η, ϕ, and strictest working point passed for r_{core} and r_{had}. 70

4.1 Resolution of the $m_{c\bar{c}}$ distribution with the standard jet calibration and with the additional muon-in-jet correction, for a simulated signal sample in the analysis [57, 58]. 79
4.2 Jet-tagging efficiencies in data for different flavoured jets as a function of the jet transverse momentum. The uncertainty band corresponds to the total uncertainty associated with the c-tagging efficiency calibration in data [57, 58]. 83
4.3 Data-to-simulation calibration scale factors for each jet flavour [57, 58]. 84

5.1 Background composition in simulation of the 0-lepton channel signal and control regions [57]. 106
5.2 Background composition in simulation of the 1-lepton channel signal and control regions [57]. 107
5.3 Background composition in simulation of the 2-lepton channel signal and control regions [57]. 108
5.4 Z+jets flavour fractions from fit of nominal to alternative MC. . . . 117
5.5 Treatment of a jet energy scale systematic uncertainty. 128
5.6 Observed and expected 95% CL upper limits on $\mu_{VH(H\to c\bar{c})}$. The individual channel limits come from a fit with five POIs: two diboson signals and one $VH(H \to c\bar{c})$ POI per channel [57]. 130
5.7 Examples of post-fit $m_{cc\bar{c}}$ distributions for the signal regions with 2 jets and $p_{\mathrm{T}}^{V} > 150$ GeV [57]. 133
5.8 Background-subtracted post-fit $m_{c\bar{c}}$ distributions for the sum of all signal regions in all channels with one or two c-tags categories [57]. 134
5.9 Parametrisation of the $VH(H \to c\bar{c})$ signal strength as a function of $|\kappa_c|$ [57, 58]. 139

5.10 Observed and expected constraints from a profile likelihood scan on $|\kappa_c|$ at 95% CL [57]. 139

5.11 Observed and expected 68% and 95% CL contours and respective best fit values for $\mu_{VH(H\to c\bar{c})}$ and $\mu_{VH(H\to b\bar{b})}$ [57]. 142

5.12 Observed and expected constraints on κ_c and κ_b at 68% and 95% CL [57]. 143

5.13 Observed and expected constraints from a profile likelihood scan on $|\kappa_c/\kappa_b|$ at 95% CL, where κ_b is a free parameter. A scenario with an equal coupling of the Higgs boson to charm and bottom quarks is represented by green lines, corresponding to the ratio $|\kappa_c/\kappa_b| = m_b/m_c$ [57]. 144

5.14 Expected limits on $\mu_{VH(H\to c\bar{c})}$ and constraints on κ_c for an extrapolation of the Run 2 analysis to the HL-LHC scenario with $\sqrt{s} = 14$ TeV and a total integrated luminosity of 3000 fb^{-1}. The individual channel likelihoods (limits) come from a fit with three (five) POIs: one $VH(H \to c\bar{c})$ POI per channel (and two diboson signals) [65]. . 147

5.15 Expected constraints on κ_c and κ_b and κ_c/κ_b for an extrapolation of the Run 2 analysis to the HL-LHC scenario with $\sqrt{s} = 14$ TeV and a total integrated luminosity of 3000 fb^{-1} [65]. 149

A.1 Post-fit distributions of the four 0-lepton signal regions. The total signal-plus-background prediction is shown by the solid black line and includes the $H \to c\bar{c}$ signal scaled to the best-fit value of $\mu_{VH(c\bar{c})} = -9$. The $H \to c\bar{c}$ signal is also shown as an unfilled histogram scaled to 300 times the SM prediction. The post-fit uncertainty is shown as the hatched background including correlations between uncertainties. The ratio of the data to the sum of the post-fit signal plus background is shown in the lower panel [57]. 164

A.2 Post-fit distributions of the four 1-lepton signal regions. The total signal-plus-background prediction is shown by the solid black line and includes the $H \to c\bar{c}$ signal scaled to the best-fit value of $\mu_{VH(c\bar{c})} = -9$. The $H \to c\bar{c}$ signal is also shown as an unfilled histogram scaled to 300 times the SM prediction. The post-fit uncertainty is shown as the hatched background including correlations between uncertainties. The ratio of the data to the sum of the post-fit signal plus background is shown in the lower panel [57]. 165

A.3 Post-fit distributions of the eight 2-lepton signal regions. The total signal-plus-background prediction is shown by the solid black line and includes the $H \to c\bar{c}$ signal scaled to the best-fit value of $\mu_{VH(c\bar{c})} = -9$. The $H \to c\bar{c}$ signal is also shown as an unfilled histogram scaled to 300 times the SM prediction. The post-fit uncertainty is shown as the hatched background including correlations between uncertainties. The ratio of the data to the sum of the post-fit signal plus background is shown in the lower panel [57]. 166

Definitions of acronyms

ATLAS A Toroidal LHC Apparatus

BDT Boosted Decision Tree

BR Branching ratio

BSM Beyond Standard Model

CERN European Organization for Nuclear Research

CL Confidence level

CR Control region

CMS Compact Muon Solenoid

DT Direct tagging

EM Electromagnetic

eFEX e/γ Feature extractor

HL-LHC High-Luminosity Large Hadron Collider

HLT High Level Trigger

MC Monte Carlo simulation

LHC Large Hadron Collider

L1Calo Level 1 calorimeter sub-trigger

LO Leading Order

NN Neural network

POI Parameter of interest

PV Primary vertex

QCD Quantum Chromodynamics

SCT Semiconductor Tracker

SM Standard Model

TOB Trigger object

TRT Transition Radiation Tracker

TT Truth flavour tagging

SR Signal region

WP Working point

Chapter 1

Introduction

The known elementary particles in the universe can be divided in two main groups, the bosons and the fermions, which are pictorially represented in Figure 1.1.

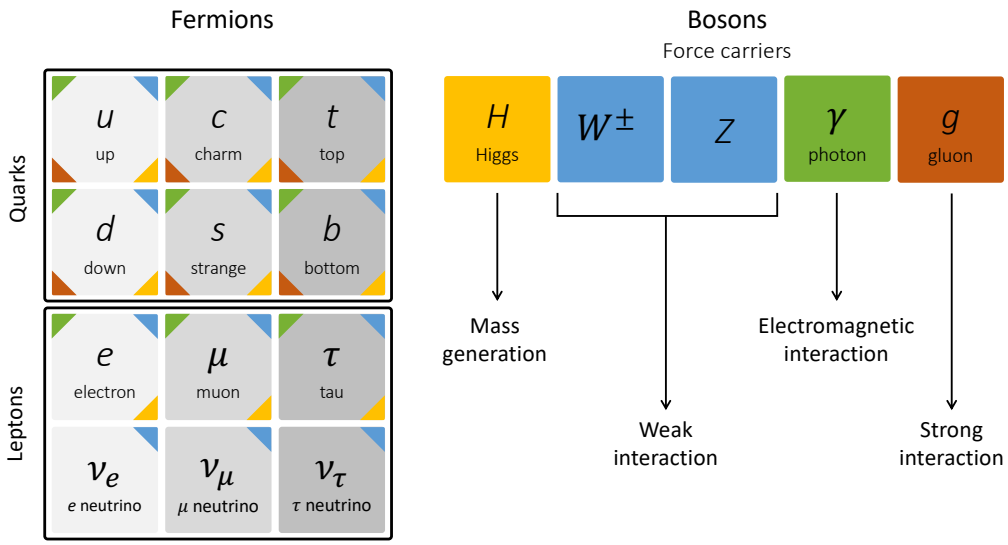


Figure 1.1: Elementary particles.

Bosons are the force carriers of the different interactions as illustrated in the fig-

ure. These interactions are described by the Standard Model of particle physics (SM), later presented in Section 2.1, with the Higgs boson in particular being responsible for the mass generation of the elementary particles.

Fermions comprise quarks and leptons, with each having three generations of particles. The ever increasing tones of grey of the fermion boxes in the figure represent the associated increased mass of each generation. The fermion masses [1] are distributed across several orders of magnitude, as can be seen in Figure 1.2, following no apparent pattern.

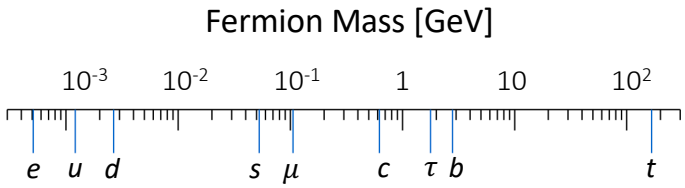


Figure 1.2: Range of observed fermion masses.

The Higgs-fermion interactions were so far observed experimentally for top quarks, bottom quarks and τ-leptons, and evidence was found for the Higgs-muon coupling. These SM interactions do not explain however the observed pattern of fermions masses, highlighting the need to measure all Higgs couplings to fermions. This endeavour would allow to verify the SM Higgs-fermion prediction or potentially find deviations hinting for Standard Model extended theories.

The Higgs-charm coupling constitutes the next step in this measurement effort, with the Higgs decay to charm quarks providing the most direct probe of this coupling. To this effect data are analysed to search for this process, using events collected at the ATLAS experiment [2] in the Large Hadron Collider (LHC) [3],

the world's largest and most powerful particle accelerator located at CERN.

Protons in the LHC travel at almost the speed of light, colliding inside (among others) the ATLAS detector, where Higgs bosons are created in the collision and decay into particles that leave specific signatures in the detector. These signatures allow for the reconstruction of the particles in dedicated systems in the detector, enabling the development of analyses targeting the Higgs boson decays, in particular in this case to charm quarks.

Not all collision events can, however, be stored in ATLAS due to limited storage capacity (among other technical reasons), which led to the development of a trigger system to filter the most promising events. This trigger system comprises a fast-response hardware-based (Level 1) component and a more comprehensive software-based (high-level trigger) component.

Upgrades to some of the ATLAS dedicated systems are taking place in order to maintain and improve where possible the experimental performance. One of such systems is the Level 1 calorimeter (L1Calo) sub-trigger, which will use more detailed input information and new hardware and firmware to find trigger candidates for electrons, photons and other physics objects. A bitwise framework was developed in the context of the L1Calo upgrade, mimicking the firmware behaviour, and allowing for the monitoring and study of the trigger performance. The upgraded L1Calo system had also to be reflected in the ATLAS monitoring framework, which allows for the detection of problems in the software or hardware when collecting collision data.

This thesis describes the latest ATLAS search for the Higgs decay to charm quarks, covering the relevant experimental and theoretical context. Chapter 2 of this thesis presents an overview of the Standard Model of particle physics and possible extensions, with a focus on the Higgs boson couplings to fermions. A description of proton-proton collisions, essential to the study of the Standard Model and the Higgs boson is also included. Chapter 3 describes the LHC and the ATLAS experiment, covering also the work developed in the context of the L1Calo trigger upgrade, namely in the bitwise and monitoring frameworks. Chapter 4 presents the techniques used to reconstruct the different physics objects used in analyses. Chapter 5 is dedicated to the search for the Higgs-charm coupling in ATLAS, covering all the methods and results. Conclusions are presented in Chapter 6.

Chapter 2

High Energy Physics

The Standard Model of particle physics encodes the current understanding of the known elementary particles in nature, and describes their interactions, which are governed by three of the four known fundamental forces: strong, weak and electromagnetic. The Standard Model does not describe the gravitational force. A central part of the Standard Model is the Higgs mechanism, which leads to the generation of the masses of the massive fundamental particles in the Standard Model, with the potential exception of neutrinos, where the masses are not yet understood.

This chapter presents an overview of the Standard Model, with particular emphasis on the Higgs theory and phenomenology. A description of proton-proton collisions, which are used in studies for the properties of the Higgs boson, is also included in this chapter. The framework used in the Higgs studies for improved knowledge of the Higgs couplings to other particles is later introduced in the

chapter, followed by an overview of the searches for the Higgs coupling to charm quarks. A brief summary of related Beyond the Standard Model models is also included.

2.1 Standard Model of Particle Physics

The complete set of elementary particles, pictorially represented in Figure 2.1, is the basis of the Standard Model of particle physics (SM), that through the efforts of theorists and experimentalists for several decades has found a solid way to describe the strong, electromagnetic and weak forces through gauge field theories. It is however not complete, as it doesn't explain the asymmetry between matter and anti-matter in the universe or the nature of the dark matter in the universe, for instance.

The SM description uses a quantum field theory, where each type of particle is described by a different field. Each particle is characterised by a mass, charge and spin. Bosons (particles of integer spin) are classified as scalar if they have spin 0, and gauge/vector bosons if they have spin 1, while fermions are particles with half-integer spin.

The SM is ruled by a $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry, which imposes constraints on the terms allowed in the Lagrangian - the Lagrangian being a formulation in which it is possible to describe the physics in the form of different terms. The SM Lagrangian includes gauge boson self-interactions,

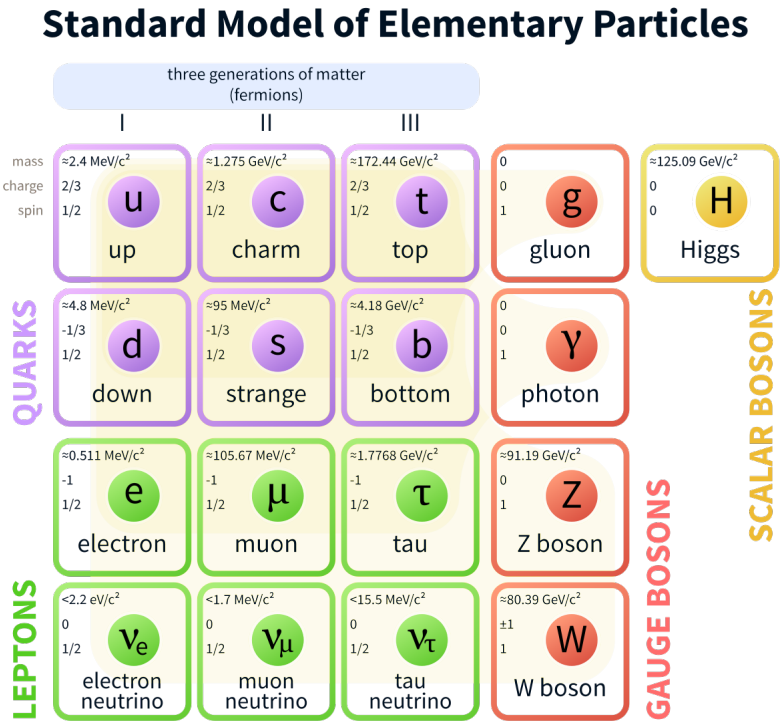


Figure 2.1: The Standard Model particles [4].

along with the interactions of the gauge bosons with fermions, and the propagators of these latter particles. While the full Lagrangian is invariant under the $SU(3)_C \times SU(2)_L \times U(1)_Y$, the vacuum is only $SU(3)_C \times U(1)_{EM}$ invariant (EM stands for electromagnetic), due to the spontaneous symmetry breaking, resulting from the Brout-Englert-Higgs mechanism, described later in this section. The SM Lagrangian accounts in addition for the couplings of the fermions to the Brout–Englert-Higgs field (Higgs field for simplicity hereafter). The interactions of the Higgs field with the gauge bosons, its propagators, and the Higgs potential are also included in the SM.

Gauge bosons are the carriers of the interactions: gluons are the mediators of the strong force, photons of the electromagnetic force and the W^+, W^- and Z

bosons being responsible for the weak force [5]. In the SM, gluons and photons are massless while the Z, W^+ and W^- bosons are massive.

Fermions divide in two categories: six quarks and six leptons. There are three families or generations of quarks (composed of an "up-type" and "down-type" quark), the only difference between each family and the next one being the increasing mass of the particles in each family. Likewise for the leptons, each family containing a charged lepton and a neutrino, the mass of the charged leptons (but not necessarily that of the neutrinos) growing from one family to the next.

Quarks and leptons are subject to the electromagnetic and weak interactions, which are described in the unified electroweak theory under the $SU(2)_L \times U(1)_Y$ gauge symmetry, that conserves the weak isospin I_3. The $SU(2)_L$ symmetry introduces three gauge bosons (W^a_μ, $a = 1, 2, 3$), while the $U(1)_Y$ symmetry introduces one gauge boson (B_μ). After symmetry breaking these fields combine to define the physical W, Z and photon fields.

Left-handed components of the fermions are arranged in isospin doublets with $I_3 = \pm 1/2$, and the right-handed components in singlets with $I_3 = 0$. The electric charge Q is defined as $Q = I_3 + \frac{Y}{2}$, where Y is the hypercharge (which couples to the $U(1)_Y$ B_μ field). The isospin, hypercharge and electric charge numbers for the fermions are shown in Table 2.1. Leptons with neutral electric charge, referred to as neutrinos, feel only the weak interactions. Moreover, quarks and charged leptons are massive particles, and some neutrinos were experimentally demonstrated to have a non-zero mass, leading to the 2015 Nobel Prize in Physics,

though the origin of their mass is still the subject of research nowadays.

Table 2.1: Isospin I_3, hypercharge Y and electric charge Q for fermions.

Fermion	Particle type	I_3	Y	Q
Leptons	$\nu_{e,L}$, $\nu_{\mu,L}$, $\nu_{\tau,L}$	1/2	-1	0
	$e_{e,L}$, $e_{\mu,L}$, $e_{\tau,L}$	-1/2	-1	-1
	$e_{e,R}$, $e_{\mu,R}$, $e_{\tau,R}$	0	-2	-1
Quarks	u_L, c_L, t_L	1/2	1/3	2/3
	d_L, s_L, b_L	-1/2	1/3	-1/3
	u_R, c_R, t_R	0	4/3	2/3
	d_R, s_R, b_R	0	-2/3	-1/3

Quarks are also charged under the strong interactions. The confined nature of QCD, the $SU(3)_C$ theory that describes the strong force, dictates that quarks can only appear in colourless combinations, $i.e$, conserving the colour charge. The *Lund* string model [6, 7] describes this in the following way: the separation of two colour charges leads to an energy increment, with the energy stored in the field between them, resembling a string being stretched. The continued separation and increased energy may lead to the production of colour-anticolour pairs of quarks out of the vacuum (breaking the string), which combine with the original quarks. This is a recurring process, as long as there is enough energy, leading to only colourless clusters existing, which are called "hadrons". As gluons also carry colour charge the mechanism works in the same way. Both quarks and gluons can also radiate gluons, leading to extra emissions.

Baryons are hadrons constituted by three quarks (corresponding to the combination of colour charges: red, green and blue or the inverse case, anti-red/anti-

green/anti-blue), and have half-integer spin. If comprised of two quarks, the hadron is called a "meson", having integer spin, and the colour combinations: red/anti-red, green/anti-green or blue/anti-blue.

In proton-proton colliders the resulting hadrons are grouped in "jets" according to their distance and momentum, via jet clustering algorithms. Jets correspond to the only way to study partons, as free partons are unreachable due to the QCD confinement. The jet clustering process must be independent of "soft" (low energy; infrared) and collinear (low angle with respect to the original quark or gluon) emissions by the quarks and gluons. This is due to divergences in higher-order perturbative calculations in such emissions in QCD [8].

Higgs Mechanism

The $SU(2)_L \times U(1)_Y$ symmetry forbids mass terms for the gauge bosons and the fermions, as these terms would not be invariant under the respective gauge transformations.

In the gauge bosons case, due to the vector field transformations, mass terms are prohibited. In the fermion case, the weak $SU(2)_L$ interaction only transforms particles with left-handed components, and leaves the right-handed particles unchanged, therefore not allowing simple fermion mass terms.

The weak gauge bosons acquire mass via the spontaneous symmetry breaking due to the Higgs mechanism, that also introduces a scalar boson, the Higgs boson

[9, 10, 11]. The masses of the charged leptons and quarks can be generated via interactions with the Higgs field.

A new particle consistent with the Standard Model Higgs boson was discovered in 2012 by the ATLAS and CMS experiments at CERN [12, 13], marking the beginning of studies of its properties.

The Higgs mechanism first introduces the smallest fundamental $SU(2)_L$ scalar Higgs doublet ϕ, with hypercharge $Y = +1$, defined as

$$\phi = \begin{pmatrix} \phi^+ \\ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \\ \phi_3 + i\phi_4 \end{pmatrix} . \tag{2.1}$$

It is then necessary to add the Higgs potential, generally expressed as

$$V(\phi) = \mu^2|\phi|^2 + \lambda|\phi|^4 , \tag{2.2}$$

where μ and λ are almost free parameters. The parameter λ must be positive in order to have vacuum stability, $i.e$, there exists an absolute minimum of the potential. On the other hand, μ^2 can be positive or negative.

If $\mu^2 > 0$, the vacuum has a single minimum at $\phi = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, and this represents a trivial case, as there is no symmetry breaking. With $\mu^2 < 0$ there is an infinite number of vacua that satisfy the condition

$$\frac{dV}{d\phi} = 0 \Leftrightarrow |\phi| = \sqrt{\frac{-\mu^2}{\lambda}} = v\,, \tag{2.3}$$

and for this case, the vacuum then has a value different from zero, the so-called vacuum expectation value (VEV or v).

The vacuum configuration with $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = v$ is chosen in this case,

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \\ v \end{pmatrix}\,, \tag{2.4}$$

introducing a symmetry breaking.

The system is then studied under small perturbations around this minimum, with four shifted fields being introduced, H, θ_1, θ_2 and θ_3, resulting in

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} \theta_1 + i\theta_2 \\ \\ v + H + i\theta_3 \end{pmatrix}\,. \tag{2.5}$$

Rewriting the Lagrangian in terms of the shifted fields results in massive gauge fields, and an additional massive scalar, the Higgs boson. In particular, the W^1 and W^2 fields mix to form the W^+ and W^- bosons, as defined in Equation 2.6, and the W^3 and B fields mix to form the Z boson and the photon, defined respectively, as:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu)\,, \tag{2.6}$$

$$Z_\mu = \frac{gW^3_\mu - g'B_\mu}{\sqrt{g^2+g'^2}}\,, \tag{2.7}$$

$$A_\mu = \frac{gW^3_\mu + g'B_\mu}{\sqrt{g^2+g'^2}}\,, \tag{2.8}$$

where g and g' are the $SU(2)_L$ and $U(1)_Y$ coupling constants.

The masses for the gauge bosons are then defined as:

$$M_W = \frac{1}{2}vg\,, \tag{2.9}$$

$$M_Z = \frac{1}{2}v\sqrt{g^2+g'^2}\,, \tag{2.10}$$

$$M_A = 0\,, \tag{2.11}$$

for the W bosons, Z boson and photon, respectively. The photon remains massless, as the vacuum is invariant under the $U(1)_{EM}$ symmetry. In fact, there exists charge conservation of the vacuum, with $Q(\phi_0) = 0$, as $Y = +1$ and $I_3 = -\frac{1}{2}$.

The Higgs boson also has an associated mass, not predicted by the SM:

$$M_H = \sqrt{2\lambda v^2} = \sqrt{-2\mu^2}\,. \tag{2.12}$$

The latest combination of the ATLAS and CMS measurements resulted in a observed Higgs mass of $m_H = 125.09 \pm 0.24$ GeV [14].

The fermion masses can be obtained by introducing the interactions of these particles with the Higgs field, the so-called Yukawa couplings (y_f). While the previous mass terms in the Lagrangian for the fermions were not invariant under the $SU(2)_L \times U(1)_Y$ symmetry, it is possible to construct terms combined with the Higgs boson, which become invariant under this symmetry, and are therefore allowed in the Lagrangian. The Yukawa coupling is proportional to the fermion mass M_f, via $y_f = \sqrt{2}\frac{M_f}{v}$. The study of muon decay measurements led to a vacuum expectation value of v of 246 GeV, where the coupling strength of the muon to the W boson can be measured in this decay, and a relation with the vacuum expectation value can be established.

2.2 Proton-proton collisions and hadronic products

Proton-proton collisions with large momentum transfer are referred to as hard scattering processes or events [15], and their cross section can be expressed as:

$$\sigma_{pp\to X} = \sum_{i,j} \int_0^1 \int_0^1 dx_i\, dx_j f_i(x_i, \mu_F) f_j(x_j, \mu_F) \hat{\sigma}_{ij\to X}(x_i, x_j, \mu_F, \mu_R)\,, \tag{2.13}$$

where the indices i, j run over all the partons within the proton, $x_{i,j}$ correspond to the fraction of the initial proton longitudinal momentum carried by each respec-

tive parton, and $f_{i,j}(x_{i,j}, \mu_F)$ are the parton distribution functions (PDFs). These represent the probability to find a parton carrying a fraction x of the initial proton longitudinal momentum at a factorisation scale μ_F. The partonic cross-section $\hat{\sigma}_{ij\to X}(x_i, x_j, \mu_F, \mu_R)$ denotes the cross section for the scattering of partons i and j, resulting in a final state X. μ_R is the renormalisation scale.

The factorisation scale μ_F is introduced due to divergences in soft and collinear gluons emissions, and corresponds to a cut-off in the PDFs at which emissions with an energy below the scale are absorbed in the PDF. The μ_F scale is commonly taken as the scale of the process, generally denoted as Q, and its uncertainties are estimated by taking a factor half or double of the central value. In order to handle ultraviolet divergences a renormalisation scale, μ_R, is introduced. The value of the QCD coupling is influenced by this scale, and its dependence can be expressed in terms of a renormalisation group equation. The uncertainties for μ_R, as for μ_F, are usually estimated by choosing $\mu_R^2 = (x_\mu \mu_F)^2$ with $x_\mu = \frac{1}{2}, 1, 2$.

The soft and collinear regime, which is not described by the hard scatter process, can be approximated via parton shower algorithms. These take into account the probability of final states with an extra parton emission, and translate the evolution in momentum transfer from the high scales of the hard scatter to the hadronisation energy scales. Unstable hadrons then decay into lighter hadrons.

For a proton-proton collider the cross-section can be written as:

$$\sigma_{pp\to X} = \sum_{i,j} \int_0^1 \int_0^1 dx_i\, dx_j\, f_i(x_i, \mu_F) f_j(x_j, \mu_F) \int d\phi_X \frac{1}{2\hat{s}} |\mathcal{M}_{ij\to X}(\phi_X, \mu_F, \mu_R)|^2 , \quad (2.14)$$

with ϕ_X being the phase-space of the final state X, and with the matrix element, $\mathcal{M}_{ij\to X}$, corresponding to the description of the process according to the interactions permitted by the SM Lagrangian. The simplest description is in leading order (LO) in the strong/electroweak coupling. A next-to-leading order (NLO) calculation corresponds to a description with the emission and absorption of a gluon or quark included where physically possible. The same logic applies to higher orders. $\hat{s} = x_i x_j s$ is the square of the partonic system (X) mass, where $\sqrt{s}$ is the collider centre-of-mass energy.

Besides the hard-scattering process a proton-proton collision at the LHC may also lead to "pile-up" or "underlying" events. Pile-up events correspond to additional proton-proton collisions occurring in the same bunch-crossing as the collision of interest, or in bunch-crossings just before and after the hard-scattering interaction. These events are modelled using simulated events that pass the loosest possible trigger requirements. Underlying events describes the activity of the softer radiation in the event, coming from particles originating from multiple-parton interactions, initial- or final-state radiation of gluons, or the remnants of the beams. A diagram with the different contributions in a proton-proton collision is shown in Figure 2.2.

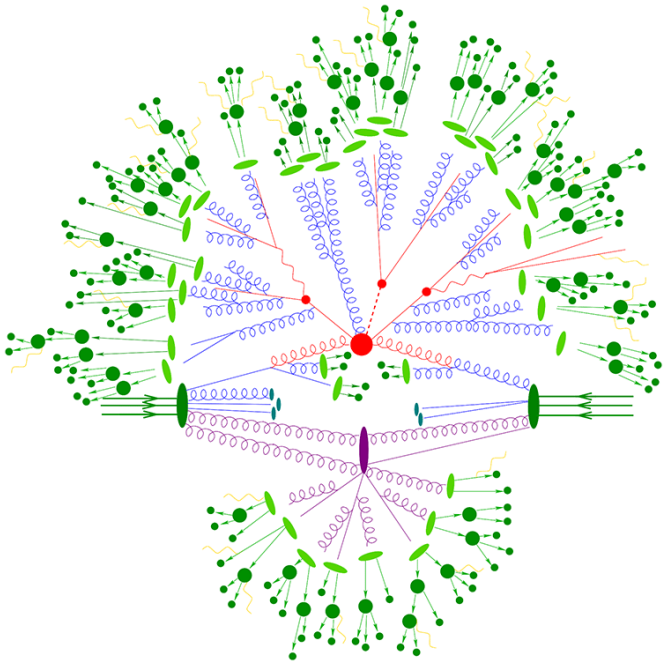


Figure 2.2: Schematic of a proton-proton collision. The hard process is represented in red, the parton shower in blue, hadron production in light green blobs and their decay products in dark green. Gluon radiation in the (red) gluons leading to the hard-scatter process correspond to initial state radiation, while gluon emission in the (red) hard-scattering process represent final state radiation. The underlying event is shown in purple [16].

Heavy flavour quarks and hadrons

Heavy flavoured (bottom and charm) quarks hadronise after being produced, with the heavy hadron having a generally longer lifetime than light hadrons due to the associated weak decay, and therefore travelling some millimetres in the detector. At 13 TeV *pp* collisions an heavy flavoured hadron travels a few millimetres from the primary vertex (PV), after which it decays at a secondary vertex (SV) into other sub-products. Tracks from this vertex can be extrapolated back to the primary vertex, defining an "impact parameter". A diagram representing an heavy flavour quark evolution is shown in Figure 2.3, where d_0 is the impact

parameter for an extrapolated track.

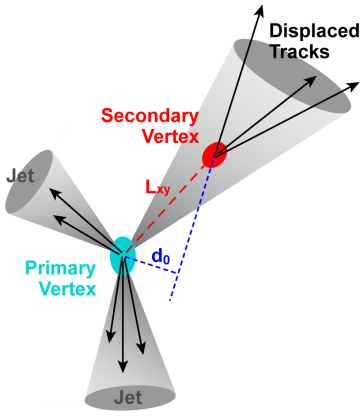


Figure 2.3: Schematic diagram of heavy flavour hadron decay [17]

The bottom and charm hadrons and resulting jets are however characterised by having slightly different properties, with more extensive differences when compared to light flavour (up, down, strange quarks and gluons) jets, translating into different signatures in the detectors. These properties [18] are exploited by flavour tagging algorithms, and are described below:

- Properties of b-jets (represented diagrammatically in Figure 2.4);
  - b-quark fragments into b-hadron carrying around 80% of the jet energy;
  - High b-hadron decay product multiplicities (around 5 charged particles per decay);
  - Most b-hadrons ($\approx$ 90%) decay into c-hadrons;
  - b-hadron decay vertex often displaced from the PV by a few mm;
  - c-hadron decay vertex often displaced from the SV by a further few mm;

– Tracks from both of these vertices often have large impact parameters.

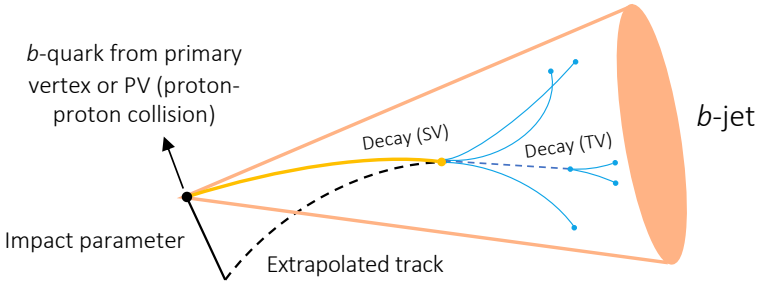


Figure 2.4: Anatomy of a b-jet.

- Properties of c-hadrons (represented diagrammatically in Figure 2.5);
  - c-quark fragments into c-hadron carrying around 55% of the jet energy;
  - 2 to 3 times lower c-hadron decay product multiplicities than for b-hadrons (around 2 charged particles/decay);
  - c-hadron decay vertex often displaced from the PV by a few mm;
  - Tracks from this vertex often have large impact parameters.
- Light flavour jets (represented diagrammatically in Figure 2.6);
  - Light quark hadronises into many light hadrons sharing the jet energy;
  - Tracks from this vertex most often have impact parameters consistent with zero;
  - Long-lived light hadrons can be produced, but are more likely to decay cms away from the PV.

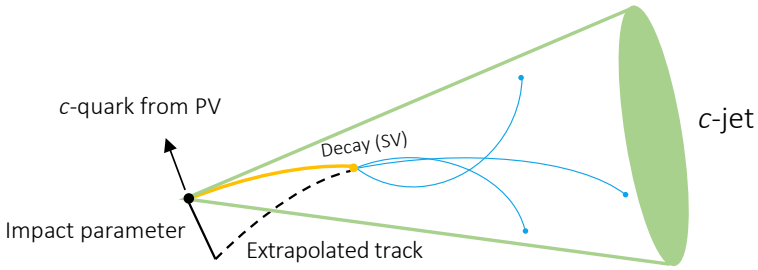


Figure 2.5: Anatomy of a c-jet.

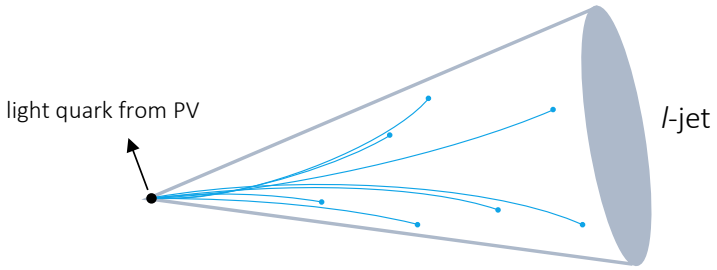
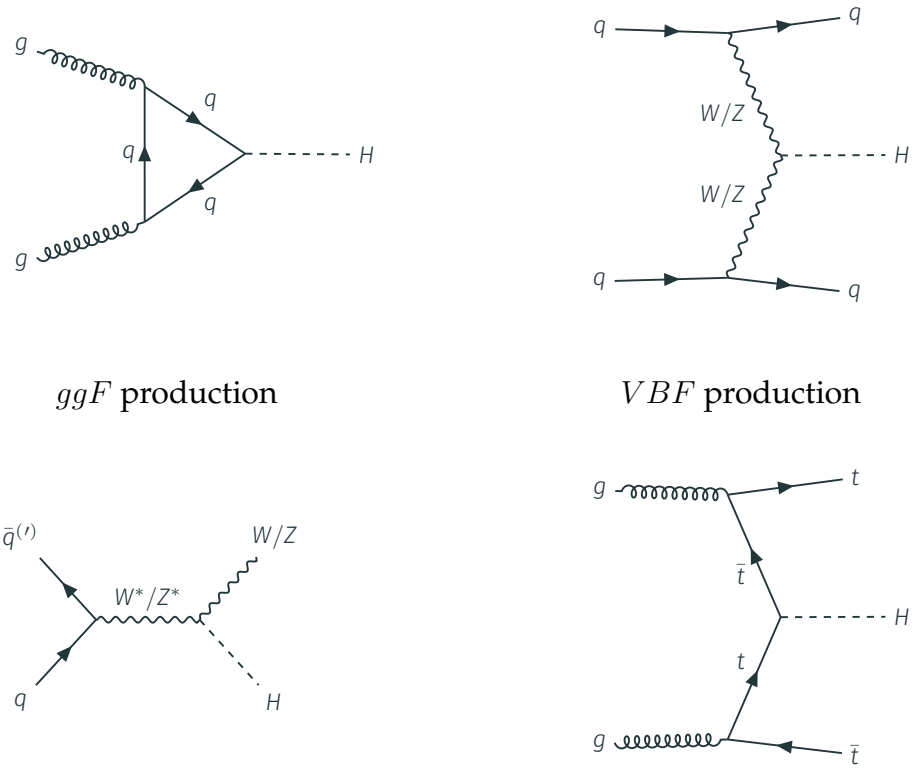


Figure 2.6: Anatomy of a light flavour jet.

2.3 Higgs boson Production and Couplings

The Higgs boson has four main production modes in proton-proton collisions, of which representative examples are shown diagrammatically via Feynman diagrams in Figure 2.7. The top-left diagram represents gluon-gluon fusion (ggF) production, where initial state gluons interact via a quark loop, usually comprised of top quarks (given their larger mass), to produce a Higgs boson. The top-right figure depicts vector-boson fusion (VBF) processes, where two vector bosons (Z or W bosons) are radiated from quarks and interact, producing a

Higgs. Represented on the bottom-left diagram is VH production, with the production of a Higgs boson in association with a vector boson. The bottom-right figure shows the $t\bar{t}H$ channel, with the production of a Higgs boson in association with two top quarks.



ggF production | VBF production

VH production | $t\bar{t}H$ production

Figure 2.7: Feynman diagrams for the main Higgs boson production modes.

The evolution of the cross-sections as a function of the energy can be found in Figure 2.8, and the associated cross-section values for each of the production modes for a Higgs boson with $m_H = 125$ GeV at a centre-of-mass energy of 13 TeV in proton-proton collisions are shown in Table 2.2.

The branching ratio of a Higgs boson decay into a pair of particles is defined as

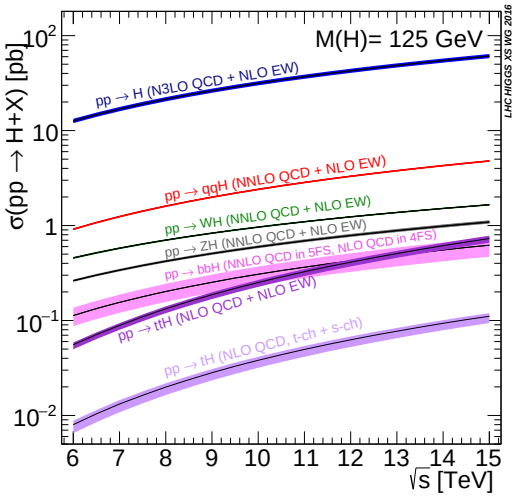


Figure 2.8: SM cross-sections for Higgs boson production modes as a function of center-of-mass energy, with $m_H = 125$ GeV [19].

Table 2.2: SM cross-sections for main production modes of a Higgs boson with $m_H = 125$ GeV in pp collisions [1].

Production Mode	Cross section (pb)
ggF	48.6
VBF	3.78
WH	1.37
ZH	0.88
$t\bar{t}H$	0.50

the ratio between the partial width of the decay, Γ, with respect to the total Higgs width:

$$\text{BR}\,(H \to X\bar{X}) = \frac{\Gamma(H \to X\bar{X})}{\Sigma_i\,\Gamma(H \to X_i\bar{X}_i)}\,. \tag{2.15}$$

The Standard Model predictions for the Higgs boson branching ratios are shown in Figure 2.9 and Table 2.3. The expected total Higgs decay width in the SM is $\Gamma_H = 4$ MeV. The decays of the Higgs to massless particles require the presence of a loop of massive particles, as the Higgs only couples to massive fermions and bosons. This is the case for the $H \to gg$, $H \to \gamma\gamma$ and $H \to Z\gamma$ (due to the massless photon) decays.

The Higgs Yukawa couplings to fermions have been probed since the discovery of the Higgs boson, starting with the interactions with the third generation of

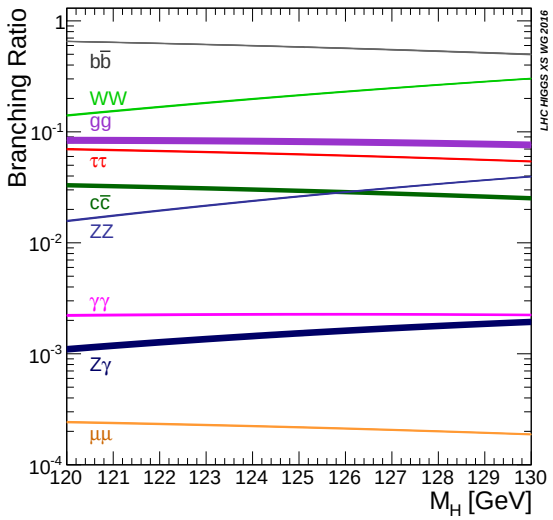


Figure 2.9: SM branching ratios for Higgs decays as a function of its mass [19, 20].

Table 2.3: SM branching ratios for a Higgs boson with $m_H = 125$ GeV [19, 20].

Decay	Branching Ratio
$H \to b\bar{b}$	58.20%
$H \to WW$	21.40%
$H \to gg$	8.19%
$H \to \tau\tau$	6.27%
$H \to c\bar{c}$	2.89%
$H \to ZZ$	2.62%
$H \to \gamma\gamma$	0.23%
$H \to Z\gamma$	0.15%
$H \to \mu\mu$	0.03%

fermions given their larger masses and therefore stronger couplings and larger branching ratios.

The ATLAS and CMS collaborations observed the Higgs coupling to top [21, 22] and bottom quarks [23, 24], and τ-leptons [25, 26], and evidence for the Higgs boson decay to muons was found by CMS [27], with ATLAS [28] finding a 2σ excess over the background-only prediction.

2.4 Kappa Framework

The searches for the Higgs couplings can be re-interpreted in terms of coupling modifiers in the κ-framework [19, 20], in order to measure deviations of the Higgs couplings from the Standard Model based on the leading-order contributions to each Higgs production and decay process.

The production cross section and decay branching fraction for a Higgs boson process can be factorised as:

$$\sigma(i \to H) \times \text{BR}\,(H \to f) = \frac{\sigma_i \times \Gamma_f}{\Gamma_H}\,. \tag{2.16}$$

Where σ_i is the production cross section, Γ_f is the partial width for a Higgs boson decay into a pair of f particles and Γ_H is the total width of the Higgs boson.

Introducing the Yukawa coupling modifiers, $y_n \to \kappa_n y_n$, the process cross section times branching fraction becomes:

$$\sigma(i \to H) \times \text{BR}\,(H \to f) = \frac{\sigma_i^{SM}\kappa_i^2 \times \Gamma_f^{SM}\kappa_f^2}{\Gamma_H^{SM}\kappa_H^2}, \kappa_H^2 = \frac{\Sigma_j \Gamma_j^{SM} \times \kappa_j^2}{\Gamma_H^{SM}}\,, \tag{2.17}$$

where κ_n are the coupling modifiers. In the SM these correspond to 1.

As an example it is shown in Figure 2.10 a $H \to c\bar{c}$ process, in which two coupling modifiers are introduced, κ_V being the modified coupling of the Higgs to the vector bosons, and κ_c the Higgs-charm coupling modifier.

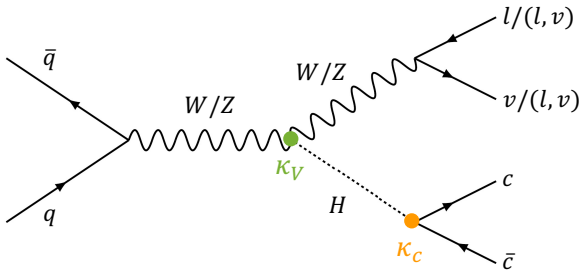


Figure 2.10: Example of Higgs coupling modifiers in a $H \to c\bar{c}$ process.

In this case Equation 2.17 becomes:

$$\sigma(VH) \times \text{BR}(H \to c\bar{c}) = \sigma(VH)_{SM} \times \text{BR}(H \to c\bar{c})_{SM} \times \frac{\kappa_V^2 \kappa_c^2}{\kappa_H^2}. \quad (2.18)$$

The ATLAS collaboration combined the analyses for the $H \to \gamma\gamma$, $H \to ZZ^*$, $H \to WW^*$, $H \to \tau\tau$, $H \to b\bar{b}$, $H \to \mu\mu$, $H \to Z\gamma$ decay modes and searches for the Higgs decays into invisible final states [29] for improved constraints on the coupling modifiers.

The fitted values for the coupling modifiers using a parameterisation assuming no new particles in loops and decays are in Table 2.4. These results assume no beyond-the-Standard-Model contributions to the Higgs production and decay, and invisible or undetected Higgs boson decays are not considered. Invisible decays correspond to Higgs decays resulting in objects that do not interact with the detector, resulting in missing transverse energy in the events. The SM expectation for this branching ratio is of 0.1% from $H \to ZZ^* \to 4\nu$ decays. Undetected decays are decays resulting in particles not yet resolved and not associated to large missing transverse energy contributions in the final state. The SM contribution to this effect is of 11%, coming mainly from Higgs decays to gluons. In addition, this parametrisation assumes κ_c to vary as κ_t and κ_s as κ_b, and the coupling modifiers to the first generation of fermions are assumed to follow the SM expectation, $i.e$, are set to the unity. Finally, all coupling modifiers are assumed to be positive.

Including in the parametrisation effective coupling modifiers for the Higgs decay

Table 2.4: Fit results for κ_Z, κ_W, κ_b, κ_t, κ_τ and κ_μ [29].

Parameter	Result
κ_Z	0.99 ± 0.06
κ_W	1.03 ± 0.05
κ_b	0.88 ± 0.11
κ_t	0.92 ± 0.06
κ_τ	0.92 ± 0.07
κ_μ	$1.07\ ^{+\ 0.25}_{-\ 0.31}$

to photons, $Z\gamma$ or gluons, which can only happen via loops of massive particles, translates into the fitted values in Figure 2.11. All coupling modifiers are assumed to be positive, except κ_t that can be positive or negative. The beyond-the-Standard-Model (BSM) contribution to the branching ratio of invisible (undetected) decays is defined as $B_{\text{i.}}$ ($B_{\text{u.}}$). Two scenarios are considered, one in which there are no BSM contributions to the Higgs width, $i.e$, $B_{\text{i.}} = B_{\text{u.}} = 0$, and another where the BSM contributions to the undetected ($B_{\text{u.}}$) and invisible decays ($B_{\text{i.}}$) are free parameters in the fit model. The $B_{\text{i.}}$ contribution is constrained via the search for Higgs decays into invisible final states, and $B_{\text{u.}}$ via the conditions $\kappa_W \leq 1$ and $\kappa_Z \leq 1$ (models with extended Higgs sectors, as the ones discussed in Section 2.6, typically include couplings to the W/Z bosons of smaller magnitude with respect to the SM predictions). Moreover, $B_{\text{u.}}$ is assumed to be positive. In the scenario in which $B_{\text{i.}}$ and $B_{\text{u.}}$ are free parameters of the model, the constraints $B_{\text{i.}} < 0.09$ and $B_{\text{u.}} < 0.16$ are observed at 95% confidence level (CL).

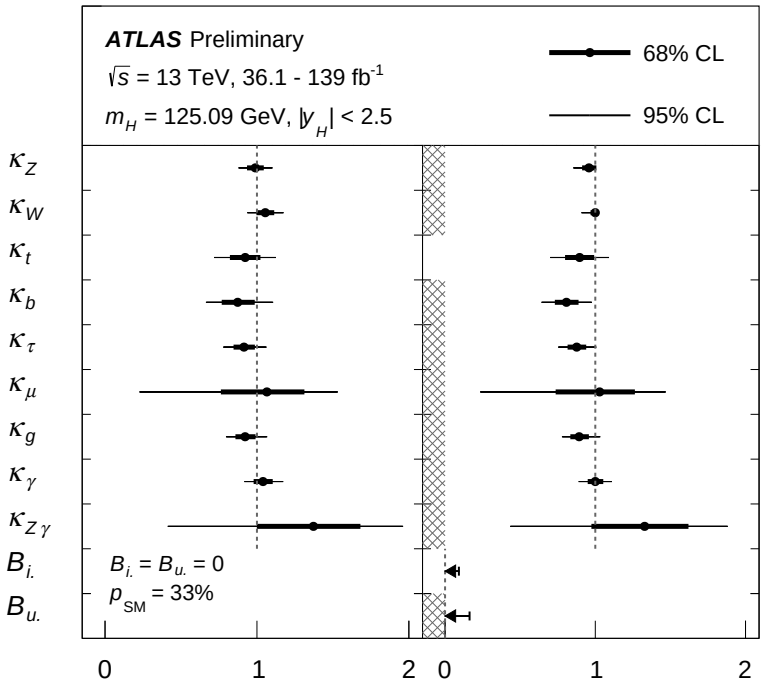


Figure 2.11: Best-fit values and uncertainties for Higgs boson coupling modifiers per particle type with effective photon, $Z\gamma$ and gluon couplings and either $B_{\text{i.}}$ = $B_{\text{u.}}$ = 0 (left), or $B_{\text{i.}}$ and $B_{\text{u.}}$ included as free parameters in the model (right) [29].

2.5 Searches for the Higgs-charm coupling

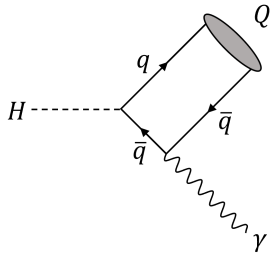
The measurement of the Higgs-charm coupling represents the next natural step in the endeavour to measure all fermion interactions with the Higgs. The branching fraction for the Higgs boson to charm decay, $H \to c\bar{c}$, 2.89% [30] in the Standard Model, is one of largest expected contributions to the Higgs boson width yet to be established experimentally. Furthermore, the SM predicts the Higgs-charm Yukawa coupling, y_c, to be around five times smaller than the Higgs coupling to bottom quarks, y_b [1, 31, 32]:

$$y_c = \sqrt{2}\, m_c(\mu = m_H)/v \simeq \sqrt{2}\, 0.6\,\text{GeV}/246\,\text{GeV} \simeq 0.2 \times y_b\,. \tag{2.19}$$

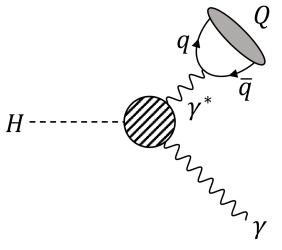
Different probes of the $Hc\bar{c}$ coupling have been proposed and investigated by both the ATLAS and CMS experiments, targeting directly or indirectly the coupling under different assumptions [33, 34, 35].

2.5.1 Exclusive $H \to J/\psi\,\gamma$ decays

This process is a rare decay but experimentally a clean probe of the Higgs coupling to charm quarks and is characterised by an interference between direct and indirect contributions, represented diagrammatically in Figure 2.12. The direct amplitude is sensitive to the magnitude and sign of the Higgs-charm coupling, while the indirect amplitude constitutes the dominant contribution to the Higgs boson width, by around a factor 20, and is not sensitive to the charm Yukawa coupling. The expected SM branching ratio is of BR($H \to J/\psi\,\gamma$) $= (3.01 \pm 0.16) \times 10^{-6}$. Moreover, given the production rate is dominated by the indirect contribution, the sensitivity to the coupling is limited.



Direct contribution



Indirect contribution

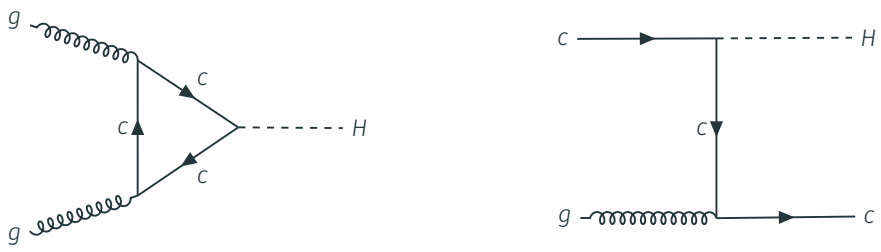
Figure 2.12: $H \to J/\psi\,\gamma$ decay. Q is a vector bound state of quarks and the hatched circle represents a set of one-loop diagrams.

The latest ATLAS search for this decay found an observed limit of BR($H \to J/\psi\,\gamma$) $< 3.5 \times 10^{-4}$ at 95% CL [33]. Due to the observed branching ratio being two orders

of magnitude above the SM prediction, it is only possible to set constraints on the coupling modifier at very large values, that contradict the Higgs rate measurements from the LHC.

2.5.2 Charm quark initiated Higgs boson production

The production of the Higgs boson in charm quark initiated processes provides sensitivity to the Higgs-charm (and Higgs-bottom) coupling, via measurements of the total production cross-section and kinematic features. Examples of these processes are shown in Figure 2.13. One example of useful kinematic variables is the p_{T}^{H} spectrum, whose shape will be distorted in case of an $Hc\bar{c}$ coupling different from the SM prediction. Only changes to the shape of the p_{T}^{H} differential cross-section are considered and interpreted in the kappa framework, as the rate has a Higgs width dependence hard to factor out.



Charm loop in $gg \to Hg$ process $Hc\bar{c}$ coupling in $gc \to Hc$

Figure 2.13: Higgs boson production involving Higgs-charm coupling.

The latest ATLAS $H \to ZZ^* \to 4\ell$ [34] and $H \to \gamma\gamma$ [35] cross-section measurements were recently statistically combined [36]. The observed constraints on κ_b

and κ_c, from likelihood scans on each coupling modifier and leaving the other free to float, were of $-2.0 < \kappa_b < 7.4$ and $-8.6 < \kappa_c < 17.3$ at 95% CL. These constraints are not as stringent as the ones set by inclusive Higgs decays searches, but provide complementary sensitivity to the sign of the coupling modifiers.

2.5.3 Inclusive $H \to c\bar{c}$ decays

The Higgs decay to charm quarks is the most direct probe of the Higgs-charm coupling. The searches for the $H \to c\bar{c}$ decay are built around the use of c-jet tagging algorithms and target generally the production of the Higgs boson in association with a Z or W boson, VH production ($V = W/Z$), with the vector bosons decaying leptonically, as represented in Figure 2.14.

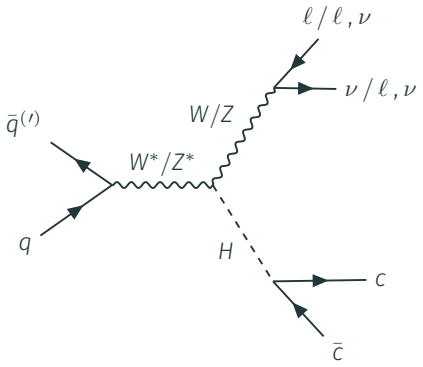


Figure 2.14: $H \to c\bar{c}$ decay in VH production.

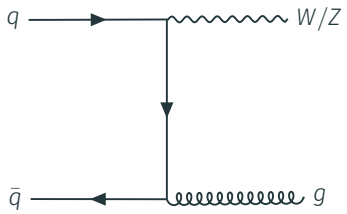
The VH production channel is targeted in these searches as the W/Z boson decays into leptons allow for a convenient trigger strategy, with suppression of QCD hadronic backgrounds, and it provides an enhancement of the ratio between the Higgs signal and background yields with respect to the inclusive Higgs production case.

The first constraint on κ_c via searches for the $H \to c\bar{c}$ decay was set by ATLAS in the analysis described in this thesis.

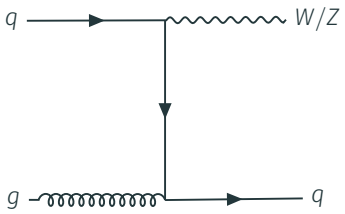
Main backgrounds

The signature of the $VH(H \to c\bar{c})$ signal can be mimicked by other processes producing the same particles in the final state, or particles leading to similar signatures in the detector. As the jet flavour tagging algorithms do not have a 100% rejection efficiency of b-jets and light flavour jets, processes with these objects can enter the analysis. In addition, these processes can contain real final state leptons, from a W or Z boson decay, which can satisfy the leptonic requirements in the analysis. Alternatively, jets can simulate a lepton signature with a $\mathcal{O}(1\%)$ rate. The processes resembling the signal are denoted as "backgrounds" and are briefly described on the following points:

- V+jets processes (represented in Figure 2.15)
  - In these processes the final state is comprised of a vector boson (W or Z boson), which decays leptonically, and in addition a gluon or quark with radiated gluons that give origin to jets containing hadrons of the same flavour, or jets with mixed flavour hadrons.
- Top processes (represented in Figure 2.16)
  - Top quarks decay to a b-quark and a W boson with a branching fraction of 99.7% [1].



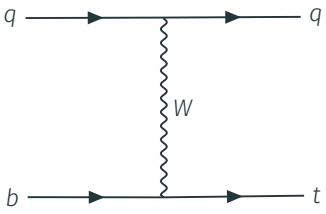
V boson + same flavour jets



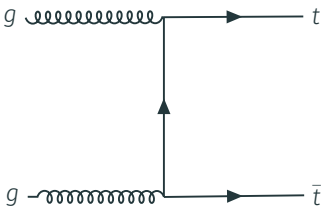
V boson + with mixed flavour jets

Figure 2.15: Representative V+jets processes.

– The main top production modes are via "single-top" processes, or in $t\bar{t}$ pairs.



Single-top production

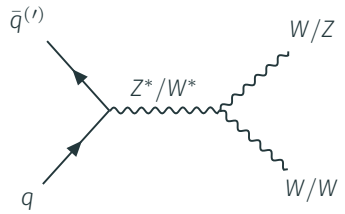


$t\bar{t}$ production

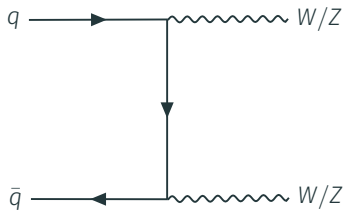
Figure 2.16: Representative top processes.

- Diboson processes (represented in Figure 2.17)

  – Two vector bosons in the final state (ZZ, WZ or WW) are present, with one decaying leptonically and the other hadronically.



WZ and WW processes



WW, ZW and ZZ processes

Figure 2.17: Representative diboson processes.

2.6 Beyond the Standard Model

The interaction of the Higgs boson with the charged fermions does not explain, however, the observed distribution of their masses, with the masses for the quarks and charged leptons spanning several orders of magnitude with no apparent pattern. The measurement of all couplings of the Higgs to fermions will therefore test the Standard Model predicted interaction or potentially help find hints for new physics that could explain the origin of the different fermion masses.

The Higgs-charm coupling can be used as a tool to this effect, as although it is rather small, particularly when compared to the Higgs-bottom coupling, it is susceptible to significant modifications in some new physics scenarios [37, 38, 39].

Some of these extensions to the SM decouple the mass acquisition of the fermions, with the third generation of fermions obtaining predominantly their masses via the Higgs mechanism, and the first and second generations via a different mechanism.

One hypothesis is to consider two Higgs-doublet models (2HDM), in which the potential is still invariant under the same symmetries as in the SM but it is built with two complex Higgs doublets (instead of the single one in the SM, which represents the simplest case to obtain spontaneous symmetry breaking). As a result, new Higgs bosons are introduced: two neutral scalars (h and H, with $m_h < m_H$, where h is the SM Higgs boson), two charged Higgs scalars (H^+ and H^-) and a neutral pseudo-scalar A. In some of these models the SM Higgs field ϕ can interact only/predominantly with the third generation of fermions as in the Higgs mechanism, with the other generations obtaining their masses via a mixing of the SM Higgs field ϕ and the additional Higgs field ϕ'.

Another option is to consider that the mass of the first and second generation of fermions arise or suffer contributions from new physics (scalars, vector-like quarks, etc) that couple to the fermions, with these effects being treated in an effective field theory approach. In this context the Standard Model is treated as a low-energy representation of a higher energy theory, and the extended SM Lagrangian takes the form:

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{c_i^{(6)}}{\Lambda^2} \mathcal{O}_i^{(6)} \,, \tag{2.20}$$

where $\mathcal{L}_{SM}$ is the SM Lagrangian, Λ is the energy scale of new physics (usually around a few TeV) and $c_i^{(6)}/\Lambda^2$ are the "Wilson coefficients", that specify the strength of the new physics interactions (in the SM these coefficients are zero).

$\mathcal{O}_i^{(6)}$ are dimension-six operators encoding the new interactions (the SM includes only dimension-four operators in order to be renormalisable). Dimension-five operators are excluded as they lead to lepton number violation. Higher dimension operators either violate further conservation rules or are negligible.

These SM extensions predict enhancements of the Higgs decay to charm quarks with respect to the SM prediction from a factor three to scenarios in which the Higgs-charm coupling has a similar strength to the Higgs-bottom one.

Chapter 3

The ATLAS Experiment

The ATLAS experiment [2] is a general-purpose detector designed to study the proton-proton and heavy ion collisions collected in the Large Hadron Collider (LHC) [3] at CERN, the European Organization for Nuclear Research.

This chapter starts by describing the LHC particle accelerator, followed by a description of the ATLAS experiment and its sub-systems. A later section covers the upgrade of one particular component of the ATLAS detector, as well as the work performed on the context of this thesis. Finally, this chapter covers the future upgrades of the ATLAS experiment and the LHC.

3.1 The Large Hadron Collider at CERN

The LHC is a proton-proton and heavy ions accelerator and collider, installed at the French–Swiss border near Geneva in a 27 km circumference tunnel, previously used in the Large Electron Positron (LEP) operation [40]. The LHC accelerates two counter-rotating beams of protons (heavy ions) via a multi-step operation, shown diagrammatically in Figure 3.1.

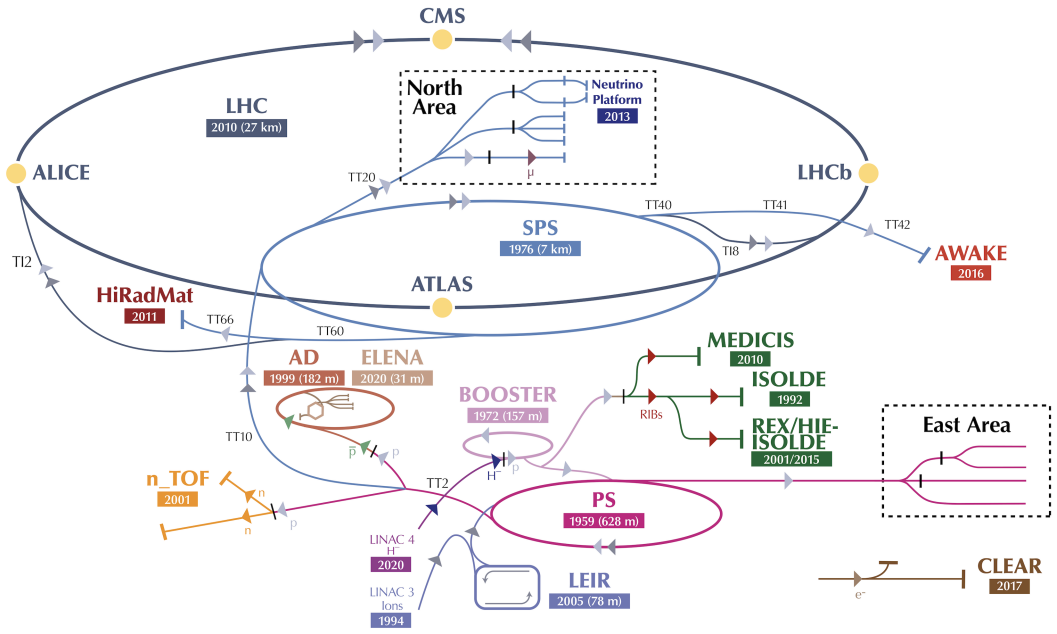


Figure 3.1: Schematic view of the CERN accelerator complex [2].

Protons from ionised hydrogen gas (or heavy ions) are first accelerated to 50 MeV by the linear accelerator LINAC4 (LINAC3 and Low Energy Ion Ring). These beams are then accelerated to 1.4 GeV in the Proton Synchrotron Booster and their energy is further increased afterwards to 25 GeV in the Proton Synchrotron (PS). The bunch structure is arranged in the PS. The beams are then fed into the Super Proton Synchrotron (SPS) in order to reach the LHC injection level energy

of 450 GeV. In the LHC, radiofrequency cavities operated at 400 MHz accelerate the particles and keep them in the desired operating energy, compensating energy losses from synchrotron radiation. The LHC is designed to run beams of 2808 bunches, each with around 10^{11} protons, with a bunch spacing of 25 ns, and collide them with a 40 MHz collision rate. The LHC Run 2 operated in a pile-up environment of around 40 additional collisions besides the hard-scattering event.

The LHC tunnel is also equipped with 1232 8.3 T superconducting dipole magnets to bend the beams, quadrupole magnets to focus the beams in the ring, and additional multipoles magnets to further focus the beams for collisions and to separate them afterwards, having in total arout 9000 magnets.

There are four main crossing points in the LHC, each corresponding to a different experiment. LHCb [41] and ALICE [42] focus on b/c-hadron physics and heavy ion collisions, respectively, while ATLAS and CMS [43] are two general purpose detectors, focusing on Higgs boson physics, improvements of the Standard Model understanding, and dark matter and other new physics searches. Both make use of different technologies in the respective detectors, providing in this way complementary approaches to physics results. The opposing beams collide in the centre of the detectors at a slight crossing angle, as to avoid parasitic long-range beam-beam interactions.

Proton-proton collisions took place at centre-of-mass energies of 7 and 8 TeV in the first LHC operation run (Run 1) during 2011 and 2012, after which there was a two-year shutdown (long shutdown 1) for maintenance and upgrades. The

LHC Run 2 operated at a centre-of-energy of 13 TeV from 2015 to 2018, with a shutdown taking place from 2019 to 2022.

Instantaneous luminosity is defined in Equation 3.1, with f_{coll} being the bunch crossing frequency, n_1 and n_2 the number of particles in each bunch and σ_x and σ_y being the transverse profiles of the beams, assuming they are Gaussian. The LHC Run 2 operated at a maximum instantaneous luminosity of around 2×10^{34} $\text{cm}^{-2}\text{s}^{-1}$.

$$\mathcal{L} = f_{coll} \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \,. \tag{3.1}$$

The integrated luminosity constitutes another parameter of the LHC machine, being defined as the integral of instantaneous luminosity over a period of time. It is usually expressed in inverse femtobarn ($1\ \text{fb}^{-1} = 10^{39}\ \text{cm}^{-2}$). The luminosity ($\mathcal{L}_{int}$) relates the processes' cross-sections σ (usually expressed in pico or femtobarns) and rate of potential occurrences of those processes in collisions via:

$$rate = \mathcal{L}_{int} \times \sigma \,, \tag{3.2}$$

where the cross-section for a process is dependent on the centre-of-mass energy.

3.2 The ATLAS Detector

ATLAS (A Toroidal LHC ApparatuS) [2] is a general-purpose cylindrical detector, providing a broad geometric coverage around the collision's interaction point. It is located around 100 m underground, and is 44 m long, with a diameter of 25 m, weighing about 7 kilo-tonnes. The ATLAS detector is comprised of multiple sub-systems, each with a particular function, and is represented in Figure 3.2.

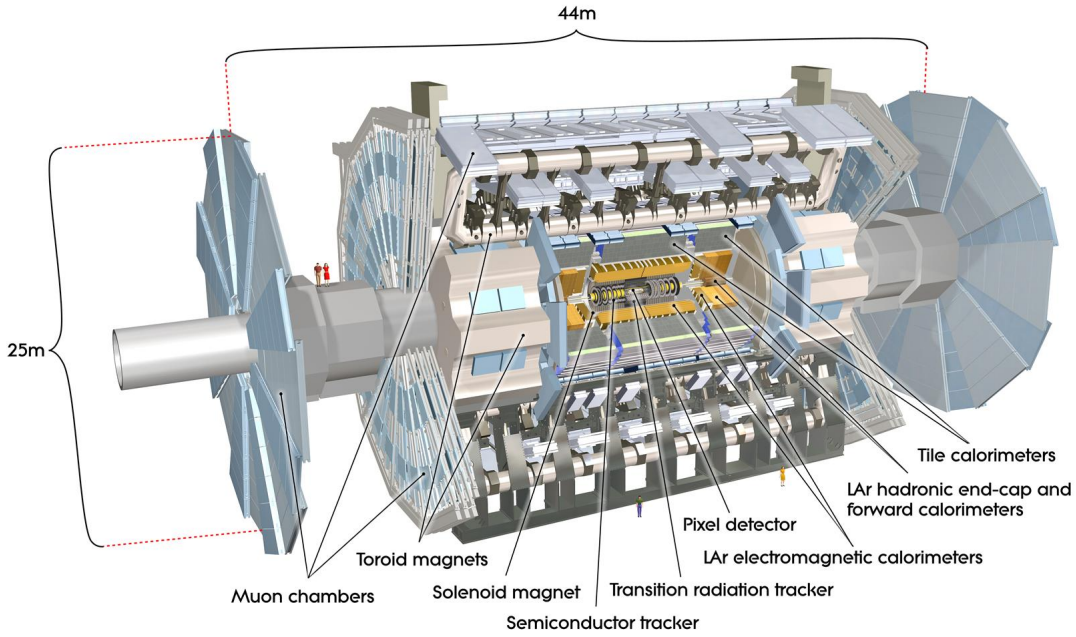


Figure 3.2: The ATLAS detector [2].

At its centre is the inner detector (ID) surrounded by a superconducting solenoid magnet creating a 2 T magnetic field, which bends the particles in the plane perpendicular to the beam line. The interaction of the particles with the detector environment produce "hits" in the detector, with sequences of hits constituting tracks. The curvature of a particle's passage through the detector due to the magnetic field and resulting track allow for the momentum and charge of the parti-

cle to be measured. The ID consists of a high-granularity silicon pixel detector located around the beam pipe, allowing for the measurement of the impact parameter of charged-particle tracks and the position of primary and secondary charged particle vertices, a semiconductor tracker (SCT) that contributes to precision tracking, and further out a transition radiation tracker (TRT) providing tracking and electron identification.

The ID is surrounded by the electromagnetic and hadronic calorimeters, that measure the energy of charged and neutral particles. The innermost electromagnetic calorimeter (ECAL or Liquid Argon - LAr) is a lead/liquid argon system primarily responsible for the detection of photons and electrons. Further out are the hadronic calorimeters, using liquid argon or scintillating tiles as active materials, and iron, copper or tungsten as absorbers.

The last layer of the ATLAS detector corresponds to the muon spectrometer, which comprises the muon system and toroid magnets (1.5 to 5.5 Tm of bending power in the central region $\eta < 2.5$ and of 1 to 7.5 Tm in the higher η regions of the detector). The deflection of the muons in the magnetic field allows for their momentum to be determined.

It is not possible to store all the collision events with a bunch-crossing rate of 40 MHz due to technical limitations, such as storage capacity and readout bandwidth, leading to the implementation of an online trigger. The trigger is therefore a system which identifies potentially interesting interactions and selects them to be read out.

The ATLAS trigger [2, 44] is divided in two layers: Level 1 (L1) and High Level Trigger (HLT). L1 is purely hardware-based, and is responsible for reducing the event rate from 40 MHz to 100 kHz. It is composed of muon (L1Muon) and calorimeter sub-triggers (L1Calo), whose information is used to find trigger objects (TOBs), which are sent to a Central Trigger, where the decision of accepting or rejecting the event is made, based on the objects provided to it. The HLT uses offline-like algorithms on top of the trigger objects to reconstruct the events and select the ones containing interesting features, further lowering the rate to around 1 kHz.

ATLAS uses a right-handed coordinate system with the proton-proton interaction point in the centre of the detector as its origin. The z-axis is aligned with the direction of the beam pipe, with the y-axis pointing upwards and the x-axis pointing towards the centre of the LHC. The polar (θ) and azimuthal (ϕ) angles are as usually defined from and around the z-axis.

Pseudo-rapidity, η, given by Equation 3.3, is the approximation of rapidity for high energy regimes. This variable is particularly useful as there are several possible scenarios when two proton beams meet, as they can collide head-on, miss each other, or just deflect. The particles resulting from the interaction between the quarks and gluons inside the protons will be spread in different angles, justifying the use of an angle-dependent variable. This is also preferred over the angle itself, as differences in pseudo-rapidity are Lorentz invariant, and particle production in collisions is roughly constant as a function of rapidity. Small (high)

$|\eta|$ values correspond to the central (forward) part of the detector.

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{3.3}$$

3.2.1 Inner Detector

The inner detector is the system closest to the beam line in the ATLAS detector, and is responsible for the reconstruction of the trajectories, vertices and momenta of charged particles. It comprises four components for high-precision measurements: the insertable B-layer (IBL), the pixel detector, the SCT and the TRT. The ID is divided into a barrel and end-cap regions, covering in total the pseudo-rapidity range $|\eta| < 2.5$. Particles are detected via the ionisation of the detector materials, or the creation of electron-hole pairs in the silicon detectors, that are separated using an electric field, and the generated charge is read. A diagram of the ID barrel and end-caps is shown in Figure 3.3 and Figure 3.4, respectively.

The silicon pixel detector is composed of three layers in the barrel region, covering $|\eta| < 1.5$, and two end-caps, one on each side of the detector and each with three disks perpendicular to the beam direction, covering the range $1.5 < |\eta| < 2.5$. The three layers in both the barrel and end-caps allow for three space points for track reconstruction of charged particles produced in the interaction point. The barrel and the end-caps contain 1456 and 288 sensor modules, respectively, with 46080 readout pixels in each module, resulting in around 80 million readout

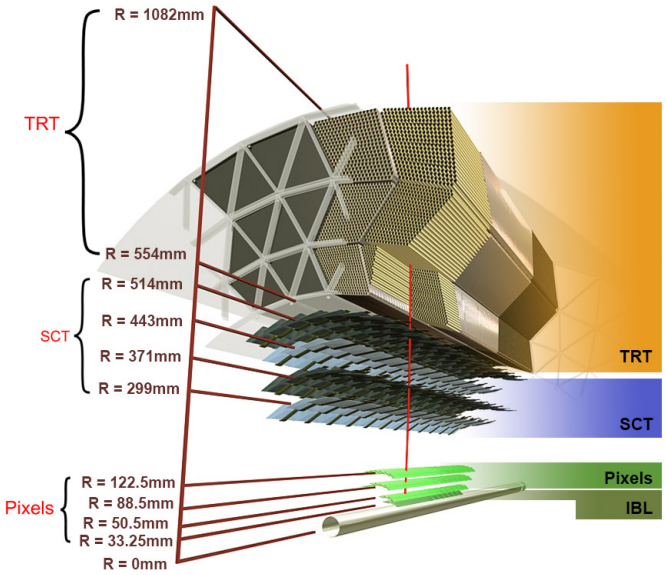


Figure 3.3: The ATLAS inner detector barrel region [45].

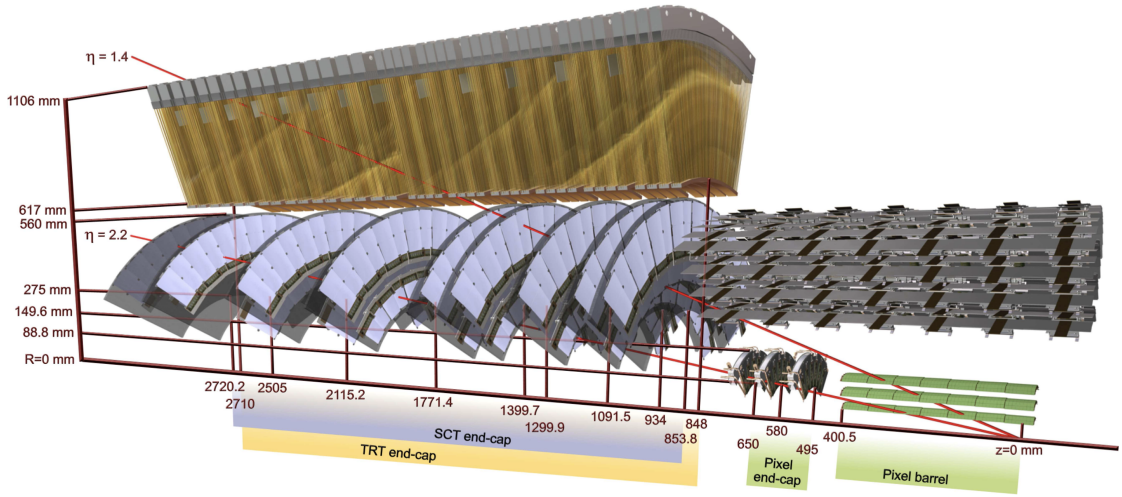


Figure 3.4: The ATLAS Inner Detector end-cap regions [2] before the IBL installation.

channels. Each pixel has a nominal area of 50 $\times$ 400 μm^2, with a resolution of 10 μm in the r-ϕ plane and of 115 μm in the z-direction in the barrel region, and of 10 μm in the r-ϕ plane and of 115 μm in the r-direction in the end-caps.

The additional insertable B-layer was installed between the existing pixel detector and the beam pipe during the first LHC long shut-down [46]. It provides im-

proved tracking precision, facilitating in particular b-jet tagging in the detector, while also recovering performance lost by dead pixels in the pixel detector due to exposure to radiation. The IBL adds more 26880 pixel cells to the ID system, each with a nominal pixel size of 50×250 μm^2, and covers $|\eta| < 2.5$.

The SCT consists of silicon microstrips detectors, aligned parallel to the beam pipe in four layers in the barrel region, and radially in nine end-cap disks. It contains around 6.3 million readout channels, covering the pseudo-rapidity range $|\eta| < 2.5$. The SCT has hit resolutions of 17 μm in r-ϕ and 580 μm in z (r) in the barrel (end-cap disks).

The TRT has around 300,000 straws, arranged longitudinally in 73 layers in the barrel, and radially in 160 layers in the end-caps. The TRT has typically at least 36 hits for a charged particle with a transverse momentum above 5 GeV. It measures the ionisation of a gas mixture (mainly composed of xenon, with smaller percentages of carbon dioxide and oxygen) in the straws. The TRT contributes to the identification of electron and positrons by detecting their emitted X-ray photons as they cross the straws. The TRT has a resolution of 130 μm in r-ϕ per straw and covers $|\eta| < 2$.

3.2.2 Calorimeters

The ATLAS calorimetry system consists of an electromagnetic and hadronic part, being responsible for the measurement of the energy deposited by charged and

neutral particles, providing information in addition about their position and direction. The calorimeters allow in addition for the determination of the missing transverse energy in the event.

Transverse energy, or E_T, refers to the component of energy, treated as a vector, in the plane perpendicular to the beam direction[1], while missing E_T corresponds to the E_T imbalance in the event. This imbalance allows to estimate the E_T of particles which do not interact with the detector, such as neutrinos.

The calorimeters are sampling detectors, consisting of layers of absorber material alternated with active materials. The particles interact with the absorbers when crossing the detector, losing energy and forming showers. The electromagnetic (hadronic) calorimeter is designed to contain electromagnetic (hadronic) showers. The layout of the ATLAS calorimeters is shown in Figure 3.5.

The LAr electromagnetic (EM) calorimeter consists of accordion-shaped lead absorber plates and electrode plates interleaved with liquid argon. The accordion design leads to a full coverage in ϕ without any cracks, and allows for the extraction of the signal at the front or back of the electrodes. Showers are created as particles interact with the absorber plates. The showers then ionise the liquid argon, and the free electrons are collected by the electrodes. The EM calorimeter provides full symmetric coverage in ϕ and is divided in a barrel and two endcaps regions, covering $|\eta| < 1.475$ and $1.375 < |\eta| < 3.200$, respectively. A presampler calorimeter is installed in the region $|\eta| < 1.8$, consisting of a thin liquid

[1]The unknown rapidity of the collision system makes the transverse energy/momentum more sensitive to the physics than its total.

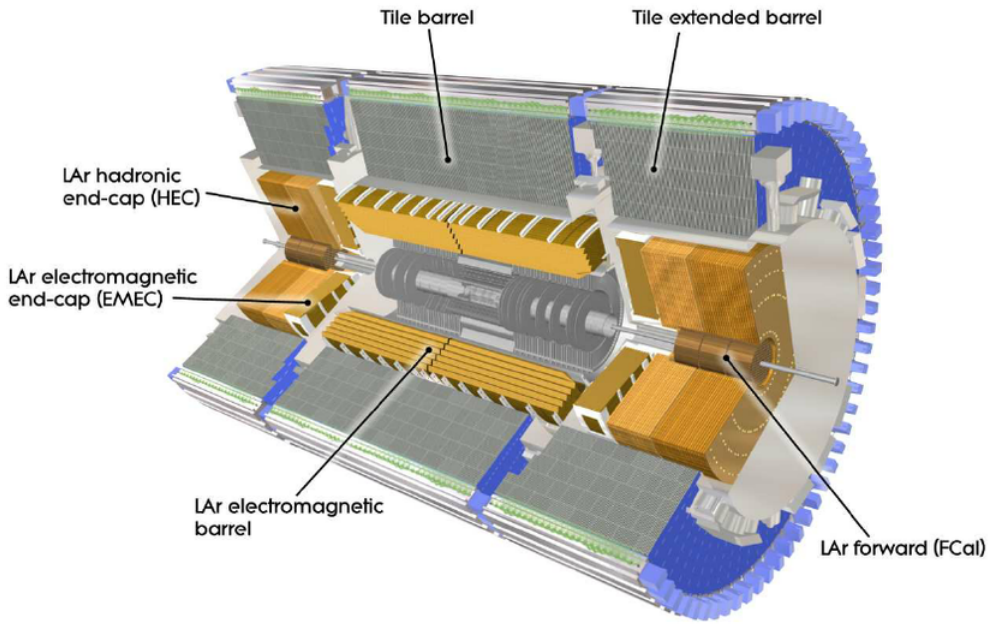


Figure 3.5: The ATLAS calorimeters [2].

argon layer, to estimate the energy lost by particles crossing the inner detector and solenoid. The EM calorimeter has a total thickness of more that 22 and $24X_0$ radiation lengths[2] in the barrel and end-caps, respectively. Its typical resolution is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$.

Apart from the pre-sampler the EM barrel calorimeter is divided longitudinally in three layers, as shown in Figure 3.6, allowing for the measurement of both the energy and direction of the electromagnetic shower. The first layer has the highest granularity, being split in $\Delta\eta \times \Delta\phi = 0.0031 \times 0.01$ cells, and providing separation of nearby photons ($e.g$ from π^0 meson decays in two photons). The second layer has a depth equivalent of 16 radiation lengths, being segmented in $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ cells, and it is expected to contain the majority of the electromagnetic shower. The third layer is the coarser in granularity, with

[2]A radiation length, X_0, is the mean distance over which a particle loses $\frac{1}{e}$ of its energy.

$\Delta\eta \times \Delta\phi = 0.05 \times 0.025$ cells. A trigger tower has the size of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The end-cap EM regions have three (two) layers between $1.5 < |\eta| < 2.5$ ($1.375 < |\eta| < 1.5$ and $2.5 < |\eta| < 3.2$, transition region and higher η region) and in general coarser granularity both in η and ϕ.

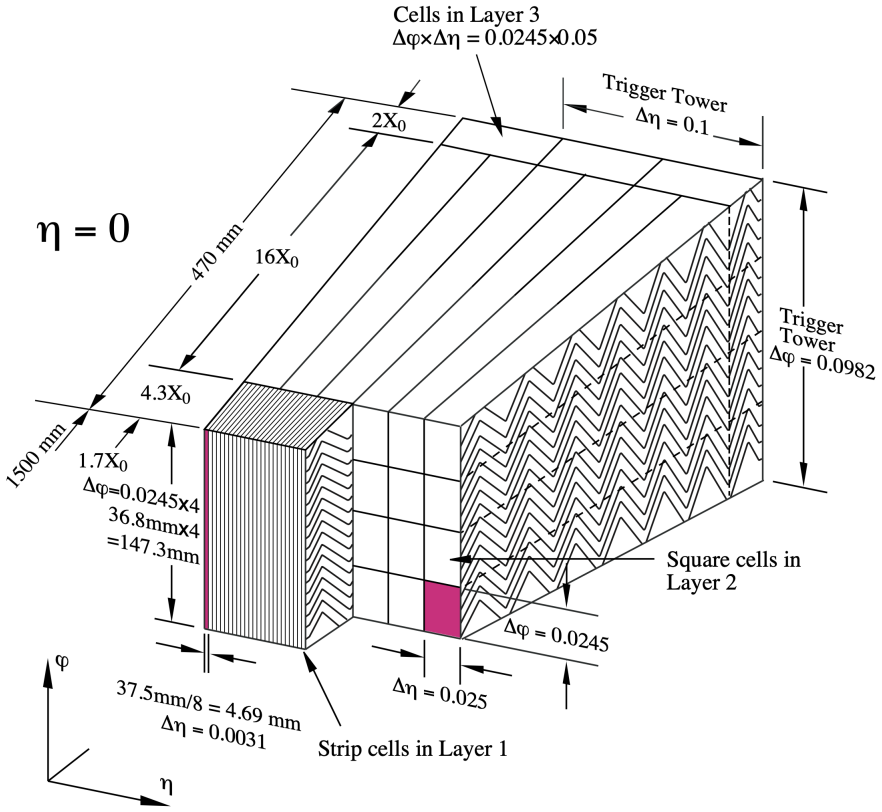


Figure 3.6: A module of the LAr electromagnetic calorimeter in the barrel region [2].

The hadronic calorimeter is comprised of three parts: the tile calorimeter (TileCal), two hadronic end-caps (HEC) and two forward calorimeters (FCal). It has a total thickness of at least 11 interaction lengths λ[3]. The TileCal and HEC have an energy resolution of $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$, while the FCAL has a resolution of $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$.

[3]An interaction length, λ, is the mean distance over which an hadron loses $\frac{1}{e}$ of its energy.

The TileCal uses scintillating tiles as active material and steel plates as absorber. The hadronic showers cause the tiles to emit ultraviolet (UV) light that is collected by wavelength-shifting optical fibres. The fibres then convert the UV light in visible light, and send it to photo-multiplier tubes. The top of the fibres opposite to the photo-multiplier is coated with an aluminium mirror to maximise the efficiency of the light collected from the tiles. The HEC (FCal) use liquid argon as active material, and copper (copper or tungsten) plates as absorbers. The detection of particles in the HEC and FCal works in the same way as for the EM calorimeter.

The TileCal has a total coverage of $|\eta| < 1.7$. It is divided in a central barrel, covering $|\eta| < 1.0$, and two extended barrels, which cover $0.8 < |\eta| < 1.7$. The TileCal is segmented in three layers, with depths of 1.4, 4.1 and 1.8 (1.5, 2.6, and 3.3) interactions lengths in the first, second and third layers of barrel (extended barrels). The first and second layers have cell sizes of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ and the third layers has cells twice as coarse in η. Hadronic showers are expected to be broader than electromagnetic ones, hence the possibility to afford a general coarser granularity in the hadronic calorimeter. The HEC has two wheels in each end-cap and covers the pseudo-rapidity region $1.5 < |\eta| < 3.2$. It has $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ cells for $|\eta| <$2.5, with the granularity halved for larger η values. The forward calorimeter covers $3.1 < |\eta| < 4.9$.

3.2.3 Muon Spectrometer

The muon spectrometer is responsible for the triggering and momentum measurement of muons deflected by large superconducting toroid magnets. This procedure is employed as muons have less radiation losses due to their larger mass, and interact minimally with the other detector components.

The magnetic bending is provided by a large barrel toroid up to $|\eta|$ <1.4, while between $1.6 < |\eta| < 2.7$ muons are deflected by two end-cap toroids. In the transition region the bending of muon tracks is achieved through a combination of barrel and end-cap fields. The muon spectrometer is equally divided in a barrel region and two end-caps, and consists of three cylindrical layers of chambers in the barrel, and three layers in the end-caps, with chambers in wheels perpendicular to the beam direction.

It is instrumented with trigger and high-precision tracking chambers. The muon triggering is performed in $|\eta| < 2.4$ via the use of resistive plate chambers (RPC) and thin gap chambers (TGC). The RPCs cover the region $|\eta| < 1.05$, while the TGCs operate in $1.05 < |\eta| < 2.4$. Muon tracking is provided by monitored drift tubes (MDT) in $|\eta| < 2.7$, and by cathode strip chambers (CSC) providing supplementary coverage between $2.0 < |\eta| < 2.7$. CSC chambers are used in the very forward regions as they are capable of coping with higher interaction rates, a condition in which the MDTs lose performance. The layout of the muon spectrometer is presented in Figure 3.7.

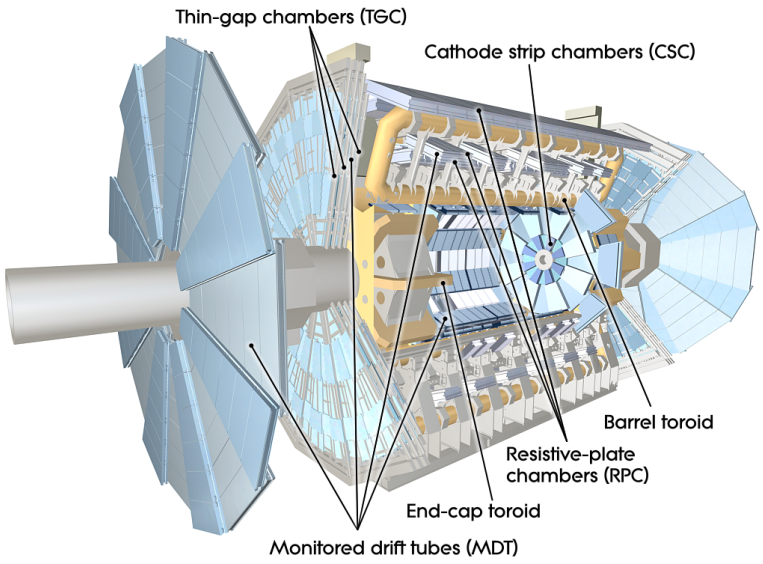


Figure 3.7: The ATLAS muon system [2].

The MDT chambers consist of two layers of aluminium drift tubes filled with argon and carbon dioxide and an anode wire made of tungsten-rhenium for readout of the ionisation of the gas mixture by traversing particles. MDTs have an hit resolution in the bending plane of 80 μm per tube and 35 μm per chamber. The CSC chambers are multi-wire proportional chambers with copper cathodes strips and anode wires oriented radially. The CSCs have a resolution of 40 μm per chamber. The RPCs are two parallel gaseous electrode plate detectors, with the gap filled with a gas mixture mainly composed of $C_2H_2F_4$. The TGCs are chambers filled with a CO_2 and n-pentane gas mixture and consist of a layer of wires held at a high voltage and sandwiched between two cathode planes. Copper pads and strips perpendicular to the wires provide the readout.

3.3 ATLAS Run 3 Level-1 Calorimeter Trigger Upgrade

The Level-1 calorimeter (L1Calo) trigger is responsible for the processing of signals from the electromagnetic and hadronic calorimeters in order to identify electron/photon and tau candidates, jet objects and missing transverse energy, and check if they pass certain trigger algorithms. These algorithms can impose thresholds on electron, photon or jet transverse energy and isolation (as later explained). For this system custom hardware processors based on FPGA technology are used.

An upgrade of the L1Calo trigger for the LHC Run 3 operation is set to finish in May 2022 and will allow for improved trigger performance in an increased pile-up environment, while exploiting a new and finer granularity trigger path digitised on the electromagnetic detector, provided by the LAr Digital Processing System (DPS), allowing for an increment in efficiency of the trigger algorithms.

The Run 3 L1Calo system is represented in yellow in Figure 3.8, and will run in parallel with the old Run 2 "legacy" system (represented in blue) for the first year of operation, after which only the upgraded architecture will be kept.

The legacy system receives analogue signals from the electromagnetic calorimeter, and digitises and calibrates them in the pre-processor. This information is then sent to the jet energy or cluster processors, where jets or electron/photon/τ-lepton candidates are identified through dedicated algorithms, resulting in Trigger Objects (TOBs) comprising the location, transverse energy and type of object identified. The TOBs are sent to the L1Topo trigger, allowing for selections for

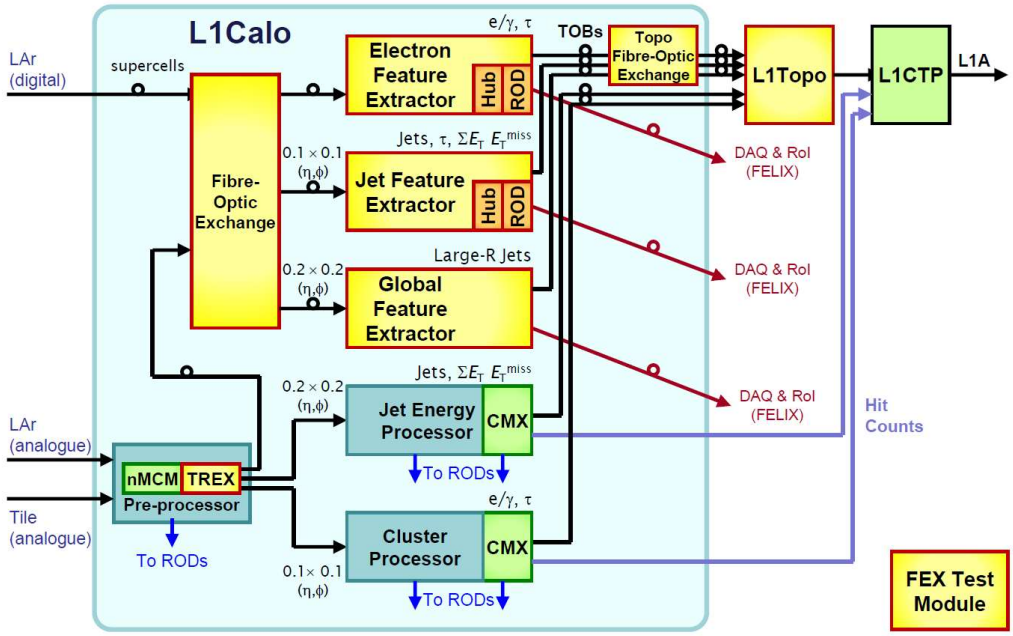


Figure 3.8: Overview of the Level-1 trigger system planned for Run 3 [47].

instance based on angles as well as E_T. L1Topo creates combined trigger objects which are transmitted to the Central Trigger Processor (CTP) that, based on the criteria that were satisfied or failed, decides whether the event is accepted or not.

The Run 3 system works in a conceptually similar way but without need for a pre-processor for the electromagnetic signals, as those signals reaching L1Calo are already digitised by the DPS. Before the upgrade the signals from the electromagnetic calorimeter sent to the L1Calo trigger were encoded in 0.1×0.1 (in η and ϕ coordinates) electromagnetic and hadronic trigger towers, formed by analogue summation of calorimeter cells. In Run 3, however, the DPS will provide electromagnetic information from up to ten "supercells" within each tower, where a supercell is the sum of four or eight calorimeter cells, as shown in Fig-

ure 3.9. Layer-0 and layer-3 (pre-sampler and back sample of the electromagnetic calorimeter) of each tower will contain one supercell, while the first and second layer will have a finer granularity, with four supercells each. The hadronic towers continue to be provided by the legacy pre-processor, constituting each a single cell.

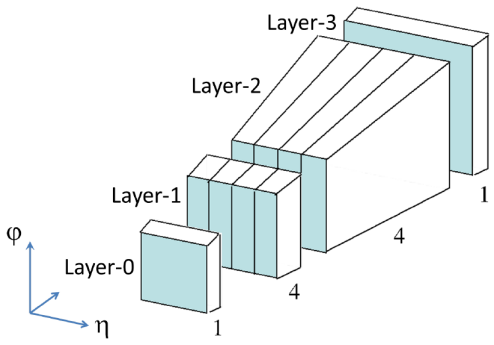


Figure 3.9: The trigger granularity from each 0.1 × 0.1 trigger tower after the upgrade of the electromagnetic calorimeter electronics, resulting in ten supercells. [47].

The information in the trigger towers is analysed by the new Feature Extractor (FEX) modules. These are responsible for producing TOBs, with each FEX having a different goal. The electron Feature Extractor (eFEX) will identify electron/photon and τ-lepton candidates, the jet Feature Extractor (jFEX) will search for jet objects and missing transverse energy, and the global Feature Extractor (gFEX) will identify large radius jets. The jFEX and gFEX modules do not use the finer granularity described above. Trigger objects will include the position and energy information of the candidate object, as well as the tightest isolation discriminant thresholds satisfied by the candidate. The TOBs are passed to L1Topo as input to topological algorithms, and the results are sent to the CTP.

The eFEX system, in particular, will consist of 24 modules covering $|\eta| \leq 2.5$ and the whole ϕ region, with each eFEX returning up to six TOBs. Up to six TOBs will be returned by the electron/photon (e/γ) algorithm and up to six other TOBs by the tau algorithm.

The e/γ algorithm is a 3×3 trigger tower window sliding algorithm (in η and ϕ coordinates, respectively), with the central tower supercells being part of a "core" region, surrounded by an "environment" region. This window is depicted in Figure 3.10. As the e/γ algorithm uses a sliding window all towers are considered as "core" regions at a certain point of the algorithm. The e/γ TOBs are then built from the layer-2 of the central tower in the 3×3 tower window. A "seed" finder procedure compares each of the 4 core cells with the surrounding cells and looks for a local maximum in transverse energy, E_T, following the conditions shown in Figure 3.11. If two maxima exist and have the same E_T, the one to the right is kept.

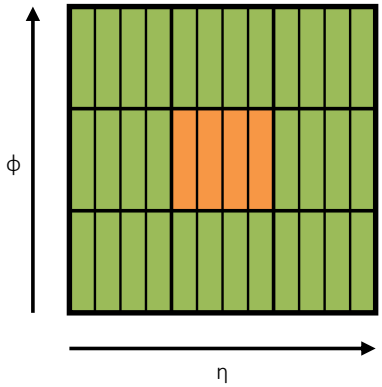


Figure 3.10: 3×3 (in η and ϕ) tower window in layer-2 for e/γ algorithm. Core (environment) supercells represented in orange (green).

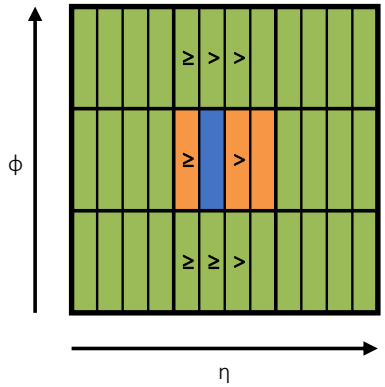


Figure 3.11: Seed finder conditions for each of the four supercells in the central tower. Seed candidate represented in blue.

A 3 $\times$ 2 supercell cluster (in η and ϕ) is formed by taking, in addition to the seed, its closest neighbours in the η direction, as well as the most energetic neighbours in ϕ. The group of three ϕ supercells directly above and below the seed are considered, and the one with the highest associated E_T is selected. If top and bottom ϕ neighbours have the same energy the top ones are chosen by convention. An example for one candidate seed is shown in Figure 3.12. The transverse energy of this layer-2 object is summed with the respective values in the corresponding cells from layer-0, layer-1 and layer-3, as represented in Figure 3.13, to give the TOB's total transverse energy.

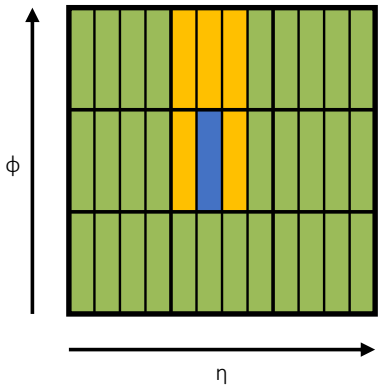


Figure 3.12: 3×2 (in η and ϕ) supercell cluster in layer-2. Core supercells represented in orange and environment supercells in green. Seed candidate represented in blue.

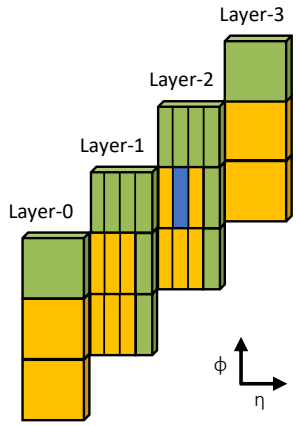


Figure 3.13: Total TOB transverse energy. Core supercells represented in orange and environment supercells in green. Seed candidate represented in blue.

The dimensions of the clusters take into account the fact that photons and electrons produce narrow electromagnetic showers (and therefore narrower clusters of E_T), while jets are more spread out in space and are associated to higher rates. Consequently, it is possible to discriminate electrons and photons from jets, especially given the finer granularity of the electromagnetic signals provided by the

DPS. The seed is identified in layer-2 as that is where most of the electron energy is deposited.

The eFEX system is under UK responsibility, with Birmingham participating in the testing of the boards, as well as in the development of the firmware, trigger algorithms and offline and online software frameworks. The firmware engineers are responsible for implementing the trigger algorithms in the boards. The online software, on the other hand, concerns the testing and monitoring of the systems during operation and data-taking periods. Distinctively, the offline software is for reproducing the functionality of the feature extractors with extremely accurate precision (bitwise simulation - mimicking the boards down to the FPGA level), allowing the monitoring of the trigger performance and the modelling of the trigger performance in physics studies.

3.3.1 Bitwise Simulation

Within the scope of the offline software a new bitwise framework was developed reflecting the upgraded L1Calo system with the same accuracy as the firmware. This framework was validated with an alternative offline implementation.

One important aspect to reproduce from the electronics was the multi-linear digitisation scheme of the energy of the eFEX supercells. The electromagnetic calorimeter electronics send the supercells E_T values compressed into a multi-linear code to the eFEX boards, which have to be linearised before use in the eFEX algo-

rithms. This code is associated with different energy ranges, as shown in Figure 3.14. Each of these corresponds to an energy step of different size: 25, 50, 100, 200, 400 and 102,400 MeV, with the lower E_T range associated to the smaller energy step. A requirement of $E_T > 100$ MeV is applied in the supercells to suppress electronic noise.

Higher E_T values are associated to a coarser granularity, as from a trigger point of view very energetic objects are more likely to correspond to objects of interest and so require less stringent criteria. Special codes are allocated for cases with invalid or saturated data. The latter correspond to events with very energetic tower supercells, and in line with the above reasoning, these events pass the trigger straight away.

The supercells are used to produce TOBs, as described above, which will contain information on the isolation criteria satisfied. These requirements exist due to limited bandwidth, imposing that only the relevant objects should be kept, and consist of specific algorithms (or variables) that discriminate the objects of interest from the rest. Electron/gamma TOB candidates in the eFEX will be checked against three variables, R_η, R_{had} and w_{stot}.

The most discriminating isolation variable corresponds to R_η, which is defined as in Equation 3.4 and represented in Figure 3.15. It compares a 3×2 (in η and ϕ) supercell cluster versus a 7×3 supercell cluster in layer-2. The ratio of these values is subtracted from unity so as to have a cut below a certain threshold, in accordance to other variables like R_{had}. This variable exploits the narrower nature

CODE	Code meaning / Energy range, MeV	Range size, MeV	Comment
0	<EMPTY>		no data available
1	< -750		Linear region #1
2	[-750 . . . -725)	25	
3	[-725 . . . -700)	25	
	. . .		
95	[1,575. . . 1,600)	25	
96	[1,600 . . . 1,650)	50	Linear region #2
97	[1,650 . . . 1,700)	50	
	. . .		
191	[6,350 . . . 6,400)	50	
192	[6,400 . . . 6,500)	100	Linear region #3
193	[6,500 . . . 6,600)	100	
	. . .		
383	[25,500 . . . 25,600)	100	
384	[25,600 . . . 25,800)	200	Linear region #4
385	[25,800 . . . 26,000)	200	
	. . .		
767	[102,200 . . . 102,400)	200	

CODE	Code meaning / Energy range, MeV	Range size, MeV	Comment
768	[102,400 . . . 102,800)	400	Linear region #5
769	[102,800 . . . 103,200)	400	
	. . .		
1011	[199,600 . . . 200,000)	400	
1012	[200,000 . . . 302,400)	102,400	Linear region #6 (COARSE)
1013	[302,400 . . . 404,800)	102,400	
	. . .		
1019	[916,800 . . . 1,019,200)	102,400	
1020	>= 1,019,200		
1021	<RESERVED>		reserved code
1022	<INVALID>		invalid data
1023	<SATURATED>		saturated data

Figure 3.14: eFEX energy encoding scheme.

of electromagnetic showers produced by the electrons to discriminate against the jet backgrounds, as an electron is expected to have most of its energy contained within the smaller cluster, while a jet will typically result in significant energy deposits also in the larger cluster. Hence, the ratio of the smaller and larger clusters should be close or equal to 1 for an electron/gamma candidate, and the R_η value should be close or equal to 0. If $R_\eta < t_\eta$ then the condition is satisfied for that threshold working point. The typical t_η threshold values for well isolated clusters are ~ 0.02.

$$R_\eta = 1 - \frac{E_{\text{Cluster}}}{E_{\text{Environment}} + E_{\text{Cluster}}} = 1 - \frac{E_{Layer-2\ (3\times2)}}{E_{Layer-2\ (7\times3)}}\,. \tag{3.4}$$

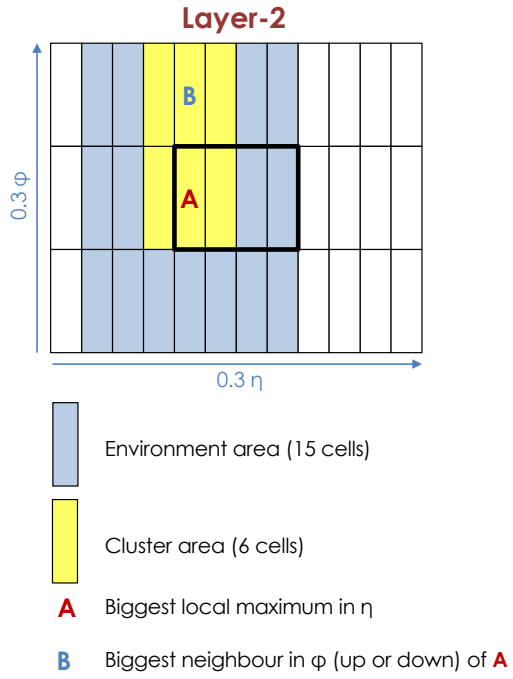


Figure 3.15: The eFEX R_η cluster (yellow) and environment (blue) in layer-2, where the seed cell **A** is the biggest local maximum in η and the one above **B** is the biggest neighbor in ϕ (up or down) of the seed **A**.

An example of an R_η distribution for a simulated signal sample of a Z boson decaying to a pair of electrons is shown in Figure 3.16. The expected behaviour of the variable for real electrons can be seen in this plot, with a peak at very low values of R_η. On the other hand, the distribution for a simulated background sample of multi jets, shown in Figure 3.17, is characterised by a broad shape, due to jets being mistakenly taken as electrons and naturally not having the requested isolation.

In the firmware, however, ratios are not implemented due to the associated com-

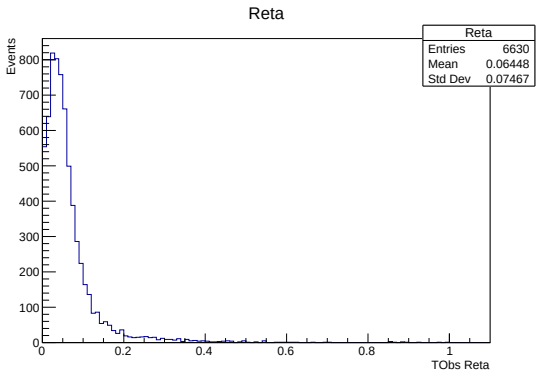


Figure 3.16: R_η distribution for a simulated signal sample of a Z boson decaying to a pair of electrons.

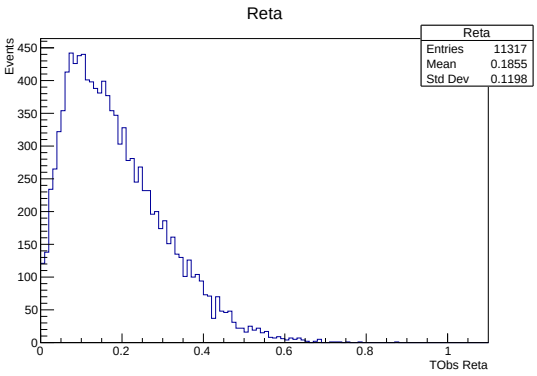


Figure 3.17: R_η distribution for a simulated background sample of multi jets.

plexity of division operations. The $R_\eta < t_\eta$ condition is therefore redefined as follows, where the parameter $C_\eta = (1 - t_\eta)/t_\eta$:

$$E_{\text{Cluster}} > C_\eta \times E_{\text{Environment}} \,. \tag{3.5}$$

Figure 3.18 shows the values of the C_η parameter as a function of the $1 - R_\eta$. Comparisons with the thresholds mentioned above in the text can be obtained by calculating $t_\eta = 1 - R_\eta$. Typical values of C_η are approximately 100.

The R_{had} hadronic condition takes the ratio between the energy deposited in the hadronic calorimeter to the total energy of the electromagnetic and hadronic towers, as defined in Equation 3.6 and shown in Figure 3.19. Electrons are expected to be associated with very small or null values of R_{had}, as all the energy of the electron candidate should normally be deposited in the electromagnetic part of the tower.

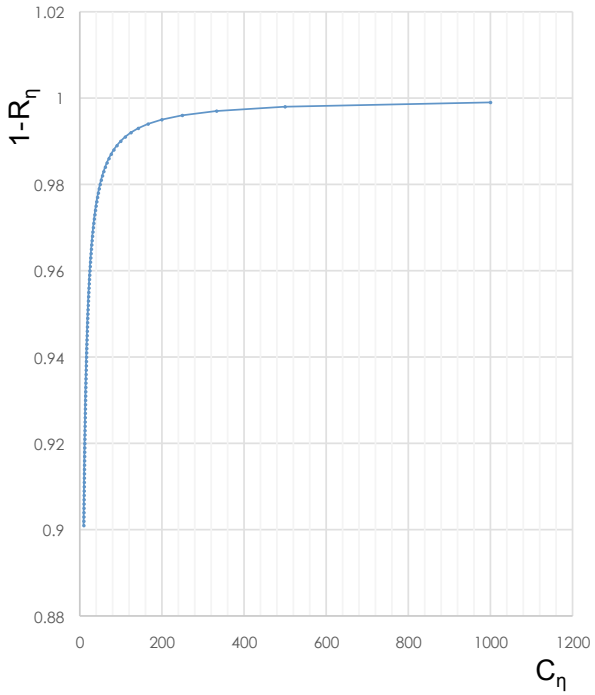


Figure 3.18: $1 - R_\eta$ ratio as a function of C_η.

$$R_{\text{had}} = \frac{E_{\text{had}}}{E_{\text{had}} + E_0 + E_1 + E_2 + E_3} \,. \tag{3.6}$$

The centre of the cluster is chosen in layer-2 according to the seed **A** found by the seed finding algorithm. The E_n areas are defined in a $0.1\eta \times 0.3\phi$ area in layers- 0 and 3, and 3×3 supercells in layers- 1 and 2, as shown in Figure 3.19. Towers in layers- 0 and 3 have only one supercell each, and the ones containing the most energy deposited around the seed will always be the ones centred in η, given the seed is always on the central tower of the window. The energy of the hadronic layer, E_{had} is defined as a $0.3\eta \times 0.3\phi$ area, as each tower has only one hadronic supercell. If $R_{\text{had}} < t_{\text{had}}$, the condition is satisfied for a certain working point. The typical t_{had} threshold values of well isolated clusters are ~ 0.04.

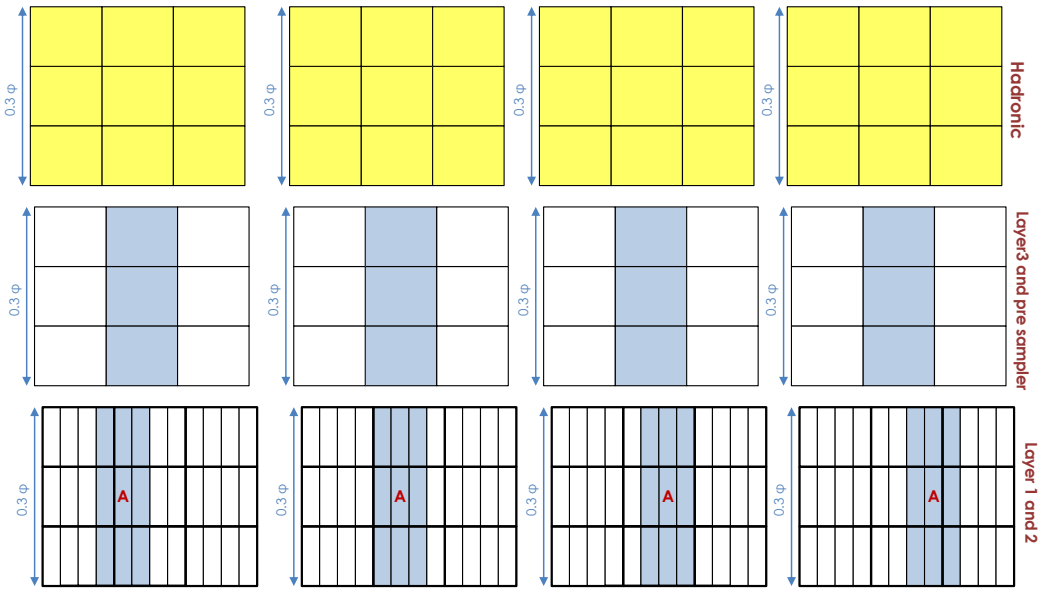


Figure 3.19: Cell selection for the R_{had} hadronic condition.

The R_{had} distribution for a simulated signal sample of a Z boson decaying to a pair of electrons is shown in Figure 3.20, with a sharp peak at values around zero. The distribution for a simulated background sample of multi jets, shown in Figure 3.21, contains a large tail resulting of jets misidentified as electrons.

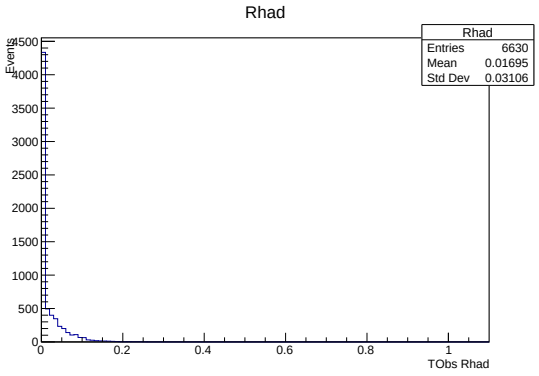


Figure 3.20: R_{had} distribution for a simulated signal sample of a Z boson decaying to a pair of electrons.

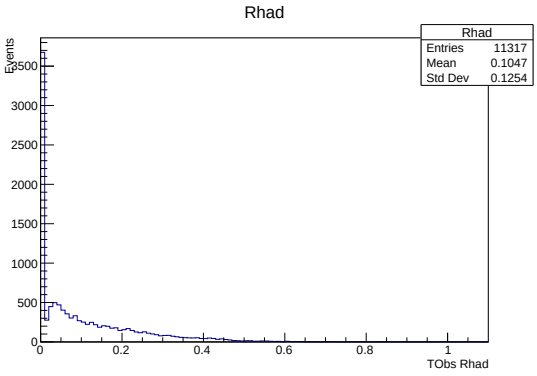


Figure 3.21: R_{had} distribution for a simulated background sample of multi jets.

Again, for the actual firmware implementation the $R_{\text{had}} < t_{\text{had}}$ condition is redefined, considering the parameter $D_{\text{had}} = (1 - t_{\text{had}})/t_{\text{had}}$:

$$D_{\text{had}} = \frac{1 - R_{\text{had}}}{R_{\text{had}}} = \frac{E_0 + E_1 + E_2 + E_3}{E_{\text{had}}} \,. \tag{3.7}$$

Using D_{had} the same condition on R_{had} becomes:

$$E_0 + E_1 + E_2 + E_3 > D_{\text{had}} \times E_{\text{had}} \,. \tag{3.8}$$

The D_{had} redefined parameter as a function of the R_{had} is shown in Figure 3.22. The typical value of D_{had} is approximately 20.

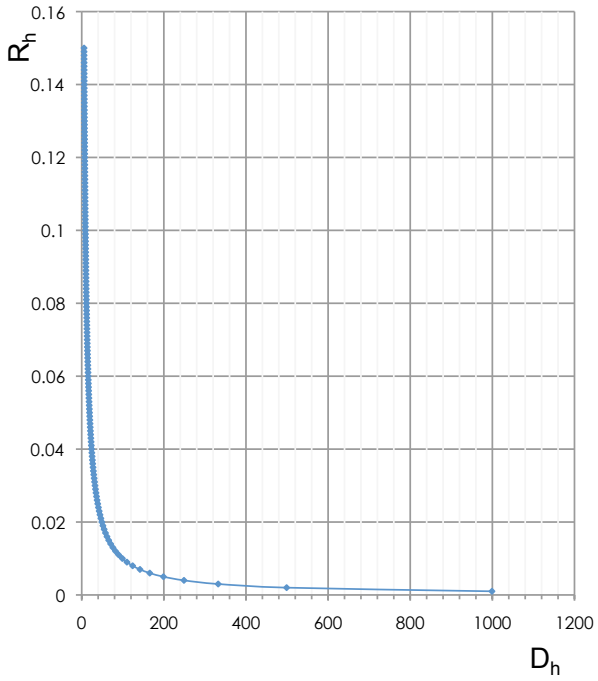


Figure 3.22: D_{had} parameter as a function of the R_{had} parameter.

The cluster width w_{stot} is defined as in Equation 3.9, where E_i is an energy sum

over five supercells in η in layer-1 and i_{max} is the index of cell **A** defined in the R_η condition.

$$w^2_{\text{stot}} = \frac{\sum E_i (i - i_{max})^2}{\sum E_i} \,. \tag{3.9}$$

The $(i - i_{max})^2$ coefficient can be 0, 1 or 4, as supercells can participate in the sum according to the following:

$i = i_{max}$: does not participate in the sum

$i = i_{max} \pm 1$: participates in the sum as is (multiplied by 1)

$i = i_{max} \pm 2$: participates in the sum shifted by 2 bits (multiplied by 4)

This privileges electron/gamma candidates with narrower shower tails in layer-1, multiplying by increasing factors wider energy deposits. Again if $w_{\text{stot}} < t_{\text{stot}}$, the condition is satisfied at a certain working point.

A schematic diagram of this condition is presented in Figure 3.23. The area of this condition does not depend on **B**. Moreover, cells indicated by 4 are multiplied by 4 (bit shifted by 2) in the sum.

The division is replaced by a multiplication in the same fashion as the other cluster conditions, with the $w_{\text{stot}} < t_{\text{stot}}$ condition becoming:

$$\sum E_i > 1/Threshold \times \sum E_i (i - i_{max})^2 \,. \tag{3.10}$$

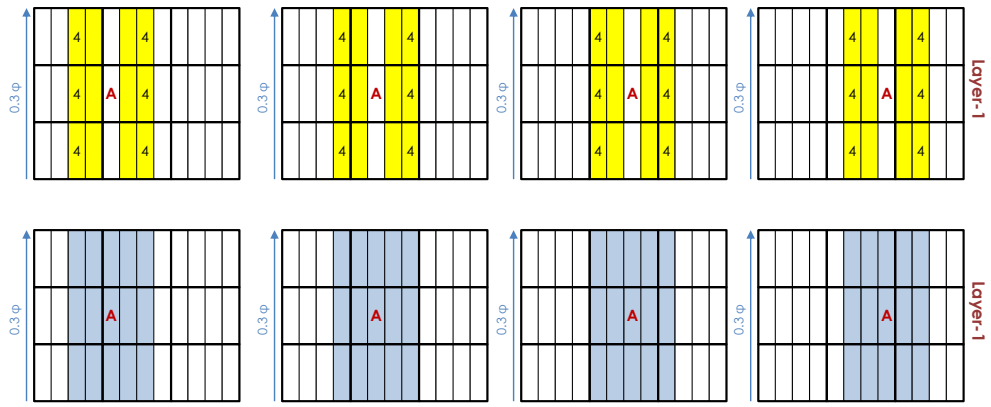


Figure 3.23: The w_{stot} isolation condition.

The firmware and equivalent bitwise isolation conditions can therefore be rewritten, as described above, in the general form:

$$\text{Core} \geq t_{\text{var}} \cdot \text{Environment}\,, \tag{3.11}$$

with the "core" corresponding to the left-hand side of Equations 3.5, 3.8 and 3.10 and the "environment" to the right-hand side component of these equations multiplying the respective variable threshold value t_{var}.

Four working points (WPs) are defined for the isolation variables based on their firmware-converted core, environment and threshold values: *fail*, *loose*, *medium* and *tight*. The *tight* WP has the strictest threshold and the *loose* WP the looser threshold. The *fail* working point is set when the object does not pass more stringent conditions. Also if the right-hand side of Equation 3.11 overflows, $i.e$, it corresponds to a transverse energy value above what can be stored in the electronics, or the environment overflows, the *fail* working point is directly set for all

the isolation conditions. If the left-hand side of Equation 3.11 overflows or if the object has E_T above an adjustable high threshold (60 GeV at the moment), the *tight* working point is directly satisfied for all variables. The working points are encoded using a pair of bits, with the *fail*, *loose*, *medium* and *tight* working points corresponding to 0, 1, 2 and 3 respectively.

The TOB is encoded in a 32-bit word, with the total word containing bits correspondent to:

- Position: η and ϕ coordinates within the FPGA, FPGA number, seed supercell number, bunch crossing number;
- E_T of object;
- Isolation variable results: indicating tightest working point condition passed for each isolation variable;
- Up/Down cluster orientation flag: determined by layer-2 object taking supercells from tower above or below central tower (see example in Figure 3.12).

3.3.2 Monitoring

The data taking periods require dedicated shifts from technical experts to ensure that the events collected at the detector were recorded under stable beam conditions and with all relevant detector sub-systems operating normally.

In order to assist this a monitoring framework exists in the ATLAS athena monitoring software, providing the shifters with quick and simple histograms from the data acquired and allowing quick detection of possible problems in the software or hardware. The monitoring software also compares the response of the system with the results of the bitwise simulation to identify potential issues.

These histograms plot the information contained within the TOBs created by the trigger system mentioned earlier, namely the position, transverse energy and tightest isolation working points (WPs) passed by the object. The position and transverse energy of the supercells in the trigger towers can be equally plotted and verified.

Examples of the monitored variable distributions are shown for the e/γ and τ-lepton trigger objects in Figures 3.24 and 3.25, respectively. These histograms were produced from the bitwise simulation output using a $Z \to \ell\ell$ ($\ell = e, \mu$) simulated sample, given the absence of data, with the bitwise framework being in the debugging process at the time. The distributions presented nevertheless reveal hints in the right direction, with the trigger objects having associated tighter working points as expected for the used simulated sample. R_η can be seen as providing the most discriminating power, followed by R_{had} and w_{stot}.

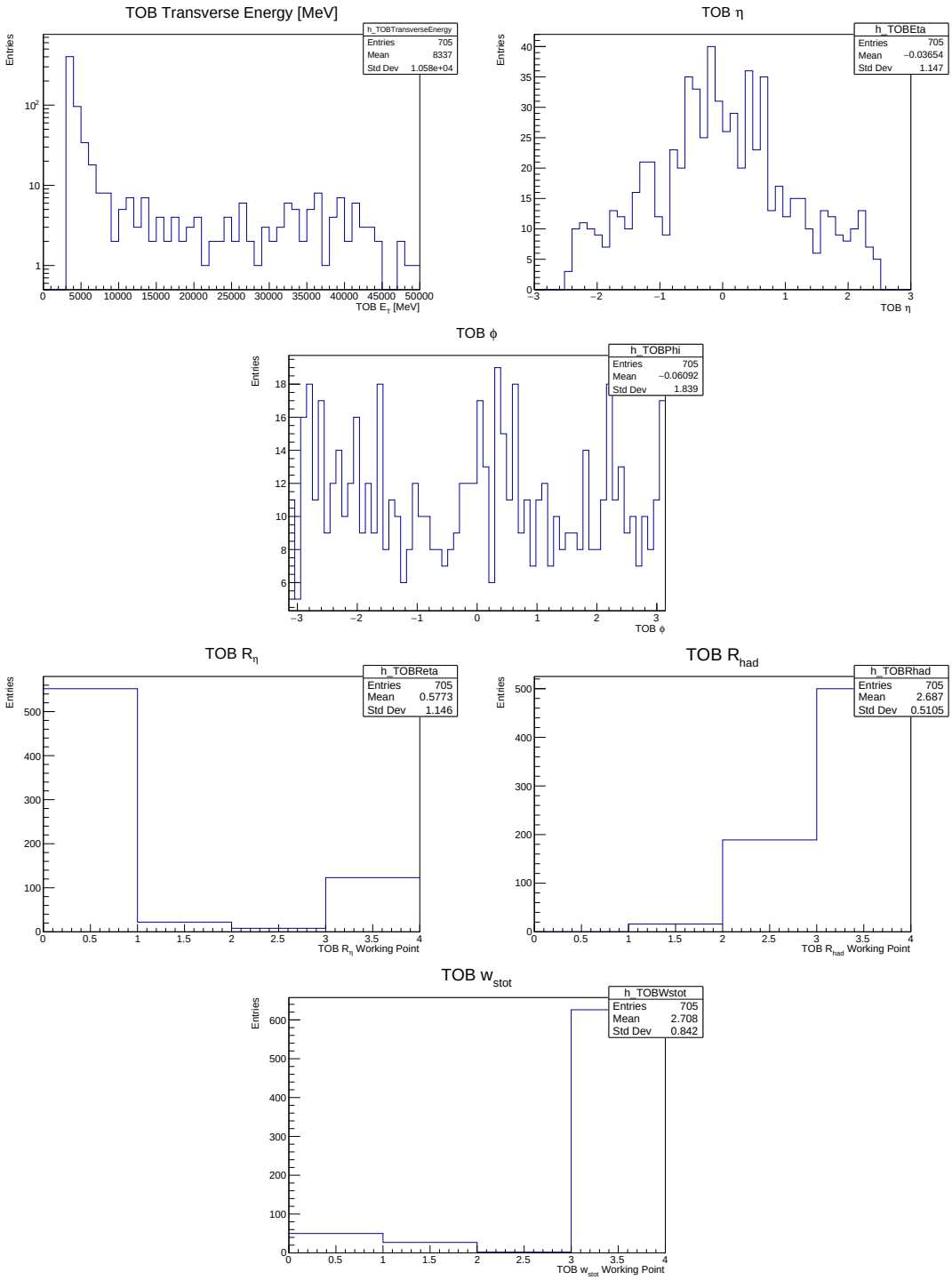


Figure 3.24: Selection of e/γ TOB monitored variables. From left to right, and top to bottom: E_T, η, ϕ, and strictest working point passed for R_η, R_{had} and w_{stot}.

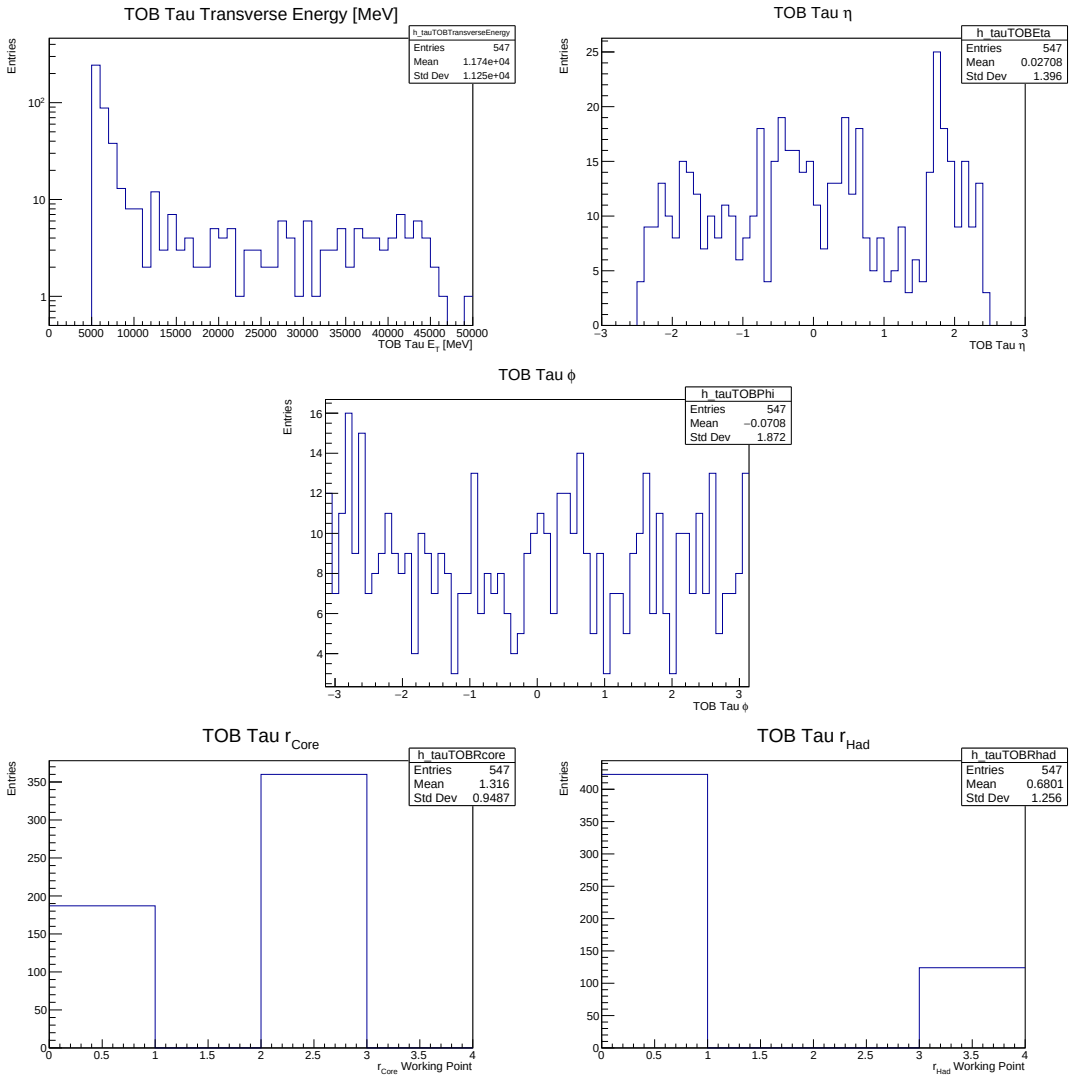


Figure 3.25: Selection of τ-lepton TOB monitored variables. From left to right, and top to bottom: E_T, η, ϕ, and strictest working point passed for r_{core} and r_{had}.

3.4 The ATLAS detector at the High-Luminosity Large Hadron Collider

The LHC will undergo a major upgrade between 2026 and 2028, leading to the High-Luminosity Large Hadron Collider (HL-LHC) [48]. The HL-LHC is expected to start operating in 2029 at a centre-of-mass energy of $\sqrt{s} = 14$ TeV and an increased instantaneous luminosity of $5 - 7.5 \times 10^{34}$ cm^{-2}s^{-1}. It aims to collect $250 - 350$ fb^{-1} of proton-proton collision data per year, up to a total of $3000 - 4000$ fb^{-1} in about twelve years of operation. The HL-LHC will operate at a reduced 25 ns bunch spacing, with an increased average number of $140 - 200$ simultaneous interactions per bunch predicted.

The increased instantaneous luminosity will be achieved predominantly via the replacement of the LHC superconducting quadrupole magnets by more powerful ones, which will allow stronger focusing of the beams in the crossing points. A crab crossing scheme will also be introduced, tilting the proton bunches by giving them a transverse momentum, and in that way maximizing the overlap area of the two bunches, resulting in an increased probability of head-on collisions of the beams. The main LHC dipole magnets, responsible for bending the proton trajectories at the LHC, will also be replaced by new ones capable of providing a magnetic field of 11 T (versus the 8.3 T at the LHC).

The increased data statistics will allow for a reduction of the statistical uncertainties in the analyses, as well as their systematic errors, due to larger data samples

available for the calibration procedures.

The ATLAS detector will be upgraded for the HL-LHC in order to cope with the increased radiation and pile-up, while maintaining an optimal performance [49]. The ATLAS HL-LHC hardware-based trigger system is designed to reduce the 40 MHz event rate to 1 MHz (100 kHz at the current LHC operation), with the upgraded HLT then expected to further reduce the event rate to 10 kHz. The inner detector will be replaced by an all-silicon high granularity tracker, the ITK, covering $|\eta| < 4$ (while the LHC ID covers $|\eta| < 2.5$). Its associated lower mass will allow for a reduction of the effect of photon conversions and multiple parton interactions. A high granularity timing detector will be installed in front of the end-cap calorimeters, covering $2.4 < |\eta| < 4.0$. This detector will improve the reconstruction of physics objects in the forward region, reducing the effect of pile-up and improving the matching of tracks to the primary vertex. The muon system will see its trigger efficiency improved, via the installation of more muon chambers.

Chapter 4

Physics Objects

The $VH(H \rightarrow c\bar{c})$ analysis considers various physics objects that are identified and reconstructed in the detector according to their specific signatures using dedicated techniques. This section presents an overview of these methods and procedures employed in the ATLAS experiment.

4.1 Tracks and Interaction vertices

The passage of charged particles creates hits in the inner detector, allowing for the reconstruction of tracks. The procedure starts from the innermost layers of the ID, with hits in the pixel detectors being used to create of three-dimensional space points. Sets of three space points form seeds, which are exploited by a combinatorial Kalman filter [50] to forms track candidates. The reconstruction procedure then expands to the TRT and matching hits are combined with the

silicon ones. The same seed can contribute to at most two track candidates, with ambiguities, incomplete or fake tracks (based on measurements originating from several particles) are resolved/suppressed by placing requirements on the track candidates. Reconstructed tracks must have $p_{\mathrm{T}} > 400$ MeV, at least seven silicon hits and at most one silicon hit shared by multiple tracks.

The primary interaction point, or primary vertex, is reconstructed from tracks in an iterative procedure. Tracks likely stemming from the primary vertex must satisfy requirements based on their transverse (d_0) and longitudinal (z_0) impact parameters (shortest distances between the track and the primary vertex in the transverse or z-direction, respectively), as well as the respective σ_{d_0} and σ_{z_0} uncertainties. Tracks more than 7σ away from any vertex define a new one, with each vertex being required to contain at least two tracks. The primary vertex is then defined as the vertex with the highest sum of squared transverse momenta of the associated tracks.

4.2 Electrons

Electrons are reconstructed from energy clusters with $E_T > 2.5$ GeV in the electromagnetic calorimeter that are matched to at least one ID track [51]. Tracks must be associated to the primary vertex. In case of multiple tracks associated to the energy cluster, the closest one is selected.

Isolation requirements are placed on the electrons, in view of rejecting "fake"

electrons coming from hadrons or photon conversions. These include a likelihood discriminant taking into account the shape of the electromagnetic showers in the calorimeter and the properties of the tracks matched to the energy cluster. For this discriminant three likelihood identification working points (WPs) are defined and used in analyses, a *loose*, *medium* and a *tight* one, each with a larger fake electron rejection than the previous.

Electrons must also satisfy track and calorimeter isolation requirements, with again two working points being used - *loose* and *tight*. The two isolation requirements are defined as follows:

- Track-based isolation: the transverse momentum of tracks with $p_{\mathrm{T}} > 1$ GeV and $|\eta| < 2.5$, and contained within a p_{T}-dependent ΔR ($\Delta R_{ij} \equiv \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$) cone around the electron track, is summed, with the electron track contribution not taking part of the sum. The cone radius is defined as $\Delta R = \min(10/p_{\mathrm{T}}^{electron}[\mathrm{GeV}], 0.2)$ and therefore decreases for higher p_{T} values. The ratio of this sum to the electron p_{T} then has to satisfy a threshold.
- Calorimeter-based isolation: the transverse energy of topological clusters contained within a $\Delta R = 0.2$ cone around the electron cluster is summed, with the electron E_T contribution excluded from the sum. The E_T sum or the ratio of the sum to the electron p_{T} have to pass the working points thresholds.

Additional criteria are applied to ensure the matching of electrons to the primary vertex and reduce pile-up contributions.

4.3 Muons

Muons can be reconstructed from information in the ID, calorimeters and muon spectrometer (MS), taking each contribution individually or combined between the different detectors [52]. Muons reconstructed from a combination of the MS information with ID tracks are referred to as *combined* muons. Muons reconstructed with an ID track matched to a MS partically reconstructed track are denominated as *segment tagged*. A *stand alone* muon is reconstructed from just a MS track. Muons can also be reconstructed from the combination of an energy deposit in the calorimeters and an ID track.

Muon quality criteria are applied, which consider the components of the detector crossed by the muons, as well as from how many of their layers the muons were reconstructed. Three muon quality working points are defined for analyses, a *loose*, *medium* and a *tight* one. The former uses all types of muons, while the *medium* working point uses only stand alone or combined muons. The *tight* working point considers only combined muons.

Track and calorimeter isolation requirements must be satisfied for the muons in criteria equivalent to those of the electrons. An additional track-based isolation criteria is considered for muons, with the transverse momentum of all tracks con-

tained within a $\Delta R = 0.2$ cone around the muon track being summed, with the muon track contribution not taking part of the sum. This sum then has to satisfy the working point threshold.

Additional requirements are applied to muons to establish their match with the primary vertex.

4.4 Jets

Jets are reconstructed as topological energy clusters in the calorimeters [53], using the anti-k_t jet algorithm [54], which is infrared and collinear safe, with a radius parameter of $R = 0.4$. Jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ in the analysis are classified as *central* jets, while jets with $p_{\mathrm{T}} > 30$ GeV and $2.5 < |\eta| < 4.5$ are designated as *forward* jets.

A jet energy scale (JES) calibration [55] procedure is applied on reconstructed jets, which corrects the jet direction to point to the primary vertex, corrects for pile-up contributions, and calibrates the jet four-vector to the particle-level energy scale, correcting the average energy scale based on the reconstructed energy and η of the jet. An additional step corrects the jet energy response for its dependence on the distribution of energy within the jet, the distribution of energy across different calorimeter layers, and the fraction of charged particles within the jet. Residual in-situ corrections for differences between the data and simulation are also implemented.

The jet vertex tagger (JVT) [56] likelihood-based discriminant is used to further suppress jets from pile-up, using primary vertex information, and jet and track transverse momenta. It is only available for central jets, as there is no track information above $|\eta| > 2.5$. The working point used in the analysis corresponds to an average efficiency of 92% for the jet tagging to the primary vertex.

The jet energy and direction are corrected for jets with nearby muons, to account for possible semi-leptonic decays of heavy flavour hadrons. This "muon-in-jet" correction involves only muons as any possible electron energy is deposited in the calorimeter and is already taken into account for in the jet, while this is not the case for muons as they do not interact substantially with the calorimeters. In this correction the muon and jet four-vectors are summed, and translates in an improvement of up to 6% in the jet mass resolution in the $VH(H \to c\bar{c})$ analysis, with one example illustrated in Figure 4.1.

4.5 Missing Transverse Energy

The $VH(H \to c\bar{c})$ analysis considers events with neutrinos in the 0- and 1-lepton channels. These particles do not interact with the detector and therefore can not be directly reconstructed. To obtain an estimate of their contribution the missing transverse energy in the event, $E_{\mathrm{T}}^{\mathrm{miss}}$, is computed. This quantity follows the principle of conservation of the total (transverse) momentum in the event, and consists of the negative vector sum of the transverse momenta of all the recon-

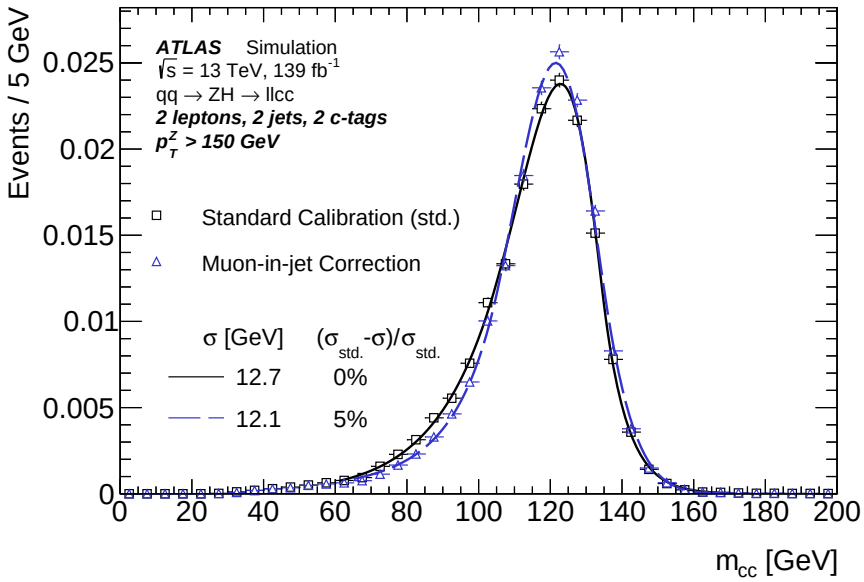


Figure 4.1: Resolution of the $m_{c\bar{c}}$ distribution with the standard jet calibration and with the additional muon-in-jet correction, for a simulated signal sample in the analysis [57, 58].

structed objects in the event (electrons, muons, τ-leptons, photons, jets), and a soft track-based term, that considers tracks matched to the primary vertex in the event and not associated to any object previously considered [59].

The missing transverse momentum, $p_{\mathrm{T}}^{\mathrm{miss}}$, is constructed from mainly tracking information, being less sensitive to pile-up effects. It uses only the objects/inner detector tracks matched to the primary vertex, with the exceptions of the electron and muon momentum, whose estimates are obtained from the more precise calorimeter and muon spectrometer measurements. It neglects however contributions from neutral particles.

4.6 Overlap Removal

The same signatures in the detector may be reconstructed at times as two different objects. In order to avoid this double counting an "overlap removal" procedure is applied in a sequence of steps, where objects removed at a certain step are not further taken into account in later steps. The overlap removal sequence is as follows:

1. τ-leptons within $\Delta R < 0.2$ of electrons are removed.

2. Removal of τ-leptons if within $\Delta R < 0.2$ of muons.

3. Any electron sharing inner detector tracks with combined muons is removed, as it can be the result of muon radiation.

4. Jets within $\Delta R < 0.2$ of electrons are removed. Electrons are removed if within $\Delta R < 0.4$ of the surviving jets, as they can be associated to semi-leptonic decays of heavy flavour hadrons.

5. Any jet within $\Delta R < 0.2$ of a muon is removed, if it has less than three associated tracks (as such a low multiplicity can be the result of muon radiation), or if the muon carries more than half the jet momentum and more than 70% of the sum of p_{T} of tracks associated to the jet. Muons are removed if within $\Delta R < 0.4$ of the surviving jets, following the same reasoning as for the electrons.

6. Removal of jets within $\Delta R < 0.2$ of τ-leptons.

4.7 Heavy Flavour Hadron Tagging

Jets originating from heavy flavour (which contain bottom and charm quarks) hadrons can be identified, or tagged, using dedicated algorithms that exploit their different properties. Not only can these algorithms distinguish charm from bottom jets, but also these with respect to light flavour (up, down, strange quarks, and gluons) jets.

Two jet flavour tagging algorithms [60] are used in the $VH(H \rightarrow c\bar{c})$ analysis, namely the DL1 c-tagger algorithm, for c-jet tags, and MV2c10, a c-tagger, for b-jet vetoes with a 70% b-jet efficiency working point, which corresponds to what is used in the ATLAS $VH(H \rightarrow b\bar{b})$ analysis [61].

DL1 corresponds to a deep neural network, while MV2c10 is a boosted decision tree, and both take as input impact parameters, secondary vertex information ($e.g.$ invariant mass, number of tracks, distance to primary vertex) and jet kinematic variables (p_{T} and η). These two algorithms are trained using hybrid $t\bar{t}$ and Z' simulated samples (a Z' is a more massive Z boson predicted in some BSM models), with jets with $p_{\mathrm{T}} < 250$ GeV coming from $t\bar{t}$ events, and jets with $p_{\mathrm{T}} > 250$ GeV from hadronically decaying Z' events, due to the falling $t\bar{t}$ spectrum above that p_{T}^{V} threshold. Jets in simulation are tagged according to their Monte Carlo generator information, and following a staged procedure. A simulated jet is classified as a b-jet if a b-hadron with $p_{\mathrm{T}} > 5$ GeV is found within $\Delta R < 0.3$ of the jet. If a b-hadron is not found, but instead a c-hadron, and sat-

isfying those criteria, the jet is classified as a c-jet. The jet is labeled as a τ-jet if instead a τ-lepton is found within $\Delta R < 0.3$ of the jet. All the remaining jets are classified as light flavour jets.

MV2c10 outputs the likelihood of a jet to be b-tagged or not, $i.e$, provides discrimination of b-jets with respect to charm and light flavour jets, while DL1 outputs the probability for a jet to be tagged as a c-, b- or light flavour jet. The MV2c10 lower threshold for a jet to be b-tagged corresponds to a 70% b-jet efficiency[1], as mentioned above. The three DL1 probability outputs are combined in a single discriminant:

$$DL_1 = \ln \frac{p_c}{f p_b + (1-f) p_l}\,, \tag{4.1}$$

where p_c, p_b and p_l are the probabilities for a jet to be c, b or light flavour tagged, and f is a weight correlating the b and light flavour jet tagging probabilities, allowing for a better b-jet or light flavour jet rejection. Given MV2c10 provides already an efficient b-jet rejection, the optimisation of DL1 on the basis of the significance of the $VH(H \to c\bar{c})$ signal leads to a preference for a better light flavour jet rejection.

The dedicated working point built from these two algorithms has an associated efficiency of 27%, 8% and 1.6% for c, b and light flavour jets, respectively. The efficiencies in data for each flavour of jets are shown in Figure 4.2. Furthermore,

[1]This working point is associated to a c- and light flavour jet tagging efficiency of 11% and 0.3%, respectively.

in the analysis jet is considered to be "c-tagged" if it passes both the c-tag and b-veto conditions, *i.e*, have a c-tag with a b-veto.

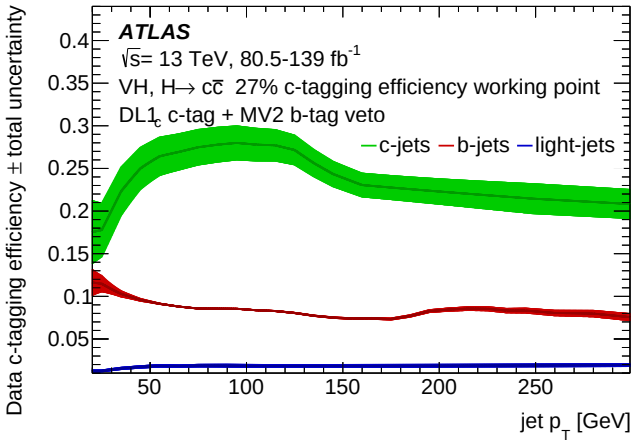


Figure 4.2: Jet-tagging efficiencies in data for different flavoured jets as a function of the jet transverse momentum. The uncertainty band corresponds to the total uncertainty associated with the c-tagging efficiency calibration in data [57, 58].

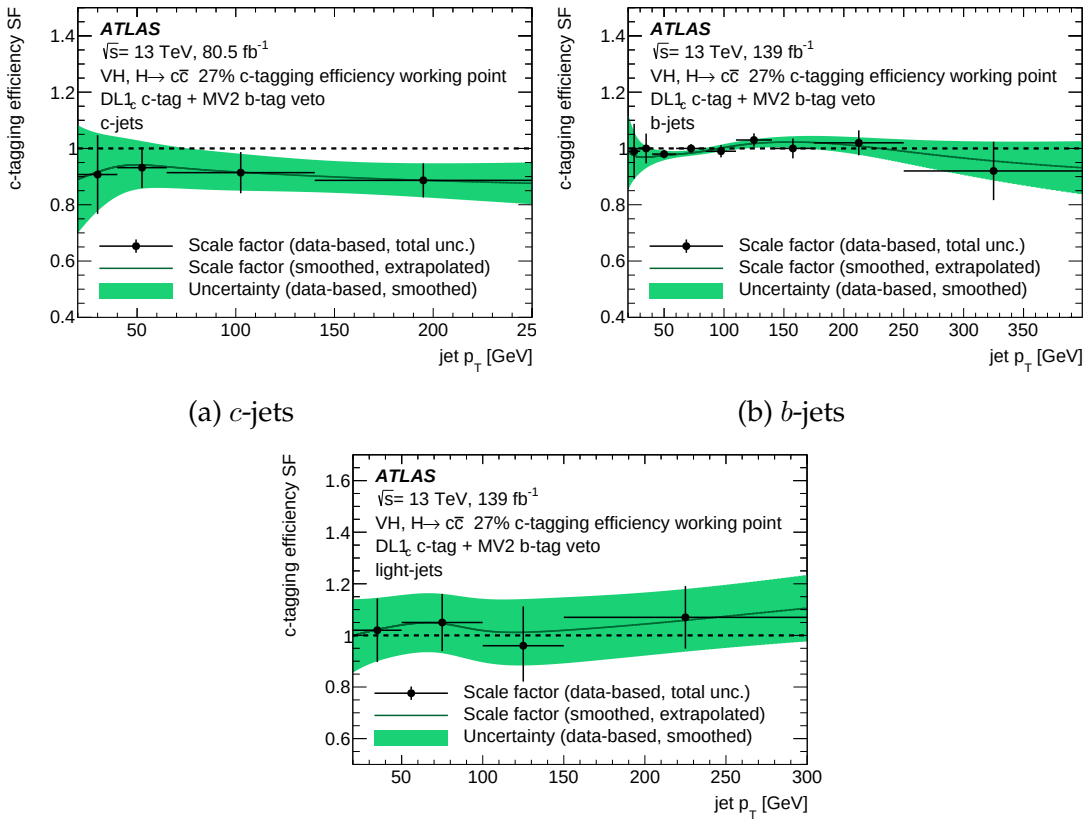
The jet tagging efficiencies in simulation (ϵ_{MC}) are corrected to reflect the respective efficiencies in data (ϵ_{data}), being measured as a "scale factor":

$$SF = \frac{\epsilon_{data}}{\epsilon_{MC}} . \tag{4.2}$$

These scale factors are derived from dedicated data-driven studies for each jet flavour, using specific control samples and methods identical to those applied to b-tagging algorithms [60, 62, 63]. PYTHIA8.2.30 is used in the derivation of the default calibrations. The same derivation procedure applies for the jet tagging inefficiencies, corresponding to jets failing the flavour tagging requirement.

The calibration scale factors for c-jets are measured with a sample of semi-leptonic decaying $t\bar{t}$ events and have a typical precision of around 10% for jets with $p_{\mathrm{T}} <$

250 GeV. A sample of di-leptonic decaying $t\bar{t}$ events is used to derive scale factors for b-jets, with a typical precision of a few percent for $p_{\mathrm{T}} < 250$ GeV. Scale factors for light flavour jets are measured with a sample of Z+jets events, in which the flavour tagging algorithm receives the inverse of the input variables, making it therefore sensitive to light flavour jets as opposed to heavy flavoured ones. The light flavour jets scale factors have a typical precision of around 15% for $p_{\mathrm{T}} < 250$ GeV. All these scale factors are generally consistent with unity, as shown in Figure 4.3.



(a) c-jets (b) b-jets

(c) Light flavour jets

Figure 4.3: Data-to-simulation calibration scale factors for each jet flavour [57, 58].

The samples used in the analysis are not only generated with the use of PYTHIA8.2.30, but also with other MC generators which have different hadronisation models,

namely SHERPA2.2.1 and HERWIG7.1.3. The particular hadronisation models of each generator will translate into different flavour tagging efficiencies, leading to the derivation of simulation-to-simulation scale factors to account for these differences and to allow for the working point to be calibrated for the various generators. These scale factors are derived for each jet flavour, and are obtained by comparing the efficiencies in each alternative MC sample to the ones in PYTHIA8.2.30, and are generally of the order of 5-15%.

The usual tagging procedure, "direct tagging", resulting from selecting events with jets that pass the specific thresholds of the flavour tagging algorithms, and the associated relatively low efficiency for c-jets and high rejection of b- and light flavour jets, leads to a significant loss of simulated events and statistical power. In order to reduce the MC statistical uncertainties associated with direct tagging an alternative procedure is implemented for the flavour tagging of the background processes, denominated "truth flavour tagging", in which the tagging (in-)efficiencies of the jets are applied as event weights. The truth flavour tagging (TT) implementation is validated by comparisons to the direct tagging simulation.

The TT event weights are calculated using efficiency maps, which parametrise the tagging efficiency as a function of the jet p_{T} and η, and are derived using a sample of $t\bar{t}$ MC events. The truth flavour c-tagging +b-veto procedure is only implemented for the two leading jets in p_{T}, with the b-veto on the remaining jets being applied using direct tagging. The TT event weights are therefore defined

as:

$$\text{TT weight (1 } c\text{-jet events)} = \epsilon_1 \times (1 - \epsilon_2) + (1 - \epsilon_1) \times \epsilon_2 \,, \tag{4.3}$$

$$\text{TT weight (2 } c\text{-jet events)} = \epsilon_1 \times \epsilon_2 \,, \tag{4.4}$$

where ϵ_1 and ϵ_2 are the tagging efficiencies of each of the two p_{T} leading jets.

The implementation of the truth flavour tagging procedure is further detailed in Section 5.6.1, comprising the associated corrections and uncertainties.

Chapter 5

Search for Higgs Boson Decays to Charm Quarks

The latest direct search for Standard Model Higgs boson decays to a pair of charm quarks using the ATLAS detector [57] is presented in this chapter, probing the Higgs boson Yukawa couplings to the second generation of fermions. This analysis improves the results from the previous search performed by the ATLAS Collaboration [64], with the most recent search using a larger dataset of proton-proton collisions and improved analysis design and techniques. An extrapolation of the latest ATLAS analysis to the expected HL-LHC scenario [65] is also presented in this chapter.

5.1 Introduction

The first search for Higgs boson decays to charm quarks was performed by the ATLAS collaboration in 2018 [64], using an integrated luminosity of 36.1 fb^{-1} of proton-proton (pp) collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, recorded in 2015 and 2016 with the ATLAS detector. The production of a Higgs boson produced in association with a Z boson was targeted, with the Z boson decaying into charged leptons, namely electrons and muons. The presence of vector bosons in the final state allows for suppression of multi-jet backgrounds, while their decays into leptons allow for a convenient trigger strategy. The ZW and ZZ processes were also measured, for validation of the analysis.

In this analysis the events are required to have two charged leptons of the same flavour, defining a 2-lepton channel. The two leptons need to have opposite charge in case of dimuon events, while this requirement is not specified for dielectron events due to the non-negligible charge misidentification rate of electrons.

The phase-space is further split in terms of the transverse momentum of the vector boson (V = Z, W) produced in association with the Higgs boson, p_{T}^{V}, resulting in two regimes: $75 < p_{\mathrm{T}}^{V} < 150$ GeV and $p_{\mathrm{T}}^{V} > 150$ GeV.

Events with two or more jets are considered, and are classified depending on the number of c-tagged jets, with two categories being used, namely 1 and 2 c-tags regions. The tagging of the jets uses two boosted decision tree-based algorithms, using as input impact parameters, secondary vertex information and jet kine-

matic variables. The algorithm outputs one parameter discriminating b- versus c-quarks and another discriminating c- versus $light$-quarks, with associated tagging efficiencies of 41%, 25% and 5% for c, b and light flavour jets respectively. Only the two highest-p_{T} c-jets are used to reconstruct the Higgs boson invariant mass, $m_{c\bar{c}}$.

In order to reduce statistical uncertainties due to the limited size of the simulated samples, the respective events are weighted by the tagging efficiency of the jets, instead of applying a direct cut on the flavour tagging algorithm's discriminants. These efficiencies are parametrised as a function of the jet flavour, p_{T} and η, as well as the angular separation between the two jets $\Delta R(jet_1, jet_2)$.

A joint binned maximum-profile-likelihood fit to the $m_{c\bar{c}}$ distribution is performed in four categories, with 15 bins in each category ranging from 50 GeV to 200 GeV. The parameter of interest, μ, is extracted, corresponding to the signal strength, defined as the ratio between the measured signal yield and the SM prediction.

Systematic uncertainties are included in the fit as nuisance parameters, comprising detector-related systematics, signal and background modelling uncertainties and statistical uncertainties in the simulation samples. These can modify the shape and/or normalisation of the distributions. The normalisation of the Z+ jets background is allowed to float freely in each fit category and is obtained from the data.

The fitted $ZH(H \rightarrow c\bar{c})$ signal strength is $\mu_{ZHcc} = -69 \pm 101$. An upper limit

of 110 (150^{+80}_{-40}) is observed (expected) at 95% CL using a modified frequentist CL_s method [66]. The dominant uncertainties on μ come from the background modelling, efficiency of the jet flavour tagging algorithms and jet energy scale and resolution, with the Z+jets normalisation being the largest uncertainty.

The CMS collaboration published two searches for Higgs boson decays to charm quarks, the first in 2020 [67] and the second in 2022 [68], using 35.9 fb^{-1} and 138 fb^{-1} of pp collisions collected at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, respectively. Both analyses target the production of the Higgs boson in association with a W or Z boson, which decay leptonically, resulting in three analysis channels: 0-lepton ($Z(\rightarrow \nu\nu)H$), 1-lepton ($W(\rightarrow \ell\nu)H$) and 2-lepton ($Z(\rightarrow \ell^+\ell^-)H$). A brief description of the most recent CMS analysis is presented below.

The phase-space of the analysis is divided in two regions, defined by a threshold in the transverse momentum of the vector boson, p_{T}^{V}, defining a "resolved-jet" and a "merged-jet" category. The former looks for Higgs boson candidates reconstructed from two separate and well distinguished jets, while the latter targets the case where, due to the high p_{T} of the Higgs boson, the two jets become collimated in a single large jet.

Events in the resolved-jet regime are categorised according to the flavour tagging of the jets, using a neural network (NN) to discriminate the jet flavours, resulting in two c-tagging discriminators, one against bottom quarks and the other against light flavour ones. The associated efficiencies are of 43%, 15% and 4% for c, b and light flavour jets, respectively.

A boosted decision tree (BDT) is trained to separate the signal from the background in each of the lepton channels, using event kinematic observables, Higgs and vector boson properties and jet flavour tagging information. The BDT discriminator is used for the final signal extraction in the signal regions. The normalisations of the V+jets and $t\bar{t}$ backgrounds are obtained from fits to data in dedicated control regions.

In the merged-jet regime a BDT is used in each lepton channel to distinguish between VH production and the backgrounds, taking as input event kinematic information. Large jets are then passed to a NN algorithm for identification of the large jet sub-structure, discriminating between Higgs and W/Z boson production, as well as between the flavour of the boson decay particles, separating $H \to c\bar{c}$ decays from others. The value of the NN discriminant then defines the signal and control regions. The Z+jets, W+jets and $t\bar{t}$ normalisations are allowed to float freely in the fit.

Systematic effects can affect the shapes of the discriminants, as well as the mass distributions and the background normalisations. Other theoretical uncertainties, detector systematics and signal and background modelling uncertainties are also taken into account in the fit.

The combination of the two analysis topologies allows for an observed (expected) upper limit on the $VH(H \to c\bar{c})$ cross section times branching ratio of 14 ($7.6^{+3.4}_{-2.3}$) to be set at 95% CL. The main expected sensitivity is associated to the merged-jet topology. The main uncertainties on μ come from the theoretical modelling,

the efficiencies of the jet flavour tagging algorithms, the limited size of the dataset and the jet energy scale and resolution. The analysis sets additionally an observed (expected) constraint of $1.1 < |\kappa_c| < 5.5$ ($|\kappa_c| < 3.4$) at 95% CL.

The full Run 2 ATLAS search for the decay of the Higgs boson to charm quarks pairs makes use of an integrated luminosity of 139 fb^{-1} of *pp* collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, collected between 2015 and 2018 with the ATLAS detector [57]. The analysis is split in three channels, corresponding to different leptonic decays of the vector bosons produced in association with the Higgs boson. The 0-lepton channel looks for $Z(\rightarrow \nu\nu)H$ production, the 1-lepton channel targets $W(\rightarrow \ell\nu)H$ production and the 2-lepton channel searches for $Z(\rightarrow \ell^+\ell^-)H$ production, with the charged leptons ℓ = e, μ being considered.

5.2 Data and Simulation Samples

This analysis uses a data sample corresponding to an integrated luminosity of 139 fb^{-1} of $\sqrt{s} = 13$ TeV *pp* collision collisions [69], collected between 2015 and 2018. The events considered in the analysis were recorded under stable beam conditions and with all relevant detector sub-systems operating normally [70].

A combination of missing transverse momentum triggers [71] was used to collect the data in the 0-lepton channel and the 1-lepton muon sub-channel, while events in the 1-lepton electron sub-channel were selected using single electron triggers [72]. The missing transverse momentum triggers provide a better ef-

ficiency than single muon triggers in the 1-lepton muon sub-channel, due to a better geometrical acceptance, leading to its use. Both the aforementioned single electron triggers and single muon triggers [73] are used for the 2-lepton channel. The trigger thresholds can be found in Table 5.1. Please note that the full trigger efficiency is not present at exactly the trigger threshold, but at larger values. Different thresholds are considered for triggering on the same physics object for the following reason: lower thresholds require tighter requirements on the object being triggered in order to keep a sustainable trigger rate; considering in addition triggers with relaxed thresholds allows for looser selections being applied on the objects, and to increase the statistics.

Table 5.1: Thresholds for each data taking period of the triggers used in the analysis.

Triggered object	Data Period	Threshold
$E_{\mathrm{T}}^{\mathrm{miss}}$	2015	70 GeV
	2016	90/110 GeV
	2017-2018	110 GeV
Single electron	2015	24/60/120 GeV
	2016-2018	26/60/140 GeV
	2018	300 GeV
Single muon	2015	20 GeV
	2015-2018	60 GeV
	2016-2018	26 GeV

Monte Carlo (MC) simulated samples were produced for signal and background processes using the GEANT4 [74] package for the full ATLAS detector simulation [75]. These simulated events were reconstructed using the same algorithms used to process the data. Samples were simulated with a matrix element

(ME) generator and interfaced with an appropriate parton distribution function (PDF) for the modelling of the hard scattering process. Events were then interfaced to PYTHIA 8.230 [76], SHERPA 2.2.1 [77, 78, 79] or SHERPA 2.2 to model the parton shower (PS), underlying event (UE) and multiple-parton interactions (MPI), while applying the appropriate set of tuning parameters. All samples were normalised to their theoretical cross-sections, with any higher order corrections applied (the cross-section orders mentioned hereafter refer to the order of the cross-section calculation used for process normalisation in QCD, unless otherwise stated). The details of the event generators used for each signal and background sample are shown in Table 5.2.

Table 5.2: Signal and background processes and their corresponding MC generators used in the analysis. NNLO and ((N)NLL stand for next-to-next-to-leading order and ((next-to-)next-to-)leading log, respectively [57].

Process	ME generator	ME PDF	PS and hadronisation	Tune	Cross-section order
$qq \to VH$ $(H \to c\bar{c}/b\bar{b})$	POWHEG BOX v2 [80, 81] + GOSAM [89] + MINLO [90, 91]	NNPDF3.0nlo [82]	PYTHIA 8.212 [76]	AZNLO [83]	NNLO(QCD) +NLO(EW) [84, 85, 86, 87, 88]
$gg \to ZH$ $(H \to c\bar{c}/b\bar{b})$	POWHEG BOX v2	NNPDF3.0nlo	PYTHIA 8.212	AZNLO	NLO+NLL [92, 93]
$t\bar{t}$	POWHEG BOX v2 [94]	NNPDF3.0nlo	PYTHIA 8.230	A14 [95]	NNLO +NNLL [96, 97, 98, 99, 100, 101, 102]
t/s-channel single top	POWHEG BOX v2 [103]	NNPDF3.0nlo	PYTHIA 8.230	A14	NLO [104, 105]
Wt-channel single top	POWHEG BOX v2 [106]	NNPDF3.0nlo	PYTHIA 8.230	A14	Approx. NNLO [107, 108]
V+jets	SHERPA 2.2.1 [77, 78, 79]	NNPDF3.0nlo	SHERPA 2.2.1	Default	NNLO [109]
$qq \to VV$	SHERPA 2.2.1	NNPDF3.0nlo	SHERPA 2.2.1	Default	NLO
$gg \to VV$	SHERPA 2.2.2	NNPDF3.0nlo	SHERPA 2.2.2	Default	NLO

In particular for V+jets (diboson) processes, events with up to two (one) partons are generated at next-to-leading order (NLO), while events with three or four extra (two or more) partons are generated at leading order (LO). The V+jets samples are split according to the transverse momentum of the vector boson (p_{T}^{V}) and total hadronic momentum of the event (H_T) and are generated in slices in max(H_T

, p_V^T) to improve the statistics in each region (higher generation efficiency per slice). These samples are generated (except for the highest slices) with various filters and vetoes placed on the jet flavour composition, in order to improve the statistical precision on the W/Z boson production.

Alternative $VH(H \to c\bar{c})$, $VH(H \to b\bar{b})$, V+jets, diboson and top samples are generated in order to perform comparisons against the prediction from the nominal MC samples mentioned above, from which systematic modelling uncertainties are derived. In these samples different hard-scattering or parton shower descriptions may be used, or variations of the renormalisation (μ_R) and factorisation (μ_F) scales may be taken.

Alternative samples for the $VH(H \to c\bar{c})$ and $VH(H \to b\bar{b})$ processes are generated with a POWHEG BOX v2 + HERWIG 7 prescription, and by varying the μ_R and μ_F scales in the nominal generator.

Diboson alternative samples are obtained by generating events with POWHEG BOX v2 interfaced to PYTHIA 8.230 and again by varying the μ_R and μ_F scale in the nominal generator.

The alternative V+jets samples are simulated with the MADGRAPH 5_AMC@NLO generator at LO, with events with up to four additional partons, interfaced to PYTHIA 8.230. Additionally samples are generated by varying the μ_R and μ_F scale in the nominal generator.

For the top processes three different alternative generator configurations are sim-

ulated, using POWHEG BOX v2 + HERWIG 7, MADGRAPH 5_aMC@NLO + PYTHIA 8.230 and the nominal generator accounting for radiation in the initial or final states.

The QCD multi-jet background is estimated with a data-driven approach, deriving templates from an enriched control region which are then applied in the 1-lepton channel. This background is found to be negligible in the 0- and 2-lepton channels and hence no dedicated background component is considered. The procedure for 1-lepton consists first of selecting events that fail the lepton track and calorimeter-based isolation criteria, meant to separate 'fake' leptons (from non-leptonic sources as photon conversions or misidentified jet signatures) from 'real' ones. Templates are then derived from fits to data in transverse W mass[1], m_{T}^{W}, distributions, given the good discrimination between QCD and electroweak events provided by this variable. The electroweak MC background is subtracted from the data in these distributions, and the remaining difference defines the multi-jet background shape and normalisation.

5.3 Object Selection

Electrons in all lepton channels must satisfy the electron *loose* likelihood identification and *loose* track isolation working points, and have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.47$ (to avoid the end of the inner detector coverage and consequently pos-

[1] $m_{\mathrm{T}}^{W} = \sqrt{2p_{\mathrm{T}}^{\ell}p_{\mathrm{T}}^{\nu}(1-\cos(\phi^{\ell}-\phi^{\nu}))}$, where $E_{\mathrm{T}}^{\mathrm{miss}}$ is used as an approximation for p_{T}^{ν}.

sible lost information). The 1-lepton channel additionally requires electrons with the *tight* WPs passed, and having $p_{\mathrm{T}} > 27$ GeV, in order to reduce the multi-jet background contribution. Electrons in the 2-lepton channel satisfying the *loose* WPs must also have a $p_{\mathrm{T}} > 27$ GeV. The electron-primary vertex matching and pile-up reduction criteria are the same for all channels.

The muon *loose* quality and track isolation working points are required for muons in all lepton channels, and these muons must have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.7$. The 1-lepton channel requires muons also satisfying the *medium* WPs, following the same reasoning as for the electrons, and muons must have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. Muons in the 2-lepton channel satisfying the *loose* WPs are required to have $p_{\mathrm{T}} > 27$ GeV. The additional criteria concerning the matching of the muons to the primary vertex are common to all lepton channels.

Hadronically decaying τ-leptons reconstructed with the anti-k_t jet algorithm with a radius parameter value of $R = 0.4$ are required to satisfy the *medium* quality working point and must have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$, excluding the transition region of $1.37 < |\eta| < 1.52$ between the barrel and end-cap sections of the electromagnetic calorimeter.

Jets are reconstructed using the anti-k_t jet algorithm, with a radius parameter of $R = 0.4$. *Central* and *forward* jets are used in the analysis. These jets are calibrated and required to match with the primary vertex as described in Section 4.4. Jet energy and direction corrections are applied when suitable, as well as the overlap removal procedure.

5.4 Event Selection

Events are categorised according to the number of *loose* charged leptons in three channels: 0-, 1- and 2-lepton. Events in the 0-lepton channel must contain zero *loose* charged leptons. The 1-lepton channel selects events with one *loose* charged lepton that also satisfies the tighter criteria described in Section 5.3. The 2-lepton channel contains events with two *loose* charged leptons, that must also satisfy the tighter p_{T} requirement mentioned in the previous section. The categorisation of the analysis in different lepton channels allows for optimised selections in each channel, maximising the analysis sensitivity to the signal.

A further signal sensitivity enhancement results from the categorisation of the events in the lepton channels according to the total number of jets that they contain. The 0- and 1-lepton channels use 2-jet and 3-jet events in separate categories, while the 2-lepton channel has one category for events with two jets and another for events with three or more jets, accommodating possible initial or final state radiation from the $H \to c\bar{c}$ decay. This higher multiplicity of jets is not possible to be considered in the analysis in the 0- and 1-lepton channels due to a high contamination of those events with $t\bar{t}$ backgrounds, contrary to what happens in the 2-lepton channel.

The phase-space of the analysis is also split in terms of the transverse momentum of the vector boson ($V = Z, W$) produced in association with the Higgs boson, p_{T}^{V}, as the signal-to-background ratio is enhanced for larger transverse momentum.

All lepton channels include a $p_{\mathrm{T}}^{V} > 150$ GeV regime, with the 2-lepton channel having also an additional $75 < p_{\mathrm{T}}^{V} < 150$ GeV category. The lower p_{T}^{V} regime is inaccessible in the 0-lepton channel and 1-lepton muon sub-channel, as the lower boundary of the p_{T}^{V} window is located closely to the trigger threshold and therefore full trigger efficiency is not possible. The 1-lepton electron sub-channel also does not make use of the lower p_{T}^{V} regime, in order to avoid large multi-jet contributions. The p_{T}^{V} quantity is reconstructed as the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, in the 0-lepton channel. In the 1-lepton channel the p_{T}^{V} corresponds to the magnitude of the vector sum of the E_{T}^{miss} and the charged lepton p_{T}. In the 2-lepton channel the p_{T}^{V} is the magnitude of the vector sum of the two charged lepton transverse momenta.

The Higgs candidate mass is reconstructed from the two leading *central* jets in p_{T}, with at least one having $p_{\mathrm{T}} > 45$ GeV. The two Higgs candidate jets are referred to as *signal* jets. Events with one of the *signal* jets c-tagged define the 1 c-tag categories, while events with both *signal* jets c-tagged constitute the 2 c-tag categories. Additional jets to the two forming the Higgs boson candidate have to satisfy a b-tag veto. The $VH(H \to c\bar{c})$ signal regions are therefore orthogonal to those in the $VH(H \to b\bar{b})$ analysis [61], that use two b-tagged jets, allowing for a combination of the two analyses.

This selection defines the common event categorisation for the signal sensitive regions, hereinafter designated as "signal regions", and is summarised in Table 5.3.

Table 5.3: Summary of event categorisation of signal sensitive regions.

Channel	Tag Categories	Number of Jets	p_{T}^{V}
0-lepton	1 and 2 c-tags	2 and 3 jets	$p_{\mathrm{T}}^{V} > 150$ GeV
1-lepton			
2-lepton		2 and 3+ jets	$75 < p_{\mathrm{T}}^{V} < 150$ GeV
			$p_{\mathrm{T}}^{V} > 150$ GeV

Requirements on the angular distance between the two *signal* jets, ΔR, are applied on the signal regions in order to optimise the analysis sensitivity to the $VH(H \to c\bar{c})$ signal. These exploit the fact that jets originating from Higgs bosons are expected to have lower ΔR separation with respect to decay products coming mainly from $t\bar{t}$ and V+jets processes, in which the leading p_{T} jets may come from different initial particles. In addition, the high-p_{T} of the Higgs boson candidates results in more collimated jets. Scans on the statistical significance of the signal (S) with respect to these backgrounds (B), $S/\sqrt{B}$, for different ΔR values resulted in the following thresholds for each p_{T}^{V} range:

- $75 < p_{\mathrm{T}}^{V} < 150$ GeV: $\Delta R < 2.3$,
- $150 < p_{\mathrm{T}}^{V} < 250$ GeV: $\Delta R < 1.6$,
- $p_{\mathrm{T}}^{V} > 250$ GeV: $\Delta R < 1.2$.

A summary of the signal region event selection in lepton channels is presented in Table 5.4. This event selection results in an efficiency to the $VH(H \to c\bar{c})$ signal of around 1.0-1.7%, with the 0-lepton channel being the most sensitive to the signal,

followed by the 2-lepton and 1-lepton channels, respectively.

5.4.1 0-lepton channel Signal Region Event Selection

Events are required to have $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV, at which point the selected triggers are approximately 75%–90% efficient, depending on the year. The full efficiency plateau is reached at about 200 GeV.

The scalar sum of jet transverse momenta, H_{T} is required to be larger than 120 (150) GeV in 2-jet (3-jet) events in order to remove a region with a dependence of the trigger efficiency on the number of jets.

The missing transverse momentum, $p_{\mathrm{T}}^{\mathrm{miss}}$, is required to be larger than 30 GeV to remove the contribution from non-collision backgrounds.

Angular separation requirements, described in Table 5.4, are implemented to reject QCD multi-jet background events, that otherwise could enter the analysis due to mismeasured jet energies in the calorimeter. The surviving multi-jet contamination after implementing these cuts is rendered a negligible contribution to the total background in the 0-lepton channel.

5.4.2 1-lepton channel Signal Region Event Selection

Events in the electron sub-channel are required to have $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV to remove the existent contribution from backgrounds with jets faking electron signatures.

The transverse mass of the reconstructed W boson, m_{T}^{W}, is required to be less than 120 GeV in order to reduce the $t\bar{t}$ background contribution to the total background.

5.4.3 2-lepton channel Signal Region Event Selection

Events in the 2-lepton channel are required to have leptons with the same flavour, and opposite charges in the dimuon case. This latter requirement is not present in the dielectron sub-channel due to a higher probability of charge misidentification.

The invariant mass of the dilepton system, $m_{\ell\ell}$, must be consistent with the mass of the Z boson, $81 < m_{\ell\ell} < 101$ GeV.

Table 5.4: Summary of the signal region event selection in the 0-, 1- and 2-lepton channels. Jet1 and jet2 refer to the two *signal* jets and H refers to the jet1–jet2 system [57].

Common selections	
Central jets	≥ 2
Signal jet p_{T}	≥ 1 signal jet with $p_{\mathrm{T}} > 45$ GeV
c-jets	One or two c-tagged *signal* jets
b-jets	No b-tagged non-*signal* jets
Jets	$2, 3$ (0- and 1-lepton); $2, \geq 3$ (2-lepton)
p_{T}^{V} regions	$75 - 150$ GeV (2-lepton) > 150 GeV
ΔR(**jet1**, **jet2**)	$75 < p_{\mathrm{T}}^{V} < 150$ GeV: $\Delta R \leq 2.3$ $150 < p_{\mathrm{T}}^{V} < 250$ GeV: $\Delta R \leq 1.6$ $p_{\mathrm{T}}^{V} > 250$ GeV: $\Delta R \leq 1.2$
0-lepton channel	
Trigger	E_T^{miss}
Leptons	No *loose* leptons
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 150 GeV
$p_{\mathrm{T}}^{\mathrm{miss}}$	> 30 GeV
H_{T}	> 120 GeV (2 jets), > 150 GeV (3 jets)
$\min\lvert\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{jet})\rvert$	$> 20°$ (2 jets), $> 30°$ (3 jets)
$\lvert\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{H})\rvert$	$> 120°$
$\lvert\Delta\phi($**jet1, jet2**$)\rvert$	$< 140°$
$\lvert\Delta\phi(\mathbf{E_T^{miss}}, \mathbf{p_T^{miss}})\rvert$	$< 90°$
1-lepton channel	
Trigger	e sub-channel: single electron μ sub-channel: $E_{\mathrm{T}}^{\mathrm{miss}}$
Leptons	One *tight* lepton and no additional *loose* leptons
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 30 GeV (e sub-channel)
m_{T}^{W}	< 120 GeV
2-lepton channel	
Trigger	Single lepton
Leptons	Exactly two *loose* leptons Same flavour, opposite charge for $\mu\mu$
$m_{\ell\ell}$	$81 < m_{\ell\ell} < 101$ GeV

5.4.4 Control Regions

Three sets of control regions (CRs) are defined to constrain the modelling of the main backgrounds, namely the $t\bar{t}$ and V+jets processes. These use events not considered in the signal regions and are enriched in the targeted background process. The CRs are split in jet and/or p_{T}^{V} categories to maximise the statistical power.

Control regions containing events failing the ΔR requirements in the signal regions are used to constrain the modelling of the V+jets background shapes and normalisations, particularly at higher masses (exploiting the correlation of ΔR with $m_{c\bar{c}}$). Only events with $\Delta R < 2.5$ are considered, due to limited statistics for events with larger angular distances between the two *signal* jets. These regions are referred to as "high-ΔR" control regions and are implemented for each corresponding signal region.

Top control regions are defined in all lepton channels to constrain the modelling of $t\bar{t}$ and single-top processes. In the 0- and 1-lepton channels, events with one c-tag in the two *signal* jets and a third additional b-tagged jet are selected, while events with an $e\mu$ pair with opposite charges are used in the 2-lepton channel top control region. The 0- and 1-lepton channel have each just one top CR category (1 c-tag, 3 jets), while the 2-lepton top $e\mu$ CR is split in terms of jet multiplicity and p_{T}^{V} regime.

Events with no c-tagged Higgs candidate jets and no additional b-tagged jets define the "0 c-tag" control regions in the 1- and 2-lepton channels, and are used

to constrain the normalisation of the V+jets light flavour component. These CRs are split in terms of jet multiplicity and p_{T}^{V} regime (only applicable in the 2-lepton channel).

5.5 Background Composition

The Z+jets backgrounds are split in three flavour dependent components: $Z+hf$ (heavy flavour jets - $Z+\{cc,bb\}$), $Z+mf$ (mixed flavour jets - $Z+\{cl,bc,bl\}$) and $Z+lf$ (light flavour jets - $Z+\{ll\}$). The W+jets backgrounds are grouped accordingly, but with the inclusion of W+jets processes in which the W boson decays to an hadronically-decaying τ-lepton (τ_h) and neutrino, $W(\rightarrow \tau_h\nu)+$light/$c$/$b$-jet, as this is a non-negligible component in the 0-lepton channel. The $W(\rightarrow \tau_h\nu)+c/b$-jet processes enter the mixed flavour component and the $W(\rightarrow \tau_h\nu)+$ light-jet processes the light flavour one. Top processes are split in two components in the 0- and 1-lepton channel, one denoted as Top ($other$) containing jets from a W boson decay (also referred to as resonant production) in $t\bar{t}$ and single-top processes, and a Top (b) component for jet pairs consisting of a b-jet from a top quark decay and other jet in the event (also referred to as non-resonant production). The $t\bar{t}$ process is treated inclusively in the 2-lepton channel, being separate from the two components in the 0- and 1-lepton channels. The diboson background consists of all the ZZ, WW and WZ processes apart from the $VZ(Z\rightarrow c\bar{c})$ and $VW(W\rightarrow c\bar{l})$ which are considered separately.

The Z+jets, W+jets and $t\bar{t}$ processes constitute the main contributions to the total background in the 0-lepton channel signal regions, with the tagging scheme leading to 1 c-tag categories having a larger proportion of W/Z processes with light flavour jets associated, and 2 c-tags regions with larger contributions from vector bosons produced in association with heavy flavour jets. The higher multiplicity of jets in some of the signal regions results also in a higher $t\bar{t}$ contamination. The diboson VZ and VW processes constitute sub-dominant contributions in this lepton channel. The background compositions are similar in the ΔR CRs, but containing events with higher $m_{c\bar{c}}$ values. The simulated background composition of all the signal and control regions in the 0-lepton channels after the respective event selections is presented in Figure 5.1.

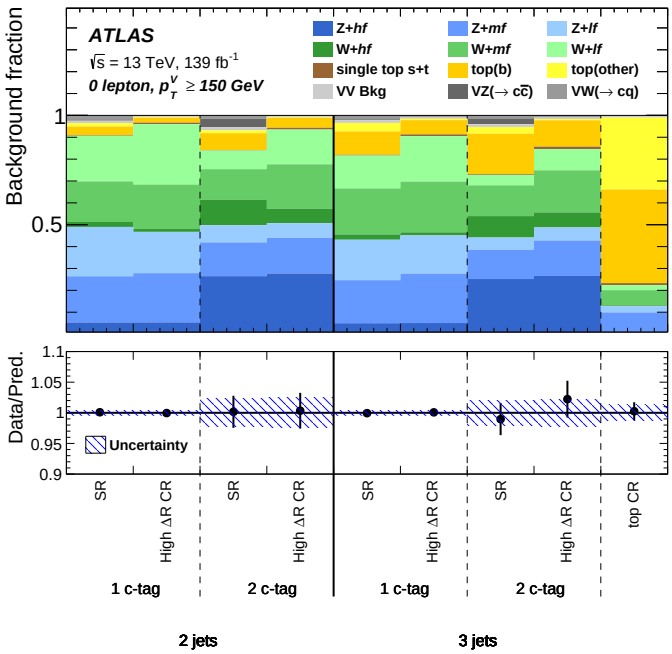


Figure 5.1: Background composition in simulation of the 0-lepton channel signal and control regions [57].

The main backgrounds of the 1-lepton channel consist of the W+jets and top pro-

cesses, as shown in Figure 5.2. The tagging categories have the same light/heavy flavour pattern described for the 0-lepton channel. Again regions with larger jet multiplicity have larger contributions from top processes. Additionally in the 1-lepton channel there are smaller multi-jet contributions, along with VZ and VW processes.

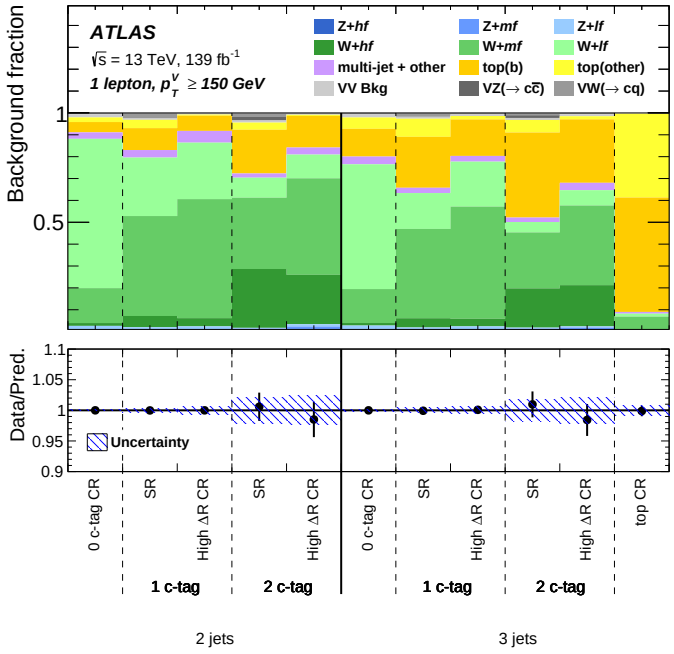
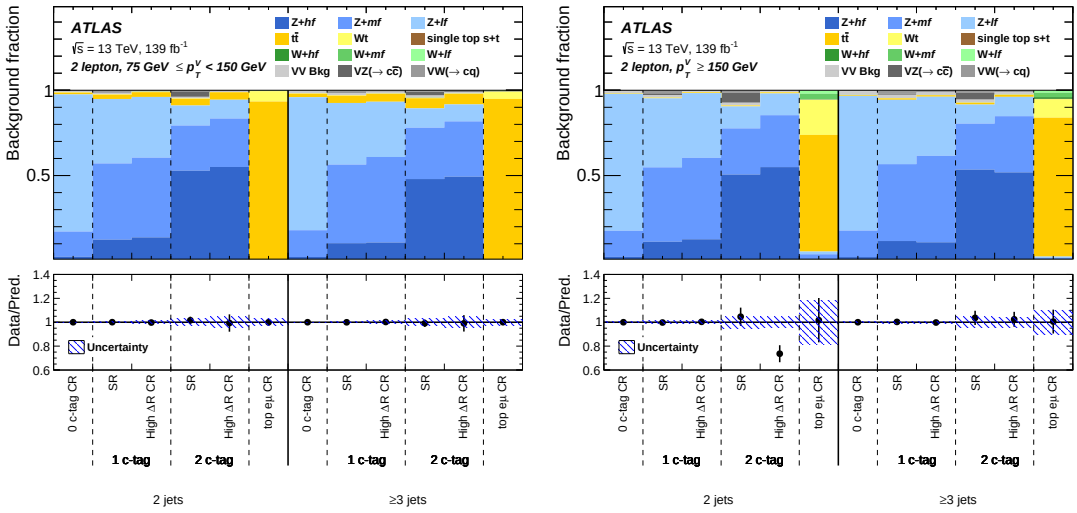


Figure 5.2: Background composition in simulation of the 1-lepton channel signal and control regions [57].

The 2-lepton channel signal regions are completely dominated by the Z+jets processes, with sub-dominant contributions from VZ and VW diboson and top quark processes (the latter in particular in the low p_{T}^{V} regime). The simulated background composition of the 2-lepton channel signal and control regions is shown in Figure 5.3.



(a) 75< p_{T}^{V} <150 GeV (b) 2-lepton channel, $p_{\mathrm{T}}^{V} > 150$ GeV

Figure 5.3: Background composition in simulation of the 2-lepton channel signal and control regions [57].

5.6 Systematic Uncertainties

The sources of systematic uncertainties affecting the $VH(H \to c\bar{c})$ analysis can be divided in two main groups, one related to the detector performance and reconstruction of the physics objects, constituting the experimental uncertainties, and another group related to the theoretical modelling of the signal and background processes.

These systematic uncertainties can affect the normalisation and shape of the signal and background processes. A normalisation effect will translate into an uniform increase or decrease of the number of events in the $m_{c\bar{c}}$ distribution, while a shape effect will change the distribution of the events in the bins of the mass distribution, but keeping the same number of events. Some systematic uncertainties can have both normalisation and shape effects.

5.6.1 Experimental Uncertainties

A summary of the different experimental uncertainties for the physics objects is presented in Table 5.5.

Table 5.5: Summary of the experimental systematic uncertainties applied in the $VH(H \to c\bar{c})$ analysis.

Physics object	Systematic uncertainty
Luminosity	Uncertainty on total integrated luminosity
Pile-up	Uncertainty on average number of interactions per bunch-crossing
Electrons/muons	Trigger efficiency uncertainty
	Uncertainties on reconstruction, identification and isolation efficiencies
	Energy scale and resolution uncertainty
Muons	Track-to-vertex association efficiency uncertainty
Jets	Energy scale and resolution uncertainty
	Uncertainties on jet flavour response and composition
$E_{\mathrm{T}}^{\mathrm{miss}}$	Trigger efficiency uncertainty
	Energy scale and resolution uncertainty
Flavour tagging	c-jet tagging efficiency uncertainty
	b-jet tagging efficiency uncertainty
	Light jet tagging efficiency uncertainty
	τ-lepton jet tagging efficiency uncertainty

The integrated luminosity for the 2015–2018 period has an associated uncertainty of 1.7% [69], obtained via a dedicated measurement by the LUCID (LUminosity measurement using Cerenkov Integrating Detector) detector [110]. The pile-up distribution in the simulated samples is corrected to match the conditions found in data, with a scale factor of 3%.

Electrons and muons have uncertainties on their energy scale and resolution, as well as systematics associated with the respective trigger and reconstruction efficiencies, obtained from comparisons between data and simulation in events with

final states containing leptons.

Jets have associated uncertainties related to their energy scale (JES) and resolution (JER). The former are derived from comparisons between data and simulation with variation of parameters as the event selection or the Monte Carlo generator. Differences in the jet response and simulated flavour composition for jets initiated by (light and heavy flavour) quarks and gluons are also accounted for in additional jets systematic uncertainties. Pile-up related uncertainties are also taken into account for jets. JER uncertainties are derived via estimates of the jet energy resolution from calorimeter observables and comparisons of simulation to data [111].

Missing transverse energy scale and resolution systematics are obtained from comparisons between data and simulation for the same final state selection, with the uncertainties coming only from the track-based soft term, as uncertainties on the other objects used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ computation are already included in their respective systematics. There are additionally missing transverse energy uncertainties related to the trigger efficiency.

Flavour tagging uncertainties for c-, b-, τ-lepton and light jets are considered, and obtained by comparing the tagging efficiencies in simulation to data.

As mentioned in Section 4.7 a truth flavour tagging (TT) procedure is employed in the analysis, applying the jet tagging efficiencies as event weights instead of using the direct tagging (DT) approach of cutting on the flavour tagging algorithm

discriminant. The TT implementation has two sets of uncertainties associated to correct for differences between the TT and DT $m_{c\bar{c}}$ distributions.

The first one is a ΔR correction applied for V+jets processes, derived for each jet and dependent on its flavour. The minimum ΔR between a c-jet and any other jet in the event is taken, with the correction being derived for the cases where the c-jet corresponds to a generator-level b-, c- or light flavour jet. The correction then results from fits to the ratio of the DT and TT ΔR distributions. The associated uncertainty corresponds to the full size of the correction.

With this correction the truth flavour event weights from Equations 4.3 and 4.4 take the form:

$$\text{TT weight (1 } c\text{-jet events)} = c_1\epsilon_1 \times (1 - c_2\epsilon_2) + (1 - c_1\epsilon_1) \times c_2\epsilon_2 \,, \tag{5.1}$$

$$\text{TT weight (2 } c\text{-jets events)} = c_1\epsilon_1 \times c_2\epsilon_2 \,, \tag{5.2}$$

with c_1 and c_2 corresponding to the ΔR per-jet corrections.

This correction arises from the fact that the tagging efficiency weights are derived from a simulated sample of $t\bar{t}$ events, and residual differences may exist for different processes, and in addition the loss of performance in the jet flavour tagging algorithms for two nearby jets. For this latter case, tracks, vertices and energy deposits can be shared between the two jets, degrading the jet reconstruc-

tion and identification, with this aspect impacting the DT performance but not being accounted for in the TT per-jet efficiencies.

Small remaining DT/TT non-closure in the $m_{c\bar{c}}$ distributions introduces a second TT uncertainty, correcting the TT yields to match the DT ones via an additional normalisation-only systematic uncertainty.

5.6.2 Signal and Background Modelling

The $VH(H \rightarrow c\bar{c})$ analysis makes use of simulated signal and background samples (except for the multi-jet background) as described in Section 5.2, with the respective Monte Carlo generators being referred to as the "nominal" ones. If using alternative generators different signal and background predictions may be obtained, due to for instance different hadronisation descriptions or orders of perturbation theory corrections considered. Variations of the factorisation and renormalisation scales in the nominal generation will also result in different predictions.

Modelling systematic uncertainties are derived to account for these differences by comparing the alternative generators and scale variations to the nominal description of the signal and background processes, from which the following uncertainties can be obtained:

- Normalisation uncertainties
- Acceptance ratios

- Flavour composition ratios
- Channel extrapolations
- Signal region/control region extrapolations
- $m_{c\bar{c}}$ shape uncertainties

Normalisation uncertainties reflect the relative difference on total yield predictions, taking into account the cross sections and branching ratios of the respective processes. These are considered in the analysis for sub-dominant diboson and VH processes.

Acceptance ratios reflect the relative differences of the various simulation setups in the predictions of categories in the fit with different jet multiplicity or vector boson transverse momentum regimes considered. These uncertainties are defined as double ratios between two regions (denoted here simply by 1 and 2), as given by

$$\text{Acceptance ratio} = \sqrt{\sum_i \left(\frac{(\frac{n_1}{n_2})_i}{(\frac{n_1}{n_2})_{nominal}} - 1 \right)^2}, \tag{5.3}$$

with region 2 having the highest precision for a specific process. $(\frac{n_1}{n_2})_{nominal}$ corresponds to the yield ratio using the nominal MC prediction, and $(\frac{n_1}{n_2})_i$ the yield ratio using one the following uncertainty sources i considered: an alternative MC generator, or the nominal generator with variation of the μ_R or μ_F scale. The three contributions are added in quadrature for the final ratio.

Flavour composition ratios account for different predictions of the proportion

of each background flavour/process in the categories. These are computed in the same way as the acceptance ratios mentioned above, and are derived with respect to each dominant background component ($e.g.$ flavour composition ratios are derived for $Z + bl$ and $Z + bc$ with respect to the dominant $Z + cl$ component in the $Z + mf$ backgrounds).

The $m_{c\bar{c}}$ shape uncertainties account for differences in the binned $m_{c\bar{c}}$ predictions, and are obtained from ratios of the normalised distributions for nominal and alternative generators or scale variations in the signal or control regions. The systematic uncertainty is parametrised as a fit to the ratio of the alternative over nominal normalised distributions. Shapes of all processes were analysed in each separate c-tag, p_{T}^{V} and jet multiplicity categories. Different shapes were combined in a region or between regions if in agreement within their statistical uncertainty, with the merged nominal shape being compared to the merged alternative one for the derivation of the final shape uncertainty.

VH processes

Theoretical uncertainties in the WH and ZH production cross-section are considered in the fit, as well as uncertainties on the $H \to c\bar{c}$ branching ratio [112, 113]. Normalisation, acceptance and shape uncertainties are derived for the $qq \to ZH(H \to c\bar{c})$, $qq \to WH(H \to c\bar{c})$ and $gg \to ZH(H \to c\bar{c})$ signal processes separately. A summary of the signal systematic uncertainties can be seen in Table 5.6.

Table 5.6: Systematic uncertainties on the $VH(H \to c\bar{c})$ signal processes.

Source	Uncertainty
σ_{WH}	[-2.05%, +1.98%]
σ_{ZH}	[-3.48%, +4.11%]
$\text{BR}_{H\to c\bar{c}}$	[-1.99%, +5.53%]
$qq \to ZH(H \to c\bar{c})$ normalisation	6%
$qq \to WH(H \to c\bar{c})$ normalisation	6%
$gg \to ZH(H \to c\bar{c})$ normalisation	31%
N_{jet} acceptance, $qq \to ZH(H \to c\bar{c})$	8%–12%
N_{jet} acceptance, $qq \to WH(H \to c\bar{c})$	6%
N_{jet} acceptance, $gg \to ZH(H \to c\bar{c})$	19%–56%
p_{T}^{V} acceptance, $qq \to ZH(H \to c\bar{c})$ (2-lepton channel only)	2%
p_{T}^{V} acceptance, $gg \to ZH(H \to c\bar{c})$ (2-lepton channel only)	5%

Normalisation uncertainties are included for the $WH(H \to b\bar{b})$ and $ZH(H \to b\bar{b})$ backgrounds, corresponding to the uncertainties on the results of the latest ATLAS measurement [61]. Additional p_{T}^{V} and jet multiplicity acceptance uncertainties are also considered, using the same uncertainties obtained for the $VH(H \to c\bar{c})$ signal. $VH(H \to b\bar{b})$ inclusive shape uncertainties are derived from the $m_{c\bar{c}}$ distributions and included in the fit.

Diboson processes

Normalisation, p_{T}^{V} and jet acceptance uncertainties are considered for the diboson background, $VZ(Z \to c\bar{c})$ and $VW(W \to c\bar{l})$ processes, being derived for separate WW, ZZ and WZ components, but inclusively in their decays. Shape uncertain-

ties are derived separately for the $VZ(Z \to c\bar{c})$ and $VW(W \to c\bar{l})$ processes and the remaining diboson backgrounds.

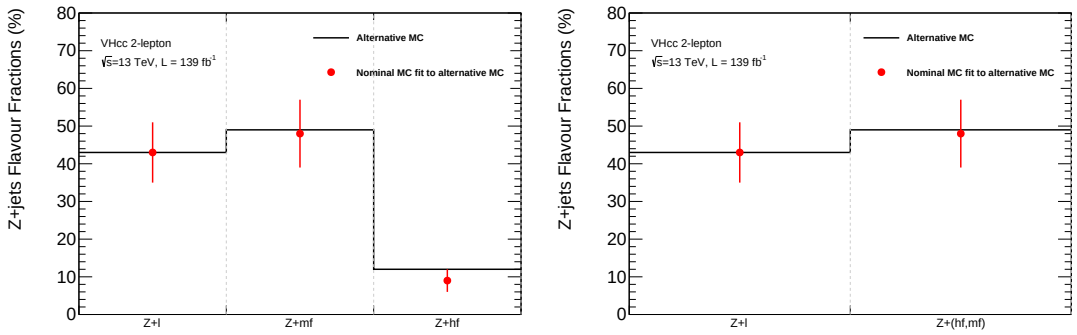
V+jets processes

A study of the Z+jets flavour modelling was performed in the context of the analysis, in order to understand which flavour scheme of the jets produced in association with the Z boson would result in a better control of this background. Two options were tested in the 2-lepton channel, given the highest purity of this background in the channel:

- Three Z+jets categories: $Z + hf$ ($Z + \{cc, bb\}$), $Z + mf$ ($Z + \{cl, bc, bl\}$) and $Z + lf(\{ll\})$
- Two Z+jets categories: $Z + (hf, mf)$ ($Z + \{cc, bb, cl, bc, bl\}$) and $Z + lf(\{ll\})$

with the second setup being considered given the $Z+hf$ and $Z+mf$ background components have similar shapes in the 2-lepton channel when taken inclusively in p_{T}^{V}, c-tag and jet categories.

The modelling evaluation was performed by fitting the nominal MC events to the alternative MADGRAPH 5_AMC@NLO + PYTHIA 8 simulated data (being used as pseudodata) for the two fit setups mentioned above, and evaluating which would better recover the Z+jets flavour fractions of the alternative MC generator. The flavour fractions after the fit to the alternative MC for both fit setups can be seen in Figure 5.4.



(a) Three Z+jets categories (b) Two Z+jets categories

Figure 5.4: Z+jets flavour fractions from fit of nominal to alternative MC.

An equal modelling performance was observed, with both setups returning compatible flavour fractions for the fit of the nominal simulated sample to the alternative one. The fit setup with three Z+jets categories was adopted to facilitate the merging with the other analysis channels.

The normalisations of all the Z+jets and W+jets flavour split components are allowed to float freely in the fit and are obtained from the data. The Z+jets (W+jets) floating normalisations are determined commonly in the 0- and 2-lepton (0- and 1-lepton) channels, with these processes constituting a negligible background in the 1-lepton (2-lepton channel). The $Z+hf$ and $Z+mf$ floating normalisation are decorrelated in the p_{T}^{V} regions, $i.e$, there is one independent floating uncertainty for each component in each of the transverse momentum regimes, being constrained by the signal and ΔR control regions. The inclusion of the 0 c-tag control region allows for the $Z+lf$ background normalisation to be equally obtained from data, floating separately in p_{T}^{V} and jet multiplicity categories ($i.e.$ having four parameters). The $W+lf$ normalisations are floated separately in the different jet multiplicity regions.

Acceptance uncertainties are considered to account for effects in the modelling of different jet multiplicity categories (for processes with floating normalisations correlated between jet regions), and acceptance effects between the lepton channels. Flavour composition ratios are taken for the sub-dominant contribution in each V+jets flavour component with respect to the dominant one, and are derived in the lepton channel offering the most precise estimate (0-lepton for Z+jets and 1-lepton for W+jets). The 0-lepton channel includes an additional flavour acceptance uncertainty between the τ-lepton components and the respective dominant W+jets flavour component. The acceptance effect between the signal and control regions for the τ-lepton components of the W+jets background is also included, as this effect can not be derived from its shape in the ΔR regions (as for the other background components, which are mentioned later) due to low statistics.

Shape uncertainties are considered for each of the W+jets and Z+jets components, accounting for differences in shape of the predictions of the different generators in the signal and control regions, and include a normalisation acceptance effect between the two sets of regions.

Top processes

The normalisations of the three top background components are floated in the fit, with the Top (b) and Top ($other$) components being determined in a correlated way between the the 0- and 1-lepton channel, and a separate $t\bar{t}$ component being considered only in the 2-lepton channel. Each of the components is constrained

by the respective top control regions. The 2-lepton channel top normalisation is decorrelated in p_{T}^{V} and jet multiplicity categories.

Shape and acceptance uncertainties are included in the 0- and 1-lepton channels to account for differences in the shapes of the top processes and the normalisations of different jet multiplicity categories, lepton channels (0-lepton/1-lepton) and signal and control regions. An uncertainty associated to the $t\bar{t}$ and Wt contributions to the top components is also included. The $t\bar{t}$ background has no acceptance uncertainties, given the floating normalisation is decorrelated in p_{T}^{V} and jet categories, and no shape associated in the 2-lepton channel, as it represents a small contribution to the total background.

A summary of the modelling and free-floating systematics uncertainties affecting the background yields in the fit is shown in Table 5.7. The values in the table correspond to the size of the uncertainty affecting the yield of each background. A range is displayed in the cases where the size of an acceptance systematic uncertainty varies between analysis regions. For simplicity the decorrelation schemes of the floating normalisations of the main backgrounds in jet and/or p_{T}^{V} categories are not shown.

Table 5.7: Summary of the background modelling systematic uncertainties considered. CR and SR stand for control region and signal region [57].

$\boldsymbol{VH(\to b\bar{b})}$	
$WH(\to b\bar{b})$ normalisation	27%
$ZH(\to b\bar{b})$ normalisation	25%
Diboson	
$WW/ZZ/WZ$ acceptance	10%/5%/12%
p_{T}^{V} acceptance (2-lepton channel only)	4%
N_{jet} acceptance	7%–11%
Z**+jets**	
$Z+hf$ normalisation	Floating
$Z+mf$ normalisation	Floating
$Z+lf$ normalisation	Floating
$Z+bb$ to $Z+cc$ ratio	20%
$Z+bl$ to $Z+cl$ ratio	18%
$Z+bc$ to $Z+cl$ ratio	6%
N_{jet} acceptance	10%–37%
High-ΔR CR to SR	12%–37%
0- to 2-lepton ratio	4%–5%
W**+jets**	
$W+hf$ normalisation	Floating
$W+mf$ normalisation	Floating
$W+lf$ normalisation	Floating
$W+bb$ to $W+cc$ ratio	4%–10%
$W+bl$ to $W+cl$ ratio	31%–32%
$W+bc$ to $W+cl$ ratio	31%–33%
$W\to\tau\nu(+c)$ to $W+cl$ ratio	11%
$W\to\tau\nu(+b)$ to $W+cl$ ratio	27%
$W\to\tau\nu(+l)$ to $W+l$ ratio	8%
N_{jet} acceptance	8%–14%
High-ΔR CR to SR	15%–29%
$W\to\tau\nu$ SR to high-ΔR CR ratio	5%–18%
0- to 1-lepton ratio	1%– 6%
Top quark (0- and 1-lepton)	
Top(b) normalisation	Floating
Top(other) normalisation	Floating
N_{jet} acceptance	7%–9%
0- to 1-lepton ratio	4%
SR/top CR acceptance ($t\bar{t}$)	9%
SR/top CR acceptance (Wt)	16%
Wt / $t\bar{t}$ ratio	10%
Top quark (2-lepton)	
Normalisation	Floating
Multi-jet (1-lepton)	
Normalisation	20%–100%

5.7 Statistical Analysis

A binned profile likelihood fit was performed to the data, using the $m_{c\bar{c}}$ distribution as the fit discriminant. The fit is performed across the three lepton channels, and measuring simultaneously the signal strengths for the $VH(\rightarrow c\bar{c})$, $VW(\rightarrow cq)$ and $VZ(\rightarrow c\bar{c})$ processes. Each signal strength, also referred to as a *parameter of interest* (POI) hereafter, is defined as the ratio between the measured signal yield and the SM prediction:

$$\mu = \frac{(\sigma \times BR)_{Observed}}{(\sigma \times BR)_{SM}} \,. \tag{5.4}$$

5.7.1 Likelihood Definition

The likelihood function can be written as the product of three terms, taking into account not only the expected signal and background yields, but also the statistical and systematic errors:

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\theta}) = \mathcal{L}_{Poiss}(\boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \cdot \mathcal{L}_{Syst}(\boldsymbol{\alpha}) \cdot \mathcal{L}_{Stat}(\boldsymbol{\gamma}) \,. \tag{5.5}$$

The first term in the likelihood function corresponds to the Poisson probability to estimate the $\boldsymbol{\mu} = \{\mu_{VHcc}, \mu_{VZcc}, \mu_{VWcl}\}$ signal strengths from the data, with N_i being the number of data events in bin i, and s_i (b_i) the expected number of signal (background) events in bin i, with these depending on the nuisance parameters

$\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\tau}\}$:

$$\mathcal{L}_{Poiss}(\boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\tau}) = \prod_{i \,\epsilon\, \text{bins}} \text{Poisson}(N_i | \boldsymbol{\mu} \cdot s_i(\boldsymbol{\alpha}) + b_i(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\tau}))\,. \tag{5.6}$$

The $\boldsymbol{\alpha} = \{\alpha_1, ..., \alpha_{n_{Syst}}\}$ set of nuisance parameters encodes the effect of the systematic uncertainties in the signal and background normalisations and shapes, which are estimated from the data or auxiliary measurements. The constraints on the systematic uncertainties are considered as a Gaussian penalty term in the likelihood, with a mean value of zero and variance of one:

$$\mathcal{L}_{Syst}(\boldsymbol{\alpha}) = \prod_{\alpha \,\epsilon\, \boldsymbol{\alpha}} \frac{1}{\sqrt{2\pi}} e^{-\alpha^2/2}\,. \tag{5.7}$$

The $\boldsymbol{\gamma} = \{\gamma_1, ..., \gamma_{n_{bins}}\}$ NPs are associated to the statistical uncertainties on the background yields, resulting from the limited sizes of the simulated samples. Each bin of the total background prediction in each distribution is taken as an independent nuisance parameter γ_i, with nominal value of 1 and variance of σ_i/b_i. These uncertainties are small for the signal simulated events and are therefore not considered. The $\boldsymbol{\gamma}$ contribution is included in the likelihood as a Gaussian constraint:

$$\mathcal{L}_{Stat}(\gamma) = \prod_{i \,\epsilon\, \text{bins}} \text{Gaussian}(\beta_i | \gamma_i \beta_i, \sqrt{\gamma_i \beta_i})\,, \tag{5.8}$$

where $\beta_i = 1/\sigma_{b_i}^2$ and σ the relative statistical uncertainty on the expected total background prediction.

The $\boldsymbol{\tau}$ nuisance parameters (NPs) correspond to the *free floating* normalisations of chosen backgrounds in the fit, which are obtained directly from the data.

A test statistic is defined to measure the agreement between the background-only hypothesis ($\mu = 0$) and the data, and is defined as:

$$t_{\boldsymbol{\mu}} = -2 \ln \lambda(\boldsymbol{\mu}) \,, \tag{5.9}$$

being constructed from the profile likelihood ratio [114]:

$$\lambda(\boldsymbol{\mu}) = \frac{\mathcal{L}(\boldsymbol{\mu}, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\theta}})} \,, \tag{5.10}$$

where $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\theta}}$ are the parameters that maximise the likelihood, and $\hat{\hat{\boldsymbol{\theta}}}$ are the NPs values that maximise the likelihood for a given signal strength. Values of the profile likelihood ratio vary between zero and one, with values close to one implying a good agreement between the data and the background-only hypothesis.

The incompatibility of the data with the background-only hypothesis can be expressed with the p-value (p_0), which describes the probability of finding an observed test statistic of equal or greater incompatibility with the $\mu = 0$ prediction. By convention the discovery of a signal is claimed if the background-only hy-

pothesis is rejected with a p-value equal or smaller to 2.87×10^{-7}, corresponding to a significance of 5 standard deviations (σ). The significance can be defined as the number of standard deviations necessary for a Gaussian distributed variable to fluctuate in one direction to give a certain p-value. In the impossibility of rejecting the $\mu = 0$ hypothesis, upper limits on the signal strength are set using the CL_s method [66].

5.7.2 Fit Model

The binned profile likelihood fit is performed to the data simultaneously in the three lepton channels, using the $m_{c\bar{c}}$ distribution as the fit discriminant. The fit is performed across 16 signal regions and 28 control regions, as outlined in Table 5.8.

Table 5.8: Summary of the signal and control regions used in the analysis. ΔR_{SR}^{cut} is the upper cut of $\Delta R(\text{jet1}, \text{jet2})$ based on the signal optimisation described in Section 5.4.

			SR		Top CR	High ΔR CR		0 c-tag CR
0-lepton	Description		Table 5.4		3rd jet b-tagged	$\Delta R_{SR}^{cut} < \Delta R(\text{jet1}, \text{jet2}) < 2.5$		0 c-tag *signal* jets
	Categories	$p_{\mathrm{T}}^{V} > 150$ GeV	1 c-tag, 2 jets	2 c-tags, 2 jets		1 c-tag, 2 jets	2 c-tags, 2 jets	-
			1 c-tag,3 jets	2 c-tags, 3 jets	1 c-tag, 3 jets	1 c-tag, 3 jets	2 c-tags, 3 jets	
1-lepton	Description		Table 5.4		3rd jet b-tagged	$\Delta R_{SR}^{cut} < \Delta R(\text{jet1}, \text{jet2}) < 2.5$		0 c-tag *signal* jets
	Categories	$p_{\mathrm{T}}^{V} > 150$ GeV	1 c-tag, 2 jets	2 c-tags, 2 jets		1 c-tag, 2 jets	2 c-tags, 2 jets	0 c-tag, 2 jets
			1 c-tag, 3 jets	2 c-tags, 3 jets	1 c-tag, 3 jets	1 c-tag, 3 jets	2 c-tags, 3 jets	0 c-tag, 3 jets
2-lepton	Description		Table 5.4		Different flavour leptons	$\Delta R_{SR}^{cut} < \Delta R(\text{jet1}, \text{jet2}) < 2.5$		0 c-tag *signal* jets
	Categories	$75 < p_{\mathrm{T}}^{V} < 150$ GeV	1 c-tag, 2 jets	2 c-tags, 2 jets	1 c-tag, 2 jets	1 c-tag, 2 jets	2 c-tags, 2 jets	0 c-tag, 2 jets
			1 c-tag, 3+ jets	2 c-tags, 3+ jets	1 c-tag, 3+ jets	1 c-tag, 3+ jets	2 c-tags, 3+ jets	0 c-tag, 3+ jets
		$p_{\mathrm{T}}^{V} > 150$ GeV	1 c-tag, 2 jets	2 c-tags, 2 jets	1 c-tag, 2 jets	1 c-tag, 2 jets	2 c-tags, 2 jets	0 c-tag, 2 jets
			1 c-tag, 3+ jets	2 c-tags, 3+ jets	1 c-tag, 3+ jets	1 c-tag, 3+ jets	2 c-tags, 3+ jets	0 c-tag, 3+ jets

The $m_{c\bar{c}}$ distributions are fitted in the [50, 210] GeV range in the signal regions (SRs) and 0 c-tag and top control regions (CRs), while in the high ΔR control

region the fit range is instead [100, 350] GeV. In the SRs and most of the CRs, the $m_{c\bar{c}} < 50$ GeV region of the mass distributions is excluded due to a larger mis-modelling of the V+jets backgrounds, where the majority of the associated jets are produced via a gluon decay. The upper cut at 210 GeV is justified by the available data statistics, which become scarce after that point. The high ΔR CRs aim to constrain backgrounds at higher $m_{c\bar{c}}$, hence its adjusted fit range.

The binning of the $m_{c\bar{c}}$ distributions is chosen such that is it is close to the experimental resolution of the Higgs candidate invariant mass distribution (of around 10 GeV), and minimising statistical fluctuations of the data and simulation. All signal regions have a bin width of 10 GeV, apart from the 2-lepton channel high p_{T}^{V} categories where, due to lower statistics, the bin width is increased to 15 GeV. The 0 c-tag CRs and 2-lepton top CRs are designed to extract background normalisations (and no shape information), and therefore only a single bin is used in the respective categories. On the other hand, the top CRs in the 0- and 1-lepton channel use a bin width of 10 GeV. The high ΔR CRs have a 20 GeV binning in order to reduce statistical fluctuations.

The fit model considers three signal processes:

- $VH\,(H \to c\bar{c})$
- $VW\,(W \to cq)$: validation of the 1 c-tag categories
- $VZ\,(Z \to c\bar{c})$: validation of the 2 c-tags categories

The background processes entering the analysis are grouped in the fit into the

following components:

- Three Z+jets components, depending on the quark flavour of the associated jets:
  - $Z + hf$ (heavy flavour): $Z + \{cc, bb\}$
  - $Z + mf$ (mixed flavour): $Z + \{cl, bc, bl\}$
  - $Z + lf$ (light flavour): $Z + \{ll\}$
- Three W+jets components, depending on the quark flavour of the associated jets:
  - $W + hf$ (heavy flavour): $W + \{cc, bb\}$
  - $W + mf$ (mixed flavour): $W + \{cl, bc, bl, c\tau, b\tau\}$
  - $W + lf$ (light flavour): $W + \{ll, l\tau\}$
- Three top components, depending on the processes considered:
  - Top (b): $t\bar{t}$ and single-top non-resonant production
  - Top ($other$): $t\bar{t}$ and single-top resonant production
  - $t\bar{t}$: exclusive to the 2-lepton channel
- $VV\ bkg$ (backgrounds): all $WW,\ ZZ$ and WZ processes but the signal ones
- Multi-jet: only in the 1-lepton channel

Systematic uncertainties are included in the fit as nuisance parameters as mentioned before, and these comprise the detector systematics, theoretical signal and

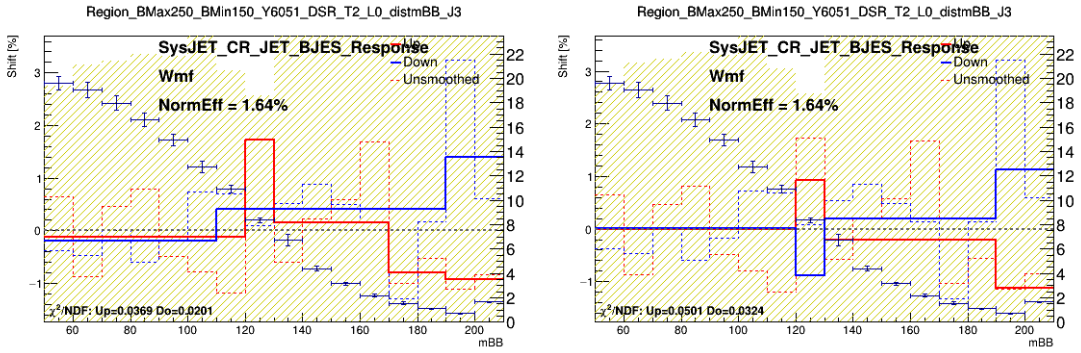
background modelling uncertainties and statistical uncertainties in the simulation samples for the signals and backgrounds. The systematic uncertainties considered in the fit can modify the shape and/or normalisation of the $m_{c\bar{c}}$ distributions.

Smoothing and pruning of nuisance parameters

A *smoothing* procedure is applied in two steps on systematic uncertainties showing large statistical fluctuations resulting from the limited size of the simulated samples used to estimate their impact. In a first iteration bins in the varied distribution relative to the nominal one are merged until the number of local maxima is at most one. In the second step bins are further grouped until the statistical uncertainty of the nominal distribution for each of the merged bins is less than 5%. Irregular distributions may persist after the smoothing procedure, in which case the two smoothed "up" and "down" variations ($\pm 1\sigma$) of the systematic uncertainty can be averaged and symmetrised.

An example of these procedures is shown in Figure 5.5 for a b-jet energy scale systematic for a W+jets sample. In these figures the original up and down variations of the systematic are shown as dotted lines, and the smoothed variations are presented as thick lines. In addition, the points with error bars correspond to the W+jets distribution and its statistical uncertainty is shown as a yellow hatched band.

Nuisance parameters with a negligible impact in regions in the fit are discarded



(a) Only smoothing procedure applied on systematic

(b) Smoothing, averaging and symmetrisation of systematic

Figure 5.5: Treatment of a jet energy scale systematic uncertainty.

in a *pruning* procedure applied in all fit regions, in order to help both the performance and stability of the fit. Normalisation and shape uncertainties for a given sample in a region are neglected if the variation of the template is less than 0.5% in all the bins. The normalisation uncertainty is also dropped if both up and down variations have the same sign ($i.e$, both increase or decrease the nominal normalisation). If only one of the variations is non-zero the shape uncertainty is removed. Additionally, if a sample represents less than 2% of the total background in a region, the signal over background is less than 2% in that region, and the effect of the shape and normalisation uncertainties is less than 0.5%, the shape and normalisation uncertainties of that sample are neglected in that region.

5.8 Results

5.8.1 Signal Strengths and Significances

The measured signal strengths for the three signal processes are:

$$\mu_{VH(H\to c\bar{c})} = -9 \pm 10\,(\text{stat.}) \pm 12\,(\text{syst.})\,, \tag{5.11}$$

$$\mu_{VW(W\to cq)} = 0.83 \pm 0.11\,(\text{stat.}) \pm 0.21\,(\text{syst.})\,, \tag{5.12}$$

$$\mu_{VZ(Z\to c\bar{c})} = 1.16 \pm 0.32\,(\text{stat.}) \pm 0.36\,(\text{syst.})\,. \tag{5.13}$$

The $VW(W \to cq)$ and $VZ(Z \to c\bar{c})$ signal strengths are found to be in agreement with the Standard Model expectations, providing a validation of the analysis. The $VH(H \to c\bar{c})$ signal strength, however, is consistent with both zero and unity, and no evidence for the Standard Model signal is found. The correlation between the $VH(H \to c\bar{c})$ and $VW(W \to cq)$ signal strengths is 17%, while the $VH(H \to c\bar{c})$ and $VZ(Z \to c\bar{c})$ POIs are 16% correlated. The two diboson signal strengths are 17% anti-correlated. The probability of compatibility with the SM, defined as all three POIs being equal to unity, is 84%.

The expected and observed standard deviations over the background-only prediction for the diboson signals are shown in Table 5.9.

Table 5.9: Expected and observed significances for the $VW(W \to cq)$ and $VZ(Z \to c\bar{c})$ processes.

Process	Expected significance	Observed significance
$VW(W \to cq)$	3.8 σ	4.6 σ
$VZ(Z \to c\bar{c})$	2.6 σ	2.2 σ

5.8.2 Limit on $\mu_{VH(H\to c\bar{c})}$

An observed (expected) upper limit on $\mu_{VH(H\to c\bar{c})}$ of 26 (31^{+12}_{-8}) is set at 95% CL, using a modified frequentist CL_s method [66]. The limits for the combined and individual lepton channels, with the latter coming from a fit with the $VH(H \to c\bar{c})$ signal strength decorrelated between lepton channels, are shown in Figure 5.6.

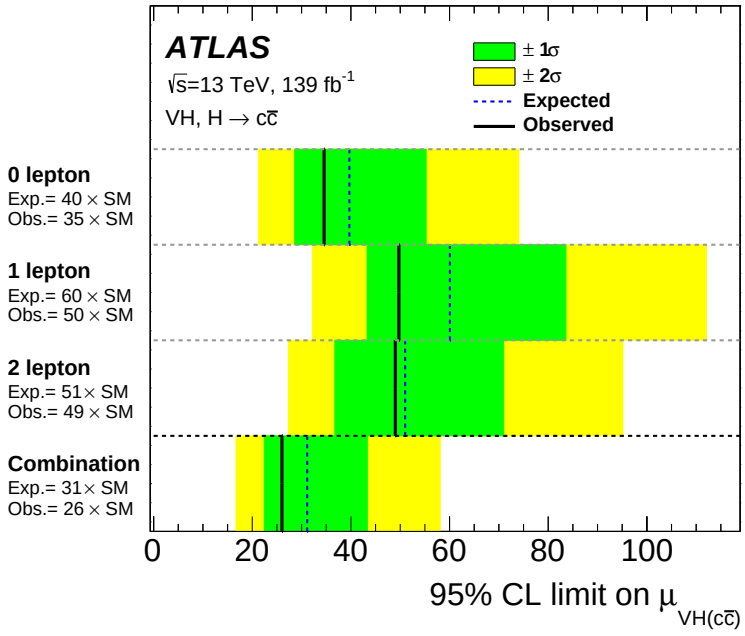


Figure 5.6: Observed and expected 95% CL upper limits on $\mu_{VH(H\to c\bar{c})}$. The individual channel limits come from a fit with five POIs: two diboson signals and one $VH(H \to c\bar{c})$ POI per channel [57].

In order to understand the improvement of this analysis with respect to the previous ATLAS one, which used 36.1 fb^{-1} of pp collisions collected in 2015 and

2016 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, a fit to the 36.1 fb^{-1} dataset was performed, using the same signal regions as the previous iteration, and equally targeting only the 2-lepton channel. This translates into a 36% better expected limit, mainly due to the improved jet flavour tagging performance. Adding to this fit setup the new 2-lepton channel signal and control regions leads to a 43% improvement in the expected limit compared to the previous analysis, with most of the improvement coming from the better flavour tagging performance, but also with contributions from the event categorisation in jet categories (+6%), the new control regions (+10%), and a small loss in sensitivity associated to the measurement of the diboson signal strengths (-7%), as it introduces/enchances the correlations with the $VH(H \to c\bar{c})$ POI. The addition of the total 139 fb^{-1} dataset, in conjunction with 0- and 1- lepton channels, represents a factor five improvement in the expected limit with respect to the previous ATLAS iteration. A comparison of the fit setups of the two analyses is shown in Table 5.10.

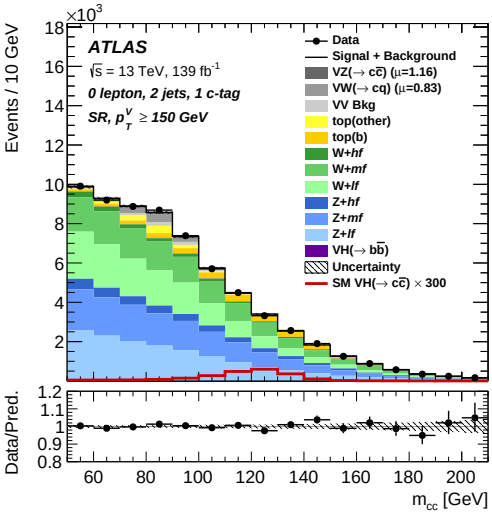
Table 5.10: Comparison between 36.1 fb^{-1} and 139 fb^{-1} ATLAS analyses.

Feature	36.1 fb^{-1}	139 fb^{-1}
Flavour tagging (FTAG)	MV2c100 (b vs c) + MV2cl100 (c vs l) (c-tag)	DL1 (c-tag) + MV2c10 (b-veto)
Tagging efficiencies	c-jets (41%), b-jets (25%), light-jets (5%)	c-jets (27%), b-jets (8%), light-jets (1.6%)
FTAG calibrations	36.1 fb^{-1}	139 fb^{-1}, 80 fb^{-1} for c-jets
Lepton channels	2-lepton	0-, 1- and 2- lepton
Jet categories	2+ jets	2 and 3(+) jets
p_{T}^{V} regimes	75 GeV $< p_{\mathrm{T}}^{V} <$ 150 GeV $p_{\mathrm{T}}^{V} >$ 150 GeV	75 GeV $< p_{\mathrm{T}}^{V} <$ 150 GeV (only in 2L) $p_{\mathrm{T}}^{V} >$ 150 GeV
Signal regions	1 and 2 c-tags	1 and 2 c-tags
Control regions	-	Top $e\mu$ (2L), Top (0L/1L), High ΔR, 0 c-tag
Main backgrounds treatment	Floating Z+jets normalisations in each category	Common floating normalisations
$VH(H \to b\bar{b})$ treatment	SM background Signal region overlap	SM background Orthogonality in signal regions
$VH(H \to b\bar{b})$ fraction in 2 c-tag	6%	0.7%

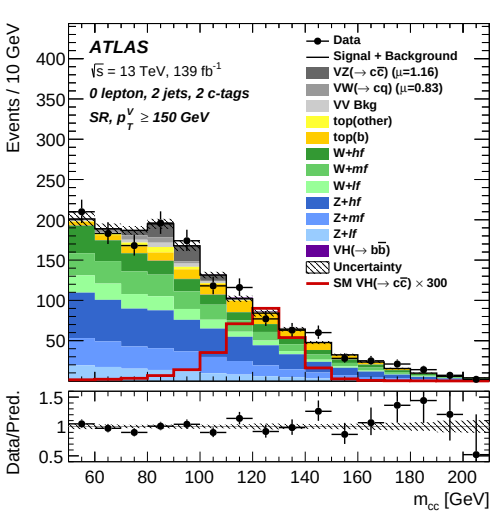
5.8.3 Post-fit $m_{c\bar{c}}$ distributions

Post-fit $m_{c\bar{c}}$ distributions for selected signal regions of the three lepton channels, for categories with one or two c-tags, two jets, and $p_{\mathrm{T}}^{V} > 150$ GeV, are shown in Figure 5.7. The background contributions are represented as filled histograms, with the expected SM $H \to c\bar{c}$ signal multiplied by a factor of 300 as a thick red line superimposed on top. A solid black line corresponds to the sum of the signal and background contributions, including a Higgs signal scaled to the best-fit value of $\mu_{VH(H\to c\bar{c})} = -9$. The combined statistical and systematic uncertainty on this sum is represented by a hatched band. The data are shown as black points. The lower panels in each distribution correspond to the ratio between the data and the sum of the fitted signal and background. The post-fit $m_{c\bar{c}}$ distributions for all signal regions can be found in Appendix A.

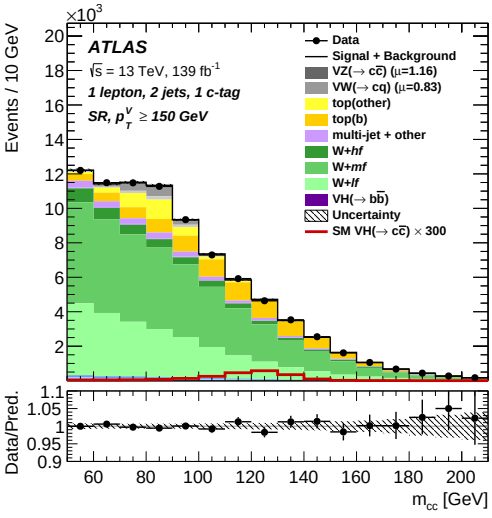
The post-fit $m_{c\bar{c}}$ distributions for the sum in all lepton channels of all the signal regions in the one or two c-tag categories, with all the backgrounds subtracted, are shown in Figure 5.8. The fitted diboson and Higgs signals are present as filled histograms, according to their observed signal strengths. In addition, a thick red line shows the expected SM $H \to c\bar{c}$ signal scaled to the observed 95% CL upper limit of $\mu_{VH(H\to c\bar{c})} = 26$. The data are represented as black dots, and the uncertainty on the fitted backgrounds is shown as a hatched band.



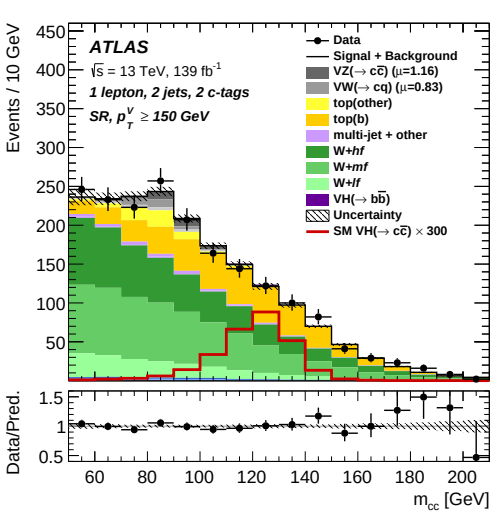
(a) 0-lepton channel, 1 c-tag



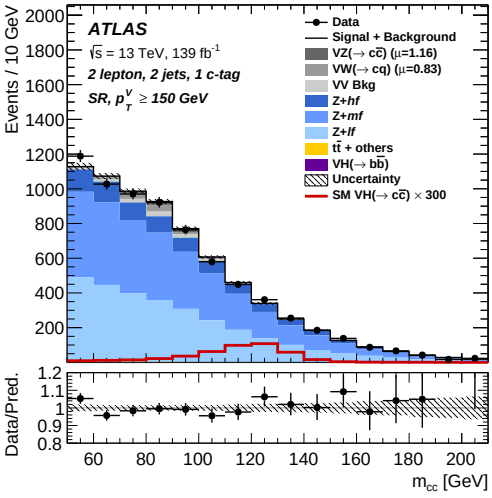
(b) 0-lepton channel, 2 c-tags



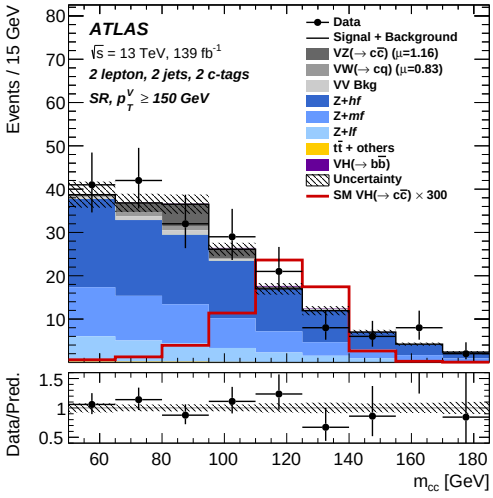
(c) 1-lepton channel, 1 c-tag



(d) 1-lepton channel, 2 c-tags

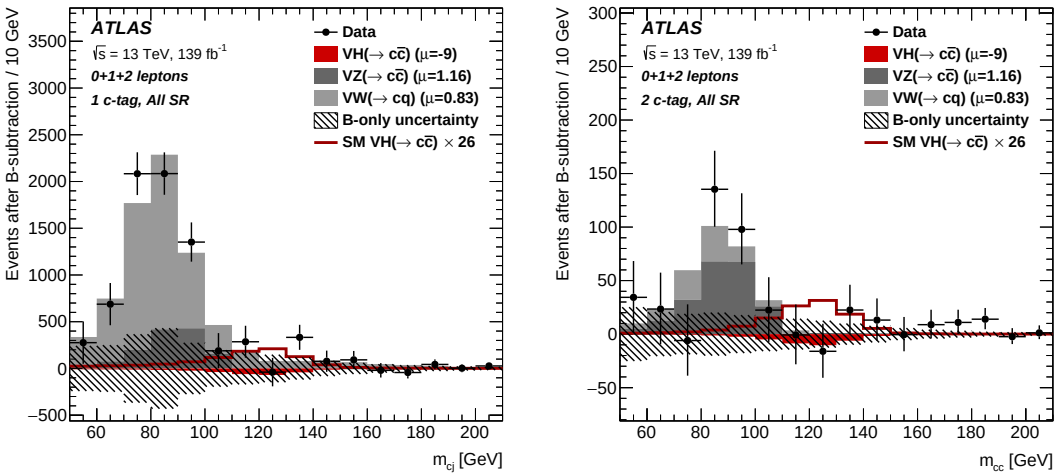


(e) 2-lepton channel, 1 c-tag



(f) 2-lepton channel, 2 c-tags

Figure 5.7: Examples of post-fit $m_{c\bar{c}}$ distributions for the signal regions with 2 jets and $p_{\mathrm{T}}^{V} > 150$ GeV [57].



(a) 1 c-tag (b) 2 c-tags

Figure 5.8: Background-subtracted post-fit $m_{c\bar{c}}$ distributions for the sum of all signal regions in all channels with one or two c-tags categories [57].

5.8.4 Breakdown of uncertainties

The sources of uncertainty for each signal strength are divided in two groups, statistical and systematic errors. Both comprise different sets of uncertainties as described in Section 5.6. The impact of the statistical uncertainties is obtained by subtracting the squared uncertainty obtained from a fit with all the experimental nuisance parameters fixed to their post-fit values to the squared total uncertainty. Conversely, the total systematic impact on the total uncertainty of each signal strength is obtained by subtracting the squared statistical uncertainty from the squared total uncertainty,

$$\sigma_{syst.} = \sqrt{\sigma^2_{total} - \sigma^2_{stat.}}\,. \tag{5.14}$$

The impact of each set of systematic uncertainties is evaluated in a similar way, by fixing the respective nuisance parameters in the fit and then subtracting the square of the obtained uncertainty to the squared total uncertainty.

The statistical and systematic uncertainties are of the same magnitude for the $VH(H \rightarrow c\bar{c})$ POI, while the diboson signal strengths are more limited by the systematic uncertainties. The main sources of uncertainties on $\mu_{VH(H\rightarrow c\bar{c})}$ come hierarchically from the background modelling of the V+jets and top processes, the statistical uncertainty resulting from the limited size of the simulated samples, and the truth-tagging procedure. It should be noted that the use of truth-flavour tagging represents nevertheless a 10% improvement on the $\mu_{VH(H\rightarrow c\bar{c})}$ expected

limit with respect to a scenario using a direct tagging procedure. The hierarchy of the contributions to the diboson POI's uncertainties are similar to the $\mu_{VH(H\to c\bar{c})}$ case. A complete breakdown of the sources of uncertainties on the $VH(H \to c\bar{c})$, $VW(W \to cq)$ and $VZ(Z \to c\bar{c})$ signal strengths is presented in Table 5.11. The mean of the absolute values is shown in the table in cases where the upward and downward systematic variations have different values.

Table 5.11: Breakdown of absolute contributions to the uncertainty in the fitted values of $\mu_{VH(H\to c\bar{c})}$, $\mu_{VW(W\to cq)}$ and $\mu_{VZ(Z\to c\bar{c})}$. The sum in quadrature of uncertainties from different sources may differ from the total due to correlations [57].

Source of uncertainty		$\mu_{VH(c\bar{c})}$	$\mu_{VW(cq)}$	$\mu_{VZ(c\bar{c})}$
Total		15.3	0.24	0.48
Statistical		10.0	0.11	0.32
Systematic		11.5	0.21	0.36
Statistical uncertainties				
Signal normalisation		7.8	0.05	0.23
Other normalisations		5.1	0.09	0.22
Theoretical and modelling uncertainties				
$VH(H\to c\bar{c})$		2.1	< 0.01	0.01
Z+jets		7.0	0.05	0.17
Top quark		3.9	0.13	0.09
W+jets		3.0	0.05	0.11
Diboson		1.0	0.09	0.12
$VH(H\to b\bar{b})$		0.8	< 0.01	0.01
Multi-jet		1.0	0.03	0.02
Simulation samples size		4.2	0.09	0.13
Experimental uncertainties				
Jets		2.8	0.06	0.13
Leptons		0.5	0.01	0.01
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.2	0.01	0.01
Pile-up and luminosity		0.3	0.01	0.01
Flavour tagging	c-jets	1.6	0.05	0.16
	b-jets	1.1	0.01	0.03
	light-jets	0.4	0.01	0.06
	τ-jets	0.3	0.01	0.04
Truth-flavour tagging	ΔR correction	3.3	0.03	0.10
	Residual non-closure	1.7	0.03	0.10

5.8.5 Kappa Interpretation

The $VH(H \rightarrow c\bar{c})$ signal strength is interpreted within the kappa framework, being re-parametrised in terms of the κ_c coupling modifier for the Higgs boson-charm quark interaction. All other coupling modifiers are set to unity, $i.e$, their SM expectation, and no beyond-the-Standard-Model contributions to the Higgs boson width are assumed. Moreover, only modifications to the Higgs boson decay are considered.

Taking Equation 2.18, with $\mu_{VH(H\rightarrow c\bar{c})}$ defined as the ratio of the observed and expected cross-section times branching ratio, leads to the following $VH(H \rightarrow c\bar{c})$ signal strength as a function of κ_c:

$$\mu_{VH(H\rightarrow c\bar{c})}(\kappa_c) = \frac{\kappa_c^2}{1 + BR^{SM}_{H\rightarrow c\bar{c}}(\kappa_c^2 - 1)} \tag{5.15}$$

using the Standard Model branching ratio for the Higgs decay to charm quarks.

In this parametrisation only analyses sensitive to signal strengths closer to the SM expected value are able to set constraints on the Higgs-charm coupling modifier, as for large values of κ_c the parametrisation takes the approximate form $\mu_{VH(H\rightarrow c\bar{c})}(\kappa_c) \approx \frac{1}{BR^{SM}_{H\rightarrow c\bar{c}}}$, reaches a limit at around $\mu \sim 35$. This parametrisation is shown in Figure 5.9.

A profile likelihood scan on the re-parametrised signal strength allows for an observed (expected) constraint of $|\kappa_c| \leq 8.5$ (12.4) to be set at 95% CL, as shown

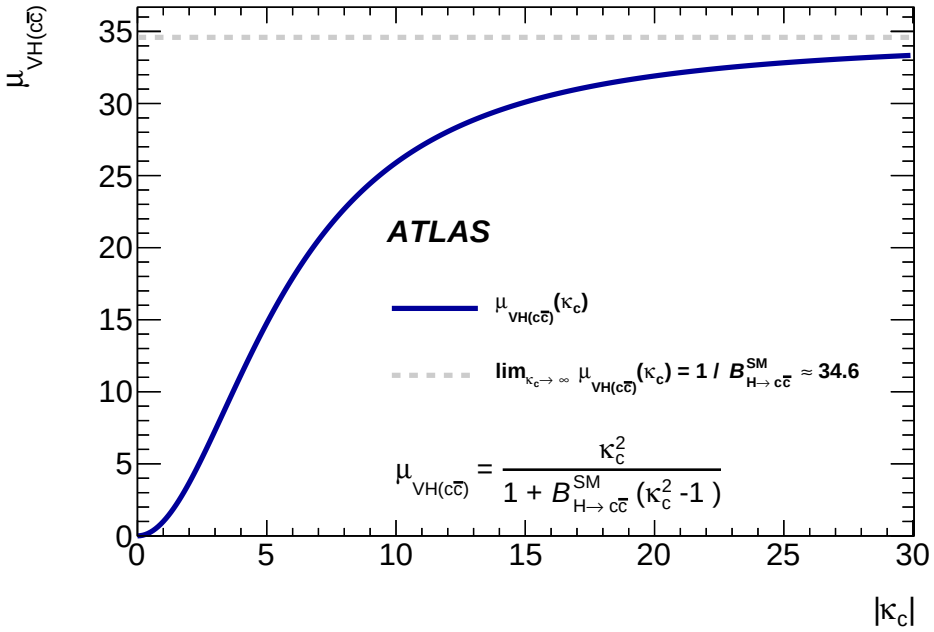


Figure 5.9: Parametrisation of the $VH(H \to c\bar{c})$ signal strength as a function of $|\kappa_c|$ [57, 58].

in Fig. 5.10.

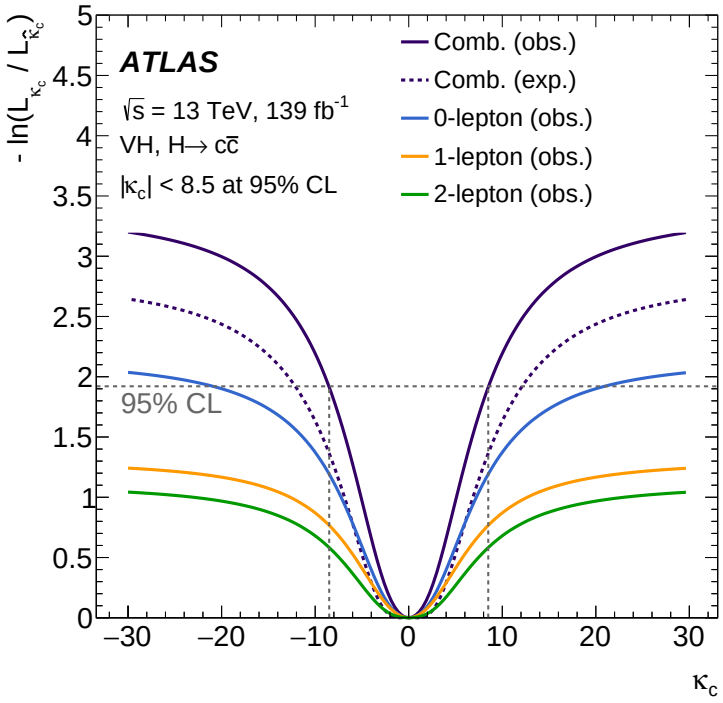


Figure 5.10: Observed and expected constraints from a profile likelihood scan on $|\kappa_c|$ at 95% CL [57].

5.9 Combination of $VH(H \to c\bar{c})$ and $VH(H \to b\bar{b})$ analyses

The $VH(H \to c\bar{c})$ analysis is combined with the $VH(H \to b\bar{b})$ measurement [61], given the similar analysis features between both. This combination allows for the correlation of systematic uncertainties, as well as improved constraints on the coupling modifiers of the Higgs decay to bottom and charm quarks. The ATLAS $VH(H \to b\bar{b})$ analysis targets VH-produced events, in which the Higgs boson decays to two bottom quarks. Three lepton channels are defined corresponding to the different decay modes of the vector bosons, matching the $VH(H \to c\bar{c})$ categorisation. The object and event selection is also the same between the two analyses.

As mentioned before, the signal regions in the $VH(H \to b\bar{b})$ measurement are constructed using two b-tagged jets, using the MV2C10 algorithm, which corresponds to the one providing the b-jet vetoes in the $VH(H \to c\bar{c})$ search. The information in the $VH(H \to b\bar{b})$ signal regions is exploited by a boosted decision tree (BDT), while the analogous charm analysis performs linear cuts in relevant variables. The main differences between the two analyses are outlined in Table 5.12.

The common experimental systematic uncertainties are correlated between the two analyses, except for the jet flavour tagging uncertainties, due to different calibration procedures used for b- and c-tagging in each analysis. The background

Table 5.12: Main differences between the $VH(H \to c\bar{c})$ and $VH(H \to b\bar{b})$ analyses.

Feature	$VH(H \to c\bar{c})$	$VH(H \to b\bar{b})$
Flavour Tagging	DL1 (c-tag) + MV2C10 (b-veto)	MV2C10 (b-tag)
Tagging Categories	1 or 2 c-tag+b-veto SRs (0 c-tag CRs)	2 b-tag SRs
Analysis Design	Cut-based analysis	BDT to discriminate signal and background
Discriminating Variable	Invariant di-jet mass	BDT discriminant

normalisations and modelling uncertainties are also uncorrelated between the two searches, given different templates and strategies were used in each.

The likelihood is parametrised in the combination as the product of the individual likelihoods of the two analyses, with two POIs: $\mu_{VH(H\to c\bar{c})}$ and $\mu_{VH(H\to b\bar{b})}$. The previous diboson POIs in the $VH(H \to c\bar{c})$ analysis become floating normalisation factors.

The fitted signal strengths for the combination (and the comparison with the individual "*indv.*" analyses ones) are:

$$\mu_{VH(H\to c\bar{c})} = -9 \pm 10\,(\text{stat.}) \pm 11\,(\text{syst.})\,, \tag{5.16}$$

$$\mu_{VH(H\to b\bar{b})} = 1.06 \pm 0.12\,(\text{stat.})^{+0.15}_{-0.13}\,(\text{syst.})\,, \tag{5.17}$$

$$\mu_{VH(H\to c\bar{c})\,(indv.)} = -9 \pm 10\,(\text{stat.}) \pm 12\,(\text{syst.})\,, \tag{5.18}$$

$$\mu_{VH(H\to b\bar{b})\,(indv.)} = 1.02^{+0.12}_{-0.11}\,(\text{stat.})^{+0.14}_{-0.13}\,(\text{syst.})\,, \tag{5.19}$$

with the $\mu_{VH(H\to c\bar{c})}$ and $\mu_{VH(H\to b\bar{b})}$ 12% anti-correlated. The signal strengths are consistent with the ones obtained for the individual analyses. The 68% and 95%

CL contours are shown in Figure 5.11.

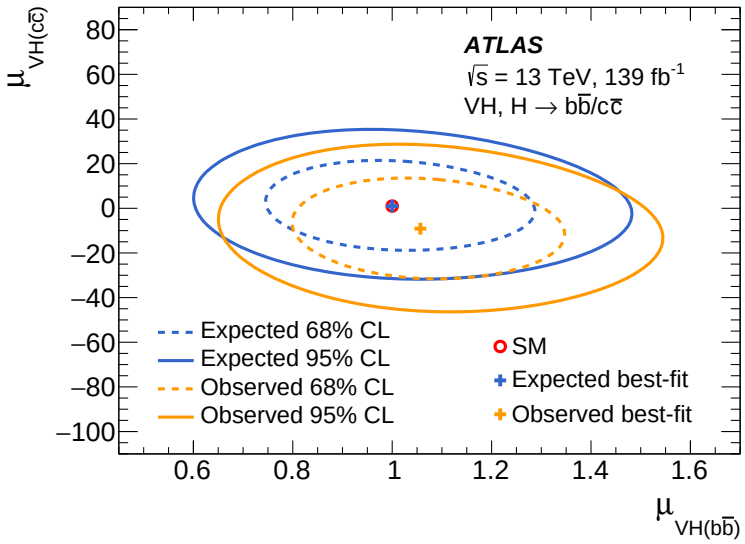


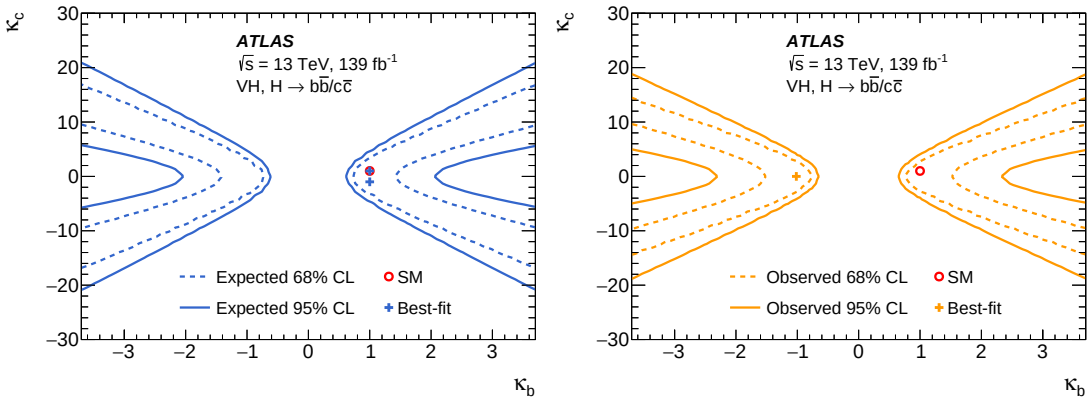
Figure 5.11: Observed and expected 68% and 95% CL contours and respective best fit values for $\mu_{VH(H\to c\bar{c})}$ and $\mu_{VH(H\to b\bar{b})}$ [57].

There exists a small overlap of events in the $VH(H \to c\bar{c})$ control regions with the $VH(H \to b\bar{b})$ analysis, namely of around 6% in the top CRs in the 0- and 1-lepton channels, and of around 4% in the 0 c-tag(+b-veto) CR in the 1- and 2-lepton channels. The removal of these events from the $VH(H \to c\bar{c})$ control regions in the combination was performed in a test and proved not to affect the results. Correlating the background normalisations between the two analyses also has no impact on results.

The likelihood is re-parametrised in terms of both the charm (κ_c) and bottom (κ_b) coupling modifiers, with all the other coupling modifiers set to their SM predictions. Only modifications to the Higgs boson decay are considered, and the Higgs boson width is not constrained. Gluon-initiated ZH production has no explicit parametrisation as a function of κ_c, being only taken into account for κ_b. This,

however, only has an impact for large values of κ_b ($\kappa_b \gtrsim 10$). Charm initiated VH production is neglected as its impact is negligible for $\kappa_c \leq 50$ [115] (above the low values of κ_c probed by this combination).

The observed and expected constraints on κ_c and κ_b are presented in Figure 5.12. The charm and bottom coupling modifiers compensate for each other's effect in most of the cases, via the Higgs boson width, creating the diagonal limits. The observed best fit value is $(\kappa_c, \kappa_b) = (-1.02, 0)$. The difference in the likelihood value between the best fit value and $(\kappa_c, \kappa_b) = (1.02, 0)$ is just 0.02, with the analysis having little power to constrain the sign of κ_b. The small likelihood asymmetry comes from the b-quark loop contributions to gluon-initiated ZH production.



(a) Expected (b) Observed

Figure 5.12: Observed and expected constraints on κ_c and κ_b at 68% and 95% CL [57].

A profile likelihood scan parametrised in terms of κ_c/κ_b and with κ_b free to float, allows for an observed (expected) constraint of $|\kappa_c/\kappa_b| \leq 4.5\ (5.1)$ at 95% CL to be set, as shown in Fig. 5.13. Given the ratio of the bottom- and charm-quark masses being equal to $m_b(m_H)/m_c(m_H) = 4.578 \pm 0.008$ [116], this result, under the Higgs width assumption, constrains the Higgs boson coupling to charm quarks to be

weaker than its coupling to bottom quarks at 95% CL.

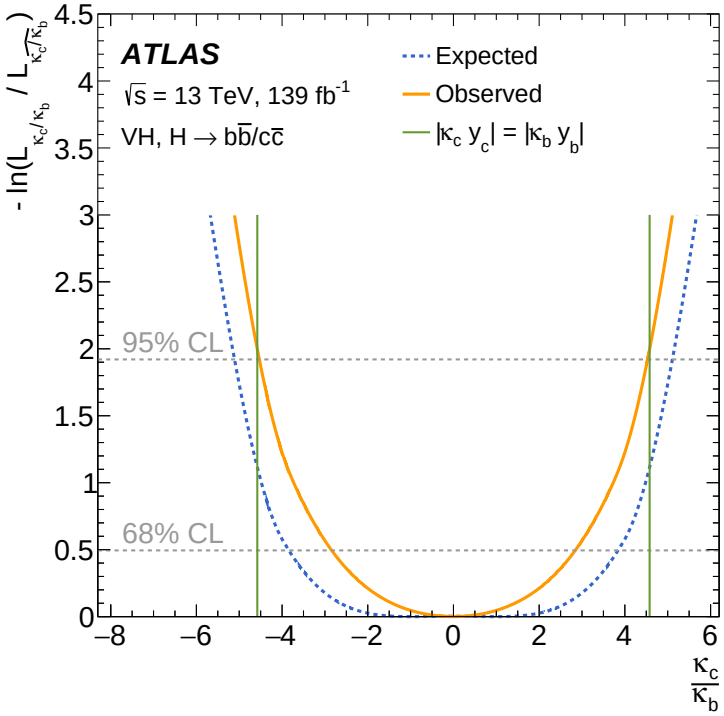


Figure 5.13: Observed and expected constraints from a profile likelihood scan on $|\kappa_c/\kappa_b|$ at 95% CL, where κ_b is a free parameter. A scenario with an equal coupling of the Higgs boson to charm and bottom quarks is represented by green lines, corresponding to the ratio $|\kappa_c/\kappa_b| = m_b/m_c$ [57].

5.10 ATLAS sensitivity to $H \to b\bar{b}$ and $H \to c\bar{c}$ decays in VH production at the HL-LHC

The $VH(H \to c\bar{c})$ and $VH(H \to b\bar{b})$ Run 2 analyses are extrapolated to the High-Luminosity LHC (HL-LHC) scenario [65], based on an expected centre-of-mass energy $\sqrt{s} = 14$ TeV and a total integrated luminosity of 3000 fb^{-1}.

The normalisations of the signal and background expected yields are scaled for

the increase in integrated luminosity from 139 fb^{-1} to 3000 fb^{-1} and for the increase in the centre-of-mass energy from 13 to 14 TeV. A summary of the centre-of-mass scale factors, which are derived inclusively in p_{T}^{V} using the expected cross-sections [19] and are applied per process, are summarised in Table 5.13.

Table 5.13: Centre-of-mass scale factors from 13 to 14 TeV for signal and background expected yields [65].

Process	Scale factor
$qq \to WH(H \to c\bar{c}/b\bar{b})$	1.10
$qq \to ZH(H \to c\bar{c}/b\bar{b})$	1.11
$gg \to ZH(H \to c\bar{c}/b\bar{b})$	1.18
$t\bar{t}$ $gg \to ZZ$	1.16
$qq \to VV$ V+jets single-top	1.10

Systematic errors are taken into account by scaling their Run 2 values according to the expected increased statistics and improvements in analysis techniques associated with the larger dataset [117, 118]. Theory, background modelling and jet flavour tagging uncertainties (with the exception of the light-jets component in the $VH(H \to c\bar{c})$ analysis) are scaled by a factor 0.5. Missing transverse energy associated uncertainties are also halved. The luminosity uncertainty is scaled by 0.58, representing a reduction from the Run 2 1.7% value to 1%. The remaining experimental uncertainties are kept at their Run 2 levels given their small impact on the analysis. Statistical uncertainties from the limited size of the simulated samples are not considered, nor truth-flavour tagging uncertainties, based

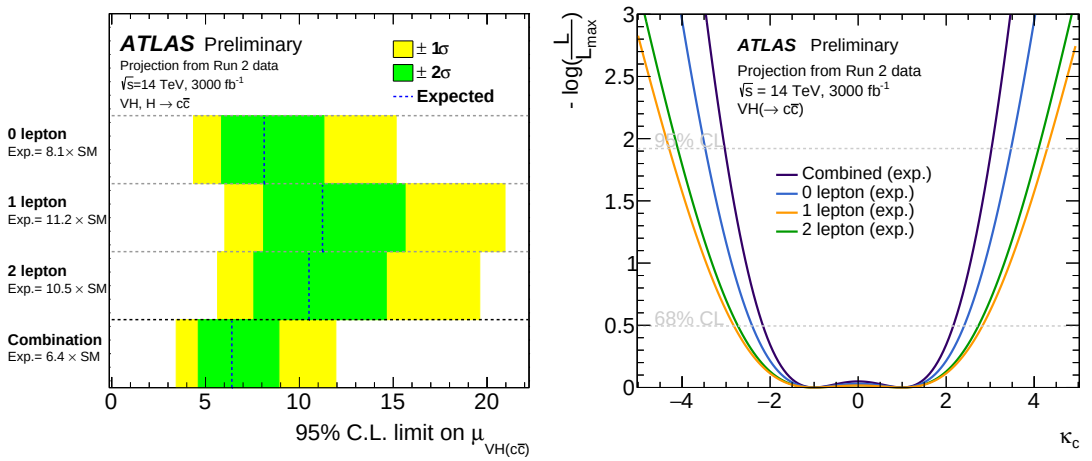
on the expectation of having in the future large enough samples to render these uncertainties negligible. A summary of the scale factors applied to systematic uncertainties is shown in Table 5.14.

Table 5.14: Scale factors applied to systematic uncertainties to reflect expected improvements in the HL-LHC scenario. A scale factor of 1 represents no improvement foreseen [65].

Uncertainties	Scale Factor
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.5
Lepton	1
Jet	1
Flavour tagging c-, b- and τ-jets	0.5
Flavour tagging light-jets (MV2c10 in $VH(H \to b\bar{b})$)	0.5
Flavour tagging light-jets (DL1 in $VH(H \to c\bar{c})$)	1.0
Luminosity	0.58
Signal modelling	0.5
Background modelling	0.5
MC statistics	0
Truth-tagging uncertainties (in $VH, H \to c\bar{c}$)	0

The extrapolation to the HL-LHC scenario of the individual $VH(H \to c\bar{c})$ analysis allows for an expected upper limit on $\mu_{VH(H\to c\bar{c})}$ of 6.4×SM to be set at 95% CL, for an expected signal strength of $\mu_{VH(H\to c\bar{c})} = 1.0 \pm 3.2 = 1.0 \pm 2.0\,(\text{stat.})^{+2.6}_{-2.5}\,(\text{syst.})$. Re-interpreting the likelihood in terms of the Higgs-charm coupling modifier allows for an expected constraint of $|\kappa_c| \leq 3$ at 95% CL to be obtained. The respective limits and likelihood scans are shown in Figure 5.14.

The main sources of uncertainty on the expected $VH(H \to c\bar{c})$ signal strength come hierarchically from the theoretical and background modelling uncertain-



(a) Expected 95% CL upper limits on $\mu_{VH(H\to c\bar{c})}$.

(b) Expected profile likelihood scans on κ_c.

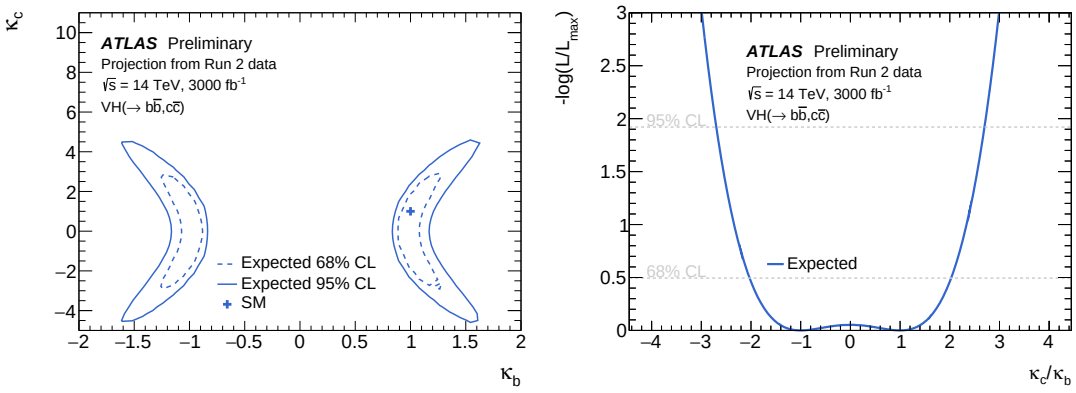
Figure 5.14: Expected limits on $\mu_{VH(H\to c\bar{c})}$ and constraints on κ_c for an extrapolation of the Run 2 analysis to the HL-LHC scenario with $\sqrt{s} = 14$ TeV and a total integrated luminosity of 3000 fb^{-1}. The individual channel likelihoods (limits) come from a fit with three (five) POIs: one $VH(H \to c\bar{c})$ POI per channel (and two diboson signals) [65].

ties for Z+jets, top processes and W+jets, followed by the flavour tagging uncertainties. A complete breakdown of the sources of the expected uncertainties is presented in Table 5.15. The mean of the absolute values is shown in the table in cases where the upward and downward systematic variations have different values.

The extrapolation of the combination of the $VH(H \to c\bar{c})$ and $VH(H \to b\bar{b})$ analyses allows for an expected constraint of $|\kappa_c/\kappa_b| \leq 2.7$ at 95% CL to be set. The profile likelihood scans on the coupling modifiers are shown in Figure 5.15.

Table 5.15: Breakdown of contributions to the uncertainty in the fitted values of $\mu_{VH(H\to c\bar{c})}$ from a fit to a simulated dataset scaled to $\sqrt{s} = 14$ TeV and an integrated luminosity of 3000 fb^{-1}. The sum in quadrature of different sources of uncertainty may differ from the total due to correlations [65].

Source of uncertainty		$\Delta\mu_{VH}^{c\bar{c}}$
Total		3.21
Statistical		1.97
Systematics		2.53
Statistical uncertainties		
Data statistics only		1.59
Floating normalisations		0.95
Theoretical and modelling uncertainties		
$VH(H \to c\bar{c})$		0.27
Z+jets		1.77
Top-quark		0.96
W+jets		0.84
Diboson		0.34
$VH(H \to b\bar{b})$		0.29
Multi-Jet		0.09
Experimental uncertainties		
Jets		0.59
Leptons		0.20
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.18
Pile-up and luminosity		0.19
Flavour tagging	c-jets	0.61
	b-jets	0.16
	light-jets	0.51
	τ-jets	0.19



(a) Expected profile likelihood scans on κ_c and κ_b.

(b) Expected profile likelihood scans for κ_c/κ_b, with κ_b free to float.

Figure 5.15: Expected constraints on κ_c and κ_b and κ_c/κ_b for an extrapolation of the Run 2 analysis to the HL-LHC scenario with $\sqrt{s} = 14$ TeV and a total integrated luminosity of 3000 fb^{-1} [65].

Chapter 6

Conclusion

The properties of a Higgs boson with mass of 125 GeV have been studied by the LHC experiments since its discovery in 2012, in order to understand its compatibility with the Standard Model prediction. One of such properties is the Higgs coupling to fermions, with the SM predicting the mass generation of all fermions (apart from neutrinos) via the Higgs mechanism. The interactions of the Higgs to the third generation of fermions have been observed by both the ATLAS and CMS experiments, and evidence for the coupling of the Higgs to muons was found by CMS.

In the endeavour to test the SM all the Higgs-fermion couplings must be measured, and in that way allow to understand whether these interactions correspond to the SM predictions or show any deviation. Such discrepancies are found in extensions of the SM, predicting different mass generation mechanisms for the third and remaining generations of fermions. The natural next step in the effort

to measure all Higgs couplings to fermions now lies on the searches for Higgs decays to charm quarks.

The optimal performance of the LHC and ATLAS experiments are crucial in the effort to better understand the SM and in particular the Higgs, with upgrades taking place on several sections of the detector, including the hardware-based level 1 calorimeter sub-trigger system. Such an upgrade is comprised of several elements, from the actual hardware boards to the firmware running on them and the respective simulations. This thesis describes the work performed in the context of this upgrade in a firmware-mimicking simulation and trigger monitoring frameworks.

The ATLAS collaboration performed a $VH(H \to c\bar{c})$ analysis using 139 fb^{-1} of proton-proton collision data collected at the LHC at $\sqrt{s} = 13$ TeV, which is reported in this thesis. Through the use of multivariate jet flavour tagging algorithms, an optimised event selection and categorisation, and a larger dataset, this search improves by a factor five on the previous limit obtained by the ATLAS experiment, setting an observed (expected) upper limit on $\mu_{VH(H\to c\bar{c})}$ of 26 (31^{+12}_{-8}) at 95% CL. This direct search for the Higgs-charm coupling allows for an observed (expected) constraint of $|\kappa_c| \leq 8.5$ (12.4) to be set at 95% CL.

The extrapolation of the individual $VH(H \to c\bar{c})$ analysis and of its combination with the $VH(H \to b\bar{b})$ search demonstrates room for future improvements, as the Run 2 analyses extrapolated to HL-LHC conditions are not sensitive enough to test the SM predictions. Changes on the analysis design, better flavour tagging

performance and further reductions of the background modelling uncertainties were not considered in the extrapolation, and are possibilities for the future which would result in improved sensitivity at the HL-LHC.

References

[1] R. L. Workman and Others (Particle Data Group). "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[2] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[3] Lyndon Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[4] Wikimedia - MissMJ. *PBS NOVA, Fermilab, Office of Science, United States Department of Energy*. Particle Data Group.

[5] D. H. Perkins. *Introduction to High Energy Physics*. 4^{th}. Cambridge University Press, 2000.

[6] B. Andersson et al. "Parton fragmentation and string dynamics". In: *Physics Reports* 97.2 (1983), pp. 31–145.

[7] Roger Barlow. "Jets in high-energy interactions". In: *Reports on Progress in Physics* 56.9 (1993), pp. 1067–1143. DOI: 10.1088/0034-4885/56/9/001. URL: https://doi.org/10.1088/0034-4885/56/9/001.

[8] Gavin P Salam. "Elements of QCD for hadron colliders". In: (2010). Lectures given at the 2009 European School of High-Energy Physics, Bautzen, Germany, 14-27 June 2009. DOI: 10.5170/CERN-2010-002.45. arXiv: 1011.5131. URL: https://cds.cern.ch/record/1281947.

[9] P.W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12.2 (1964), pp. 132–133. DOI: https://doi.org/10.1016/0031-9163(64)91136-9.

[10] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

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

[12] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[13] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett. B* 716 (2012), p. 30. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[14] ATLAS and CMS Collaborations. "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments". In: *Phys. Rev. Lett.* 114 (2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex].

[15] John C. Collins, Davison E. Soper, and George Sterman. "Factorization of Hard Processes in QCD". In: (2004). DOI: 10.48550/ARXIV.HEP-PH/0409313. URL: https://arxiv.org/abs/hep-ph/0409313.

[16] Stefan Hoeche. *Simulations*. URL: https://theory.slac.stanford.edu/our-research/simulations.

[17] A. Heinson and T. R. Junk. "Observation of Single Top Quark Production". In: *Annual Review of Nuclear and Particle Science* 61 (2011). arXiv:1101.1275 [hep-ex].

[18] Andrew Chisholm. *Introduction to Heavy Flavour Jet Tagging with ATLAS*. Presented at ATLAS $H \to b\bar{b}$ and Flavour Tagging Workshop. 2017. URL: https://indico.cern.ch/event/655628/contributions/2670400/attachments/1518249/2370617/ASC_FTagHbbWS_BTagging.pdf.

[19] LHC Higgs Cross Section Working Group, D. de Florian, C. Grojean, F. Maltoni, C. Mariotti, A. Nikitenko, M. Pieri, P. Savard, M. Schumacher, R. Tanaka (Eds.) "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector". In: (2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[20] LHC Higgs Cross Section Working Group et al. "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties". In: *CERN-2013-004* (2013). DOI: 10.5170/CERN-2013-004. arXiv: 1307.1347 [hep-ph].

[21] ATLAS Collaboration. "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector". In: *Phys. Lett. B* 784 (2018), p. 173. DOI: 10.1016/j.physletb.2018.07.035. arXiv: 1806.00425 [hep-ex].

[22] CMS Collaboration. "Observation of $t\bar{t}H$ Production". In: *Phys. Rev. Lett.* 120 (2018), p. 231801. DOI: 10.1103/PhysRevLett.120.231801. arXiv: 1804.02610 [hep-ex].

[23] ATLAS Collaboration. "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector". In: *Phys. Lett. B* 786 (2018), p. 59. DOI: 10.1016/j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex].

[24] CMS Collaboration. "Observation of Higgs Boson Decay to Bottom Quarks". In: *Phys. Rev. Lett.* 121 (2018), p. 121801. DOI: 10.1103/PhysRevLett.121.121801. arXiv: 1808.08242 [hep-ex].

[25] ATLAS Collaboration. "Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 99 (2019), p. 072001. DOI: 10.1103/PhysRevD.99.072001. arXiv: 1811.08856 [hep-ex].

[26] CMS Collaboration. "Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector". In: *Phys. Lett. B* 779 (2018), p. 283. DOI: 10.1016/j.physletb.2018.02.004. arXiv: 1708.00373 [hep-ex].

[27] CMS Collaboration. "Evidence for Higgs boson decay to a pair of muons". In: *JHEP* 01 (2021), p. 148. DOI: 10.1007/JHEP01(2021)148. arXiv: 2009.04363 [hep-ex].

[28] ATLAS Collaboration. "A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector". In: *Phys. Lett. B* 812 (2021), p. 135980. DOI: 10.1016/j.physletb.2020.135980. arXiv: 2007.07830 [hep-ex].

[29] ATLAS Collaboration. *Combined measurements of Higgs boson production and decay using up to 139 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*. ATLAS-CONF-2021-053. 2021. URL: https://cds.cern.ch/record/2789544.

[30] A. Djouadi, J. Kalinowski, and M. Spira. "HDECAY: a program for Higgs boson decays in the Standard Model and its supersymmetric extension". In: *Comput. Phys. Commun.* 108 (1998), p. 56. DOI: 10.1016/S0010-4655(97)00123-9. arXiv: hep-ph/9704448.

[31] A. Gizhko et al. "Running of the charm-quark mass from HERA deep-inelastic scattering data". In: *Physics Letters B* 775 (2017), pp. 233–238. DOI: https://doi.org/10.1016/j.physletb.2017.11.002.

[32] Javier Aparisi et al. "m_b at m_H: The Running Bottom Quark Mass and the Higgs Boson". In: *Phys. Rev. Lett.* 128 (12 2022), p. 122001. DOI: 10.1103/PhysRevLett.128.122001.

[33] ATLAS Collaboration. "Searches for exclusive Higgs and Z boson decays into $J/\psi\,\gamma$, $\psi(2S)\,\gamma$, and $\Upsilon(nS)\,\gamma$ at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 786 (2018), p. 134. DOI: 10.1016/j.physletb.2018.09.024. arXiv: 1807.00802 [hep-ex].

[34] ATLAS Collaboration. "Measurements of the Higgs boson inclusive and differential fiducial cross sections in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 80 (2020), p. 942. DOI: 10.1140/epjc/s10052-020-8223-0. arXiv: 2004.03969 [hep-ex].

[35] ATLAS Collaboration. "Measurements of the Higgs boson inclusive and differential fiducial cross-sections in the diphoton decay channel with pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 08 (2022), p. 027. DOI: 10.1007/JHEP08(2022)027. arXiv: 2202.00487 [hep-ex].

[36] ATLAS Collaboration. "Measurement of the total and differential Higgs boson production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector by combining the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ decay channels". In: (2022). arXiv: 2207.08615 [hep-ex].

[37] F. J. Botella et al. "What if the masses of the first two quark families are not generated by the standard model Higgs boson?" In: *Phys. Rev. D* 94.11 (2016), p. 115031. DOI: 10.1103/PhysRevD.94.115031. arXiv: 1602.08011 [hep-ph].

[38] Diptimoy Ghosh, Rick Sandeepan Gupta, and Gilad Perez. "Is the Higgs mechanism of fermion mass generation a fact? A Yukawa-less first-two-generation model". In: *Phys. Lett. B* 755 (2016), pp. 504–508. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2016.02.059.

[39] Shaouly Bar-Shalom and Amarjit Soni. "Universally enhanced light-quarks Yukawa couplings paradigm". In: *Phys. Rev. D* 98 (2018), p. 055001. DOI: 10.1103/PhysRevD.98.055001.

[40] *LEP design report*. Copies shelved as reports in LEP, PS and SPS libraries. Geneva: CERN, 1984. URL: https://cds.cern.ch/record/102083.

[41] The LHCb Collaboration. "The LHCb Detector at the LHC". In: *Journal of Instrumentation* 3.08 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005. URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08005.

[42] The ALICE Collaboration. "The ALICE experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002. URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08002.

[43] CMS Collaboration. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[44] ATLAS Collaboration. "Operation of the ATLAS trigger system in Run 2". In: *JINST* 15 (2020), P10004. DOI: 10.1088/1748-0221/15/10/P10004. arXiv: 2007.12539 [hep-ex].

[45] Karolos Potamianos. "The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider". In: *Proceedings, EPS-HEP 2015*. DOI: 10.48550/arXiv.1608.07850. arXiv: 1608.07850 [hep-ph].

[46] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report*. ATLAS-TDR-19; CERN-LHCC-2010-013. 2010. URL: https://cds.cern.ch/record/1291633.

[47] ATLAS Collaboration. *ATLAS TDAQ System Phase-I Upgrade: Technical Design Report*. ATLAS-TDR-023; CERN-LHCC-2013-018. 2013. URL: https://cds.cern.ch/record/1602235.

[48] I. Béjar Alonso, O. Brüning, P. Fessia, M. Lamont, L. Rossi, L. Tavian, M. Zerlauth (Eds.) *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2020. DOI: 10.23731/CYRM-2020-0010. URL: https://cds.cern.ch/record/2749422.

[49] C. Antel (ATLAS Collaboration). *Upgrade and Physics Plans for ATLAS in LHC Run 4*. 2021. URL: https://cds.cern.ch/record/2775629.

[50] R. Frühwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(87)90887-4. URL: https://www.sciencedirect.com/science/article/pii/0168900287908874.

[51] ATLAS Collaboration. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 79 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [hep-ex].

[52] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 76 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[53] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

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

[55] ATLAS Collaboration. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 96 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[56] ATLAS Collaboration. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *Eur. Phys. J. C* 76 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].

[57] ATLAS Collaboration. "Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector". In: *Eur. Phys. J. C* 82 (2022), p. 717. DOI: 10.1140/epjc/s10052-022-10588-3. arXiv: 2201.11428 [hep-ex].

[58] ATLAS Collaboration. *Auxiliary figures and tables from The European Physical Journal C 82.8 (2022), p. 717 paper*. 2022. URL: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2021-12/#auxstuff.

[59] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 78 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[60] ATLAS Collaboration. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].

[61] ATLAS Collaboration. "Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector". In: *Eur. Phys. J. C* 81 (2021), p. 178. DOI: 10.1140/epjc/s10052-020-08677-2. arXiv: 2007.02873 [hep-ex].

[62] ATLAS Collaboration. *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*. ATLAS-CONF-2018-001. 2018. URL: https://cds.cern.ch/record/2306649.

[63] ATLAS Collaboration. *Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV*. ATLAS-CONF-2018-006. 2018. URL: https://cds.cern.ch/record/2314418.

[64] ATLAS Collaboration. "Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment". In: *Phys. Rev. Lett.* 120 (2018), p. 211802. DOI: 10.1103/PhysRevLett.120.211802. arXiv: 1802.04329 [hep-ex].

[65] ATLAS Collaboration. *Extrapolation of ATLAS sensitivity to $H \to b\bar{b}$ and $H \to c\bar{c}$ decays in VH production at the HL-LHC*. ATL-PHYS-PUB-2021-039. 2021. URL: https://cds.cern.ch/record/2788490.

[66] Alexander L. Read. "Presentation of search results: the CL_S technique". In: *J. Phys. G* 28 (2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313.

[67] CMS Collaboration. "A search for the standard model Higgs boson decaying to charm quarks". In: *JHEP* 03 (2020), p. 131. DOI: 10.1007/JHEP03(2020)131. arXiv: 1912.01662 [hep-ex].

[68] CMS Collaboration. "Search for Higgs boson decay to a charm quark-antiquark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: (2022). arXiv: 2205.05550 [hep-ex].

[69] 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.

[70] ATLAS Collaboration. "ATLAS data quality operations and performance for 2015-2018 data-taking". In: *JINST* 15 (2020), P04003. DOI: 10.1088/1748-0221/15/04/P04003. arXiv: 1911.04632 [physics.ins-det].

[71] ATLAS Collaboration. "Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking". In: *JHEP* 08 (2020), p. 080. DOI: 10.1007/JHEP08(2020)080. arXiv: 2005.09554 [hep-ex].

[72] ATLAS Collaboration. "Performance of electron and photon triggers in ATLAS during LHC Run 2". In: *Eur. Phys. J. C* 80 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex].

[73] ATLAS Collaboration. "Performance of the ATLAS muon triggers in Run 2". In: *JINST* 15 (2020), P09015. DOI: 10.1088/1748-0221/15/09/p09015. arXiv: 2004.13447 [hep-ex].

[74] GEANT4 Collaboration, S. Agostinelli, et al. "GEANT4 – a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.

[75] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C* 70 (2010), p. 823. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[76] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[77] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[78] Stefan Höche et al. "QCD matrix elements and truncated showers". In: *JHEP* 05 (2009), p. 053. DOI: 10.1088/1126-6708/2009/05/053. arXiv: 0903.1219 [hep-ph].

[79] Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

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

[81] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[82] Richard D. Ball et al. "Parton distributions for the LHC run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[83] ATLAS Collaboration. "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *JHEP* 09 (2014), p. 145. DOI: 10.1007/JHEP09(2014)145. arXiv: 1406.3660 [hep-ex].

[84] M. L. Ciccolini, S. Dittmaier, and M. Krämer. "Electroweak radiative corrections to associated WH and ZH production at hadron colliders". In: *Phys. Rev. D* 68 (2003), p. 073003. DOI: 10.1103/PhysRevD.68.073003. arXiv: hep-ph/0306234 [hep-ph].

[85] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders". In: *Phys. Lett. B* 579 (2004), pp. 149–156. DOI: 10.1016/j.physletb.2003.10.112. arXiv: hep-ph/0307206.

[86] Oliver Brein et al. "Top-quark mediated effects in hadronic Higgs-Strahlung". In: *Eur. Phys. J. C* 72 (2012), p. 1868. DOI: 10.1140/epjc/s10052-012-1868-6. arXiv: 1111.0761 [hep-ph].

[87] Ansgar Denner et al. "HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders". In: *Comput. Phys. Commun.* 195 (2015), pp. 161–171. DOI: 10.1016/j.cpc.2015.04.021. arXiv: 1412.5390 [hep-ph].

[88] Oliver Brein, Robert V. Harlander, and Tom J. E. Zirke. "vh@nnlo – Higgs Strahlung at hadron colliders". In: *Comput. Phys. Commun.* 184 (2013), pp. 998–1003. DOI: 10.1016/j.cpc.2012.11.002. arXiv: 1210.5347 [hep-ph].

[89] Gavin Cullen et al. "Automated one-loop calculations with GoSam". In: *Eur. Phys. J. C* 72 (2012), p. 1889. DOI: 10.1140/epjc/s10052-012-1889-1. arXiv: 1111.2034 [hep-ph].

[90] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. "MINLO: multi-scale improved NLO". In: *JHEP* 10 (2012), p. 155. DOI: 10.1007/JHEP10(2012)155. arXiv: 1206.3572 [hep-ph].

[91] Gionata Luisoni et al. "$HW^{\pm}/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO". In: *JHEP* 10 (2013), p. 083. DOI: 10.1007/JHEP10(2013)083. arXiv: 1306.2542 [hep-ph].

[92] Lukas Altenkamp et al. "Gluon-induced Higgs-strahlung at next-to-leading order QCD". In: *JHEP* 02 (2013), p. 078. DOI: 10.1007/JHEP02(2013)078. arXiv: 1211.5015 [hep-ph].

[93] Robert V. Harlander et al. "Soft gluon resummation for gluon-induced Higgs Strahlung". In: *JHEP* 11 (2014), p. 082. DOI: 10.1007/JHEP11(2014)082. arXiv: 1410.0217 [hep-ph].

[94] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction". In: *JHEP* 09 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph].

[95] ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. ATL-PHYS-PUB-2014-021. 2014. URL: https://cds.cern.ch/record/1966419.

[96] M. Beneke et al. "Hadronic top-quark pair production with NNLL threshold resummation". In: *Nucl. Phys. B* 855 (2012), pp. 695–741. DOI: 10.1016/j.nuclphysb.2011.10.021. arXiv: 1109.1536 [hep-ph].

[97] Matteo Cacciari et al. "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation". In: *Phys. Lett. B* 710 (2012), pp. 612–622. DOI: 10.1016/j.physletb.2012.03.013. arXiv: 1111.5869 [hep-ph].

[98] Peter Bärnreuther, Michal Czakon, and Alexander Mitov. "Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$". In: *Phys. Rev. Lett.* 109 (2012), p. 132001. DOI: 10.1103/PhysRevLett.109.132001. arXiv: 1204.5201 [hep-ph].

[99] Michal Czakon and Alexander Mitov. "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels". In: *JHEP* 12 (2012), p. 054. DOI: 10.1007/JHEP12(2012)054. arXiv: 1207.0236 [hep-ph].

[100] Michal Czakon and Alexander Mitov. "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction". In: *JHEP* 01 (2013), p. 080. DOI: 10.1007/JHEP01(2013)080. arXiv: 1210.6832 [hep-ph].

[101] Michal Czakon, Paul Fiedler, and Alexander Mitov. "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$". In: *Phys. Rev. Lett.* 110 (2013), p. 252004. DOI: 10.1103/PhysRevLett.110.252004. arXiv: 1303.6254 [hep-ph].

[102] Michal Czakon and Alexander Mitov. "Top++: A program for the calculation of the top-pair cross-section at hadron colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[103] Simone Alioli et al. "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions". In: *JHEP* 09 (2009), p. 111. DOI: 10.1088/1126-6708/2009/09/111. arXiv: 0907.4076 [hep-ph].

[104] M. Aliev et al. "HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR". In: *Comput. Phys. Commun.* 182 (2011), pp. 1034–1046. DOI: 10.1016/j.cpc.2010.12.040. arXiv: 1007.1327 [hep-ph].

[105] P. Kant et al. "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions". In: *Comput. Phys. Commun.* 191 (2015), pp. 74–89. DOI: 10.1016/j.cpc.2015.02.001. arXiv: 1406.4403 [hep-ph].

[106] Emanuele Re. "Single-top Wt-channel production matched with parton showers using the POWHEG method". In: *Eur. Phys. J. C* 71 (2011), p. 1547. DOI: 10.1140/epjc/s10052-011-1547-z. arXiv: 1009.2450 [hep-ph].

[107] Nikolaos Kidonakis. "Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-". In: *Phys. Rev. D* 82 (2010), p. 054018. DOI: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451 [hep-ph].

[108] Nikolaos Kidonakis. "Top Quark Production". In: *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)* (JINR, Dubna, Russia, July 15–28, 2013), pp. 139–168. DOI: 10.3204/DESY-PROC-2013-03/Kidonakis. arXiv: 1311.0283 [hep-ph].

[109] Charalampos Anastasiou et al. "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order". In: *Phys. Rev. D* 69 (2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266.

[110] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017.

[111] ATLAS Collaboration. "Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 81 (2020), p. 689. DOI: 10.1140/epjc/s10052-021-09402-3. arXiv: 2007.02645 [hep-ex].

[112] 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". In: *CERN-2011-002* (2011). DOI: 10.5170/CERN-2011-002. arXiv: 1101.0593 [hep-ph].

[113] 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". In: *CERN-2012-002* (2012). DOI: 10.5170/CERN-2012-002. arXiv: 1201.3084 [hep-ph].

[114] Glen Cowan. *Statistics for Searches at the LHC*. 2013. DOI: 10.48550/ARXIV.1307.2487. URL: https://arxiv.org/abs/1307.2487.

[115] Gilad Perez et al. "Constraining the charm Yukawa and Higgs-quark coupling universality". In: *Phys. Rev. D* 92 (2015), p. 033016. DOI: 10.1103/PhysRevD.92.033016.

[116] MILC A. Bazavov et al. (Fermilab Lattice and TUMQCD Collaborations). "Up-, down-, strange-, charm-, and bottom-quark masses from four-flavor lattice QCD". In: *Phys. Rev. D* 98 (Sept. 2018), p. 054517. DOI: 10.1103/PhysRevD.98.054517. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.054517.

[117] ATLAS Collaboration. *Expected performance of the ATLAS detector under different High-Luminosity LHC conditions*. ATL-PHYS-PUB-2021-023. 2021. URL: https://cds.cern.ch/record/2765851.

[118] ATLAS Collaboration. *Expected performance of the ATLAS detector at the High-Luminosity LHC*. ATL-PHYS-PUB-2019-005. 2019. URL: https://cds.cern.ch/record/2655304.

Appendix A

Post-fit m_{cc} distributions of analysis signal regions

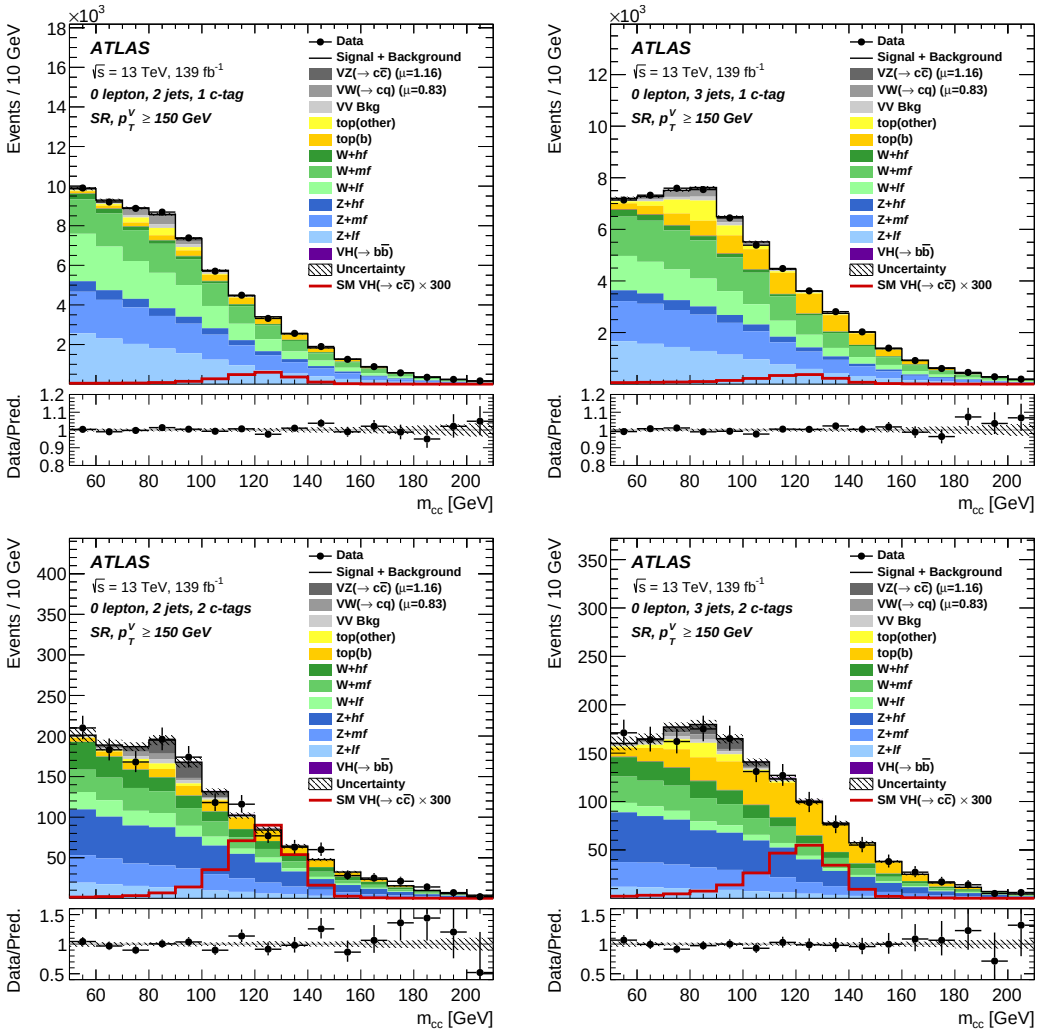


Figure A.1: Post-fit distributions of the four 0-lepton signal regions. The total signal-plus-background prediction is shown by the solid black line and includes the $H \to c\bar{c}$ signal scaled to the best-fit value of $\mu_{VH(c\bar{c})} = -9$. The $H \to c\bar{c}$ signal is also shown as an unfilled histogram scaled to 300 times the SM prediction. The post-fit uncertainty is shown as the hatched background including correlations between uncertainties. The ratio of the data to the sum of the post-fit signal plus background is shown in the lower panel [57].

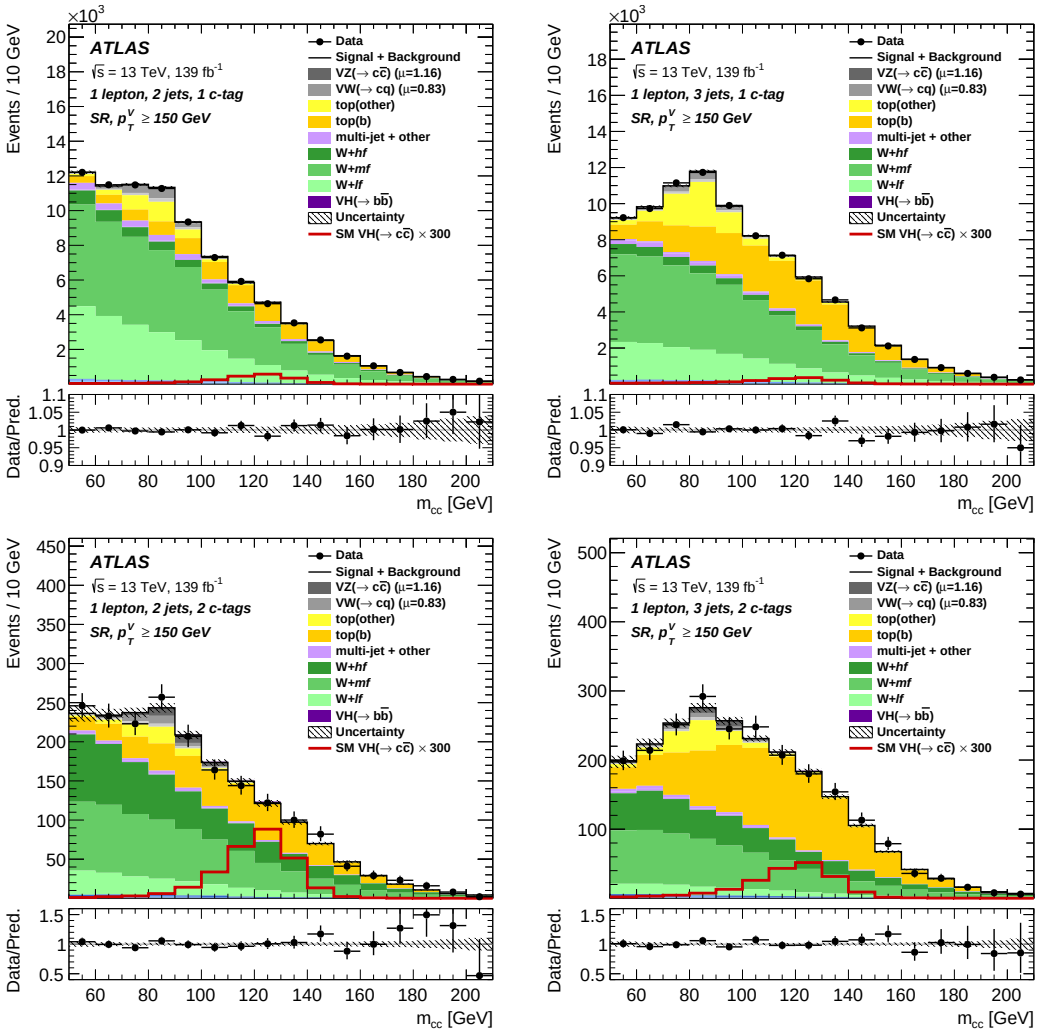


Figure A.2: Post-fit distributions of the four 1-lepton signal regions. The total signal-plus-background prediction is shown by the solid black line and includes the $H \to c\bar{c}$ signal scaled to the best-fit value of $\mu_{VH(c\bar{c})} = -9$. The $H \to c\bar{c}$ signal is also shown as an unfilled histogram scaled to 300 times the SM prediction. The post-fit uncertainty is shown as the hatched background including correlations between uncertainties. The ratio of the data to the sum of the post-fit signal plus background is shown in the lower panel [57].

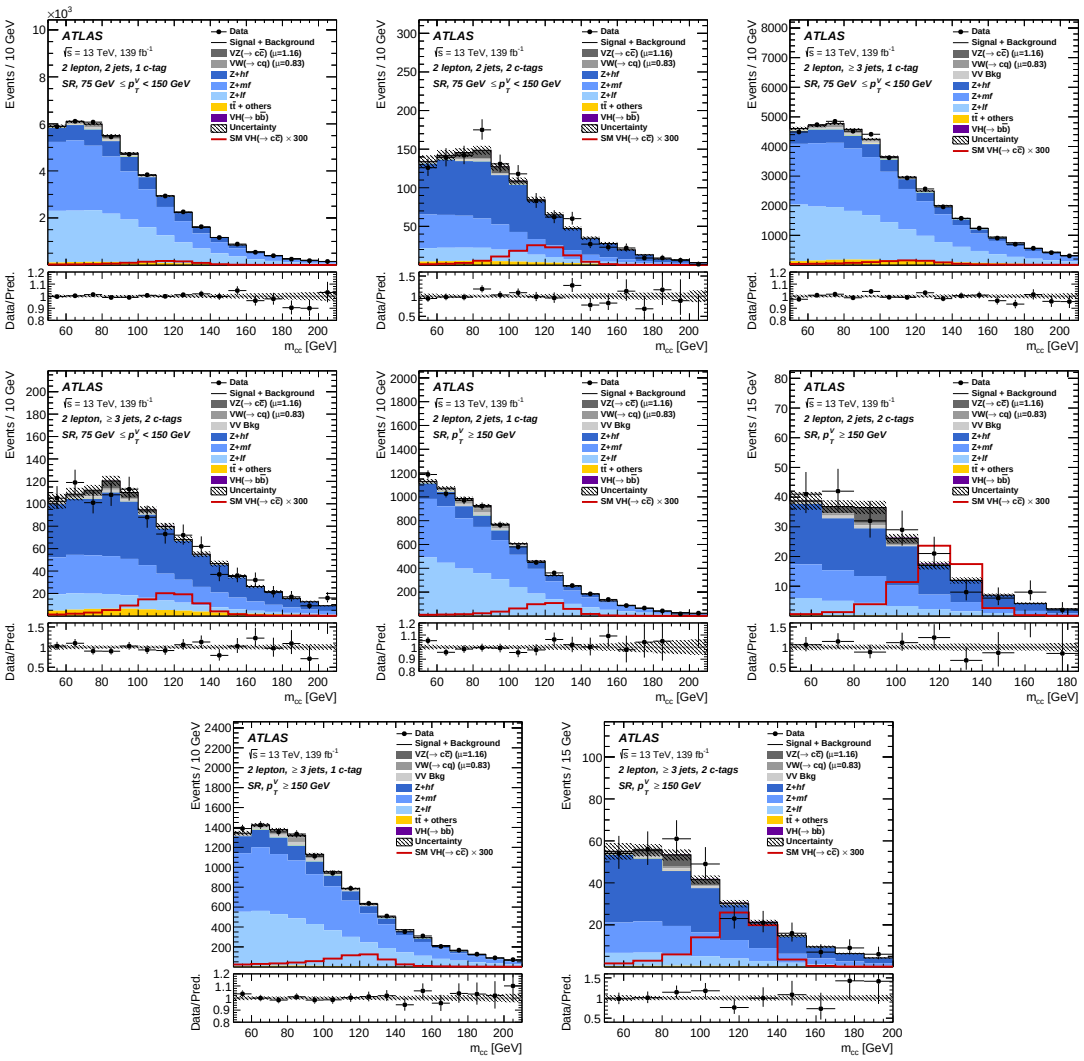


Figure A.3: Post-fit distributions of the eight 2-lepton signal regions. The total signal-plus-background prediction is shown by the solid black line and includes the $H \to c\bar{c}$ signal scaled to the best-fit value of $\mu_{VH(c\bar{c})} = -9$. The $H \to c\bar{c}$ signal is also shown as an unfilled histogram scaled to 300 times the SM prediction. The post-fit uncertainty is shown as the hatched background including correlations between uncertainties. The ratio of the data to the sum of the post-fit signal plus background is shown in the lower panel [57].