Search for the decay of the Higgs boson to charm quarks in p-p collisions at $\sqrt{s} = 13$ TeV in the LHC-ATLAS experiment

This dissertation is submitted in fulfillment of the
requirements for the degree of *Doctor of Philosophy*

Lakmin Vindula Bandara Wickremasinghe

Nanjo Group, Department of Physics,
Graduate School of Science, Osaka University

February 28th, 2025

Abstract

This thesis describes the latest search on the Higgs boson decay to a charm and anti-charm quark pair ($H \to c\bar{c}$), where the Higgs boson is produced in association with a vector boson. The associated production of a Higgs boson with a vector boson gives the best sensitivity to probe the Higgs-charm coupling strength, since the leptonic decay of the vector bosons can be used to trigger this process. The $H \to c\bar{c}$ search is based on the full proton-proton collision dataset recorded by the ATLAS experiment from 2015 to 2018. The analysis utilizes a new jet-flavor tagging algorithm and new analysis techniques to increase the signal sensitivity compared to the previous $H \to c\bar{c}$ search by the ATLAS experiment. The new analysis techniques include a multivariate analysis to discriminate signal and background events, and an event weighting method to effectively use all the simulated events, which helps in reducing the statistical error attributed to simulated samples. In the new analysis, an observed (expected) upper limit at 95% confidence level of 12 (11) times the standard model prediction on the $H \to c\bar{c}$ signal strength is obtained. The expected upper limit is a factor of three times improvement compared to the previous ATLAS result.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The discovery of the Higgs boson by the ATLAS and CMS experiments in 2012 provided the experimental confirmation of the Higgs mechanism in the standard model of particle physics. In the standard model, the coupling between the Higgs boson to fermions manifest as Yukawa interactions between the Higgs field and the fermion fields. While the Yukawa coupling to top and bottom quarks have been observed by the ATLAS experiment [1], out of the remaining four quarks (namely up, down, strange and charm), the Yukawa coupling to charm quarks remains the only coupling that could be observed at the Large Hadron Collider (LHC).

This chapter first provides a brief description on the theoretical background of the Higgs mechanism in the standard model, and the physics at the LHC. Then the Higgs boson Yukawa coupling measurements are discussed, with a focus on the Higgs boson to charm quark decay search at the LHC. Finally, an overview is given on the latest Higgs boson to charm quark decay analysis described in this thesis, including the personal contributions of the author.

1.1 Theoretical introduction

The standard model of particle physics is a gauge theory of the symmetry group $\mathrm{U}(1)_\mathrm{Y} \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{SU}(3)$[1], where the symmetry group $\mathrm{U}(1)_\mathrm{Y} \times \mathrm{SU}(2)_\mathrm{L}$ acts on the electroweak gauge theory, and SU(3) represents the symmetry group for the gauge theory of quantum chromodynamics (QCD). The symmetry group $\mathrm{U}(1)_\mathrm{Y} \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{SU}(3)$ has a total of twelve gauge bosons as depicted in Figure 1.1; the photon, three weak bosons, and eight gluons.

In the standard model of particle physics, the matter in the universe are composed of the spin-half elementary particles known as fermions. Fermions are further classified into quarks and leptons, and each particle has an anti-matter counter part. In total there are 24 fermions in the standard model, as depicted in Figure 1.1. Further description on the standard model and the Higgs mechanism[2] can be found in Appendix A.

---

[1] U stands for unitary, and ”1” is the 1-dimensional representation. SU stands for special unitary, and ”2” and ”3” denotes the two-dimensional and three-dimensional representations, respectively.

[2] Section 1.1.1 and Appendix A is mainly compiled by referring to [2] and [3].

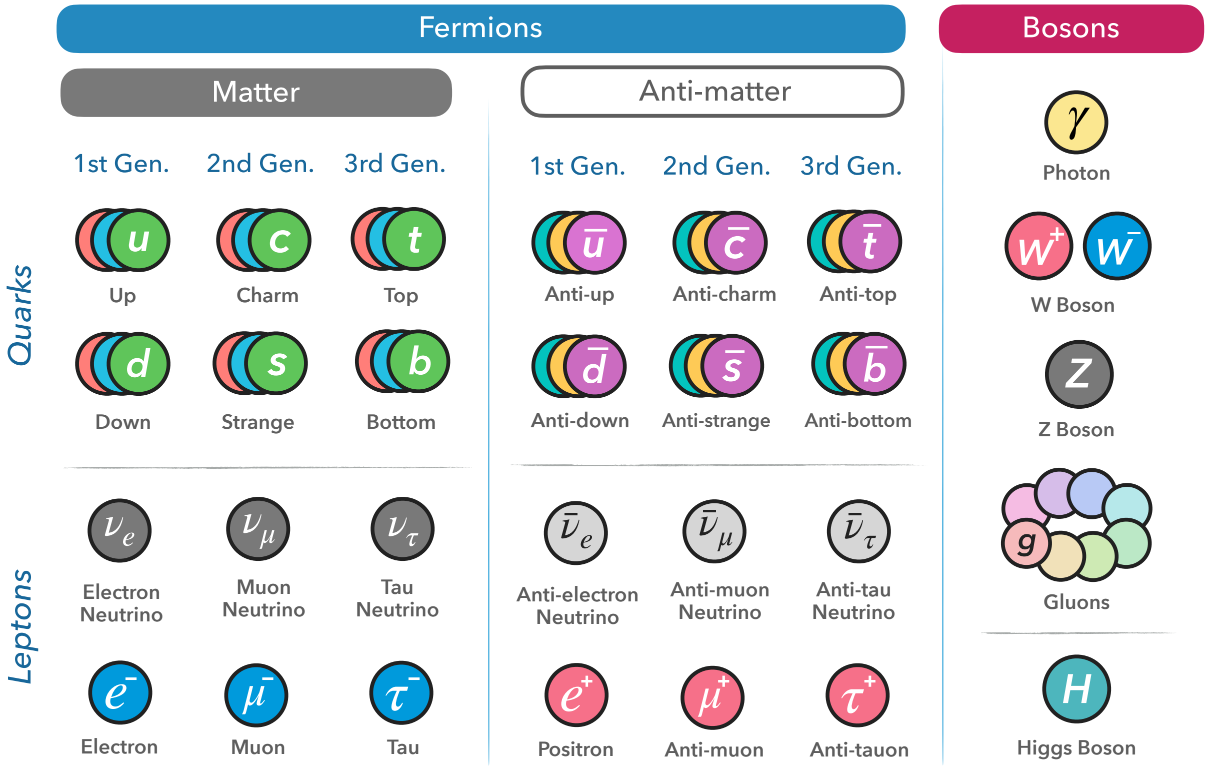


Figure 1.1: The standard model particles.

1.1.1 The Higgs mechanism

The Higgs mechanism is the spontaneous symmetry breaking of the $SU(2)_L \times U(1)_Y$ local gauge symmetry of the electroweak theory, generating the mass of the weak bosons. After symmetry breaking, the electromagnetic symmetry should be left as the true symmetry at low energy. Hence, the gauge theory should demonstrate the behavior

$$SU(2)_L \times U(1)_Y \xrightarrow[\text{symmetry breaking}]{\text{spontaneous}} U(1)_{QED}$$

The minimal Higgs model contains two complex scalar fields, placed in a weak isospin doublet[3]; one of the scalar fields must be neutral (ϕ^0) in order to generate the masses of the electroweak gauge bosons, while the other scalar field must be charged (ϕ^+) to give the longitudinal degrees of freedom of the W^+ and W^- bosons.

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.1}$$

The Lagrangian for this doublet is

$$\mathcal{L} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \tag{1.2}$$

where $V(\phi)$ is the Higgs potential given by

---

[3] The upper and lower components of the doublet differ by one unit of charge.

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2, \tag{1.3}$$

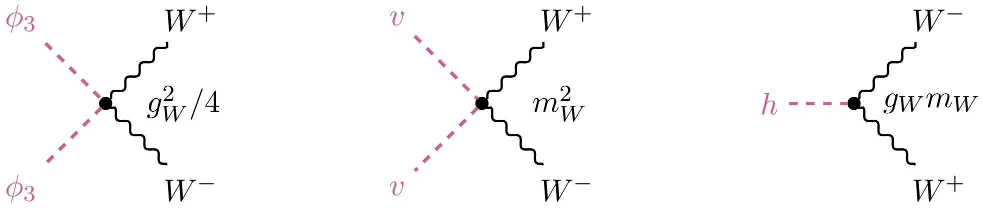
and $D_\mu = \partial_\mu + i\frac{g_W}{2}\boldsymbol{\sigma}\cdot\mathbf{W}_\mu + i\frac{g'}{2}B_\mu$ is an appropriate covariant derivative that respects the $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ local gauge symmetry of the electroweak model. $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ represents the three Pauli matrices, and linear combinations of the fields $\mathbf{W}_\mu = \left(W_\mu^{(1)}, W_\mu^{(2)}, W_\mu^{(3)}\right)$ and B_μ form the gauge fields attributed to the weak bosons and the photon. The relevant ϕ_3 field term of $(D_\mu\phi)^\dagger(D^\mu\phi)$ can be expressed as,

$$\begin{aligned}(\partial_\mu\phi_3)(\partial^\mu\phi_3) + \frac{1}{4}g_W^2\left(W_\mu^{(1)} + iW_\mu^{(2)}\right)\left(W^{(1)\mu} - iW^{(2)\mu}\right)\phi_3^2 \\ + \frac{1}{4}\left(g_W W_\mu^{(3)} - g'B_\mu\right)\left(g_W W^{(3)\mu} - g'B^\mu\right)\phi_3^2.\end{aligned} \tag{1.4}$$

The first term in Equation 1.4 corresponds to the kinetic term of the scalar, and the second and third terms correspond to the interaction terms of the gauge bosons. The second term of Equation 1.4 is expressed using the W^+ and W^- fields as

$$\mathcal{L} = \frac{1}{4}g_W^2 W^+W^-\phi_3^2. \tag{1.5}$$

The scalar field ϕ_3 originally couples to the gauge boson fields W^+ and W^- as depicted in Figure 1.2(a).



(a) $W^+W^-\phi_3\phi_3$ interaction (b) W^+W^-vv interaction (c) W^+W^-h interaction

Figure 1.2: The interaction between the scalar fields with the gauge boson fields W^+ and W^-.

After spontaneous symmetry breaking[4], the field ϕ_3 is taken as $\phi_3 \to v + h$, where h is a physical Higgs boson arising from a small perturbation from the vacuum expectation value v. With this, Equation 1.5 can be expanded as

$$\mathcal{L} = \frac{1}{4}g_W^2 v^2 W^+W^- + \frac{1}{2}g_W^2 vhW^+W^- + \frac{1}{4}g_W^2 h^2 W^+W^-. \tag{1.6}$$

The first term (depicted in Figure 1.2(b)) can be interpreted as the mass term the W boson, with $m_W = \frac{1}{2}g_W v$. Hence, Equation 1.6 can be expressed as

$$\mathcal{L} = m_W^2 W^+W^- + g_W m_W hW^+W^- + \frac{1}{4}g_W^2 h^2 W^+W^-, \tag{1.7}$$

---

[4] The fields ϕ_1, ϕ_2, and ϕ_4 are chosen to be zero after spontaneous symmetry breaking.

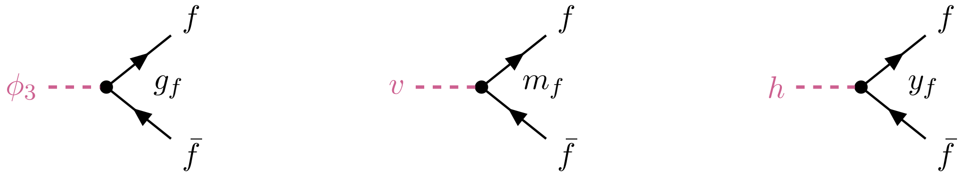
where the second term (depicted in Figure 1.2(c)) represents the coupling of the Higgs field to W boson fields, with a coupling strength of $g_{hWW} = g_W m_W = \frac{2m_W^2}{v}$. The third term represents the quartic coupling of the Higgs field and the W boson field. Similarly, the coupling between the Higgs field and the Z boson field is obtained as $g_{hZZ} = g_Z m_Z$.

1.1.2 The Higgs boson coupling to fermions

The Higgs mechanism can also be extended to explain the mechanism that generates the mass of fermions. Similar to the Lagrangian term of Equation 1.5, the coupling between the fermion field and the scalar field ϕ_3 (depicted in Figure 1.3(a)) can be expressed as

$$\mathcal{L} = \frac{1}{\sqrt{2}} g_f \phi_3 f \bar{f}, \tag{1.8}$$

where f represents the fermion field, and $\bar{f}$ represents the adjoint fermion field.



(**a**) $f\bar{f}\phi_3$ interaction (**b**) $f\bar{f}v$ interaction (**c**) $f\bar{f}h$ interaction

Figure 1.3: The interaction between the scalar fields with the fermion boson fields.

After spontaneous symmetry breaking, Equation 1.8 can be expanded as

$$\mathcal{L} = \frac{1}{\sqrt{2}} g_f v f \bar{f} + \frac{1}{\sqrt{2}} g_f h f \bar{f}. \tag{1.9}$$

The first term, depicted in Figure 1.3(b), can be interpreted as the mass term of the fermions, and the second term, depicted in Figure 1.3(c), is the coupling term between fermions and the Higgs field. Letting the fermion mass $m_f = \frac{1}{\sqrt{2}} g_f v$, and $y_f = \frac{m_f}{v}$, Equation 1.9 can be written as,

$$\mathcal{L} = m_f f \bar{f} + y_f f \bar{f}, \tag{1.10}$$

where y_f is referred to as the Yukawa coupling strength between the Higgs field and the fermion field f.

1.2 Physics at the Large Hadron Collider

The Large Hadron Collider (LHC) collides protons with protons, and the hard scattering process occurs between the partons (quarks or gluons) in the proton. Although the collisions occur on the center-of-mass frame of the colliding protons, this is not the same as the center-of-mass of the colliding partons, since partons carry a certain momentum fraction x of the protons. Hence, post collision objects are boosted by a net momentum of $(x_1 - x_2)\sqrt{s}/2$ along the beam axis, where x_1, x_2 are the momentum fractions of the two colliding partons (respect to the parent proton), and $\sqrt{s}$ is the center of mass energy of the colliding proton-proton system. With respect to the momentum in the beam direction, the colliding partons have negligible momentum transverse to the beam axis. Therefore, the transverse momentum (p_T) of post collision objects are taken to be conserved.

Parton distribution functions (discussed in Section 3.2) describe the probability of partons inside a proton participating in a hard scattering interaction, as a function of the parton momentum fraction x. In the low x region, the momentum fraction x of gluons and sea quarks is more compared to valence quarks, and this fraction increases with the increase in momentum transfer Q^2. Therefore, processes involving gluons as the colliding partons are dominant at the LHC.

1.2.1 Standard model physics processes

The ATLAS experiment is able to measure different standard model processes from proton-proton collisions at the Large Hadron Collider, and Figure 1.4 shows a summary of the production cross-section measurements of different physics processes as a function of the center-of-mass energy $\sqrt{s}$.

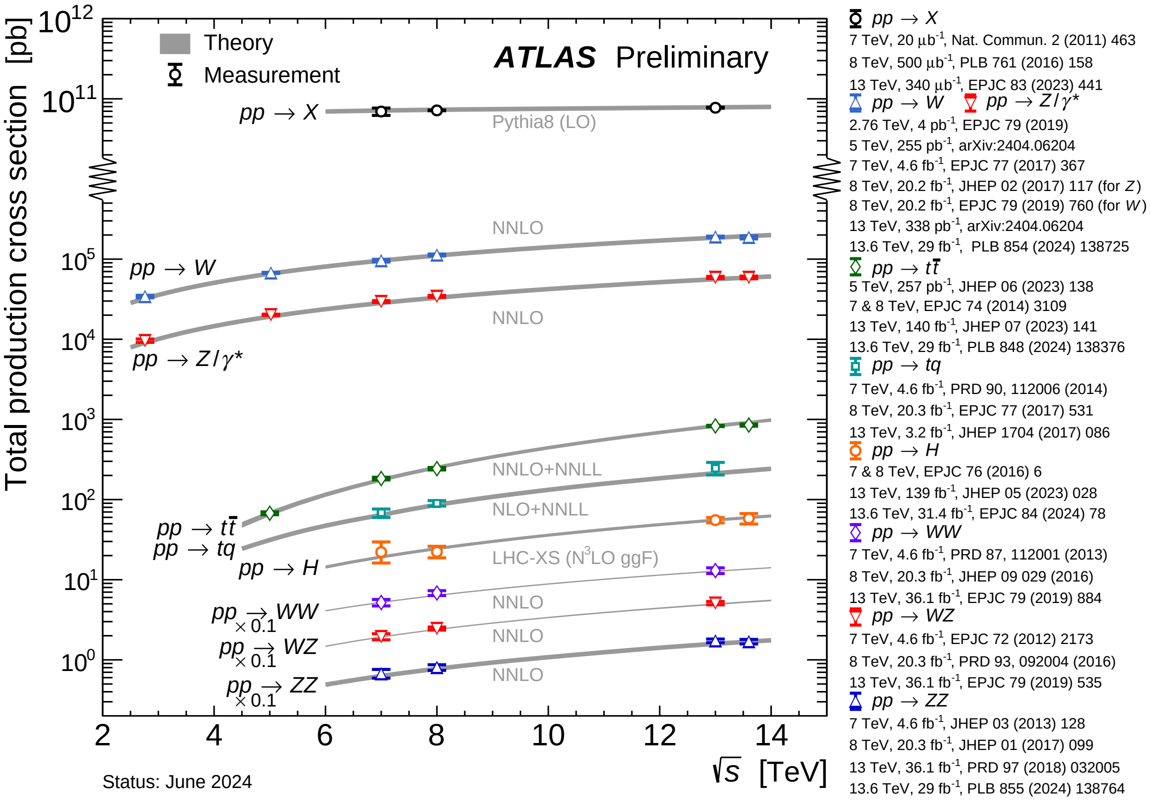


Figure 1.4: A summary of the ATLAS measurements on total production cross-section measurements of standard model processes as a function of the center-of-mass energy $\sqrt{s}$ [4].

$pp \rightarrow X$ in Figure 1.4 denotes the total cross section from proton-proton collisions, where roughly 80% of the cross section is due to proton-proton inelastic scattering [5]. A large fraction of the $pp \to X$ inelastic scattering cross section is from the QCD multi-jet processes[5] [6].

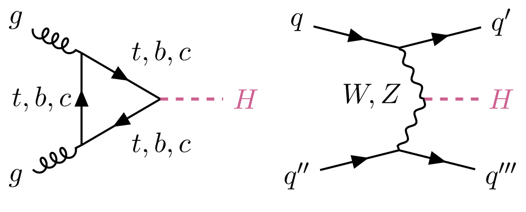
A vector boson with jets, also referred to as the V+jets process, is denoted by $pp \to W$ and $pp \to Z$ in Figure 1.4, and is six orders smaller in cross-section compared to $pp \to X$ [7]. The next dominant physics processes are the top-quark pair production (referred to as the $t\bar{t}$ process) and the single top-quark production (referred to as the single-top process), denoted by $pp \to t\bar{t}$ and $pp \to tq$ in Figure 1.4, which is roughly two orders of magnitude smaller in cross-section than V+jets production [8, 9, 10, 11].

The production cross-section for di-boson $pp \to WW$ process is roughly an order smaller compared to $pp \to t\bar{t}$, while the di-boson processes $pp \to WZ$ ($pp \to ZZ$) has roughly half (an order smaller) cross section compared with $pp \to WW$ [12, 13, 14]. Finally, the Higgs boson production cross sections are at a similar magnitude compared to the $pp \to WZ$ cross section, and is denoted by $pp \to H$ in Figure 1.4 [15].

More details on the standard model processes relevant to this thesis are discussed in Section 3.2, and Appendix B lists additional details on the cross section measurements of the above mentioned processes.

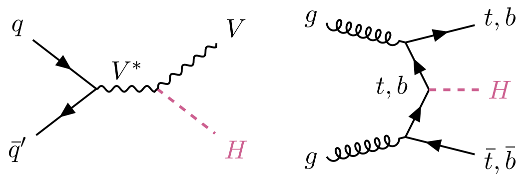
1.2.2 Higgs boson production

The dominant processes of Higgs boson production at the LHC are illustrated in Figures 1.5(a)∼(d), and Figure 1.5(e) shows the production cross section for the different processes at the LHC, as a function of the center-of-mass energy $\sqrt{s}$ of proton-proton collisions.



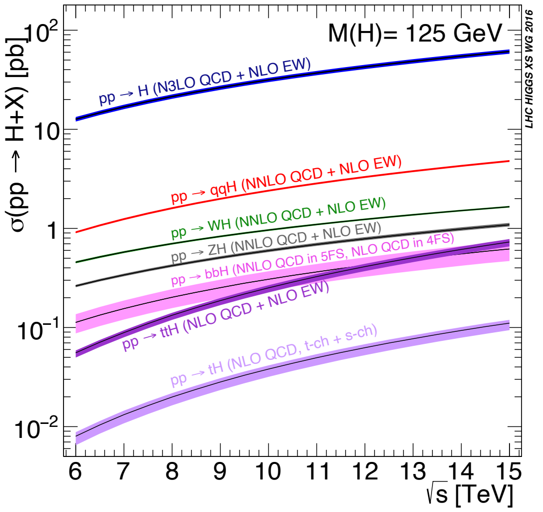
(a) ggF channel

(b) VBF channel



(c) VH channel

(d) ttH and bbH channels



(e) Higgs boson production cross section.

Figure 1.5: The leading order Feynman diagrams for the dominant production modes of a single Higgs boson at the large hadron collider, and the predicted cross section (based on theoretical calculations) for these processes as a function of the center-of-mass energy $\sqrt{s}$ of proton-proton collisions. The Higgs boson mass is assumed to be 125 GeV in these predictions [16]

$pp \to H$ in Figure 1.5(e) denotes the gluon-gluon fusion (ggF) channel, which is the dominant

---

[5] The QCD partons created in physics processes from proton-proton collisions will be observed as a jet of particles in the detector, due to fragmentation and hadronization.

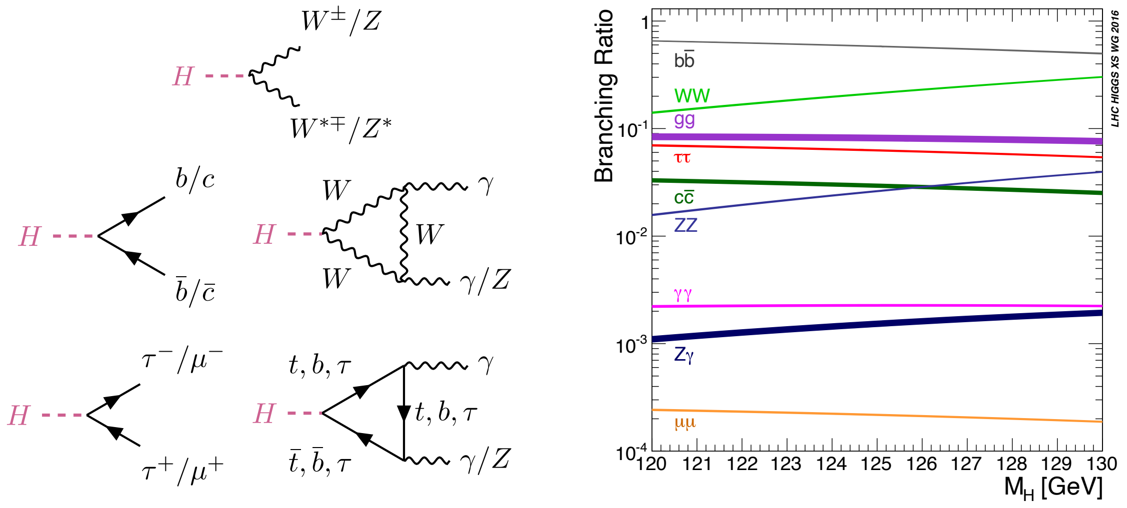
production mode for Higgs bosons at the Large Hadron Collider. The next dominant production mode for Higgs bosons is the through the vector boson fusion (VBF) channel, denoted by $pp \to qqH$ in Figure 1.5(e), which is roughly 13 times smaller in cross section compared to the the ggF channel. In the same order of magnitude in cross-section is the associated production of a Higgs boson with a vector boson (VH channel, also known as Higgstrahlung), denoted by $pp \to WH$ and $pp \to ZH$ in Figure 1.5(e). Finally, in a similar order of magnitude to $pp \to ZH$ is the Higgs boson production in association with top and bottom quarks (ttH and bbH channels), denoted by $pp \to bbH$ and $pp \to ttH$ in Figure 1.5(e).

Table 1.1 details the cross sections of the dominant Higgs boson production modes at a center-of-mass energy $\sqrt{s} = 13$ TeV, and the Higgs boson mass $M_H = 125$ GeV.

Table 1.1: The production cross sections based on theoretical calculations for a Higgs boson at mass $M_H = 125$ GeV, in pp collisions at a center-of-mass energy of 13 TeV. The uncertainties shown are the total of theoretical, PDF and α_s uncertainties [16, 17]

Production process	Cross section σ	Calculation order
ggF channel	$48.58\ ^{+2.72}_{-3.59}$ pb	N^{3}LO QCD + NLO EW
VBF channel	$3.78\ ^{+0.08}_{-0.08}$ pb	NNLO QCD + NLO EW
WH channel	$1.37\ ^{+0.03}_{-0.03}$ pb	NNLO QCD + NLO EW
ZH channel	$0.88\ ^{+0.04}_{-0.03}$ pb	NNLO QCD + NLO EW
ttH channel	$0.50\ ^{+0.03}_{-0.05}$ pb	NLO QCD + NLO EW
bbH channel	$0.49\ ^{+0.10}_{-0.12}$ pb	NNLO QCD

1.2.3 Higgs boson decay



(a) The dominant Higgs boson decay modes.

(b) The Higgs boson branching fraction.

Figure 1.6: The Higgs boson decay processes with the largest branching fraction, and the branching fraction as a function of the Higgs boson mass [16, 18].

Once a Higgs boson is produced, it will decay almost instantly through the dominant decay channels shown in Figure 1.6(a). Figure 1.6(b) shows the branching fractions as a function of the Higgs boson mass in the range 120 GeV to 130 GeV.

In the mass range shown in Figure 1.6(b), the dominant decay mode is the Higgs boson decay to a pair of bottom quarks, followed by the decay to a pair of W bosons. The decay to τ-leptons has the largest branching fraction within the lepton family. The Higgs boson decay to charm quarks which is the focus of this thesis, has the second largest branching fraction between quarks.

Table 1.2 shows the branching fraction and the relative uncertainty for the Higgs boson decay modes. The eight modes shown in the table account to almost 92 % of the Higgs boson decays.

Table 1.2: The branching fractions (based on theoretical calculations) and the relative uncertainty for the decay channels of a Higgs boson of mass $M_H = 125$ GeV [16, 17, 18].

Decay channel	Branching fraction
$H \to b\bar{b}$	$58.2^{+1.2}_{-1.3}\%$
$H \to W^{\pm}W^{*\mp}$	21.4% ±1.5%
$H \to \tau^+\tau^-$	6.27% ±1.6%
$H \to c\bar{c}$	$2.89^{+5.5}_{-2.0}\%$
$H \to ZZ^*$	2.62% ±1.5%
$H \to \gamma\gamma$	0.227% ±2.1%
$H \to Z\gamma$	0.153% ±5.8%
$H \to \mu^+\mu^-$	0.0218% ±1.7%

1.2.4 The kappa framework

When measuring the strength of gauge coupling to vector bosons, and the Yukawa coupling to fermions, the production cross sections and branching fractions cannot be treated independently since observed processes involves the effects of both. Therefore, the κ-framework (kappa framework) is introduced for a consistent treatment of coupling strengths in both production and decay. The aim of this framework is to determine if resonances at 125 GeV indeed match the properties of the SM Higgs boson, or to establish deviations from the SM behavior, which rules out the SM being sufficiently significant.

In the κ-framework, cross section times branching fraction $\sigma_i \times B_f$ for Higgs boson production and decay process $i \to H \to f$ is parameterized using a set of parameters κ (known as *coupling strength modifiers*) [18];

$$\sigma_i \times B_f = \frac{\sigma_i(\boldsymbol{\kappa}) \times \Gamma_f(\boldsymbol{\kappa})}{\Gamma_H(\boldsymbol{\kappa})}, \tag{1.11}$$

where Γ_H is the total width of the Higgs boson decay, and Γ_f is the partial width for the Higgs boson decay to final state f. The κ parameters affect the Higgs boson coupling strengths, without altering the kinematic distributions of the given process. The coupling strength modifiers

κ_i for a Higgs boson production process $i \to H$, and κ_f for a Higgs boson decay process $H \to f$ is given by[6]

$$\kappa_i^2 = \frac{\sigma_i}{\sigma_i^{\rm SM}}, \qquad \kappa_f^2 = \frac{\Gamma_f}{\Gamma_f^{\rm SM}}. \tag{1.12}$$

The κ parametrization takes into account the total Higgs boson decay width Γ_H depends on all the decay modes in the present measurements, currently undetectable (to gluons, light quarks or neutrinos) or invisible decays, as well as decays into non-standard model particles. Hence, the Higgs boson total decay width is expressed using the κ-framework as

$$\Gamma_H(\boldsymbol{\kappa}, B_{\rm inv}, B_{\rm undect}) = \kappa_H^2(\boldsymbol{\kappa}, B_{\rm inv}, B_{\rm undect})\Gamma_H^{\rm SM}, \tag{1.13}$$

where

$$\kappa_H^2(\boldsymbol{\kappa}, B_{\rm inv}, B_{\rm undect}) = \frac{\Sigma_j B_f^{\rm SM}\kappa_j^2}{(1 - B_{\rm inv} - B_{\rm undect})}. \tag{1.14}$$

In the above equations, $B_{\rm inv}$ and $B_{\rm undect}$ denote the branching fractions to beyond the standard model (BSM) states, and undetectable states respectively. When determining the coupling modifiers of the top (κ_t), bottom (κ_b), and charm quarks (κ_c), the $\overline{MS}$ running mass evaluated using lattice QCD at a renormalization scale equal to the Higgs boson mass of 125 GeV is used [19].

1.3 Higgs boson Yukawa coupling measurements

The discovery of a particle consistent with the standard model Higgs boson was announced by the LHC ATLAS and CMS experiments in 2012 [21, 22], based on the data taken during the Run 1 period of the LHC from 2011 to 2012, with the center-of-mass energies of 7 TeV in 2011, and 8 TeV in 2012. Subsequent measurements using the Run 1 data showed evidence the particle has spin zero, and it is a CP-even quantum state [23], confirming the discovered particle is the Higgs boson predicted by the standard model.

During the Run 2 period (from 2015 and 2018) of the LHC where the center-of-mass energies increased to 13 TeV, more precise measurements on the Higgs boson properties were done owing to the large dataset equivalent to almost 30 times more Higgs bosons compared to Run 1. The increase in data is partially due to the increase in the center-of-mass energy from 8 TeV to 13 TeV, that also increased the Higgs boson production cross section as seen in Figure 1.5(b). Figure 1.7 shows the measurements during the Run 2 period, of the gauge couplings and Yukawa couplings between the Higgs boson and the standard model particles expressed using the κ-framework. All measurements are consistent with the standard model within uncertainties.

---

[6] Note that for the ggF and VBF production channels explained in Section 1.2.2, $\kappa_{\rm g}^2$ and $\kappa_{\rm VBF}^2$ is expressed as a function of the corresponding quark or vector boson couplings.

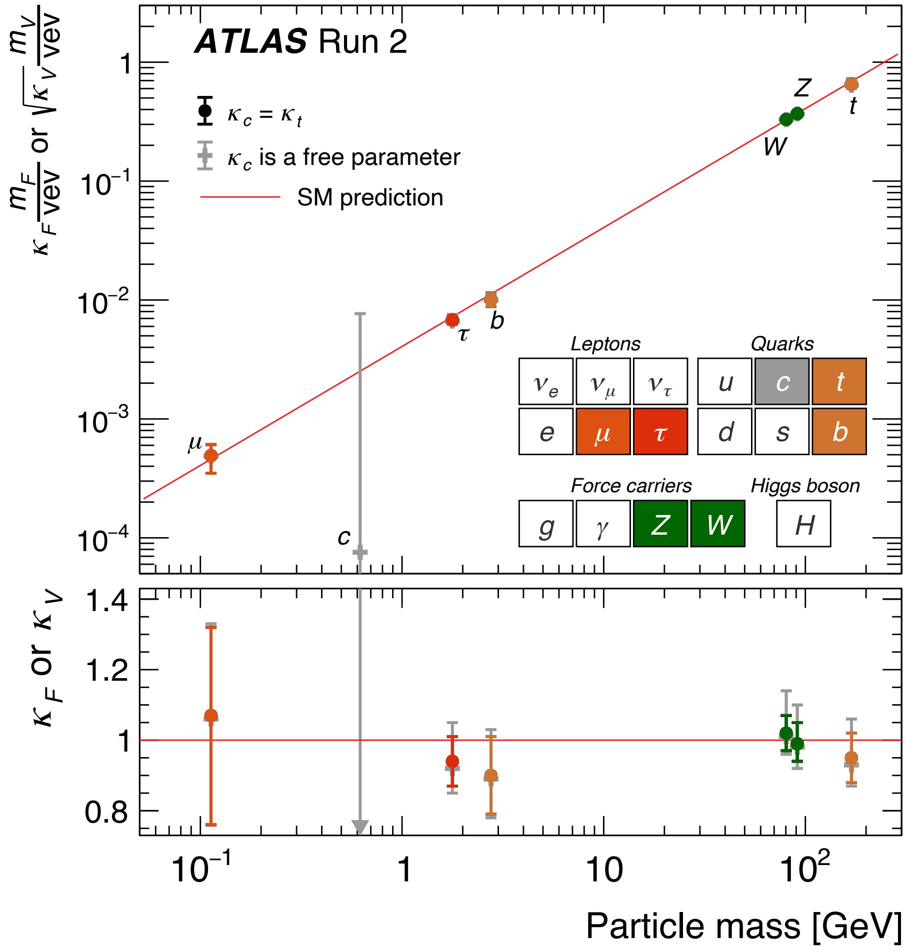


Figure 1.7: The measurements by the ATLAS experiment on the coupling strength between the Higgs boson and the standard model particles [20]. The couplings are interpreted using the κ framework [18].

1.3.1 Search for the Higgs boson decay to charm quarks

The Higgs boson decay to a pair of charm quarks is still unobserved at the LHC due to the challenges in measuring it. However, since $H \to c\bar{c}$ has a sizable contribution to the total Higgs boson decay width Γ_H at the current precision, an accurate determination of the $H \to c\bar{c}$ coupling strength is important.

The ATLAS experiment has probed the Higgs-charm Yukawa coupling using the Run 2 period data in several approaches. One approach is through the radiative rare decay $H \to J/\psi\gamma$ (Figure 1.8(a)), which has a theoretical branching fraction of $2.99^{+0.16}_{-0.15} \times 10^{-6}$, and is a clean probe of the Higgs-charm Yukawa coupling owing to the better signal to background separation [24]. No significant events were observed above the standard model backgrounds, and a 95% confidence level upper limit of 3.5×10^{-4} on the branching fraction was obtained. Another approach is through the charm quark initiated Higgs boson production, and the associated production of charm quarks and a Higgs boson decaying to a pair of photons $(gc \to Hc, H \to \gamma\gamma)$, shown in Figure 1.8(b) [25]. Since no significant excess above the standard model background was observed, a 95% confidence level observed (expected)[7] upper limit on the cross section was obtained at 10.4 pb (8.6 pb). The CMS experiment has also done similar searches in the $H \to J/\psi\gamma$ and $gc \to Hc, H \to \gamma\gamma$ channels [27, 28].

(a) The $H \to J/\psi\,\gamma$ process

(b) The $gc \to Hc$ process

Figure 1.8: The approaches taken to probe the Higgs-charm Yukawa coupling strength.

An alternative approach to probe the Higgs boson decay to charm quarks is through the inclusive $H \to c\bar{c}$ decay process (of Figure 1.6(a)) that has a branching fraction of 2.89%. However, this approach is also challenging due to the difficulty in separating the $H \to c\bar{c}$ signal events from the overwhelming background processes in the Large Hadron Collider. Results in this approach are usually expressed by 95% CL upper limits on the signal strength, defined as

$$\mu_{i\to H,H\to c\bar{c}} \equiv \frac{(\sigma_{i\to H} \times BR_{H\to c\bar{c}})_{\text{Measured}}}{(\sigma_{i\to H} \times BR_{H\to c\bar{c}})_{\text{SM expectation}}}. \tag{1.15}$$

The CMS experiment searched for the $H \to c\bar{c}$ decay using the ggF production channel (of Figure 1.5(a)) which achieved an observed (expected) upper limit on the signal strength of 47 (39) times the standard model prediction on the signal strength, at 95% confidence level [29].

Currently, the most sensitive production channel for searching for the $H \to c\bar{c}$ mode is the VH production channel, since the leptons from the vector boson decay can be used to select this specific signal and reject other background processes. Using this approach, the ATLAS experiment obtained an observed (expected) upper limit of 24 (31) times the standard model on the $VH, H \to c\bar{c}$ process [30]. This search for the $VH, H \to c\bar{c}$ channel was based on the full Run 2 proton-proton collision dataset of 139 fb^{-1}, and obtained an observed (expected) 95% confidence level limits on κ_c to be $|\kappa_c| < 8.5$ (12.4) times the standard model. The CMS experiment also released their latest result, based on a Run 2 dataset of 138 fb^{-1}, for the same $VH, H \to c\bar{c}$ process [31]. They achieved an observed (expected) upper limit of 14 (7.6) times the standard model, a factor of four times better expected limit than the ATLAS measurement, and a factor of five times better than previous CMS result obtained using 35.9 fb^{-1} of proton-proton collision data [32]. This new CMS measurement obtained an observed (expected) constrain on κ_c to be $1.1 < |\kappa_c| < 5.5$ ($|\kappa_c| < 3.4$).

Furthermore, the measurement by the ATLAS experiment shown in Figure 1.7 combined the results of the ATLAS $VH, H \to c\bar{c}$ analysis with direct measurements on other coupling strength measurements, and obtained an observed (expected) upper limit of $|\kappa_c| < 5.7(7.6)$ [20].

---

[7] Expected upper limits are obtained through a fit to a theoretically built data representation assuming the standard model, while observed upper limits are obtained through a fit to the actual data [26].

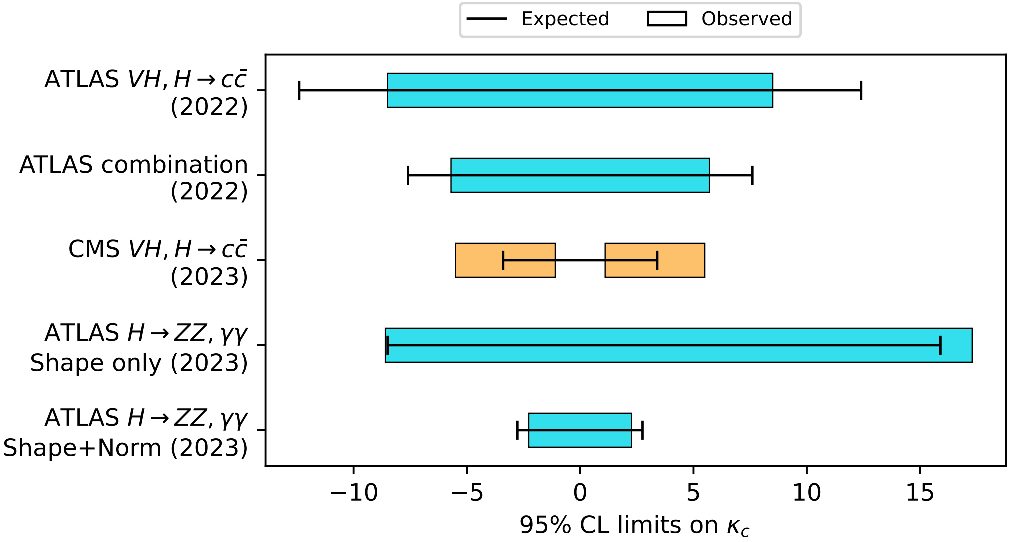


Figure 1.9: A comparison of the constraints on κ_c from several latest measurements. The *ATLAS* $VH, H \to c\bar{c}$, *CMS* $VH, H \to c\bar{c}$, *ATLAS combination*, and *ATLAS* $H \to ZZ, \gamma\gamma$ are the analyses described in [30], [31], [20], and [15], respectively.

While the constraints on κ_c mentioned above were obtained from direct $H \to c\bar{c}$ searches, it is also possible to obtain indirect constraints on κ_c through modifications to the Higgs boson production cross sections. An analysis by the ATLAS experiment, targeting the $H \to ZZ^* \to 4l$ and $H \to \gamma\gamma$ final states, probed κ_c in the ggF production channel with a charm quark loop, and the charm quark induced production $c\bar{c} \to H$. Since deviations from $\kappa_c = 1$ is expected to modify the transverse mass of the Higgs boson (p_{T}^H) distributions at low p_{T}^H [33], κ_c was probed in the fiducial cross section measurement as a function of p_{T}^H [15].

Considering only modifications to the shape of the measured p_{T}^H distributions by varying κ_c (referred to as *shape only*), the analysis obtained an observed (expected) constrain on κ_c as $-8.6 < \kappa_c < 17.3$ ($-8.5 < \kappa_c < 15.9$), which has a weaker constraint compared to the constraint obtained from the direct searches (i.e. the $VH, H \to c\bar{c}$ process). However, considering modifications to the total production cross-sections, branching ratios, and the Higgs boson decay width (referred to as *shape* + *norm*), the analysis obtained an observed (expected) constrain on κ_c as $|\kappa_c| < 2.27$ ($-2.77 < \kappa_c < 2.75$), the most stringent constraints on κ_c to date. Furthermore, a combination of the $H \to ZZ^* \to 4l$ and $H \to \gamma\gamma$ fiducial cross section measurements together with the measurements of the ATLAS $VH, H \to c\bar{c}$ analysis, obtained a 95% CL interval of $-2.47 < \kappa_c < 2.53$.

Since the indirect measurements (described above) and the direct measurements in the $VH, H \to c\bar{c}$ process are complementary, it is also important to probe κ_c directly in the $VH, H \to c\bar{c}$ process. The ATLAS experiment re-analyzed the $VH, H \to c\bar{c}$ process using the Run 2 dataset with new c-tagging algorithms (generally referred to as *jet-flavor tagging* algorithms), and other analysis techniques described in Section 1.4.2. The results of this latest analysis is presented in this thesis.

1.4 Overview of the new $VH, H \to c\bar{c}$ analysis

The new $VH, H \to c\bar{c}$ analysis is a part of the combined $VH, H \to b\bar{b}/c\bar{c}$ analysis that simultaneously searches for the Higgs boson to bottom and charm quark decay processes, where the Higgs boson is produced associated with a vector boson (i.e the WH and ZH channels). The $VH, H \to b\bar{b}/c\bar{c}$ analysis is a re-analysis of the previous three independent Run 2 analyses for the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ processes, that used a proton-proton collision dataset of 139 fb^{-1}[34, 35, 30]. Figure 1.10 shows the overview of the combined three analyses; the $VH, H \to b\bar{b}$ analysis was originally split into two independent analyses using the p_{T} of the vector boson, based on the topology of the Higgs boson decay.

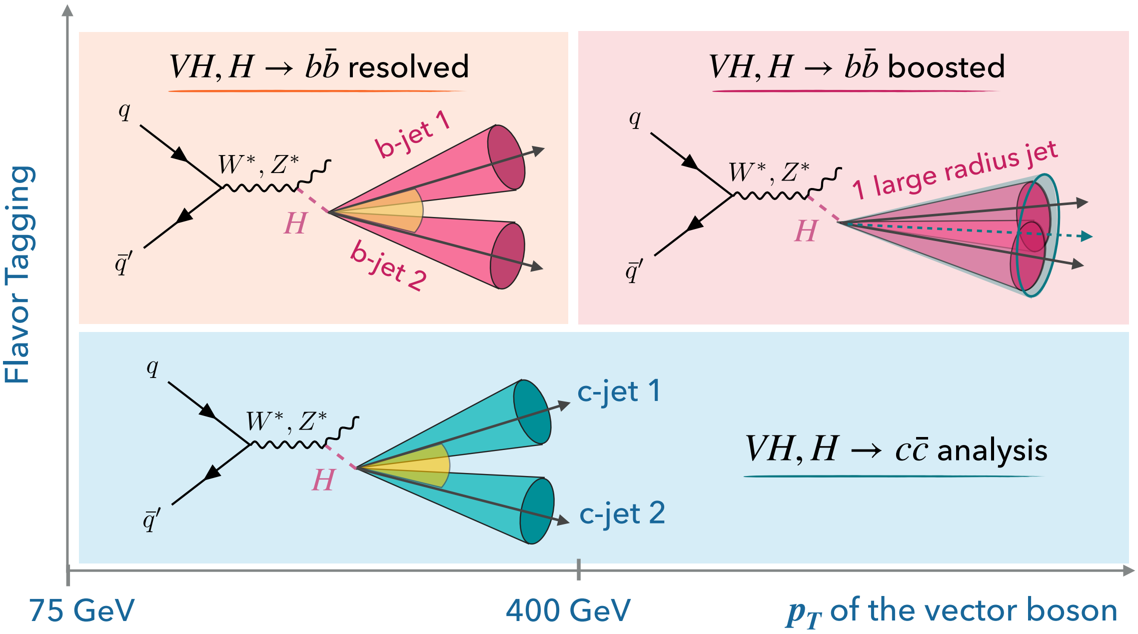


Figure 1.10: The overview of the $VH, H \to b\bar{b}/c\bar{c}$ analysis.

A motivation for the combination of the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ analyses is to simultaneously measure the signal strength μ of the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ processes, as defined in Equation 1.16 and 1.17. The simultaneous measurement also allows the determination of the relative strength of Yukawa coupling between the Higgs boson and bottom/charm quarks.

$$\mu_{VH,H\to b\bar{b}} \equiv \frac{(\sigma_{VH} \times BR_{H\to b\bar{b}})_{\text{Measured}}}{(\sigma_{VH} \times BR_{H\to b\bar{b}})_{\text{SM expectation}}}. \tag{1.16}$$

$$\mu_{VH,H\to c\bar{c}} \equiv \frac{(\sigma_{VH} \times BR_{H\to c\bar{c}})_{\text{Measured}}}{(\sigma_{VH} \times BR_{H\to c\bar{c}})_{\text{SM expectation}}}. \tag{1.17}$$

In addition to the simultaneous measurement of the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ signal strengths, determining the contribution of dominant standard model background processes in the $VH, H \to b\bar{b}$ (or $VH, H \to c\bar{c}$) phase spaces that can be extrapolated to the $VH, H \to c\bar{c}$ (or $VH, H \to b\bar{b}$) phase spaces, and harmonizing on different analysis techniques which were useful in the previously three independent analyses, comes as additional benefits. These are briefly discussed in Section 1.4.2, and in more detail in this thesis.

1.4.1 The signal and background processes

As discussed previously, the WH and ZH channels give the best sensitivity for $H \to b\bar{b}$ and $H \to c\bar{c}$ searches, since this process can be triggered using the leptonic decay of the vector bosons. The signal is probed in three channels (as depicted in Figure 1.11), depending on the leptonic final state of the vector boson.

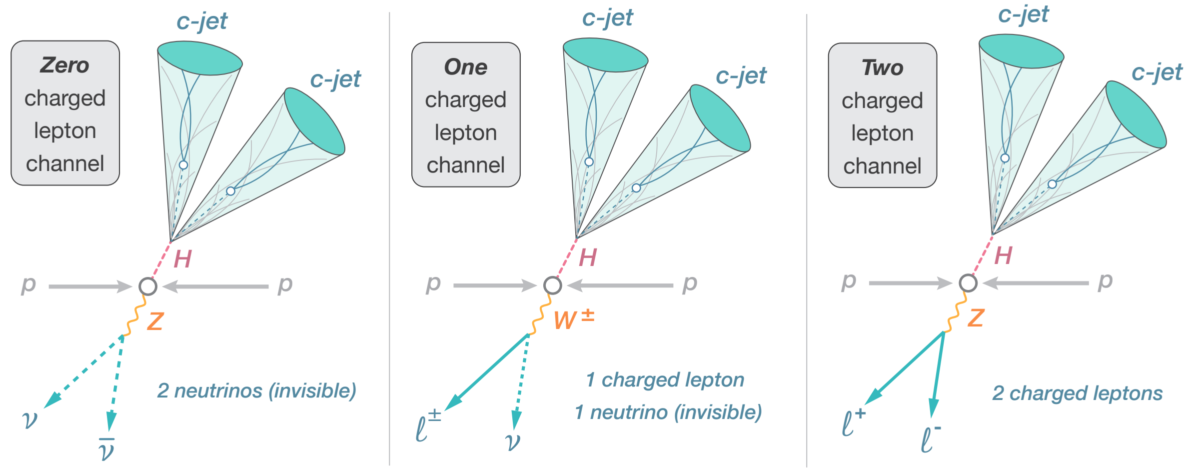


Figure 1.11: The three different charged lepton channels used in the search for $VH, H \to c\bar{c}$.

When probing the $VH, H \to b\bar{b}/c\bar{c}$ signal processes, it is necessary to consider the contribution of the other standard model processes (shown in Figure 1.4) in the signal phase space. Such processes, referred to as *background processes*, can be classified to two main types in the current analysis; the top-quark processes and vector boson processes as illustrated in Figure 1.12.

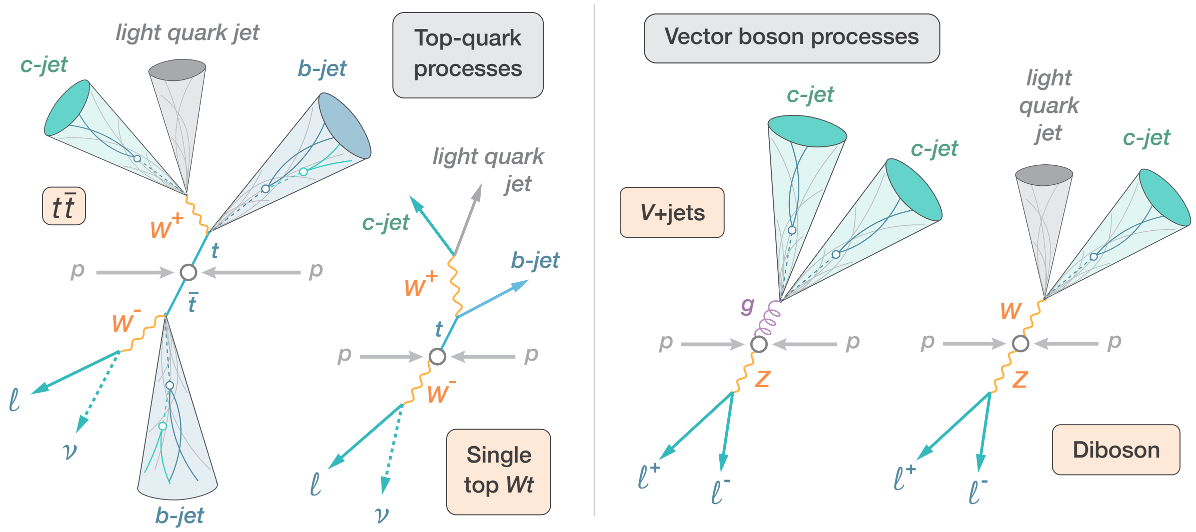


Figure 1.12: Examples of the main standard model background processes in the analysis.

1.4.2 The $VH, H \to b\bar{b}/c\bar{c}$ analysis strategy

The $VH, H \to b\bar{b}/c\bar{c}$ analysis uses the same Run 2 proton-proton collision data set as the previous $VH, H \to c\bar{c}$ analysis [30], which now accounts for 140 fb^{-1} of data. The candidate events for the analysis are selected using the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) triggers, the single electron triggers, and the single muon triggers, as discussed in more detail in Section 3.1.1.

The recorded data of the different detector components in triggered events are then reconstructed into different objects; electrons, muons, hadronically decaying τ leptons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. In the $VH, H \to b\bar{b}/c\bar{c}$ analysis, the most important objects are the jets initiated by the b or c quark pairs, since they are used to reconstruct the Higgs boson candidate. Hence, it is important to identify the initial partons that create the jets with a good accuracy, and this is achieved using the jet-flavor tagging algorithms discussed in more detail in Section 4.5.

The $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ signal regions undergo further events selections to enhance the signal over background significance (referred to *significance* in the thesis), as described in Section 5.1. The selected events are then fed to a multivariate analysis (MVA) algorithm to further improve the signal to background separation (described in Chapter 7), and the MVA output is used as the final discriminant to extract the signal strength $\mu_{VH,H\to c\bar{c}}$, simultaneously with $\mu_{VH,H\to b\bar{b}}$. Furthermore, to estimate the contribution of different background processes in the signal region, several control regions are defined, as described in Section 5.3.

The signal and most background processes are generated using monte-carlo (MC) simulations (described in Chapter 3), and these events are used to build the templates in the signal and control regions. The signal strengths $\mu_{VH,H\to c\bar{c}}$ and $\mu_{VH,H\to b\bar{b}}$ are then extracted by fitting all the templates using a profile likelihood fit, and is described in detail in Chapter 9.

The $VH, H \to b\bar{b}/c\bar{c}$ analysis introduces several improvements listed below, to increase the signal sensitivity compared to the previous Run 2 $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ analyses [34, 35, 30].

- A new jet-flavor tagging algorithm (known as the DL1r tagger [36]) used in the new analysis gives better separation between jets with b-hadrons (b-jets), c-jets, and other jets. Based on this new jet-flavor tagging algorithm, new working points are defined, and are optimized to get the best signal over background significance for the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ processes. The new algorithm is described in more detail in Section 4.5.

- New control regions to constrain the Top-quark processes and V+jets processes are defined in the new analysis, while the previous ΔR control regions are optimized to increase the sensitivity in signal regions even further. More details can be found in Chapter 5.

- A new event weighting method based on graph neural networks (described in Chapter 6) for effectively using more simulated samples is used, and this helps to reduce the systematic uncertainty attributed to the size of the simulated sample set. The old method utilized in the previous analyses was attributed additional systematics uncertainties due to several limitations in the accuracy of the event weighting [30]. The new event weighting method, which has a better accuracy, does not require such additional uncertainties.

- A new analysis region, at low transverse momentum of the vector boson, is added in the one charged lepton channel, increasing the signal acceptance.

- A multivariate analysis is used to separate signal processes from the background processes, further increasing the sensitivity in the signal regions. This is described in more detail in Chapter 7.

- Modeling uncertainties, acceptance between regions, and shape effects are re-estimated in the analysis. For estimating uncertainties due to shape effects, a new tool based on neural networks, the Calibrated Likelihood Ratio Estimator (CARL) [37] is used. More on this can be found in Section 8.2.

1.4.3 Thesis overview and personal contribution

Chapter 1 has given a brief introduction to the motivation and strategy of the $VH, H \to b\bar{b}/c\bar{c}$ analysis. Next, Chapter 2 introduces the Large Hadron Collider (LHC) and the ATLAS experiment. Chapter 3 explains the data and monte-carlo (MC) samples used in the analyses. Then Chapter 4 explains the different objects that are used in the analyses, followed by the event selection and event categorization described in Chapter 5. Chapter 6 explains the event weighting method used for effectively using the simulated sample set. Chapter 7 explains the multivariate analysis used for further separating the signal from background. The various systematic uncertainties are described in Chapter 8, and Chapter 9 explains the statistical analysis used to extract the signal strengths $\mu_{VH,H\to c\bar{c}}$ simultaneously with $\mu_{VH,H\to b\bar{b}}$. Finally Chapter 10 discusses the results of this analysis and the interpretation in the κ-framework.

My main contributions to the $VH, H \to b\bar{b}/c\bar{c}$ analysis are listed below.

- Studies of the signal yield in the signal regions, and the ΔR control regions for different flavor tag regions. This is discussed in Section 5.3.2, and Appendix E.2.
- Implementation of the graph neural network based event weighting method, and validating its performance. Further, several fine tuning studies were done to bring the best performance while avoiding additional systematic uncertainties on this method. This is discussed in Chapter 6.
- Studies of different variables for the multivariate analysis, which brings better signal vs background separation. This is discussed in Section 7.2.
- Estimation of the modelling acceptance uncertainties for the di-boson background process. This is discussed in Section 8.2.5.
- Studies on the simultaneous fit to extract the $\mu_{VH,H\to c\bar{c}}$ and $\mu_{VH,H\to b\bar{b}}$, which is presented in Chapter 10.

Chapter 2

The Large Hadron Collider and the ATLAS experiment

This chapter describes the two important experimental instruments used for collecting the data used in this analysis: the Large Hadron Collider and the ATLAS experiment located in CERN, Switzerland.

2.1 The Large Hadron Collider

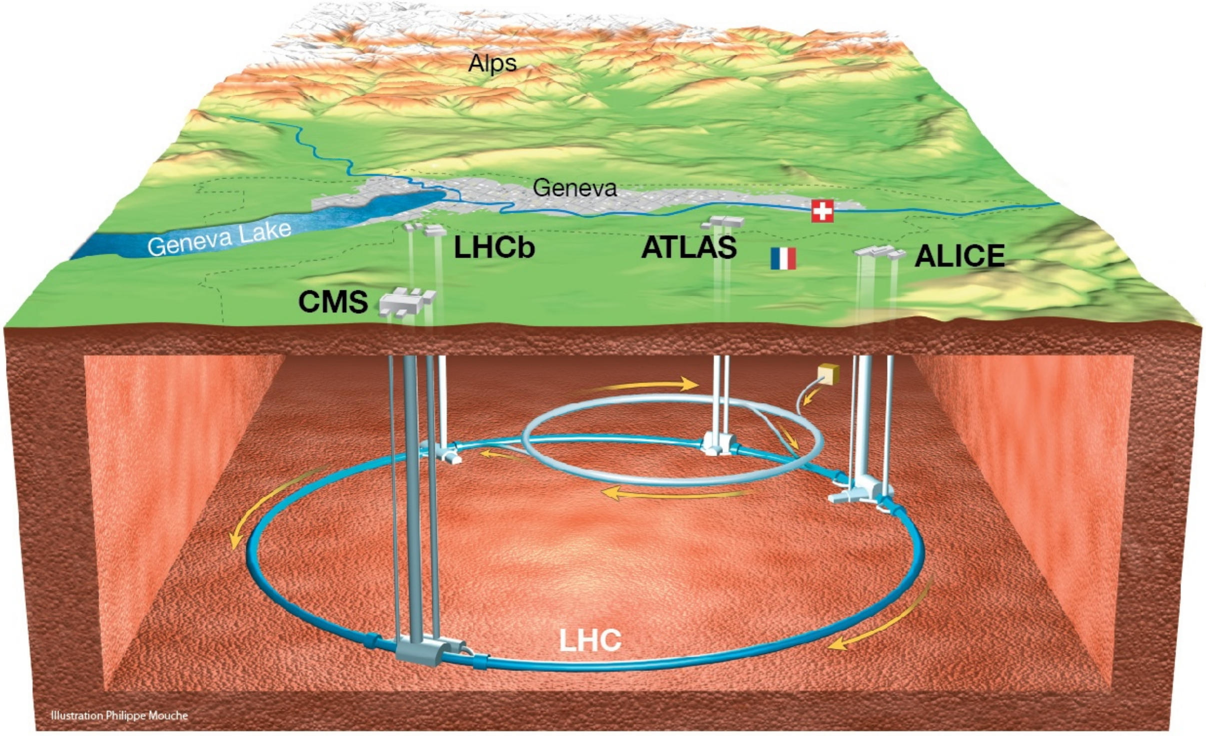


Figure 2.1: The LHC and its four main detectors ALICE, ATLAS, CMS, and LHCb [38].

The Large Hadron Collider (LHC) is the largest particle accelerator in the world, which is housed inside an underground tunnel of 26.7 km in circumference [39]. The LHC tunnel was initially built for the large electron positron collider (LEP), and is situated in an average depth of 100 m[1] below the French-Swiss border near Geneva, Switzerland. The LHC is built with two beam-pipes which carries protons (or heavy-ions) in opposite directions, and the two beam pipes cross at four interaction regions where the four LHC experiments are located: ALICE [40], ATLAS [41], CMS [42], and LHCb [43]. And overall view of the LHC and its four main detectors is shown in Figure 2.1.

The Run 1 data collection period of the LHC was from 2009 to 2013, where the LHC operated at a center-of-mass energy of $\sqrt{s} = 7$ TeV and 8 TeV [44]. The LHC then underwent a two year period of long shutdown for the repair and upgrade of the LHC and the four LHC-detectors. The Run 2 data taking period of the LHC started in 2015 at a center-of-mass energy of $\sqrt{s} = 13$ TeV, and continued till the end of 2018 [45]. After another three and a half years of a long shutdown, the LHC started operation at a center-of-mass energy of $\sqrt{s} = 13.6$ TeV in the summer of 2022, and plans to continue operation till the summer of 2026 [46, 47].

2.1.1 Production, acceleration and guiding of the proton beams

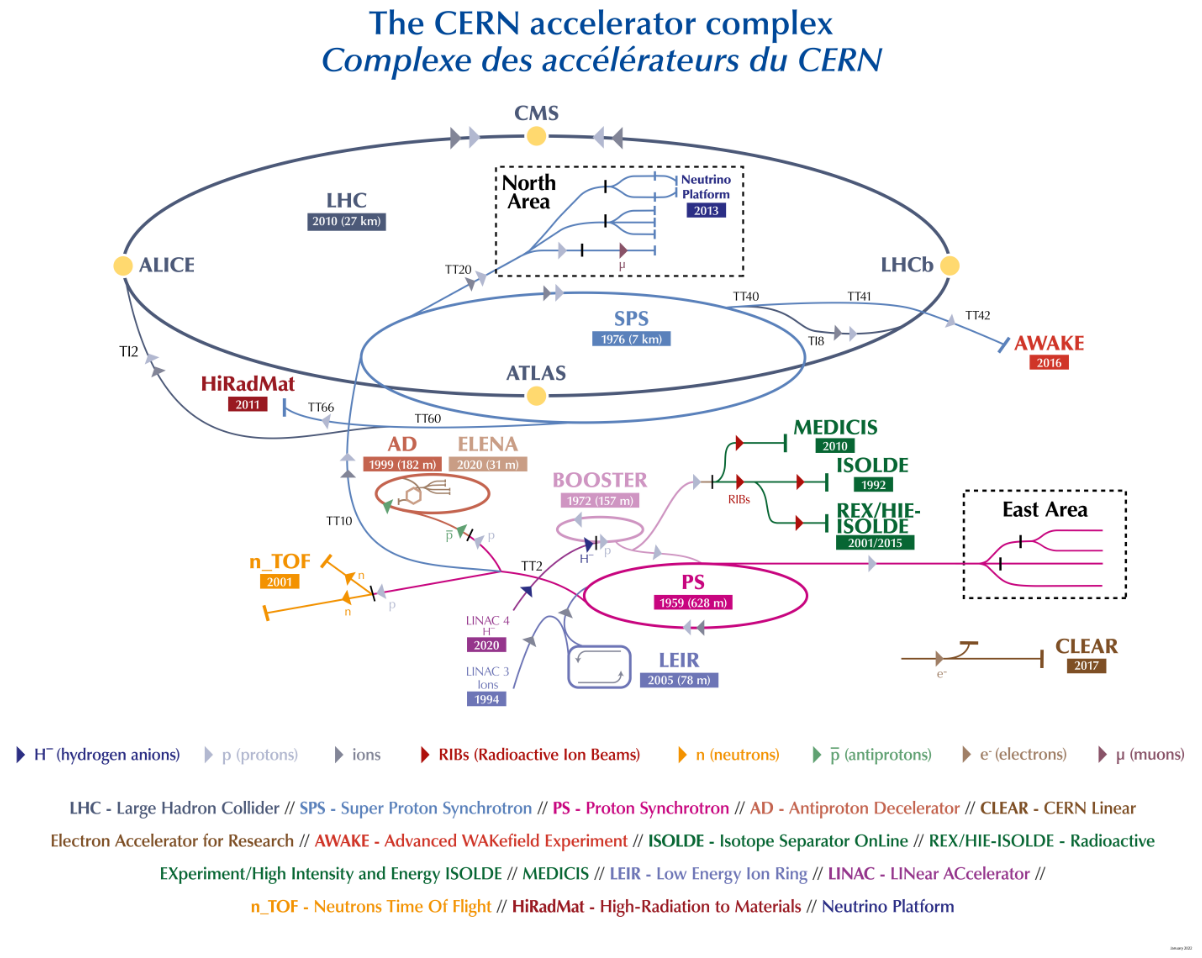


Figure 2.2: The CERN accelerator complex and the LHC [48].

---

[1] The depth of the tunnel varies from 50 m (towards lake Geneva) to 175 m (under the Jura mountains), with an inclination of 1.4% towards the lake Geneva.

During the Run 2 period of the LHC (from 2015 to 2018), protons for the LHC were produced from hydrogen atoms where electrons are striped off using a strong electric field. These protons were accelerated to 50 MeV using the 30 m Linac2[2] linear accelerator, and then injected into a 157 m Proton Synchrotron Booster (PSB) ring and accelerated to an energy of 1.4 GeV. The beam from the PSB is then injected in to the 628 m Proton Synchrotron (PS), which accelerates the protons to an energy of 25 GeV before injecting to the 6.9 km Super Proton Synchrotron (SPS). The SPS gives the final stage of acceleration of the proton beam to 450 GeV before injecting the beam to the LHC [49].

The acceleration of the LHC beams are provided by 16 superconducting radio-frequency (RF) cavities (eight per proton beam) that are housed in four cryomodules. Each cavity is driven by a high-power klystron at a frequency of 400 MHz, and can reach a maximum voltage of 2 MV per proton beam; this corresponds to an acceleration of 16 MV per proton beam in one revolution around the LHC[50, 51]. The acceleration to 6.5 TeV per beam (14 times the injected energy) is achieved in about 20 minutes, with the protons having circulated the LHC ring more than 10 million times. Once the beam has reached the required energy, ideally protons tuned with the 400 MHz RF frequency that have reached the desired energy will not be accelerated. However, protons with slightly different energies arriving earlier or later will be accelerated or decelerated to reach the desired energy. Therefore, the proton beam is grouped to *bunches*, with each bunch containing 1.2×10^{11} protons, and each beam containing 2808 proton bunches.

The beams are guided through the LHC ring using thousands of magnets. These include 1232 dipole magnets that are 15 m in length, and designed to allow a current of 11,850 A to flow through, in order to generate a 8.3 T magnetic field that is able to bend the 7 TeV beams. To achieve this, the magnets are made of NbTi cables that reach super-conductivity below 10 K, and therefore, the magnets are cooled to a temperature of 1.9 K using liquid Helium [52]. Additionally, 392 quadrupole magnets between 5 m to 7 m in length are used for focusing the beams [53]. There are also magnets with higher multipoles used to correct small imperfections in the beam geometry, and stronger quadrupole magnets used to squeeze the beams before colliding them, to increase the chances of a proton-proton collision.

2.1.2 Luminosity and pile-up

The instantaneous luminosity $\mathcal{L}$ of a particle accelerator is related to the production cross-section of a physics process σ as per the equation:

$$\frac{dN}{dt} = \mathcal{L} \times \sigma, \tag{2.1}$$

where dN/dt is the rate of particle interactions of the physics process. Assuming round and identical beams at the interaction point, the instantaneous luminosity for the LHC can be expressed as

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

---

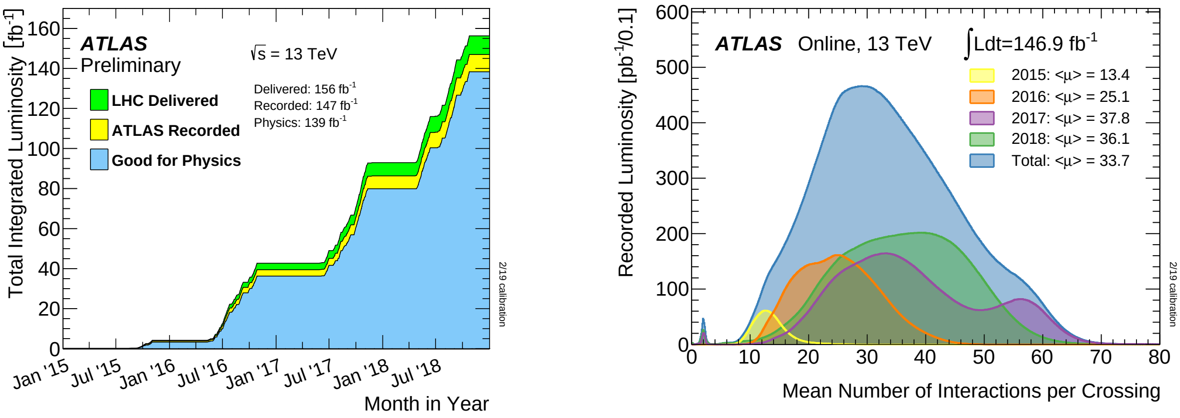
[2] The Linac2 was replaced by the the Linac4 after the Run 2 period. The Linac4 accelerates H^- ions to 160 MeV before injecting it to the PSB.

The description of each parameter and the corresponding values during Run 2 is shown in Table 2.1. Based on the values given in Table 2.1, a good approximation on the nominal LHC instantaneous luminosity is obtained[3] to be $\sim 1 \times 10^{34}$ cm^{2}s^{-1}.

Table 2.1: Nominal luminosity parameters of the LHC during Run 2 data taking period [53, 54]. The geometric reduction factor F is a term which accounts for the reduction in beam overlap (due to the crossing angle between the beams) with respect to head-on collisions.

	Luminosity parameter	**Value**
γ	Relativistic gamma factor	$\approx$ 7000
f_{rev}	Bunch revolution frequency	11245 Hz
n_b	Number of colliding bunches	2808
N_b	Particle density per bunch	1.2 $\times 10^{11}$
F	Geometric reduction factor	0.84
ϵ_n	Normalized beam emittance	3.75 μm
β^*	Horizontal beam dispersion function	0.55 m

The integrated luminosity $\hat{\mathcal{L}}$ is calculated by integrating the instantaneous luminosity over a given time period $t - t_0$ as $\hat{\mathcal{L}} = \int_{t_0}^{t} \mathcal{L}(\tau) d\tau$. The integrated luminosity during the Run 2 data taking period is shown in Figure 2.3(a). In the preliminary calibration of the luminosity, the total integrated luminosity was determined to be 139 fb^{-1} seen in Figure 2.3(a). However, after a more refined calibration [55], the luminosity value was determined to be 140.1 fb^{-1}, which is also the value used in the analysis discussed in this thesis.



(a) Cumulative integrated luminosity.

(b) Mean number of interactions per bunch crossing.

Figure 2.3: The left plot shows the cumulative integrated luminosity delivered to ATLAS, recorded by ATLAS, and finally selected as good for physics. The right plot shows the luminosity-weighted distribution for the mean number of interactions per bunch crossing $\langle \mu \rangle$ [56].

During a bunch crossing, there are multiple proton-proton interactions taking place, which is referred to as *pile-up*. The mean number of pile-up interactions varied mostly between 8 to 70 during the Run 2 period, as seen in Figure 2.3(b).

---

[3] The peak luminosity during Run 2 was 2.1 $\times 10^{34}$ cm^{2}s^{-1}, but varied in each LHC fill. Further, the instantaneous luminosity expressed in Equation 2.2 is an approximation as described in [49].

2.2 The ATLAS experiment

The ATLAS experiment, short for **A** **T**oroidal **L**HC **A**pparatu**S**, is 46 m in length, 25 m in diameter, and covers nearly the entire solid angle around the collision point as seen in Figure 2.4. The ATLAS experiment is made up of mainly three main sub-detector systems; the inner detector, the calorimeters, and the muon spectrometer.

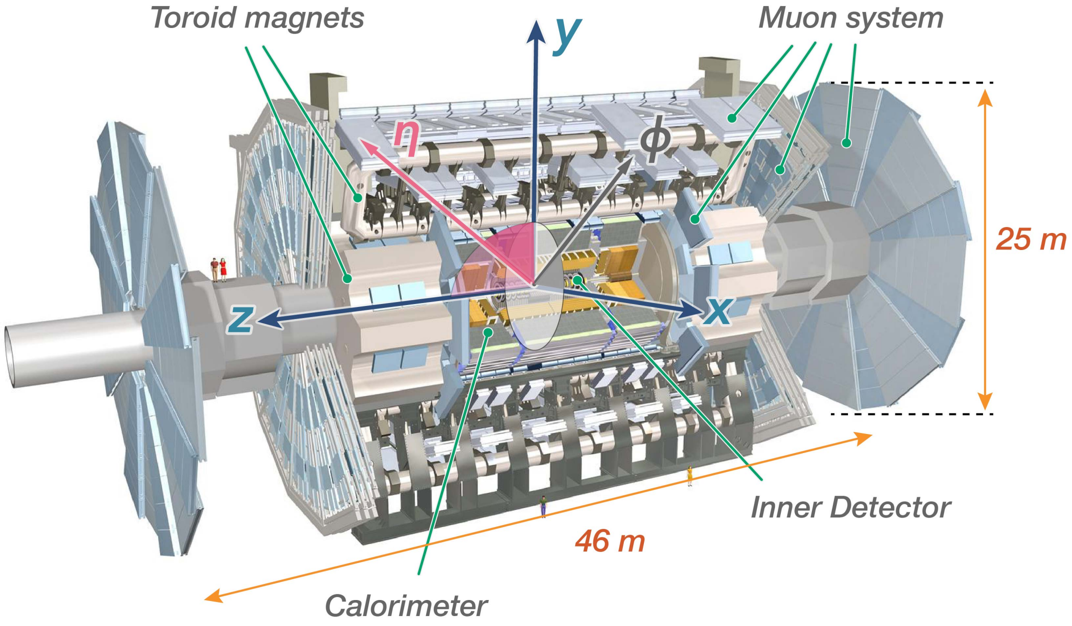


Figure 2.4: The overall view of the ATLAS experiment [57].

The ATLAS detector uses a right-handed coordinate system, with the origin of the axes at the center of the detector, which is also defined as the nominal interaction point. The z-axis is taken along the beam axis, the y-axis points upwards, and the x-axis points towards the center of the LHC from the nominal interaction point. Polar coordinates (r, ϕ) are used in the transverse $x-y$ plane, with ϕ being the azimuthal angle in the clockwise direction around the z-axis as shown in Figure 2.4. The polar angle θ is used to define the pseudorapidity $\eta = -\ln\tan(\theta/2)$, that is equal to the rapidity y (given in Equation 2.3) in the relativistic limit.

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right), \tag{2.3}$$

where p_z is the z-component of the momentum. The momentum component in the transverse $x-y$ plane is denoted as $\vec{p_{\mathrm{T}}}$. Further, the angular distance between two vectors is a commonly used quantity in the ATLAS experiment, which is defined using the quadratic sum of the differences in ϕ and η as defined in Equation 2.4.

$$\Delta R_{1,2} = \sqrt{\Delta\phi_{1,2}^2 + \Delta\eta_{1,2}^2} = \sqrt{(\phi_1-\phi_2)^2 + (\eta_1-\eta_2)^2} \tag{2.4}$$

2.2.1 Inner detector

The inner detector is the innermost subsystem of the ATLAS detector, and is made up of the pixel detector which is closest to the proton-proton collisions, followed by the semiconductor tracker (SCT), and the transition radiation tracker (TRT) as seen in Figure 2.5. A solenoid magnet that provides a 2 T magnetic field surrounds the pixel, SCT and TRT systems. From the center of the beam pipe, the inner detector extends radially from 33.25 mm (the first layer of the pixel detector) to 108 cm (the outer end of the TRT), and covers the $\eta < 2.5$ region.

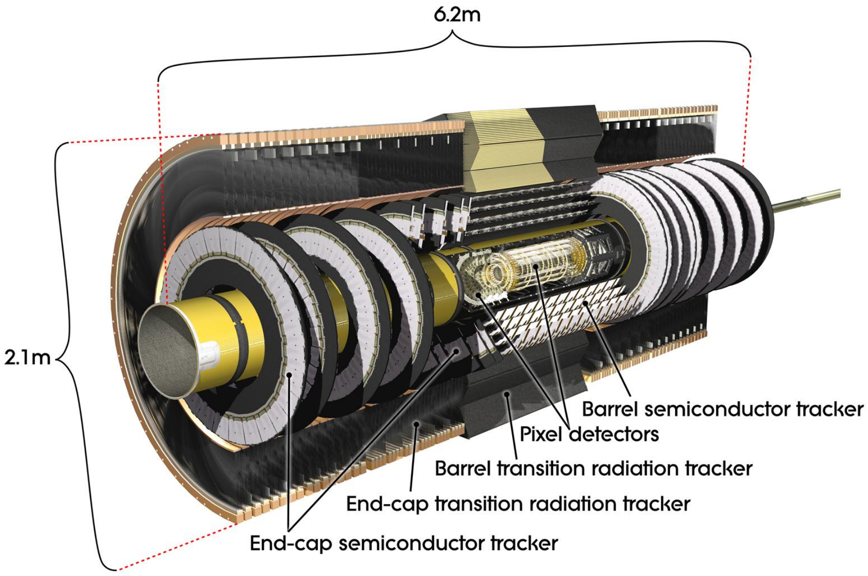


Figure 2.5: The pixel, SCT, and TRT sub-systems of the ATLAS inner detector [58].

The main purpose of the inner detector is to measure the momentum of charged particles through the curved trajectories in the 2 T solenoid field. The design inverse p_{T} resolution of charged pions [59] in the inner detector is given by ,

$$\sigma\left(\frac{1}{p_{\mathrm{T}}}\right)\cdot p_{\mathrm{T}} = 0.036\% \cdot p_{\mathrm{T}} \oplus \frac{1.3\%}{\sqrt{\sin\theta}}, \tag{2.5}$$

where p_{T} is measured in GeV, and the first term accounts for the intrinsic error on track parameters, while the second term accounts for the error in multiple scattering [60].

Pixel detector

The pixel detector is made up of silicon pixel detector modules arranged in four longitudinal layers (i.e *barrel* layers), and four vertical layers (i.e *end-cap* layers). The ATLAS experiment was designed originally with the three outermost barrels and the four end-cap layers, consisting of 1744 silicon pixel modules [61, 62]. One pixel module is composed of 16 front-end (FE) chips bump bonded to a sensor having a pixel resolution of 50 μm × 400 μm, and the FE chips are wire-bonded to a flexible printed circuit board (PCB), as illustrated in Figure 2.7(a)

The inner most pixel module layer, referred to as the insertable B-layer (IBL), was added for Run 2 during the first long shutdown of the LHC [63]. The IBL silicon pixel modules have a higher granularity (50 μm $\times$ 250 μm) compared to the rest of the pixel modules, and improves the precision and robustness of track reconstruction[4] [64]. In total there are 280 IBL pixel modules arranged on 14 longitudinal staves, at a radius of 33.25 mm around the beam pipe. The other three barrel layers extend radially up to 12.25 cm as depicted in Figure 2.6.

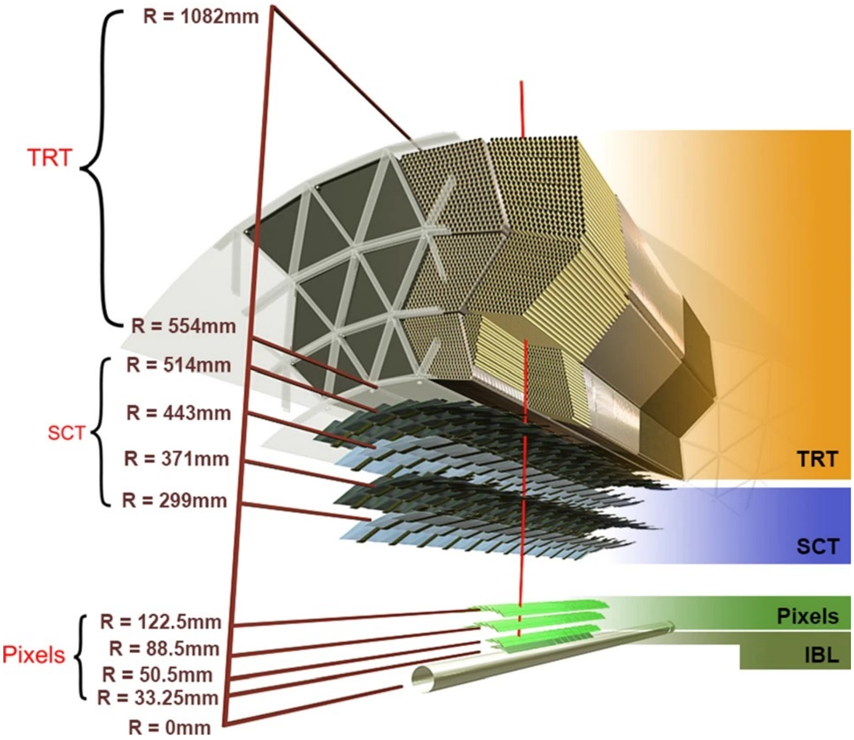


Figure 2.6: The illustration of the ATLAS inner detector barrel region [62].

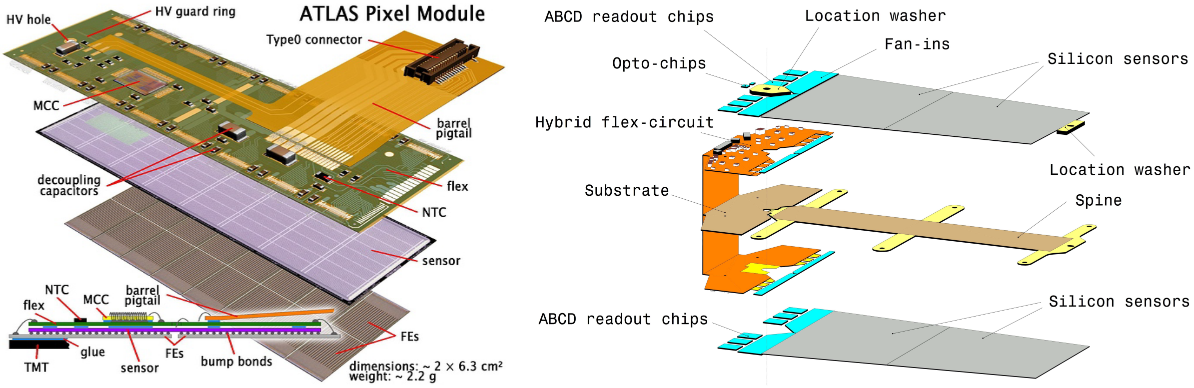
Semiconductor tracker

The semiconductor tracker (SCT) is made up of 4088 silicon strip modules, arrange in four barrel layers and 18 end-cap disks [65]. The SCT barrel layers have rings of 12 modules along the z-axis, totalling 2112 barrel modules. Each barrel modules consists of four 285 μm thick p^{+}-on-n type silicon sensors, with 768 active micro-strips that are $\sim$ 6 cm in length, placed at a 80 μm pitch to each other [66]. A pair of sensors are wire-bonded to each other to create 12 cm long strips, and then glued back to back at a stereo angle of 40 mrad to create one 6 cm $\times$ 12 cm SCT barrel module. The layout of a SCT module is depicted in Figure 2.7(b).

Each SCT end-cap has 988 modules in total, arranged in rings of the nine disks per end-cap [67]; the number of rings vary depending on the position of the disks. Most end-cap modules have similar dimension to barrel modules, but are trapezoidal in shape in order to accommodate the modules in circular disks. Further, the end-cap modules in the inner rings of the disks 2-6, and the middle ring of disk 8, consists of only one sensor on each side and therefore is only half the length of other SCT modules. Signal amplification, shaping and digitization is done using the readout chips, with each SCT module having 12 readout chips (one chip per 128 strips) [68].

---

[4] For jet-flavor tagging (discussed more in Section 4.5), the IBL was estimated to improve the light-flavor jet rejection by a factor of four in b-tagging working points.



(**a**) An ATLAS pixel module.

(**b**) An end-cap SCT module.

Figure 2.7: The pixel and SCT modules of the ATLAS inner detector [41].

Transition radiation tracker

The transition radiation tracker (TRT) is the outermost sub-system of the inner detector, and is made of 298,304 carbon-fibre-reinforced kapton straw tubes of 4 mm in diameter [69]. The straw tubes are filled with a gas mixture (70% Xe, 27% CO_2, 3% O_2), and has 70 μm thick walls that are kept at a potential of -1530 V. The TRT operates as a drift chamber, where charged particles traversing the straw tube deposits about 2.5 keV of energy in the active gas through ionization, and produces detectable signals. Additionally, extra information for particle identification is provided by the transition radiation X-rays created when a charge particle passes the polypropylene or polyethylene fibers in which the straw tubes are embedded in, further ionizing the gas in the straw tubes. The intensity of transition radiation is proportional to the Lorentz factor $\gamma = E/mc^2$ of the particle [70], and therefore, in addition to tracking the TRT is used for identification of different particles with similar energies, such as electrons from pions.

Similar to SCT, the TRT is also composed of one barrel section and two end-caps, and the coverage is in the range $|\eta| < 2.0$. In the TRT barrel, 142.4 cm long tubes are arranged into 96 barrel modules in three layers, while 39 cm long tubes are arranged to 40 disks in each TRT end-cap. The TRT has a position resolution of 130 μm in the $r - \phi$ plane in the barrel region.

Solenoid magnet

The superconducting solenoid magnet is made from high-strength Al-stabilized NbTi conductor, and placed just outside the inner detector as seen in Figure 2.12(b). It is designed to provide a 1.998 T axial magnetic field at a nominal current of 7.730 kA. To minimize the energy loss of particles before reaching the calorimeter, the solenoid magnet was designed to have a minimum amount of material; approximately 0.66 radiation lengths (X_0) at normal incidence [41]. The magnetic flux is returned by the steel plates in the hadronic calorimeter and its support structure.

2.2.2 Electromagnetic and hadronic calorimeters

The ATLAS calorimeter is composed of the electromagnetic calorimeter and hadronic calorimeters used to measure the energy of neutral and charged particles, and the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ [41]. The calorimeters provide complete coverage in the ϕ-direction, and in the range $|\eta| < 4.9$. Both calorimeters are sampling calorimeters. Regions made of passive high-density material are used to induce electromagnetic or hadronic showers, and a region with an active material is used to measure the energy deposited by the showers.

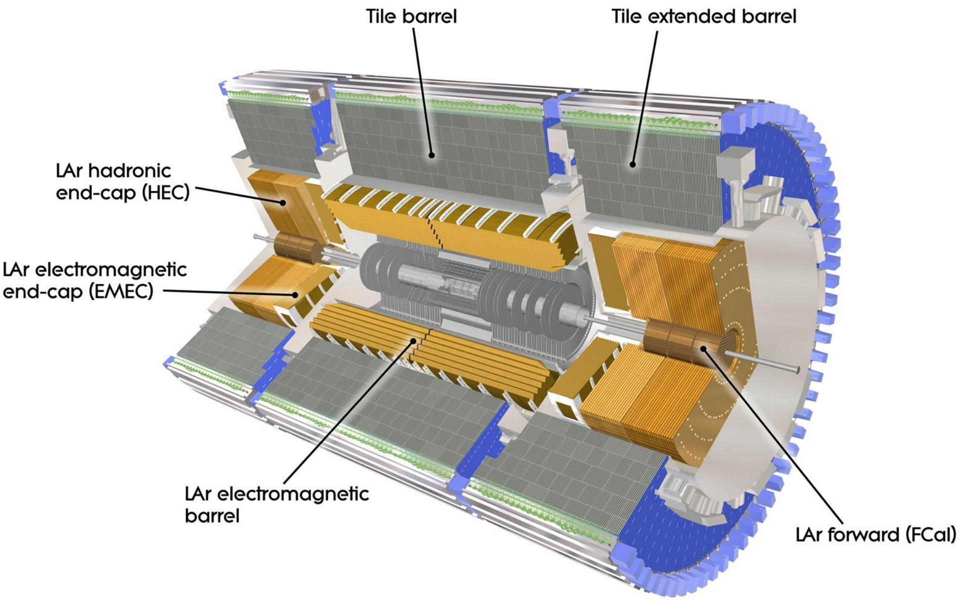


Figure 2.8: The ATLAS electromagnetic and hadronic calorimeters [71].

The electromagnetic calorimeter has greater than 22 radiation length (X_0), and the hadronic calorimeter has around 10 interaction length (λ) in the different regions of the detector. This ensures a good containment of the electromagnetic and hadronic showers, while limiting the punch-through of particles into the muon spectrometer.

The energy resolution of the calorimeters can be expressed using a *stochastic term* $a/\sqrt{E}$ which accounts for stochastic fluctuations in shower development in the calorimeter, a *noise term* b/E accounting for the electronic and pileup noise measured from calibration runs, and a *constant term* c reflecting local non-uniformities in the response of the calorimeter, as given in Equation 2.6 [41]. Here, the energy E is measured in GeV.

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c. \tag{2.6}$$

The design resolution considers only the stochastic and noise terms, which is $[a = 10\%, c = 0.7\%]$ for the electromagnetic calorimeter, and $[a = 50\%, c = 3\%]$ for the hadronic calorimeter. A forward hadronic calorimeter has different parameters $[a = 100\%, c = 10\%]$ from the rest of the hadronic calorimeters. Finally, the combined energy resolution of the calorimeter[5] obtained using test beam studies has a value $[a = 52\%, b = 1.6$ GeV, $c = 3.0\%]$.

---

[5] Accurately, the shown combined resolution is for the ECal and the TileCal, for pions.

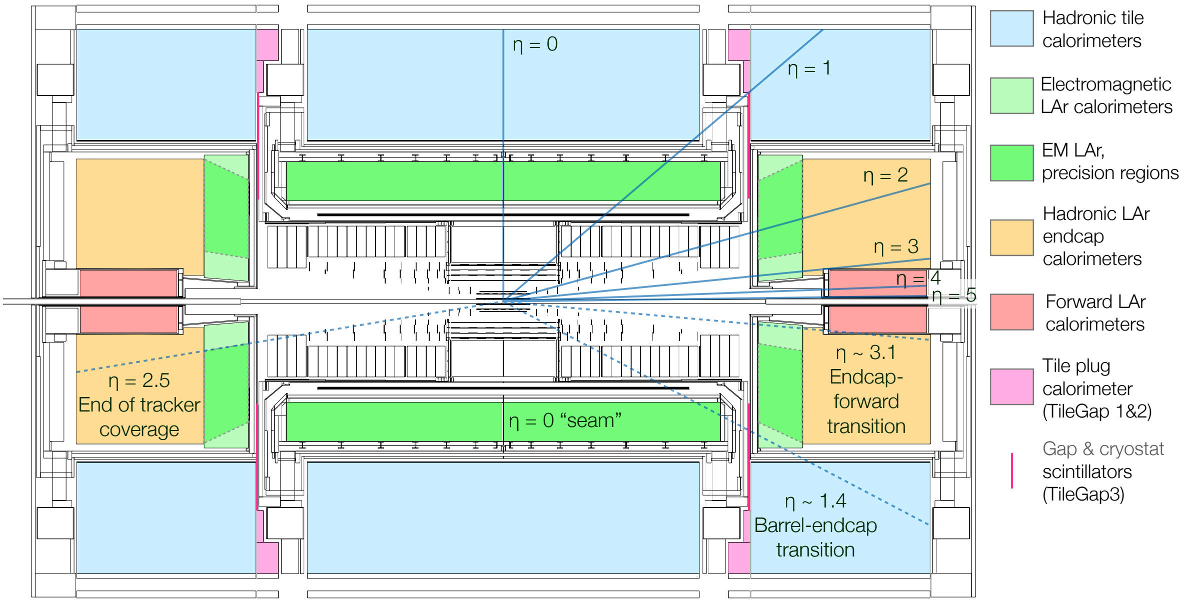


Figure 2.9: A schematic of the cross-sectional view of the ATLAS electromagnetic and hadronic calorimeters [72].

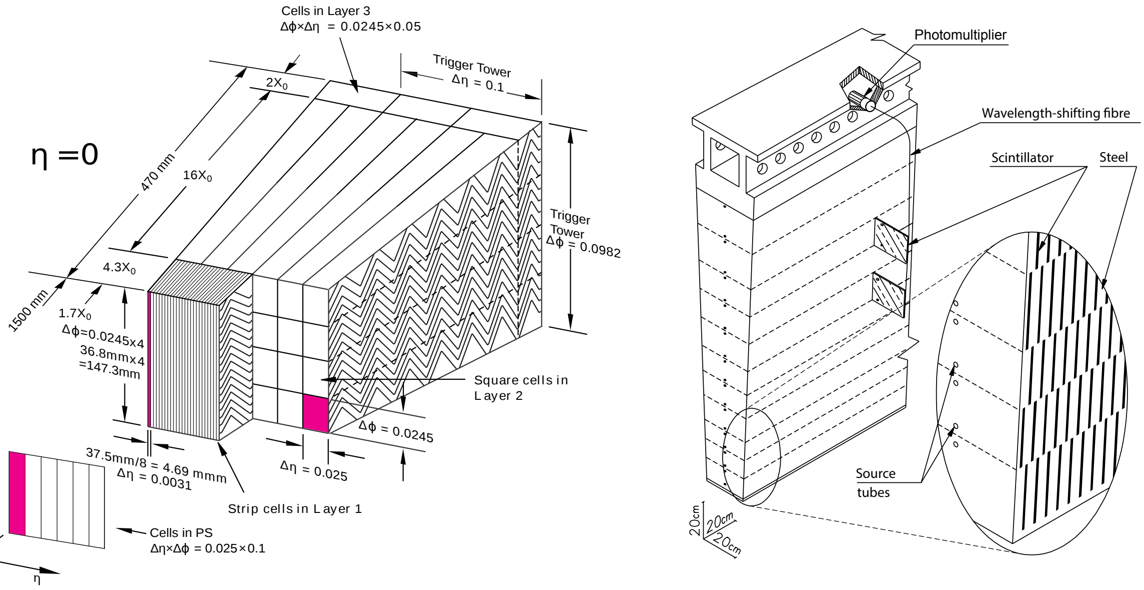
Electromagnetic calorimeter

The electromagnetic calorimeter (ECal) uses sheets of lead as the passive material, while liquid argon is used as the active material. The sheets of lead and the copper electrodes have an accordion shape, which provides a complete coverage in the ϕ-direction without any gaps. The ECal is divided into the barrel region covering $|\eta| < 1.475$, and two end-cap regions covering $1.375 < |\eta| < 3.2$, with all three regions housed in its own cryostat. The barrel to end-cap transition region in $1.37 < |\eta| < 1.52$ contains significant additional inactive material, and therefore not used in most physics analysis.

The region $|\eta| < 2.5$ is dedicated to precision physics, where the ECal is segmented into three layers in depth as seen in Figure 2.10(a). The first layer contains strips finely grated in the η direction, and offers excellent discrimination between isolated photons and pairs of collimated photons coming from a $\pi^0 \to \gamma\gamma$ decay. For electrons, and photons with high transverse energy E_T, most of the energy is deposited in the second layer. The third layer provides energy measurements in the tails of the electromagnetic shower. In the front of the calorimeter $|\eta| < 1.8$ region, a thin pre-sampler detector is used to correct the energy lost by electrons and photons between the collision point and the calorimeter.

Hadronic calorimeter

The hadronic calorimeter is placed directly outside the ECal, and is composed of three types of calorimeters; the tile calorimeter (TileCal), the hadronic end-cap calorimeter (HEC), and the forward calorimeter (FCal). The main barrel region of the TileCal covers the $|\eta| < 1.0$ region, while the extended barrels cover the $0.8 < |\eta| < 1.7$ region. The TileCal uses steel as the passive material, while scintillating tiles are used as the active material. The main barrel and



(**a**) The ATLAS ECal module at $\eta = 0$.

(**b**) The ATLAS TileCal module.

Figure 2.10: The illustration of the ECal and the TileCal modules of the ATLAS experiment [41].

extended barrel are divided azimuthally into 64 modules, with each module extending radially from 2.28 m to 4.25 m as depicted in Figure 2.10(b). The scintillation photons (from the tiles) are collected using wavelength shifting fibers placed on either side of the modules, and the fibers are read out using two separate photomultiplier tubes at the outer radius of each module.

The HEC consists of two independent wheels per end-cap, and uses copper as the passive medium and liquid argon as the active medium. The HEC is placed directly behind the ECal end-caps, and provides calorimetry in the $1.5 < |\eta| < 3.2$ region, slightly overlapping with the TileCal, and the FCal which covers the $3.1 < |\eta| < 4.9$ region. The FCal is approximately 10 interaction lengths deep, and is made of three modules. The first module uses copper as the passive medium, and is optimized for electromagnetic measurements. The remaining two modules use tungsten as the passive medium, and predominantly measures the energy from hadronic interactions. All three modules use liquid argon as the active medium.

2.2.3 Muon spectrometer

The layout of the ATLAS muon spectrometer is shown in Figure 2.11. The muon spectrometry in the ATLAS experiment is based on the precision tracking of deflected muons in the large superconducting air-core toroidal magnets. The precision tracking as well as the fast triggering of muons are done using dedicated muon detectors.

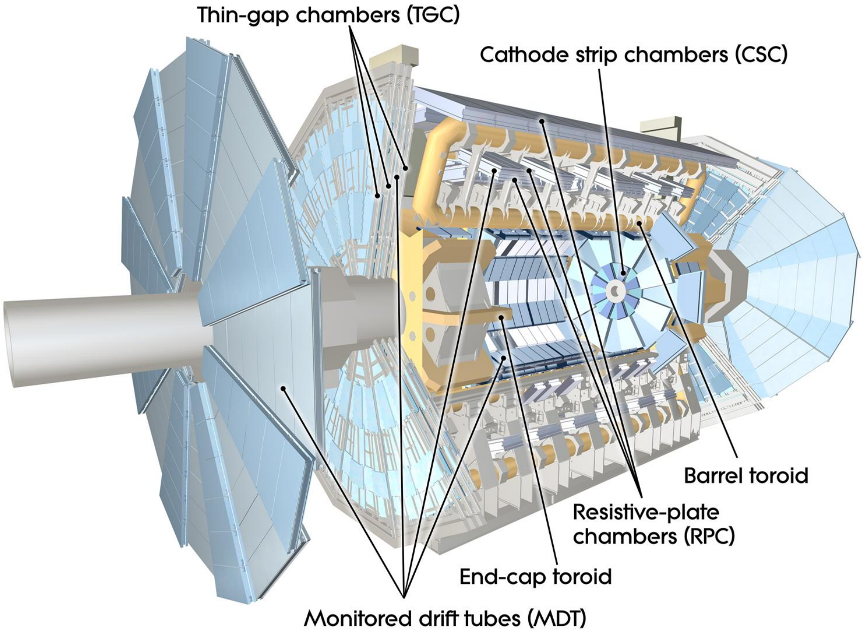


Figure 2.11: The muon system and the toroidal magnets of the ATLAS muon spectrometer [73].

Over most of the η-range of the ATLAS detector, the precision measurement of the track coordinates in the bending direction of the magnetic field (in the η coordinate) is done by monitored drift tube (MDT) chambers. In the $2 < |\eta| < 2.7$ region, Cathode Strip Chambers (CSC's) were used[6] due to their capability to handle the demanding trigger rate and tolerance against the background conditions at high η [75].

The trigger system covers the range $|\eta| < 2.4$ utilizing resistive plate chambers (RPC's) in the barrel region, and thin gap chambers (TGC's) in the end-cap regions. These trigger chambers serve three purposes in the ATLAS detector: to provide the proton bunch crossing identification, to provide well-defined p_{T} thresholds, and to measure muon coordinates in the orthogonal direction (the ϕ coordinate) to the precision-tracking chambers.

The muon spectrometer has a design momentum resolution $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}}$ of $\approx 3.5\%$ for muons with $p_{\mathrm{T}} = 10$ GeV, which decreases to $\approx 2.5\%$ for muons with $p_{\mathrm{T}} = 100$ GeV, and then increases to $\approx 11\%$ for $p_{\mathrm{T}} = 1$ TeV muons [76].

The toroid magnets

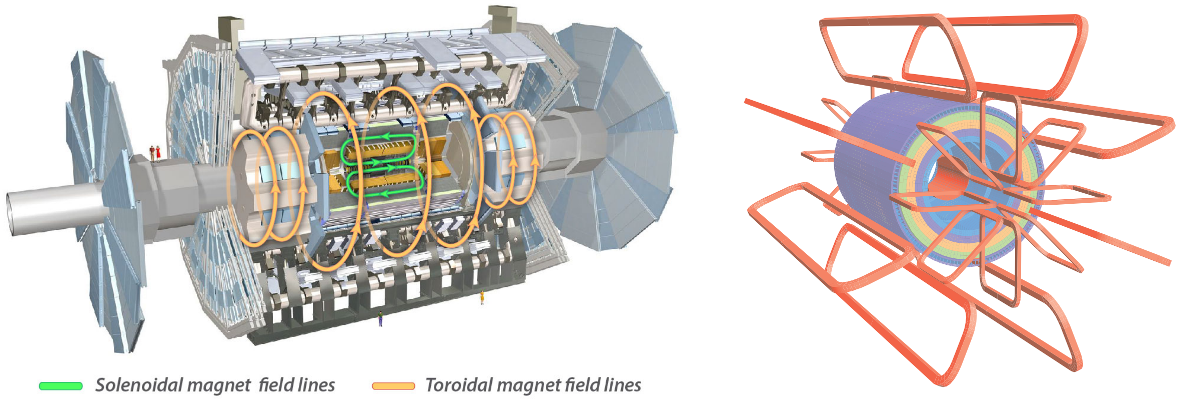
Over the range $|\eta| < 1.4$, the magnetic field for bending the muons is provided by the large barrel toroid. In the range $1.6 < |\eta| < 2.7$, the magnetic field is provided by two smaller end-

---

[6] CSC's were replaced by the micromegas (MM) and small-strips thin gap chamber (sTGC) technologies in the New Small Wheel that was installed into the ATLAS detector after the Run 2 of the LHC [74].

cap magnets as seen in Figure 2.12(a). In the transition region between $1.4 < |\eta| < 1.6$, the magnetic field is a combination of the barrel and end-cap fields [41]. The air toroid magnets provide a field that is mostly orthogonal to the muon trajectory, while minimizing the muon momentum degradation from multiple scattering effects.

As seen from Figure 2.12(b), the toroid magnets are made from eight coils assembled symmetrically around the beam axis. The end-cap toroids are rotated 22.5° with respect to the barrel toroids, so that it can be inserted into the barrel toroid system and provide a radial overlap of the magnetic field.



(**a**) The magnetic field lines in the ATLAS detector.

(**b**) The skeleton of the ATLAS magnets.

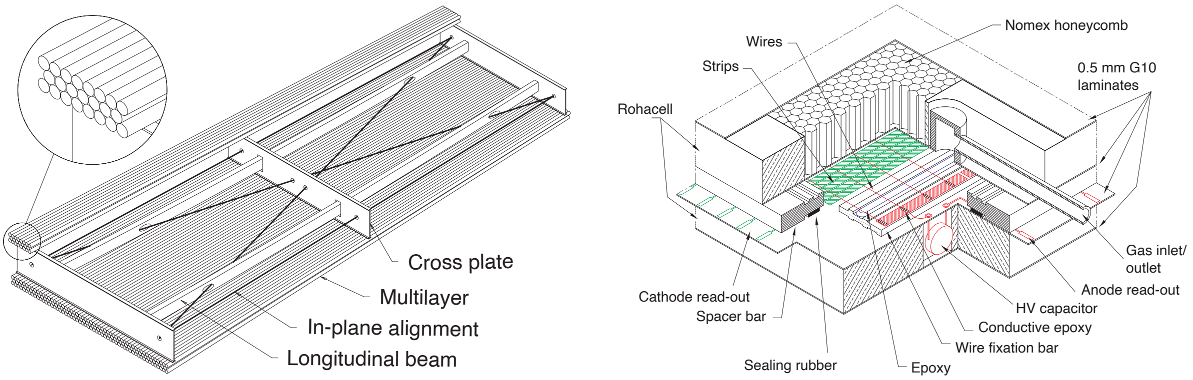
Figure 2.12: The ATLAS magnet system. In the right figure, the orange region within the blue cylinder (the TileCal) represents the central solenoid which generates a 2 T magnetic field for the inner detector, while the outer orange loops represent the toroid magnets [41].

The performance of the toroid magnets bending power is characterized by the field integral $\int B \, \mathrm{d}l$, where B is the magnetic field component normal to the muon direction, and integral is computed along an infinite-momentum muon trajectory (i.e a straight line), between the innermost and outermost muon-chamber planes. The barrel toroids provide a bending power of 1.5 Tm to 5.5 Tm, and the end-cap toroids provide a bending power of 1 Tm to 7.5 Tm.

Precision-tracking muon chambers

The MDT chambers contain pressurized drift tubes with a diameter of 29.970 mm, operating with Ar(93%) and CO_2(7%) gas at 3 bar. Electrons are generated due to ionization from muons passing through, and are collected at the central tungsten-rhenium wire with a diameter of 50 μm that is kept at a potential of 3080 V. The MDT chambers are rectangular in the barrel and trapezoidal in the end-cap, with the shape chosen to optimize the solid angle coverage. In most chambers, three to four layers of tubes are separated by a mechanical structure as seen in Figure 2.13(a). The position resolution per MDT layer was measured during Run 2 to be 81.7 ± 2.2 μm, consistent with the design value of 80 μm [77].

Figure 2.13(b) shows a cutaway view of a single CSC layer. The CSC's are multiwire proportional chambers with wires (of pitch 2.54 mm) oriented in the radial direction, and uses two segmented cathode strips (of pitch 5.08 mm) to readout the position information in the muon bending direction, and orthogonal to the bending direction. The CSC is composed of a gas mixture of Ar(30%), CO_2(50%) and CF_4(20%), and the ionization and subsequent avalanche produced in the gas induces a charge in the cathode strips which can be read out. The resolution of the CSC is 40 μm in the bending plane, and 5 mm in the non-bending plane.



(a) The monitored drift tube (MDT) chambers.

(b) The cathode strip chamber (CSC) layer.

Figure 2.13: The ATLAS muon chambers [76].

The MDT's and in the barrel region are located in three layers, while in the end-cap regions, they are located in four wheels per end-cap. The barrel *inner layer* is placed before the barrel toroidal coils at a radius of $r = 5$ m from the interaction point (IP). The *middle layer* and is placed between the barrel toroidal coils at $r = 7.5$ m from the IP, and the *outer layer* is placed on the barrel toroidal coils, at $r = 10$ m from the IP as seen in Figure 2.14. In the end-cap region, the MDT's are placed together with CSC's in the *inner small wheel* at $|z| = 7.4$ m from the IP, and in three more layers; an *extra layer* at $|z| = 10.8$ m, the *middle big wheel* at $|z| = 14$ m, and the *outer big wheel* at $|z| = 21.5$ m from the IP.

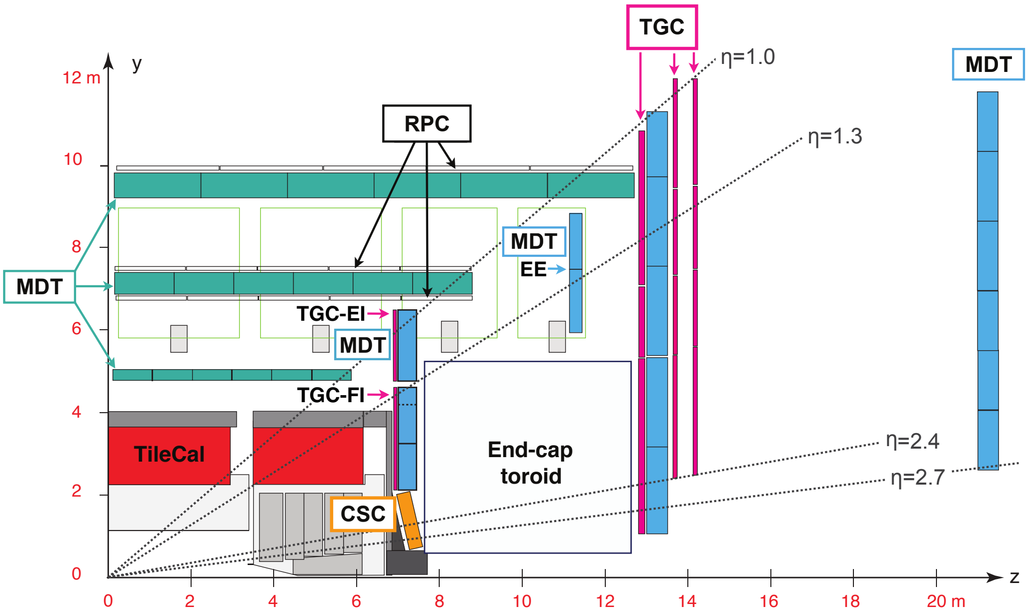


Figure 2.14: A schematic of the quarter-section of the muon system [78]. The Forward Inner and End-cap Inner TGC chambers are marked TGC-FI and TGC-EI respectively. The Extended End-cap MDT chambers are marked EE.

Fast-triggering muon chambers

The RPC is also a gaseous detector similar to other muon chambers, but uses parallel electrode plates instead of wires. Two resistive plates, made of phenolic-melaminic plastic laminate, are kept parallel to each other at a distance of 2 mm using insulated spacers as seen in Figure 2.15. The space between the resistive plates is filled with a gas mixture of $C_2H_2F_4$(94.7%), Iso-C_4H_{10}(5%), and SF_6(0.3%). The 4.9 kV/mm electric field between the plates creates an avalanche along the ionizing particle tracks towards an anode, while the signal is then read out via the capacitive coupling to metallic strips. These metallic strips are mounted on the outer surface of the resistive plates, and one layer has strips in the same direction to the MDT's, while the other layer has strips in the transverse direction to MDT's. The measured RPC timing resolution is approximately 2 ns during the Run 2 period, which is sufficient enough to distinguish the 25 ns bunch crossings in the LHC proton beams [79].

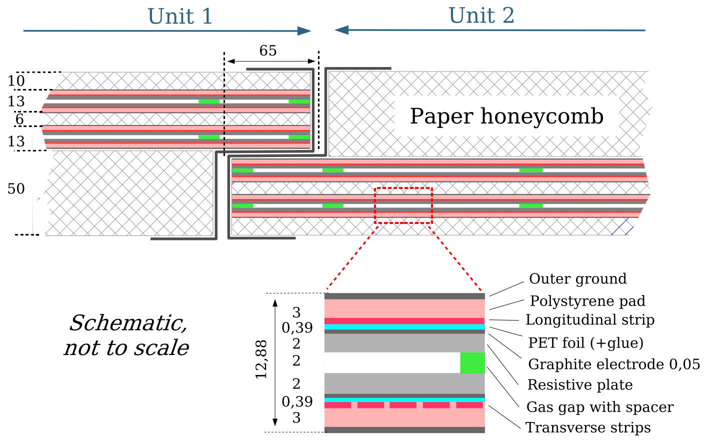


Figure 2.15: The cross-sectional view of the ATLAS RPC muon chambers [41].

TGC's are multi-wire proportional chambers (similar to CSC's) with wire planes (anode) sandwiched by two 1.6 mm FR4 plates[7]. The two FR4 plates are coated with graphite cathode on the anode facing side, and segmented into copper strips on the other side to provide readout in the azimuthal ϕ-plane. The wires are kep at a voltage of 2900 ± 100 V, and the graphite coating provides the ground plane. The TGC is filled with a highly quenching gas mixture of CO_2(55%) and n-C_5H_{12}(45%), is operated at a gas gain of $\sim 3 \times 10^5$. A characteristic of the TGC is that the wire-to-cathode distance of 1.4 mm is smaller than the wire-to-wire distance of 1.8 mm, leading to a fast pulse with a short tail and a timing jitter within 25 ns.

The RPC's are located in three layers in the barrel; two layers of RPC's (RPC1 and RPC2) sandwich the MDT middle layer, while the third layer (RPC3) is located close to the MDT outer layer as seen in Figure 2.14. The TGC's are located in four layers in the end-caps, with the fourth layer added to increase the trigger robustness due to higher background in the forward region. Two TGC layers are located in front of the MDT middle big wheel, while one TGC layer is placed after the middle big wheel. The fourth TGC layer is located between the calorimeter and the end-cap toroidal magnet.

---

[7] FR4 is a glass-reinforced epoxy material used for printed circuit boards, where FR stands for flame retardant.

2.2.4 Forward detectors

In addition to the main detector systems discussed so far, there are four types of *forward detectors* currently present in ATLAS: the luminosity cherenkov integrating detector (LUCID), the zero-degree calorimeter (ZDC), the absolute luminosity for ATLAS (ALFA), and the ATLAS forward proton (AFP) detectors [41, 80]. LUCID is the main relative luminosity monitor in ATLAS, and is located at a distance of ±17 m from the IP. The ZDC is located at a distance of ±140 m from the IP, near the point where the LHC beam-pipe divides into two separate pipes. The primary purpose of the ZDC is to detect forward neutrons in heavy-ion collisions. AFP is installed at a distance of ±220 m from the interaction point to measure protons in the forward direction. The furthest from the IP (at ±240 m) is ALFA, which determines the proton beam absolute luminosity via elastic scattering at small angles.

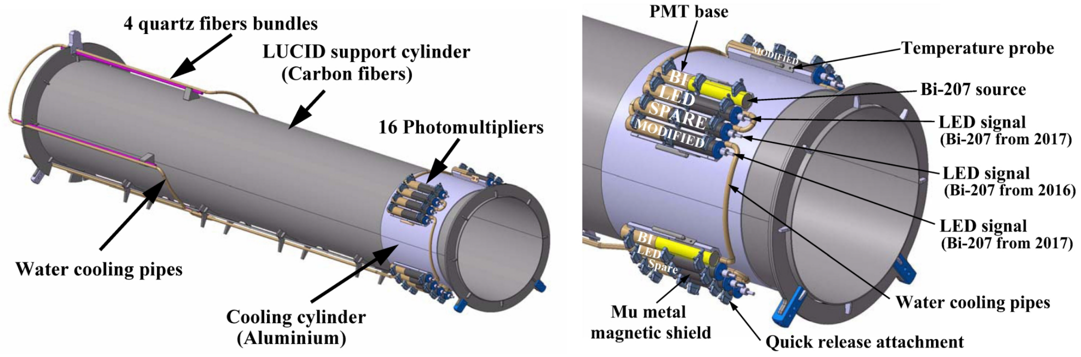


Figure 2.16: The LUCID forward detector in ATLAS [81].

While luminosity can be determined by several detectors, LUCID is the only detector which is primarily dedicated to online luminosity monitoring and luminosity measurement in ATLAS. The LUCID detector was initially installed for the Run 1 data taking period, but was upgraded to the LUCID-2 detector for the Run 2 data taking period [81]. The LUCID-2 detector consists of several small Cherenkov detectors placed on a 1550 mm support cylinder made of carbon fiber as shown in Figure 2.16. LUCID-2 is made up of 16 small photo-multiplier tubes (PMT's) of diameter 10 mm, and four bundles of quartz fiber read out using four additional PMT's, per side (i.e altogether 40 PMT's are used). It uses thin quartz windows of photo-multipliers as the Cherenkov medium, while small amounts of ^{207}Bi sources deposited in these windows allows for the measurement of the gain stability of the photo-multipliers. This results in a fast and accurate luminosity determination that is stable over long periods of data taking.

2.2.5 Trigger and data acquisition system

The trigger and data acquisition (TDAQ) system of the ATLAS experiment (depicted in Figure 2.17) is responsible for deciding which proton-proton bunch-crossing events are kept for offline analysis. Due to the limited long term storage and processing capabilities, only a fraction of events in the order of 10^{-5} can be retained, with the rest being discarded and unavailable for further analysis.

During the Run 1 data taking period, the TDAQ operated in three levels of event selection; level-1 (L1), level-2 (L2), and the event filter. The TDAQ system underwent substantial upgrades for the Run 2 period to handle higher instantaneous luminosities and the higher pile-up compared to Run 1. As a consequence, the TDAQ system was operated in two levels during Run 2; the L1, and the high-level trigger (HLT) that merged the L2 and the event filter [82].

The events are selected by *trigger chains* that consists of the L1 trigger item and a series of HLT algorithms that reconstruct physics objects and apply kinematic selections to them. Each chain is designed to select particular physics signatures of electrons, photons, muons, τ-leptons, jets, missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), the total energy ($E_{\mathrm{total}}$), or B-meson candidates.

To control the rate of accepted events, a pre-scale value can be applied (referred to as pre-scaling); for a pre-scale value of $n \geq 1$, an event has a probability of $1/n$ of being accepted. However, the primary triggers used to collect data for physics analyses are un-prescaled (i.e $n = 1$), and cover all the physics signatures listed above. Further, the main goal of the ATLAS Run 2 trigger selection were to maintain un-prescaled single-electron and single-muon trigger p_{T} thresholds of around 25 GeV, to ensure the majority of events collected are leptonic W-boson and Z-boson decays; these triggers are also used in this thesis as described in Section 3.1.1.

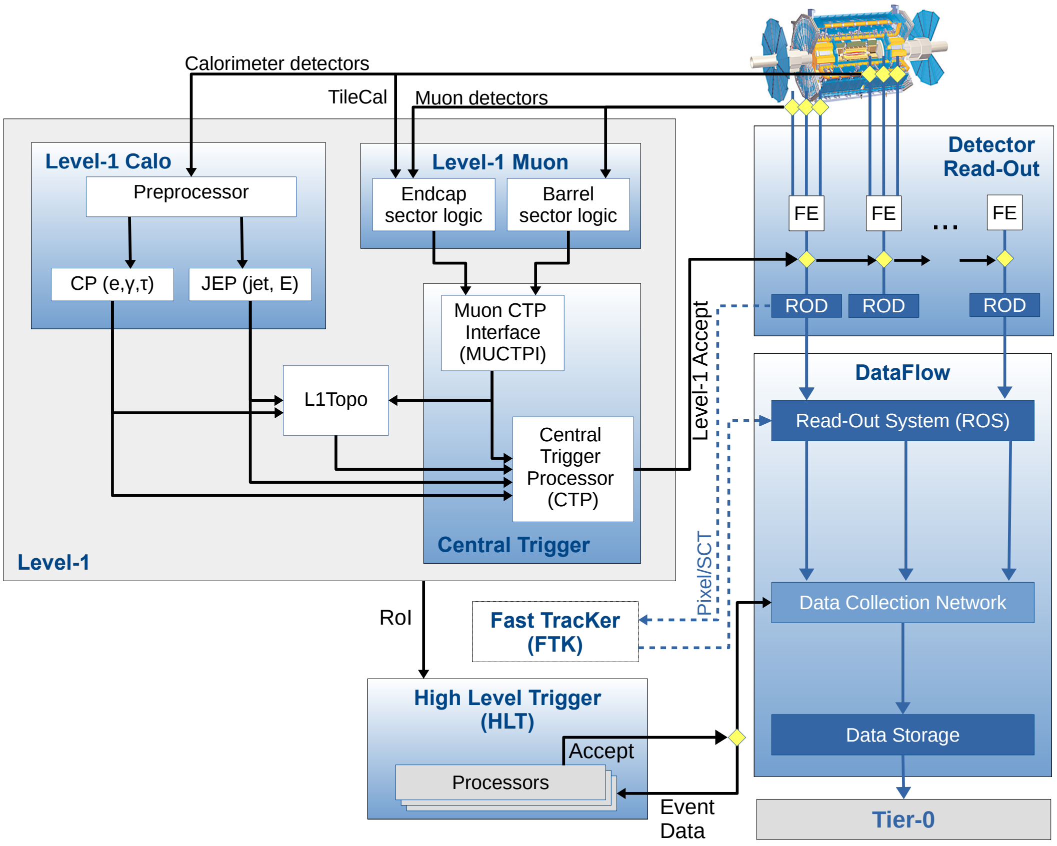


Figure 2.17: The trigger and data acquisition (TDAQ) system for the Run 2 data taking period of the ATLAS experiment [82].

The Level-1 trigger

The L1 trigger is a hardware-based trigger implemented in firmware running on custom electronics, and issue triggers based on reduced-granularity information from the calorimeters and the fast-triggering muon chambers. The L1 decision reaches the front-end electronics within 2.5 μs after the proton bunch crossings (occurring at 40 MHz), and accepts events of interest at a maximum rate of 100 kHz.

The *L1 calorimeter* (L1-Calo) trigger takes signals from the calorimeter as input, digitizes and calibrates them (pre-processing), and sends it to the cluster processor (CP), and the jet/energy sum processor (JEP). The CP identifies electron, photon, and τ-lepton candidates above a set threshold, while the JEP identifies jet candidates and produces global sums of total transverse energy (E_T^{total}) and missing transverse energy ($E_{\text{T}}^{\text{miss}}$).

The *L1 muon* (L1-Muon) trigger uses hits from the RPC's in the barrel, and TGC's in the end-caps to determine the deviation of the hit-pattern from a muon of infinite momentum (i.e a straight line). To reduce the trigger rate in the end-cap regions, the L1-Muon trigger applies coincidence requirements between the outer-inner TGC chambers, and between the TGC-TileCal. Before sending the trigger information downstream, the L1-Muon Central Trigger Processor Interface (MUCTPI) is used to pre-sort the trigger input data.

The *L1 topological* (L1-Topo) trigger was included for the Run 2 data taking period, as it allows topological angular and kinematic selections providing a large improvement in the background rejection with minimal to no signal loss [83]. The L1 trigger decision is made by the *central trigger processor* (CTP), which receives inputs from the L1-Calo, L1-Muon through MUCTPI, L1-Topo, as well as several detector subsystems such as LUCID-2, ZDC, and the Minimum Bias Trigger Scintillators[8] (MBTS) [84].

For each L1-accepted event, the event data from the front-end (FE) electronics of all the ATLAS sub-detectors are read out. The data is first sent to the read-out drivers (ROD's) that perform the initial processing and formatting of the data, and then to the read-out system (ROS's) to buffer the data. The data from the ROS is sent to the HLT only when requested by the HLT. Further, the L1 triggers identifies *regions-of-interest* (ROI's) in η and ϕ within the detector to be investigated by the HLT.

The high level trigger

The HLT is the second stage of the trigger chain, and is a software-based trigger which runs on a server farm with 5×10^4 processing units. A reconstruction sequence uses dedicated fast-trigger algorithms to provide early rejection, while more precise, CPU-intensive algorithms (similar to ones used in the offline reconstruction) are used for the final selection. The HLT software is largely based on the ATLAS offline software *Athena* [85], which is also the main software used in the analysis described in this thesis. The HLT decision is made within an average time of 500 ms and accepts data at an average rate of 1.2 kHz. Finally, the data gets stored at a throughput of 1.2 GBps into permanent storage for further *offline* analysis.

---

[8] The MBTS gives primary triggers for selecting events with the smallest bias for low luminosity runs.

Chapter 3

Data and simulated samples

The ATLAS experiment collects data from proton-proton collisions based on different triggers. From the collected data, the ATLAS offline software reconstructs physics objects as detailed in Chapter 4. The various physics processes from proton-proton collisions are simulated using Monte-Carlo methods using different event generators. The simulated list of particles are then passed through the ATLAS detector simulation and digitized to simulate the detector response. Then the same reconstruction software used for data is used to reconstruct the digitized detector responses into physics objects. The flow of processing data and the simulated samples are illustrated in the diagrams of Figure 3.1.

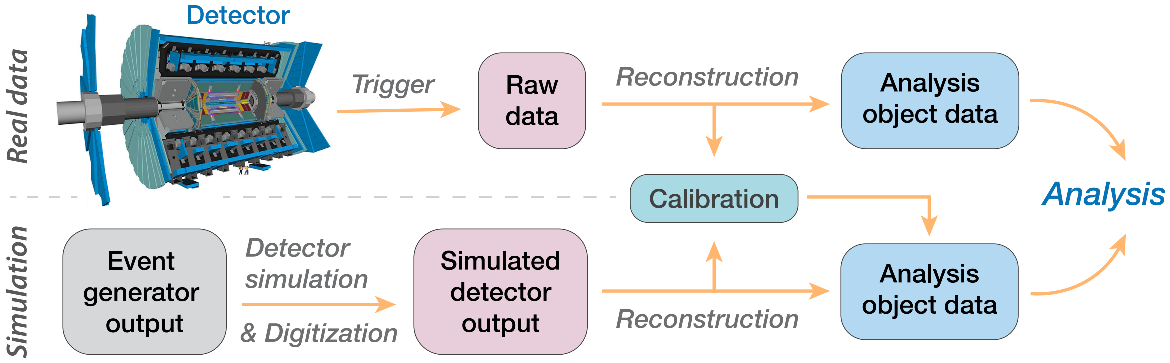


Figure 3.1: The flow of real data and simulated data in the ATLAS experiment.

3.1 The data samples

The data used in this analysis was collected during the Run 2 of the LHC, in the years 2015, 2016, 2017, and 2018. During the data taking period, it was made sure that all detector components were working in a good operating condition, and the events selected were those which pass stringent data quality checks [56]. After data quality selections, the integrated luminosity correspond to 140.1±1.2 fb^{-1} at $\sqrt{s} = 13$ TeV.

3.1.1 Trigger selection

The candidate events for the $VH, H \to b\bar{b}/c\bar{c}$ process are selected from events that are triggered by the leptonic decay of the vector boson. Table 3.1 summarizes the list of triggers used in the analysis. Additional explanation on the trigger selection can be found in the Appendix C.

Table 3.1: The trigger selections and the events selections for reconstructing the vector boson.

Lepton channel	Final state leptons	Trigger
Zero charged lepton channel	Neutrinos	$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger
One charged lepton channel	Electrons	Single electron trigger
	Muons ($p_{\mathrm{T}}^{V} < 150$ GeV) Muons ($p_{\mathrm{T}}^{V} > 150$ GeV)	Single muon trigger $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger
	τ-leptons	$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger
Two charged lepton channel	Electrons	Single electron trigger
	Muons ($p_{\mathrm{T}}^{V} < 250$ GeV) Muons ($p_{\mathrm{T}}^{V} > 250$ GeV)	Single muon trigger $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger

In the zero charged-lepton channel, since the target decay mode of the vector boson is $Z \to \nu\bar{\nu}$, events are selected using the lowest un-prescaled $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers. In the one charged-lepton channel, the target decay mode of the vector boson is $W^{\pm} \to e^{\pm}\nu_e$, $W^{\pm} \to \mu^{\pm}\nu_{\mu}$, or $W^{\pm} \to \tau^{\pm}\nu_{\tau}$ (where the τ lepton decays hadronically), and in the two charged-lepton channel, the target vector boson decay is $Z \to e^+e^-$ or $Z \to \mu^+\mu^-$. Hence, in the one and two charged-lepton channels, events are collected using the lowest un-prescaled single electron triggers and single muon triggers[1]. Further, $Z \to \mu^+\mu^-$ events with large transverse momentum of the vector boson (p_{T}^{V}) are triggered using the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger as explained below.

$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the momentum imbalance in the transverse direction, and is calculated using a set of energy deposits in the calorimeter. Muons are treated similar to neutrinos in the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger calculations since they deposit only a small fraction of their total energy in the calorimeter. Therefore, events having muons with high p_{T} are collected using the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger for the one and two charged-lepton channels. At L1, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated based on digitized signals from calorimeter cells, and accepts events with $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV mostly throughout the Run 2 data taking period.

The HLT recalculates $E_{\mathrm{T}}^{\mathrm{miss}}$ for events accepted by the L1, where different algorithms are utilized to mitigate pileup effects, and improve the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution [86]. The overall (L1 & HLT) $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger chains have different thresholds from 2015 to 2018, varying between 50 GeV to 110 GeV.

---

[1] Adding di-lepton triggers was studied in the previous $VH, H \to b\bar{b}$ analysis, and this would only bring 2% gain on the signal acceptance. To avoid complexity in calibrating a combination of single and di-lepton triggers, only single lepton triggers are used in this analysis.

However, events used in the analysis are required to have passed a trigger threshold of $E_{\mathrm{T}}^{\mathrm{miss}} = 150$ GeV to ensure a good trigger efficiency; the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency is 90% at 150 GeV, and reaches 100% at around 200 GeV.

For leptonically decaying τ-leptons, the final states are not distinguishable from the electron and muon channels. However, hadronically decaying τ-leptons have a $E_{\mathrm{T}}^{\mathrm{miss}}$ accounting for the neutrino in the τ-lepton decay, and several jets in the final state. Hence, hadronically decaying τ-leptons events are collected using the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers.

Single electron trigger

The reconstructed energy in the transverse direction (E_{T}) at L1 is required to satisfy a nominal threshold of $E_{\mathrm{T}} > 20$ GeV in 2015, and $E_{\mathrm{T}} > 22$ GeV in 2016 to 2018 [87]. Hadron activity is suppressed using requirements on the reconstructed E_{T}. Isolation criteria (in 2016-2018) were required to select *prompt* electrons from the primary vertex. The hadron activity selection and the isolation criteria are not applied at L1 for reconstructed $E_{\mathrm{T}} > 50$ GeV.

The HLT electron reconstruction is based on two steps: a fast reconstruction step, where electrons are required to have calorimeter clusters matched to tracks reconstructed from a fast tracking algorithm, and a precision reconstruction step, where precision tracks within the region-of-interest (RoI) are extrapolated to the second layer of the electromagnetic calorimeter and required to match clusters within $\Delta\eta < 0.05$ and $\Delta\phi < 0.05$ radians. Then electron candidates are selected using a multivariate likelihood discriminant giving four operating points, with varying signal to background rejection criteria similar to the offline criteria discussed in Section 4.2.

The electrons that are selected from the full (L1 & HLT) trigger chain are required to have a E_{T} threshold of 24 GeV for 2015, and 26 GeV for 2016 to 2018. The trigger efficiency for electrons is typically around 70% at the E_{T} thresholds, and reach more than[2] 90% for $E_{\mathrm{T}} > 50$ GeV.

Single muon trigger

In the L1 muon trigger, the p_{T} of the muon is estimated from the degree of deviation of hit patterns in the muon chambers compared to a muon of infinite momentum, based on several programmable p_{T} thresholds [78]. The number of muons passing each threshold is used in the conditions for the global L1 trigger. The L1 trigger decision in the barrel region is based on the coincidence of hits from three concentric RPC stations, for three different p_{T} thresholds [3]. The L1 trigger decision on the end-cap region is based on the coincidence of hits in the TGC stations in the middle layer (i.e the Big Wheel). For the events which are accepted by L1, the p_{T} thresholds and the corresponding ROIs are sent to the HLT.

The muon HLT selects events in two stages similar to the electron HLT; fast reconstruction algorithms are executed in the first stage, followed by muon algorithms similar to the offline reconstruction. The RoIs identified by the L1 trigger enables the fast algorithms to select regions of the detector where interesting features reside. The muon standard-alone (SA) algorithms refines the L1 candidate by constructing the muon track using MDT hit information within the RoI. The muon track is then back-extrapolated to the interaction point, and combined with tracks recon-

---

[2] 2016 is an exception, where the trigger efficiencies are slightly lower.

[3] Note that there are three more (low) p_{T} thresholds that are not mentioned since they are not used in this analysis.

structed from the inner detector to form a combined (CB) muon candidate. If no CB muons are formed, muon candidates are searched for by extrapolating ID tracks to the muon detectors.

The default L1 single-muon triggers used to collect the data used in the analysis were unprescaled, and had a p_{T} threshold of 20 GeV. At the HLT, these triggers require at least one CB muon with $p_{\mathrm{T}} \geq 26$ GeV, or $p_{\mathrm{T}} \geq 50$ GeV. For the trigger requiring a threshold of 26 GeV, a medium isolation was also required. However, in 2015 another trigger with a L1 p_{T} threshold of 15 GeV, and HLT p_{T} threshold of 20 GeV was also used to collect data. Figure 3.2 shows the event display of a candidate $ZH \rightarrow \mu^+\mu^- c\bar{c}$ event.

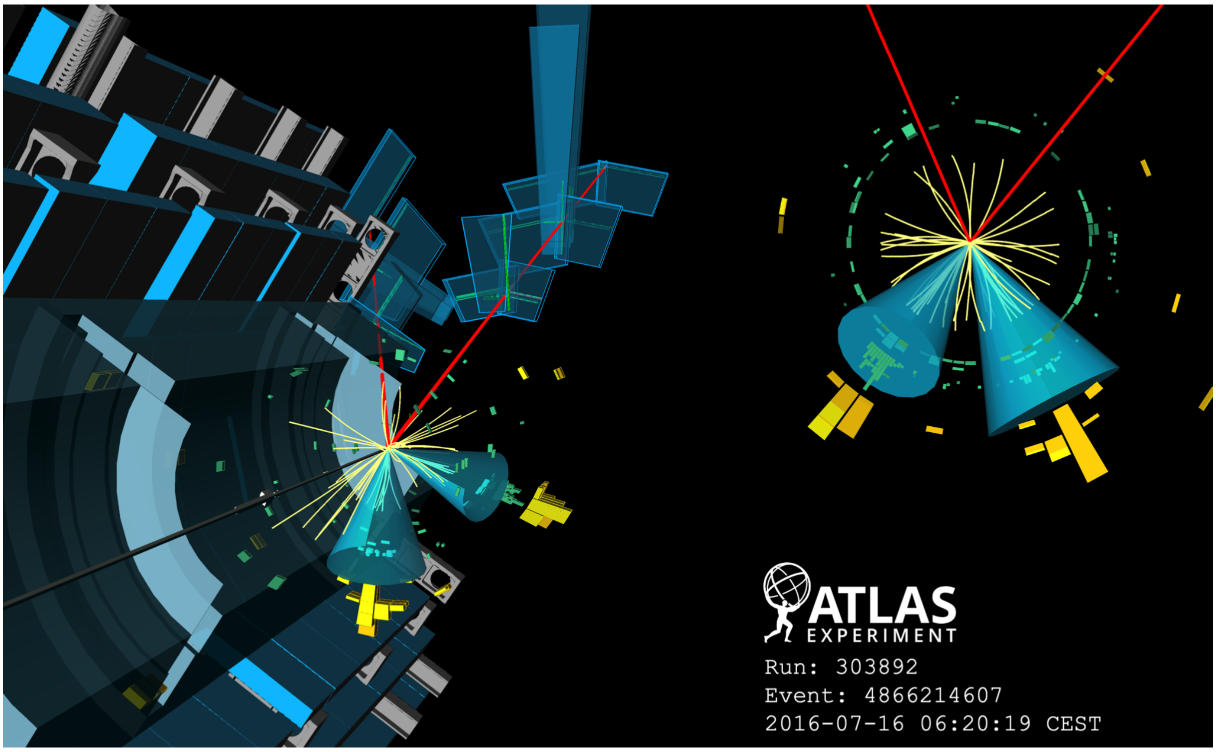


Figure 3.2: Candidate event for the $ZH \rightarrow \mu^+\mu^- c\bar{c}$ process [88].

3.2 The simulated samples

Event simulations are based on Monte Carlo (MC) methods, where different MC generators are used to model the physics processes used in the analysis. The simulated list of particles from the MC generators are then passed through the ATLAS detector simulation [89, 90], and then reconstructed with Athena (the ATLAS offline reconstruction software) [85].

Figure 3.3 is an illustration of the various steps involved in the simulation of a proton-proton collision event.

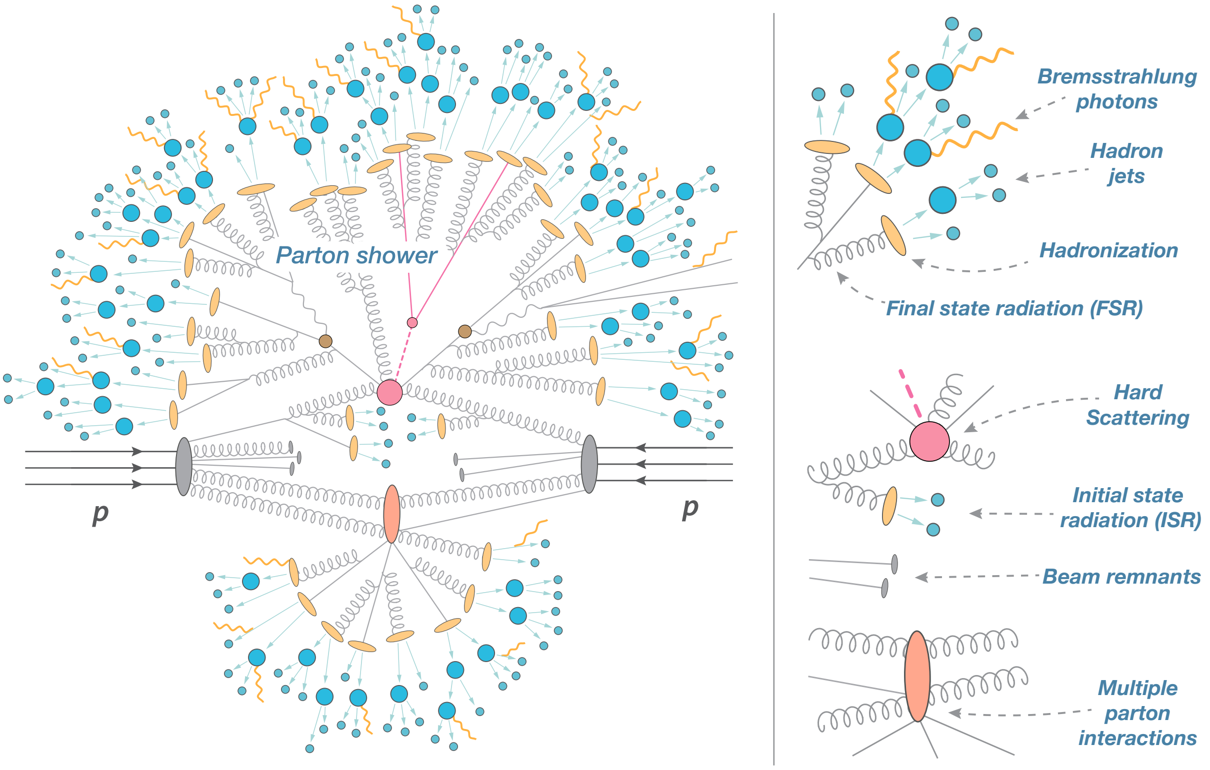


Figure 3.3: The illustration of a proton-proton collision event in the point of view of a Monte-Carlo event generator. The illustration is based on the original source at [91].

The hard scatter process occurs from a high momentum transfer Q^2 between the colliding partons in a proton, and is usually the process of interest due to the possibility of creating massive particles like the Higgs boson. At the LHC, several hard scattering processes occur in one bunch crossing, which is referred to as pile up, and is simulated by overlaying several hard scatter processes [92]. Before the hard scattering, partons can undergo initial state radiation (ISR), while final state radiation (FSR) occurs for hard scattered partons.

Highly energetic particles produced from the hard scatter process are then used for parton showering (PS). Such particles do not undergo hadronisation directly since they are still too energetic (at perturbative QCD scales), and undergoes energy loss through FSR till it has low enough energy (at 1 GeV order) to undergo hadronisation. At the hadronisation stage, partons are bound together to form hadrons, which is simulated mainly using the Lund string model [93] or the cluster fragmentation model [94].

The remaining partons in the proton that do not undergo hard scattering (beam remnants), and additional hard scattering in the same proton-proton collisions (multiple parton interactions, or MPI), are classified as the underlying event (UE) [95]. Parton distribution function (PDF) $f_i^H(x, Q^2)$ describes the probability that a parton i of hadron H participates in a hard scattering interaction, as a function of the partons momentum x (with respect to H) and the square of momentum transfer Q^2 during interaction. Figure 3.4 shows an example of a parton distribution function when $Q = 10$ GeV.

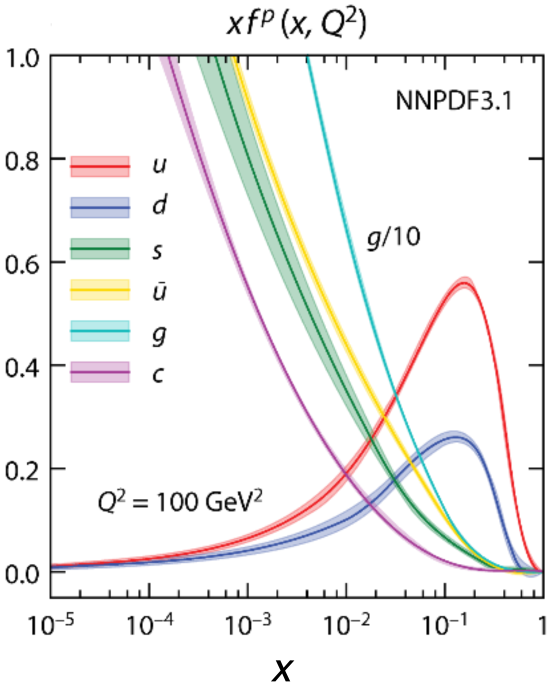


Figure 3.4: Parton distribution functions (PDF's) for different partons inside a proton, as a function of the partons momentum fraction x. The PDF's shown are the NNPDF3.1 PDF set at NNLO, for an energy scale of $Q = 10$ GeV. The image is modified from [96].

The QCD factorization theorem states that the hadronic interaction cross section is a convolution of the partonic cross sections such that,

$$\sigma_{AB\to k} = \iint \sum_{i\in A}\sum_{j\in B} f_i^A(x_1, Q^2) f_j^B(x_2, Q^2)\sigma_{ij\to k}\, dx_1 dx_2, \tag{3.1}$$

where $f_i^A(x_1, Q^2)$, $f_j^B(x_2, Q^2)$ are the PDF's of the two colliding partons A and B, and $\sigma_{ij\to k}$ is the production cross section of particle k through the interaction of two partons i and j, obtained using Fermi's golden rule,

$$\sigma_{ij\to k} \propto \int |\mathcal{M}_{ij\to k}(\boldsymbol{\Phi_k})|^2\, d\boldsymbol{\Phi_k}, \tag{3.2}$$

where $\boldsymbol{\Phi_k}$ is the allowed phase space for the final decay k. $|\mathcal{M}_{ij\to k}(\boldsymbol{\Phi_k})|$ is the transition amplitude which is proportional to the matrix element (ME), and is calculated inclusively with higher order perturbative corrections. Equation 3.1 sums over all partons in the collision between A and B that can result in the $ij \to k$ process.

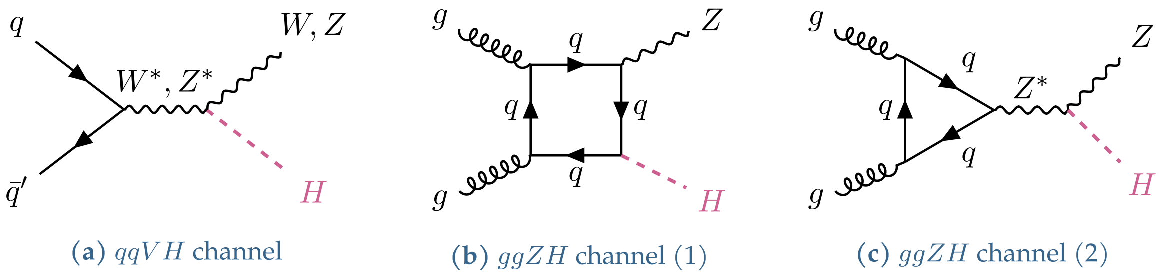
The factorization scale μ_F is used for the consistent treatment of perturbative and non perturbative effects when calculating the hadronic interaction cross section of Equation 3.1, where μ_F modifies the PDF as $f_i^H(x, Q^2, \mu_F^2)$. The renormalization scale μ_R is used to describe the energy scale of QCD interactions, where the strong coupling constant α_s is expressed as a function of the renormalization scale as $\alpha_s(\mu_R^2)$, and μ_R is taken to be close to the scale of the momentum transfer Q in a given process [17]. The ME is computed for a fixed μ_F and μ_R values, but samples by altering the μ_F and μ_R values are also created for estimating modeling uncertainties as discussed in Section 8.2.

The different event generation stages described above can be simulated using one MC event generator, or using multiple MC event generators. The MC event generators used in the current analysis are listed in Table 3.2, and they are explained in detail for the signal and background categories in the following sections. There are also alternate signal and background MC samples that are generated to asses the modeling uncertainties, as detailed in Table 8.8.

Table 3.2: The nominal MC generators used for simulating the signal and background processes. ME, PDF, PS, Had. and UE stands for the matrix element, parton distribution function, parton shower model, hadronisation model, and the underlying event, respectively.

Process	ME generator	ME PDF	PS and Had.	UE tune
$qq \to VH$	POWHEG BOX v2 + MINLO	NNPDF3.0NLO	PYTHIA 8.245	AZNLO
$gg \to VH$	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.307	AZNLO
Top-quark	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230	A14
V+jets	SHERPA 2.2.11	NNPDF3.0NNLO	SHERPA 2.2.11	Default
$qq \to VV$	SHERPA 2.2.11	NNPDF3.0NNLO	SHERPA 2.2.11	Default
$gg \to VV$	SHERPA 2.2.2	NNPDF3.0NNLO	SHERPA 2.2.2	Default

3.2.1 The signal processes



(a) $qqVH$ channel (b) $ggZH$ channel (1) (c) $ggZH$ channel (2)

Figure 3.5: The $qqVH$ and $ggZH$ signal processes.

The $VH, H \to b\bar{b}/c\bar{c}$ signal processes are classified into the three charged lepton channels of Figure 1.11. The dominant production modes are the quark initiated processes of Figure 3.5(a), but in the zero and the two charged lepton channel, there is also the contribution of the gluon initiated processes depicted in Figures 3.5(b) and 3.5(c).

The cross section of the VH signal processes are summarized in Table 3.3, where it is seen that the gluon-initiated $ggZH$ production cross-section is roughly six times smaller compared to the quark initiated $qqZH$ production cross-section.

Table 3.3: The production cross sections of the WH and ZH processes at a Higgs boson mass $M_H = 125$ GeV, in pp collisions at a C.O.M energy of 13 TeV [16]. Note that each cross section given in this table is the sum of the cross sections for all three lepton flavors.

Production process	Cross section σ	Order
$qq \to W(\to l\nu)H$	462.27 fb	NNLO(QCD)+ NLO(EW)
$qq \to Z(\to \nu\bar{\nu})H$	153.05 fb	
$qq \to Z(\to l^+l^-)H$	77.04 fb	
$gg \to Z(\to \nu\bar{\nu})H$	24.57 fb	NLO + NLL
$gg \to Z(\to l^+l^-)H$	12.42 fb	NLO + NLL

The signal production processes are simulated using the Powheg generator [97] with multi-scale improved NLO (MiNLO), interfaced with the Pythia 8 MC generator [98]. The Powheg generator models the matrix element with the NNPDF3.0NLO PDF set [99], and reweighted to the PDF4LHC15NLO PDF set. The Pythia 8 generator is used to model the parton shower, underlying event, and multiple parton interactions, applying the tuned parameters of the AZNLO tune (defined in [100]) due to the better agreement with measurement.

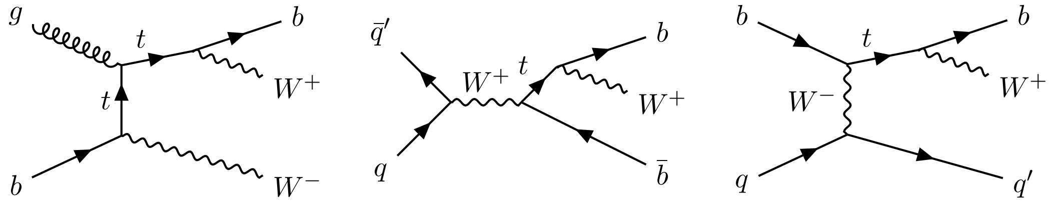
When modeling the signal processes, the mass of the Higgs boson is set to 125 GeV, and the decay of the Higgs boson is forced to the $H \to b\bar{b}$ or $H \to c\bar{c}$ modes.

3.2.2 The background processes

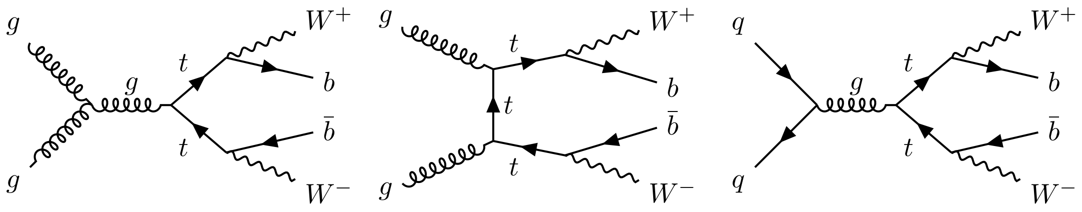
Most of the the background processes discussed in Section 1.4.1 are modeled using MC methods. The two exceptions are the QCD multi-jet process, and the top quark decay to two leptons in the $VH, H \to b\bar{b}$ phase space, which are derived using data driven methods.

Top quark processes

The top quark processes are mainly categorized into the two categories; the single top quark processes depicted in Figures 3.6(a)∼(c), and top pair production processes (referred to as $\boldsymbol{t\bar{t}}$) depicted in Figures 3.6(d)∼(f). Top-quark backgrounds mainly enter the $VH, H \to b\bar{b}$ phase space when the selected events contain the two b-jets from the top-quark pair decay, and either one (two) leptons from the decay of a W-boson (two W-bosons). In the $VH, H \to c\bar{c}$ phase space, different top-quark processes can be picked up due to the limitation in separating c-jets from b-jets and light-quark jets. An example of a top-quark process entering the $VH, H \to c\bar{c}$ phase space is when selecting a c-jet from a $W \to cs$ decay, a b-jet from a top-decay, and the leptonic decay of the other W-boson.



(a) Single top Wt (b) Single top s-channel (c) Single top t-channel



(d) $t\bar{t}$ process (1) (e) $t\bar{t}$ process (2) (f) $t\bar{t}$ process (3)

Figure 3.6: The single and double top quark background processes.

Table 3.4 shows the production cross section of the top quark processes, where the dominant process is $t\bar{t}$.

Table 3.4: The production cross sections of the top quark processes [8, 101, 102, 103]

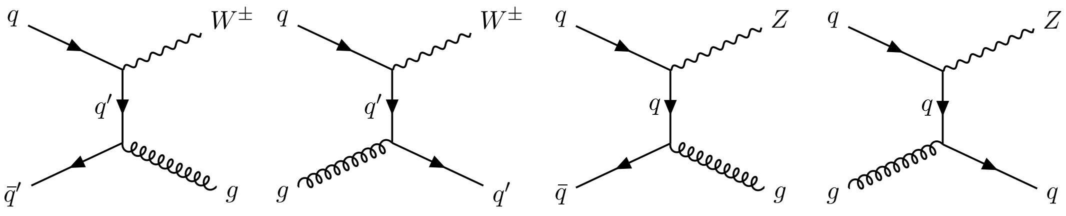
Production process	**Cross section σ**	**Order**
$t\bar{t}$ process	832 pb	NNLO + NNLL
Top t-channel process	214.2 pb	NNLO
Top Wt process	79.3 pb	Approx. NNLO
Top s-channel process	10.32 pb	NLO

Top pair production ($t\bar{t}$) and single top processes are simulated using Powheg generator interfaced with the Pythia 8.230 generator. Similar to the VH signal, the Powheg generator models the matrix element with the NNPDF3.0NLO PDF set. The parton shower, hadronisation, and the underlying event are modeled using Pythia with the NNPDF3.0NLO PDF's and the A14 set of tuned parameters[104].

The interference between the $t\bar{t}$ and single-top Wt processes are treated in two ways when modelling [105]. One way is using the *diagram subtraction* (DS) method, where the interference is removed from the NLO cross section by subtracting a $t\bar{t}$ contribution term. Another way is the *diagram reduction* (DR) method, where the resonant $t\bar{t}$ diagrams causing the interference are removed from the single-top Wt ME amplitudes. The current analysis uses single-top Wt samples created using the DS method, since DS is seen to have better agreement with data compared to DR.

Vector boson plus jets processes

The vector boson plus jets, referred to as V+jets, is the dominant background process in the $VH, H \to c\bar{c}$ analysis. The V+jets processes are mainly split into two categories; the Z+jets and the W+jets processes, as depicted in Figure 3.7.



(a) W + jets process (1) (b) W + jets process (2) (c) Z + jets process (1) (d) Z + jets process (2)

Figure 3.7: A single vector boson plus jets background processes.

The production cross sections of the V+jets processes are given in Table 3.5. The one charged lepton channel is dominated by the W+jets processes, while two charged lepton channel is dominated by Z+jets processes, due to the similar final states with the $VH, H \to b\bar{b}/c\bar{c}$ signal process. The zero charged lepton channel has contributions from Z+jets background processes, as well as the contributions from W+jets processes where the electron or muon of the W-boson decay is not identified. For V+jets, the $b\bar{b}$ and $c\bar{c}$ final states are mainly originating from gluon splitting.

Table 3.5: The production cross sections of the single vector boson plus jets process at NNLO QCD [106]. For $W^\pm \to l^\pm\nu$ and $Z \to l^+l^-$, the cross sections shown are for each lepton flavor. For $Z \to \nu\bar{\nu}$ + jets, the cross section shown is the total for all flavors obtained by scaling $Z \to l^+l^-$ cross-section with the measured branching ratios of the Z-boson decay[17].

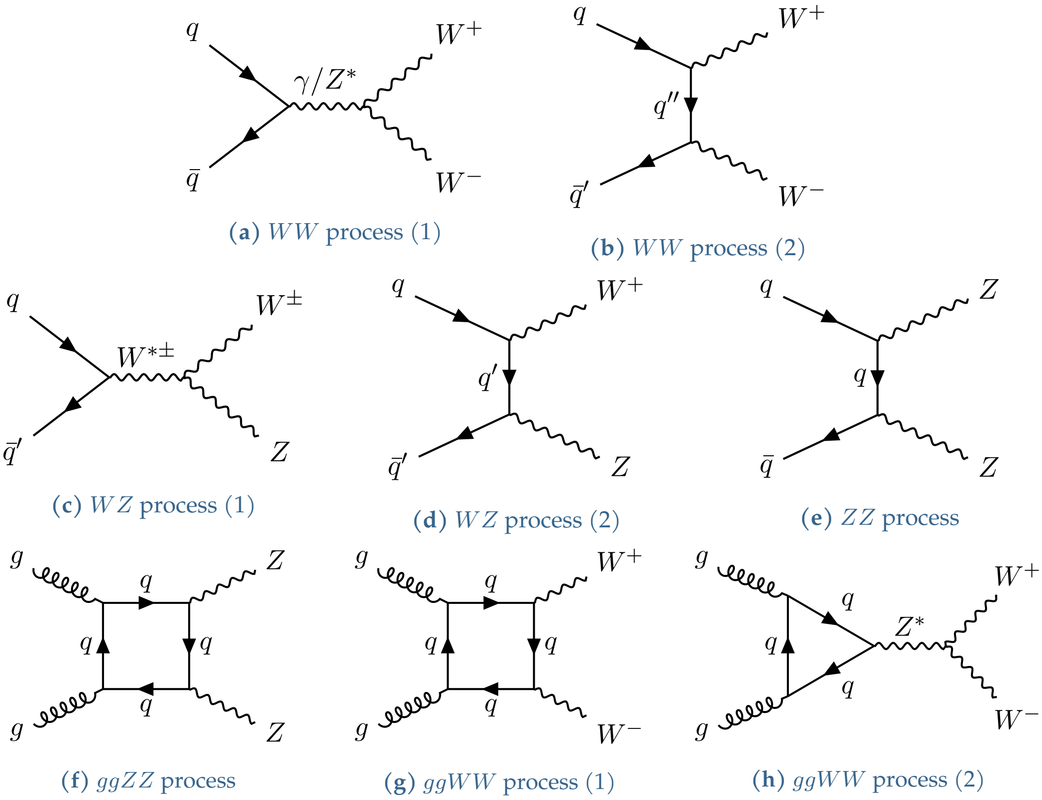
Production process	Cross section σ	Order
$Z \to \nu\bar{\nu}$ + jets	12282 pb	NNLO
$W^\pm \to l^\pm\nu$ + jets	20080 pb	NNLO
$Z \to l^+l^-$ + jets	2067 pb	NNLO

The V+jets processes are simulated using the SHERPA 2.2.11 generator [106, 107]. The matrix elements are calculated to NLO (LO) accuracy for ≤ 2 jets (between 3 to 5 jets) using the SHERPA 2.2.11 generator, utilizing the *five-flavor* scheme; in the five-flavor scheme, the light-quarks u, d, s, and the heavy quarks c and b are treated as massless, and included in the ME and PDF calculations [108]. For simulating the parton shower and the hadronisation, the default SHERPA parton shower [109] and hadronisation [110] models are used, with the built in Hessian NNPDF3.0NNLO PDF set.

Virtual loop NLO electroweak (EW) corrections are combined with the NLO QCD prediction using the additive scheme, since this shows the smallest deviation from the NLO-QCD only prediction. Another two scheme of combination; multiplicative and exponentiated schemes are used to assess the modeling systematic uncertainties as discussed in Section 8.2.3.

Diboson processes

The diboson is a sub-dominant background process, and can be mainly categorized into five types; the WW, WZ, ZZ, $ggZZ$, and $ggWW$, as depicted in Figure 3.8. Since diboson has a similar final state to the VH signal, there are contributions from diboson processes in all the charged lepton channels.



(a) WW process (1)
(b) WW process (2)
(c) WZ process (1)
(d) WZ process (2)
(e) ZZ process
(f) $ggZZ$ process
(g) $ggWW$ process (1)
(h) $ggWW$ process (2)

Figure 3.8: Double vector boson background processes.

Table 3.6 shows the production cross-sections for the diboson processes.

Table 3.6: The production cross sections of the diboson processes [111, 112].

Production process	Cross section σ	Order
$qq \to W^+W^-$	109.28 pb	NLO
$qq \to W^\pm Z$	20.85 pb	NLO
$qq \to ZZ$	17.51 pb	NLO
$gg \to W^+W^-$	151.6 fb	NLO
$gg \to ZZ$	4.27 fb	NLO

The $q\bar{q}$ initiated diboson processes are simulated using the Sherpa 2.2.11 generator, while the Sherpa 2.2.2 is used for simulating the gg initiated process. The NNPDF3.0NNLO PDF set is used for both the matrix element calculation and parton shower modeling.

3.2.3 The ATLAS detector simulation and digitization

After the event generation step, the generated events are passed through the ATLAS detector simulation, which is based on the Geant4 simulation toolkit [113], to simulate the detector response to physics interactions with the various sub-detectors of the ATLAS detector.

The ATLAS detector geometry is constructed in the Geant4 format. Since the detector simulation, digitization and the reconstruction are done separately, the same detector geometry is used. Figure 3.9 shows an example of a visualization of the ATLAS geometry constructed from the GeoModel description, for a Higgs boson event decaying to four muons.

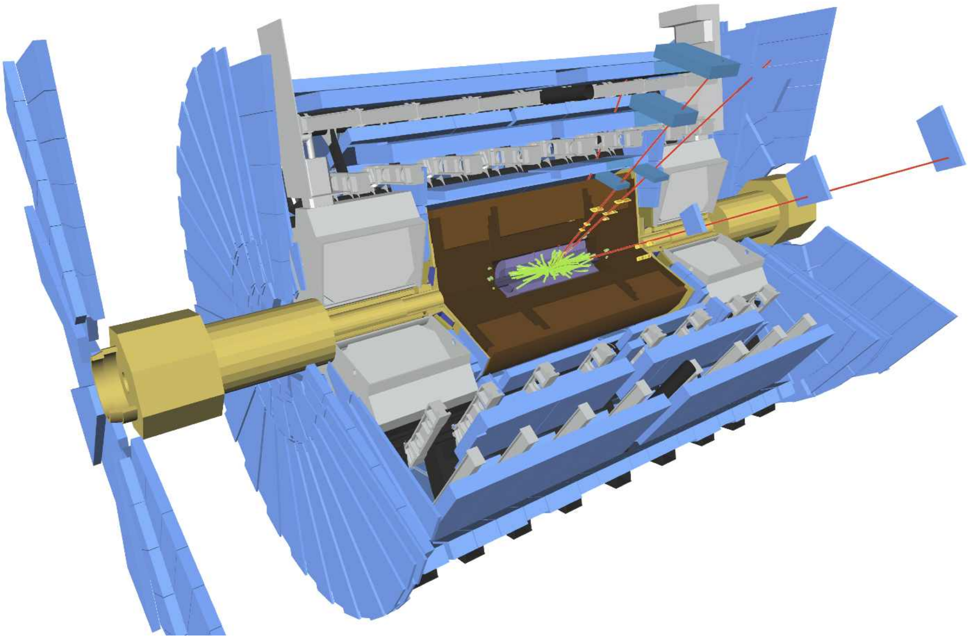


Figure 3.9: Visualization of a Higgs boson to four muon event in the ATLAS detector simulation. The inner detector track are shown in green, the energy deposits in the calorimeters are shown in yellow, and the muon tracks are shown in red [89].

In the detector simulation, transformations such as the vertex position smearing in the proton-proton collision region is made to represent real collision events. While geometrical cuts such as simulating only the $|\eta| < 6$ region are crucial to save computation time, cuts specific to particle production are also important. Neutrons are removed after a time cut of 150 ns, neutrinos are removed as soon as they are created, and range cuts (usually 1 mm) control the creation of secondary electrons and photons during bremsstrahlung and pair creation. The output of the detector simulation is a hit file, containing the simulation configuration, the truth information of the particle, and a collection of hits in each sub-detector; the hits are records of energy deposition with position and time information.

The Geant4 detector response in the form of hits in the sub-detectors, are then passed through a simulation of the readout electronics and digitized. Typically digits are produce when the voltage or current on a particular readout channel rises above a pre-configured threshold. While some sub-detectors will record a signal shape in time, others will simply record a hit if the signal goes above the threshold in a particular time window. The different properties of the sub-detectors like charge collection, electronic noise and channel-by-channel dependencies are also modelled in sub-detector specific digitization algorithms. These algorithms are tuned using measurements from lab tests, beam tests, cosmic ray tests, as well as collision data. Unresponsive channels and noise rates are referred from a database, is considered to reproduce conditions seen in a particular collision run [89].

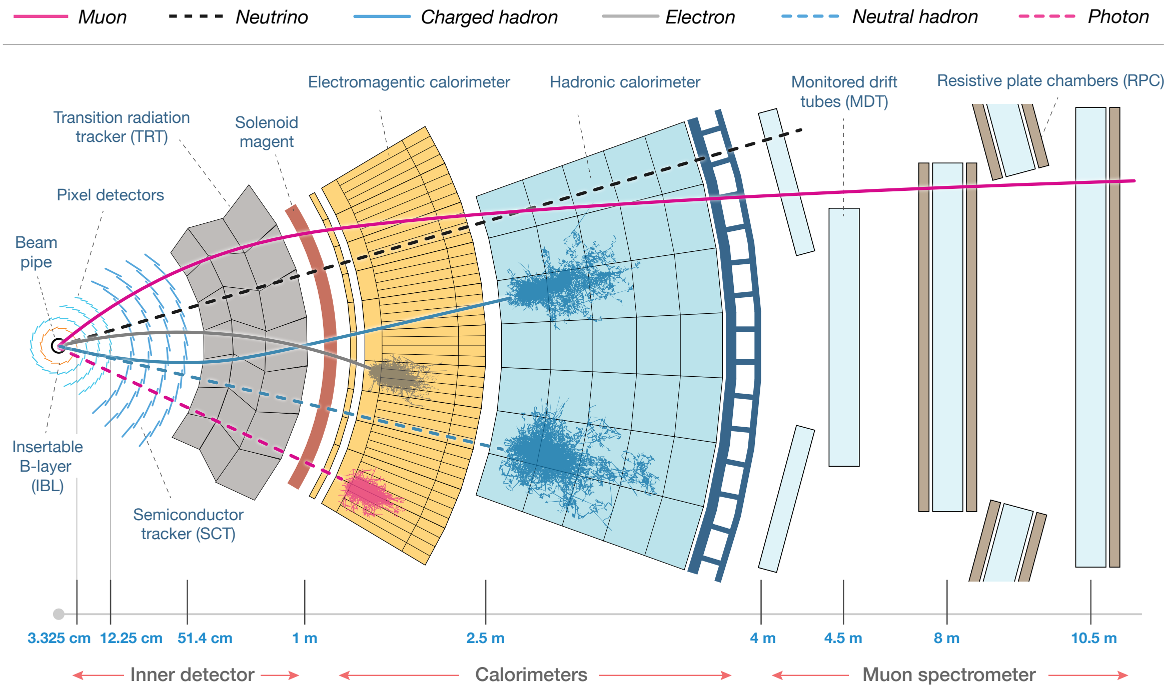
In addition to the multiple hard-scattering pile-up, there are also pile-up effects from the beam-halo resulting from interactions of the proton beams and upstream accelerator elements, cavern background due to particles such as neutrons escaping the ATLAS detector and causing spallation, etc. Most sub-detector responses are affected by such pile-up effects, and therefore pile-up effects also need to be simulated in the detector simulation; this is done by overlaying hits from various simulated events.

In the last step, data and simulated events undergo object reconstruction, which is explained in depth in Chapter 4.

Chapter 4

Physics object reconstruction

This chapter explains the reconstruction of physics objects used in the ATLAS experiment, as illustrated in Figure 4.1. The object reconstruction criteria specific to the $VH, H \to b\bar{b}/c\bar{c}$ analysis is described at the end of this chapter, in Section 4.9



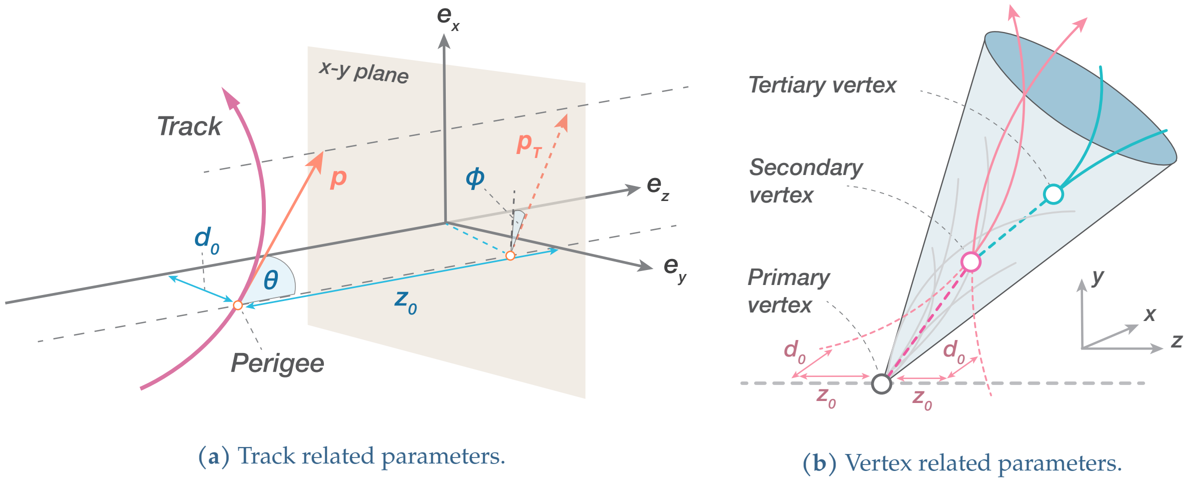
Note: This illustration shows a section of the barrel region of the ATLAS detector, and the scaling is non-linear in order to show all sub-detector systems.

Figure 4.1: How particles are detected in the ATLAS experiment.

The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, electrons, muons, and hadronically decaying τ-leptons are used in the reconstruction of the vector boson candidate, while jets (after jet flavor tagging) are used in the reconstruction of the Higgs boson candidate. The reconstruction of photons are not discussed in this chapter, since photons are not used in the $VH, H \to b\bar{b}/c\bar{c}$ analysis, but some description can be found in Appendix D.2 and [114].

4.1 Tracks and vertices

Tracks measured in the inner detector are used to reconstruct the interaction vertices. The highest sum of squared transverse momenta of associated tracks is selected as the *primary vertex* of the hard scattering process.



(a) Track related parameters.

(b) Vertex related parameters.

Figure 4.2: An illustration of the definition of track and vertex related parameters used in ATLAS.

4.1.1 Track reconstruction and selection

In the ATLAS inner detector, charged particles from the proton-proton collisions undergo helical trajectories due to the 2 T solenoidal magnetic field applied. Therefore charged particle tracks are described using five helical parameters as illustrated in Figure 4.2(a), and explained below.

- **Transverse impact parameter d_0** is the transverse distance between the perigee and the beam axis. The perigee is the point of closet approach of a track to the beam axis.
- **Longitudinal impact parameter z_0** is the longitudinal distance between the perigee and the transverse plane (i.e $x - y$ plane) containing the primary vertex.
- **Track azimuthal angle ϕ** is the azimuthal angle of the track direction at the perigee.
- **Track polar angle θ** is the polar angle of the track direction at the perigee.
- **Charge-momentum ratio q/p** is defined by the curvature of the track due to the 2 T solenoidal magnetic field.

The track reconstruction will first reconstruct tracks in the pixel detector and the semi-conductor tracker (SCT), before extending it to the transition radiation tracker (TRT) [115]. The first step is a clusterization of pixel and SCT hits, from which 3D *space-points*[1] are created. Next, track seeds are formed from the sets of three space-points[2]. A combinatorial Kalman filter then builds track candidates from chosen track seeds, also incorporating the additional space-points from

---

[1] In the pixel detector, a cluster of two to three pixels represent a space point, while in SCT, clusters from both sides of a strip detector are combined to obtain a 3D measurement

[2] This maximizes the possible number of tack combinations since a requirement on more space points will reduce the amount of combinations to build track candidates.

other pixel/SCT layers. Further selection on track quality, such as a χ^2 fit and a shared-cluster criterion, is required to select tracks more likely to represent a charged primary particle.

Fake track reconstruction in the inner detector can happen due to combinatorial effects from high pile-up, when track seeds are formed. To suppress fake tracks, tracks failing to meet a set of quality criteria will be rejected; a track is required to have $p_{\mathrm{T}} > 400$ MeV, $|\eta| < 2.5$, ≥ 7 pixel and SCT clusters, ≤ 2 holes in the pixel and SCT detectors (with ≤ 1 hole in the pixel detector), $|d_0| < 2.0$ mm, and $|z_0 \sin\theta| < 3.0$ mm. The track reconstruction and selection criteria described above has a good efficiency for primary particles (for example, more than 99% reconstruction efficiency for muons), while rejecting fake tracks.

For jets with $p_{\mathrm{T}} < 400$ GeV, the combined efficiency of the track reconstruction and selection criteria (evaluated through simulation) is around 94% in the $|\eta| < 1.2$ region, and 86% in the forward $|\eta| > 1.2$ region. This reconstruction efficiency decreases with increasing jet transverse momentum [115].

4.1.2 Vertex reconstruction

The reconstruction of the primary vertex is crucial to the $VH, H \to b\bar{b}/c\bar{c}$ analysis, since it is used to identify the jets with b-quarks and c-quarks from the Higgs boson decay. To reconstruct the primary vertex, the longitudinal impact parameters z_0 of all tracks with respect to the centre of the beam spot are used to produce a single seed at the location of the estimated mode in z, and iteratively finds the most likely value [116]. Then tracks compatible with the seed are grouped together for fitting, and an adaptive vertex fitting algorithm [117] is used to estimate the position and uncertainty of the vertex. If more than one primary vertex (PV) is reconstructed, the vertex with the highest $\sum(p_{\mathrm{T}})^2$ of tracks is chosen as the primary vertex. A longitudinal resolution of 30 μm, and a transverse resolution ranging[3] between 10 μm to 12 μm is obtained for the primary vertex.

4.2 Electrons

The reconstruction, identification, and isolation of electrons involve information from the inner detector, the electromagnetic calorimeter (ECal), and the hadron calorimeter (HCal). Electromagnetic *topo-clusters* are clusters of energy deposits in topologically connected ECal and HCal cells, used in the reconstruction of electrons and photons [114]. Topo-clusters are required to have an electromagnetic energy (i.e the energy from ECal cells) above 400 MeV, and more than 50% of the total cluster energy to be the electromagnetic energy.

4.2.1 Electron reconstruction

The electron reconstruction algorithm first prepares tracks and topo-clusters it will use for reconstruction. The topo-clusters are matched to one or several inner detector tracks, and the algorithm refits the track-cluster matching accounting for bremsstrahlung. The matched tracks are required to satisfy a $|d_0|/\sigma_{d_0} < 5$, and $|z_0|\sin\theta < 0.5$ mm, where σ_{d_0} is the uncertainty in d_0.

Topo-clusters are then used to form dynamic, variable-size clusters of energy deposits in the ECal, known as *superclusters*. Dynamic clusters are useful over fixed-size clusters (used in AT-

---

[3] Depends on the LHC running conditions that determine the beam spot size.

LAS before 2017), since dynamic clusters can change in size to recover energy from bremsstrahlung photons or from electron-photon conversions. Further, the energy calibration for the dynamic clustering algorithm can still achieve a similar linear energy response to fixed-sized clustering algorithms. Figure 4.3 shows the simulated reconstruction efficiency for clusters, tracks, and electron candidates as a function of transverse energy of the electron E_{T}. More details on electron reconstruction can be found in Appendix D.1, and [114].

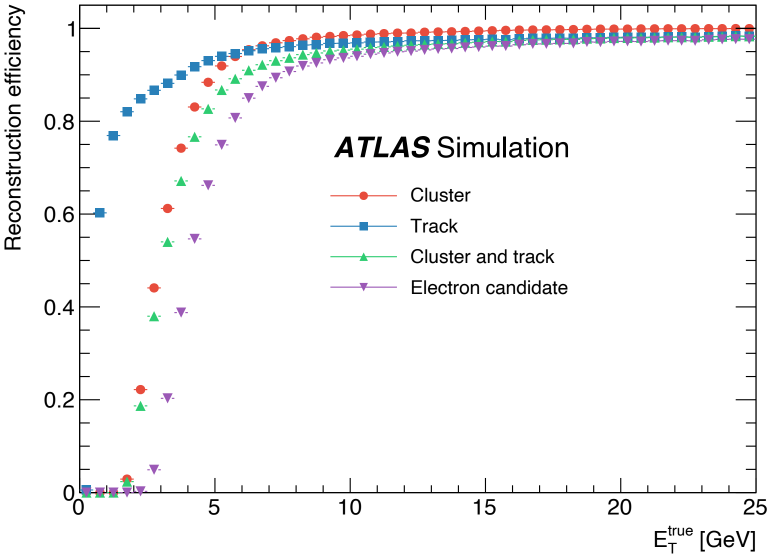


Figure 4.3: The simulated reconstruction efficiency for clusters, tracks, and electron candidates as a function of transverse energy of the electron E_{T} [114].

4.2.2 Electron identification and isolation

Reconstructed electron candidates pass through several quality criteria known as *identification* selections, in order to improve the purity of the prompt electrons coming from the primary vertex. Prompt electrons are identified using a likelihood based method, which uses quantities derived from the inner detector and the calorimeter[4], that are able to discriminate prompt isolated electrons from fake electrons; hadronic jets, converted photons, or electrons from heavy flavour hadron decays [118].

In ATLAS, three different identification operating points are defined to be used for analyses; Tight, Medium, and Loose. The identification efficiencies in the energy range of 20 GeV$<E_{\mathrm{T}}<$50 GeV are 93%, 87%, and 79% for the Loose, Medium and Tight operating points, respectively, and gradually increase from low to high E_{T}. Despite the reduced efficiency for Medium and Tight operating points compared to the Loose operating point, they have a better rejection of background processes; a factor of 1.9 (3.1) better background rejection for Medium (Tight) operating points with respect to the Loose operating point.

Non-prompt electrons, which are not from the primary vertex, are emitted by other physics objects like semi-leptonically decaying heavy hadrons in jets. While some fraction of non-prompt

---

[4] Such as the number of hits in the pixel and SCT detectors, ratio of the energy measured in the first ECal layer to the total electromagnetic energy, etc.

electrons are rejected from identification selections, an *isolation* selection is additionally required to further separate prompt electrons from and non-prompt electrons. The isolation can be applied based on the energy deposits in the calorimeter (known as *calorimeter isolation*), or using tracks of nearby charged particles (known as *track isolation*), or from a combination of both the calorimeter energy deposits and tracks (More details in Appendix D.1).

4.2.3 Electron calibration

For simulated electrons, the digitized energy output from the ECal need to be calibrated using data. The energy of an electron candidate is built by combining the energy of the cells in the presampler and the first three electromagnetic calorimeter layers[5] assigned to the supercluster. The electron energy is calibrated as described in [119] using a boosted decision tree regression algorithm.

The difference in *energy scale* between the data and simulation (parameterized by a multiplicative term), and the mis-modelling in *energy resolution* (parameterized by a constant additive term) is derived for different η regions. Electron energy scale corrections are applied to data, while the energy resolution corrections are applied to the simulated samples. Both corrections are determined using $Z \to ee$ events, and in the low η region the energy scale correction is around 1 (corresponding to no correction), while the energy resolution correction is around 0.05%.

4.3 Muons

Muon candidates are reconstructed primarily based on information from the inner detector and the muon spectrometer (MS) tracking detectors. Information from the calorimeters is also used in determination of track parameters, to account for cases of large energy loss in calorimeters, and for muon spectrometer independent selection of inner detector tracks of the muon candidate[120].

4.3.1 Muon reconstruction

Muon reconstruction starts with the short straight-line[6] local track segments reconstructed from hits in individual muon spectrometer (MS) stations[120]. Segments in the different stations are combined into preliminary track candidates, and then information from precision measurements in the bending plane (from MS tracking detectors) are combined with a second coordinate (from MS trigger detectors) to obtain a 3D track candidates. Finally, a global χ^2 fits the muon trajectory through the magnetic field, taking into account the muon interactions with the detector material and effect of possible misalignment between detector chambers.

The reconstruction of muons using the full detector information is done next, following five main reconstruction strategies that lead to different *types* of muons.

- **Combined (CB) muons** are reconstructed by matching MS tracks to inner detector tracks, and then performing a combined track fit based on the inner detector and MS hits, while

---

[5] However, in the transition region $1.4 < |\eta| < 1.6$ between the ECal barrel and endcap, the energy measured in the scintillator between calorimeter cryostats is also added.

[6] Muons are straight lines in the $r - \phi$ plane, since the toroid magnet bends muons in the $r - z$ plane.

also accounting for the energy loss in the calorimeters.

- **Inside-out (IO) muons** are reconstructed using an algorithm which extrapolates inner detector tracks to the MS, and searches for ≥ 3 loosely-aligned MS hits. The inner detector hits, the energy loss in the calorimeters, and the MS hits are then used in a combined track fit.
- **MS extrapolated (ME) muons** are reconstructed when a MS track cannot be matched to an inner detector track, and the muon candidate parameters are extrapolated to the beam line. Such muons are used to extend the acceptance to $|\eta| < 2.7$, beyond the inner detectors acceptance of $|\eta| < 2.5$.
- **Segment tagged (ST) muons** are reconstructed by extrapolating an inner detector track to the MS, and requiring the a tight angular matching to at least one reconstructed MS segment. A successfully matched inner detector track is identified as a muon candidate, and the muon parameters are taken directly from the inner detector track fit.
- **Calorimeter tagged (CT) muons** are reconstructed by extrapolating inner detector tracks to calorimeter energy deposits consistent with a minimum-ionizing particle. Such energy deposits are used to tag the inner detector track as a muon, and the muon parameters are taken directly from the inner detector track fit.

CB and IO muons cover the majority of the acceptance region of $|\eta| < 2.5$. However, a gap in the muon spectrometer in $|\eta| < 0.1$ leads to loss of reconstruction efficiency for CB muons, and therefore CT and ST muons are used in this region. Finally, the acceptance is extended to $|\eta| < 2.7$ using ME muons, and a subset of CB muons that are reconstructed in the $2.5 < |\eta| < 2.7$ region by combining MS tracks with short tracks segments from pixel and SCT detectors.

4.3.2 Muon identification and isolation

Similar to electrons, muons are also required to pass a set of quality criteria referred to as *identification* criteria to increase the purity of high-quality muon candidates. The muon identification is done through a set of requirements, known as *selection working points*, on several variables based on the inner detector and muon spectrometer. Several selection working points are defined, but in general the *Loose*, *Medium* and *Tight* working points are used (More details are found in Appendix D.3). The efficiency for the selections (as a function of η) is shown in Figure 4.4.

The *Loose* selection working point accepts all five types of muons described in Section 4.3.1, with about 97% of prompt muons selected being CB and IO muons, and approximately 1.5% are CT and ST muons in the $|\eta| < 0.1$ region. The *Medium* selection working point accepts only CB and IO muons within the inner detectors acceptance of $|\eta| < 2.5$, while in the $2.5 < |\eta| < 2.7$ region, ME muons and some CB muons are accepted. 98% of prompt muons passing the Medium selection working point are CB muons.

The muon isolation criteria, similar to electrons, aims to discriminate prompt muons from non-prompt muons[7] coming from hadronic sources by measuring the amount of hadronic activity

---

[7] Similar to prompt electrons, prompt muons are muons coming from the primary vertex, and non-prompt muons are from hadron decays in flight, pile-up interactions, and cosmic muons.

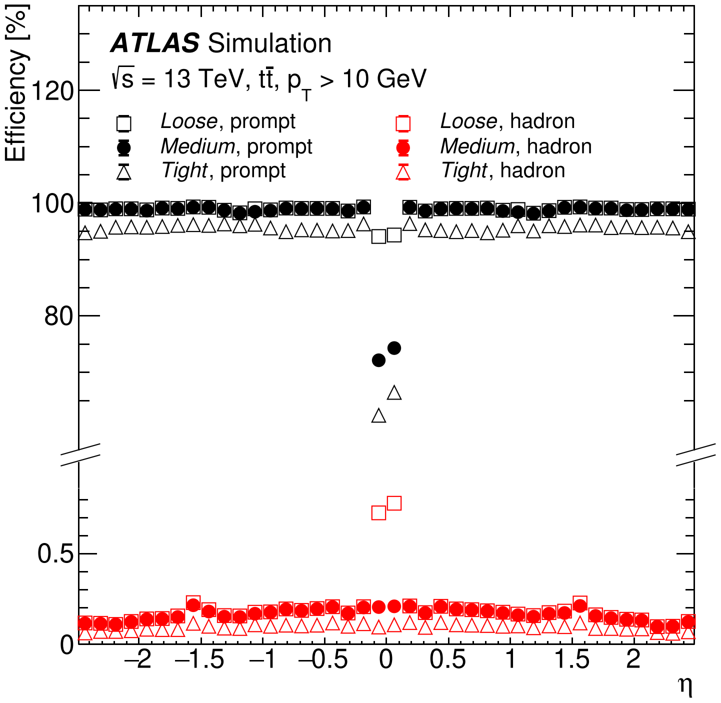


Figure 4.4: Simulated identification efficiency for muons and hadrons as a function of η [120].

in the vicinity of a muon. Similar to electrons, there are different muon *isolation working point* criteria. More details are found in Appendix D.3.

4.3.3 Muon calibration

The first step in calibrating the muon momentum is to measuring and correcting charge dependant biases in the scale of muon momentum measurement. Then, $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events are used to measure the muon momentum resolution and scale, and compare that with the prediction by simulation. Then a calibration procedure is applied to the simulated events in order to improve the agreement between the simulation and data [121].

The muon momentum calibration is performed according to Equation 4.1

$$p_{\mathrm{T}}^{\mathrm{Cor}} = \frac{p_{\mathrm{T}}^{MC} + \sum_{n=0}^{1} s_n(\eta,\phi)\left(p_{\mathrm{T}}^{MC}\right)^n}{1 + \sum_{m=0}^{2} \Delta r_m(\eta,\phi)\left(p_{\mathrm{T}}^{MC}\right)^{m-1} g_m}, \tag{4.1}$$

where p_{T}^{MC} is the uncorrected muon transverse momentum, g_m are normally distributed random variables, and terms $s_n(\eta,\phi)$ and $\Delta r_m(\eta,\phi)$ describe the muon momentum scale and momentum resolution corrections. The term $s_0(\eta,\phi)$ corrects the energy loss effects in the detector, and $s_1(\eta,\phi)$ corrects the inaccuracy in the magnetic field description. The denominator of Equation 4.1 accounts for momentum resolution smearing on the relative muon p_{T}, which can be parameterised as,

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{r_0}{p_{\mathrm{T}}} \oplus r_1 \oplus r_2 \cdot p_{\mathrm{T}} \tag{4.2}$$

with $\oplus$ denoting the sum in quadrature. r_0 accounts for the fluctuations of energy loss in traversed material, r_1 accounts for the effects such as multiple scattering or local magnetic field modeling, and r_2 accounts for effects such as spatial resolution of hits or detector misalignments. The corrections are derived in 18 pseudo-rapidity η regions as described in [121].

4.4 Jets

The QCD partons created in physics processes from proton-proton collisions will be observed as a jet of particles in the detector, due to fragmentation and hadronization. These will manifest themselves as particle tracks in the inner detector and energy deposits in the calorimeters.

4.4.1 Jet reconstruction

Objects from the Particle Flow (PFlow) algorithm, which combines the calorimeter energy deposits with inner detector tracks, are combined into jets using the anti-k_t algorithm. This section gives a brief overview of these two algorithms, but details can be found in Appendix D.6 and the given references.

Before being used in jet reconstruction, calorimeter cells are first clustered into massless 3D *topo-clusters*. Cells are added to the topo-cluster according to the ratio of the cell energy to the expected noise in each cell, using threshold controlling the growth of each topo-cluster [72]. The resulting energy of the topo-cluster is defined at the electromagnetic (EM) scale, that is the baseline calorimeter energy scale.

Object definition using the particle flow algorithm

The particle flow algorithm combines the capabilities of the inner detector in reconstructing charged particles, and the calorimeter's ability to reconstruct both charged and neutral particles [59]. However, since the inner detector has an acceptance for $|\eta| < 2.5$, jets reconstructed in the forward region $(2.5 < |\eta| < 4.5)$ uses topo-clusters from the calorimeter to reconstruct particle flow jets.

Particle flow requires a special procedure; to avoid double counting the energy of a particle during reconstruction, the correct signal in the calorimeter needs to be identified for the tracks used for reconstruction. Therefore, for particle tracks that are determined to be used for reconstruction, the corresponding energy deposits in the calorimeter are subtracted. Above $p_{\mathrm{T}}^{\mathrm{trk}} = 100\,\mathrm{GeV}$ no track information is used, and the the particle flow algorithm relies on the good calorimeter performance at high energies. Final jet reconstruction is done using an ensemble of particle flow objects consisting the remaining calorimeter energy deposits and the tracks that are matched to the hard interaction.

Jet clustering using the anti-k_t algorithm

Jet reconstruction is done using the anti-k_t algorithm [122], using the four-vector objects reconstructed with the Particle Flow algorithm. The anti-k_t algorithm is an algorithm used to iteratively merge particle flow objects, which is *infrared safe* and *collinear safe*; an algorithm is deemed infrared safe if the reconstructed jet properties, such as the four-momentum, are stable in the case of emission of additional soft (i.e low energy) particles such as gluons, and collinear safe implies that the reconstructed jet properties are stable with respect to the number of emitted collinear particles.

k_t stands for the transverse momentum p_{T}, where the term k_t^{2p} is used when determining the smallest distance between clusters. The p parameter can take various values, and $p = -1$ which translates to $k_t^{-2} \to 1/p_{\mathrm{T}}^2$ is referred to as the anti-k_t algorithm. Using a value of $p = -1$ ensures that the particle flow objects are associated to hard (high-energy) objects compared to soft objects.

4.4.2 Jet calibration

The fully reconstructed jets will next be calibrated through a dedicated jet calibration procedure that aims to determine the jet energy scale (JES) accurately, while achieving the best possible energy and angular resolution. Figure 4.5 shows an illustration of the jet energy scale calibration procedure implemented in ATLAS, which is done in four main steps. At each step, the corrections/calibrations are applied to the reconstructed 4-momentum of the jets, scaling the jet p_{T}, energy, and the mass [72].

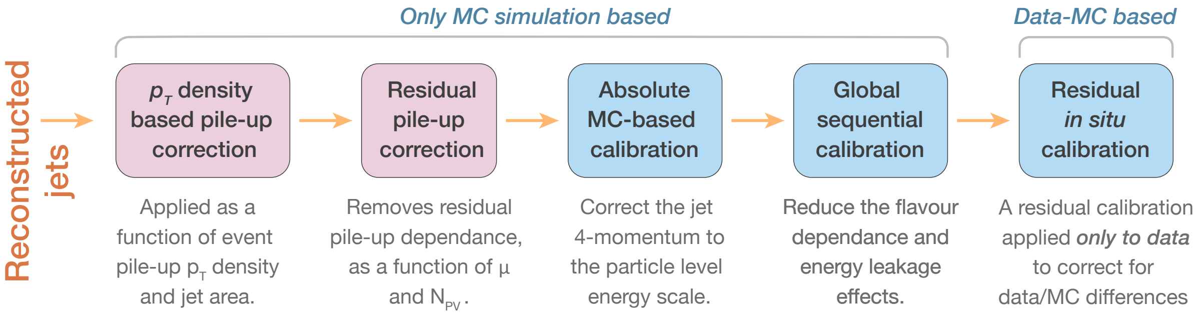
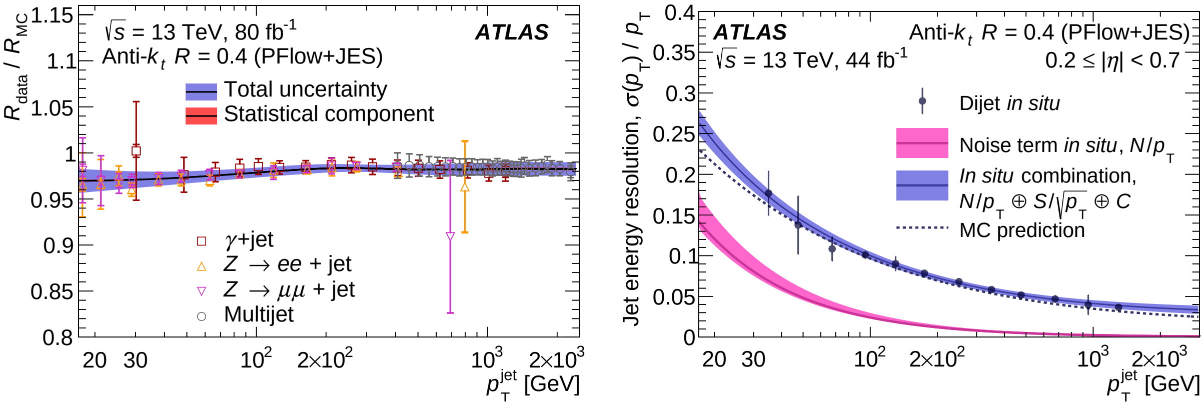


Figure 4.5: An illustration of the different stages in jet energy scale calibration. Each calibration is applied to the 4-momentum of the jet. The illustration is based on the original diagram in [72]

The first step in the calibration procedure involves two *pile-up corrections* to remove the excess energy due to additional proton-proton interactions within the same (in-time) bunch crossing, or nearby (out-of-time) bunch crossings. The first pile-up correction is based on the jet area (A) and the transverse momentum density (ρ) of the event. The second pile-up correction is a residual correction derived from the MC simulation, and parameterized as a function of the mean pile-up (μ) and the number of reconstructed primary vertices in the event (N_{PV}).

Next, the *absolute JES calibration* step corrects the 4-momentum of the jets such that the jet agrees in energy and direction with truth jets from MC simulations. A *global sequential calibration* then improves the jet p_{T} resolution and the associated uncertainties by removing the dependance of the reconstructed jet response on the observables constructed using the information of the inner detector, calorimeter, and the muon spectrometer. Finally an in-situ calibration is applied only to data, to correct for residual differences in the Data and the MC simulation. The insitu-calibration scale factors are derived using well measured objects such as photons and Z-bosons as a reference.

Once the full JES calibration is done, the jet energy resolution (JER) expressed as in Equation 4.3, is measured from objects including photons, Z-bosons, and calibrated jets.



(**a**) Jet response ratio as a function of p_T.

(**b**) JER as a function of p_T.

Figure 4.6: Left plot shows the jet response ratio as a function of p_T for fully calibrated (JES) particle flow (PFlow) jets, that is obtained at the final *in-situ* calibration step. The right plots shows the relative JER as a function of p_T for fully calibrated PFlow jets[72].

$$\frac{\sigma(p_T)}{p_T} = \frac{N}{p_T} \oplus \frac{S}{\sqrt{p_T}} \oplus C, \tag{4.3}$$

where N is the noise term, S is the stochastic coefficient and C is a constant term.

4.4.3 Jet selections to suppress pile-up and fake jets

The Jet Vertex Tagger (JVT) algorithm is used to suppress jets arising from pile-up interactions, while correctly identifying jets attributed to the hard-scattering vertex. The JVT procedure builds a multivariate discriminant for each jet within $|\eta| < 2.4$ based on tracks ghost-associated with the jet [123]. In simulation, the efficiency for picking jets arriving from the primary vertex is estimated to be 92%. However, since the rate of pile-up jets with $p_T \geq 60$ GeV is sufficiently small, the JVT requirement is not applied for jets with $p_T \geq 60$ GeV. Due to the lack of tracking information in the forward region (i.e $|\eta| > 2.5$), the JVT algorithm cannot be used for forward jets. Hence, another algorithm known as the forward Jet Vertex Tagger (fJVT) is applied for forward jets with $p_T < 120$ GeV [124].

4.5 Jet flavor tagging

After the reconstruction of jets, it is important to identify the initial particle that the jet is attributed to. This is possible due to the different substructure in jets containing b-hadrons, c-hadrons, or τ-leptons, and the jets can be classified into a b-jet, c-jet, light-flavor jets[8], or τ-jets. This classification of jets in b-jets, c-jets, and light-jets is called jet flavor tagging (FTAG), and the identification of τ-jets falls under the framework described in Section 4.6. More details on jet flavor tagging can be found in Appendix D.7.

The ATLAS experiment uses a set of algorithms for identifying b-jets and c-jets, referred to as

---

[8] Light-flavor jets are hadronic jets initiated by light quarks u, d, s, or gluons.

flavor-tagging algorithms [36]. These algorithms exploit the long lifetime[9], high mass and high decay multiplicity of b-hadrons and c-hadrons. The long lifetime of b-hadrons will lead to at least one measurable displaced vertex a few millimeters from the primary vertex.

Jet flavor tagging is based on the vertices, tracks, and jets reconstructed in the ATLAS detector; Only tracks with p_{T} larger than 500 MeV is considered for flavor tagging, with further selections applied to reject fake or poorly reconstructed tracks. The primary vertex (PV) serves as the reference point from which tracks and vertex displacements are computed, and each event is required to have at least one PV, while jets are reconstructed from particle flow objects using the anti-k_t algorithm, as explained in Section 4.4. It should be noted that jets with $p_{\mathrm{T}} < 20$ GeV or $|\eta| > 2.5$ are not considered for jet flavor tagging, since low-p_{T} jets are outside the valid calibration range, and high $|\eta|$ jets are outside the tracking fiducial volume.

Tracks are matched to jets by requiring a maximum allowed ΔR between the track momenta defined at the perigee, and the jet axis; since decay products of high-p_{T} heavy hadrons are more collimated, the ΔR requirement varies as a function of p_{T}. The requirement is $\Delta R = 0.45$ at $p_{\mathrm{T}} = 20$ GeV, and reaches the requirement of $\Delta R = 0.26$ at $p_{\mathrm{T}} = 150$ GeV. If more than one jet fulfills the matching criteria, tracks are matched to the jet to which the tracks are closest in ΔR.

4.5.1 Jet flavor tagging algorithms

In ATLAS, jet flavor tagging takes a two stage approach. First, low level flavor tagging algorithms are used to reconstruct characteristic features of b-jets and c-jets through two complementary approaches; one approach uses properties of individual tracks associated with a hadronic jet, while another approach combines tracks to explicitly reconstruct displaced vertices. To maximise the flavor tagging performance even further, the output values of the low level flavor tagging algorithms are used for high level flavor tagging algorithms that are based on multivariate classifiers [36].

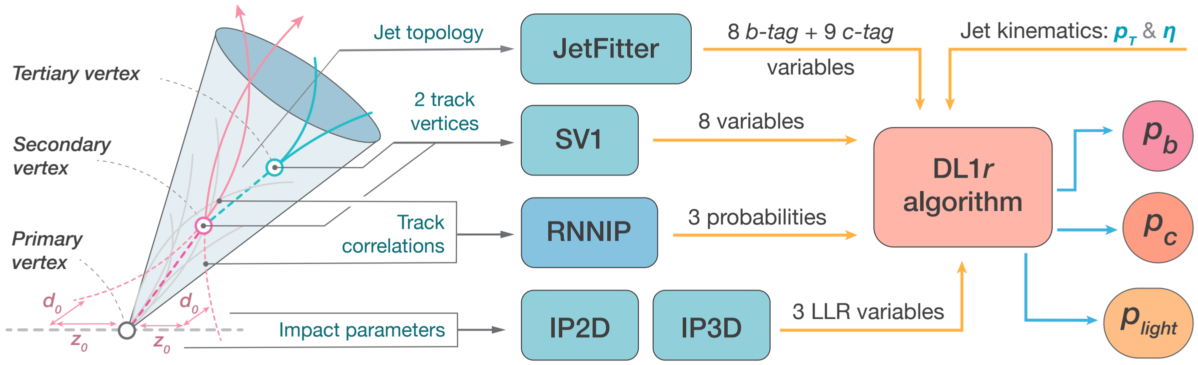


Figure 4.7: An illustration of the flavor tagging algorithms used in ATLAS

Low level flavor tagging algorithms fall into two broad categories. The first category consists of the `IP2D` and `IP3D` algorithms, and are based on exploiting the large impact parameters of the tracks from b-hadron and c-hadron decays. The `RNNIP` algorithm, a new algorithm used in the $VH, H \to b\bar{b}/c\bar{c}$ analysis, uses a recurrent neural network that learns track impact parameter

---

[9] Since b-hadrons have a mean lifetime of $< c\tau > \approx 450\ \mu m$, they have a measurable mean flight length $< l > = \beta\gamma c\tau$ in the detector.

correlations to further improve the jet flavor discrimination. The second category consists of the `SV1`, that attempts to reconstruct a secondary vertex inside a jet, and the JetFitter algorithms, which aims to reconstruct the full b-hadron and c-hadron decay chains.

Figure 4.7 illustrates the flow of jet flavor tagging in ATLAS, which is explained more in the following sub-sections.

Low level flavor tagging algorithms

The `IP2D` tagger uses the *signed transverse impact parameter* significance (d_0/σ_{d_0}) of tracks to create a discriminant that can be used to tag b-jets. The `IP3D` tagger uses both the signed transverse impact parameter significance (d_0/σ_{d_0}) of tracks, and the longitudinal impact parameter significance (z_0/σ_{z_0}) in a 2D template to account for their correlation. Probability density functions (PDF), obtained from reference MC simulated histograms based on the signed transverse and longitudinal impact parameter significance of tracks associated with b-jet, c-jet and light-jets, are used to calculate the ratio of b-jet, c-jet and light-jet fractions on a per-track basis. Using the PDF's, log likelihood ratio (LLR) discriminants are defined as the sum of the probability ratios for each jet-flavor hypotheses.

In heavy hadron decays, charged particles emerge from secondary or tertiary vertices with large impact parameters. Since the impact parameters of such charged particle tracks are intrinsically correlated[10], the `RNNIP` algorithm is trained to exploit the correlation in order to better discriminate between b-jets, c-jets, and light-jets [36].

The secondary vertex tagging algorithm (`SV1`) reconstructs a single displaced vertex in a jet. First, possible two-track vertices that can be built from tracks associated with the jet are identified, while rejecting two-track vertices compatible with K^0_S meson or Λ baryon decays, photon conversion to e^+e^-, and hadron interaction with the detector material. The `SV1` algorithm runs iteratively on all tracks of the identified two track vertices, combining each track using a χ^2 approach to fit one secondary vertex, with an invariant mass less than 6 GeV. This approach typically assigns the decay products of b-hadrons and c-hadrons to a single common secondary vertex. Finally, eight discriminating variables associated with the `SV1` algorithm will be used as inputs to the high level flavor tagging algorithms.

The JetFitter, also known as the topological multi-vertex finding algorithm, exploits the topological structure of b-hadron and c-hadron (referred to as heavy hadrons) decays inside the jet to reconstruct the full decay chain [36]. Similar to the `SV1` algorithm, there are eight b-jet discriminating variables associated with the JetFitter. However, by further exploiting c-hadron decay properties such as only a single reconstructed secondary vertex, intermediate charged decay multiplicity and moderate decay length with respect to b-hadron decays, nine c-jet discriminating variables associated with the JetFitter are defined. Altogether, these 17 variables serve as inputs to the high level flavor tagging algorithms.

High level flavor tagging algorithms

High level flavor tagging algorithms are based on machine learning based classifiers like boosted decision trees and neural networks. The high level flavor tagging algorithm used in the current

---

[10] In heavy hadron decays, if one track with large impact parameter is found, it is likely to find another track with a large impact parameter as well.

analysis is based on a deep feed-forward neural network, and commonly referred to as the DL1r algorithm.

The inputs to the DL1r algorithm, the output of the low level flavor taggers and the jet kinematic variables, have varying levels of correlation. In general, for heavy flavors the `IP2D`, `IP3D`, `SV1` and JetFitter algorithms have a large correlation between them. However, this is significantly reduced for light-flavor jets, and at high p_{T} regions. In addition to this, the output probabilities from the `RNNIP` algorithm are not strongly correlated between each other. The DL1r algorithm is trained to exploit this correlation differences, in order to reach the best flavor tagging performance. Furthermore, the use of the `RNNIP` algorithm is the specialty of the DL1r algorithm, where r denotes for the `RNNIP` algorithm.

The DL1r algorithm has a multi-dimensional output corresponding to probabilities for jet to be a b-jet, c-jet, or a light-flavor jet. Since all the jet flavors are treated equally during training, the trained network can be used for both b-jet tagging and c-jet tagging. The DL1r discriminant for b-jet tagging is defined as

$$D_{\mathrm{DL1r}} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\mathrm{light}}}\right) \tag{4.4}$$

where p_b, p_c and p_{light} represent the probabilities for jet to be a b-jet, c-jet, or a light-flavor jet, respectively. f_c denotes the effective fraction of c-jets in the background training samples. Similarly, the DL1r discriminant for c-jet tagging is defined as

$$D^c_{\mathrm{DL1r}} = \ln\left(\frac{p_c}{f_b \cdot p_b + (1 - f_b) \cdot p_{\mathrm{light}}}\right) \tag{4.5}$$

where f_b denotes the effective fraction of b-jets in the background training sample. The D_{DL1r} and D^c_{DL1r} discriminants are used to define optimized working points for tagging b-jets or c-jets, also known as working points. The efficiency of tagging a flavor f in a specific working point is defined as

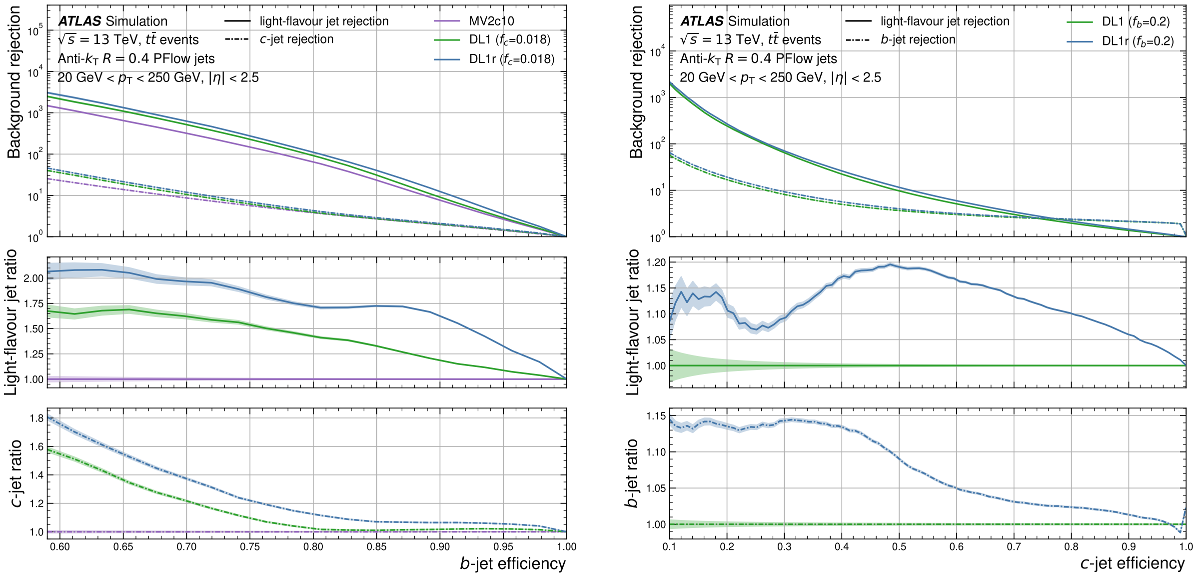
$$\epsilon_f = \frac{\text{The number of jets of flavor } f \text{ in a specific working point}}{\text{Total number of jets of flavor } f \text{ in all working points}} \tag{4.6}$$

Using ϵ_f, the rejection for the corresponding jet flavors f is quantified as $1/\epsilon_f$.

Figure 4.8(a) shows the efficiency of b-jet tagging (ϵ_b), and Figure 4.8(b) shows the efficiency c-jet tagging (ϵ_c), together with the rejection power for different flavors f. Compared to the previous DL1 algorithm (not using `RNNIP` for the input) or MV2 algorithm (based on a boosted decision tree, and not using `RNNIP` for the input), the new DL1r algorithm is seen to have a good improvement in performance for both b-jet tagging and c-jet tagging due to improved background rejection at the same b-jet/c-jet tagging efficiencies.

4.5.2 Calibration of the high-level flavor tagging algorithms

The imperfect description of the detector response and physics modelling effects in MC simulations requires the measurement of the flavor tagging performance with real collision data.



(**a**) b-tagging performance vs the b-jet efficiency.

(**b**) c-tagging performance vs the c-jet efficiency

Figure 4.8: The top panel of the left (right) plot shows the b-jet (c-jet) tagging performance as a function of the b-jet (c-jet) efficiency. The bottom panels show the rejection power of the algorithm for c-jet (b-jet) and light-flavor jet backgrounds [36].

However, due to varying levels of statistics for the jet flavors b, c and light in the regions optimized for flavor tagging, each jet flavor needs to be calibrated in dedicated procedures.

The calibration derives a scaling factor given in Equation 4.7 for the flavor f, to account for the difference between data and the MC simulation. These scaling factors are then applied to simulated events to calibrate the flavor tagging effects.

$$\mathrm{SF}_f = \frac{\epsilon_f^{\mathrm{data}}}{\epsilon_f^{\mathrm{MC}}} \tag{4.7}$$

b-jet tagging efficiency calibration

Given the abundance of top pair production ($t\bar{t}$) in proton-proton collisions at the LHC, and the $t \to Wb$ branching ratio is almost 100%, an abundant source of b-jets can be obtained by selecting the $t\bar{t}$ process, which then can be used to extract the b-jet tagging efficiency (ϵ_b). A very pure sample of $t\bar{t}$ is obtained through requiring both W-bosons decay leptonically to electrons or muons plus a neutrino, referred to as the di-leptonic $t\bar{t}$ process.

For calibrating the scale factor, $t\bar{t}$ events are binned into *nine bins* according to the p_T of the leading jet (i.e jet with the highest p_T, labeled as j_1), and the p_T sub-leading jet (i.e jet with the second highest p_T, labeled as j_2). Then in each p_T bin, events are assigned to a signal region (SR) and control regions (CR). The combinations with non b-jets from the top decay, labeled l, are bl, lb, ll, are constrained from the CRs, which have a larger non-bb sample fraction compared to the SR. Then finally the SR events are classified into the different b-jet tagging working points of the DL1r algorithm [125], to measure the ϵ_b in each DL1r working point as a function of p_T through a combinatorial likelihood approach.

In ATLAS there are four default regions used for b-jet tagging with different ϵ_b. In these regions, ϵ_b is 60%, 70%, 77%, and 85% where 70% (for example) implies that 70% of the true b-jets are estimated to be correctly tagged as b-jets in this working point. Regions with less b-jet tagging efficiency have better c-jet and light-flavor jet rejection, making these regions ideal for obtaining pure b-jet samples. Figure 4.9 shows the data to MC scale factors for the measured b-jet tagging efficiencies ϵ_b in four different DL1 algorithm working points, with the associated uncertainties.

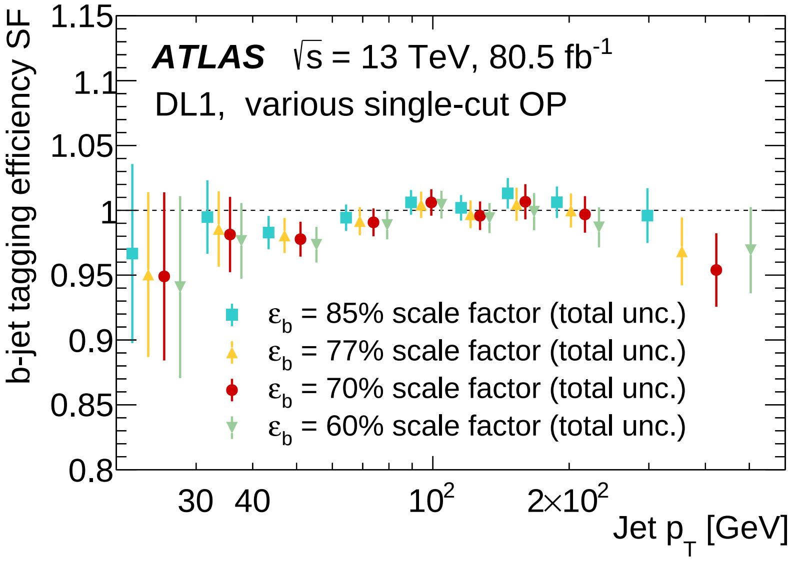


Figure 4.9: The measured ϵ_b for the 70% working point of the DL1 algorithm [125]

c-jet tagging efficiency calibration

Similar to calibrating b-jet tagging efficiency, c-jet tagging efficiency calibration also uses $t\bar{t}$ events, but one of the W-bosons is required to decay hadronically, while the other W-boson is required to decay leptonically to electrons or muons plus a neutrino (referred to as the semi-leptonic $t\bar{t}$ process) [126].

After event selections are applied to reduce the non-semi-leptonic $t\bar{t}$ processes, an algorithm using a likelihood function assigns the four jets to the $t\bar{t}$ decay topology; either the b-jets from the top decay, or the jets from the W-decay. After further selections, including a tight b-jet selection for the jets from the top decay, a sample that is 99.8% pure in true b-jets coming from a top-decay is obtained, with a good reduction in the non-$t\bar{t}$ background for the jets coming from the W-boson.

The jets from the W-boson decay are then classified into four p_{T} regions (20-40 GeV, 40-65 GeV, 65-140 GeV, and 140-250 GeV), and working points defined using the DL1r discriminant. Next, the ϵ_b (ϵ_{light}) scale factors are applied to the true b-jets (light-flavor jets), and the remaining c-jets are corrected to real collision data through a χ^2 minimization fit to obtain the c-jet scaling factors. This same calibration procedure was used when calibrating working point defined using the D^c_{DL1r} discriminant.

Light-flavor jet tagging efficiency calibration

Calibrating the light-flavor jet tagging efficiencies ϵ_{light} (also referred to as mis-tagging efficiencies) is challenging as a consequence of the good rejection of light-jets in the flavor tagging working points; there is limited statistics of light-flavor jets to determine ϵ_{light} accurately. Therefore, a sample enriched in light-flavor jets is constructed using a modified DL1r algorithm, known as the DL1r$^{\text{Flip}}$ algorithm, that achieves a lower ϵ_b and ϵ_c with a moderate ϵ_{light}, to enhance the light-flavor jet ratio in the sample.

The modified DL1r tagger method, commonly referred to as the *negative tag method*, first determines the negative tag rate $\epsilon^{\text{data}}_{\text{neg,light}}$ in a data sample enriched in light-flavor jets, and then extracts the real mis-tagging rate ϵ_{light} in the relevant flavor tagging working points [127]. For extracting $\epsilon^{\text{data}}_{\text{neg,light}}$, a Z+jets sample with the Z-boson decaying to two leptons is used.

The negative tag method is based on the assumption that light-flavor jets are mis-tagged as b-jets due to the finite resolution of the reconstructed inner detector track trajectories and impact parameters. Under this assumption, signed impact parameters d_0 and z_0 are expected to be almost symmetric around zero for light-flavor jets, but not for b-hadrons or c-hadrons due to their longer decay lengths as seen in Figure 4.10. Hence, the probability for light-flavor jets to be mis-tagged remains almost the same after flipping the sign of the impact parameters, and a good approximation of the light-flavor jet mis-tag rates can be obtained by running the `IP2D` and `IP3D` algorithms with the flipped signs.

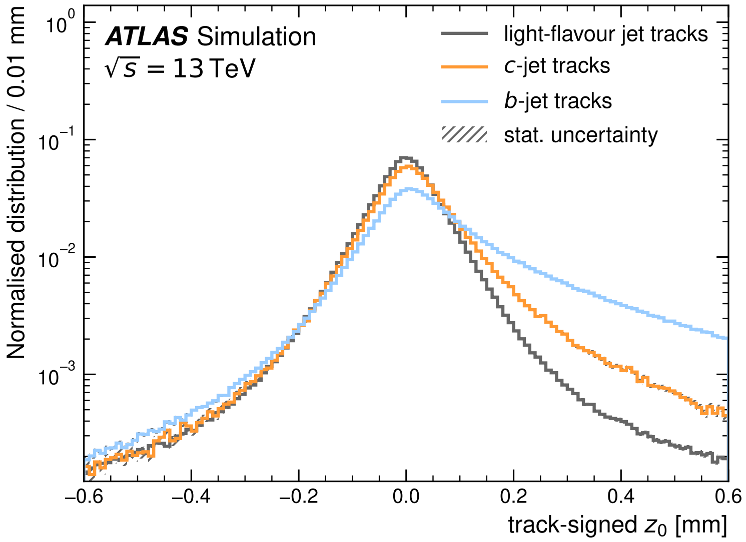


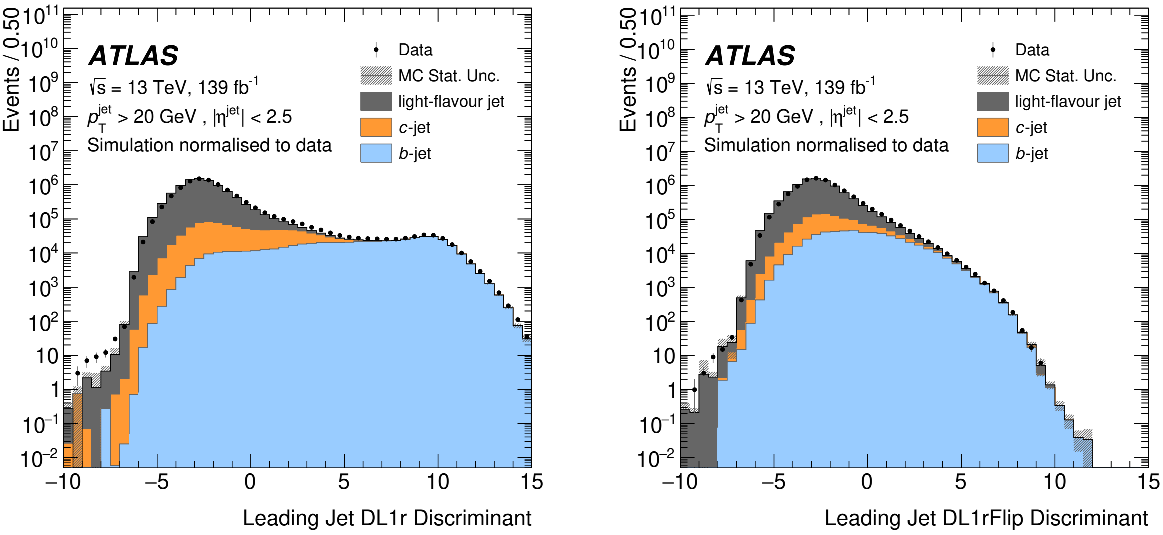
Figure 4.10: The signed-longitudinal impact parameter z_0 [127].

Similar features are expected for secondary vertices and signed decay length significance (which are seeded from tracks), and hence, the negative tag method can be extended to `SV1` and JetFitter algorithms. It should be noted that the flavor tagging algorithms are not retrained, and uses the same input types but the modified values that are determined from the flipped version of the low level flavor tagging algorithms[11].

---

[11] The flipped versions of the low level flavor tagging algorithms are referred to as `IP2DNeg`, `IP3DNeg`, `SV1Flip`, and `JetFitter-Flip`

The DL1r$^{\text{Flip}}$ algorithm achieves a similar ϵ_{light} in the flavor tagging working points, but with less b-jet and c-jet fractions. This is seen in the comparison of the DL1r and DL1r$^{\text{Flip}}$ discriminant scores in Figure 4.11, where less b-jets and c-jets are there at larger DL1r$^{\text{Flip}}$ discriminant score.



(a) Different jet flavors in DL1r discriminant bins.

(b) Different jet flavors in DL1r$^{\text{Flip}}$ discriminant bins.

Figure 4.11: The different jet flavors in the DL1r and DL1r$^{\text{Flip}}$ discriminant bins [127].

Similar to ϵ_b and ϵ_c calibration, the ϵ_{light} calibrations are performed independently in different p_{T} bins for the different working points of the DL1r algorithm, but utilizing the DL1r$^{\text{Flip}}$ discriminant templates. The $\epsilon^{\text{data}}_{\text{neg},b}$, $\epsilon^{\text{data}}_{\text{neg,light}}$, and the ϵ_{light} scale factors are extracted from a binned fit to the secondary vertex mass (m_{SV}) templates, while $\epsilon^{\text{data}}_{\text{neg},c}$ and the corresponding scale factor is fixed to the MC simulation value with a 30% uncertainty due to limited sensitivity. The extracted ϵ_{light} scale factors are used in the end, with an extrapolation uncertainty from DL1r$^{\text{Flip}}$ to DL1r.

4.6 Tau leptons

τ leptons decay either leptonically as $\tau \to \nu_\tau \nu_l l$, or hadronically as $\tau \to \nu_\tau + hadrons$ after a mean decay length of 87 μm. Hence, due to the short decay length, most τ leptons decay before reaching the innermost layer of the pixel detector system, and only the decay products of the τ leptons can be observed. Leptonic decays will produce isolated electrons or muons, with a missing transverse energy ($E_{\text{T}}^{\text{miss}}$) due to neutrinos. Hadronic decays, accounting to 65% of the τ lepton decays, will mostly include one charged pion (one prong) or three charged pions (three prong). Therefore, a hadronic jet from a τ lepton decay will have a signature corresponding to that of a narrow jet with one or three tracks in the inner detector. Since the neutrino from the hadronic τ lepton decay cannot be reconstructed, a set of all the visible τ lepton decay products is referred to as $\tau_{\text{had-vis}}$.

$\tau_{\text{had-vis}}$ candidates are reconstructed as jets formed using the anti-k_t algorithm, with the distance parameter $R = 0.4$. After reconstruction and calibration, a recurrent neural network (RNN) based on track and calorimeter information is used to identify the true $\tau_{\text{had-vis}}$ from fake $\tau_{\text{had}-\text{vis}}$ that could originate from quark or gluon initiated jets [128]. Similar to electrons and muons, several working points are defined for the $\tau_{\text{had}-\text{vis}}$ identification. More details on τ lepton reconstruction is found in Appendix D.5.

4.7 Overlap removal

After the reconstruction and calibration of leptons and jets, the *overlap removal* procedure is used to avoid double counting multiple objects being reconstructed from the same detector responses. The overlap removal is done in the following sequence of steps, where objects considered at a given step are ones that were not removed in the previous step.

- **$\boldsymbol{\tau}$-leptons vs electrons**: If a τ-lepton and and electron satisfies $\Delta R(\tau, e) < 0.2$, the τ-lepton is removed from the event.
- **$\boldsymbol{\tau}$-leptons vs muons**: If a τ-lepton and a muon satisfies $\Delta R(\tau, \mu) < 0.2$, the τ-lepton is removed from the event. However, if the τ-lepton has a $p_{\text{T}} > 50$ GeV, and the muon is not a *combined muon*, then the τ-lepton is not removed.
- **Electrons vs muons**: If a *combined muon* shares an inner detector track with an electron, then the electron is removed from the event. If a *calorimeter tagged muon* shares an inner detector track with an electron, then the muon is removed from the event.
- **Electrons vs jets**: If an electron and jet satisfies $\Delta R(\text{jet}, \text{e}) < 0.2$, then the jet is removed from the event. For the remaining jets, if $\Delta R(\text{jet}, \text{e}) < 0.4$, then the electron is removed from the event.
- **Muons vs jets**: If a muon and jet satisfies $\Delta R(\text{jet}, \mu) < 0.2$, or if the muons inner detector track is ghost associated to a jet having less than three tracks with $p_{\text{T}} > 500$ MeV[12], the jet is removed from the event. For the remaining jets, if $\Delta R(\text{jet}, \mu) < \min(0.4, 0.04 + 10\ \text{GeV}/\text{p}_{\text{T}}^{\mu})$, then the muon is removed from the event.
- **$\boldsymbol{\tau}$-leptons vs jets**: If a τ-lepton and jet satisfies $\Delta R(\tau, \text{jet}) < 0.2$, the jet is removed from the event.

4.8 Missing transverse energy

The missing transverse momentum $p_{\text{T}}^{\text{miss}}$ (commonly referred to as $E_{\text{T}}^{\text{miss}}$) is an important observable in ATLAS, since it can account for the momentum carried by undetected particles, like neutrinos, in the ATLAS detector.

The missing transverse momentum in ATLAS is calculated in two types [129]. The first type is called the *hard term*, $\boldsymbol{p}_{\text{T}}^{\text{hard term}}$, which is calculated using hard objects; reconstructed and calibrated electrons, photons, τ-leptons, muons, and jets. The other type is called the *soft term*, $\boldsymbol{p}_{\text{T}}^{\text{soft term}}$, which is calculated using reconstructed charged particle tracks that are not associated with a hard object. The total missing transverse momentum vector $\boldsymbol{p}_{\text{T}}^{\text{miss}} = (p_x^{\text{miss}}, p_y^{\text{miss}})$ is calculated as the negative vector sum of $\boldsymbol{p}_{\text{T}}^{\text{hard term}}$ and $\boldsymbol{p}_{\text{T}}^{\text{soft term}}$ as

---

[12] Such tracks could be due to the final state radiation from the muon.

$$\begin{aligned}
\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} &= -\left(\boldsymbol{p}_{\mathrm{T}}^{\text{hard term}} + \boldsymbol{p}_{\mathrm{T}}^{\text{soft term}}\right) \\
&= \underbrace{-\left(\sum_{\substack{\text{selected}\\\text{electrons}}} \boldsymbol{p}_{\mathrm{T}}^{e} + \sum_{\substack{\text{selected}\\\text{photons}}} \boldsymbol{p}_{\mathrm{T}}^{\gamma} + \sum_{\substack{\text{selected}\\\tau-\text{leptons}}} \boldsymbol{p}_{\mathrm{T}}^{\tau} + \sum_{\substack{\text{selected}\\\mu}} \boldsymbol{p}_{\mathrm{T}}^{\mu} + \sum_{\substack{\text{selected}\\\text{jets}}} \boldsymbol{p}_{\mathrm{T}}^{\text{jet}}\right)}_{\text{hard term}} \\
&\quad \underbrace{-\sum_{\substack{\text{unused}\\\text{tracks}}} \boldsymbol{p}_{\mathrm{T}}^{\text{tracks}}}_{\text{soft term}} .
\end{aligned} \tag{4.8}$$

With the calculated $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$, the magnitude $p_{\mathrm{T}}^{\mathrm{miss}}$ ($\equiv E_{\mathrm{T}}^{\mathrm{miss}}$) is defined as

$$p_{\mathrm{T}}^{\mathrm{miss}} \equiv E_{\mathrm{T}}^{\mathrm{miss}} = |\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}| = \sqrt{\left(p_x^{\mathrm{miss}}\right)^2 + \left(p_y^{\mathrm{miss}}\right)^2}. \tag{4.9}$$

and another observable called the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, S_T, is defined using the scalar sum of all jet p_{T} as

$$S_T = \frac{E_{\mathrm{T}}^{\mathrm{miss}}}{\sqrt{H_T}} \qquad \text{where} \qquad H_T = \sum_i p_{T,i}. \tag{4.10}$$

The validation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is done mainly using $Z \to ll$+jets events, since this topology is ideal to evaluate fake $E_{\mathrm{T}}^{\mathrm{miss}}$ effects due to no real sources of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the process. Further, semi-leptonic $t\bar{t}$ events, or $W \to \mu\nu$+jets events are used to evaluate the presence of real $E_{\mathrm{T}}^{\mathrm{miss}}$.

4.9 Object selection in the $VH, H \to b\bar{b}/c\bar{c}$ analysis

The leptons used in the $VH, H \to b\bar{b}/c\bar{c}$ analysis are electrons, muons, and hadronically decaying τ-leptons that are used to reconstruct the vector boson candidate, while jets are used to reconstruct the Higgs boson candidate. The object reconstruction follows the general procedures described so far, and analysis specific selections are explained in this section.

4.9.1 Leptons

After the initial reconstruction explained in Sections 4.2 to 4.6, a selection criteria detailed in Table 4.1 is used to define the lepton candidates to reconstruct the vector boson candidate. To ensure high trigger efficiency, leptons selected using the single electron and single muon triggers are required to have an offline reconstructed p_{T} which is 5% larger than the trigger threshold. Therefore, the lepton p_{T} is required to be greater than 27 GeV.

Table 4.1: The lepton selection requirements for the $VH, H \to b\bar{b}/c\bar{c}$ analysis.

Lepton	Label	p_{T}	η	Identification	Isolation
Electrons	VH-Loose	>7 GeV	$\lvert\eta\rvert$<2.47	Loose	Calorimeter and tracks
	ZH-Signal WH-Signal	>27 GeV		Tight	Only calorimeter
Muons	VH-Loose	>7 GeV	$\lvert\eta\rvert$<2.7	Loose	Calorimeter and tracks
	ZH-Signal WH-Signal	>27 GeV >27 GeV ($p_{\mathrm{T}}^{V} < 150$ GeV) >25 GeV ($p_{\mathrm{T}}^{V} > 150$ GeV)	$\lvert\eta\rvert$<2.5	Medium	Only tracks
τ leptons	-	>20 GeV	$\lvert\eta\rvert$<2.5	Loose	1-prong | 3-prong

Using the selections given in Table 4.1, electrons and muons are classified into *VH-Loose*, *ZH-Signal*, and *WH-Signal* lepton labels, and Section 5.1.1 describes how the labels are used in selecting the vector boson candidate of the $VH, H \to b\bar{b}/c\bar{c}$ process. For the identification of the hadronically decaying τ-leptons, the *loose* working point of the RNN τ identification algorithm is used; this working point has 85% (75%) τ-jet identification efficiency, with a background rejection factor of 21 (90) for 1-prong (3-prong) events.

4.9.2 Jets

The jets used in the the analysis are *small radius jets*, reconstructed with the anti-k_t algorithm with $R = 0.4$. Small radius jets are used to reconstruct the Higgs candidate for the $VH, H \to c\bar{c}$ analysis, and for the $VH, H \to b\bar{b}$ analysis in the low p_{T}^{V} region[13]. The jets are further classified into,

---

[13] There are two more jet collections used in the high p_{T}^{V} region, but is not discussed in detail since that analysis is out of scope of this thesis.

- **Signal jets** - These are jets in the central region of the detector, where $|\eta| < 2.5$. These jets are required to have p_{T} >20 GeV.
- **Forward jets** - These are jets in the forward region of the detector, for $2.5 < |\eta| < 4.5$. These jets are required to have a p_{T} >30 GeV.

To reduce the background not associated with the primary vertex, the jet-vertex-tagging technique (discussed in Section 4.4.3) is applied to central (forward) jets with $p_{\mathrm{T}} < 60(120)$ GeV. Jet cleaning is also applied in all the charged lepton channels to reduce the background from non-collision processes.

Jet energy and momentum corrections

Due to the semi-leptonic decay of b-hadrons and c-hadrons, and different energy response of b-jets and c-jets compared to light-jets, corrections are applied to the pair of small radius jets chosen to build the Higgs boson candidate; the steps to build the Higgs boson candidate is explained in Section 5.2.

For the *muon vs jets* overlap removal procedure explained in Section 4.7, the 4-vector of the muon removed from the $\Delta R(\text{jet}, \mu) < \min(0.4, 0.04 + 10\ \text{GeV}/\text{p}_{\text{T}}^{\mu})$ selection is added to the 4-vector of the jet. In the two charged lepton channel, two more techniques are employed to improve the mass resolution of the Higgs boson candidate. One technique applied only for events with less than four jets, exploits the good energy and momentum resolution of reconstructed electrons and muons, and corrects the jet 4-vector utilizing the event topology information of the transverse plane balance using a kinematic likelihood fit. Another technique, applied for events with more than three jets, aims to add back the jet from final state radiation.

The combined improvement in the Higgs mass resolution of all these techniques yield up to 12% in the 0-lepton and 1-lepton channels, and up to 40% improvement in the 2-lepton channel.

4.9.3 Jet flavor tagging

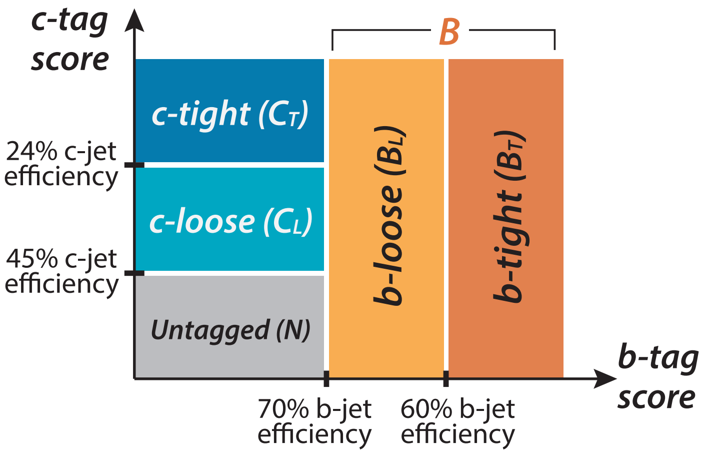


Figure 4.12: The 2D flavor tagging scheme used in the $VH, H \to b\bar{b}/c\bar{c}$ analysis.

The analysis uses the DL1r high-level flavor tagging algorithm, introduced in Section 4.5, to classify jets into the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ working points. The DL1r flavor tagging algorithm has significantly improved tagging performance for b-jets and c-jets, compared to the (DL1) algorithm used in the previous $VH, H \to c\bar{c}$ analysis. For example, as seen from Figure 4.8(b) for the 40% c-jet tagging efficiency, there is a 1.17 times better b-jet rejection and 1.13 times better light-flavor jet rejection for the DL1r algorithm compared with the DL1 algorithm.

The working points for the $VH, H \to b\bar{b}/c\bar{c}$ analysis are defined using a custom 2D scheme, where one dimension is the D_{DL1r} discriminant of Equation 4.4, and the other is the D^{c}_{DL1r} discriminant of Equation 4.5. Figure 4.12 depicts this custom 2D scheme (known as the 2D flavor tagging scheme), where D_{DL1r} and D^{c}_{DL1r} are used as scores to classify jets into five working points.

In the 2D flavor tagging scheme, the working point B is optimized for tagging b-jets, while the two bins C_T and C_L are optimized for tagging c-jets. The corresponding jet-flavor tagging efficiencies for a flavor f (ϵ_{f}) in each of the DL1r working points is given in Table 4.2. The calibration of ϵ_b, ϵ_c, and $\epsilon_{\mathrm{light}}$ scale factors for these working point are also done according to the same methods explained in Section 4.5.2.

Table 4.2: The jet-tagging efficiencies of b-jets, c-jets, light-jets, and τ-jets, in different bins of the 2D flavor tagging scheme.

Bin	**Bin name**	ϵ_b	ϵ_c	$\epsilon_{\mathrm{light}}$	ϵ_τ
1	c-loose (C_L)	11.5%	20.5%	6.5%	18.5%
2	c-tight (C_T)	4.8%	24.2%	0.9%	19.5%
3	b-loose (B_L)	11.2%	5.2%	0.13%	1.7%
4	b-tight (B_T)	58%	2.65%	0.051%	0.49%

Chapter 5

The event selection and event categorization

The search for the $VH, H \to c\bar{c}$ process is challenging due to the large background from other standard model processes, as discussed in Chapters 1 and 3. Therefore event selections aimed at rejecting the background processes while maximising the signal is essential. After event selections, a jet-flavor tagging selection will categorize the events further into the $VH, H \to c\bar{c}$ and $VH, H \to b\bar{b}$ signal regions. Finally, a categorization based on $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ will classify events to a signal region to further maximise the significance.

5.1 The general event selections

The events that pass the trigger selection (discussed in Section 3.1.1) will first undergo selections to reconstruct the vector boson candidate, and are channel specific in order to increase the significance in each channel.

5.1.1 Vector boson candidate selections

The reconstruction of the vector boson is specific to each charged lepton channel, since the final state of the vector boson is different. The leptons used in the reconstruction of the vector boson are mainly neutrinos, electrons and muons. A summary of all the vector boson selections, together with the trigger selections, are summarized in Table 5.1.

Zero charged lepton channel

The expected signal process in the zero charged lepton channel is $ZH \to \nu\bar{\nu}c\bar{c}$ (and $ZH \to \nu\bar{\nu}b\bar{b}$), and the neutrinos appear as a missing transverse momentum in the detector. Therefore, the transverse momentum of the Z-boson is represented by the $E_{\mathrm{T}}^{\mathrm{miss}}$. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be larger than 150 GeV to ensure a good trigger efficiency, as discussed in Section 3.1.1.

Scale factors in the range 0.9 to 1 are applied to events using the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger to account for the Data and MC difference in the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger turn on region from 150 GeV to 200 GeV. Further, a criterion adopted from the previous analysis, where the scalar sum (S_{T}) of the p_{T} of the jets in the event are required to be $S_{\mathrm{T}} > 120$ GeV$(S_{\mathrm{T}} > 150$ GeV$)$ for events with 2 jets(3 or more jets).

The S_{T} cut removes the mis-modelling in simulation due to non-trivial dependence on the jet activity.

Table 5.1: The trigger selections and the events selections for reconstructing the vector boson.

Channel	Event Selection	Description
Zero charged lepton channel	Leptons	0 VH-loose leptons
	$E_{\mathrm{T}}^{\mathrm{miss}}$ S_{T}	> 150 GeV > 120 GeV (2 jets), > 150 GeV (≥ 3 jets)
One charged lepton channel	Leptons	1 WH-signal lepton Veto ≥ 1 VH-loose lepton
	S_{T} (μ channel with $E_{\mathrm{T}}^{\mathrm{miss}}$ trig.)	> 120 GeV (2 jets), > 150 GeV (≥ 3 jets)
Two charged leptons channel	Leptons	2 VH-loose leptons ($\geq 1 ZH$-signal lepton) Same flavour, but opposite in charge (for μ)
	m_{ll}	81 GeV $< m_{ll} <$ 101 GeV

One charged lepton channel

The expected signal process in the one charged lepton channel is $WH \to l\nu c\bar{c}$ (and $WH \to l\nu b\bar{b}$), and the transverse momentum of the W-boson (p_{T}^{W}) is reconstructed as the vectorial sum of the lepton transverse momentum and the $E_{\mathrm{T}}^{\mathrm{miss}}$. The reconstructed p_{T}^{W} is required to satisfy $p_{\mathrm{T}}^{W} > 75$ GeV. For the muon events obtained using the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, the same selections on S_{T} as described in Section 5.1.1 is required.

The final state of the process $WH \to \tau\nu_{\tau}\ c\bar{c}$ is similar to $VH \to \nu\bar{\nu}c\bar{c}$ when the τ-lepton decays hadronically (as τ-jets), and therefore this process can end up in the zero charged lepton channel. Events with one or more reconstructed τ-jets are classified in the one charged lepton channel, with the W boson reconstructed as the vectorial sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the τ-jet with the larger p_{T}.

Two charged leptons channel

The expected signal process in the two charged leptons channel is $ZH \to l^{-}l^{+}c\bar{c}$ (and $ZH \to l^{-}l^{+}b\bar{b}$), where the two leptons are required to have the same flavour, and the lepton with the larger p_{T} is required to satisfy $p_{\mathrm{T}} > 25$ GeV ($p_{\mathrm{T}} > 27$ GeV) for electrons (muons). Further, for di-muon events, the muons are required to have opposite charge. This requirement is not applied for di-electron events since the misidentification rate of the electron charge is non-negligible[1]. To suppress backgrounds having non-resonant lepton pairs, a cut on the reconstructed di-lepton mass is also applied; 81 GeV$< m_{ee}(m_{\mu\mu})$ <101 GeV.

---

[1] The reconstruction of the electron charge relies on the measurement of the curvature of its associated track in the inner detector. Hence, mis-identification can happen due to the incorrect determination of the track curvature, or by the choice of an incorrect track. The probability for charge misidentification can vary from a few percent to around 8% in higher η regions [114].

5.1.2 Event selections to suppress background processes

Several event selections (detailed in Table 5.2) are applied in the analysis to suppress the QCD multi-jet and the top quark process backgrounds.

Supressing QCD multi-jet processes

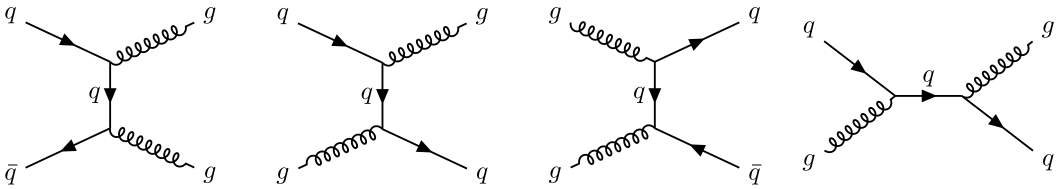


Figure 5.1: Examples of leading order multi-jet processes from proton-proton collisions.

Figure 5.1 shows some examples of leading order diagrams for QCD multi-jet processes at the LHC. QCD multi-jets are created back-to-back in the transverse plane, and therefore specific cuts to exploit the multi-jet topology can suppress this background greatly in the analysis phase space. Further, QCD multi-jet background contributes to the analysis phase space mainly from the high $E_{\mathrm{T}}^{\mathrm{miss}}$ of mis-measured jets in the calorimeters. In the two charged leptons channel, requiring the reconstructed mass of the two charged leptons to match the Z boson mass can reduce the QCD multi-jet contribution to a negligible level.

In the zero charged lepton channel, QCD multi-jet background can be efficiently removed, to around 1% of the total background, by requiring below selections using the azimuthal angle ϕ.

- The difference in azimuthal angle between the two jets forming the Higgs candidate, $|\Delta\phi(j_1, j_2)|$, to be less than $140°$.
- The difference in azimuthal angle $\min\left(|\Delta\phi(|E_{\mathrm{T}}^{\mathrm{miss}}, j_i)|\right)$ between the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector and the closest jet to be more than $20°$ $(30°)$ for events with 2 jets ($\geq$ 3 jets), since the multi-jet contribution is confined to low values of $\min\left(|\Delta\phi(|E_{\mathrm{T}}^{\mathrm{miss}}, j_i)|\right)$.
- The difference in azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and the Higgs candidates four-vector, $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, H)|$, to be greater than $120°$.

In the one charged lepton channel, the QCD multi-jet backgrounds are suppressed by requiring the transverse mass of the reconstructed W-boson,

$$m_{\mathrm{T}}^{W} = \sqrt{2p_{\mathrm{T}}^{l}E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos(\Delta\phi(l, E_{\mathrm{T}}^{\mathrm{miss}}))}, \tag{5.1}$$

to be larger than 20 GeV in the low p_{T}^{V} region (75 GeV< p_{T}^{V} <150 GeV). Further,

- For the electron channel, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be larger than 30 GeV.
- For hadronically decaying τ-jet events re-classified to the one charged lepton channel from the zero charged lepton channel, the multi-jet background is suppressed by the same three cuts on the azimuthal angle described above, while also requiring $m_{\mathrm{T}}^{W} > 10\,\mathrm{GeV}$ to remove the fake τ contamination from jets.

Supressing top processes in the one charged lepton channel

Two additional event selections in the one charged lepton channel are aimed at reducing the dominant background processes from the decay of the top quark, where one of the W bosons decays leptonically.

- When one of the W-bosons from the $t\bar{t}$ process decays in the electron or muon channel, and the other W-boson decays to a τ-lepton and a neutrino, and the τ-lepton subsequently decays hadronically, such events can end up in the $VH, H \to c\bar{c}$ (or $VH, H \to b\bar{b}$) signal regions. Hence, events with one or more τ-jets in the one charged lepton channel are discarded. It should be noted that this selection is independent to the re-classification of τ-leptons from the zero charged lepton channel, and does not discard the re-classified events.
- Events with more than one additional jet are rejected in the $VH, H \to c\bar{c}$ signal regions (described in Section 5.2), since top quark processes are likely to contain more than 3 jets in the event.

Table 5.2 summarizes all the general selections listed above.

Table 5.2: Events selections to suppress QCD multijet and top quark processes.

Channel	Event Selection	Description
Zero charged lepton channel	m_{T}^{W}	> 10 GeV (for ≥ 1 had. τ events)
	$\lvert\Delta\phi(j_1, j_2)\rvert$	$< 140^\circ$
	$\min\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, j_i)\rvert$	$> 20^\circ$ (2 jets), $> 30^\circ$ (≥ 3 jets)
	$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, H)\rvert$	$> 120^\circ$
One charged lepton channel	Hadronic τ veto	No hadronically decaying τ's
	$E_{\mathrm{T}}^{\mathrm{miss}}$ (e channel)	> 30 GeV
	m_{T}^{W}	> 20 GeV (for $75\mathrm{GeV} < p_{\mathrm{T}}^{V} < 150\mathrm{GeV}$)

After the above event selections, 8.4% of the $ZH \to \nu\bar{\nu}c\bar{c}$ events recorded by the ATLAS experiment (i.e in 140.1 fb^{-1}) will remain in the zero-charged lepton channel, and the contribution of $ZH \to \nu\bar{\nu}c\bar{c}$ to the one and two charged leptons channels is negligible. Further, around 1% of $WH \to l\nu c\bar{c}$ events will also end up in the zero charged lepton channel, before the migration of hadronically decaying τ-leptons to the one charged lepton channel. 7.9% of $WH \to l\nu c\bar{c}$ events in 140.1 fb^{-1} will be retained in the one charged lepton channel, while less than 2% of the total $ZH \to l^-l^+c\bar{c}$ events end up in the one charged lepton channel (which is negligible compared to $WH \to l\nu c\bar{c}$). Finally, the two charged leptons channel retains 12.9% of the $ZH \to l^-l^+c\bar{c}$ events in 140.1 fb^{-1}, with the contribution of $WH \to l\nu c\bar{c}$ to be negligible. The acceptance for $VH, H \to b\bar{b}$ events in the different charged lepton channels is also similar to the $VH, H \to c\bar{c}$ events mentioned above.

5.2 The jet-flavor tagging selection

To reconstruct the $H \to c\bar{c}$ (or $H \to b\bar{b}$) candidates (also referred to as the *Higgs boson candidate*), ideally only two jets that are identified as c-jets (or b-jets) are required. However, due to limitations in jet flavor tagging, and since there are cases where there are more than two jets in the final state that come from initial state radiation (ISR), final state radiation (FSR), or pile-up events, the jet-flavor tagging selection depicted in Figure 5.2 is required to reconstruct the Higgs candidate, and is defined using the custom 2D flavor tagging scheme discussed in Section 4.9.3.

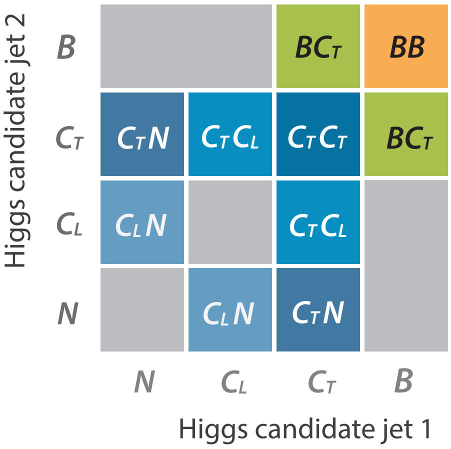


Figure 5.2: The flavor tag regions used in the $VH, H \to b\bar{b}/c\bar{c}/$ analysis to classify events based on the two Higgs candidate jets.

Before the jet-flavor tagging selection, the jets in each event are assigned to the flavor tag bin in the order B, C_T, C_L, N. Therefore, the bin B has the highest priority, and jets that are not assigned B will be assigned C_T, C_L, or N, with C_T having the highest priority. Finally, jets that are not assigned to any of the bins B, C_T, or C_L, fall into the bin N.

The jets that are assigned to the bin B are called b-tagged, and jets assigned to bins C_T, C_L are called c-tagged. If an event contains exactly two b-tagged jets, then this event enters the BB region shown in Figure 5.2. From all two jet events having true b-jets, such as the $VH, H \to b\bar{b}$ process with exactly two b-jets, the BB region will contain roughly 50% of the events[2]. For events with greater than two jets, the BB region still has a higher purity of true b-jets, and therefore, the BB region is defined as the signal region for the $VH, H \to b\bar{b}$ process.

Events that do not contain any b-tagged jets will be selected to define regions such as C_TC_T, C_TC_L, C_TN, etc. However, the definition of the $VH, H \to c\bar{c}$ signal region using the non b-tagged regions is slightly complex. As seen from Table 5.3, the C_TN, C_LN, or NN regions contain most of the $VH, H \to c\bar{c}$ signal, but these regions also contain a lot of background.

Therefore, considering also the contributions from background processes in the flavor tag regions, the poisson significance in each flavor tag regions (calculated using Equation 5.2 [26]) is used as a figure-of-merit in determining the $VH, H \to c\bar{c}$ signal regions.

---

[2] This is evident from Table 4.2, where the selection efficiency for two b-jets in the BB region can be calculated as $\epsilon_b(B) \times \epsilon_b(B) = 69.2\% \times 69.2\% = 47.9\%$

Table 5.3: Relative acceptance of $VH, H \to c\bar{c}$ signal processes in the different flavor tag regions.

Lepton Channel	**Flavor tag region**									
	$\boldsymbol{BB}$	$\boldsymbol{BC_T}$	$\boldsymbol{BC_L}$	$\boldsymbol{BN}$	$\boldsymbol{C_TC_T}$	$\boldsymbol{C_TC_L}$	$\boldsymbol{C_TN}$	$\boldsymbol{C_LC_L}$	$\boldsymbol{C_LN}$	$\boldsymbol{NN}$
Zero	0.67%	3.33%	3.35%	7.73%	4.04%	8.44%	19.94%	4.46%	20.97%	27.07%
One	0.66%	3.05%	3.18%	7.97%	3.68%	7.91%	20.02%	4.20%	21.19%	28.14%
Two	0.85%	3.65%	3.73%	8.02%	3.94%	8.64%	19.20%	4.757%	20.80%	26.39%

$$S = \sqrt{\sum_i 2\left[(s_i + b_i)\ln\left(1 + \frac{s_i}{b_i}\right) - s_i\right]}, \tag{5.2}$$

where s_i and b_i are the total number of signal and background[3] events, respectively, contained in the i_{th} bin of the Higgs candidate mass distribution[4]. Table 5.4 shows the poisson significance in each flavor tag region for the $VH, H \to c\bar{c}$ process, and it is seen that the C_TC_T, C_TC_L, and the C_TN regions have the highest significance for the $VH, H \to c\bar{c}$ process. Hence, the $\boldsymbol{C_TC_T + C_TC_L + C_TN}$ region is defined as the signal region of the $VH, H \to c\bar{c}$ process.

Table 5.4: The poisson significance for the $VH, H \to c\bar{c}$ processes in the different flavor tag regions.

Lepton Channel	**Flavor tag region**									
	$\boldsymbol{BB}$	$\boldsymbol{BC_T}$	$\boldsymbol{BC_L}$	$\boldsymbol{BN}$	$\boldsymbol{C_TC_T}$	$\boldsymbol{C_TC_L}$	$\boldsymbol{C_TN}$	$\boldsymbol{C_LC_L}$	$\boldsymbol{C_LN}$	$\boldsymbol{NN}$
Zero	0.002	0.014	0.011	0.011	**0.068**	**0.059**	**0.056**	0.023	0.027	0.015
One	0.002	0.011	0.008	0.009	**0.056**	**0.048**	**0.046**	0.021	0.026	0.015
Two	0.002	0.018	0.011	0.010	**0.051**	**0.044**	**0.038**	0.017	0.019	0.010

The two b-tagged jets, labelled as j_1 and j_2, are used to build the $VH, H \to b\bar{b}$ candidate. Here j_1 is the jet with the larger transverse momentum (p_{T}) (referred to as the *leading jet*), and is required to satisfy $p_{\text{T}} > 45$ GeV. j_2 is the sub-leading jet (i.e the jet with the second largest p_{T}) which is required to satisfy $p_{\text{T}} > 20$ GeV. Similar to the BB region, in the C_TC_T, C_TC_L, and the C_TN regions, j_1 (j_2) is assigned to the leading (sub-leading) jet used to build the Higgs boson candidate. For events containing more than two jets, the additional jets are referred to as *non-Higgs boson candidate jets*, are ordered in p_{T} and labeled j_i, where $i > 2$. Additional jets are also required to satisfy $p_{\text{T}} > 20$ GeV.

Table 5.5 summarizes all the jet-flavor tagging related selections, including selections listed below to avoid regions that show data and simulation disagreement, and far from the $VH, H \to c\bar{c}$ (and $VH, H \to b\bar{b}$) signal regions.

---

[3] Note that QCD multi-jet contribution is not included in the calculation of the poisson significance for the one charged lepton channel.

[4] The Higgs candidate mass distribution is suitable to determine the significance since it provides a good signal to background discrimination, due to the $VH, H \to c\bar{c}$ process peaking at 125 GeV.

- The Higgs boson candidate mass, which is the invariant mass reconstructed from j_1 and j_2 is required to satisfy $m_H \equiv m_{j_1 j_2} > 50$ GeV. This is due to difficulty in simulating the gluon splitting in the V+jets process at low m_H, which was identified in the previous analysis.
- The angular separation between Higgs candidate jets j_1 and j_2 defined as $\Delta R(j_1, j_2) = \sqrt{\Delta\phi(j_1, j_2)^2 + \Delta\eta(j_1, j_2)^2}$, is required to satisfy $\Delta R(j_1, j_2) < \pi$.

Table 5.5: Common events selections and lepton channel specific events selections for reconstructing the Higgs boson candidate. 0L, 1L, 2L stand for the zero, one, and two charged leptons channels. H_{cand} is an abbreviation for higgs candidate, and FTAG is an abbreviation for flavor tag.

Type	**Event Selection**	$VH, H \to b\bar{b}$	$VH, H \to c\bar{c}$
Common	Number of jets	≥ 2 signal jets	
	Signal region	BB	$C_T C_T + C_T C_L + C_T N$
	p_{T} of leading H_{cand} jet (j_1)	> 45 GeV	
	p_{T} of sub-leading H_{cand} jet (j_2)	> 20 GeV	
	H_{cand} mass m_H	> 50 GeV	
	Solid angle $\Delta R(j_1, j_2)$	$< \pi$	
0L channel	Number of non-H_{cand} jets	≤ 2 jets	≤ 1 jet
	Non H_{cand} jet (j_i) FTAG bin	Not B or C_T	Not B
1L channel	Number of non-H_{cand} jets	≤ 1 jet	
	Non H_{cand} jet (j_i) FTAG bin	Not B or C_T	Not B
2L channel	Non H_{cand} jet (j_i) FTAG bin	Not B	Not B

After the above mentioned selections, the contribution of the signal and background processes in the flavor tag regions used in the analysis are shown in Figure 5.3, where it is seen that the dominant background contribution to the $VH, H \to c\bar{c}$ signal region is from V+jets processes, and the second dominant contribution comes from top quark processes.

5.2.1 Control regions for top and V+jets background

As seen from Figure 5.3, the $\boldsymbol{BC_T}$ region is enriched with top quark processes, and the $\boldsymbol{C_L N}$ region is enriched with V+light jet processes. Therefore, these two regions are used as control regions for the top quark and V+light jet processes in the analysis.

Top quark process control region

The top quark processes with a b-jet and c-jet (or light-jet), from the same hadronically decaying top quark, is a dominant background in the zero and one charged lepton channel of the $VH, H \to c\bar{c}$ signal regions (as seen from Figure 5.3). Hence, the top bq control region is de-

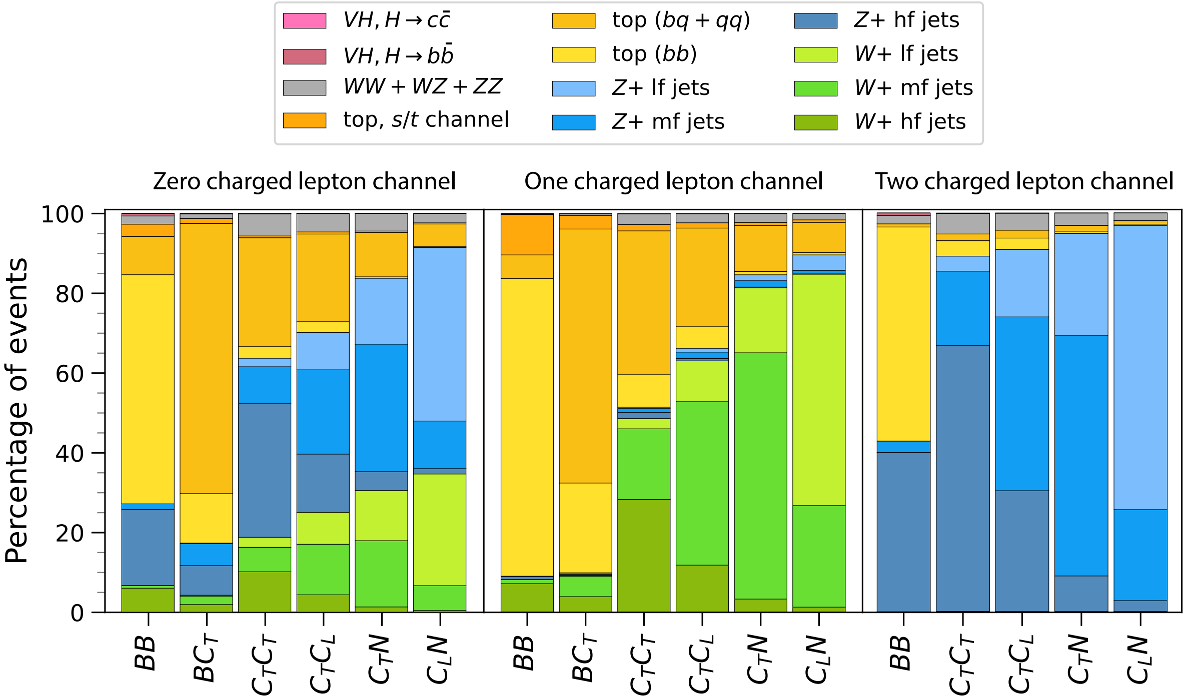


Figure 5.3: The percentage of events in the flavor tag regions of the zero, one, and two charged leptons channels. The yields are obtained using simulated events that pass the event selections, but before additional categorization. The flavor combination $bq + qq$ is an abbreviation for $(bc + bl + cl + cc + ll)$ flavors, and hf, mf, lf are abbreviations for the heavy flavor combination $(bb + cc)$, mixed flavor combination $(bc + bl + cl)$, and the light flavor combination (ll), respectively.

fined to constrain this background, and consists of events with at least one b-tagged jet and one tight c-tagged jet (i.e the BC_T region). It is important to note that this region is completely orthogonal and avoids overlap with the BB and $C_TC_T + C_TC_L + C_TN$ signal regions. The Higgs candidate is reconstructed from the leading p_{T} jet of the jets that are classified into the bin B, and the leading p_{T} jet of the jets that are classified into bin C_T.

In the two charged leptons channel, the top quark background arises mainly from the di-leptonic $t\bar{t}$ process. Therefore, this control region is defined in the two charged leptons channel, with the same selection as that of the signal region, but requires an electron and muon (instead of the same flavor leptons). This significantly increases the fraction of top quark events since other backgrounds are expected to give same flavor leptons from the Z-boson decay. This region provides the top quark background prediction in the $VH, H \to b\bar{b}$ signal region in a data-driven method. In the $VH, H \to c\bar{c}$ signal regions, due to low statistics, the top quark $e\mu$ control region is used to constrain the normalization on the simulated background templates.

V+light jet process control region

The V+light jet control region consists of Higgs candidate jets with one being loose c-tagged, and one being not tagged (i.e the C_LN region). Since this region is dominant in the V+light flavor component as seen from Figure 5.3, it is used to constrain the W+light flavor processes from the one charged lepton channel, and Z+light flavor processes from the two charged leptons channel.

5.3 The event categorization

After the general event selections and the jet-flavor tagging selections, the events are categorized into different regions based on the number of selected jets in the event, the transverse momentum of the vector boson (p_{T}^{V}), signal regions, and additional control regions. This categorization is aimed at maximizing the $H \rightarrow c\bar{c}$ and $H \rightarrow b\bar{b}$ sensitivity in the signal regions, while constraining the modeling of background processes in the control regions. This section gives a detailed explanation on the event categorization, while a visual representation of the categorization is shown in Figure 5.4. Plots showing the contribution of the different background processes in the categorized regions can be found in Appendix E.3.

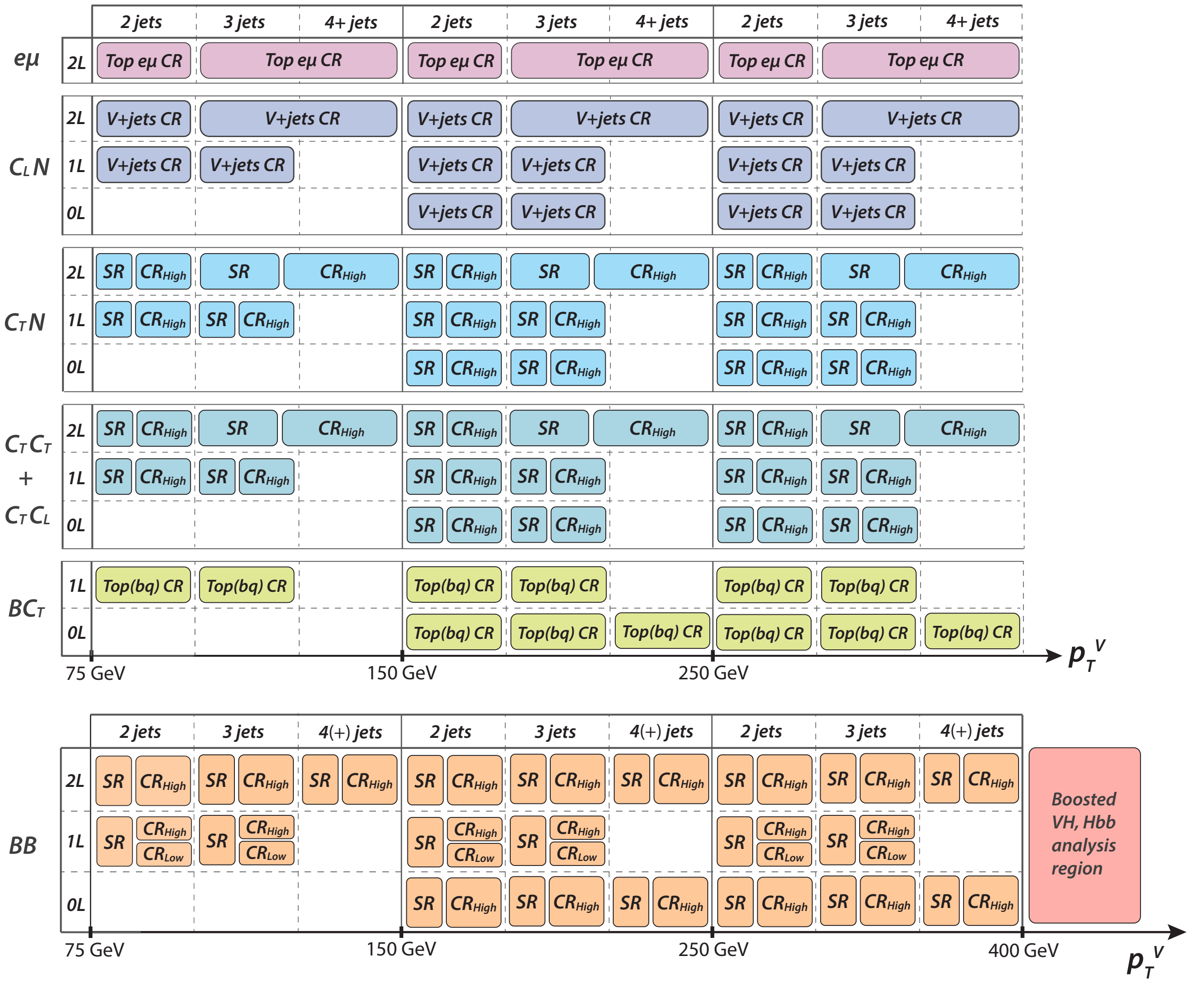


Figure 5.4: The $VH, H \rightarrow c\bar{c}$ and $VH, H \rightarrow b\bar{b}$ analysis regions. *SR* and *CR* stands for the signal region and the control region, respectively.

5.3.1 Number of jets and p_{T}^{V} categorization

The number of selected jets (referred to as *n-jets*) in the event is the sum of the signal jets ($|\eta| < 2.5$), and forward jets ($2.5 < |\eta| < 4.5$). The $VH, H \rightarrow c\bar{c}$ signal regions are categorized in the jet multiplicities of 2 jets, 3 jets, 4 jets, or ≥ 5 jets, depending on the charged lepton channel. In the zero and one charged lepton channels, events with more than 3 jets are rejected due to the large

$t\bar{t}$ background. In the two charged leptons channel, there is no requirement on the number of selected jets since the $t\bar{t}$ background is small; the background is from $t\bar{t}$ di-lepton events, where two jets are expected to b-jets. The n-jet categorization in $VH, H \to b\bar{b}$ signal regions are also similar, except for the zero charged lepton channel, where the 4 jet category is also included due to sufficient sensitivity.

As described in Section 4.9.2, the signal jets are required to have a $p_{\mathrm{T}} > 20$ GeV. Further, to reduce the signal migration between the jet multiplicity bins, the non-Higgs candidate signal jets are required to have a $p_{\mathrm{T}} > 30$ GeV. It is important to note that this cut does not reject events, but re-categorizes such events from higher jet multiplicity to lower jet multiplicity.

The $VH, H \to c\bar{c}$ signal regions are further categorized in three regions based on p_{T}^{V}; 75 GeV $< p_{\mathrm{T}}^{V} < 150$ GeV, 150 GeV $< p_{\mathrm{T}}^{V} < 250$ GeV, and $p_{\mathrm{T}}^{V} > 250$ GeV regions. The p_{T}^{V} can be used in place of the transverse momentum of the Higgs boson (p_{T}^{H}), since the Higgs boson and the vector boson in the VH process are created back to back in the transverse plane. However, using p_{T}^{V} is preferred due to the better resolution in the reconstructed p_{T}^{V}, compared with the p_{T}^{H} reconstructed using jets.

The p_{T}^{V} region categorization in the $VH, H \to b\bar{b}$ signal region (the BB region) is similar to the $VH, H \to c\bar{c}$ signal regions, but the $p_{\mathrm{T}}^{V} > 250$ GeV region is split further into 250 GeV $< p_{\mathrm{T}}^{V} <$ 400 GeV, 400 GeV $< p_{\mathrm{T}}^{V} < 600$ GeV, and $p_{\mathrm{T}}^{V} > 600$ GeV regions[5].

5.3.2 The signal region and ΔR control regions

The opening angle $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ (of Equation 2.4) between the Higgs candidate jets is expected to follow the relation $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2}) \approx 2m_H/p_{\mathrm{T}}^{H}$, where m_H and p_{T}^{H} are the Higgs candidate mass and transverse momentum, respectively. Replacing p_{T}^{H} with p_{T}^{V}, it is possible to introduce a continuous cut based on $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$, which can be used to separate the $H \to c\bar{c}$ and $H \to b\bar{b}$ processes from background processes which does not show such a relation. Hence, the signal regions for the $VH, H \to c\bar{c}$ process ($C_T C_T + C_T C_L + C_T N$) and $VH, H \to b\bar{b}$ process (BB) are split further into a ΔR signal region and ΔR control regions, based on the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ of the Higgs candidate jets. Note that the ΔR signal region will be referred to as simply the *signal region* in the rest of the text in this thesis.

In the previous $VH, H \to c\bar{c}$ analysis, the split of the regions were based on a fixed $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ value which resulted in 80% of signal events entering the signal region below the fixed $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cut [30]. In the new $VH, H \to b\bar{b}/c\bar{c}$ analysis, the split of the signal and the ΔR regions is further optimized using a continuous cut based on $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$, which retains more $VH, H \to c\bar{c}$ signal events in the signal region. There are two types of ΔR control regions in the analysis. The high ΔR control region is used to constrain the shape and normalization of the V+jets and $t\bar{t}$ background processes, while the low ΔR control region is used to constrain the shape and normalization of the W+jets in the one charged lepton channel.

---

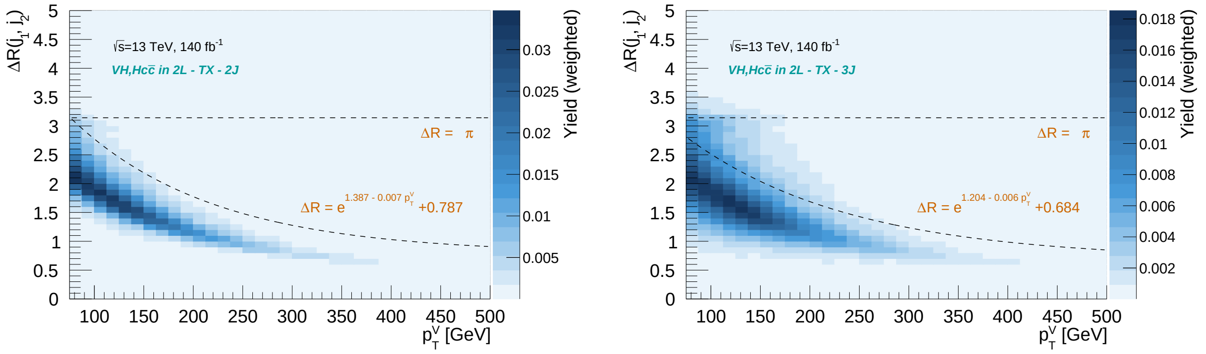
[5] The 400 GeV $< p_{\mathrm{T}}^{V} <$ 600 GeV and $p_{\mathrm{T}}^{V} >$ 600 GeV regions fall under the boosted $VH, H \to b\bar{b}$ analysis, which is not discussed in this thesis. However, due to the concern of double counting the data from $VH, H \to c\bar{c}$ signal regions in the $p_{\mathrm{T}}^{V} > 250$ GeV region, and the boosted $VH, H \to b\bar{b}$ signal region, the overlap of events was investigated and was concluded to be negligible; 99.6 % of $VH, H \to c\bar{c}$ signal regions events do not overlap with the boosted $VH, H \to b\bar{b}$ signal regions.

The high ΔR control region

The continuous cut functions of Table 5.6 are derived in the two charged leptons channel of the $VH, H \rightarrow b\bar{b}$ signal region, where 95% (85%) of signal events are taken to end up below the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})\,^{high}_{cut}$ cut for the 2 jet (≥3 jets) regions. These continuous cuts are then applied to all charged lepton channels in both the $VH, H \rightarrow b\bar{b}$ signal region and $VH, H \rightarrow c\bar{c}$ signal regions, to split the events to the signal region and the high ΔR control region. Figure 5.5 illustrates these cuts for the $VH, H \rightarrow c\bar{c}$ signal process in the 2-lepton channel.

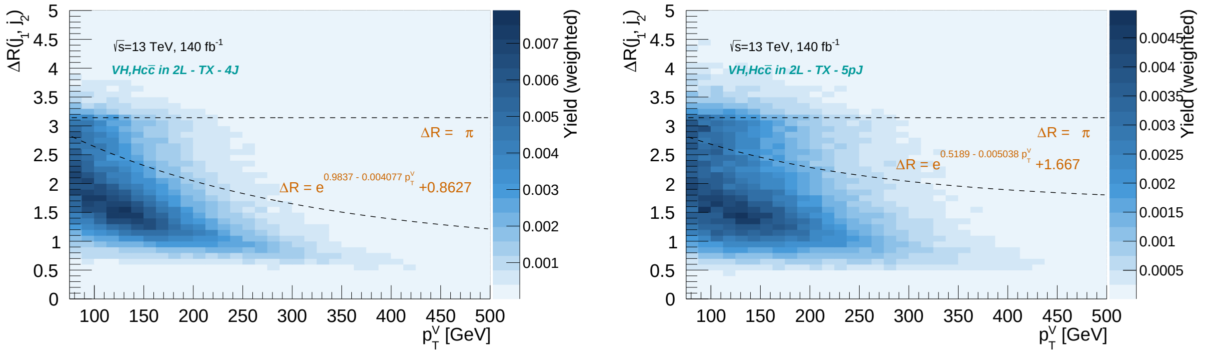
Table 5.6: The $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cut used to separate the signal and the high ΔR control region.

Jet category	**Cut function**
2 jets	$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})\,^{high}_{cut} = 0.787 + e^{1.387-0.00700\times p_{\mathrm{T}}^{V}}$
3 jets	$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})\,^{high}_{cut} = 0.684 + e^{1.204-0.00600\times p_{\mathrm{T}}^{V}}$
4 jets	$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})\,^{high}_{cut} = 0.863 + e^{0.984-0.00408\times p_{\mathrm{T}}^{V}}$
≥ 5 jets	$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})\,^{high}_{cut} = 1.667 + e^{0.519-0.00504\times p_{\mathrm{T}}^{V}}$



(a) $ZH \rightarrow l^-l^+c\bar{c}$ with 2-jets in the C_TC region

(b) $ZH \rightarrow l^-l^+c\bar{c}$ with 3-jets in the C_TC region.



(c) $ZH \rightarrow l^-l^+c\bar{c}$ with 4-jets in the C_TC region.

(d) $ZH \rightarrow l^-l^+c\bar{c}$ with ≥5-jets in the C_TC region.

Figure 5.5: The $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cut to separate the signal region and the high ΔR control region. The C_TC region is the merged $C_TC_T + C_TC_L$ regions. The cut at $\Delta R = \pi$ is added to remove the region with Data-MC disagreement, as explained in Section 5.2.

The motivation behind harmonization of the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cuts between the charged lepton channels, and the $VH, H \to c\bar{c}$ and $VH, H \to b\bar{b}$ signal regions is to correlate the SR-CR extrapolation uncertainties and normalisation factors between different regions. Further, it was investigated how much signal events are contained in the signal regions of the zero and one charged lepton channels, and the $VH, H \to c\bar{c}$ signal regions. It is seen that the for the BB and $C_T C_T$ regions, 95% (85%) of the 2 jet (≥3 jets) signal is contained in the signal region, in all charged lepton channels. However, for the $C_T C_L$ and $C_T N$ regions, jets not originating from the Higgs boson are likely to get tagged, giving a larger $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$. The details of this study can be found in Appendix E.2.

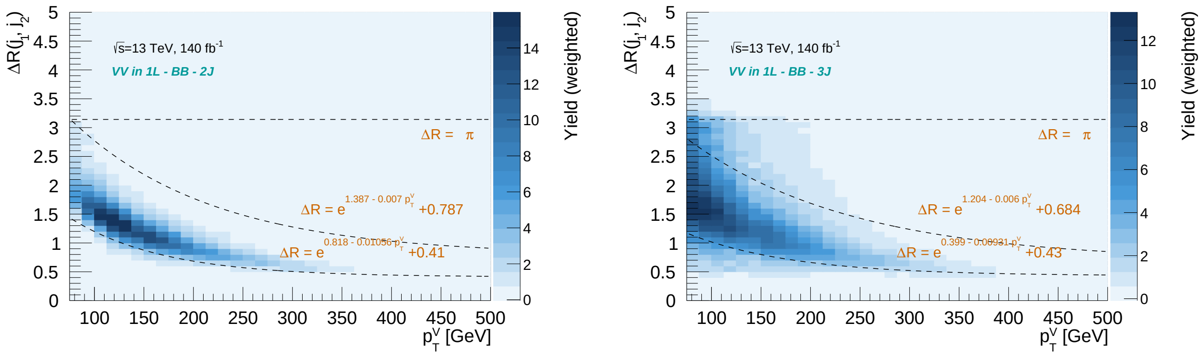
The low ΔR control region

The low ΔR control region is only defined in the one charged lepton channel of the BB tag region (i.e the $VH, H \to b\bar{b}$ signal region), and is below the continuous cuts given in Table 5.7. This cut is defined such that 90% of the $VZ, Z \to b\bar{b}$ events remain in the signal region, since the same signal region is also used in the di-boson validation analysis.

Table 5.7: The continuous $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cut which is applied to separate the signal and the low ΔR control region.

Jet category	Cut function
2 jets	$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})\,^{low}_{cut} = 0.410 + e^{0.818-0.01056\times p_{\mathrm{T}}^{V}}$
3 jets	$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})\,^{low}_{cut} = 0.430 + e^{0.399-0.00931\times p_{\mathrm{T}}^{V}}$
4 jets	$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})\,^{low}_{cut} = 0.411 + e^{1.204-0.0006\times p_{\mathrm{T}}^{V}}$
≥ 5 jets	$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})\,^{low}_{cut} = 0.501 + e^{1.192-0.00753\times p_{\mathrm{T}}^{V}}$

Figure 5.6 shows the low ΔR cuts in the one charged lepton channel BB region for di-boson events, where it is seen most events remain in the signal region.



(**a**) di-boson with 2-jets in the BB tag region

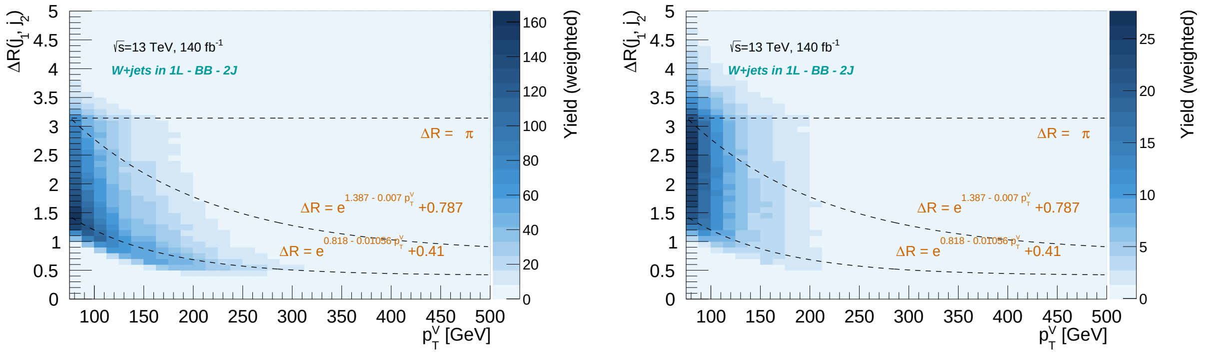
(**b**) di-boson with 3-jets in the BB tag region.

Figure 5.6: The $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cuts in the one charged lepton channel of the BB tag region. The low ΔR control region is defined for the events that enter the region below the bottom curve. Above this region is the signal region, and above the signal region is the high ΔR control region.

The low ΔR control region is enriched in $V+$ heavy flavor jets $(bb + cc)$, with low mixed flavor $(bc + bl + cl)$ jet contamination. Hence, the low ΔR control region is used to constrain the $W+$

heavy flavor background component in the BB, $C_TC_T + C_TC_L$ and C_TN signal regions. The low ΔR control region is not used in the zero and two charged leptons channels to constrain the $Z+$ heavy flavor component, since the impact was found to be negligible.

Figure 5.7 shows how the $W+$ heavy flavor component is distributed more closer to the lower ΔR cut, while the $W+$ mixed flavor component is distributed at relatively higher $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ values.



(**a**) $W + hf$ process with 2-jets in the BB tag region (**b**) $W + mf$ process with 2-jets in the BB tag region

Figure 5.7: The distribution of W+jets in the one charged lepton channel of the BB regions. hf and mf refer to *heavy flavor* and *mixed flavor* respectively.

Chapter 6

Truth flavor tagging for enhancing simulated statistics

As discussed in Section 5.2, the Higgs candidate selection will be based on the jet-flavor tagging selection. Using the $DL1_r$ bins C_T or C_L to define the $VH, H \to c\bar{c}$ working points will reject most events with true b-jets and light-jets, and increase the relative fraction of true c-jets. A similar scenario holds true for the $DL1_r$ bin B used to define the $VH, H \to b\bar{b}$ working point, where it is necessary to reject events with true c-jets and light-jets, and increase the relative fraction of true b-jets.

While this *direct flavor tagging* selection is essential to reject the background events and enhance the signal events in the selected data sample, it is not ideal for the simulated samples. The high rejection factor (or in other words, low flavor tagging efficiency as seen in Table 4.2) for true b-jets and light-jets in the $VH, H \to c\bar{c}$ working points (or c-jets and light-jets in the $VH, H \to b\bar{b}$ working point) lead to non-smooth distributions with large statistical uncertainties attributed to the simulated samples. The top diagram in Figure 6.1 shows an illustration of the direct flavor tagging creating a statistically limited distribution.

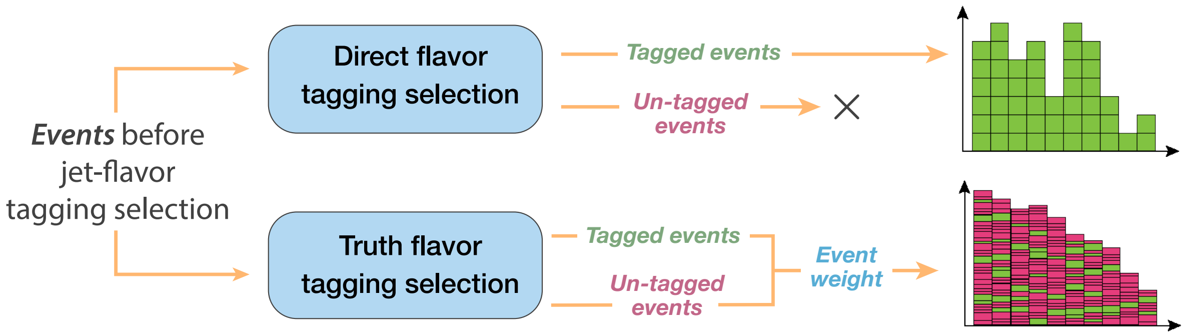


Figure 6.1: An illustration of the usual direct flavor tagging selection, and the event weighting technique that is also known as truth flavor tagging. Green represents events which originally pass the flavor tagging selection, and red represents events which originally failed the flavor tagging selection. While only the green events enter the final distribution for *direct flavor tagging* (top illustration), all the green and red events will be used to fill the final distribution in the *truth flavor tagging* technique (bottom illustration).

An alternate method to the direct flavor tagging selection for simulated samples is an event weighting technique illustrated in the bottom diagram of Figure 6.1. The event weighting method effectively uses all the simulated samples, and re-weights them based on the probability of the event entering a particular working point. This probability is calculated using the per jet flavor tagging efficiency estimated using truth flavor information of the jet, and hence, this method is referred to as *truth flavor tagging*.

6.1 Truth flavor tagging theory

The flavor tagging efficiency for a jet of flavor f to be tagged in a working point defined using a classifier trained on a set of variables denoted as **x** can be denoted as $\epsilon_f(\mathbf{x})$. In the $VH, H \rightarrow c\bar{c}$ analysis, this classifier is the DL1$_\mathrm{r}$ flavor tagging algorithm, and the variables **x** are tracks or secondary vertices introduced in Section 4.5. In truth flavor tagging, it is assumed that ϵ_f (denoted below as ϵ for simplicity) can be parametrised as a function of a set of variables $\boldsymbol{\theta}$ (for example, p_T or η of a certain jet) [130].

The parametrised flavor tagging efficiency $\epsilon(\mathbf{x}|\boldsymbol{\theta})$ can be used to define a probability (or *weight*) for each simulated event to pass the flavor tagging selection. This probability is expressed as

$$P = \sum_{i}^{C} \epsilon_{tag}(T_i, \mathbf{x}, \boldsymbol{\theta}) \cdot \epsilon_{not-tag}(\cancel{T}_i, \mathrm{x}, \boldsymbol{\theta}), \tag{6.1}$$

where i is the jet under consideration, C is the total number of jets in the event, and

$$\epsilon_{tag}(T_i, \mathbf{x}, \boldsymbol{\theta}) = \prod_{j \in T_i} \epsilon_j(\mathbf{x}|\boldsymbol{\theta}) \qquad \text{and} \qquad \epsilon_{not-tag}(\cancel{T}_i, \mathrm{x}, \boldsymbol{\theta}) = \prod_{\mathrm{j} \in \cancel{T}_\mathrm{i}} (1 - \epsilon_\mathrm{j}(\mathbf{x}|\boldsymbol{\theta})). \tag{6.2}$$

The set of tagged jets are represented as T_i, and the set of non-tagged jets are represented as $\cancel{T}_i$. Once the probability P is calculated for each event, then each event is weighted by P before filling the final kinematic distributions.

6.1.1 Determining the event weight and permutation

For events with two jets, the probability P can be simply expressed as the product of the parametrised flavor tagging efficiencies; $P = \epsilon_{jet_1}(\mathbf{x}|\boldsymbol{\theta}) \cdot \epsilon_{jet_2}(\mathbf{x}|\boldsymbol{\theta})$. After the probability is computed, this probability will be used as a weight and applied to the permutation of jets chosen to build the Higgs candidate. For events with only two jets, it suffices to simply apply the weight P to the event, since there exists only one permutation.

For events with three jets, the probability is expressed as shown in Equation 6.3, which takes into account the three permutations for tagging two jets in three jets. Similarly this calculation can be extended to events with four or more jets.

$$
\begin{aligned}
P(\text{2tag3jet}|\boldsymbol{\theta}) =& P(\text{j}_1 \text{ is tagged}|\boldsymbol{\theta}) \cdot P(\text{j}_2 \text{ is tagged}|\boldsymbol{\theta}) \cdot P(\text{j}_3 \text{ is not tagged}|\boldsymbol{\theta}) + \\
& P(\text{j}_1 \text{ is tagged}|\boldsymbol{\theta}) \cdot P(\text{j}_2 \text{ is not tagged}|\boldsymbol{\theta}) \cdot P(\text{j}_3 \text{ is tagged}|\boldsymbol{\theta}) + \\
& P(\text{j}_1 \text{ is not tagged}|\boldsymbol{\theta}) \cdot P(\text{j}_2 \text{ is tagged}|\boldsymbol{\theta}) \cdot P(\text{j}_3 \text{ is tagged}|\boldsymbol{\theta}) \\
=& \epsilon_{j_1}(\mathbf{x}|\boldsymbol{\theta}) \cdot \epsilon_{j_2}(\mathbf{x}|\boldsymbol{\theta}) \cdot (1 - \epsilon_{j_3}(\mathbf{x}|\boldsymbol{\theta})) + \epsilon_{j_1}(\mathbf{x}|\boldsymbol{\theta}) \cdot (1 - \epsilon_{j_2}(\mathbf{x}|\boldsymbol{\theta})) \cdot \epsilon_{j_3}(\mathbf{x}|\boldsymbol{\theta}) \\
& + (1 - \epsilon_{j_1}(\mathbf{x}|\boldsymbol{\theta})) \cdot \epsilon_{j_2}(\mathbf{x}|\boldsymbol{\theta}) \cdot \epsilon_{j_3}(\mathbf{x}|\boldsymbol{\theta})
\end{aligned}
\tag{6.3}
$$

However, the choice of the event to which the weight will be applied is a bit more complex for events with three jets; since there exists three permutations as seen in Equation 6.3, it is necessary to determine which event the weight will be applied to[1].

Therefore, a permutation will be randomly sampled from a 1D distribution. Figure 6.2 shows an example for an event with three jets, where the 1D distribution is constructed using the three probabilities P_1, P_2 and P_3 for each of the three permutations given in Equation 6.3; $P_1 = \epsilon_{j_1}(\mathbf{x}|\boldsymbol{\theta}) \cdot \epsilon_{j_2}(\mathbf{x}|\boldsymbol{\theta}) \cdot (1 - \epsilon_{j_3}(\mathbf{x}|\boldsymbol{\theta}))$, $P_2 = (1 - \epsilon_{j_1}(\mathbf{x}|\boldsymbol{\theta})) \cdot \epsilon_{j_2}(\mathbf{x}|\boldsymbol{\theta}) \cdot \epsilon_{j_3}(\mathbf{x}|\boldsymbol{\theta})$, and $P_3 = \epsilon_{j_1}(\mathbf{x}|\boldsymbol{\theta}) \cdot (1 - \epsilon_{j_2}(\mathbf{x}|\boldsymbol{\theta})) \cdot \epsilon_{j_3}(\mathbf{x}|\boldsymbol{\theta})$.

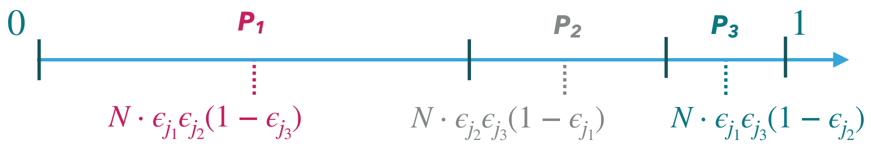


Figure 6.2: The 1D distribution when picking the permutation for truth flavor tagging.

The factor N is applied to normalize the 1D distribution to 1. Once the permutation is determined, the computed probability $P(\text{2tag3jet}|\boldsymbol{\theta})$ is applied to the permutation and filled in the kinematic distributions.

Expanding to multiple flavor tagging bins

The probability shown in Equation 6.3 for the example for three jet events only considers the case where both jets classified into the same flavor tagging bin. This holds for flavor tagging choice in the $VH, H \to b\bar{b}$ two charged lepton channel, where all jets are required to pass or fail the DL1$_\text{r}$ bin B only. However, the zero and one charged leptons channels in $VH, H \to b\bar{b}$ analysis, and all the charged lepton channels in the $VH, H \to c\bar{c}$ analysis follow a slightly more complicated selection (Table 5.5).

The probability can be expanded to include multiple DL1$_\text{r}$ bins, and Equation 6.4 shows the probability for an example where it is required for two jets tagged in bins C_T or C_L, and the third jet should not be tagged in the bin B; this represents a scenario in the $VH, H \to c\bar{c}$ analysis. Here P_1 is the same probability of the first permutation of the example in Equation 6.3, but expanded to multiple DL1$_\text{r}$ bins.

---

[1] The likelihood fit used in the analysis assumes poisson statistics in each bin of the distribution used to extract the $VH, H \to c\bar{c}$ signal strength, and events are uncorrelated between the bins. However, filling a distribution with all available permutations does not satisfy this assumption, since permutations in one event can be filled in multiple bins, and hence create a correlation between bins.

$$\begin{aligned}P_1 =&\epsilon_{j1}^{C_L}\epsilon_{j2}^{C_L}(1-\epsilon_{j3}^{C_L}-\epsilon_{j3}^{C_T}-\epsilon_{j3}^{B})+\epsilon_{j1}^{C_L}\epsilon_{j2}^{C_T}(1-\epsilon_{j3}^{C_L}-\epsilon_{j3}^{C_T}-\epsilon_{j3}^{B})\\ +&\epsilon_{j1}^{C_T}\epsilon_{j2}^{C_L}(1-\epsilon_{j3}^{C_L}-\epsilon_{j3}^{C_T}-\epsilon_{j3}^{B})+\epsilon_{j1}^{C_T}\epsilon_{j2}^{C_T}(1-\epsilon_{j3}^{C_L}-\epsilon_{j3}^{C_T}-\epsilon_{j3}^{B})\\ =&P_{1.1}+P_{1.2}+P_{1.3}+P_{1.4}.\end{aligned} \tag{6.4}$$

Since the number of permutations now increase to 12, the 1D distribution also needs to be adapted to this increase as illustrated in Figure 6.3.

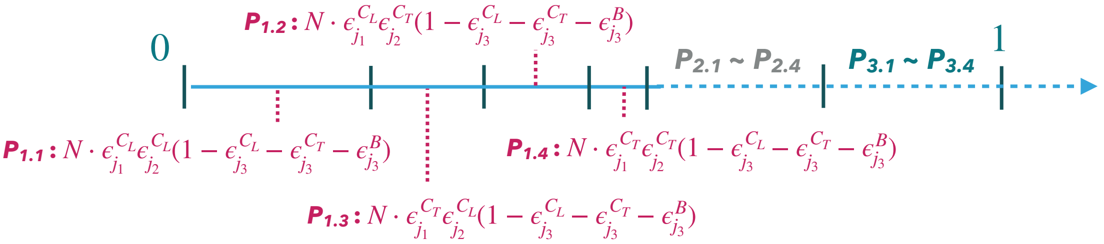


Figure 6.3: The 1D distribution when picking the permutation for truth flavor tagging, expanded to multiple $\mathrm{DL1_r}$ bins.

6.2 Parametrising the flavor tagging efficiency

The output of the $\mathrm{DL1_r}$ flavor tagging algorithm (i.e p_b, p_c and p_u explained in Section 4.5) can only be used to determine the $\mathrm{DL1_r}$ bin the jet is tagged into. However, as evident from the description in Section 6.1, *truth flavor tagging* requires the flavor tagging efficiency per jet, in all the $\mathrm{DL1_r}$ bins used to compute the probability P.

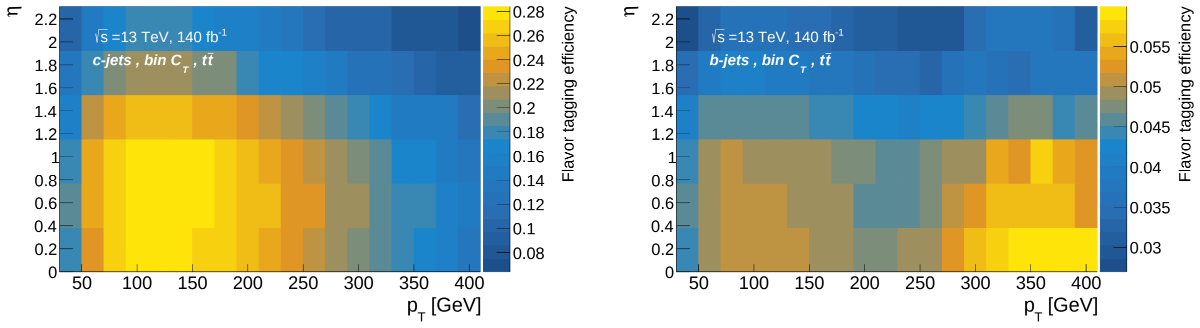
While $\epsilon_f(\mathbf{x})$ is the accurate flavor tagging efficiency required for the truth flavor tagging technique, it is not feasible to determine this directly; since there are a lot of $\mathbf{x}$ parameters, some regions are statistically in-sufficient to accurately determine $\epsilon_f(\mathbf{x})$. Hence, the flavor tagging efficiency $\epsilon_f(\mathbf{x})$ is parametrised by jet variables and events variables $\boldsymbol{\theta}$ as $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$, and the goal is to achieve a $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$ as close as possible to $\epsilon_f(\mathbf{x})$.

6.2.1 2D map based truth flavor tagging

A common method to parametrise the flavor tagging efficiency is the use of 2D map of kinematic variables. The 2D maps are based on the two most dominant $\boldsymbol{\theta}$ parameters affecting $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$ parametrisation; the transverse momentum p_T of the jet, and the η direction of the jet.

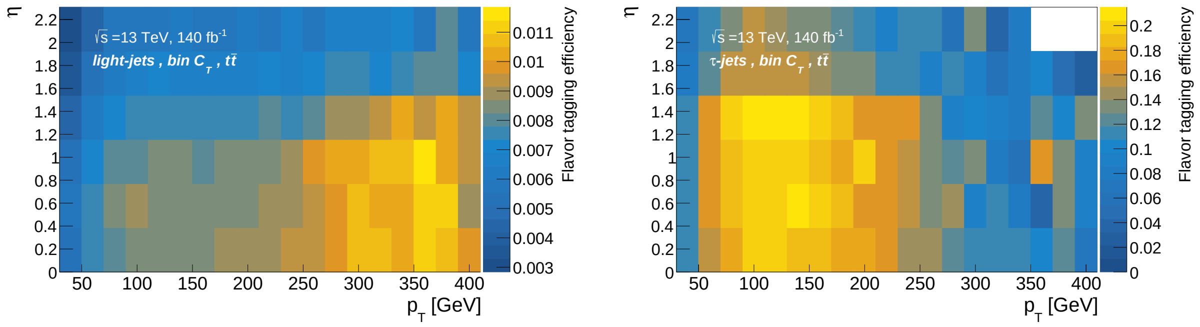
Figure 6.4 shows examples for the 2D maps in the $\mathrm{DL1_r}$ bin C_T, for the four flavors c-jets, b-jets, light-jets and τ-jets. It should be noted that the shown 2D maps are not distributions used in analyses, since prior to use a much coarser re-binning will be applied to avoid the statistical fluctuation in the statistically limited bins (for example, the rightmost bins in Figure 6.4(c)).

The 2D map based parametrisation was used for truth flavor tagging in the previous $VH, H\to c\bar{c}$ analysis, but the limitation of 2D maps in accurately parametrising $\epsilon_f(\mathbf{x}|p_\mathrm{T},\eta)$ contributed to a *normalization effect* on the kinematic distributions obtained through truth flavor tagging. In



(**a**) ϵ_f for c-jets in the C_T bin

(**b**) ϵ_f for b-jets in the C_T bin.



(**c**) ϵ_f for light-jets in the C_T bin.

(**d**) ϵ_f for τ-jets in the C_T bin.

Figure 6.4: The 2D maps used to parametrise the flavor tagging efficiency derived using simulated $t\bar{t}$ events.

addition to the normalisation effect, a $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ dependant *shape effect* was also present in the obtained distributions. Corrections to the two effects contributed to two dominant experimental uncertainties[2] in the previous analysis [30]. However, despite these systematic uncertainties, truth flavor tagging was able to bring an overall 10% increase in the significance (compared to not using truth flavor tagging), since it was able to reduce the statistical uncertainty of the simulated samples.

The normalisation effect and the$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ dependant shape effect can be attributed mainly to the below reasons.

- Maps only take into consideration the two parameters p_T and η which are the dominant contributions to the parametrisation of $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$. However, other parameters like pile-up also affect the parametrisation, and hence, limits the accuracy of parametrising $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$ only using p_T and η.
- At low $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$, which is when the two jets j_1 and j_2 are very close by, fragmentation from one jet can affect the other. When the jet under consideration is a c-jet, and the nearby jet is a b-jet, a b-hadron leaked to the c-jet could mis-tag the c-jet as a b-jet. Further, additional tracks from a b-jet can also affect the tagging efficiency of a c-jet or a light-jet nearby.
- The efficiency maps are constructed per jet flavor as seen in Figure 6.4. However, as dis-

---

[2] The impact on the $VH, H \rightarrow c\bar{c}$ signal strength is 1.7 for the normalisation correction, and 3.3 for the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ based correction. The impact of the total systematic uncertainty was 11.5.

cussed in the previous point, the presence of different flavors in an event also needs to be taken into consideration when parametrising $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$.

- Clusters in a calorimeter for a jet under consideration can be affected by energy deposits from a nearby jet at low $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$. This could shift the jet axis and affect the performance of the JETFITTER algorithm, which assumes that the jet axis is on the flight path of the heavy hadron. Such effects are not taken into consideration in the $\epsilon_f(\mathbf{x}|p_\mathrm{T}, \eta)$ parametrisation using 2D maps.

6.2.2 Graph neural network based truth flavor tagging

The limitations of parametrising $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$ using 2D maps was discussed in Section 6.2.1. However, given the need for a more accurate parametrisation of $\epsilon_f(\mathbf{x})$, the current analysis uses a new technique utilizing graph neural networks.

Parametrizing flavor tagging efficiency using graph neural networks

A graph neural network (GNN) is a type of neural network used in deep learning where data points, called *nodes*, are linked by lines, called *edges*. GNNs are powerful deep learning tools suited for data structures with data points exhibiting some relation between them [131, 132]. GNNs have grown in popularity in high energy physics in recent years, with some examples being the *ParticleNet* flavor tagging algorithm from the CMS experiment [133], and the *GNx* flavor tagging algorithms from the ATLAS experiment [134].

A GNN is a good tool for parametrising the $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$, since it is able to handle multiple input parameters $\boldsymbol{\theta}$ to parametrize $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$ more closer to $\epsilon_f(\mathbf{x})$. Since neural networks are also able to perform regression[3], this leads to smoother $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$ distributions in statistically limited regions seen in 2D maps.

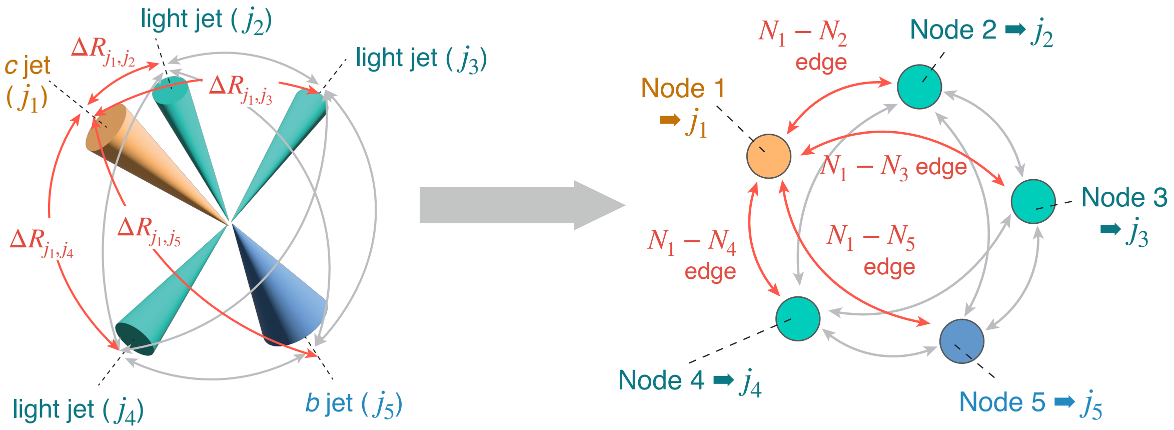


Figure 6.5: The representation of an event topology with jets using a graph. The left illustration is of an event with a c-jet, b-jet, and three light-jets. The right diagram illustrates how this event topology is expressed as a graph, with jets assigned to the nodes, and $\Delta R(\boldsymbol{j_m}, \boldsymbol{j_n})$ between the jets assigned to the respective edges.

---

[3] Regression is where an algorithm can predict a continuous outcome y, based on different input variables $\mathbf{x}$.

The event topology with jets can be represented using the nodes and edges of a graph as illustrated in Figure 6.5. Jets in the event are assigned to the nodes of the graph, while the edges are the angular separation $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ between the jets, representing the natural topology of the event. Another merit of using a GNN is the flexibility in the number of nodes, depending on the number of jets in the event. The use of $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ to define the edges takes into account the separation between the jets when training the algorithm, thereby also parametrising $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$ to $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$. This is a good improvement over the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ limitation of 2D maps.

Training the graph neural network

The inputs to the GNN are fully connected graphs, with each node representing a jet in the event with nine input features. It should be noted that the jets used in the training are only jets in the central region of $|\eta| < 2.5$. Eight out of the nine input features are variables related to each jet, while one input feature, the average pile-up, is a variable related to the event (i.e. this variable is common to all the jets).

Table 6.1: The list of input features to the GNN used for parametrisation of $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$.

Jet **features**	**Type of variable**
p_{T} of the jet η of the jet ϕ of the jet Flavour label (b, c, light, or τ) of the jet Mass of the b or c hadron with the largest p_{T} in the jet p_{T} of the b or c hadron with the largest p_{T} in the jet η of the b or c hadron with the largest p_{T} in the jet ϕ of the b or c hadron with the largest p_{T} in the jet	Jet level variable
Average number of interactions per event, $\langle\mu\rangle$	Event level variable
Angular separation between two jets, ΔR	Jet-pair variable

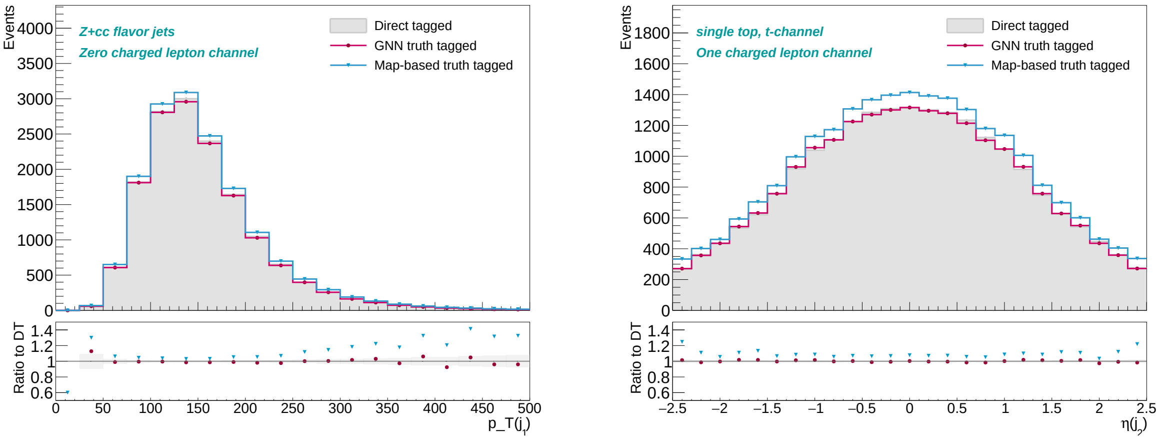
To maximise the simulated statistics used for training the GNN, the samples are merged in the charged lepton channels, p_{T}^{V} regions, and the jet multiplicity regions. Further merging is done for the physics processes made with the same MC generator, with the assumption that there is no difference in jet properties between the processes. However, merging of physics processes made using different MC generators is not possible, since different flavor tagging scaling factors (discussed in Section 4.5.2) are derived for different MC generators. Therefore, one GNN model is trained inclusively for V+jets and di-boson. For $t\bar{t}$, single-top and $VH, H \rightarrow b\bar{b}/c\bar{c}$ signal samples, three independent models are trained after observing better performance compared to one model that is inclusively trained.

The GNN is built using the Deep Graph Library [135], and trained using PyTorch [136]. The GNN model architecture is described briefly in Appendix F.1, and discussed in more detail in [130, 137]. The GNN model is then implemented into the analysis code using the Open Neural Network Exchange, commonly known as ONNX [138].

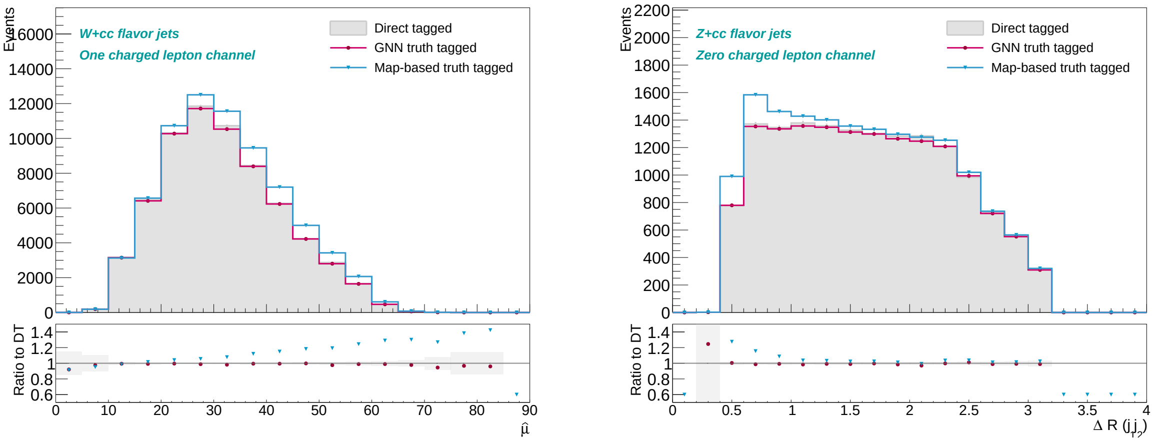
6.2.3 General performance of GNN-based truth flavor tagging

The performance of truth flavor tagging is evaluated by comparing the truth flavor tagged distributions to the direct flavor tagged distributions. If $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$ is accurately parametrised, then the yield and shape of the truth flavor tagged distributions should match to the direct flavor tagged distributions, within the statistical uncertainties of both distributions; this is referred to as the *closure*.

Figure 6.6 shows the closure of several kinematic distributions that are truth flavor tagged using the GNN model (red line) and the 2D maps (blue line) for V+jets. It is noticed that the overall closure for GNN-based truth flavor tagged distributions is better compared to the 2D map-based truth flavor tagged distributions. Further, in Figures 6.6(c)-(d), a shape-based non-closure effect is seen for the average pile-up $|\mu|$, and $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ which was also noticed in the previous $VH, H \to c\bar{c}$ analysis. However, in GNN-based truth flavor tagging such shape based non-closure effects are not seen, showing the $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$ is parametrised with sufficient accuracy for $|\mu|$ and $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ as well.



(a) $p_T(j_1)$ for $Z + cc$ jets in the 0L channel. **(b)** $\eta(j_2)$ for top t-channel in the 1L channel.



(c) $\hat{\mu}$ for $W + cc$ jets in the 1L channel. **(d)** $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ for $Z + cc$ jets in the 0L channel.

Figure 6.6: The closure between direct flavor tagging, 2D map based truth flavor tagging, and GNN based truth flavor tagging for four different θ parameters used to parametrise $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$. *0L* and *1L* are abbreviations for the *zero charged lepton* and *one charged lepton* channels.

6.3 Truth flavor tagging in the $VH, H \to b\bar{b}/c\bar{c}$ analysis

While GNN-based truth flavor tagging has a good performance as seen from the plots in Figure 6.6, it is still not perfect; some specific regions were observed to have a small degree of non-closures of up to 5%. Therefore, the analysis adopts a special strategy for truth flavor tagging as discussed in this section, and illustrated in Figure 6.7.

Truth flavor tagging is applied independently to the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ analysis regions due to the difference in tagging strategies (mainly for non-Higgs candidate jets). In the $VH, H \to b\bar{b}$ analysis regions, the choice of the $DL1r$ working point is simply the bin B, since both Higgs candidate jets need to be classified in the BB region. Truth flavor tagging in the $VH, H \to c\bar{c}$ analysis regions is more complex and was studied in depth, and is explained briefly below. Details on this study can be found in Appendix F.4.

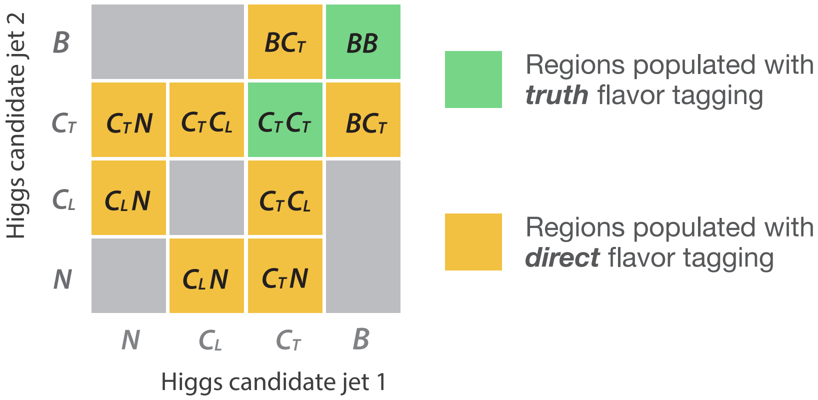


Figure 6.7: A summary of the flavor tagging selections used for MC samples. The *green* regions are populated by truth flavor tagging, and the *yellow* regions are populated with direct flavor tagging. Grey regions are also populated with direct flavor tagging, but are not used in the analysis.

The final strategy adopted for the $VH, H \to c\bar{c}$ analysis regions is to apply truth flavor tagging to the C_TC_T region only, since it is the most sensitive region to the $VH, H \to c\bar{c}$ signal; reduction of the MC statistical uncertainties are crucial in the C_TC_T region through truth flavor tagging. All other regions C_TC_L, C_TN, BC_T and C_LN requires to be populated through direct flavor tagging. However, to avoid an overlap between events from C_TC_L, C_TN, BC_T, C_LN regions entering the C_TC_T region, a percentage of events will be split for truth flavor tagging, and the remaining events are used for direct flavor tagging[4].

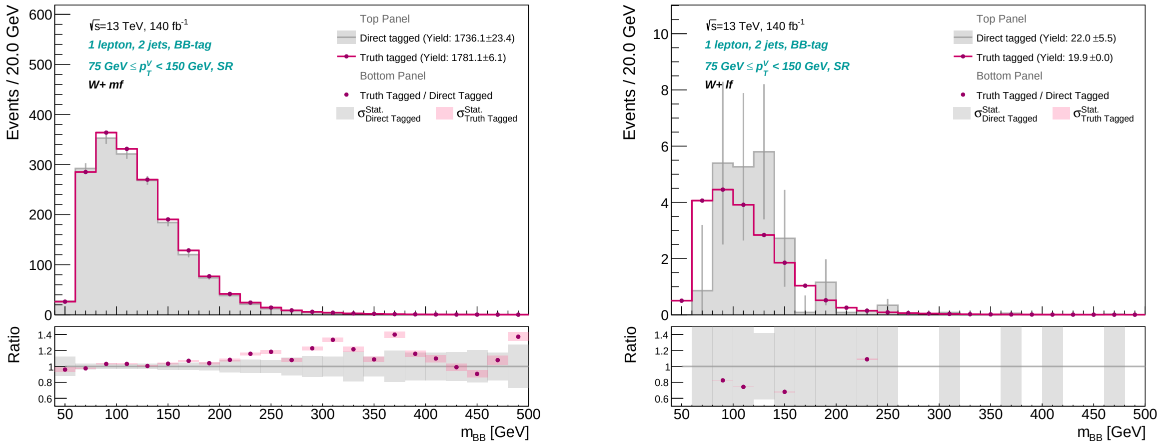
Finally, considering the available MC statistics, not all the samples required to be truth flavor tagged. Truth flavor tagging is only applied to V+jets and single top s/t channel samples, since the largest impact to MC statistical uncertainty was seen from including these two samples; a study was done using the fit framework discussed in Chapter 9, where the MC statistical uncertainty reduced by 70% when truth flavor tagging only V+jets and single top s/t channel samples, compared to 83% when truth flavor tagging all the background samples. For the VH signal samples, another reason not to use truth flavor tagging is to avoid introducing any biases to the measurement of the signal strengths $\mu_{VH,H\to b\bar{b}}$ and $\mu_{VH,H\to c\bar{c}}$, since non-closure introduced from truth flavor tagging can directly impact the signal strength measurement.

---

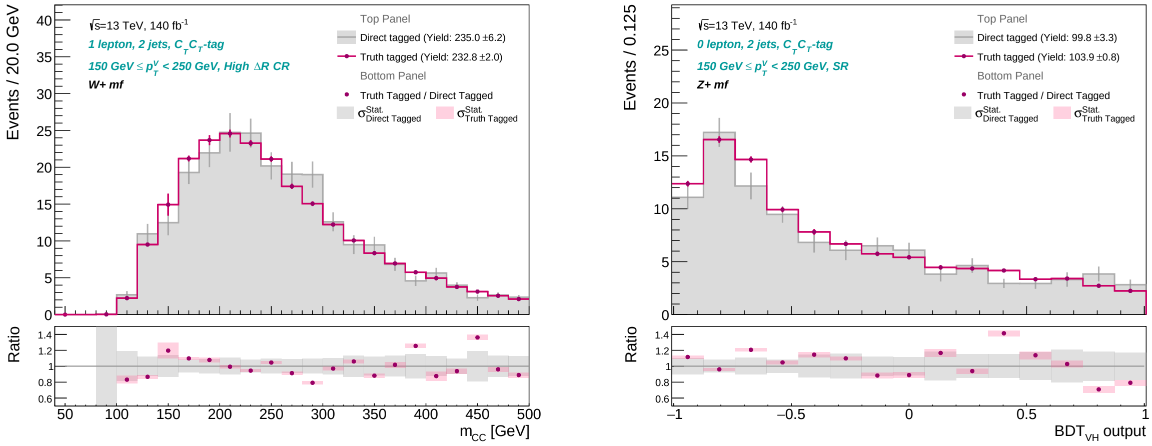
[4] Another merit of splitting the sample is that it reduces the burden on the number of times the flavor tagging selection is repeated.

6.3.1 Evaluation in the analysis categories

Truth flavor tagging is finally evaluated for the distributions in the analysis categories for BB and C_TC_T. Figures 6.8(a) and 6.8(b) show $m_{j_1j_2}$ distributions in the one charged lepton channel of the BB region. Figure 6.8(a) is another good example of how the GNN has achieved a good parametrisation of the $\epsilon_f(\mathbf{x}|\boldsymbol{\theta})$, since the truth flavor tagged distribution has a good closure with the direct flavor tagged distribution. Then, Figure 6.8(b) is a good example of how truth tagging helps in generating smooth distributions; the direct flavor tagged distribution shown has a large uncertainty due to the limited statistics and large MC event weights. Smoothing out kinematic distributions helps with the stability of the fit as well.



(a) $W + mf$ jets in the 1L BB signal region. (b) $W + lf$ jets in the 1L BB signal region.



(c) $W + mf$ jets in the 1L C_TC_T high ΔR control region. (d) $Z + mf$ jets in the 0L C_TC_T signal region.

Figure 6.8: The closure between direct flavor tagging and the GNN based truth flavor tagging in the BB and C_TC_T working points. Figures (a) to (c) shows the reconstructed mass $m_{j_1j_2}$ of the Higgs candidate jets, and (d) shows the BDT output distribution. mf, lf, *0L* and *1L* are abbreviations for *mixed flavor*, *light flavor*, *zero charged lepton* and *one charged lepton* channels.

Figures 6.8(c) and 6.8(d) shows two distributions in the C_TC_T region, where similar good closure is seen. Further, it is also worth noting that while the truth flavor tagged distributions are modelled smoother, the MC statistical uncertainties are also smaller for the truth flavor tagged distributions, which helps in reducing the impact of the MC statistical uncertainty when measuring the signal strengths. Additional closure distributions can be found in Appendix F.

Chapter 7

The multivariate analysis

The signal regions in the $VH, H \to c\bar{c}$ analysis are dominated by the background processes, and hence, methods to gain a better signal to background separation is crucial. This can be achieved through applying further selections to maximise the significance, but can lead to a reduction in the overall signal yields. On the other hand, multivariate analysis (MVA) algorithms are useful tools in separating signal events from background events, while also maximising the signal yield.

MVA approach based on a boosted decision tree (BDT) algorithm was used in the previous $VH, H \to b\bar{b}$ analysis [34] to extract the $VH, H \to b\bar{b}$ signal strength, while in the previous $VH, H \to c\bar{c}$ analysis, the $VH, H \to c\bar{c}$ signal strength was extracted from the invariant mass of the Higgs boson candidate jets $m_{j_1 j_2}$ [30]. Since the current analysis simultaneously measures the signal strengths of the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ processes, the MVA based approach is harmonized for the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ analyses, with the same set of input variables used to train the BDT algorithm. The output of the BDT algorithm will be used as the final discriminant to extract the $VH, H \to c\bar{c}$ signal strength in the signal regions.

The output of a dedicated BDT training is used for extracting the signal strengths of the $VZ, Z \to c\bar{c}$ process as well, and serves as a good validation of the MVA framework used for the $VH, H \to c\bar{c}$ process. Further, a dedicated BDT training is done on the low ΔR control region of the one charged lepton channel BB region. This is to separate the W+heavy flavor[1] events from the other backgrounds, since low ΔR control region is used to constrain the W+heavy flavor jet component. More information regarding the multivariate analysis in the $VH, H \to b\bar{b}/c\bar{c}$ analysis, especially on the $VH, H \to b\bar{b}$ analysis which is outside the focus of this thesis, can be found in Appendix G.

7.1 Boosted decision tree (BDT) algorithm

The MVA algorithm used in the analysis, to classify the signal and background events, is a boosted decision tree (BDT) classifier from the ROOT-based TMVA framework [139]. A decision tree is similar to a flowchart; the decision tree starts from a root node, and branches from the root node feed into internal nodes (also known as *decision nodes*) [140]. Each node classi-

---

[1] These are jets with the final state being bb or cc

fies events based on the input features that have discrimination power for the signal against the background. Finally the output is given by the terminal nodes (also known as *leaf nodes*).

Boosting is a method to improve the performance of a decision tree, where new decision trees are iteratively created on improving the performance of the previous decision tree, and the obtained set of decision trees (also known as a forest) are merged to get one classifier known as a boosted decision tree (BDT).

The BDT is trained on a simulated (MC) training sample set, with the input features assigned the kinematic input variables discussed in Section 7.2, to classify the events into signal or background. On each decision node, the events are classified to the subsequent nodes in the next layer based on the *Gini index*. The Gini index is a measure of the purity of a sample set, with a lower Gini index indicating a well separated signal to background [141].

7.1.1 Hyperparameters when optimizing the BDT training

Hyperparameters are external parameters that can be tuned when building a decision tree, in order to get the best performance while avoiding overtraining. *Overtraining* is where the algorithm has learned the training dataset too well, and does not perform well on a larger and more general data set. Hence, the available dataset is split into a subset used for training and a subset used for evaluating, and it is ensured the BDT performance is consistent between both subsets.

The hyperparameters described below need to be tuned case by case since the optimum configuration can be different based on the input statistics, input variables, or the use case of the decision tree.

The boosting methods for decision trees

The boost type is the boosting method used for building the BDT, and in the analysis the two methods utilized are *adaptive boosting* and *gradient boosting*.

In adaptive boosting (or AdaBoost) the training sample is re-weighted, where for example a signal event which ends up in the background node is given a larger weight than those ending in a signal node, and new decision trees are trained. The final BDT is obtained from the weighted average of the set of decision trees.

The gradient boosting (or GradBoost) method is similar to a function expansion approach where each tree corresponds to a summation. The parameters of this summation are determined one tree at a time by the minimization of an error function. In gradient boosting, the predictions from new decision trees are given a weighting factor before adding to the existing sequence of decision trees. This weighting factor, also known as the *learning rate*, will slow down the learning, and can enhance the performance of the BDT.

The BDT model related parameters

Parameters related to the decision tree model can be optimized to achieve the best performance for the intended use case. The number of trees used to build the decision tree is a parameter typically set to be in the range of 100 to 600 trees. The maximum depth is the maximum layers allowed for the decision tree before further splitting. Having too many layers can lead to overtraining since the statistics are branched into leaves on each subsequent layer. The number of steps (for each leaf) in determining the optimum cut before branching to new leaves is referred to as number of cuts, and to avoid leaves having a small sample set when branching, the minimum node size parameter is set for a leaf to contains more than a set percentage of training events. *Pruning* a decision tree reduces branches that does not bring much performance, and simplifies the BDT model. This is generally good to avoid overtraining the model (since large trees are likely to get overtrained), but in the current analysis no pruning method is adopted.

7.2 BDT input variables

Kinematic variables which exhibit a difference in shape for the signal and the background are used to train the BDT. Several candidate kinematic variables were tested, and based on the improvement on the significance, the list described below (and shown in Table 7.1) are utilized as the final set of input variables used to train the BDT. These variables are shown in Figures 7.1 to 7.10, where the signal and sub-dominant backgrounds are scaled by n times[2] to match the maximum of the dominant background distribution.

7.2.1 Input variables common to all channels

- $m_{j_1 j_2}$ is the invariant mass of the Higgs candidate system, and the $VH, H \to c\bar{c}$ signal is expected to have a peak at $m_{j_1 j_2}$ = 125 GeV as seen in Figure 7.1. Hence, in general $m_{j_1 j_2}$ is the variable which has the highest discriminating power when separating the signal from background. j_1 and j_2 are the leading and sub-leading Higgs candidate jets.
- $m_{j_1 j_2 j_3}$ is the invariant mass of the Higgs candidate system, and the additional leading jet. $m_{j_1 j_2 j_3}$ can be used to distinguish the $VH, H \to c\bar{c}$ signal from the hadronically decaying top-quark background, since the invariant mass for the top decay should get closer to the top-quark mass of 172.5 GeV. A $VH, H \to c\bar{c}$ signal event with final state radiation (FSR) would also have an $m_{j_1 j_2 j_3}$ of value closer to 125 GeV. The additional peak at 200 GeV for the $VH, H \to c\bar{c}$ signal arise for events having the additional leading jet originating from initial state radiation (ISR).

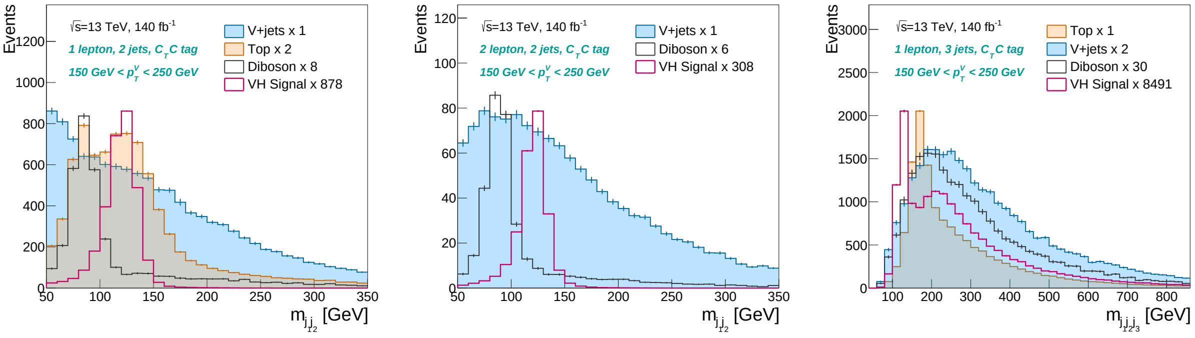


Figure 7.1: The general variables $m_{j_1 j_2}$ and $m_{j_1 j_2 j_3}$ used for training the BDT algorithm.

- $p_{\mathrm{T}}(j_1)$ and $p_{\mathrm{T}}(j_2)$ is the transverse momentum of the leading and sub-leading Higgs candidate jets. For the $VH, H \to c\bar{c}$ signal, it is likely that both jets coming from the Higgs

---

[2] An exception is for the top quark template in the two charged leptons channel which is statistically limited, and is scaled less to avoid overlap with other templates.

boson decay have a high transverse momentum (p_{T}). However, for background processes such as V+jets, if a high p_{T} leading jet is selected, it is unlikely a high p_{T} sub-leading jet is present in the event. Hence, $p_{\mathrm{T}}(j_1)$ and $p_{\mathrm{T}}(j_2)$ are effective in discriminating the $VH, H \to c\bar{c}$ signal from V+jets background.

- $\sum_{i\neq 1,2}\ p_T(j_i)$ is the sum of the transverse momentum p_{T} of the non-Higgs candidate jets that have $p_{\mathrm{T}} > 20$ GeV. Additional jets in the $VH, H \to c\bar{c}$ signal process comes from ISR, FSR, or pile-up. For the background processes, the additional jets can come from hard scattered partons or from the top quark decay (in the case of the Top quark processes), and such events tend to have a higher $\sum_{i\neq 1,2}\ p_T(j_i)$ compared to the $VH, H \to c\bar{c}$ signal.

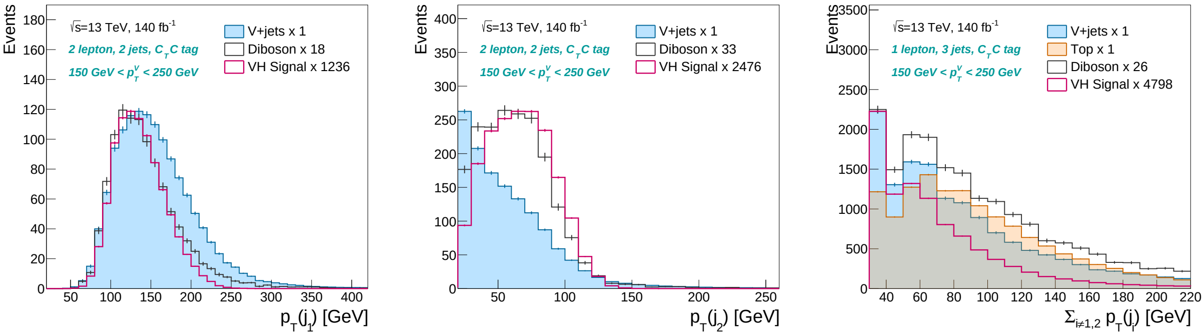


Figure 7.2: The general variables $p_{\mathrm{T}}(j_1)$, $p_{\mathrm{T}}(j_2)$, and $\sum_{i\neq 1,2}\ p_T(j_i)$ used for training the BDT algorithm.

- $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ is the $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ between the Higgs candidate jets. Since the angular separation between two jets follows the relation $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2}) \approx 2m/p_{\mathrm{T}}^{V}$, this variable is useful in discriminating the $VH, H \to c\bar{c}$ signal from the background processes as discussed in Section 5.3.2.

- $\mathrm{bin}_{\mathrm{DL1r}(j_1)}$ and $\mathrm{bin}_{\mathrm{DL1r}(j_2)}$ are the DL1r working points to which the leading and sub-leading jets fall into. The bins values are the ones shown in Table 4.2, with zero corresponding to the untagged (N) bin. This variable contains the tagging information of the jets, and hence, is useful in separating the $VH, H \to c\bar{c}$ signal from the $VH, H \to b\bar{b}$ signal, and backgrounds containing a different fractions of b-jets or light-flavor jets with respect to the $VH, H \to c\bar{c}$ signal.

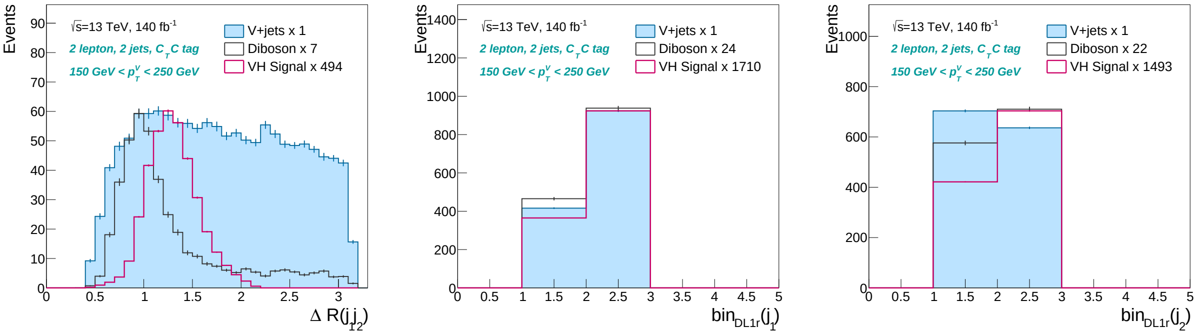


Figure 7.3: The general variables $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$, $\mathrm{bin}_{\mathrm{DL1r}(j_1)}$, and $\mathrm{bin}_{\mathrm{DL1r}(j_2)}$ used for training the BDT algorithm.

- p_{T}^{V} is the transverse momentum of the vector boson, where the $VH, H \to c\bar{c}$ signal process has relatively higher p_{T}^{V} with respect to the background. In the zero lepton channel, p_{T}^{V} is equivalent to the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. In the one lepton channel, p_{T}^{V} is the vector sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the charged lepton p_{T}. In the two lepton channel, p_{T}^{V} is the vector sum of the two charged lepton p_{T}.
- $|\Delta\phi(\boldsymbol{V}, \boldsymbol{H})|$ is the difference in azimuthal angle ϕ between the reconstructed vector boson and the Higgs candidate. For the $VH, H \to c\bar{c}$ signal process, since the Higgs boson and vector boson are created back to back in the transverse plane $x - y$, $|\Delta\phi(\boldsymbol{V}, \boldsymbol{H})|$ peaks towards 180° (or π).

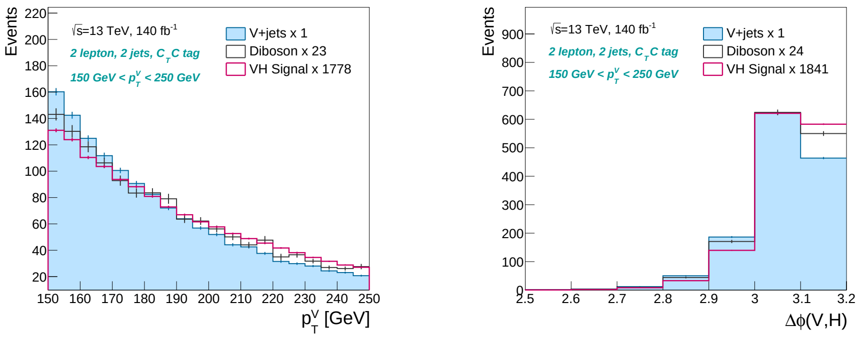


Figure 7.4: The general variables p_{T}^{V} and $|\Delta\phi(\boldsymbol{V}, \boldsymbol{H})|$ used for training the BDT algorithm.

7.2.2 Input variables specific to charged lepton channels

Input variables specific to the zero charged lepton channel

- $|\Delta\eta(\boldsymbol{j_1}, \boldsymbol{j_2})|$ is the difference in pseudo-rapidity η between the two Higgs candidate jets. This variable is effective in distinguishing the signal from the V+jets background that can have a Higgs candidate jet from ISR or FSR, and the other jet from gluon splitting. Such V+jets backgrounds are expected to have a larger $|\Delta\eta(\boldsymbol{j_1}, \boldsymbol{j_2})|$ compared to the two jets from the Higgs boson.
- m_{eff} is the scalar sum of p_{T} of all the jets, the $E_{\mathrm{T}}^{\mathrm{miss}}$ and (if present) the leading hadronic tau-jet in the event, and has more than twice the p_{T}^{V} threshold of the region under consideration. Since m_{eff} has information of the p_{T}^{V} and p_{T}^{H}, the $ZH \to \nu\bar{\nu}c\bar{c}$ signal is expected to have higher m_{eff} compared to the background.

Input variables common to the zero and one charged lepton channel

- $E_{\mathrm{T}}^{\mathrm{miss}}$ is the missing transverse energy, and is computed from the negative sum of the transverse momenta of all the hard objects; electrons, muons, hadronic τ-leptons, photons, and jets. In the zero charged lepton channel, $E_{\mathrm{T}}^{\mathrm{miss}}$ is equal to the p_{T}^{V}. In the one charged lepton channel, $E_{\mathrm{T}}^{\mathrm{miss}}$ represents the p_{T} of the neutrino.
- $\Delta R(\boldsymbol{c}, \boldsymbol{j})_{\mathrm{min}}$ is the minimum ΔR between the Higgs candidate jets and the additional jets. This variable is useful in separating the top quark processes from the $VH, H \to c\bar{c}$ process, since the additional jets in top quark processes are likely to come from the W boson decay, giving a smaller $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ between the jets. For the $VH, H \to c\bar{c}$ process, the additional

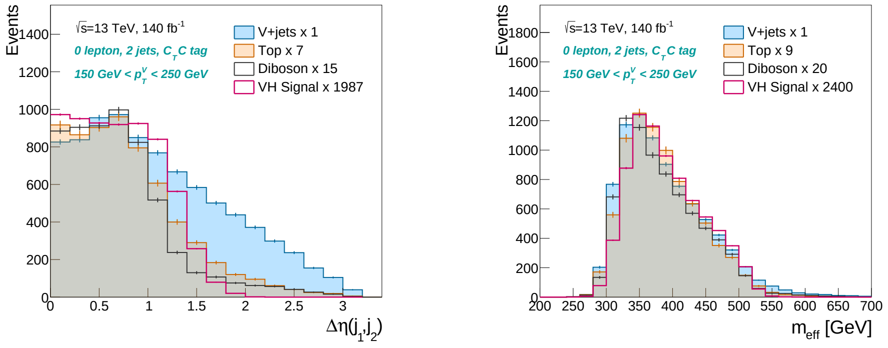


Figure 7.5: The variables $|\Delta\eta(\boldsymbol{j_1}, \boldsymbol{j_2})|$ and m_{eff} used in the zero charged lepton channel.

jets coming from ISR or pile-up are likely to have a larger $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ separation with the jets from the Higgs boson decay. The peak at low $\Delta R(\boldsymbol{c}, \boldsymbol{j})_{\text{min}}$ for the $VH, H \to c\bar{c}$ signal can be attributed to the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ between FSR jets and the c-jet which emitted the gluon radiation.

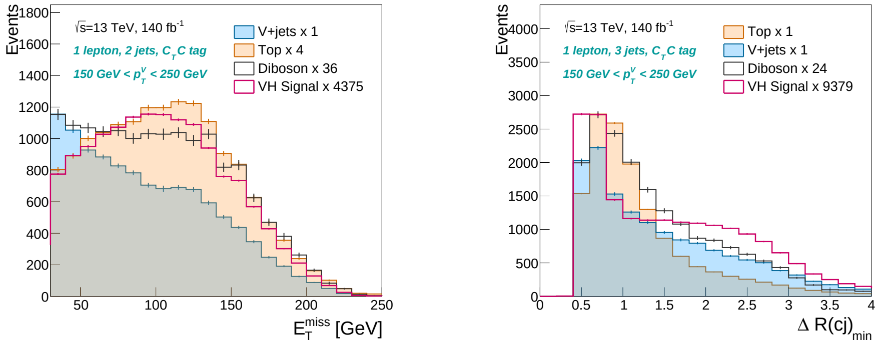


Figure 7.6: The variables $E_{\text{T}}^{\text{miss}}$ and $\Delta R(\boldsymbol{c}, \boldsymbol{j})_{\text{min}}$used in the zero and one charged lepton channels.

Input variables specific to the one charged lepton channel

- $|\Delta\phi(\boldsymbol{l}, \boldsymbol{b} \text{ or } \boldsymbol{c})|_{\text{min}}$ is the azimuthal angle ϕ between the lepton and the closest b/c-tagged jet. The $WH \to l\nu c\bar{c}$ signal has a defined angular separation between the two jets from the Higgs boson, and a defined angular separation between the neutrino and the lepton from the W boson decay. Hence, $|\Delta\phi(\boldsymbol{l}, \boldsymbol{b} \text{ or } \boldsymbol{c})|_{\text{min}}$ is expected to have a peak towards $\Delta\phi = \pi$. For W+jets, the jets from the radiated gluon does not have a well defined angular separation, giving a broader $|\Delta\phi(\boldsymbol{l}, \boldsymbol{b} \text{ or } \boldsymbol{c})|_{\text{min}}$ as seen in Figure 7.7. Hence, $|\Delta\phi(\boldsymbol{l}, \boldsymbol{b} \text{ or } \boldsymbol{c})|_{\text{min}}$ is effective in separating the the $WH \to l\nu c\bar{c}$ signal from the W+jets background.
- m_{T}^{W} is the transverse mass of the W boson candidate reconstructed using Equation 7.1. For processes with the lepton and the neutrino coming from the W boson decay, m_{T}^{W} peaks toward the W boson mass of 80 GeV. Hence, m_{T}^{W} is effective in separating the $WH \to l\nu c\bar{c}$ signal from QCD multijet processes that are confined to lower values of m_{T}^{W}.

$$m_{\text{T}}^{W} = \sqrt{2p_T^l E_{\text{T}}^{\text{miss}}(1 - \cos(\Delta\phi(l, \boldsymbol{E}_{\text{T}}^{\text{miss}})))} \tag{7.1}$$

- m_{top} is the reconstructed mass of the leptonically decaying top quark. For the reconstruction of the top quark mass, first it is necessary to reconstruct the mass of the W boson. This

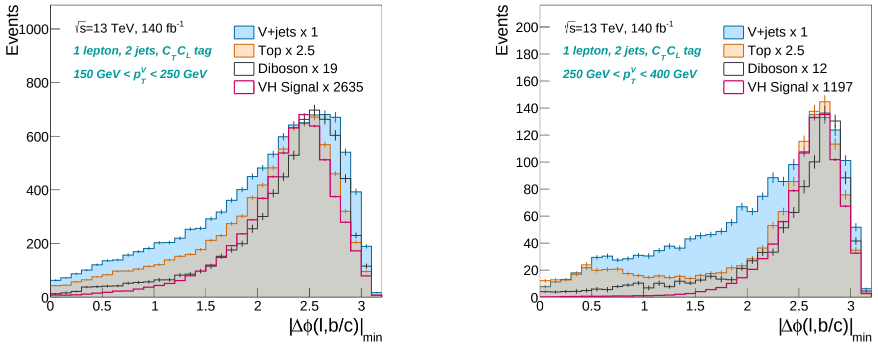


Figure 7.7: The variable $|\Delta\phi(\boldsymbol{l}, \boldsymbol{b} \text{ or } \boldsymbol{c})|_{\text{min}}$used for in the one charged lepton channel, in two different p_{T}^{V} regions.

requires the determination of the longitudinal component of the neutrino momentum p_z^{ν}. This is determined by solving the quadratic Equation 7.2, with m_W as a constraint.

$$p_z^{\nu} = \frac{1}{2(p_{\text{T}}^{l})^2}\left[p_z^l X \pm E_l\sqrt{X^2 - 4(p_{\text{T}}^{l})^2(E_{\text{T}}^{\text{miss}})^2}\right] \tag{7.2}$$

where $X = m_W^2 + 2(p_x^l E_x^{\text{miss}} + p_y^l E_y^{\text{miss}})$. Next, m_{top} is reconstructed using the Higgs candidate jet, and the solution for p_z^{ν} that minimizes the m_{top}. If $X^2 < 4(p_{\text{T}}^{l})^2(E_{\text{T}}^{\text{miss}})^2$, then the magnitude of $E_{\text{T}}^{\text{miss}}$ is corrected such that $X^2 = 4(p_{\text{T}}^{l})^2(E_{\text{T}}^{\text{miss}})^2$. The reconstructed m_{top} from the W boson and b-jet from the top quark has a peak around the top quark mass of 172.5 GeV [3]. Since the W boson and Higgs candidate jet in the $WH \to l\nu c\bar{c}$ signal are almost back to back, the reconstructed m_{top} is expected to have a higher peak (away from 172.5 GeV) owing to the larger angular separation.

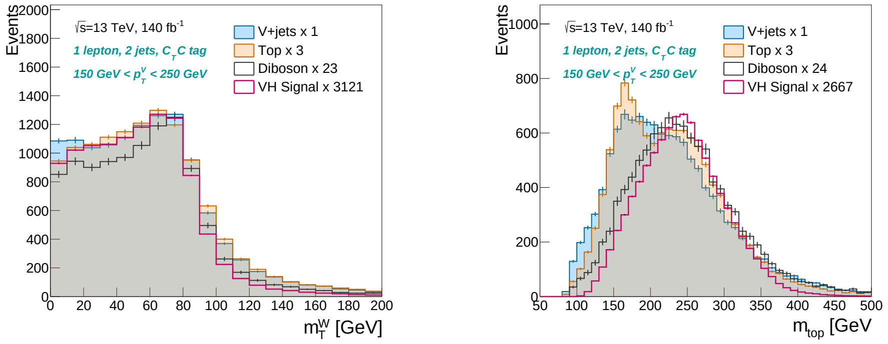


Figure 7.8: The variables m_{T}^{W} and m_{top} used in the one charged lepton channel.

Input variables common to the one and two charged lepton channel

- $|\Delta y(\boldsymbol{V}, \boldsymbol{H}_{\textbf{cand}})|$ is the the difference in rapidity between the reconstructed vector boson and the Higgs candidate. For the $VH, H \to c\bar{c}$ signal, the Higgs boson and the vector boson tend to have similar rapidities, since they are produced mostly back to back in the hard scattered process. Hence, the $VH, H \to c\bar{c}$ signal is expected to have a smaller difference in rapidity $|\Delta y(\boldsymbol{V}, \boldsymbol{H}_{\textbf{cand}})|$, and peaks towards zero.

---

[3] However, the reconstructed m_{top} using the W boson from one top quark, and a jet from the other top quark is expected to have a higher value than 172.5 GeV, similar to the $WH \to l\nu c\bar{c}$ signal.

Input variables specific to the two charged lepton channel

- m_{ll} is the invariant mass of the two lepton system. m_{ll} is useful when separating the top quark di-lepton background from $Z \to ll$ processes (such as the $ZH \to l^- l^+ c\bar{c}$ signal) that create a peak at the Z boson mass of 91 GeV.

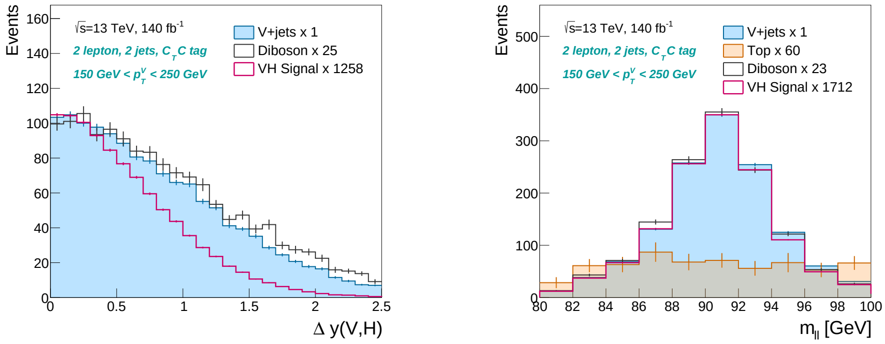


Figure 7.9: The variables $|\Delta y(\boldsymbol{V}, \boldsymbol{H}_{\mathbf{cand}})|$ used in the one and two charged lepton channels, and m_{ll} used for training the BDT in the two charged leptons channel.

- $E_T^{\text{miss}}/\sqrt{S_{\text{T}}}$ is the quasi-significance of the $E_{\text{T}}^{\text{miss}}$, where S_{T} is the scalar sum of p_{T} of the leptons and jets in the event. Top quark processes in the two charged lepton channel is expected to have a relatively larger $E_{\text{T}}^{\text{miss}}$, compared to other processes where no $E_{\text{T}}^{\text{miss}}$ is expected. Hence, this variable is effective in discriminating the $ZH \to l^- l^+ c\bar{c}$ signal from the top quark background.

- $\cos\theta(\boldsymbol{l}^-, \boldsymbol{Z})$ is the Z boson polarization sensitive angle, and is the cosine of the angular distance between the negatively charged lepton direction in the Z boson rest frame, and the flight path of the Z boson in the lab frame. The Z boson is almost 100% longitudinally polarized for the $ZH \to l^- l^+ c\bar{c}$ signal process, and the Z boson is transversely polarized for the Z+jets and di-boson processes. As a consequence, the leptons in the $ZH \to l^- l^+ c\bar{c}$ signal process are emitted mostly transverse to the Z boson direction, while the leptons in the $Z \to l^+ l^-$ decay of Z+jets and di-boson processes are emitted mostly along the direction of the Z boson [142, 143]. Hence, $\cos\theta(\boldsymbol{l}^-, \boldsymbol{Z})$ can be used to discriminate the signal with the Z+jets and di-boson background.

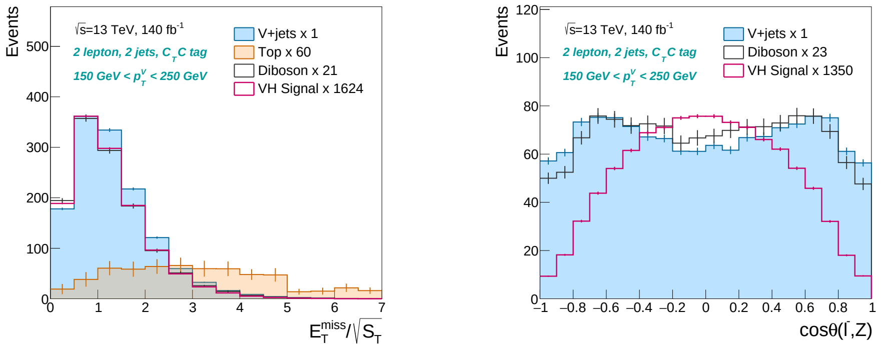


Figure 7.10: The variables $E_T^{\text{miss}}/\sqrt{S_{\text{T}}}$ and $\cos\theta(\boldsymbol{l}^-, \boldsymbol{Z})$ used in the two charged leptons channel.

The BDT variables described above were validated to confirm there is a good agreement between the data and MC, and some distributions (after performing the fit) can be found in Section 10.1.

Table 7.1: The input variables used to train the resolved $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ MVA algorithms. The check mark represents the respective variables used in the zero lepton, one lepton and two leptons channels. V stands for the vector boson, H stands for the Higgs boson candidate, and l stands for the lepton.

Variable	**Description**	**Lep. channel**		
		0L	**1L**	**2L**
$m_{j_1 j_2}$	Invariant mass of Higgs boson candidate	✓	✓	✓
$m_{j_1 j_2 j_3}$	Invariant mass of H, and the additional leading jet	✓	✓	✓
$p_{\mathrm{T}}(j_1)$	Transverse momentum (p_{T}) of the leading jet	✓	✓	✓
$p_{\mathrm{T}}(j_2)$	Transverse momentum (p_{T}) of the sub-leading jet	✓	✓	✓
$\sum_{i \neq 1,2} p_T(j_i)$	Scalar sum of the p_{T} of non-H jets	✓	✓	✓
$\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$	ΔR between the H jets	✓	✓	✓
$\mathrm{bin}_{\mathrm{DL1r}(j_1)}$	The DL1r bin of the leading jet	✓	✓	✓
$\mathrm{bin}_{\mathrm{DL1r}(j_2)}$	The DL1r bin of the sub-leading jet	✓	✓	✓
p_{T}^{V}	p_{T} of the vector boson (V)	✓	✓	✓
$\lvert\Delta\phi(\boldsymbol{V}, \boldsymbol{H})\rvert$	Difference in ϕ angle between V and H	✓	✓	✓
$\lvert\Delta\eta(\boldsymbol{j_1}, \boldsymbol{j_2})\rvert$	$\Delta\eta$ between the H jets	✓		
m_{eff}	Scalar sum of the p_{T} of all small-R jets and $E_{\mathrm{T}}^{\mathrm{miss}}$	✓		
$E_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse energy	✓	✓	
$\Delta R(\boldsymbol{c}, \boldsymbol{j})_{\mathrm{min}}$	ΔR between the closest H jet and the additional jet	✓	✓	
$\lvert\Delta\phi(\boldsymbol{l}, \boldsymbol{b} \text{ or } \boldsymbol{c})\rvert_{\mathrm{min}}$	ϕ angle between the l and the closest b/c tagged jet		✓	
m_{T}^{W}	Transverse mass of the W boson candidate		✓	
m_{top}	Reconstructed mass of the leptonically decaying top quark		✓	
$\lvert\Delta y(\boldsymbol{V}, \boldsymbol{H}_{\mathbf{cand}})\rvert$	Difference in rapidity between V and H		✓	✓
m_{ll}	Invariant mass of the di-lepton system			✓
$E_T^{\mathrm{miss}}/\sqrt{S_{\mathrm{T}}}$	Quasi-significance of the $E_{\mathrm{T}}^{\mathrm{miss}}$			✓
$\cos\theta(\boldsymbol{l}^{-}, \boldsymbol{Z})$	Z boson polarization sensitive angle			✓

7.3 Training the BDT

Maximizing the training statistics when training a BDT algorithm is essential to avoid overtraining the model. Hence, the truth flavor tagging discussed in Chapter 6 is utilized to generate more statistics to train the BDT. To further increase the statistics needed for training, the C_TC_T and C_TC_L tag regions are merged, the low and high ΔR control regions are merged to the signal region, the n-jet regions with $\geq$ 3 jets for the two lepton channel are merged, and the regions having $p_{\mathrm{T}}^{V} > 150$ GeV are merged. Hence, the BDT training regions in the $VH, H \to c\bar{c}$ analysis shown in Figure 7.11 are different from the categories shown in Figure 5.4. However, during inference of the BDT, the original categorization in Figure 5.4 is utilized.

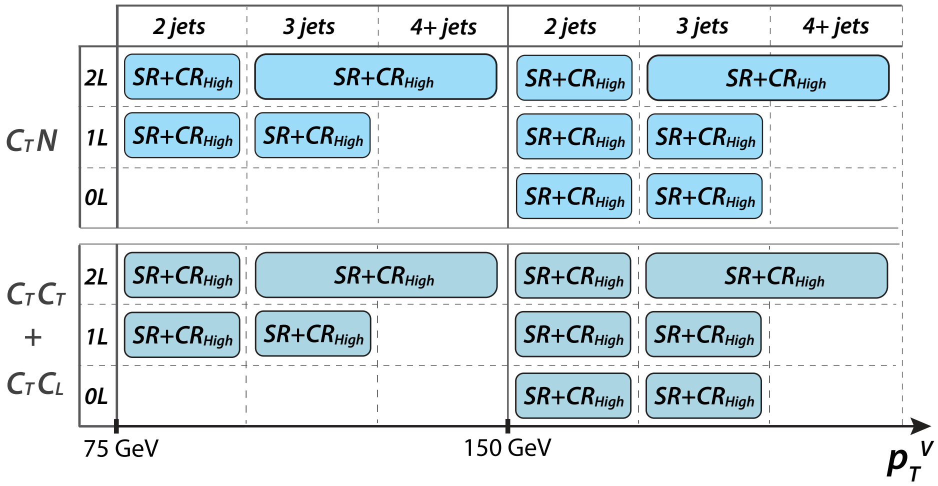


Figure 7.11: The regions used for training the BDT in the $VH, H \to c\bar{c}$ analysis, where dedicated trainings are done for each of the regions shown. 0L, 1L, 2L are abbreviations for the zero, one, and two charged lepton channels.

The BDT is trained with the $VH, H \to c\bar{c}$ process set as the signal, and all other backgrounds plus the $VH, H \to b\bar{b}$ process set as the background. However, the assignment of signal and background is different for the two other dedicated BDT trainings on the $VZ, Z \to c\bar{c}$ process and the low ΔR control region, as explained in the paragraphs below. The hyperparameters for training the BDT, given in Table 7.2 are optimized to maximise the significance in the different BDT training regions[4] and the three different BDT trainings; the $VH, H \to c\bar{c}$ process, $VZ, Z \to c\bar{c}$ process, and the low ΔR control region.

Gradient boosting method is utilized for most of the regions for training the BDT for the $VH, H \to c\bar{c}$ process, since this method is good at capturing effects outside the bulk of the distribution. Overall, the gradient boosting method gives a better signal to background separation over the adaptive boosting method. However, this method requires the region outside the bulk (i.e the tails) to be well populated, or it leads to overtraining. This is the case for $VZ, Z \to c\bar{c}$ process in all regions, and $VH, H \to c\bar{c}$ process in the $\geq$ 3 jet, 75 GeV $\leq p_{\mathrm{T}}^{V} \leq$ 150 GeV region, two charged lepton channel. Hence, the adaptive boosting method with a different *number of trees* and a *learning rate* is used in those specific BDT trainings.

When training the BDT, k-fold method with $k = 2$ is used for splitting the MC sample set into

---

[4] Each region shown in Figure 7.11 have dedicated trainings

Table 7.2: The hyperparameters used when training the BDT for the $VH, H \to c\bar{c}$ regime.

Parameter	**$VH, H \to c\bar{c}$ as signal**		**$VZ, Z \to c\bar{c}$ as signal**
	0L, 1L and some 2L regions	**2L, $\geq$ 3-jets, low p_{T}^{V}**	**0L, 1L and 2L channels**
Boost type	Gradient boost	AdaBoost	AdaBoost
Number of trees	600	200	200
Maximum depth	4	4	4
Learning rate (β)	0.5	0.15	0.15
Number of cuts	100	100	100
Minimum node size	5%	5%	5%
Separation method	Gini index	Gini index	Gini index
Pruning method	None	None	None

sub-trainings. This ensures that the all the MC samples are utilized when training the BDT, while avoid an overlap in the events used for training and evaluating the BDT.

7.3.1 BDT model for the di-boson processes

The di-boson processes $VZ, Z \to b\bar{b}$ and $VZ, Z \to c\bar{c}$ are used as a validation of the $VH, H \to b\bar{b}/c\bar{c}$ analysis framework, which includes the MVA framework. For the $VZ, Z \to c\bar{c}$ process, a dedicated BDT model is trained to classify $VZ, Z \to c\bar{c}$ events as signal, and all other processes (including the $VH, H \to c\bar{c}$, $VZ, Z \to b\bar{b}$ and $VH, H \to b\bar{b}$ processes) as background. In the final fit framework used to extract the $VZ, Z \to c\bar{c}$ signal strength, the output from the dedicated BDT model will be used as the final discriminant in the signal regions (similar to the $VH, H \to c\bar{c}$ analysis). Note that the same input variables used in training the BDT for the $VH, H \to c\bar{c}$ process are used for training the BDT for the $VZ, Z \to c\bar{c}$ process, and the same input variables are also used for the BDT trainings for $VZ, Z \to b\bar{b}$ and $VH, H \to b\bar{b}$ processes due to the harmonization.

7.3.2 BDT model for the low ΔR control region

The low ΔR control region in the BB tag region is enriched with W+heavy flavor events, and hence, this region is used to constrain the W+heavy flavor component in both the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ analyses. However, the abundance of $t\bar{t}$ events at low p_{T}^{V} degrades the purity of the W+heavy flavor events in the low ΔR control region. Therefore, through a dedicated BDT training to separate the W+heavy flavor process from the remaining background, bins with high W+heavy flavor purity can be obtained.

7.4 The MVA evaluation

The trained BDT (using the training samples) is evaluated (with the evaluation samples) to ensure no overtraining is present. Then the performance of the trained BDT is assessed based on the poisson significance of Equation 5.2, and this is used as the figure of merit to optimize the hyperparameters and the input variables discussed in the previous section. To further optimize the sensitivity in the BDT output, a special re-binning method referred to as Transformation D is applied to the default output of the BDT that originally has equidistant bins.

7.4.1 Overtraining evaluation

The evaluation samples (also known as test samples) will be classified using the trained BDT into signal and background, at a given BDT output score. If the BDT is overtrained, then the classification of the signal and background will not be the same as that of the training sample. Hence, the overtraining is assessed based on the closure of the evaluated distribution on the trained distribution. Figure 7.12 shows and example for the BDT training in the zero lepton channel, $C_TC_T + C_TC_L$ tag 2 jet, $p_{\mathrm{T}}^{V} > 150$ GeV region, where the good closure (within uncertainties) between the training and testing sub sets for the signal and background shows the BDT is not overtrained.

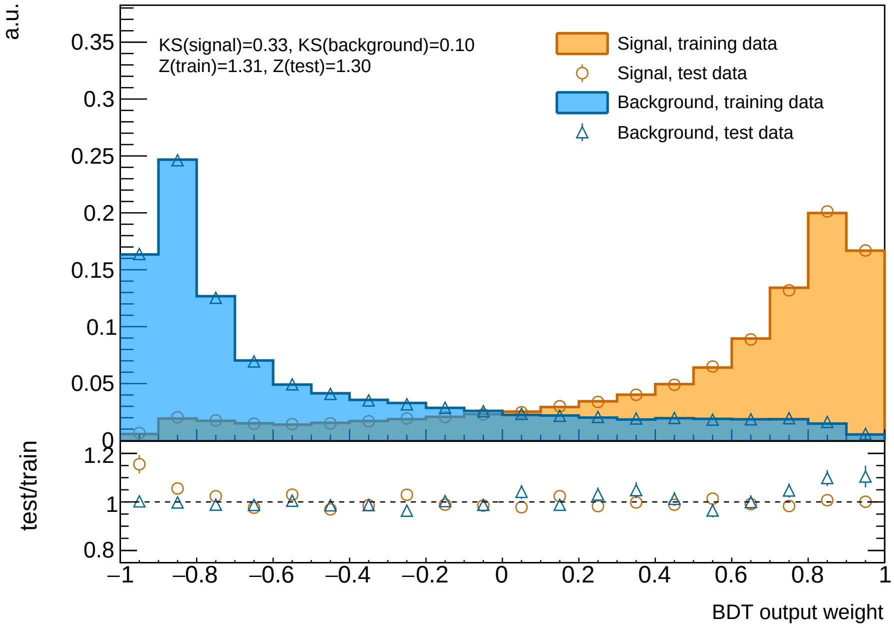
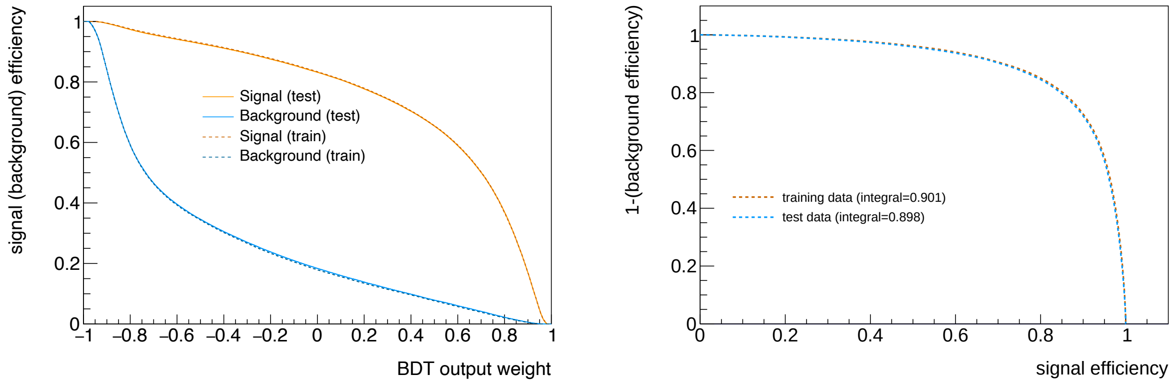


Figure 7.12: The overtraining check for the trained BDT. The plot shows the BDT output distribution for signal and background samples, that are split into the training and testing sub sets. The good closure within uncertainties is an implication that the BDT is not overtrained.

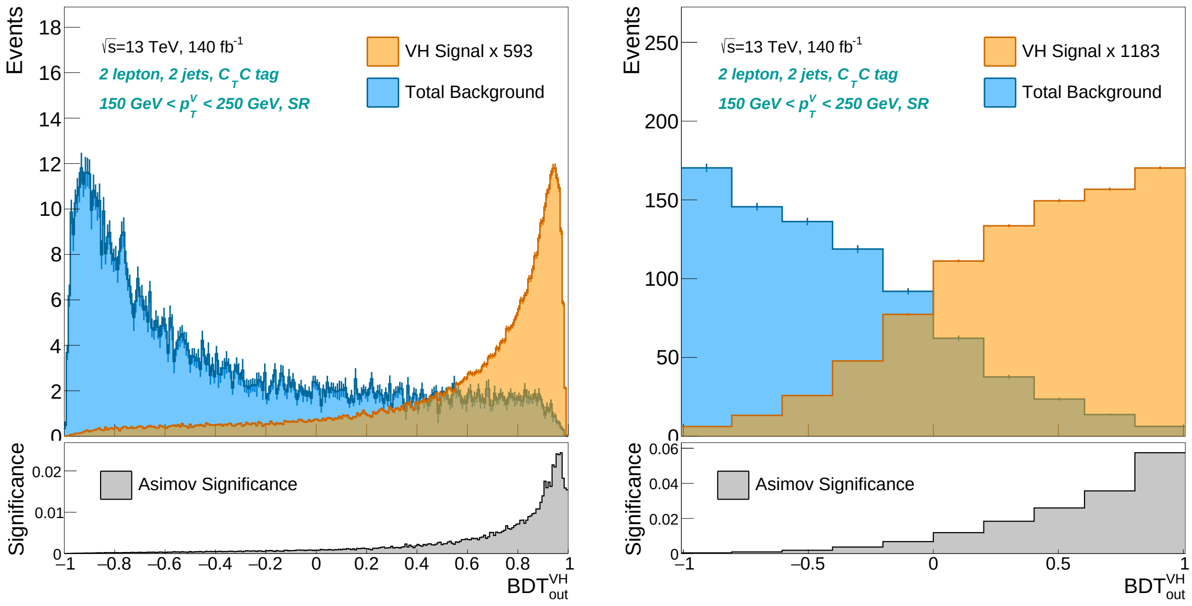
To complement the overtraining evaluation of the BDT output distributions, the signal and background efficiency curves (Figure 7.13(a)), and the ROC curves (Figure 7.13(b)) between the training and testing subsets are also assessed. The good closure for the curves between the training and testing samples indicates that the BDT is not overtrained.



(**a**) Background and signal efficiency.

(**b**) Background efficiency vs signal efficiency.

Figure 7.13: The overtraining checks for the trained BDT. The two graphs for the training sample and testing sample closing on one another indicates there is no overtraining of the BDT.



(**a**) The BDT output before Transformation D.

(**b**) The BDT output after Transformation D.

Figure 7.14: The output of the BDT, with and without Transformation D applied, for the VH signal and the total background. The top panel shows the VH signal scaled by n times to match the maximum of the total background distribution. The bottom panel shows the poisson significance calculated with $n = 1$ for the VH signal template.

7.4.2 Significance after transformation D

Prior to the calculation of the poisson significance of the BDT output distribution, a merging (or re-mapping) of bins in the distribution is done to maximise the signal and background separation[5], which is known as *Transformation D*. Through transformation D, the aim is to achieve a smoother binning for both the signal and background, while also having a fine enough binning

---

[5] The default binning of the BDT output distribution does not always provide the optimal binning since the binning is equidistant. The optimal binning, which transformation D aims to achieve, needs to take into consideration that bins with higher BDT output score have good signal-to-background ratio.

in the bins with the largest signal contribution (usually the right-most bins). During merging, the statistical uncertainty is set to be under a certain threshold for each bin. More details on the transformation D can be found in Appendix G.2.

Figure 7.14 depicts two distributions in the two charged lepton channel of the $C_TC_T + C_TC_L$ tag category, with and without the transformation D applied. The ratio panel in the below plots show the poisson significance. For comparison with the BDT output distributions, the $m_{j_1j_2}$ distribution for the same region is shown in Figure 7.15, which clearly shows how the BDT output has better signal to background separation, and hence, a better significance.

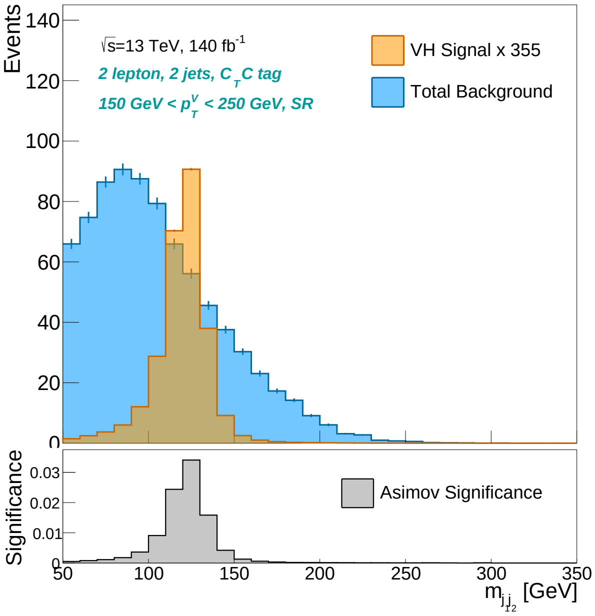


Figure 7.15: The $m_{j_1j_2}$ distribution for the VH signal and the total background. The top panel shows the VH signal scaled by n times to match the maximum of the total background distribution. The bottom panel shows the poisson significance calculated with $n = 1$ for the VH signal template.

7.4.3 The MVA performance

Since $m_{j_1j_2}$ distribution was used previous analysis to extract the $VH, H \rightarrow c\bar{c}$ signal strength, the performance of the MVA method is expressed as the relative improvement of the signal to background separation of the BDT. This improvement can be expressed as the relative increase in the poisson significance between of the BDT output distribution with respect to the $m_{j_1j_2}$ distribution. Table 7.3 shows the comparison of the BDT and $m_{j_1j_2}$ poisson significances for the training regions of Figure 7.11. Overall there is a good improvement in using the MVA method, with some regions showing improvement up to 84%.

Table 7.3: The performance of the BDT compared to the $m_{j_1 j_2}$ distribution. The **low** and **high** p_{T}^{V} regions refer to the 75 GeV $< p_{\mathrm{T}}^{V} <$ 150 GeV, and $p_{\mathrm{T}}^{V} >$ 150 GeV regions respectively.

Channel	**Regions**			**Poisson significance**		
	Tag	p_T^V	**No. of jets**	$m_{j_1 j_2}$	**BDT**	**Difference**
Zero charged lepton	$C_T N$	High	2 jets	0.047	0.073	+55%
			3 jets	0.027	0.038	+40%
	$C_T C_T + C_T C_L$	High	2 jets	0.079	0.131	+66%
			3 jets	0.040	0.067	+68%
One charged lepton	$C_T N$	Low	2 jets	0.023	0.032	+39%
			3 jets	0.012	0.015	+25%
		High	2 jets	0.038	0.061	+61%
			3 jets	0.019	0.028	+47%
	$C_T C_T + C_T C_L$	Low	2 jets	0.037	0.056	+51%
			3 jets	0.018	0.024	+33%
		High	2 jets	0.063	0.108	+71%
			3 jets	0.026	0.048	+84%
Two charged lepton	$C_T N$	Low	2 jets	0.021	0.030	+42%
			$\geq$ 3 jets	0.014	0.018	+29%
		High	2 jets	0.031	0.048	+55%
			$\geq$ 3 jets	0.021	0.029	+38%
	$C_T C_T + C_T C_L$	Low	2 jets	0.032	0.052	+63%
			$\geq$ 3 jets	0.022	0.035	+59%
		High	2 jets	0.052	0.088	+69%
			$\geq$ 3 jets	0.033	0.057	+73%

Chapter 8

Systematic uncertainties

The systematic uncertainties are mainly classified into the *experimental systematic uncertainties* and the *modeling systematic uncertainties* discussed in this chapter. Systematic uncertainties account for systematic effects which modify the signal and background distributions for the BDT output, or other kinematic variables. The modification of the overall yield in a distribution is referred to as a *normalisation effect*, while the modification of the shape of the distribution (with constant yield) is referred to as a *shape effect*.

8.1 Experimental systematic uncertainties

The experimental systematic uncertainties are uncertainties attributed to the luminosity, pile-up, trigger, objects used in the analysis (leptons, jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$), and jet-flavor tagging. These uncertainties can have both a normalisation effect and a shape effect.

8.1.1 Uncertainties on the luminosity and pile-up

The integrated luminosity is used in the scaling of the generated MC events to the expected number of events, of the different signal and background processes. Hence, the uncertainty in determining the integrated luminosity is propagated to the simulated templates of the signal and background samples.

Table 8.1: The integrated luminosity and the corresponding uncertainties for the years 2015 2018 in the Run 2 period of the LHC [55].

Year	Integrated luminosity (fb^{-1})	Uncertainty (fb^{-1})
2015	3.24	±0.04
2016	33.40	±0.30
2017	44.63	±0.50
2018	58.79	±0.64
Combined	140.07	±1.17

The absolute luminosity scale is determined through beam separation scans during dedicated LHC running periods, and then it is extrapolated to physics data taking runs based on measurements from several luminosity-sensitive detectors [55] such as the LUCID-2 detector [81]. The uncertainty on the integrated luminosity for each year is shown in Table 8.1, with the uncertainty on the total integrated luminosity of 140.07 fb^{-1} being 1.17 fb^{-1}.

The pile-up (i.e the number of interactions per bunch crossing) affects the reconstruction of physics objects such as jets, leptons, etc. Therefore, the pile-up in MC is adjusted to data by scaling the MC events using a pile-up re-weighting factor. The pile-up re-weighting has a nominal data scale factor of 1.0/1.03, and as per recommendations provided by the ATLAS experiment, an uncertainty is computed by changing the nominal data scale factor to 1.0/0.99 for the up variation, and 1.0/1.07 to get the down variation.

8.1.2 Trigger related uncertainties

As discussed in Section 3.1.1 (and detailed in Appendix C), the trigger efficiency is relatively low near the trigger thresholds, before reaching a plateau at high $E_{\mathrm{T}}^{\mathrm{miss}}$ or p_{T}. Therefore, most trigger related uncertainty terms are attributed to scale factors applied to MC events, to correct the trigger efficiency with respect to data. Altogether, there are seven trigger related uncertainty terms as explained below; three for single lepton triggers, and four for $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers.

Lepton triggers

The electron trigger systematics and muon trigger systematics can be attributed to the statistical and systematic uncertainty on the trigger scale factors. For electrons, the total of both the statistical and systematic uncertainties are considered as one systematic uncertainty, while for the muon trigger, the two systematics are considered separately.

$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger

The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors used in the analysis are derived using the $t\bar{t}$ events triggered by a single muon. To account for biases when deriving the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors, two effects are considered as systematic uncertainties for the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger. First effect is the statistical error on the dataset used to derive the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors, which is used as a systematic uncertainty on the scale factor. The second effect is the systematic uncertainty on the MC samples used, which is assessed by using alternate MC samples to derive the scale factors. Alternate scale factors are derived using $W \to \mu\nu+$ jets and $Z \to \mu^{+}\mu^{-}+$ jets events, and the difference with respect to the scale factor derived using $t\bar{t}$ events is defined as a systematic uncertainty.

The efficiency of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger used in 2017 has a kinematic dependence on the offline scalar sum of the transverse momentum of all jets in the event, $S_{\mathrm{T}} = \sum_{i=1}^{N} p_{\mathrm{T}}(j_i)$. Hence, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger scale factors as a function of S_{T} is used as an additional systematic uncertainty.

8.1.3 Uncertainties on lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction

Electrons

Systematic uncertainties for electrons takes into account the impact on the electron energy measurement depending on the transverse energy E_{T} and the pseudo-rapidity η. Table 8.2 summarizes the different types of systematic uncertainties related to the overall electron reconstruction and energy calibration.

Table 8.2: The experimental systematic uncertainties related to electron reconstruction, identification, etc.

Type	Uncertainty on
Electrons	Reconstruction efficiency Identification efficiency Isolation efficiency Energy scale Energy resolution

The systematic uncertainty for electron identification efficiency is around ±2% at $E_{\mathrm{T}} = 5$ GeV, and decreases with transverse energy, reaching lesser than ±1% for 30 GeV$< E_T <$250 GeV [118]. The complete systematic uncertainty model for calibrating the electron energy scale considers 64 systematic variations, and the achieved calibration uncertainties for the energy resolution (scale) are typically 0.05% (between 0.2% to 1%) for electrons from Z-boson decays [118]. However, the $VH, H \rightarrow b\bar{b}/c\bar{c}$ analysis is not sensitive to the electron energy scale and resolution, and hence, the uncertainty in the electron energy scale and resolution are not dominant systematic sources. Therefore, one systematic uncertainty for the energy scale, and one systematic uncertainty for the energy resolution, which wraps the total of the systematic variations is considered.

Muons

Table 8.3 summarizes the seven different types of systematic uncertainties related to the overall muon reconstruction and momentum calibration.

Table 8.3: The experimental systematic uncertainties related to muon reconstruction, identification, etc.

Type	Uncertainty on
Muons	Reco. and ID efficiency Isolation efficiency Track-to-vertex association efficiency Charge dependant local mis-alignment effects Combined momentum resolution Combined momentum scale

The muon reconstruction (reco.) and identification (ID) efficiencies are determined from $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events. For medium ID muons in the low muon p_T region, the combined statistical and systematic uncertainty is around 0.2% to 1%, and reduces to less than 0.2% in the high muon p_T region [120].

$Z \to \mu\mu$ are also used for determining the isolation and track-to-vertex association efficiencies. The uncertainty in the isolation efficiencies also differ between the working points; for example, in the *tight* working point, the combined statistical and systematic uncertainty on muon isolation is around 4% at low muon p_T, but reduces to a negligible level at high muon p_T.

$J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events are used to measure the resolution and scale of the muon momentum. Since momentum resolution and momentum scale corrections are derived for different pseudo-rapidity bins, the uncertainties are different in pseudo-rapidity regions as detailed in [121].

Hadronically decaying τ-leptons

The reconstruction efficiency and identification efficiency measurement, the τ-lepton energy scale calibration, and the related uncertainties are obtained using the $Z \to \tau\tau$ events. The four systematic uncertainties listed in Table 8.4 is attributed to the hadronically decaying τ-leptons.

Table 8.4: The experimental systematic uncertainties related to τ reconstruction, identification, etc.

Type	Uncertainty on
τ-leptons	Reconstruction efficiency RNN ID efficiency τ-lepton energy scale correction Electron veto efficiency

Missing Transverse Energy

The uncertainties affecting the scale and resolution of the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ are considered in the analysis, and are determined using $Z \to ee$ events as detailed in [129]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties are dominated by the MC modelling uncertainty on the soft term, and therefore, only the resolution and scale uncertainties affecting the soft term are considered in the analysis, as listed in Table 8.5.

Table 8.5: The experimental systematic uncertainties related to missing traverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$.

Type	Uncertainty on
$E_{\mathrm{T}}^{\mathrm{miss}}$	Longitudinal resolution of *Soft Term* Transverse resolution of *Soft Term* Scale of *Soft Term*

8.1.4 Uncertainties related to jet reconstruction

Since small radius jets are used to reconstruct the Higgs boson candidate (as described in Section 4.9.2), the total uncertainty on jet reconstruction and calibration is the dominant experimental systematic uncertainty affecting the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ signal strength measurement. The different systematic uncertainties attributed to the overall reconstruction of jets is summarized in Table 8.6, and can be attributed to mainly the uncertainty in the jet energy scale (JES), jet energy resolution (JER), and the jet vertex tagger (JVT) [72, 123, 124].

Table 8.6: The experimental systematic uncertainties related to jet reconstruction.

Type	Uncertainty on
Small Radius Jets	Jet Energy Scale for b-jets Jet Energy Scale insitu-calibrations Flavor composition on Jet Energy Scale Jet Energy Scale due to pile-up effects Jet Energy Scale due to punch-through effects Jet Energy Scale due to behavior of high-p_T jets
	Jet Energy Resolution
	The efficiency of Jet Vertex Tagger (JVT) The efficiency of the forward Jet Vertex Tagger (fJVT)

8.1.5 Uncertainties on jet-flavor tagging

Jet-flavor tagging uncertainties account for the second dominant experimental systematic uncertainties affecting the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ signal strength measurement, and are attributed to the efficiency scale factor calibration for b-jets, c-jets, light-flavor jets, and τ-jets [125, 126, 127]. The uncertainties originate from statistical uncertainties on the data and simulated samples, the uncertainty on modeling of the physics processes used to calibrate the scale factors, and the systematic uncertainties related to the detector calibration.

Table 8.7: The experimental systematic uncertainties related to jet-flavor tagging.

Type	Uncertainty on
Jet Flavor Tagging	Flavor tagging efficiency for b-jets (ϵ_b) Flavor tagging efficiency for c-jets (ϵ_c) Flavor tagging efficiency for light-flavor jets (ϵ_{light}) Flavor tagging efficiency for τ-jets (ϵ_τ)

8.2 Modeling systematic uncertainties

The modeling systematic uncertainties are uncertainties attributed to the MC based simulation of the signal and background processes used in the analysis, and are mainly categorized into three types; normalisation, shape-based, and MC statistical uncertainties. When estimating modeling uncertainties related to the parton shower (PS) and the matrix element (ME), a set of alternative generators described in Table 8.8 are used.

Table 8.8: Summary of the nominal and alternate MC generators used in the $VH, H \to b\bar{b}/c\bar{c}$ analysis, and the types of systematic effects that are calculated from the alternate generators. Had. is an abbreviation for hadronisation in the below table.

Process	Nominal Generator	Alternate Generators	Systematic Effects
$VH, H \to b\bar{b}/c\bar{c}$	Powheg + Pythia 8	Powheg + Herwig 7	PS, UE, Had., μ_R, μ_F, PDF
$t\bar{t}$ & single-top	Powheg + Pythia 8	Powheg + Herwig 7, MadGraph + Pythia 8	PS, UE, Had., ME, μ_R, μ_F, DS/DR
V+jets	Sherpa 2.2.11	MadGraph + Pythia 8, Sherpa 2.2.1	PS, ME, μ_R, μ_F, PDF, EW Correction
Di-boson	Sherpa 2.2.11	Powheg + Pythia 8, Sherpa 2.2.1	PS, ME, μ_R, μ_F, PDF, EW Correction

8.2.1 Types of modelling uncertainties

Normalisation modeling uncertainties are attributed to the uncertainty in the cross-section, branching fraction, or the acceptance between the analysis regions, and are categorized into overall normalisation uncertainties and acceptance uncertainties. *Overall normalisation uncertainties* are used to define a priori on the normalisation of signal and background templates, in order to constrain the overall yield of the process in the analysis regions. Then *acceptance uncertainties* account for the difference in the signal and background templates in the different analysis categories or signal/background processes.

Shape-based modeling uncertainties affect the shape of the BDT, $m_{j_1 j_2}$ and p_T^V templates used in the fit, directly varying the bin contents of the templates. Finally, *MC statistical uncertainties* account for the limited sizes of the simulated MC samples; as discussed in Chapter 6, truth flavor tagging is beneficial in reducing the impact of MC statistical uncertainties.

Modeling uncertainties affect the signal sensitivity, and hence, are parameterized in the fit model as nuisance parameters (more details in Chapter 9).

Overall normalisation uncertainties

The overall normalisation uncertainties account for the normalization effects from the uncertainty on cross-sections and branching fractions, and is calculated on the overall yield of a background process, affecting (in general) all the jet-multiplicity, p_T^V, and flavor-tag categories. In the $VH, H \to b\bar{b}/c\bar{c}$ analysis, overall normalisation uncertainties constrain the yield of di-boson and single-top s/t-channel processes, while V+jets, $t\bar{t}$, and single-tip Wt processes are allowed to float freely in the fit, since dedicated control regions are used to constrain the normalisation.

Acceptance uncertainties

Acceptance uncertainties account for the possible relative differences in the number of expected events in the analysis categories, due to the choice of Monte-Carlo generators and parameters. Therefore, acceptance uncertainties are derived from the comparison of the nominal MC samples with the alternate MC samples, generated with alternate generators or parameters. Since the normalisation uncertainties are used to change the overall yield of the distributions, only relative acceptance uncertainties between regions are considered. Hence, acceptance uncertainties are computed using the double ratio formula [1]

$$\text{Acceptance uncertainty}_{var_i} = \frac{(N_B/N_A)_{var_i} - (N_B/N_A)_{nom}}{(N_B/N_A)_{nom}}, \tag{8.1}$$

where N_A (N_B) are the number of events after selections in the region A (B). Region A is referred to as the *anchor* region, which has the highest purity for the signal/background process considered, and region B is the region under *test*, to which the calculated acceptance uncertainty is applied. For example, for top quark *bq* process the BC_T control region is the anchor region, while the signal region is the region under test, to which the top quark *bq* yield is extrapolated to. The total acceptance uncertainty of all the variations considered is calculated as

$$\text{Total acceptance} = \sqrt{\sum_i (\text{Acceptance uncertainty}_{var_i})^2}, \tag{8.2}$$

where i accounts for each variation. The variations here include the difference of the nominal generator to the alternate generator, but also variations of the parameters of the nominal generator. An example is for estimation of initial state radiation (ISR) or final state radiation (FSR) that will affect the jet multiplicity categorization. For the calculation of the acceptance uncertainties for ISR and FSR effects, variations in the renormalization scale factor (μ_R) and factorization scale factor (μ_F) with respect to the nominal value $[\mu_R/\mu_R^{nom}, \mu_F/\mu_F^{nom}] = [1, 1]$ are considered. This results in varying μ_R or μ_F from half the nominal value to twice the nominal value (i.e $0.5 \to 1 \to 2$), resulting in a six point variation as shown in Equation 8.3. The acceptance uncertainty for such variations will be the maximum shift with respect to the nominal value.

$$\left[\frac{\mu_R}{\mu_R^{nom}}, \frac{\mu_F}{\mu_F^{nom}}\right] = [0.5, 0.5], [0.5, 1], [1, 0.5], [1, 2], [2, 1], [2, 2], \tag{8.3}$$

If the statistical uncertainty on the calculated acceptance uncertainty is found to be larger than the acceptance uncertainty itself, a conservative approach is taken to use the statistical uncertainty in place of the acceptance uncertainty. The uncertainty in relative acceptance described below are derived for the regions in Figure 5.4, and the different background processes.

---

[1] Using the double ratio formula prevents the double counting with the normalisation uncertainties affecting the total yield.

- Relative acceptance between the different p_{T}^{V} regions.
- Relative acceptance between the signal and control regions.
- Relative acceptance between the jet multiplicity regions.
- Relative acceptance between the charged lepton channels.
- Relative acceptance between the jet-flavor compositions.
- Relative acceptance between single-top Wt process, and $t\bar{t}$ process, since these two processes are merged as one template in the final fit.

Shape-based uncertainties

Shape-based modeling uncertainties are used to quantify a systematic variation of the shape of final templates used in the fit, like the BDT output, $m_{j_1 j_2}$, and the p_{T}^{V} distributions. These uncertainties are estimated through an interpolation between the nominal distribution and an alternate distribution created by alternate methods; either through an alternate MC generator listed in Table 8.8, or through a variation of nominal parameters in the nominal generators. However, samples made through alternate methods usually lack in statistics needed for an accurate estimation of the shape-based uncertainties. To overcome this, the nominal distribution is re-weighted to match the shape of the alternate distribution, estimated using the Calibrated Likelihood Ratio Estimator (CARL) [144].

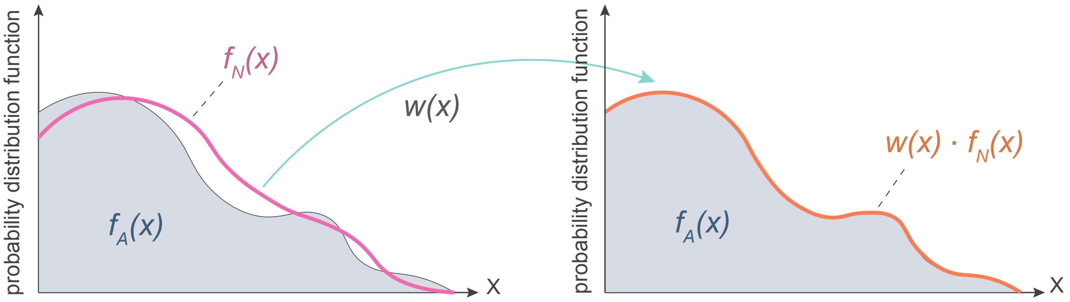


Figure 8.1: An illustration of the re-weighting of the nominal MC distribution to the alternate MC distribution using the output from the CARL neural network.

CARL is a re-weighting technique based on a deep neural network trained to discriminate between the nominal and alternate events, using the same kinematic variables listed in Table 7.1 used to train the MVA. For nominal (alternate) generators, events are sampled from a probability distribution function $f_N(x)$ $(f_A(x))$, which can be written in the machine learning nomenclature as $f(\mathrm{x}|y=0)$ $(f(\mathrm{x}|y=1))$. The output of the neural network $D(x)$, in the limit of infinite training events, approaches the probability that the event is in the alternate distribution $D(x) \to \mathcal{P}(y=1|\mathrm{x})$. This probability can be used to define a weight $w(x)$ such that

$$w(x) = \frac{f(\mathrm{x}|y=1)\mathrm{dx}}{f(\mathrm{x}|y=0)\mathrm{dx}} = \frac{\mathcal{P}(\mathrm{x}|y=1)}{\mathcal{P}(\mathrm{x}|y=0)} \xrightarrow[\text{Theorem}]{\text{Bayes}} \frac{\mathcal{P}(y=1|\mathrm{x})}{\mathcal{P}(y=0|\mathrm{x})} = \frac{\mathcal{P}(y=1|\mathrm{x})}{1-\mathcal{P}(y=1|\mathrm{x})} \approx \frac{D(x)}{1-D(x)}. \quad (8.4)$$

The weight $w(x)$ is then used to re-weight the nominal distribution, such that we obtain a similar alternate distribution as illustrated in Figure 8.1, but much larger statistics than can be used to estimate the shape-based modelling uncertainties.

8.2.2 Modeling uncertainties for the VH signal process

The uncertainties with respect to the signal can be classified into five categories.

- Uncertainties on the QCD scale.
- Uncertainty on the $H \to b\bar{b}$ and $H \to c\bar{c}$ branching fractions (Δ_{BR}), described in Table 1.2.
- Uncertainties on the parton distribution function and α_s.
- Uncertainties on parton shower, underlying event, and hadronisation models.
- Uncertainty on next-to-leading-order (NLO) electroweak corrections that modify the p_{T}^{V} distribution.

Table 8.9 lists the different uncertainties related to the modeling of the signal[2], while three categories are explained in detail below.

QCD Scale Uncertainties

The uncertainties mainly impacting the theoretical prediction of the $VH, H \to c\bar{c}$ signal production cross section can be attributed to the variations from the renormalization scale μ_R, and the factorization scale μ_F. QCD scale uncertainties are classified into the following types of uncertainties.

- The uncertainty on the inclusive cross-sections, Δ_{QCD}^{XS}, which is applied separately for the $qq \to WH$, $qq \to ZH$, and the $gg \to ZH$ processes as reported in [16].
- The **acceptance uncertainty** for the $VH, H \to c\bar{c}$ signal in the signal region, and high ΔR control region ($\Delta_{QCD}^{SR\text{-}CR}$). This uncertainty mainly varies the relative signal in the control region, and the impact on the signal region is negligible.
- The BDT **shape-based uncertainty** on the QCD scale.

The signal-control region *acceptance uncertainties* are estimated following the procedure described in [145]. This is done using the six correlated pairs of μ_R and μ_F, normalized to the same cross-section as the nominal scales μ_R^{nom} and μ_F^{nom} as given in Equation 8.3. The largest variation from all the six variations is considered as an *acceptance uncertainty*.

For the BDT *shape-based uncertainty*, the $[2, 2]$ scenario of Equation 8.3 is used to estimate the variation.

PDF and α_s Uncertainties

The uncertainties on the parton distribution function (PDF) and α_s are estimated following the recommendations in [146], considering 30 PDF Hessian uncertainty variations and a two-sided

---

[2] Note that the QCD scale uncertainties between p_{T}^{V} regions and jet multiplicity regions are not included in the table, since there uncertainties are cancelled out when measuring the signal strength of the $VH, H \to c\bar{c}$ process in the merged p_{T}^{V} and jet multiplicity regions.

α_s variation. These uncertainties are estimated for the $qq \to WH$, $qq \to ZH$, and the $gg \to ZH$ processes separately.

Parton Shower, Underlying Event, and Hadronisation Uncertainties

The uncertainty impacting the modeling of $H \to b\bar{b}$ and $H \to c\bar{c}$ decays can be attributed to the uncertainties in the parton shower, underlying event, and hadronisation models. These uncertainties are estimated from the difference between the nominal samples and alternate samples (described in Table 8.8), since the alternate signal sample is showered with Herwig 7 while the nominal sample is showered with Pythia 8. These uncertainties will be considered as an *acceptance uncertainty* on the p_{T}^{V} regions and jet multiplicity regions, and as a *shape-based uncertainty* on the BDT distribution.

Table 8.9: The uncertainties associated with the VH signal processes. The shape-based uncertainties are not included in the Table below.

Uncertainty	**Size of the uncertainty**
Δ^{XS}_{QCD} for $qq \to WH$	0.7%
Δ^{XS}_{QCD} for $qq \to ZH$	0.6%
Δ^{XS}_{QCD} for $gg \to ZH$	25%
$\Delta^{SR\text{-}CR}_{QCD}$	-
Δ_{BR} for $H \to b\bar{b}$	$+1.2 - 1.3\%$
Δ_{BR} for $H \to c\bar{c}$	$+5.5 - 2.0\%$
Δ_{PDF} for $qq \to WH$	Up to $\pm 2.5\%$
Δ_{PDF} for $qq \to ZH$	Up to $\pm 3\%$
Δ_{PDF} for $gg \to ZH$	Up to $\pm 4.5\%$
Δ_{α_s} for $qq \to WH$	Up to $\pm 2\%$
Δ_{α_s} for $qq \to ZH$	Up to $\pm 2\%$
Δ_{α_s} for $gg \to ZH$	Up to $\pm 3\%$
$\Delta_{PS\text{-}UE}$	From -5% up to $+2\%$

8.2.3 Modeling of V+jets and related uncertainties

The V+jets processes are the dominant background in the analysis, with W+jets mainly affects the one charged lepton channel, and Z+jets mainly affecting the zero and two charged lepton channel. As explained previously, the V+jets normalisation listed in Table 8.10 are allowed to float freely in the fit, since the normalisation is constrained through dedicated control regions.

The jet-flavors are merged into heavy, mixed, and light flavor components based on the similar kinematic properties, and each flavor component is floated freely in the fit, with further splitting them into the p_{T}^{V} and jet-multiplicity regions.

Table 8.10: The V+jets background template normalizations used in the fit.

Process	**Component**	**Flavors**	**Lepton channel**
Z+jets	Heavy flavor	bb, cc	0L, 2L
	Mixed flavor	bc, bl, cl	
	Light flavor	ll	
W+jets	Heavy flavor	bb, cc	0L, 1L
	Mixed flavor	bc, bl, cl	0L
		$bc, bl, cl, b\tau, c\tau$	1L
	Light flavor	ll	0L
		$ll, l\tau, \tau\tau$	1L

The high-ΔR control region of the two charged lepton channel BB tag region is used to constrain the Z+jets *heavy flavor* component. Similarly, the high-ΔR control region of the one charged lepton channel BB tag region is a candidate for constraining the W+jets *heavy flavor* component, but this region is dominated by the top quark processes[3]. However, there is sufficient W+jets *heavy flavor* component in the low-ΔR control region of the one charged lepton channel BB tag region as well, with less contamination from the top-quark processes. Therefore, the low-ΔR control region of the one charged lepton channel BB tag region is used to constrain the W+jets *heavy flavor* component.

For the V+jets *mixed flavor* component, it is constrained by the high-ΔR control region in the $C_T C_L$ and $C_T N$ tag regions. Finally, as discussed in Section 5.2.1 the V+jets *light flavor* component is constrained by the signal region[4] of $C_L N$ tag region.

The acceptance uncertainties for the V+jets processes discussed below are calculated enveloping the variations on the QCD scales μ_R, μ_F, the PS model, the ME, the PDF, and electroweak corrections. The shape-based uncertainties are also derived for the different types of systematic variations as described below.

V+jets acceptance uncertainties

The V+jets acceptance uncertainties can be mostly classified into the below four types. Since the V+jets are floated in the three flavor components (heavy, mixed, and light), the acceptance uncertainties are derived separately for the three flavor components. The acceptance uncertainties are listed in Table 8.11.

- The **charged lepton channel acceptance** uncertainty: This is the uncertainty on the extrapolation of the Z+jets yield from the two-charged lepton channel to the zero charged

---

[3] The two b-jet from the top decays are expected to have a large $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ between them, and therefore, end up in the high-ΔR control region.

[4] The $C_L N$ tag region does not have an actual signal region which is used in the fit, but this region is obtained from the same ΔR cuts satisfying the signal region in $BB, C_T C_T$,etc.

lepton channel, and the W+jets yield from the one-charged lepton channel to the zero charged lepton channel.

- The **flavor component acceptance** uncertainty: This is the acceptance uncertainty between the flavor combinations for the flavor components listed in Table 8.10. The acceptance uncertainty for the heavy flavor component in derived in the BB region, and harmonized with the other regions. Therefore, the bb component is taken as the higher purity component due to the larger contribution in the BB region. The acceptance uncertainty for the mixed flavor component is derived in the inclusive $C_T C_T + C_T C_L + C_T N$ region, and therefore, the cl component is taken as the higher purity component due to the larger contribution in this region. For the light flavor component in W+jets, ll is taken as the higher purity component to estimate the acceptance uncertainty.
- The **signal-control region acceptance** uncertainty: This is the acceptance uncertainty between the signal region and the control regions. In the calculation, the higher purity component is generally in the control regions, except in the extrapolation of the V+jets light flavor component from the signal region to the high ΔR control region. This is due to the V+light jet $C_L N$ control region having the same ΔR categorization as the signal region.
- The **jet-multiplicity region acceptance** uncertainty: The 3-jet and 4-jet W+jets components in the BB tag region of the zero-charged lepton channel is floated together. Hence, an acceptance uncertainty is considered between the 3-jet and 4-jet regions.

Table 8.11: The acceptance uncertainties associated with the V+jets processes.

Process	Acceptance uncertainty on	Size of the uncertainty
Z+jets	Extrapolation from 2L to 0L	2% to 10%
	Z + heavy flavor composition	8% to 12%
	Z + mixed flavor composition	4% to 10%
	Signal-Control region	5% to 30%
W+jets	Extrapolation from 1L to 0L	3% to 30%
	W + heavy flavor composition	4% to 25%
	W + mixed flavor composition	8% to 29%
	W + light flavor composition	9%
	Signal-Control region	2% to 50%
	Jet multiplicity region	12% to 20%

V+jets shape-based uncertainties

Four types of shaped-based modeling uncertainties are considered for the V+jets processes.

- The **CARL-based shape** uncertainties: The CARL framework is used to derive the shape variations between the nominal and alternate generators. The comparison with the alternate samples generated with MadGraph + Pythia 8 accounting for the PS and ME uncertainties.

- The **Sherpa 2.2.11 p_{T}^{V} shape** uncertainties: It was observed that the Sherpa 2.2.11 p_{T}^{V} distribution has a disagreement with data, as seen in Figure 8.2. Hence, a shape-based uncertainty derived by re-weighting the Sherpa 2.2.11 p_{T}^{V} spectrum to match the Sherpa 2.2.1 p_{T}^{V} spectrum is considered separately for the charged lepton channels, and the flavor components.
- The **QCD scale shape** uncertainties: These uncertainties are derived from the nominal Sherpa 2.2.11 samples, with the QCD scale variation of $[\mu_R/\mu_R^{nom}, \mu_F/\mu_F^{nom}] = [2, 2]$ of Equation 8.3 considered for the uncertainty, since it gives the largest variation.
- The **electroweak correction shape** uncertainties: The additive electroweak correction scheme of Figure 8.3 is used in the modeling of V+jets. Since the multiplicative electroweak correction scheme has the largest correction, this is used as a shape uncertainty. It should be noted that the impact of these uncertainties are small.

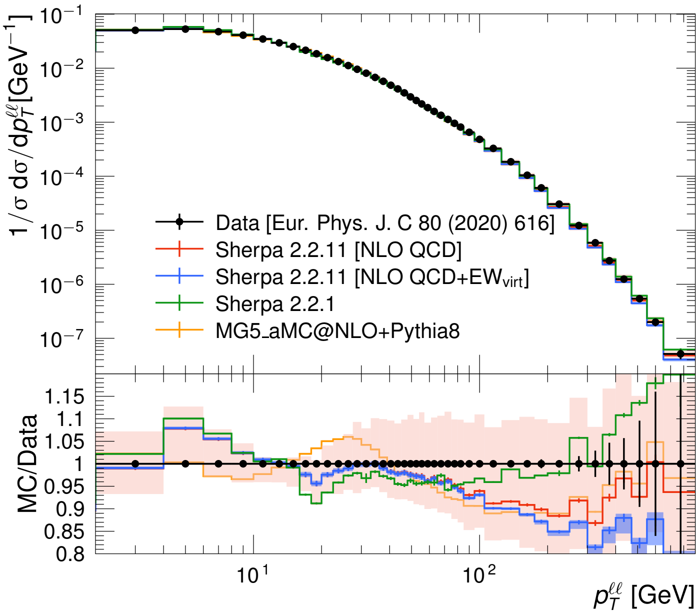


Figure 8.2: The normalized cross-section vs p_{T}^{V} spectrum using Sherpa 2.2.11 [106].

8.2.4 Modeling of top-quark processes and related uncertainties

The top quark processes are another dominant background in the $VH, H \to b\bar{b}/c\bar{c}$ analysis regions, mainly affecting the zero and one charged lepton channels. As discussed in Section 3.2.2, the top-quark processes are modelled through MC simulations using Powheg +Pythia 8, except in the two charged lepton channel of the BB tag region where the template is derived from data.

The $t\bar{t}$ and single-top Wt processes are merged into one component, but split into two flavor components in the zero and one charged leptons channels; top-bb and top-$bqqq$. The top-bb component is constrained by the high ΔR control region of the BB tag region, while the dedicated top-BC_T control region is used to constrain the top-$bqqq$ component. The top process (i.e $t\bar{t}$ and single-top Wt) in the two charged lepton channel of the C_TC_T, C_TC_L, C_TN tag regions, which is not estimated in a data driven method due to low yield, is constrained by the dedicated

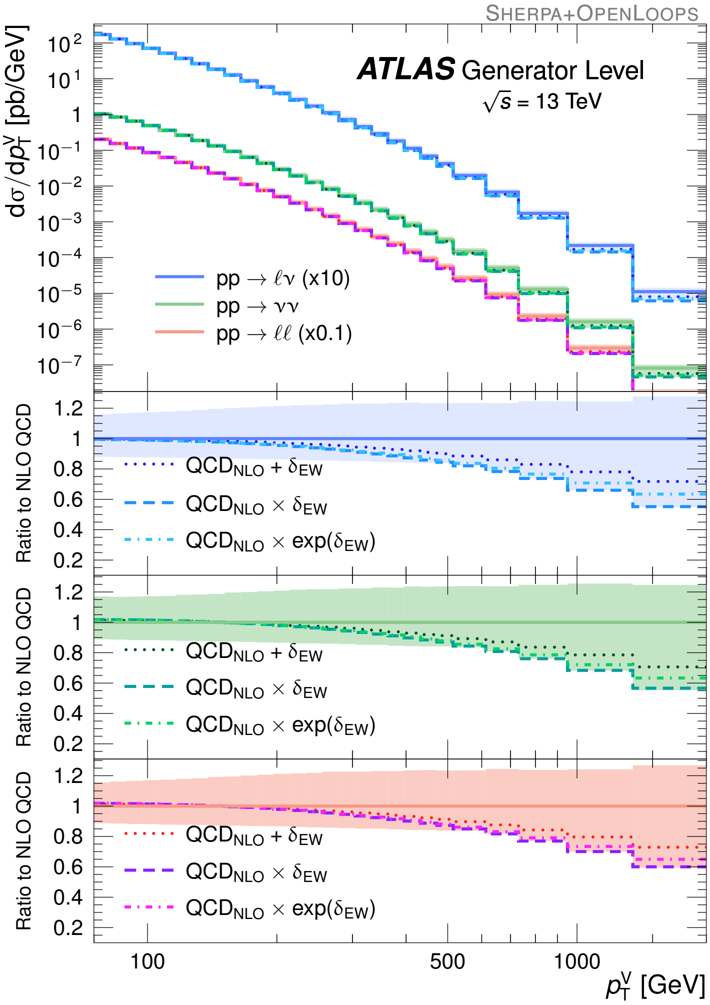


Figure 8.3: The electroweak correction schemes in the p_T^V spectrum using SHERPA 2.2.11 [106].

top-$e\mu$ control region.

The single-top s/t channel processes have a sub-dominant contribution in the analysis regions, and there is no dedicated control region to constrain this background in the analysis. A 17% acceptance uncertainty is applied to the single-top t-channel process, and since the single-top s-channel process has very small contribution, an overall cross-section uncertainty of 4.6% is applied without additional acceptance uncertainties. Table 8.12 summarizes the overall normalisation considered for the top quark processes.

Similar to V+jets, the acceptance uncertainties for the top-quark processes discussed below are calculated enveloping the variations on the QCD scales μ_R, μ_F, the PS model, and the ME. The shape-based uncertainties are also derived for the four types of systematic variations as described below.

Top-quark process acceptance uncertainties

The acceptance uncertainties for the top-quark processes fall into the categories described below. For the merged top component ($t\bar{t}$ and single-top Wt), the acceptance uncertainties are derived for each p_T^V and jet-multiplicity region, and therefore no dedicated p_T^V and jet-multiplicity

Table 8.12: The background template normalizations used in the fit. For the flavor components, *all* refers to all the flavor combinations in that category.

Process	Component	Flavors	Lepton channel	Uncertainty
Top ($t\bar{t}$ and single-top Wt-channel)	Top-bb	bb	0L, 1L	-
	Top-$bqqq$	bl, bc, cc, cl, ll	0L, 1L	-
	Top (in C_TC_T, C_TC_L, C_TN)	all	2L	-
Single-top s/t-channel	Single-top s-channel	all	0L, 1L, 2L	4.6%
	Single-top t-channel	all	0L, 1L	17%

acceptance uncertainties are considered for the merged top component.

- The **charged lepton channel acceptance** uncertainty: Since the one charged-lepton channel has the dominant top-quark contributions, the floating normalisation is mainly constrained from the one-charged lepton channel. Therefore, this acceptance uncertainty acts on the extrapolation of the top-quark process yields from the one-charged lepton channel to the zero-charged lepton channel.
- The **flavor component acceptance** uncertainty: This is the acceptance uncertainty between the flavor combinations for the top-$bqqq$ process as listed in Table 8.12, with the higher purity component decided as the top-bq process.
- The **signal-control region acceptance** uncertainty: This is the acceptance uncertainty between the signal region and the ΔR control regions. Since the majority of top-bb events and single-top t-channel events (with high p_{T}^{V}) are in the high-ΔR control region, the acceptance ratio is calculated with high-ΔR control region as the high purity region. Similarly, the majority of top-$bqqq$ events and single-top t-channel events (with low p_{T}^{V}) are in the signal region and top-BC_T control region, and therefore, the high purity regions are chosen as signal region and the top-BC_T control region.
- The **$t\bar{t}$ and single-top Wt acceptance ratio** uncertainty: $t\bar{t}$ and single-top Wt are combined into one component when fitting. Therefore, an acceptance uncertainty to account for the difference in the relative yields is considered, and is derived as $t\bar{t}$ being the higher purity component.
- The **p_{T}^{V} region acceptance** uncertainty: This acceptance uncertainty is considered only for the single-top t-channel process. Since the majority of the single-top t-channel process is in the lower p_{T}^{V} regions (75 GeV$< p_{\mathrm{T}}^{V} <$150 GeV and 150 GeV$< p_{\mathrm{T}}^{V} <$250 GeV regions), the low p_{T}^{V} regions are chosen as the higher purity regions.
- The **jet-multiplicity region acceptance** uncertainty: This acceptance uncertainty is also considered only for the single-top t-channel process. Since the 3-jet region has the majority of the single-top t-channel events, the 3-jet region is chosen as the high purity region in the acceptance ratio.

Table 8.13 summarizes the acceptance uncertainties considered for the top quark processes.

Table 8.13: The acceptance uncertainties associated with the top-quark processes. Top refers to the $t\bar{t}$ and single-top processes.

Process	Acceptance uncertainty on	Size of the uncertainty
$t\bar{t}$ and single-top Wt	1L to 0L	2% to 8%
	Top-$bqqq$ flavor composition	5% to 10%
	Signal-Control regions	2% to 10%
	Wt to $t\bar{t}$ ratio	12% to 48%
Single-top s/t channel	1L to 0L	6%
	Signal-Control regions	3% to 17%
	p_{T}^{V} regions	7% to 15%
	Jet multiplicity regions	15%

Top-quark process shape-based uncertainties

Three types of shaped-based modeling uncertainties are considered for the top-quark processes.

- The **CARL-based shape** uncertainties: CARL framework is used to derive the shape-based uncertainties for $t\bar{t}$, single-top Wt and single-top t-channel processes, between the nominal generator (Powheg +Pythia 8) and two alternate generators. The comparison with Powheg +Herwig 7 accounts for the parton shower uncertainties, while the comparison with MadGraph + Pythia 8 accounts for the matrix element differences.
- The **single-top Wt DS-DR** uncertainties: A shape uncertainty is attributed to the difference between interference of $t\bar{t}$ and single-top Wt processes, discussed in Section 3.2.2. Since the current analysis uses DS samples, the difference between DS and DR is taken as a shape-based and normalisation uncertainty.
- The **ISR and FSR shape** uncertainties: These uncertainties are derived similarly to the VH signal, with the QCD scale variation of $[\mu_R/\mu_R^{nom}, \mu_F/\mu_F^{nom}] = [2,2]$ of Equation 8.3 considered for the uncertainty since it gives the largest variation.

8.2.5 Modeling of di-boson processes and related uncertainties

Di-boson processes are a subdominant background in the analysis, and mainly classified into four categories; the WW, WZ, and the ZZ processes, with WZ process further classified into processes where the W-boson decays hadronically ans the Z-boson decays leptonically (denoted as Wh-Zl), and vice-versa (denoted as Wl-Zh).

ZZ process with a Z-boson decaying to two b-jets, and the Wl-Zh process with the Z-boson decaying to two b-jets, are the two dominant di-boson components in the BB tag region. In the C_TC_T, C_TC_L, and the C_TN tag regions, the dominant di-boson components components differ by the charged lepton channel;

- The WW, Wh-Zl, and ZZ processes dominate the zero-charged lepton channel through the decay modes like $W \to l\nu, W \to cs$, $Z \to \nu\nu, W \to cs$, and $Z \to \nu\nu, Z \to cc$.

- The WW and Wl-Zh processes dominate the one-charged lepton channel through the decay modes like $W \to l\nu, W \to cs$ and $W \to l\nu, Z \to cc$.
- The Wh-Zl and ZZ processes dominate the two-charged lepton channel through the decay modes like $Z \to ll, W \to cs$ and $Z \to ll, Z \to cc$.

Since there is no dedicated control region for constraining the di-boson processes, and overall acceptance uncertainty is derived for the different di-boson components, as shown in Table 8.14.

Table 8.14: The background template normalizations for di-boson used in the fit.

Process	**Component**	**Lepton channel**	**Uncertainty**
Di-boson	WW	0L, 1L, 2L	16%
	ZZ	0L, 1L, 2L	17%
	WZ	0L, 1L, 2L	19%
	$ggWW, ggZZ$	0L, 1L, 2L	30%

Similar to V+jets, the acceptance uncertainties for the diboson processes discussed below are calculated enveloping the variations on the QCD scales μ_R, μ_F, the PS model, the ME, the PDF, and electroweak corrections. The shape-based uncertainties are also derived for the different types of systematic variations as described below.

Di-boson acceptance uncertainties

Since the overall normalisation is applied to the overall di-boson components, four types[5] of di-boson acceptance uncertainties are considered in the analysis, which is summarized in Table 8.15. The di-boson acceptance uncertainties are derived for the four components WW, Wl-Zh, Wh-Zl, and ZZ.

- The **charged lepton channel acceptance** uncertainty: This uncertainty covers the acceptance ratio between the dominant di-boson components in the charged lepton channels. Since the WW and Wl-Zh processes have more yield in the one charge lepton channel, the higher purity component for the calculation of the acceptance ratio is the one charged lepton channel, and the uncertainty is applied to both the zero and one charged lepton channels. Similarly for the Wh-Zl and ZZ processes, the higher purity component is the two charged lepton channel, and the uncertainty is applied to both the zero and two charged lepton channels.
- The **jet-multiplicity region acceptance** uncertainty: This uncertainty is considered for the acceptance ratio between the jet-multiplicity regions. Since the 2-jet region has the higher yield, the higher purity component for the calculation of the acceptance ratio is generally the 2-jet region, and the uncertainty is then applied to other jet multiplicity regions.
- The **signal-control region acceptance** uncertainty: This is the acceptance uncertainty between the signal region and the control regions. Since majority of di-boson events are in

---

[5] There are actually 5 types of uncertainties, with the 5th one being the acceptance uncertainty between p_T^V and jet-multiplicity regions. However, this uncertainty is ignored in the measurements for the $VH, H \to c\bar{c}$ process.

the signal region, the higher purity component for the calculation of the acceptance ratio is chosen as the signal region. The uncertainty is then applied to both the signal and control regions.

- The p_{T}^{V} **acceptance** uncertainty: This uncertainty is considered for the acceptance ratio between the p_{T}^{V} regions. In the calculation of the acceptance ratio, the higher purity component is chosen as the 150 GeV< p_{T}^{V} <250 GeV region, since it is the common region in all the three charged lepton channels.

Table 8.15: The acceptance uncertainties associated with the di-boson processes.

Acceptance uncertainty on	**Size of the uncertainty**
Charged lepton channels	2% to 23%
Jet multiplicity regions	10% to 42%
p_{T}^{V} regions	3% to 16%
Signal-Control regions	6% to 50%

Di-boson shape-based uncertainties

- The **CARL-based shape** uncertainties: The CARL framework is used to derive the shape variations between the nominal generator and alternate generators. As listed in Table 8.8, two alternate generators are used for deriving the shape-based uncertainties; POWHEG +PYTHIA 8 accounts for the difference in the matrix element and the parton shower, while SHERPA 2.2.1 is used to derive a p_{T}^{V} based shape uncertainty similar to V+jets.
- The **QCD scale shape** uncertainties: These uncertainties are derived similarly to the VH signal, with the QCD scale variation of $[\mu_R/\mu_R^{nom}, \mu_F/\mu_F^{nom}] = [2,2]$ of Equation 8.3 considered for the uncertainty.
- The **PDF + α_s shape** uncertainties: These uncertainties are derived similarly to the VH signal, but applied as a combined PDF + α_s shape-based uncertainty. It should be noted that the impact of these uncertainties are negligible in the $VH, H \to b\bar{b}/c\bar{c}$ analysis.
- The **electroweak correction shape** uncertainties: Similar to V+jets, the additive electroweak correction scheme is used in the modeling of the di-boson processes. Since the multiplicative electroweak correction scheme has the largest correction, this is used as a shape uncertainty.

Chapter 9

The fit framework and statistical description of data

The signal strengths for $\mu_{VH,H\to b\bar{b}}$ and $\mu_{VH,H\to c\bar{c}}$ processes are extracted simultaneously through a likelihood fit on the signal and background templates of all categories described in Section 5.3.

9.1 The fit model

The signal strength $\mu_{VH,H\to c\bar{c}}$ and $\mu_{VH,H\to b\bar{b}}$ are extracted by maximizing the likelihood of the likelihood function given in Equation 9.1, by fitting the data to signal and background templates in the $VH, H\to c\bar{c}$ and $VH, H\to b\bar{b}$ analysis categories as described in Section 9.1.1.

$$\mathcal{L}(\vec{\mu},\vec{\theta},\vec{\gamma},\vec{\tau}) = \prod_{i\in\text{bins}} \text{Pois}(N_i|\vec{\mu}s_i(\vec{\theta}) + b_i(\vec{\theta},\vec{\gamma},\vec{\tau})) \prod_{\theta\in\vec{\theta}} \frac{1}{\sqrt{2\pi}} e^{-\theta^2} \prod_{\in\text{bins}} \text{Gauss}(\beta_i|\gamma_i\beta_i, \sqrt{\gamma_i\beta_i}) \ . \quad (9.1)$$

The first term represents the Poissonian probability terms, based on the expected and observed event yields in all the bins. Here, N_i, s_i and b_i are the number of measured data events, the number of expected signal events, and the number of expected background events in bin i. The parameter μ is the signal strength described above. This term can also be represented using the Poisson distribution as

$$\mathcal{L}(\mu) = \prod_{i\in\text{bins}} \text{Pois}(N_i|\mu s_i + b_i) = \prod_{i\in\text{bins}} \frac{(\mu s_i + b_i)^{N_i}}{N_i!} e^{-(\mu s_i + b_i)}. \quad (9.2)$$

Systematic uncertainties are introduced in the fit in the form of nuisance parameters $\vec{\theta}$. Hence the second term of Equation 9.1 represents the nuisance parameters introduced as Gaussian penalty terms with a mean value of zero and a variance of one.

$$\mathcal{L}_{\text{Aux}} = \prod_{\theta\in\vec{\theta}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}}. \quad (9.3)$$

Free floating nuisance parameters are labelled as $\vec{\tau}$, and they modify the overall normalizations of the major backgrounds, and are determined from data. Also, uncertainties related to the limited statistics of the simulated samples needs to be considered. These, known as γ-parameters, are introduced per bin and can modify the background yield in the particular bin as $b_i(\vec{\theta}) \rightarrow \gamma_i b_i(\vec{\theta})$. In the likelihood description γ-parameters are introduced as,

$$\mathcal{L}_{\text{BkgStat}}(\vec{\gamma}) = \prod_{i \in \text{bins}} \text{Gauss}(\beta_i | \gamma_i \beta_i, \sqrt{\gamma_i \beta_i}). \tag{9.4}$$

Where $\beta_i = \frac{1}{\sigma^2_{\text{rel}}}$, and σ^2_{rel} is the relative statistical uncertainty on the expected total background yield. Hence, nuisance parameters modify the expected background and signal yields as

$$s_i \rightarrow s_i(\vec{\theta}) \qquad\qquad b_i \rightarrow b_i(\vec{\theta}, \vec{\gamma}, \vec{\tau}). \tag{9.5}$$

9.1.1 The fit regions and variables

The categories used in the fit are the regions that were discussed in Section 5.3 and illustrated in Figure 5.4, including the $VH, H \rightarrow b\bar{b}$ boosted regions that are not discussed in this thesis. Different variables are used for building the signal and background templates in the fit categories, which are detailed in Table 9.1. Some regions have just one bin, and hence no variable, since these regions are only used for constraining the normalization of background components.

Table 9.1: The different variables used in the fit categories.

Region	**Variable in lepton channel**			**Tagging category**
	0L	**1L**	**2L**	
Signal region	BDT output	BDT output	BDT output	$BB \mid C_T C_T + C_T C_L \mid C_T N$
High ΔR control region	One bin	p_T^V	p_T^V	$BB \mid C_T N$
		$m_{j_1 j_2}$	$m_{j_1 j_2}$	$C_T C_T \mid C_T C_L$
Low ΔR control region	-	BDT output	-	BB
Top control region	$m_{j_1 j_2}$	$m_{j_1 j_2}$	One bin	BC_T
V+light jet control region	One bin	p_T^V	p_T^V	$C_L N$

9.1.2 The parameters of interest and signal normalization

The default fit model used in the analysis will extract two parameters of interest (POI); the signal strengths $\mu_{VH,H\rightarrow c\bar{c}}$ and $\mu_{VH,H\rightarrow b\bar{b}}$. The signal strengths are initially set to their standard model expectations (i.e $\mu_{VH,H\rightarrow c\bar{c}} = \mu_{VH,H\rightarrow b\bar{b}} = 1$), but are allowed to vary in the fit.

The default two POI can also be changed or expanded depending on the parameters of interest that need to be extracted. In the *di-boson validation* study, the fit will extract the signal strengths for the di-boson processes; $\mu_{VZ,Z\rightarrow c\bar{c}}$ and $\mu_{VZ,Z\rightarrow b\bar{b}}$. When extracting $\mu_{VZ,Z\rightarrow c\bar{c}}$ and $\mu_{VZ,Z\rightarrow b\bar{b}}$, the categories and variables are exactly the same as that of Figure 5.4 and Table 9.1, except for the signal regions where a dedicated BDT to measure the $VZ, Z \to c\bar{c}$ and $VZ, Z \to b\bar{b}$ processes are used.

The default two POI can also be de-correlated (in other words, expanded) in different categories to extract the signal strength in that category. In the analysis, de-correlated fits are done in the charged lepton channels (six POI), and the vector boson channel (four POI). This de-correlation is also be applied for the di-boson validation study, to measure the $VZ, Z \to c\bar{c}$ and $VZ, Z \to b\bar{b}$ signal strengths expanded in the charged lepton channels and vector boson channels.

Further, the uncertainty in the branching ratio for $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$, and the extrapolation uncertainty for the $ggZH, qqZH, WH$ processes between signal and control regions, are introduced as a normalization effect on the signal templates.

9.1.3 The background normalization

The background predictions in each region, other than multi-jet, have yields that are estimated from MC simulations using the theoretical cross sections and branching ratios. The uncertainty from the theoretical cross section, branching fraction, or the acceptance between regions, will manifest in the background templates as normalization effects on the template.

For the di-boson and single-top s/t-channel processes, normalization effects from the uncertainty on cross-sections and branching fractions are parameterised by a Gaussian prior on the distribution function of the nuisance parameter. Hence, the normalization of di-boson and single-top s/t-channel processes are only allowed to changed within a given uncertainty as given in Table 8.13 and Table 8.14. In the case of V+jets, $t\bar{t}$ and single-top Wt-channel processes, they have dedicated control regions to constrain the normalization, and hence the normalization is free to float without an additional likelihood penalty.

The acceptance uncertainty between the different categories, and associated with the different background components (as discussed in Section 8.2) also affect the normalization of the background templates. These uncertainties are also introduced in the fit model as nuisance parameters parameterised using a Gaussian prior.

9.1.4 Other nuisance parameters

In addition to the modelling nuisance parameters that have normalization effects, there are three other categories of nuisance parameters introduced to the fit; experimental systematic uncertainties, shaped based modelling systematic uncertainties, and the uncertainty due to limited simulated samples.

Experimental systematic uncertainties

Experimental systematic uncertainties are considered for the objects used in the analysis, and can be broadly classified into uncertainties related to $E_{\mathrm{T}}^{\mathrm{miss}}$, leptons, jets, and flavor tagging. These uncertainties are considered in the analysis by following the guidelines from dedicated groups in the ATLAS experiment which work on the specific objects, and was discussed in Section 8.1.

Modelling shape systematic uncertainties

Modelling uncertainties that affect the shapes of the distributions used in the fit (BDT output, $m_{j_1 j_2}$, p_{T}^{V}) are introduced to the fit model as nuisance parameters that vary the bin contents[1] of the signal or background template. Similar to the acceptance uncertainties or experimental systematic uncertainties, modelling shape uncertainties are also associated with an external constraint parameterised by a Gaussian prior on the distribution on the nuisance parameter.

Uncertainty due to the limited simulated samples

The uncertainty due to the limited simulated samples (also known as the *MC statistical uncertainty*), especially for the V+jets and single-top s/t templates, need to be considered in the fit model; the impact of the MC statistical uncertainty in determining the signal strength is large especially at higher BDT scores. These nuisance parameters are parameterised by Poisson priors.

9.2 Extraction of the significance and Limit

The likelihood of Equation 9.1 is maximized to determine the parameters of interest, which are the signal strengths $\mu_{VH,H\to c\bar{c}}$ and $\mu_{VH,H\to b\bar{b}}$, and their uncertainties. To test a certain hypothesis, the profile likelihood ratio is defined as

$$\lambda(\vec{\mu}) = \frac{\mathcal{L}(\vec{\mu}, \hat{\hat{\vec{\theta}}}, \hat{\hat{\vec{\gamma}}})}{\mathcal{L}(\hat{\vec{\mu}}\,, \hat{\vec{\theta}}, \hat{\vec{\gamma}})} \tag{9.6}$$

where $\{\hat{\vec{\mu}}, \hat{\vec{\theta}}, \hat{\vec{\gamma}}\}$ are the values which maximise the likelihood for all the parameters, and $\{\hat{\hat{\vec{\theta}}}, \hat{\hat{\vec{\gamma}}}\}$ are the values which maximise the likelihood for a fixed value of $\vec{\mu}$. This likelihood ratio is defined in the range [0,1], and when the ratio is closer to one, this implies that there is a good agreement between data and the chosen hypothesis.

From the likelihood ratio, the test statistic q_μ can be defined as

$$q_{\vec{\mu}} = \begin{cases} -2\ln\lambda(\vec{\mu}) & \hat{\mu} \geq \mu \\ 0 & \hat{\mu} < \mu \end{cases} . \tag{9.7}$$

For the case $\mu = 0$, the background only hypothesis with no signal (i.e the null hypothesis) is tested. Hence, it is possible to define the p-value for the $\mu = 0$ hypothesis as

$$p_{\vec{\mu}} = \int_{q_{\vec{\mu},\text{observed}}}^{\infty} f(q_{\vec{\mu}}|0)\,dq_{\vec{\mu}} \; , \tag{9.8}$$

where $f(q_{\vec{\mu}}|0)$ is the probability density function of $q_{\vec{\mu}}$ under the assumption of $\mu = 0$, and $q_{\vec{\mu},\text{observed}}$ is the value of $q_{\vec{\mu}}$ observed in data. Hence, the p-value of Equation 9.8 can be used as a measure of the probability of data being incompatible with the hypothesis $\mu = 0$. Using the

---

[1] Calculated from a linear interpolation between the nominal template and alternate template.

p-value and the inverse Gaussian cumulative distribution Φ^{-1}, the significance (the number of Gaussian standard deviations above the background) can be calculate as

$$Z = \Phi^{-1}(1 - p). \tag{9.9}$$

For the $VH, H \to c\bar{c}$ process, an upper limit at 95% confidence level on the $VH, H \to c\bar{c}$ signal strength is calculated using the test statistic defined in Equation 9.10, based on a modified frequentest approach $\mathrm{CL_s}$ method.

$$q = -2\ln\frac{\mathcal{L}_{s+b}}{\mathcal{L}_b} = -2\ln\frac{\mathcal{L}(\mu = 1, \hat{\hat{\vec{\theta}}}(1), \hat{\hat{\vec{\gamma}}}(1))}{\mathcal{L}(\mu = 0, \hat{\hat{\vec{\theta}}}(0), \hat{\hat{\vec{\gamma}}}(0))}. \tag{9.10}$$

Here $\mathcal{L}_{s+b}$ is the likelihood of the nominal signal (i.e $\mu = 1$) plus background hypothesis, and $\mathcal{L}_b$ is the likelihood of the background only hypothesis (i.e $\mu = 0$). The 95% confidence level limit corresponds to the signal strength μ with a p-value of 0.05.

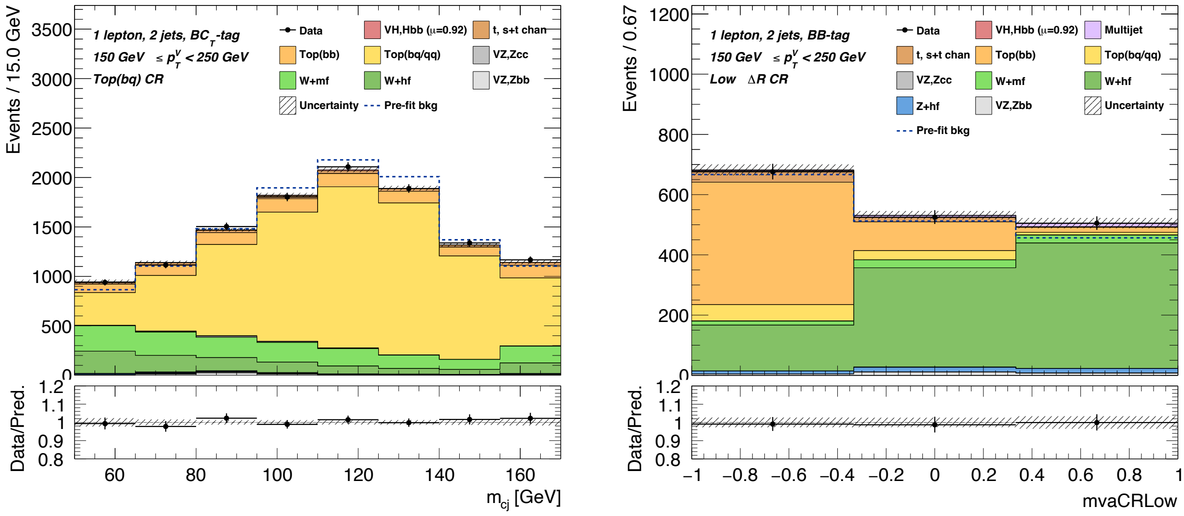
Chapter 10

The results

This chapter summarizes the results obtained using the combined fit framework discussed in the Chapter 9. The main results presented in this chapter are the measurement of the signal strength on the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ processes, and the interpretation of the measured signal strengths in the κ-framework. In addition to the main results, the signal strength measurement on the $VZ, Z \to c\bar{c}$ and $VZ, Z \to b\bar{b}$ processes, and the κ_c and κ_b measurement is also presented.

10.1 The agreement between data and background modelling

The validity of background modelling is assessed through the agreement between data and background templates after the statistical analysis (i.e post-fit). Hence, this section presents several important post-fit kinematic distributions. First, Figures 10.1 to 10.2 shows the agreement between the data and background modelling in kinematic distributions of control regions used in the fit.



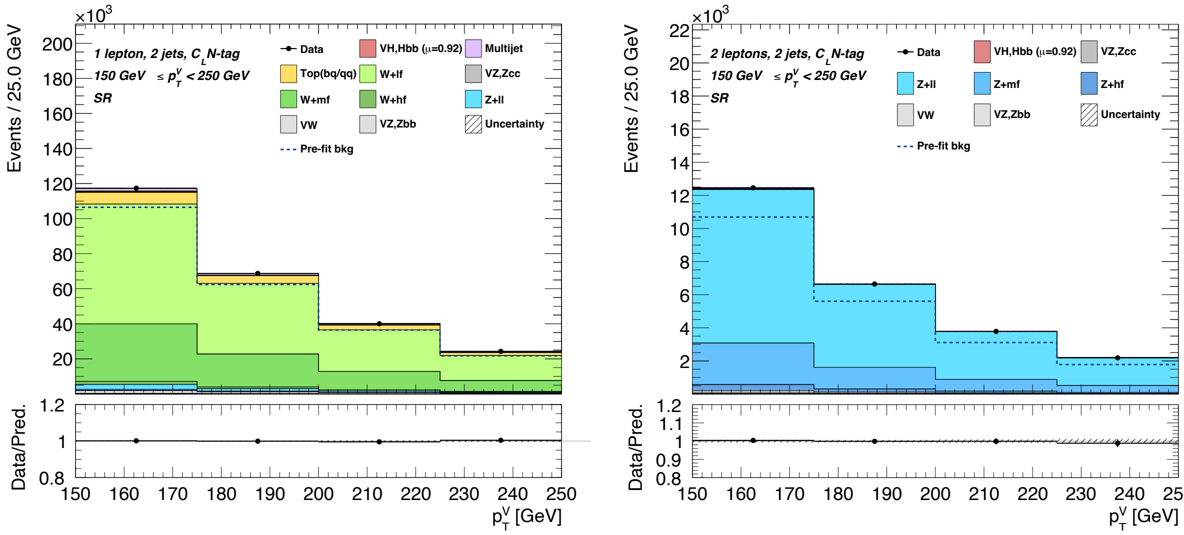
(a) top bq control region, 1L channel.

(b) W+hf control region, 1L channel.

Figure 10.1: Postfit distributions in the top and V+jets control region in the one charged lepton channel.

In all the post-fit distributions, the blue dotted line is the total signal and background distribution before fitting. The hashed band shows the total uncertainty, which includes the statistical uncertainty and all systematic uncertainties.

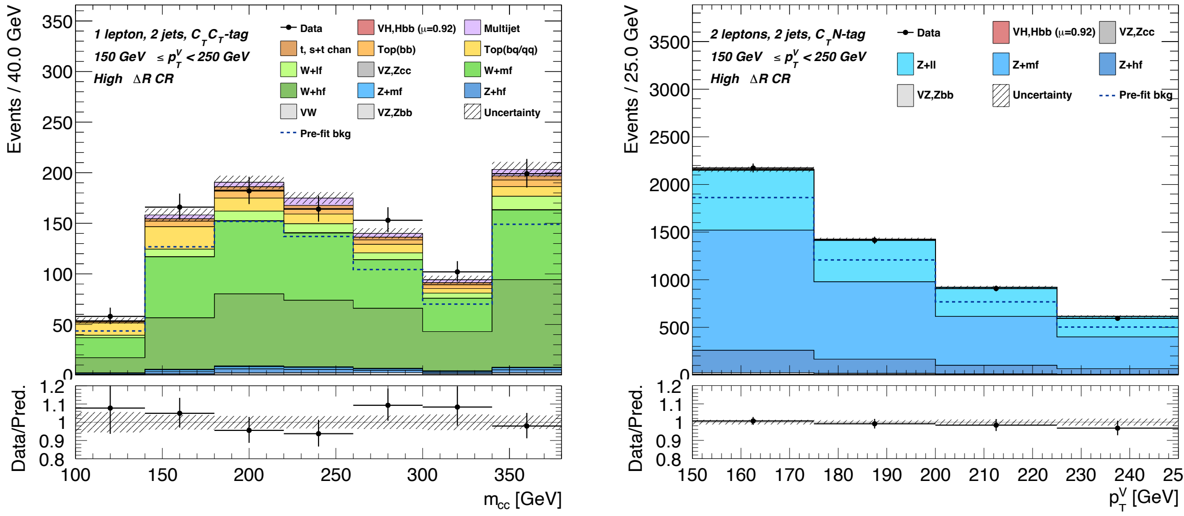
Figures 10.1(a) shows the post-fit distribution for $m_{j_1j_2}$ in the top bq control region in the one charged lepton channel, and Figures 10.1(b) shows the post-fit MVA distribution in the $W+$ heavy flavor control region. Then, Figures 10.2(a) and 10.2(b) shows the post-fit distribution for p_T^V in the V+light flavor jet control region. All post-fit distributions are seen to have a good agreement for the background modelling with data, within the total (statistical and systematic) uncertainty.



(a) V+lf control region, 1L channel.

(b) V+lf control region, 2L channel.

Figure 10.2: Post-fit distributions in the V+jets light flavor control regions in the one and two charged lepton channels.



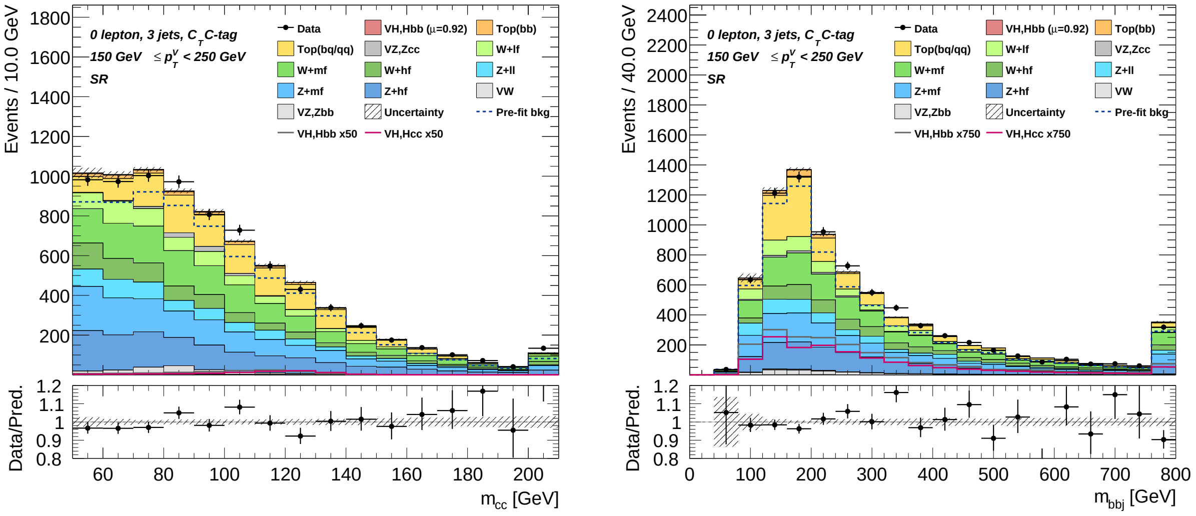
(a) 150 GeV<p_T^V <250 GeV, 2-jets, C_TC_T-tag region.

(b) 150 GeV<p_T^V <250 GeV, 2-jets, C_TN-tag region.

Figure 10.3: Post-fit distributions for the high ΔR control region in the of the one and two charged lepton channels.

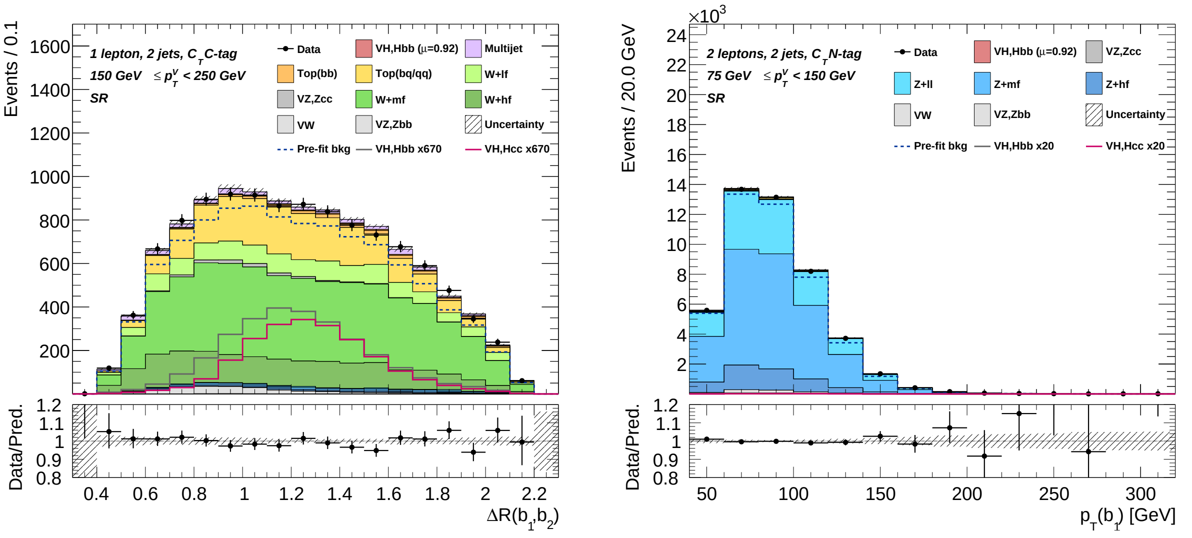
Figure 10.3 presents also the post-fit distribution for $m_{j_1 j_2}$ (p_{T}^{V}) in the high ΔR control region of the one and two charged lepton channel. The high ΔR control regions are also seen to have good agreement for background modeling with data, within the total uncertainty.

Finally, post-fit distributions for the agreement of the background modeling with data are also presented for the kinematic variables $m_{j_1 j_2}$, $m_{j_1 j_2 j_3}$, $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$, and $p_{\mathrm{T}}(j_1)$ in Figure 10.4. All plots shown are in the signal region, where such kinematic variables are used as input for the BDT, as described in Section 7.2. There is good agreement for background modeling for the post-fit distributions, confirming the validity of the kinematic variables used for the BDT training.



(a) $m_{j_1 j_2}$, $C_T C$-tag region, 0L channel.

(b) $m_{j_1 j_2 j_3}$, $C_T C$-tag region, 0L channel.



(c) $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$, $C_T C$-tag region, 1L channel.

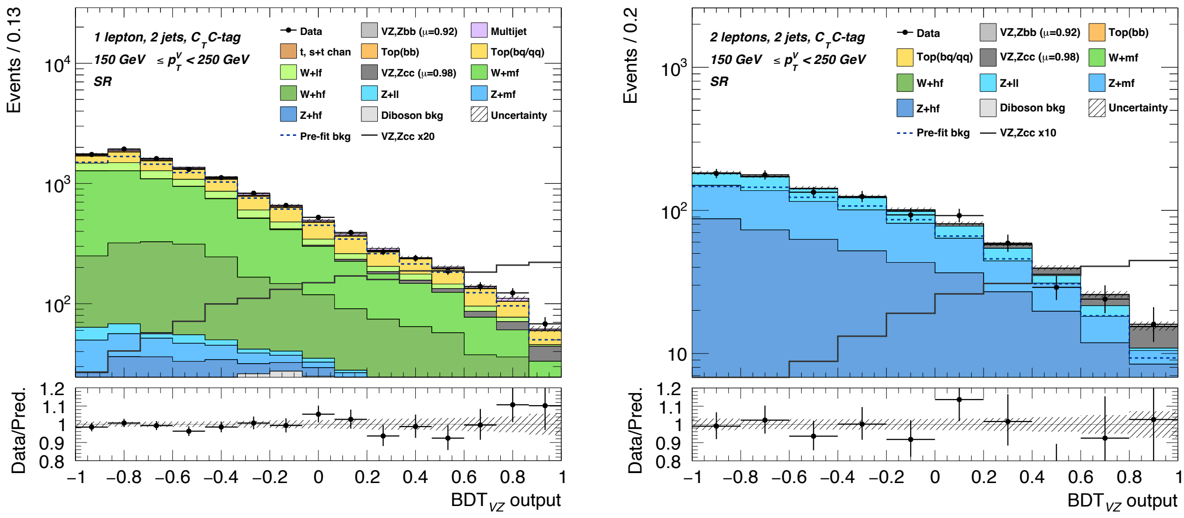
(d) $p_{\mathrm{T}}(j_1)$, $C_T N$-tag region, 2L channel.

Figure 10.4: Post-fit distributions for kinematic variables $m_{j_1 j_2}$, $m_{j_1 j_2 j_3}$, $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$, and $p_{\mathrm{T}}(j_1)$ used to train the BDT.

10.2 The validation result using diboson samples

The signal strength measurement on the $VZ, Z \to c\bar{c}$ and $VZ, Z \to b\bar{b}$ is done as a validation of the modelling of SM processes, and the signal strengths of $VZ, Z \to c\bar{c}$ and $VZ, Z \to b\bar{b}$ are measured using the same analysis strategy, but using a dedicated BDT template (i.e the BDT$_{VZ}$ output) which is obtained with diboson trained as the signal processes (instead of the $VH, H \to c\bar{c}$ or $VH, H \to b\bar{b}$ processes).

Figures 10.5 shows the post-fit distribution for the BDT$_{VZ}$ output in the one and two charged lepton channel. As seen from the distributions, the $VZ, Z \to c\bar{c}$ signal is well separated from the other diboson processes and other background processes in the high BDT$_{VZ}$ output bins, and the overall modelling agrees well with data within the total uncertainty.



(a) 150 GeV<p_T^V <250 GeV, 2-jets, C_TC-tag region. (b) 150 GeV<p_T^V <250 GeV, 2-jets, C_TC-tag region.

Figure 10.5: Examples of the post-fit distribution of the BDT$_{VZ}$ output in the one and two charged lepton channels. The black solid line shows the $VZ, Z \to c\bar{c}$ signal yield scaled by 20 (10) times in the one (two) charged lepton channel.

The measured signal strengths for the $VZ, Z \to b\bar{b}$ and $VZ, Z \to c\bar{c}$ process in all three charged lepton channels yield a central value and uncertainties as given below.

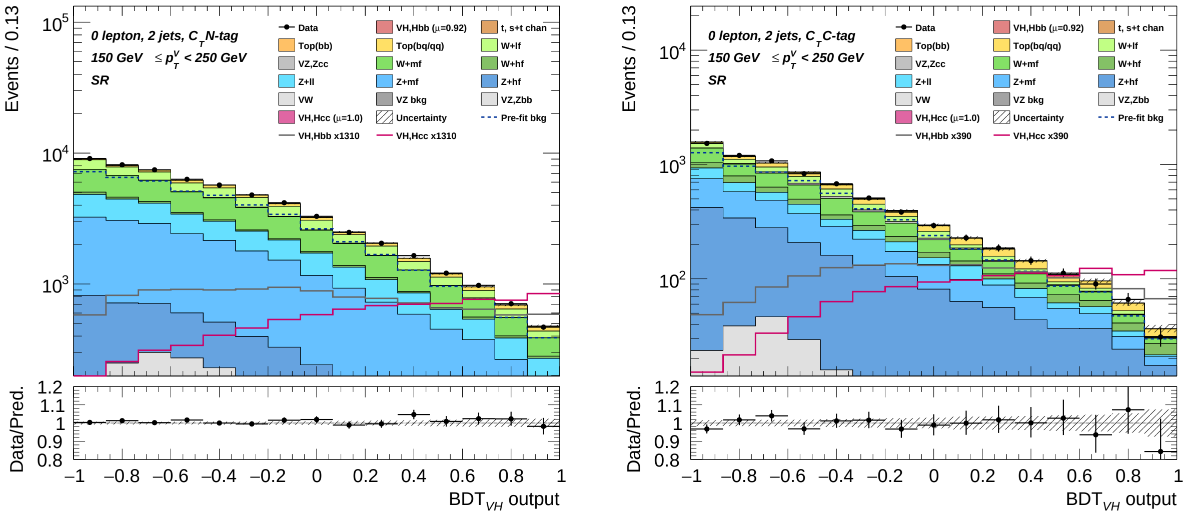
$$\mu_{VZ,Z\to b\bar{b}} = 0.92^{+0.13}_{-0.11} = 0.92\ ^{+0.05}_{-0.05}\ \text{(statistical)}\ ^{+0.12}_{-0.10}\ \text{(systematic)}$$

$$\mu_{VZ,Z\to c\bar{c}} = 0.98^{+0.25}_{-0.22} = 0.98\ ^{+0.13}_{-0.13}\ \text{(statistical)}\ ^{+0.22}_{-0.18}\ \text{(systematic)}$$

The obtained result is in good agreement with the standard model prediction, and confirms that the same analysis strategy can be adopted to measure the signal strengths of the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ processes. For the $VZ, Z \to c\bar{c}$ process, the observed (expected) significance is 5.2σ (5.3σ) over the background-only prediction, making this the first observation of the $VZ, Z \to c\bar{c}$ process in the ATLAS experiment. The first observation of the $VZ, Z \to c\bar{c}$ process in a hadron collider was achieved by the CMS experiment [31].

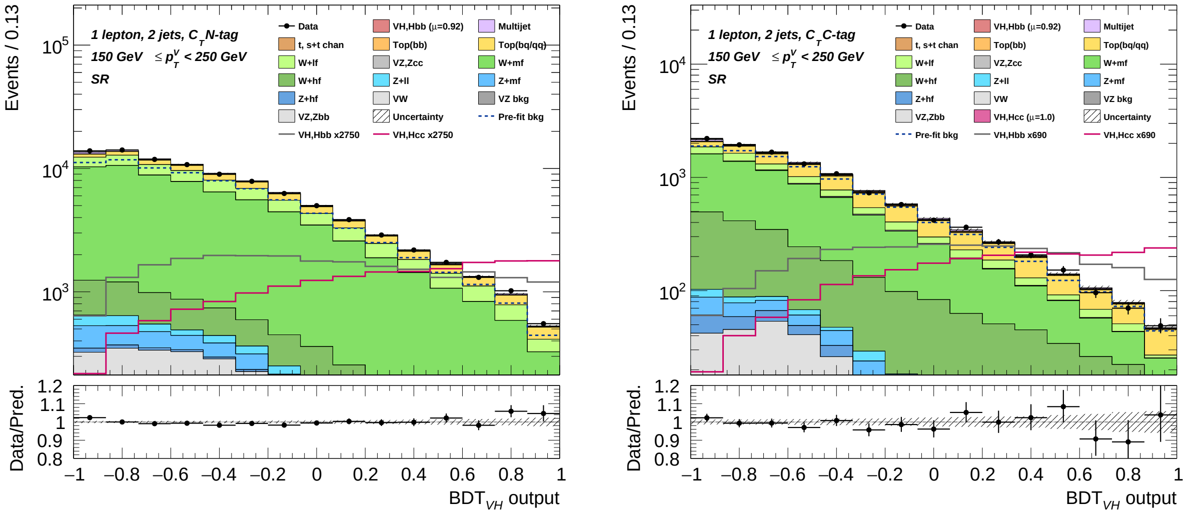
10.3 The $VH, H \to c\bar{c}$ signal strength measurement

The main result of the analysis is the measurement of the signal strengths of the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ processes through the combined fit discussed in Chapter 9. Figures 10.6 to 10.8 shows the post-fit BDT$_{VH}$ output distributions in some of the sensitive signal regions, which has good agreement between the modelling and data, within the total uncertainty.



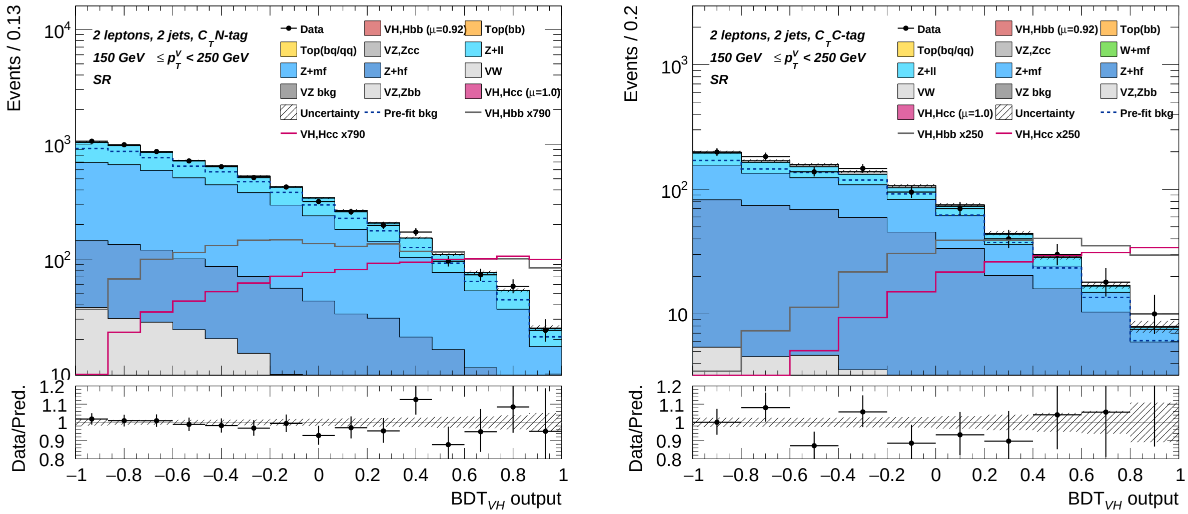
(a) 150 GeV<p_{T}^{V} <250 GeV, 2-jets, $C_T N$-tag region. (b) 150 GeV<p_{T}^{V} <250 GeV, 2-jets, $C_T C$-tag region.

Figure 10.6: BDT$_{VH}$ output post-fit distribution for the signal region in the zero charged lepton channel. The pink solid line shows the scaled $VH, H \to c\bar{c}$ signal yield, while the grey solid line shows the scaled $VH, H \to b\bar{b}$ signal yield.



(a) 150 GeV<p_{T}^{V} <250 GeV, 2-jets, $C_T N$-tag region. (b) 150 GeV<p_{T}^{V} <250 GeV, 2-jets, $C_T C$-tag region.

Figure 10.7: BDT$_{VH}$ output post-fit distribution for the signal region in the one charged lepton channel. The pink solid line shows the scaled $VH, H \to c\bar{c}$ signal yield, while the grey solid line shows the scaled $VH, H \to b\bar{b}$ signal yield.



(**a**) 150 GeV<p_{T}^{V} <250 GeV, 2-jets, $C_T N$-tag region. (**b**) 150 GeV<p_{T}^{V} <250 GeV, 2-jets, $C_T C$-tag region.

Figure 10.8: BDT$_{VH}$ output post-fit distribution for the signal region in the two charged lepton channel. The pink solid line shows the scaled $VH, H \to c\bar{c}$ signal yield, while the grey solid line shows the scaled $VH, H \to b\bar{b}$ signal yield.

The measured signal strengths for the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ processes in all three charged lepton channels yield a central value and uncertainties as shown below.

$$\mu_{VH,H\to b\bar{b}} = 0.92^{+0.16}_{-0.15} = 0.92\ ^{+0.10}_{-0.10}\ \text{(statistical)}\ ^{+0.13}_{-0.11}\ \text{(systematic)},$$
$$\mu_{VH,H\to c\bar{c}} = 1.0^{+5.4}_{-5.2} = 1.0\ ^{+4.0}_{-3.9}\ \text{(statistical)}\ ^{+3.7}_{-3.5}\ \text{(systematic)},$$

is in good agreement with the standard model prediction, within uncertainties. A correlation of 5% is seen between the two signal strengths $\mu_{VH,H\to c\bar{c}}$ and $\mu_{VH,H\to b\bar{b}}$, since there is non-negligible contribution of $VH, H \to b\bar{b}$ in the $VH, H \to c\bar{c}$ signal regions as seen in Figures 10.6 to 10.8. The two signal strengths are also measured by de-correlating in the ZH and WH processes (referred to as *four POIs*). For the four POI fit, the signal strengths are observed as

$$\mu_{ZH,H\to b\bar{b}} = 0.88^{+0.23}_{-0.20} = 0.88\ ^{+0.15}_{-0.15}\ \text{(statistical)}\ ^{+0.17}_{-0.14}\ \text{(systematic)},$$
$$\mu_{WH,H\to b\bar{b}} = 0.95^{+0.21}_{-0.20} = 0.95\ ^{+0.14}_{-0.14}\ \text{(statistical)}\ ^{+0.16}_{-0.14}\ \text{(systematic)},$$
$$\mu_{ZH,H\to c\bar{c}} = -0.5^{+7.0}_{-6.8} = -0.5\ ^{+5.6}_{-5.5}\ \text{(statistical)}\ ^{+4.1}_{-4.0}\ \text{(systematic)},$$
$$\mu_{WH,H\to c\bar{c}} = 3.3^{+9.2}_{-8.8} = -3.3\ ^{+6.3}_{-6.2}\ \text{(statistical)}\ ^{+6.7}_{-6.2}\ \text{(systematic)},$$

and is in good agreement with the standard model prediction within uncertainties. A 7% correlation between the $ZH, H \to b\bar{b}$ and $ZH, H \to c\bar{c}$ signal strengths, and a -5% correlation between the $WH, H \to b\bar{b}$ and $WH, H \to c\bar{c}$ signal strengths is seen. The results of a fit de-correlating $\mu_{VH,H\to c\bar{c}}$ and $\mu_{VH,H\to b\bar{b}}$ in the charged lepton channels (referred to as *six POIs*) is summarized in Appendix H, together with correlation plots between the signal strengths, and dominant nuisance parameters.

10.3.1 Breakdown of uncertainties

The breakdown of uncertainties on the measurement of the signal strengths of $VH, H \to c\bar{c}$ process is given in Table 10.1. Note that the sum in quadrature of the uncertainties does not add up to the total uncertainty due to correlations among the different uncertainty terms.

Table 10.1: The impact of statistical and systematic uncertainties on the measurement of $\mu_{VH,H\to c\bar{c}}$.

Category	Source of uncertainty	Impact on $\mu_{VH,H\to c\bar{c}}$
Overall	Total uncertainty	5.31
	Statistical uncertainty	3.94
	Systematic uncertainty	3.57
Statistical	Data statistics	3.67
	Top $e\mu$ control region	0.08
	Background floating normalisation	1.24
	$VH, H \to b\bar{b}$ signal strength	0.33
Systematic	MC statistics	1.62
	Luminosity	0.08
	Pile-up	0.23
	Leptons	0.25
	$E_{\mathrm{T}}^{\mathrm{miss}}$	0.23
	Jets	1.02
	Flavor tagging (b-jets)	0.29
	Flavor tagging (c-jets)	0.73
	Flavor tagging (light-flavor jets)	0.66
	Signal modelling	0.72
	Z+jets modelling	1.77
	W+jets modelling	1.42
	$t\bar{t}$ and single-top Wt modelling	1.02
	Single-top s/t modelling	0.16
	Diboson modelling	0.52
	Multijet modelling	0.55

The dominant sources of the systematic uncertainties can be attributed to the modelling of the Z+jets and W+jets processes, since Z+jets and W+jets are the dominant background sources in the $VH, H \to c\bar{c}$ phase space. Further, the uncertainty on the simulated statistics (i.e *MC statistics*) is also relatively dominant, but should be noted that truth flavor tagging discussed in Chapter 6 reduces this uncertainty by 10%.

10.3.2 The upper limit on the $VH, H \to c\bar{c}$ signal strength

The analysis presented in this thesis searched for the $VH, H \to c\bar{c}$ process, but no significant excess of the $VH, H \to c\bar{c}$ signal is observed. Hence, based on the measured value of the $VH, H \to c\bar{c}$ signal strengths, 95% CL upper limits are set for $\mu_{VH,H\to c\bar{c}}$, and the de-correlated signal strengths in the three charged lepton channels. The expected and observed upper limits[1], including the expected limits of the previous publication, are shown in the Figure 10.9.

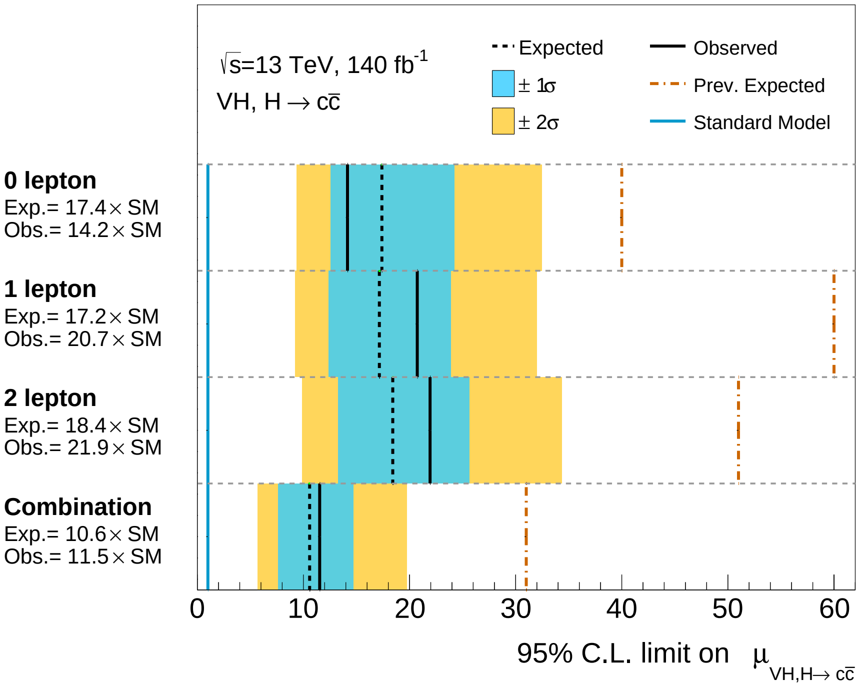


Figure 10.9: The expected and observed 95% confidence level upper limits on the $VH, H \to c\bar{c}$ signal strength. The values for each channel, and well as for the combined value in all channels are shown. Also shown as the values for the previous Run 2 measurement [30], and the standard model (SM) expectation.

The observed upper limit on the signal strength $\mu_{VH,H\to c\bar{c}}$ is 11.5 times the standard model prediction, which is almost a factor of three times improvement on the previous upper limit [30]. The observed upper limits in the zero, one and two charged lepton channels are 14.2, 20.7, and 21.9 times the standard model prediction.

While all three charged lepton channels show a good improvement with respect to the previous result, the improvement in the one charged lepton channel is the largest; a factor of 3.5 improvement in the expected upper limit. This can be attributed to the different strategies implemented in the current analysis in order to increase the sensitivity in the one charged lepton channel; the additional signal region at low p_{T}^{V}, the migration of hadronically decaying τ-lepton candidates to the one charged lepton channel, as well as better constraints on the background modelling uncertainties from the new top and W+heavy flavor control regions.

---

[1] Expected upper limits are obtained through a fit to a theoretically built data representation assuming the standard model, while observed upper limits are obtained through a fit to the actual data [26].

10.4 The interpretation in the $\boldsymbol{\kappa}$-framework

Based on the κ-framework discussed in Section 1.2.4, the signal strength $\mu_{VH,H\to c\bar{c}}$ of Equation 1.17 can be written as

$$\mu_{VH,H\to c\bar{c}} = \left(\frac{\Gamma_c/\Gamma_H}{\Gamma_c^{\mathrm{SM}}/\Gamma_H^{\mathrm{SM}}}\right)\left(\frac{\sigma_{VH}}{\sigma_{VH}^{\mathrm{SM}}}\right) = \frac{\kappa_c^2\kappa_V^2}{\kappa_H^2} = \frac{\kappa_c^2\kappa_V^2}{\Sigma_j B_f^{\mathrm{SM}}\kappa_j^2}, \tag{10.1}$$

where B_{inv} and B_{undect} of Equation 1.14 are set to zero considering only standard model decays. Further, setting all coupling modifiers other than κ_b and κ_c to their standard model prediction (i.e $\kappa = 1$), the coupling signal strength $\mu_{VH,H\to c\bar{c}}$, and similarly $\mu_{VH,H\to b\bar{b}}$, is parameterised in the κ-framework as

$$\mu_{VH,H\to c\bar{c}} = \frac{\kappa_c^2}{1 + B^{\mathrm{SM}}_{H\to b\bar{b}}(\kappa_b^2 - 1) + B^{\mathrm{SM}}_{H\to c\bar{c}}(\kappa_c^2 - 1)}, \tag{10.2}$$

$$\mu_{VH,H\to b\bar{b}} = \frac{\kappa_b^2}{1 + B^{\mathrm{SM}}_{H\to b\bar{b}}(\kappa_b^2 - 1) + B^{\mathrm{SM}}_{H\to c\bar{c}}(\kappa_c^2 - 1)}. \tag{10.3}$$

It should be also noted that the above parametrisation is done assuming the coupling modifiers κ_b and κ_c only affect the Higgs boson decays.

Figure 10.10 shows the results of the simultaneous fit on the coupling modifiers κ_b and κ_c, where a best fit value of $(\kappa_b, \kappa_c) = (0.90, 0.92)$ is observed. Figure 10.10 also shows the 68% CL and 95% CL contours that are extracted from the likelihood scan. By setting $\kappa_b = 1$, the observed (expected) constraints on κ_c at 95% CL is obtained to be $|\kappa_c| < 4.2$ ($|\kappa_c| < 4.1$).

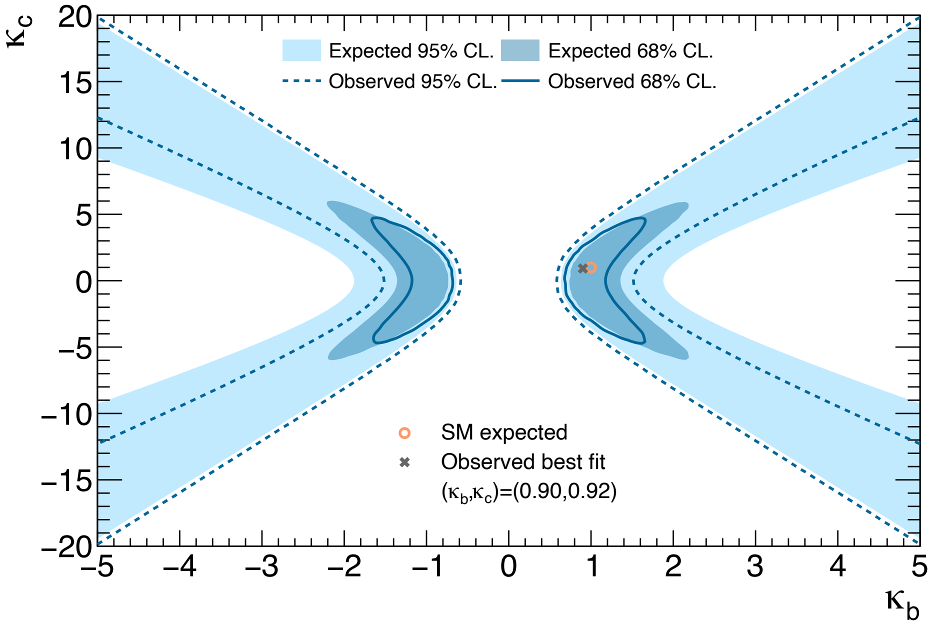


Figure 10.10: The expected (observed) constraints on κ_b and κ_c at 68% (95%) confidence level. The best fit value for (κ_b, κ_c) is (0.90,0.92).

10.4.1 The measurement of κ_c relative to κ_b

It is also possible to parameterise the signal strengths $\mu_{VH,H\to b\bar{b}}$ and $\mu_{VH,H\to c\bar{c}}$ as

$$\mu_{VH,H\to b\bar{b}} = \left(\frac{\kappa_b}{\kappa_H}\right)^2, \tag{10.4}$$

$$\mu_{VH,H\to c\bar{c}} = \left(\frac{\kappa_c}{\kappa_b}\right)^2 \left(\frac{\kappa_b}{\kappa_H}\right)^2, \tag{10.5}$$

which allows the determination of the ratio between the two coupling modifiers κ_b and κ_c. For κ_H in Equations 10.4 and 10.5, no assumptions on the Higgs boson width are taken, allowing $\mu_{VH,H\to b\bar{b}}$ to be a free floating parameter in the fit.

The result for $|\kappa_c/\kappa_b|$ is shown in Figure 10.11. The blue vertical lines represent the scenario where the modified Yukawa coupling strength for Higgs to charm quarks is equal to the Higgs to bottom quark coupling strength, allowing the $|\kappa_c/\kappa_b|$ to be expressed as the ratio between the quark masses as

$$\left|\frac{\kappa_c}{\kappa_b}\right| = \frac{y_b}{y_c} = \frac{m_b}{m_c} = 4.578, \tag{10.6}$$

where the quark masses are evaluated using lattice QCD at a renormalization scale equal to the Higgs boson mass of 125 GeV [19]. The observed 95% CL interval is $|\kappa_c/\kappa_b| < 3.6$, which is smaller than 4.578. This confirms that the Higgs to charm quark coupling strength is weaker than the Higgs to bottom quark coupling strength.

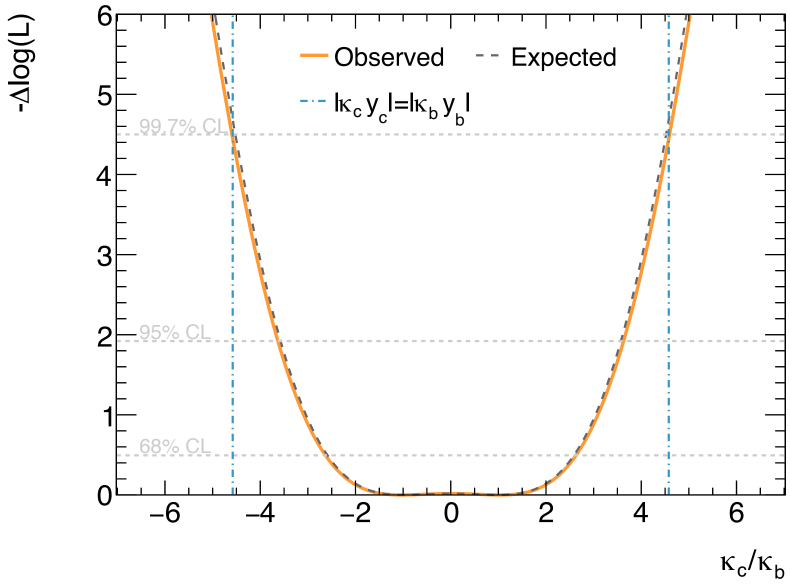


Figure 10.11: The observed and expected values of the combined profile likelihood ratio as a function of the κ_c to κ_b ratio.

Chapter 11

Discussion

Chapter 10 presented the results on the extracted signal strengths $\mu_{VH,H\to c\bar{c}}$ and $\mu_{VH,H\to b\bar{b}}$, and the interpretation of the measured values in the κ-framework. This chapter discusses additional studies done using the combined fit, and prospects on the future $H \to c\bar{c}$ searches.

11.1 Truth flavor tagging in the analysis

As discussed in Section 6.3, the strategy adopted in the analysis is to apply truth flavor tagging to the V+jets and single top s/t channel processes in the BB and C_TC_T regions. Direct flavor tagging is applied to the remaining processes, as well as the remaining regions that are not populated through truth flavor tagging. While this strategy was taken to minimize the mis-modeling bias of truth flavor tagging in the background templates, as a consequence the benefit of truth flavor tagging in reducing the MC statistical uncertainty is not maximized.

11.1.1 Evaluating the bias of truth flavor tagging

The truth flavor tagged templates used in the final fit are obtained following the strategy discussed above, and evaluated with direct flavor tagged templates as discussed in Section 6.3.1. Since the agreement with direct flavor tagged templates are not expected to be perfect, a pseudo-data fit is performed to evaluate the impact on the extracted signal strengths when the truth flavor tagged templates. In the pseudo-data fit, the real data is replaced with a data template which is constructed using the truth flavor tagged templates.

It is important to note that the diboson signal strength $\mu_{VZ,Z\to c\bar{c}}$ is used in this evaluation, since this allows a better evaluation (i.e within smaller uncertainties) on the truth flavor tagging biases on the extracted signal strength. Further, only the MC statistical uncertainties and normalisation factors are added as the nuisance parameters in the fit, to avoid other modelling or experimental nuisance parameters to absorb mis-modelling effects of truth flavor tagging. The obtained signal strength for the combined regions is

$$\mu_{VZ,Z\to c\bar{c}} = 0.989 \pm 0.145,$$

and the signal strengths de-correlated in different charged lepton channels and flavor tag regions is obtained as

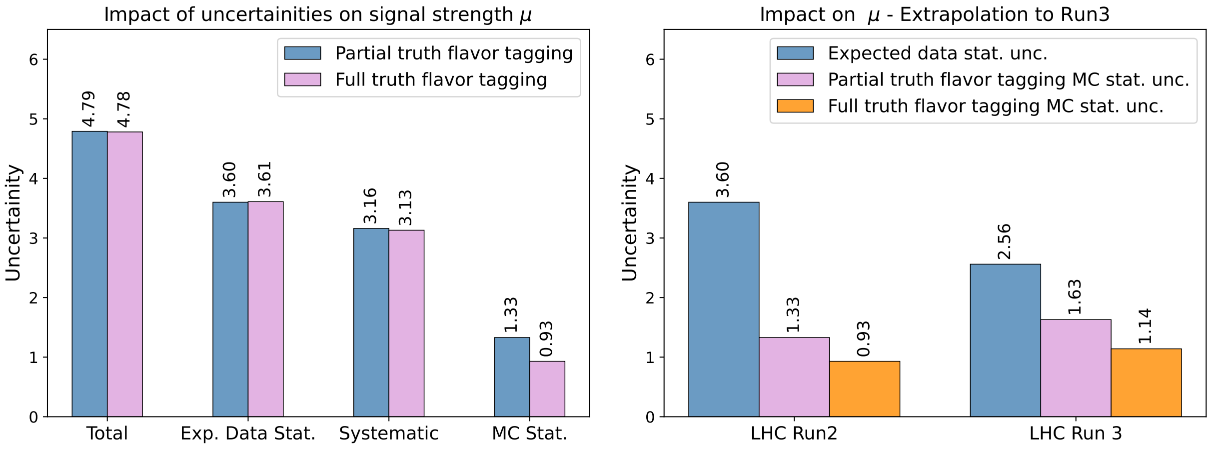
$$\mu_{VZ,Z\to c\bar{c}}(0\text{ Lepton},\ C_TC) = 1.058 \pm 0.223, \qquad \mu_{VZ,Z\to c\bar{c}}(0\text{ Lepton},\ C_TN) = 0.998 \pm 0.443,$$
$$\mu_{VZ,Z\to c\bar{c}}(1\text{ Lepton},\ C_TC) = 0.977 \pm 0.354, \qquad \mu_{VZ,Z\to c\bar{c}}(1\text{ Lepton},\ C_TN) = 0.951 \pm 0.806,$$
$$\mu_{VZ,Z\to c\bar{c}}(2\text{ Lepton},\ C_TC) = 0.898 \pm 0.269, \qquad \mu_{VZ,Z\to c\bar{c}}(2\text{ Lepton},\ C_TN) = 0.973 \pm 0.502.$$

The de-correlation is done to estimate the truth flavor tag biases for the different processes (i.e W+jets are dominant in the one charged lepton channel, and Z+jets are dominant in the two charged lepton channel), and different flavor components (i.e the V+jets mixed flavor is dominant in the C_TN region, while V+jets heavy flavor is dominant in the C_TC region).

All the obtained signal strengths are consistent with one within the uncertainties. After confirming the postfit values of the nuisance parameters are also consistent with the nominal value within the uncertainties, it was concluded that the truth flavor tagged templates used in the final fit does not have a significant bias, within the current data statistical uncertainty.

11.1.2 Prospects of truth flavor tagging in Run 3

The adopted *partial* truth flavor tagging strategy (described above) was compared to the case where all the background templates were made with truth flavor tagging (referred to as *full* truth flavor tagging below), in all the signal regions BB, C_TC_T, C_TC_L, and C_TN. It should be noted that the signal templates are populated with direct flavor tagging in both scenarios. The combined fit is done on a data representation built with the simulated samples (i.e the Asimov templates [26]), and not real data.



(a) Comparison between strategies. (b) Extrapolation to Run 3.

Figure 11.1: The left plot shows the impact on the signal strength μ from the total uncertainty, expected data statistical uncertainty, total systematic uncertainty, and the statistical uncertainty of simulated samples (i.e MC statistical uncertainty). The right plot shows the extrapolation of the impact on the signal strength μ from the expected data statistical uncertainty and the MC statistical uncertainty from Run 2 to Run 3.

The obtained impact on the signal strength μ from the total uncertainty, expected data statistical uncertainty, total systematic uncertainty, and the statistical uncertainty of simulated samples (i.e MC statistical uncertainty), is shown in Figure 11.1(a). The full truth flavor tagging strategy is clearly seen to reduce the MC statistical uncertainty, since it uses all background templates in all signal regions to be truth flavor tagged. However, the impact on the total uncertainty is seen to be negligible even if full truth flavor tagging is used, since the current analysis is limited in the statistical uncertainty on data.

An extrapolation is done from the current Run 2 period result to a similar $VH, H \to c\bar{c}$ analysis during the Run 3 period (2022 to 2026), and shown in Figure 11.1(b). The below assumptions are made when calculating the extrapolated uncertainties.

- The Run 3 period is expected to have roughly *twice* the integrated luminosity of data, compared to Run 2. This scales the expected data statistical uncertainty by $1/\sqrt{2}$.
- The Run 3 period analysis is expected to use a better jet-flavor tagging algorithm, which is expected to reject *twice* more b-jets and light-flavor jets at the same c-jet tagging efficiency[1]. This reduces the amount of MC statistics in the $VH, H \to c\bar{c}$ signal regions to roughly a third of the original sample. Hence, this scales the MC statistical uncertainty by a factor of $\sqrt{3}$.
- For the Run 3 period, it is expected to have *twice* more MC samples compared to the Run 2 period. This scales the MC statistical uncertainty by $1/\sqrt{2}$.

From Figure 11.1(b), it is seen that although the expected data statistical uncertainty is the dominating component in the Run 2 period, it is expected to get lower due to the increase in collected data in the Run 3 period. However, the MC statistical uncertainty is expected to get slightly larger due to the better jet-flavor tagging algorithms. Therefore, it is seen that using the full benefits of truth flavor tagging can be helpful in reducing the MC statistical uncertainty during the Run 3 period. This is expected to be even more important during the HL-LHC when the ATLAS experiment expects to collect roughly 20 times the amount of data of the Run 2 period. However, the use of truth flavor tagging should be accompanied with a good truth flavor tagging framework, which exhibits good modeling with respect to direct flavor tagging, in order to avoid additional uncertainties that would limit the benefits in truth flavor tagging.

11.2 Signal strength extraction from the di-jet mass distribution

One of the main improvements in the current analysis, with respect to the previous $VH, H \to c\bar{c}$ analysis, is the use of a multivariate analysis (MVA) to separate the $VH, H \to c\bar{c}$ signal from the background. The output of this MVA (i.e the BDT_{VH} output) is used to extract the signal strength, while in the previous ATLAS analysis, the di-jet mass ($m_{j_1 j_2}$) was used to extract the signal strength $\mu_{VH,H\to c\bar{c}}$ [30]. As discussed in Section 7.4, this is due to the MVA approach bringing a good improvement in significance when compared to using the $m_{j_1 j_2}$ distribution.

However, the impact of using the $m_{j_1 j_2}$ distribution for extracting the signal strengths was evaluated using the same fit framework, and comparing with the results presented in Section 10.3. Note that the only difference with respect to the nominal fit is the use of $m_{j_1 j_2}$ templates in the signal regions, where the nominal fit used the BDT_{VH} output. The measured signal strength combined (in all channels) is

---

[1] This will also impact the expected data statistical uncertainty, but it is not estimated in this extrapolation. Hence, the shown uncertainty is a conservative number.

$$\mu_{VH,H\to c\bar{c}} = 6.9 \pm 9.0,$$

and the signal strengths de-correlated in the three charged lepton channels is

$$\mu_{VH,H\to c\bar{c}}(0\text{ Lepton}) = -7.3 \pm 12.4, \qquad \mu_{VH,H\to c\bar{c}}(1\text{ Lepton}) = 18.7 \pm 14.8,$$
$$\mu_{VH,H\to c\bar{c}}(2\text{ Lepton}) = 14.2 \pm 16.1.$$

All the measured signal strengths are consistent with the standard model within uncertainties. Based on this, an observed (expected) 95% confidence level upper limit on the signal strength is obtained as 23.6 (17.8) times the standard model prediction for the combined signal strength, and 21.1 (25.9), 45.5 (29.0), and 45.5 (33.0) times the standard model prediction for the signal strengths in the zero, one, and two charged lepton channels, respectively. Figure 11.2 shows the comparison between the upper limits obtained from the MVA discriminant and the $m_{j_1j_2}$ distributions in the current analysis.

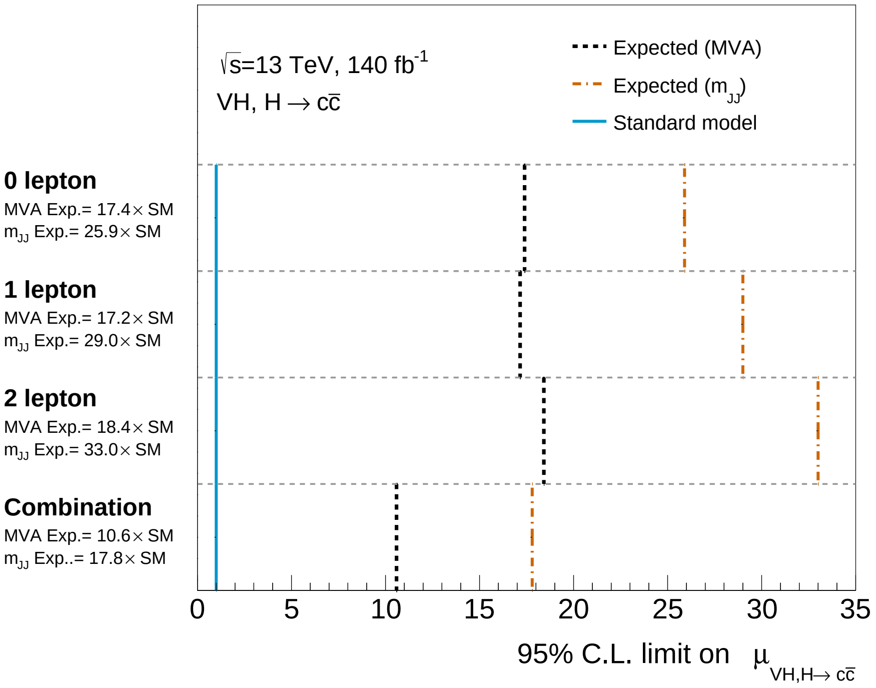


Figure 11.2: The expected and observed 95% confidence level limits on the $VH, H \to c\bar{c}$ signal strength extracted using the MVA discriminant is shown together with the expected 95% confidence level limits extracted using the $m_{j_1j_2}$ distributions. For comparison, the previous ATLAS Run 2 measurement [30] is also shown, which was also based on the measurements from a $m_{j_1j_2}$ distribution.

The combined improvement in the expected limit obtained using the MVA discriminant is 68% compared to the expected limit obtained by fitting the di-jet mass distribution, which highlights the merit of using a MVA in separating the $VH, H \to c\bar{c}$ signal process from the background.

11.3 Comparison with other measurements

As presented in Section 10.3, the observed 95% CL upper limit on the signal strength $\mu_{VH,H\to c\bar{c}}$ is 11.5 times the SM expectation. This is currently the worlds best observed upper limit on the $VH, H \to c\bar{c}$ process; the latest CMS measurement described in [31] achieved an observed upper limit of 14 times the SM expectation. However, the CMS measurement has a better expected upper limit of 7.6 times the SM expectation, a 39% better upper limit than the ATLAS measurement.

A key difference in the CMS measurement and the ATLAS measurement (presented in this thesis) can be attributed to the better sensitivity for the high p_{T}^{H} region ($p_{\mathrm{T}}^{H} > 300$ GeV), where p_{T}^{H} is the transverse momentum of the Higgs boson; in the $p_{\mathrm{T}}^{H} < 300$ GeV region, the expected upper limit is 19.0 times the SM expectation. A large part of this good sensitivity with respect to the $p_{\mathrm{T}}^{H} < 300$ GeV region can be attributed to the use of a Graph Neural Network (GNN) based jet flavor tagging algorithm [133]. The ATLAS experiment did not utilize a dedicated jet flavor tagging algorithm in the high p_{T}^{H} region, but plans to implement a GNN based flavour tagging algorithm that is expected to improve the $VH, H \to c\bar{c}$ sensitivity in the high p_{T}^{H} region in the future [147].

$VH, H \to b\bar{b}/c\bar{c}$ analysis is complementary to the constraints on κ_b and κ_c provided by the ATLAS $H \to ZZ, \gamma\gamma$ analysis, which probed κ_b and κ_c by considering modifications to the p_{T}^{H} distribution at low p_{T}^{H} (as described in Section 1.3.1). Figures 11.3 and 11.4 show the constraints on κ_b and κ_c by the $H \to ZZ, \gamma\gamma$ analysis, overlaid with the $VH, H \to b\bar{b}/c\bar{c}$ analysis result. Comparing with the $H \to ZZ, \gamma\gamma$ *shape only* result shown in Figure 11.3, it is seen that the $VH, H \to b\bar{b}/c\bar{c}$ analysis is able to provide a better constraint on κ_b and κ_c.

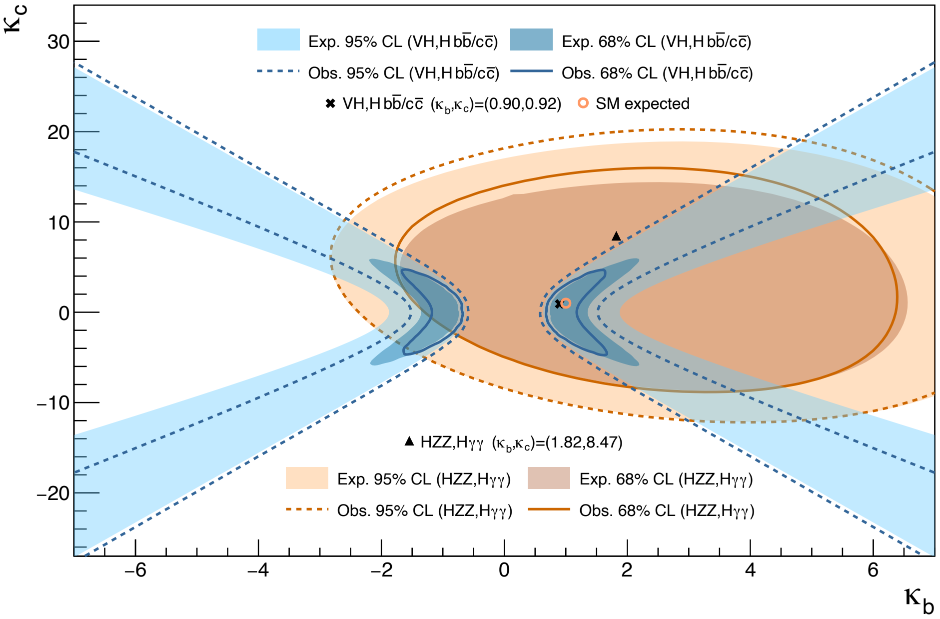


Figure 11.3: The expected (observed) constraints on κ_b and κ_c at 68% (95%) confidence level for the $VH, H \to b\bar{b}/c\bar{c}$ result vs the $H \to ZZ, \gamma\gamma$ shape only result.

This also holds true for the $H \to ZZ, \gamma\gamma$ *shape+norm* result shown in Figure 11.4, where κ_b is constrained around ±1, while κ_c is constrained around zero. Therefore, similar to the combination of the previous ATLAS $VH, H \to c\bar{c}$ measurement with the $H \to ZZ, \gamma\gamma$ measurement on κ_c providing better constraints on κ_c [15], a future combination of the current $VH, H \to c\bar{c}$ measurement with the ATLAS $H \to ZZ, \gamma\gamma$ measurement is expected further constrain κ_c.

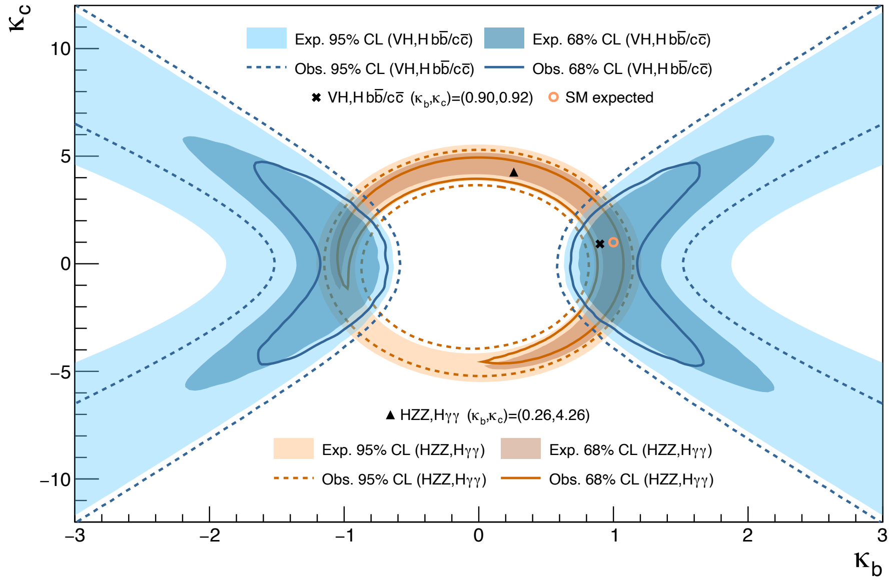


Figure 11.4: The expected (observed) constraints on κ_b and κ_c at 68% (95%) confidence level for the $VH, H \to b\bar{b}/c\bar{c}$ result vs the $H \to ZZ, \gamma\gamma$ shape+norm result.

Finally, the measurement on the 95% CL upper limit on κ_c is added to Figure 1.9, and shown in Figure 11.5. While the latest $VH, H \to b\bar{b}/c\bar{c}$ result by itself does not provide a better constraint on κ_c than the CMS $VH, H \to c\bar{c}$ result [31] or the ATLAS $H \to ZZ, \gamma\gamma$ result [15], it is consistent with the previous measurements.

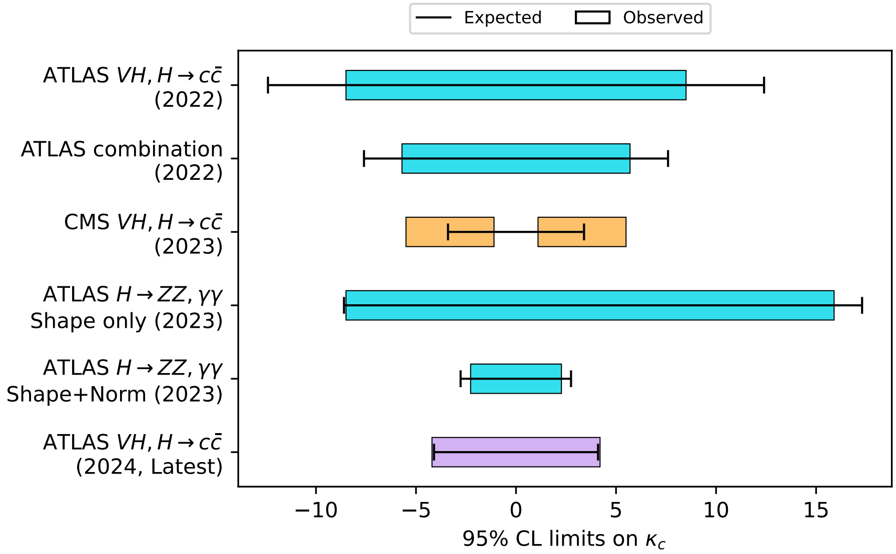


Figure 11.5: A comparison of the constraints on κ_c from obtained from the current analysis, and from several latest measurements. The *Prev. ATLAS $VH, H \to c\bar{c}$*, *CMS $VH, H \to c\bar{c}$*, *ATLAS combination*, and *ATLAS $H \to ZZ, \gamma\gamma$* are the analyses described in [30], [31], [20], and [15].

Conclusion

This thesis presented the results of the latest search on the Higgs boson decay to a $c\bar{c}$ pair, using 140.1 fb^{-1} of proton-proton collision data collected by the ATLAS experiment, at a center of mass energy of $\sqrt{s} = 13$ TeV. The analysis targeted the Higgsstrahlung (i.e VH) production channel of the Higgs boson, and the $VH, H \rightarrow c\bar{c}$ process was probed in three charged lepton final states; $ZH \rightarrow \nu\bar{\nu}c\bar{c}$, $WH \rightarrow l\nu c\bar{c}$, and $ZH \rightarrow l^{-}l^{+}c\bar{c}$. The combined signal strength for the $VH, H \rightarrow c\bar{c}$ was measured to be

$$\mu_{VH,H\rightarrow c\bar{c}} = 1.0^{+5.4}_{-5.2} = 1.0\ ^{+4.0}_{-3.9}\ \text{(statistical)}\ ^{+3.7}_{-3.5}\ \text{(systematic)},$$

which is in good agreement with the standard model prediction. The new measurement corresponds to an observed (expected) 95% confidence level upper limit on the signal strength at 11.5 (10.6) times the standard model prediction. Further, the new expected limit is a factor of three times better than the expected limit of the previous $VH, H \rightarrow c\bar{c}$ analyses, which used the same data set.

Finally, an interpretation in the κ-framework was also presented. The latest measurement obtained an observed (expected) constrain on the Yukawa coupling modifier κ_c as $|\kappa_c| < 4.2\ (4.1)$ at 95% confidence level. Further, the ratio between the coupling modifiers κ_c and κ_b was observed to be 3.6 at 95% confidence level, which is less than the ratio between the bottom and charm quark masses (4.578), confirming that the Yukawa coupling to charm quarks is weaker than that to bottom quarks.

Acknowledgements

First and foremost, I would like to express my gratitude towards my research supervisors from Osaka University in the past seven years: professor emeritus Taku Yamanaka, professor Hajime Nanjo, associate professor Tatsuya Masubuchi, and assistant professor Minoru Hirose. Their support and guidance have been instrumental in completing my dissertation and advancing my career as a researcher. Furthermore, the support I have received from my research group in Osaka University has been also vital in the past seven years. Hence, I would like to extend my gratitude to my colleagues Ryota Shiraishi, Mario Gonzalez, my juniors, and other members of the group for their everlasting support and cheering me through the past few years.

The research work described in this dissertation would not be possible without the immense support I received from the ATLAS collaboration, especially the VH(bb/cc)-legacy analysis team. There are so many people who deserve recognition, and while I cannot name everyone in just one paragraph, I would like to acknowledge the huge support I received from the analysis conveners Hannah Arnold, Yanhui Ma, Elisabeth Schopf, and the FTAG group convener Francesco Armando Di Bello. They have dedicated countless hours to guiding the team, repeatedly setting us in the right direction, and I am deeply grateful for that.

I would also like to acknowledge the Pioneering Quantum Beam Application (PQBA) program, which provided me with several opportunities to expand my horizons as a researcher. I extend my thanks to the staff of the PQBA program who supported us in numerous ways, with special acknowledgement to the immense support I received from professor Takashi Nakano and professor emeritus Tadafumi Kishimoto during the past few years. Finally, I would also like to acknowledge professor Tadayuki Takahashi of the University of Tokyo, my second mentor of the PQBA program, who has given me so many valuable pieces of advice and insights regarding research opportunities in physics.

My appreciation also goes towards the Ministry of Education, Culture, Sports, Science, and Technology-Japan (MEXT) for funding my education during my undergraduate and master's programs, and the Sato Yo International Scholarship Foundation (SISF) for funding my education during the doctoral program. I deeply appreciate not only the financial support provided to me by SISF, but also for constantly cheering me on through this journey.

Last but not least, I am deeply grateful to my parents, my two brothers, and my wife for their unwavering support and encouragement in the past decade of my higher education. Thank you for your endless patience, love, understanding, and believing in me when I myself could not see the next steps. This dissertation is as much a result of your support as it is my own effort.

Appendices

A The standard model of particle physics

This section is compiled by referring to [2] and [3].

A.1 Symmetries and gauge theories

Symmetries play an important role in the formulation of the standard model. In quantum mechanics, symmetries can be expressed by requiring all physical predictions of a theory to be invariant under the wavefunction transformation,

$$\psi \rightarrow \psi' = \hat{U}\psi.$$

From the requirement that all physical predictions are unchanged, the wavefunction normalization before and after applying symmetry transformations should be unchanged, $\langle\psi|\psi\rangle = \langle\psi'|\psi'\rangle = \langle\hat{U}\psi|\hat{U}\psi\rangle = \langle\psi|\hat{U}^\dagger\hat{U}|\psi\rangle$. Therefore, any operator corresponding to symmetry transformations should be unitary; $\hat{U}^\dagger\hat{U} = I$. A continuous symmetry operation can be built from a series of infinitesimal operations of the form,

$$\hat{U}(\epsilon) = I + i\epsilon\hat{G}, \tag{A.1}$$

where ϵ is an infinitesimally small parameter, and $\hat{G}$ is the generator of the transformation. Since U is unitary, the term $i\epsilon(\hat{G} - \hat{G}^\dagger) = 0$ for $\hat{U}^\dagger\hat{U} = I$. Hence, $\hat{G} = \hat{G}^\dagger$ and the generator is Hermitian; the eigenstates of the generator are real, and $\hat{G}$ is associated with an observable quantity.

The unitary operator $\hat{U}$ commutes with the hamiltonian $\hat{H}$, and hence $\left[\hat{H}, \hat{G}\right] = 0$. In quantum mechanics, the time evolution of an operator is given by $\frac{d}{dt}\langle\hat{G}\rangle = i\left\langle\left[\hat{H}, \hat{G}\right]\right\rangle$, and hence,

$$\frac{d}{dt}\langle\hat{G}\rangle = 0.$$

Therefore, the associated observable quantity G is conserved. Further generalizing, the unitary operator in Equation A.1 can be written as $\hat{U} = 1 + i\boldsymbol{\epsilon}\cdot\hat{\mathbf{G}}$, where $\hat{\mathbf{G}} = \{\hat{G}_i\}$ represents a set of Hermitian generators. Then this finite transformation can be expressed as a series of infinitesimal transformations using,

$$\hat{U}(\alpha) = \lim_{n\to\infty}\left(1 + i\frac{1}{n}\boldsymbol{\alpha}\cdot\hat{\mathbf{G}}\right)^n = e^{i\boldsymbol{\alpha}\cdot\hat{\mathbf{G}}}.$$

In physics, the symmetries of a theory can be categorized into two types; global symmetry and local symmetry. Global symmetries are associated with conserved quantities and can be explained using Noether's theorem[2]. Taking U(1) transformation as an example, where $\hat{U} = Ie^{i\theta}$ corresponds to a multiplication by a complex phase,

$$\psi'(x) = \hat{U}\psi(x) = e^{i\theta}\psi(x), \tag{A.2}$$

---

[2] A symmetry of a Lagrangian is associated with a conserved current.

where θ is constant in all spacetime represents a global symmetry. Local symmetry has a stronger constraint in requiring invariance with transformations that can vary from point to point in space-time,

$$\psi'(x) = e^{i\theta(x)}\psi(x). \tag{A.3}$$

A gauge theory is a type of field theory where the Lagrangian of the system is invariant under *local transformations* under the certain group of symmetry operations. As explained in section A.2 and section A.3, to ensure local symmetry it is essential to include interactions between fermion and boson fields in the Lagrangian.

In the standard model, the generators of the U(1) symmetry group are abelian, meaning they obey the commutative law of quantum mechanics, $\hat{G}_i\hat{G}_j = \hat{G}_j\hat{G}_i$. Therefore, quantum electrodynamics is an abelian gauge theory. However, the generators of the SU(2) and SU(3) symmetry groups do not obey the commutative law, and hence the theory of weak interaction and quantum chromodynamics are termed non-abelian gauge theories[3]. In non-Abelian gauge theories, the transformation properties of gauge fields are not independent, and additional gauge boson self-interaction terms need to be included for gauge invariance.

A.2 Quantum electrodynamics

The gauge symmetry under group U(1) in quantum electrodynamics refers to the invariance of the theory[4] if the fermion and gauge fields are multiplied by a complex phase $e^{i\theta(x)}$.

Under local gauge transformations, the Dirac Lagrangian for a free particle can be written (applying Equation A.3) as,

$$\begin{aligned}
\mathcal{L} \rightarrow \mathcal{L}' &= \bar{\psi}'(x)(i\gamma^\mu\partial_\mu - m)\psi'(x) \\
&= e^{-i\theta(x)}\bar{\psi}'(x)(i\gamma\partial_\mu - m)e^{i\theta(x)}\psi(x) \\
&= \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x) + \bar{\psi}(x)\gamma^\mu\psi(x)\partial_\mu\theta(x) \\
&= \mathcal{L} + j^\mu(x)\partial_\mu\theta(x),
\end{aligned}$$

where $j^\mu(x) = \bar{\psi}(x)\gamma^\mu\psi(x)$ is the vector current carried by the fermion. Hence, due to the extra term $j^\mu(x)\partial_\mu\theta(x)$, the free-particle Dirac lagrangian does not exhibit local invariance. Local invariance can be achieved by the introduction of the electromagnetic field A_μ as,

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu - iqA_\mu(x). \tag{A.4}$$

Here D_μ is defined[5] as the covariant derivative. With the introduction of A_μ, the Lagrangian becomes $\mathcal{L} - iqA_\mu(x)$, and defining the Lagrangian by replacing ∂_μ by D_μ becomes,

---

[3] The mathematical framework needed for describing non-abelian gauge theories is known as the Yang-Mills theory [148].

[4] Although the overall phase of the wave-functions are altered, the wavefunction normalization before and after applying symmetry transformations should be unchanged.

[5] Here $A_\mu = (\phi, -\mathbf{A})$ and $\partial = (\partial_0, \boldsymbol{\nabla})$

$$\mathcal{L}(\psi, A_\mu) = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + q\bar{\psi}\gamma^\mu \psi A_\mu. \tag{A.5}$$

The Lagrangian in Equation A.5 becomes invariant for the substitution,

$$A_\mu \rightarrow A'_\mu = A_\mu - \frac{1}{q}\partial_\mu \theta(x).$$

Hence, the new Lagrangian $\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi$ is invariant under the local gauge transformations. Therefore, QED is made invariant by introducing the spin-1 vector (gauge) boson field A_μ, where its quanta is called the photon.

To arrive at the complete QED Lagrangian, it is important to consider the photon mass term, and the kinetic energy term of the vector field. However, the photon mass term of the form $\frac{m^2}{2}A_\mu A^\mu$ violates the local gauge invariance[6], and the photon mass is set to zero. The Lagrangian for the free electromagnetic field is,

$$\mathcal{L} = \frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \frac{1}{4}(\partial_\nu A_\mu - \partial_\mu A_\nu)(\partial^\nu A^\mu - \partial^\mu A^\nu).$$

Hence, the complete gauge invariant QED Lagrangian for fermion and photon fields take the form,

$$\boxed{\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}.} \tag{A.6}$$

A.3 Weak interaction and the electroweak theory

The weak interaction consists of the charged-current weak interaction, and the neutral current weak interaction. The charged current weak interaction is associated with the invariance under the SU(2) local symmetry transformation,

$$\psi(x) \rightarrow \psi'(x) = e^{ig_W \boldsymbol{\alpha}(x)\cdot \mathbf{T}}\psi(x) \tag{A.7}$$

where $\mathbf{T} = \frac{1}{2}\boldsymbol{\sigma}$ denotes the three generators of the SU(2) group written in terms of the Pauli spin matrices. Similar to QED, applying this symmetry transformation on a free-particle Lagrangian does not satisfy the local gauge invariance. The local gauge symmetry is only conserved by the introduction of three gauge fields W^k_μ (where $k = 1, 2, 3$) related to three gauge bosons $\mathrm{W}^{(1)}$, $\mathrm{W}^{(2)}$ and $\mathrm{W}^{(3)}$.

Since the generators of the SU(2) transformation are 2×2 Pauli spin-matrices, the wavefunction $\psi(x)$ must be written in terms of two components,

$$\psi(x) = \begin{pmatrix} \psi^0(x) \\ \psi^-(x) \end{pmatrix}, \tag{A.8}$$

---

[6] Because $A^\mu A_\mu \rightarrow (A^\mu - \partial^\mu\theta)(A_\mu - \partial_\mu\theta) \neq A^\mu A_\mu$.

and is known as a weak isospin doublet with a total weak isospin of $I_W = \frac{1}{2}$. Since the weak charged-current couples to only left-handed chiral (LH) particle states and right-handed anti-chiral (RH) particle states, the gauge transformation of Equation A.7 can only affect LH particles and RH anti-particles. Hence, the symmetry group for the weak interaction is denoted as $\mathrm{SU}(2)_\mathrm{L}$. To achieve this transformation, the LH anti-particles and RH particles[7] are placed in a weak isospin singlet with $I_W = I_W^{(3)} = 0$.

$$e^-_R, \quad \mu^-_R, \quad \tau^-_R, \quad u_R, \quad d_R, \quad c_R, \quad s_R, \quad t_R, \quad b_R.$$

The weak isospin doublets are constructed from the weak eigenstates, with the upper state having $I_W^{(3)} = +\frac{1}{2}$ and differs by plus one unit charge compared to the bottom state (having $I_W^{(3)} = -\frac{1}{2}$),

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d' \end{pmatrix}_L, \quad \begin{pmatrix} c \\ s' \end{pmatrix}_L, \quad \begin{pmatrix} t \\ b' \end{pmatrix}_L.$$

Since weak isospin singlets are unaffected by the $\mathrm{SU}(2)_\mathrm{L}$ gauge transformation, they do not couple to the gauge bosons of the $\mathrm{SU}(2)_\mathrm{L}$ gauge symmetry. Similar to the introduction of the electromagnetic field in Equation A.4, the requirement of local gauge invariance for the weak interaction requires the introduction of three new gauge fields W^k as $\mathbf{W}(x)$,

$$\partial_\mu \to D_\mu \equiv \partial_\mu + i g_W \mathbf{T} \cdot \mathbf{W}_\mu(x). \tag{A.9}$$

The physical W bosons can be identified as linear combinations of $\mathrm{W}^{(1)}$ and $\mathrm{W}^{(2)}$,

$$\boxed{\mathrm{W}^\pm_\mu = \frac{1}{\sqrt{2}} \left(\mathrm{W}^{(1)}_\mu \mp i \mathrm{W}^{(2)}_\mu \right)}\,. \tag{A.10}$$

The component $\mathrm{W}^{(3)}$ is associated with the two neutral bosons in the electroweak theory.

Electroweak theory

The electroweak theory by Glashow, Salam and Weinberg describes the unification of the electromagnetic and weak interactions to the single electroweak interaction. In the electroweak theory, the U(1) symmetry group of QED is replaced with a new $\mathrm{U}(1)_\mathrm{Y}$ symmetry group such that,

$$\psi(x) \to \psi'(x) = \hat{U}(x)\psi(x) = e^{i g' \frac{Y}{2} \zeta(x)} \psi(x). \tag{A.11}$$

Unifying the weak interaction with QED, the symmetry group of the electroweak theory becomes $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$. To achieve local invariance under the $\mathrm{U}(1)_\mathrm{Y}$ symmetry group, the electromagnetic field A_μ is replaced by the new gauge field B_μ that couples to a new type of charge;

---

[7] Neutrinos are assumed to be massless, and thus does not have the right-handed chiral particle states; to avoid the mass term $\frac{1}{2} m \nu \nu$ in the Lagrangian.

the weak hypercharge $Y = 2\left(Q - I_{\mathrm{W}}^{(3)}\right)$, which is a linear combination of the electromagnetic charge Q and the third component of the weak isospin $W_\mu^{(3)}$. Then the photon field A_μ and the Z boson field Z_μ are expressed by the linear combination of B_μ and the neutral $W_\mu^{(3)}$ field of the weak interaction,

$$\boxed{\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_{\mathrm{W}} & \sin\theta_{\mathrm{W}} \\ -\sin\theta_{\mathrm{W}} & \cos\theta_{\mathrm{W}} \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^{(3)} \end{pmatrix}} \tag{A.12}$$

where θ_{W} is the weak mixing angle.

To achieve local gauge symmetry under the $\mathrm{SU}(2)_{\mathrm{L}} \times \mathrm{U}(1)_{\mathrm{Y}}$ symmetry group, Equation A.9 can be replaced by the appropriate covariant derivative,

$$D_\mu = \partial_\mu + ig_W \mathbf{T} \cdot \mathbf{W}_\mu + ig' \frac{Y}{2} B_\mu. \tag{A.13}$$

Similar to QED, the weak interaction also requires the gauge bosons of the symmetry group to be massless in order to preserve the symmetry of the Lagrangian, since a gauge boson mass term of the form $m^2 W_\mu W^\mu$ violates local gauge symmetry. However, this contradicts with experimental evidence where the weak bosons are observed to have mass. The masses of the weak bosons are obtained in the electroweak theory by the spontaneous symmetry breaking of the $\mathrm{SU}(2)_{\mathrm{L}} \times \mathrm{U}(1)_{\mathrm{Y}}$ local gauge symmetry through the Higgs mechanism.

A.4 The Higgs mechanism

The Higgs mechanism is the spontaneous symmetry breaking of the $\mathrm{SU}(2)_{\mathrm{L}} \times \mathrm{U}(1)_{\mathrm{Y}}$ local gauge symmetry of the electroweak theory, generating the mass of the weak bosons. After symmetry breaking, the electromagnetic symmetry should be left as the true symmetry at low energy. Hence, the gauge theory should demonstrate the behavior,

$$\mathrm{SU}(2)_{\mathrm{L}} \times \mathrm{U}(1)_{\mathrm{Y}} \xrightarrow[\text{symmetry breaking}]{\text{spontaneous}} \mathrm{U}(1)_{\mathrm{QED}}$$

The minimal Higgs model contains two complex scalar fields, placed in a weak isospin doublet [8]; one of the scalar fields must be neutral (ϕ^0) in order to generate the masses of the electroweak gauge bosons, while the other scalar field must be charged (ϕ^+) to give the longitudinal degrees of freedom of the W^+ and W^- bosons.

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{A.14}$$

The Lagrangian for this doublet is,

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi), \tag{A.15}$$

---

[8] The upper and lower components of the doublet differ by one unit of charge.

where $V(\phi)$ is the Higgs potential given by,

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2. \tag{A.16}$$

For the potential to have a finite minimum, λ should be positive. If $\mu^2 \geq 0$, the potential given in Equation A.16 will have a minimum when both fields are zero; $\phi^+ = 0$ and $\phi^0 = 0$. The corresponding Lagrangian of Equation A.15 represents a scalar particle with mass μ and a four point self-interaction term proportional to ϕ^4.

When $\mu^2 < 0$, the Higgs potential has an infinite set of minima defined by,

$$\phi^\dagger \phi = (\phi^0)^2 + (\phi^+)^2 = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}.$$

After symmetry breaking, the neutral photon is required to remain massless. Therefore the minimum of the Higgs potential must correspond to a non-zero vacuum expectation value only of the neutral scalar field ϕ^0,

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}.$$

Expanding the field ϕ_3 about the minimum as $\phi_3(x) = v + \eta(x)$,

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ v + \eta(x) + i\phi_4 \end{pmatrix}$$

Hence, after spontaneous symmetry breaking there will be a massive scalar and three massless Goldstone bosons giving the longitudinal degrees of freedom of $W^\pm$ and Z bosons. The Higgs doublet can be written in the unitary gauge will correspond to choosing the complex scalar field to be entirely real. Hence,

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{A.17}$$

The mass terms and interaction terms of the gauge bosons can be identified by replacing the derivatives in Lagrangian of Equation A.15 with appropriate covariant derivatives, such that it respects the SU$(2)_L\times$ U$(1)_Y$ local gauge symmetry of the electroweak model.

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ig_W \mathbf{T} \cdot \mathbf{W}_\mu + ig' \frac{Y}{2} B_\mu, \tag{A.18}$$

where D_μ is a 2×2 matrix acting on the two components of the weak isospin doublet, and $\mathbf{T} = \frac{1}{2}\sigma$ are the three generators of SU(2) symmetry. The weak hypercharge is $Y = 2\left(Q - I_W^{(3)}\right)$, and $I_W^{(3)} = -\frac{1}{2}$ for the lower component of the Higgs doublet to be neutral, and therefore, the Higgs doublet has a hypercharge of $Y = 1$. Applying the covariant derivative D_μ to the Higgs doublet ϕ in the unitarity gauge,

$$\begin{aligned}
D_\mu \phi &= \frac{1}{2}\left[2\partial_\mu + \left(i g_W \boldsymbol{\sigma}\cdot\mathbf{W}_\mu + i g' B_\mu\right)\right]\phi \\
&= \frac{1}{2\sqrt{2}}\begin{pmatrix} 2\partial_\mu + i g_W W_\mu^{(3)} + i g' B_\mu & i g_W\left[W_\mu^{(1)} - i W_\mu^{(2)}\right] \\ i g_W\left[W_\mu^{(1)} + i W_\mu^{(2)}\right] & 2\partial_\mu - i g_W W_\mu^{(3)} + i g' B_\mu \end{pmatrix}\begin{pmatrix} 0 \\ v+h \end{pmatrix} \\
&= \frac{1}{2\sqrt{2}}\begin{pmatrix} i g_W\left[W_\mu^{(1)} - i W_\mu^{(2)}\right](v+h) \\ \left(2\partial_\mu - i g_W W_\mu^{(3)} + i g' B_\mu\right)(v+h) \end{pmatrix}
\end{aligned}$$

The term in the Lagrangian which generates the masses of the gauge bosons is $(D_\mu\phi)^\dagger(D^\mu\phi)$.

$$\begin{aligned}
(D_\mu\phi)^\dagger(D^\mu\phi) = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) &+ \frac{1}{8}g_W^2\left(W_\mu^{(1)} + i W_\mu^{(2)}\right)\left(W^{(1)\mu} - i W^{(2)\mu}\right)(v+h)^2 \\
&+ \frac{1}{8}\left(g_W W_\mu^{(3)} - g' B_\mu\right)\left(g_W W^{(3)\mu} - g' B^\mu\right)(v+h)^2 .
\end{aligned} \tag{A.19}$$

The gauge boson masses are determined from the terms that are quadratic in gauge boson fields in Equation A.19,

$$\frac{1}{8}v^2 g_W^2\left(W_\mu^{(1)}W^{(1)\mu} + W_\mu^{(2)}W^{(2)\mu}\right) + \frac{1}{8}v^2\left(g_W W_\mu^{(3)} - g' B_\mu\right)\left(g_W W^{(3)\mu} - g' B^\mu\right). \tag{A.20}$$

In the first term of Equation A.20, the mass terms of the spin-1 fields $W_\mu^{(1)}$ and $W_\mu^{(2)}$ appear as $\frac{1}{2}m_W^2 W_\mu W^\mu$. Therefore the mass of the W boson is obtained as,

$$\boxed{m_W = \frac{1}{2}g_W v.} \tag{A.21}$$

From Equation A.21 it is seen that the W boson mass is determined by the coupling constant g_W of the SU$(2)_L$ gauge interaction and the vacuum expectation value v of the Higgs field. Next, the second term of Equation A.20 can be written as,

$$\begin{aligned}
\frac{1}{8}v^2\left(g_W W_\mu^{(3)} - g' B_\mu\right)\left(g_W W^{(3)\mu} - g' B^\mu\right) &= \frac{v^2}{8}\begin{pmatrix} W_\mu^{(3)} & B_\mu \end{pmatrix}\begin{pmatrix} g_W^2 & -g_W g' \\ -g_W g' & g'^2 \end{pmatrix}\begin{pmatrix} W^{(3)\mu} \\ B^\mu \end{pmatrix} \\
&= \frac{v^2}{8}\begin{pmatrix} W_\mu^{(3)} & B_\mu \end{pmatrix}\mathbf{M}\begin{pmatrix} W^{(3)\mu} \\ B^\mu \end{pmatrix},
\end{aligned}$$

where $\mathbf{M}$ is the mass matrix. The off diagonal elements in $\mathbf{M}$ couple with the $W^{(3)}$ and B fields, allowing them to mix. The masses of the physical gauge bosons are given by the eigenvalues of $\mathbf{M}$,

$$\det(\mathbf{M} - \lambda I) = 0 \quad \rightarrow \quad (g_W^2 - \lambda)(g'^2 - \lambda) - g_W^2 g'^2 = 0.$$

Solving the above Equation gives,

$$\lambda = 0 \qquad \text{or} \qquad \lambda = g_W^2 + g'^2$$

Hence, in the diagonal basis, the terms in the Lagrangian which represent the masses of the physical gauge boson fields A_μ and Z_μ are,

$$\frac{1}{2}\begin{pmatrix} A_\mu & Z_\mu \end{pmatrix}\begin{pmatrix} m_A^2 & 0 \\ 0 & m_Z^2 \end{pmatrix}\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \frac{1}{8}v^2 \begin{pmatrix} A_\mu & Z_\mu \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & g_W^2 + g'^2 \end{pmatrix}\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix},$$

where A_μ and Z_μ are the physical fields corresponding to the eigenvectors of $\mathbf{M}$. From the above equation, the masses of the physical gauge bosons can be identified as,

$$m_A = 0 \qquad \text{and} \qquad m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2}. \tag{A.22}$$

The first term in A.22 corresponds to a massless neutral gauge boson that can be identified as the photon. The second term corresponds to a massive neutral gauge boson which can be identified as the Z boson. Hence, it is understood that the neutral Goldstone boson of the broken symmetry has acquired mass through the Higgs mechanism, and the field corresponding to the photon remains massless.

A.5 The Higgs boson coupling to bosons and fermions

The Higgs mechanism, the spontaneous symmetry breaking of the gauge symmetry, explains how the vector bosons obtain their mass. Although not predicted in the theory, the Higgs mechanism can also be extended to explain the mechanism that generates the mass of fermions. In the standard model, while there is no direct coupling between the massless bosons (photons and gluons), there are three types of couplings between the Higgs boson and the standard model particles; the gauge coupling to the vector bosons, the Yukawa coupling to fermions, and the self coupling of the Higgs boson to itself.

The gauge coupling to vector bosons

In Equation A.19, the gauge boson fields appear in the form $VV(V+h)^2$, where $V = W^\pm, Z$. As shown in section A.4, the term VVv^2 determines the masses of the gauge bosons. The terms VVh and $VVhh$ represent the triple and quartic couplings between one or two Higgs bosons with gauge bosons.

The W^+ and W^- fields are expressed as linear combinations of $W^{(1)}$ and $W^{(2)}$ fields,

$$W^\pm = \frac{1}{\sqrt{2}}\left(W^{(1)} \mp iW^{(2)}\right).$$

Therefore, the second term of Equation A.19 can be written in terms of the the physical W^+ and W^- fields,

$$\begin{aligned}\frac{1}{8}g_W^2\left(W_\mu^{(1)}+iW_\mu^{(2)}\right)\left(W^{(1)\mu}-iW^{(2)\mu}\right)(v+h)^2 &= \frac{1}{4}g_W^2 W_\mu^- W^{+\mu}(v+h)^2 \\ &= \frac{1}{4}g_W^2 v^2 W_\mu^- W^{+\mu}+\frac{1}{2}g_W^2 v W_\mu^- W^{+\mu}h+\frac{1}{4}g_W^2 W_\mu^- W^{+\mu}hh\end{aligned}$$

The first term gives the masses of the physical $W^{\pm}$ fields. The hW^+W^- and hhW^+W^- gives rise to triple and quartic couplings of the Higgs boson to W bosons. The coupling strength at the $h-W$ vertex is therefore,

$$g_{HWW}=\frac{1}{2}g_W^2 v \equiv g_W m_W, \tag{A.23}$$

and hence, the coupling of the Higgs boson to the W-boson is proportional to the W-boson mass. Similarly it can be obtained,

$$g_{HZZ}\equiv g_Z m_Z. \tag{A.24}$$

The Yukawa coupling to fermions

Due to different transformation properties for the left and right-handed chiral states, the fermion mass term[9] in the Dirac Lagrangian (Equation A.25) does not respect the $\mathrm{SU}(2)_L\times\mathrm{U}(1)_\gamma$ gauge symmetry, and cannot be present in the standard model Lagrangian.

$$-m\bar{\psi}\psi=-m\left(\bar{\psi}_R\psi_L+\bar{\psi}_L\psi_R\right). \tag{A.25}$$

An infinitesimal SU(2) local gauge transformation on the SU(2) doublet $\phi(x)$ of Equation A.14 will have the effect $\phi\to\phi'=(I+ig_W\boldsymbol{\epsilon}(x)\cdot\mathbf{T})\phi$. The same local gauge transformations apply to the left handed doublet fermion fields as $\bar{L}\to\bar{L}'=\bar{L}(I-ig_W\boldsymbol{\epsilon}(x)\cdot\mathbf{T})$, where $\bar{L}\equiv L^\dagger\gamma^0$. Hence, the combination $\bar{L}\phi R$ and it's Hermitian conjugate $(\bar{L}\phi R)^\dagger=\bar{R}\phi^\dagger L$, of the form,

$$-g_f\left(\bar{L}\phi R+\bar{R}\phi^\dagger L\right), \tag{A.26}$$

satisfies the $\mathrm{SU}(2)_L\times\ \mathrm{U}(1)_Y$ gauge symmetry for charged leptons and down type quarks. For up type quarks, a similar term can be constructed using the conjugate doublet ϕ_c,

$$-g_f\left(\bar{L}\phi_c R+\bar{R}\phi_c^\dagger L\right). \tag{A.27}$$

Here g_f is known as the **Yukawa coupling** of the fermion to the Higgs field. After spontaneous symmetry breaking, the Higgs doublet in unitary gauge of Equation A.17 can be applied to the Lagrangian term to obtain,

---

[9] Left handed chiral fermions placed in SU(2) doublets are denoted as L, and right-handed chiral fermions placed in SU(2) singlets are denoted as R

$$\mathcal{L} = -\frac{g_f}{\sqrt{2}} v \left(\bar{f}_L f_R + \bar{f}_R f_L \right) - \frac{g_f}{\sqrt{2}} h \left(\bar{f}_L f_R + \bar{f}_R f_L \right), \tag{A.28}$$

where the first term represents the fermion mass terms, which are now introduced in a gauge invariant manner. Hence, the Yukawa couplings can be written using the mass of the fermion m_f and the vacuum expectation value v as,

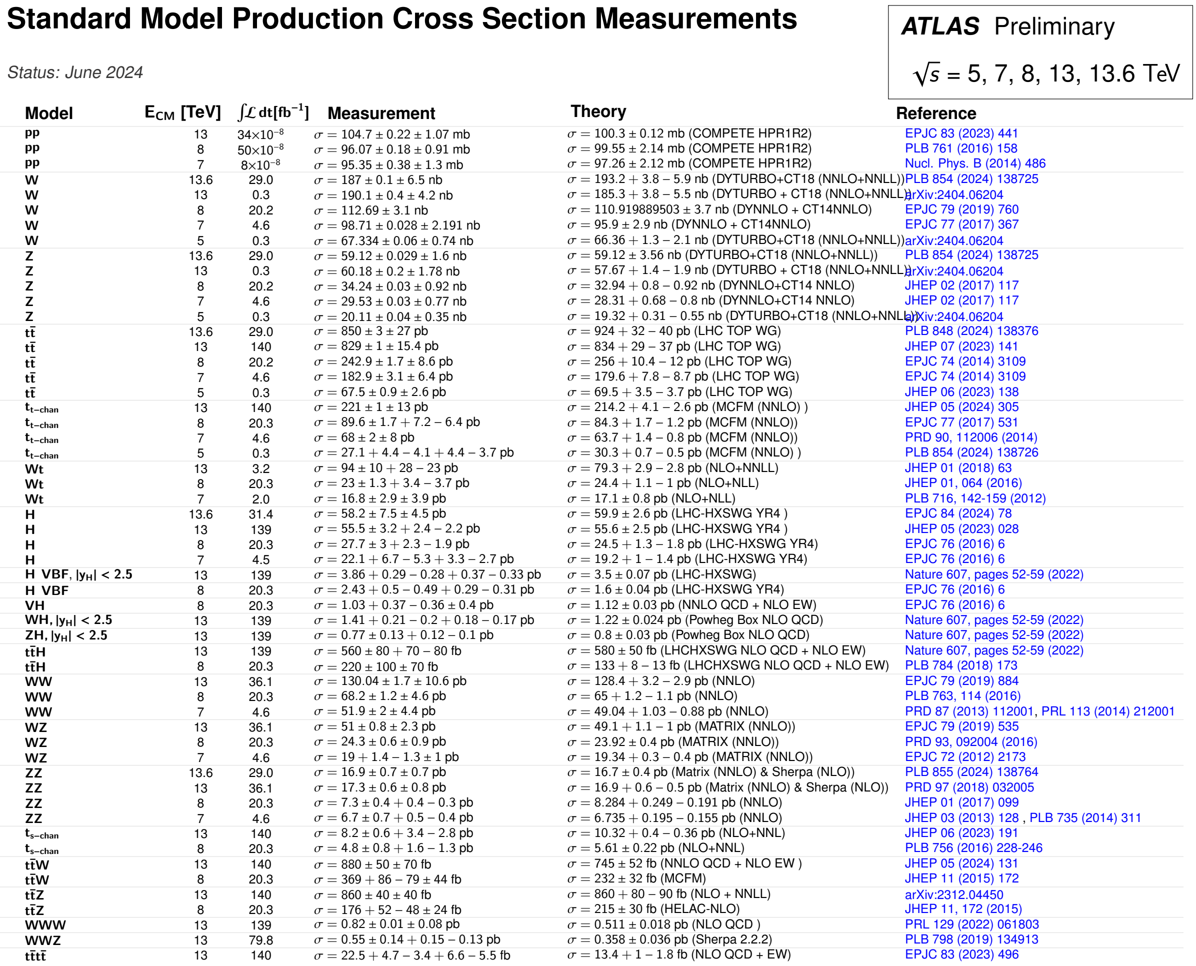
$$g_f = \sqrt{2} \frac{m_f}{v}. \tag{A.29}$$

B Standard model physics at the LHC

This section details supplementary information to Chapter 1 regarding the standard model physics at the large hadron collider.

B.1 Different physics processes of proton-proton collisions

The table in Figure B.1 lists the measured and theoretical cross-sections for different standard model processes in proton-proton at the LHC. The shown cross section values are inclusive of all the sub-processes.



Standard Model Production Cross Section Measurements

Status: June 2024

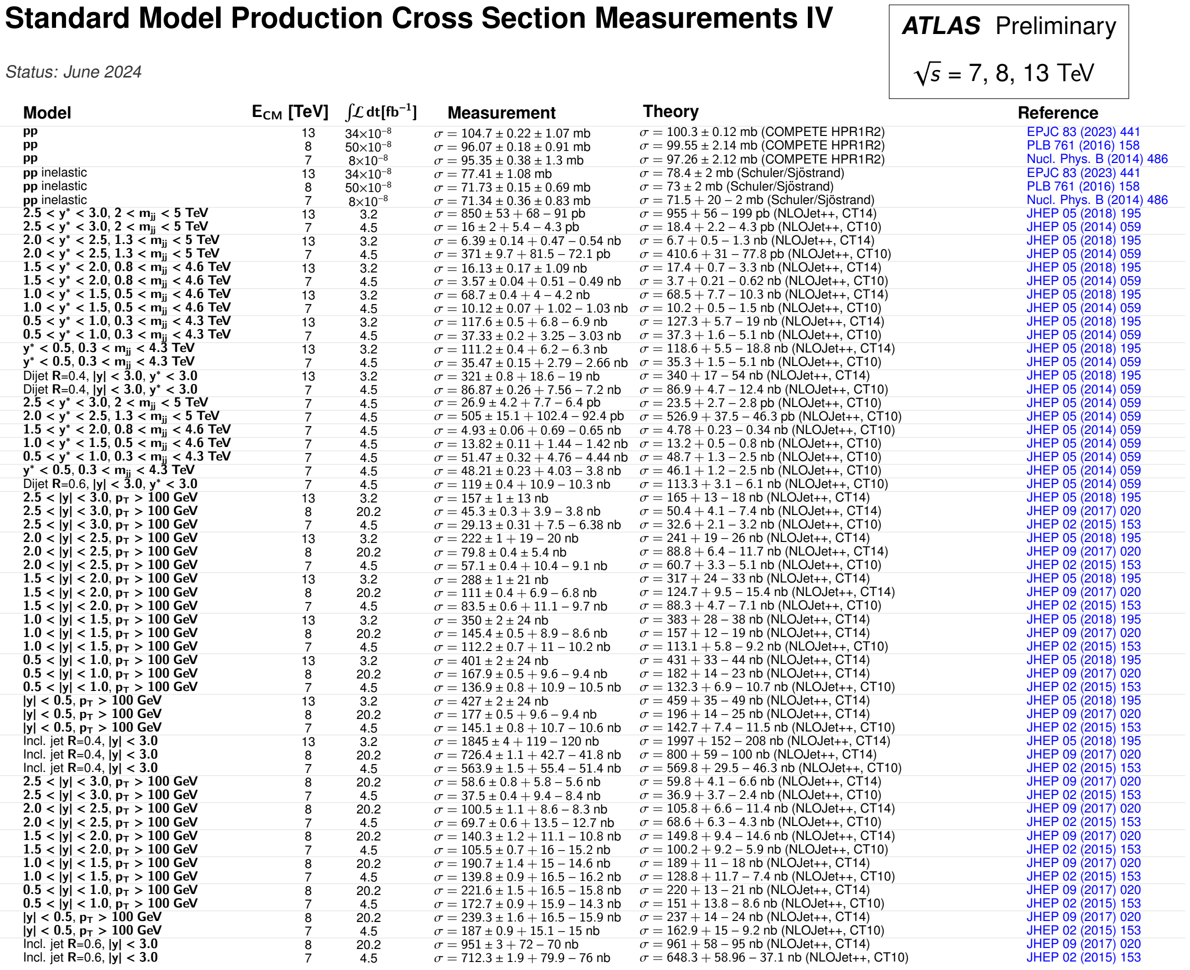
ATLAS Preliminary

$\sqrt{s}$ = 5, 7, 8, 13, 13.6 TeV

Model	E_{CM} [TeV]	$\int\mathcal{L}\,dt[fb^{-1}]$	Measurement	Theory	Reference
pp	13	34×10^{-8}	$\sigma = 104.7 \pm 0.22 \pm 1.07$ mb	$\sigma = 100.3 \pm 0.12$ mb (COMPETE HPR1R2)	EPJC 83 (2023) 441
pp	8	50×10^{-8}	$\sigma = 96.07 \pm 0.18 \pm 0.91$ mb	$\sigma = 99.55 \pm 2.14$ mb (COMPETE HPR1R2)	PLB 761 (2016) 158
pp	7	8×10^{-8}	$\sigma = 95.35 \pm 0.38 \pm 1.3$ mb	$\sigma = 97.26 \pm 2.12$ mb (COMPETE HPR1R2)	Nucl. Phys. B (2014) 486
W	13.6	29.0	$\sigma = 187 \pm 0.1 \pm 6.5$ nb	$\sigma = 193.2 + 3.8 - 5.9$ nb (DYTURBO+CT18 (NNLO+NNLL))	PLB 854 (2024) 138725
W	13	0.3	$\sigma = 190.1 \pm 0.4 \pm 4.2$ nb	$\sigma = 185.3 + 3.8 - 5.5$ nb (DYTURBO + CT18 (NNLO+NNLL))	arXiv:2404.06204
W	8	20.2	$\sigma = 112.69 \pm 3.1$ nb	$\sigma = 110.919889503 \pm 3.7$ nb (DYNNLO + CT14NNLO)	EPJC 79 (2019) 760
W	7	4.6	$\sigma = 98.71 \pm 0.028 \pm 2.191$ nb	$\sigma = 95.9 \pm 2.9$ nb (DYNNLO + CT14NNLO)	EPJC 77 (2017) 367
W	5	0.3	$\sigma = 67.334 \pm 0.06 \pm 0.74$ nb	$\sigma = 66.36 + 1.3 - 2.1$ nb (DYTURBO+CT18 (NNLO+NNLL))	arXiv:2404.06204
Z	13.6	29.0	$\sigma = 59.12 \pm 0.029 \pm 1.6$ nb	$\sigma = 59.12 \pm 3.56$ nb (DYTURBO+CT18 (NNLO+NNLL))	PLB 854 (2024) 138725
Z	13	0.3	$\sigma = 60.18 \pm 0.2 \pm 1.78$ nb	$\sigma = 57.67 + 1.4 - 1.9$ nb (DYTURBO + CT18 (NNLO+NNLL))	arXiv:2404.06204
Z	8	20.2	$\sigma = 34.24 \pm 0.03 \pm 0.92$ nb	$\sigma = 32.94 + 0.8 - 0.92$ nb (DYNNLO+CT14 NNLO)	JHEP 02 (2017) 117
Z	7	4.6	$\sigma = 29.53 \pm 0.03 \pm 0.77$ nb	$\sigma = 28.31 + 0.68 - 0.8$ nb (DYNNLO+CT14 NNLO)	JHEP 02 (2017) 117
Z	5	0.3	$\sigma = 20.11 \pm 0.04 \pm 0.35$ nb	$\sigma = 19.32 + 0.31 - 0.55$ nb (DYTURBO+CT18 (NNLO+NNLL))	arXiv:2404.06204
$\mathbf{t\bar{t}}$	13.6	29.0	$\sigma = 850 \pm 3 \pm 27$ pb	$\sigma = 924 + 32 - 40$ pb (LHC TOP WG)	PLB 848 (2024) 138376
$\mathbf{t\bar{t}}$	13	140	$\sigma = 829 \pm 1 \pm 15.4$ pb	$\sigma = 834 + 29 - 37$ pb (LHC TOP WG)	JHEP 07 (2023) 141
$\mathbf{t\bar{t}}$	8	20.2	$\sigma = 242.9 \pm 1.7 \pm 8.6$ pb	$\sigma = 256 + 10.4 - 12$ pb (LHC TOP WG)	EPJC 74 (2014) 3109
$\mathbf{t\bar{t}}$	7	4.6	$\sigma = 182.9 \pm 3.1 \pm 6.4$ pb	$\sigma = 179.6 + 7.8 - 8.7$ pb (LHC TOP WG)	EPJC 74 (2014) 3109
$\mathbf{t\bar{t}}$	5	0.3	$\sigma = 67.5 \pm 0.9 \pm 2.6$ pb	$\sigma = 69.5 + 3.5 - 3.7$ pb (LHC TOP WG)	JHEP 06 (2023) 138
$\mathbf{t_{t-chan}}$	13	140	$\sigma = 221 \pm 1 \pm 13$ pb	$\sigma = 214.2 + 4.1 - 2.6$ pb (MCFM (NNLO))	JHEP 05 (2024) 305
$\mathbf{t_{t-chan}}$	8	20.3	$\sigma = 89.6 \pm 1.7 + 7.2 - 6.4$ pb	$\sigma = 84.3 + 1.7 - 1.2$ pb (MCFM (NNLO))	EPJC 77 (2017) 531
$\mathbf{t_{t-chan}}$	7	4.6	$\sigma = 68 \pm 2 \pm 8$ pb	$\sigma = 63.7 + 1.4 - 0.8$ pb (MCFM (NNLO))	PRD 90, 112006 (2014)
$\mathbf{t_{t-chan}}$	5	0.3	$\sigma = 27.1 + 4.4 - 4.1 + 4.4 - 3.7$ pb	$\sigma = 30.3 + 0.7 - 0.5$ pb (MCFM (NNLO))	PLB 854 (2024) 138726
Wt	13	3.2	$\sigma = 94 \pm 10 + 28 - 23$ pb	$\sigma = 79.3 + 2.9 - 2.8$ pb (NLO+NNLL)	JHEP 01 (2018) 63
Wt	8	20.3	$\sigma = 23 \pm 1.3 + 3.4 - 3.7$ pb	$\sigma = 24.4 + 1.1 - 1$ pb (NLO+NLL)	JHEP 01, 064 (2016)
Wt	7	2.0	$\sigma = 16.8 \pm 2.9 \pm 3.9$ pb	$\sigma = 17.1 \pm 0.8$ pb (NLO+NLL)	PLB 716, 142-159 (2012)
H	13.6	31.4	$\sigma = 58.2 \pm 7.5 \pm 4.5$ pb	$\sigma = 59.9 \pm 2.6$ pb (LHC-HXSWG YR4)	EPJC 84 (2024) 78
H	13	139	$\sigma = 55.5 \pm 3.2 + 2.4 - 2.2$ pb	$\sigma = 55.6 \pm 2.5$ pb (LHC-HXSWG YR4)	JHEP 05 (2023) 028
H	8	20.3	$\sigma = 27.7 \pm 3 + 2.3 - 1.9$ pb	$\sigma = 24.5 + 1.3 - 1.8$ pb (LHC-HXSWG YR4)	EPJC 76 (2016) 6
H	7	4.5	$\sigma = 22.1 + 6.7 - 5.3 + 3.3 - 2.7$ pb	$\sigma = 19.2 + 1 - 1.4$ pb (LHC-HXSWG YR4)	EPJC 76 (2016) 6
H VBF, $\mathbf{\lvert y_H \rvert < 2.5}$	13	139	$\sigma = 3.86 + 0.29 - 0.28 + 0.37 - 0.33$ pb	$\sigma = 3.5 \pm 0.07$ pb (LHC-HXSWG)	Nature 607, pages 52-59 (2022)
H VBF	8	20.3	$\sigma = 2.43 + 0.5 - 0.49 + 0.29 - 0.31$ pb	$\sigma = 1.6 \pm 0.04$ pb (LHC-HXSWG YR4)	EPJC 76 (2016) 6
VH	8	20.3	$\sigma = 1.03 + 0.37 - 0.36 \pm 0.4$ pb	$\sigma = 1.12 \pm 0.03$ pb (NNLO QCD + NLO EW)	EPJC 76 (2016) 6
WH, $\mathbf{\lvert y_H \rvert < 2.5}$	13	139	$\sigma = 1.41 + 0.21 - 0.2 + 0.18 - 0.17$ pb	$\sigma = 1.22 \pm 0.024$ pb (Powheg Box NLO QCD)	Nature 607, pages 52-59 (2022)
ZH, $\mathbf{\lvert y_H \rvert < 2.5}$	13	139	$\sigma = 0.77 \pm 0.13 + 0.12 - 0.1$ pb	$\sigma = 0.8 \pm 0.03$ pb (Powheg Box NLO QCD)	Nature 607, pages 52-59 (2022)
$\mathbf{t\bar{t}H}$	13	139	$\sigma = 560 \pm 80 + 70 - 80$ fb	$\sigma = 580 \pm 50$ fb (LHCHXSWG NLO QCD + NLO EW)	Nature 607, pages 52-59 (2022)
$\mathbf{t\bar{t}H}$	8	20.3	$\sigma = 220 \pm 100 \pm 70$ fb	$\sigma = 133 + 8 - 13$ fb (LHCHXSWG NLO QCD + NLO EW)	PLB 784 (2018) 173
WW	13	36.1	$\sigma = 130.04 \pm 1.7 \pm 10.6$ pb	$\sigma = 128.4 + 3.2 - 2.9$ pb (NNLO)	EPJC 79 (2019) 884
WW	8	20.3	$\sigma = 68.2 \pm 1.2 \pm 4.6$ pb	$\sigma = 65 + 1.2 - 1.1$ pb (NNLO)	PLB 763, 114 (2016)
WW	7	4.6	$\sigma = 51.9 \pm 2 \pm 4.4$ pb	$\sigma = 49.04 + 1.03 - 0.88$ pb (NNLO)	PRD 87 (2013) 112001, PRL 113 (2014) 212001
WZ	13	36.1	$\sigma = 51 \pm 0.8 \pm 2.3$ pb	$\sigma = 49.1 + 1.1 - 1$ pb (MATRIX (NNLO))	EPJC 79 (2019) 535
WZ	8	20.3	$\sigma = 24.3 \pm 0.6 \pm 0.9$ pb	$\sigma = 23.92 \pm 0.4$ pb (MATRIX (NNLO))	PRD 93, 092004 (2016)
WZ	7	4.6	$\sigma = 19 + 1.4 - 1.3 \pm 1$ pb	$\sigma = 19.34 + 0.3 - 0.4$ pb (MATRIX (NNLO))	EPJC 72 (2012) 2173
ZZ	13.6	29.0	$\sigma = 16.9 \pm 0.7 \pm 0.7$ pb	$\sigma = 16.7 \pm 0.4$ pb (Matrix (NNLO) & Sherpa (NLO))	PLB 855 (2024) 138764
ZZ	13	36.1	$\sigma = 17.3 \pm 0.6 \pm 0.8$ pb	$\sigma = 16.9 + 0.6 - 0.5$ pb (Matrix (NNLO) & Sherpa (NLO))	PRD 97 (2018) 032005
ZZ	8	20.3	$\sigma = 7.3 \pm 0.4 + 0.4 - 0.3$ pb	$\sigma = 8.284 + 0.249 - 0.191$ pb (NNLO)	JHEP 01 (2017) 099
ZZ	7	4.6	$\sigma = 6.7 \pm 0.7 + 0.5 - 0.4$ pb	$\sigma = 6.735 + 0.195 - 0.155$ pb (NNLO)	JHEP 03 (2013) 128 , PLB 735 (2014) 311
$\mathbf{t_{s-chan}}$	13	140	$\sigma = 8.2 \pm 0.6 + 3.4 - 2.8$ pb	$\sigma = 10.32 + 0.4 - 0.36$ pb (NLO+NNL)	JHEP 06 (2023) 191
$\mathbf{t_{s-chan}}$	8	20.3	$\sigma = 4.8 \pm 0.8 + 1.6 - 1.3$ pb	$\sigma = 5.61 \pm 0.22$ pb (NLO+NNL)	PLB 756 (2016) 228-246
$\mathbf{t\bar{t}W}$	13	140	$\sigma = 880 \pm 50 \pm 70$ fb	$\sigma = 745 \pm 52$ fb (NNLO QCD + NLO EW)	JHEP 05 (2024) 131
$\mathbf{t\bar{t}W}$	8	20.3	$\sigma = 369 + 86 - 79 \pm 44$ fb	$\sigma = 232 \pm 32$ fb (MCFM)	JHEP 11 (2015) 172
$\mathbf{t\bar{t}Z}$	13	140	$\sigma = 860 \pm 40 \pm 40$ fb	$\sigma = 860 + 80 - 90$ fb (NLO + NNLL)	arXiv:2312.04450
$\mathbf{t\bar{t}Z}$	8	20.3	$\sigma = 176 + 52 - 48 \pm 24$ fb	$\sigma = 215 \pm 30$ fb (HELAC-NLO)	JHEP 11, 172 (2015)
WWW	13	139	$\sigma = 0.82 \pm 0.01 \pm 0.08$ pb	$\sigma = 0.511 \pm 0.018$ pb (NLO QCD)	PRL 129 (2022) 061803
WWZ	13	79.8	$\sigma = 0.55 \pm 0.14 + 0.15 - 0.13$ pb	$\sigma = 0.358 \pm 0.036$ pb (Sherpa 2.2.2)	PLB 798 (2019) 134913
$\mathbf{t\bar{t}t\bar{t}}$	13	140	$\sigma = 22.5 + 4.7 - 3.4 + 6.6 - 5.5$ fb	$\sigma = 13.4 + 1 - 1.8$ fb (NLO QCD + EW)	EPJC 83 (2023) 496

Figure B.1: A summary of the cross-section measurements and the theoretical predictions for standard model processes in proton-proton collisions at the LHC [4].

The tables in Figure B.2 and Figure B.3 show the measured and theoretical cross-sections for the QCD jet production, and the Higgs boson processes in proton-proton collisions at the LHC.



Standard Model Production Cross Section Measurements IV

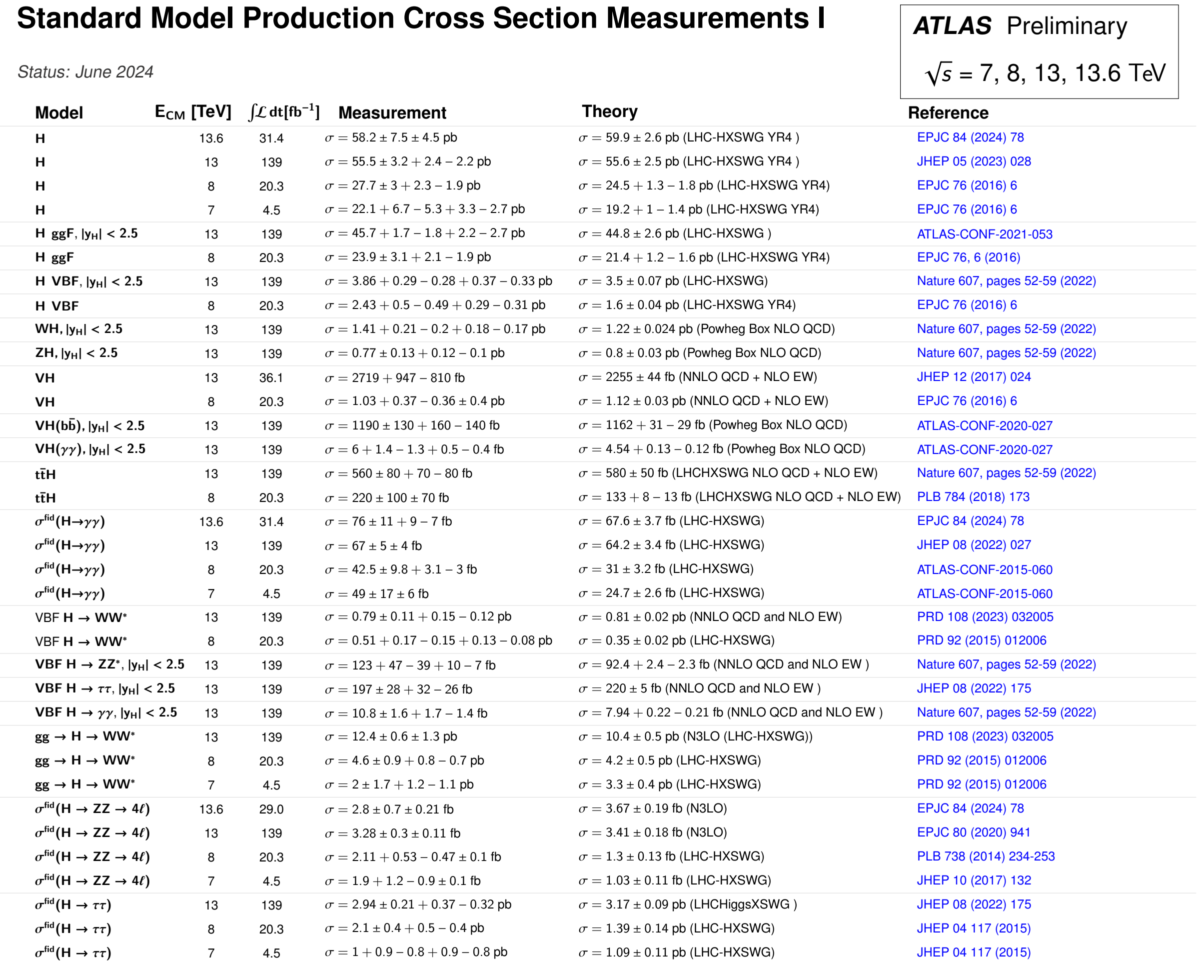
Status: June 2024

ATLAS Preliminary

$\sqrt{s}$ = 7, 8, 13 TeV

Model	E_{CM} [TeV]	$\int\mathcal{L}\,dt[fb^{-1}]$	Measurement	Theory	Reference
pp	13	34×10^{-8}	$\sigma = 104.7 \pm 0.22 \pm 1.07$ mb	$\sigma = 100.3 \pm 0.12$ mb (COMPETE HPR1R2)	EPJC 83 (2023) 441
pp	8	50×10^{-8}	$\sigma = 96.07 \pm 0.18 \pm 0.91$ mb	$\sigma = 99.55 \pm 2.14$ mb (COMPETE HPR1R2)	PLB 761 (2016) 158
pp	7	8×10^{-8}	$\sigma = 95.35 \pm 0.38 \pm 1.3$ mb	$\sigma = 97.26 \pm 2.12$ mb (COMPETE HPR1R2)	Nucl. Phys. B (2014) 486
pp inelastic	13	34×10^{-8}	$\sigma = 77.41 \pm 1.08$ mb	$\sigma = 78.4 \pm 2$ mb (Schuler/Sjöstrand)	EPJC 83 (2023) 441
pp inelastic	8	50×10^{-8}	$\sigma = 71.73 \pm 0.15 \pm 0.69$ mb	$\sigma = 73 \pm 2$ mb (Schuler/Sjöstrand)	PLB 761 (2016) 158
pp inelastic	7	8×10^{-8}	$\sigma = 71.34 \pm 0.36 \pm 0.83$ mb	$\sigma = 71.5 + 20 - 2$ mb (Schuler/Sjöstrand)	Nucl. Phys. B (2014) 486
$2.5 < y^* < 3.0$, $2 < m_{jj} < 5$ TeV	13	3.2	$\sigma = 850 \pm 53 + 68 - 91$ pb	$\sigma = 955 + 56 - 199$ pb (NLOJet++, CT14)	JHEP 05 (2018) 195
$2.5 < y^* < 3.0$, $2 < m_{jj} < 5$ TeV	7	4.5	$\sigma = 16 \pm 2 + 5.4 - 4.3$ pb	$\sigma = 18.4 + 2.2 - 4.3$ pb (NLOJet++, CT10)	JHEP 05 (2014) 059
$2.0 < y^* < 2.5$, $1.3 < m_{jj} < 5$ TeV	13	3.2	$\sigma = 6.39 \pm 0.14 + 0.47 - 0.54$ nb	$\sigma = 6.7 + 0.5 - 1.3$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$2.0 < y^* < 2.5$, $1.3 < m_{jj} < 5$ TeV	7	4.5	$\sigma = 371 \pm 9.7 + 81.5 - 72.1$ pb	$\sigma = 410.6 + 31 - 77.8$ pb (NLOJet++, CT10)	JHEP 05 (2014) 059
$1.5 < y^* < 2.0$, $0.8 < m_{jj} < 4.6$ TeV	13	3.2	$\sigma = 16.13 \pm 0.17 \pm 1.09$ nb	$\sigma = 17.4 + 0.7 - 3.3$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$1.5 < y^* < 2.0$, $0.8 < m_{jj} < 4.6$ TeV	7	4.5	$\sigma = 3.57 \pm 0.04 + 0.51 - 0.49$ nb	$\sigma = 3.7 + 0.21 - 0.62$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
$1.0 < y^* < 1.5$, $0.5 < m_{jj} < 4.6$ TeV	13	3.2	$\sigma = 68.7 \pm 0.4 + 4 - 4.2$ nb	$\sigma = 68.5 + 7.7 - 10.3$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$1.0 < y^* < 1.5$, $0.5 < m_{jj} < 4.6$ TeV	7	4.5	$\sigma = 10.12 \pm 0.07 + 1.02 - 1.03$ nb	$\sigma = 10.2 + 0.5 - 1.5$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
$0.5 < y^* < 1.0$, $0.3 < m_{jj} < 4.3$ TeV	13	3.2	$\sigma = 117.6 \pm 0.5 + 6.8 - 6.9$ nb	$\sigma = 127.3 + 5.7 - 19$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$0.5 < y^* < 1.0$, $0.3 < m_{jj} < 4.3$ TeV	7	4.5	$\sigma = 37.33 \pm 0.2 + 3.25 - 3.03$ nb	$\sigma = 37.3 + 1.6 - 5.1$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
$y^* < 0.5$, $0.3 < m_{jj} < 4.3$ TeV	13	3.2	$\sigma = 111.2 \pm 0.4 + 6.2 - 6.3$ nb	$\sigma = 118.6 + 5.5 - 18.8$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$y^* < 0.5$, $0.3 < m_{jj} < 4.3$ TeV	7	4.5	$\sigma = 35.47 \pm 0.15 + 2.79 - 2.66$ nb	$\sigma = 35.3 + 1.5 - 5.1$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
Dijet R=0.4, $\lvert y\rvert < 3.0$, $y^* < 3.0$	13	3.2	$\sigma = 321 \pm 0.8 + 18.6 - 19$ nb	$\sigma = 340 + 17 - 54$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
Dijet R=0.4, $\lvert y\rvert < 3.0$, $y^* < 3.0$	7	4.5	$\sigma = 86.87 \pm 0.26 + 7.56 - 7.2$ nb	$\sigma = 86.9 + 4.7 - 12.4$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
$2.5 < y^* < 3.0$, $2 < m_{jj} < 5$ TeV	7	4.5	$\sigma = 26.9 \pm 4.2 + 7.7 - 6.4$ pb	$\sigma = 23.5 + 2.7 - 2.8$ pb (NLOJet++, CT10)	JHEP 05 (2014) 059
$2.0 < y^* < 2.5$, $1.3 < m_{jj} < 5$ TeV	7	4.5	$\sigma = 505 \pm 15.1 + 102.4 - 92.4$ pb	$\sigma = 526.9 + 37.5 - 46.3$ pb (NLOJet++, CT10)	JHEP 05 (2014) 059
$1.5 < y^* < 2.0$, $0.8 < m_{jj} < 4.6$ TeV	7	4.5	$\sigma = 4.93 \pm 0.06 + 0.69 - 0.65$ nb	$\sigma = 4.78 + 0.23 - 0.34$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
$1.0 < y^* < 1.5$, $0.5 < m_{jj} < 4.6$ TeV	7	4.5	$\sigma = 13.82 \pm 0.11 + 1.44 - 1.42$ nb	$\sigma = 13.2 + 0.5 - 0.8$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
$0.5 < y^* < 1.0$, $0.3 < m_{jj} < 4.3$ TeV	7	4.5	$\sigma = 51.47 \pm 0.32 + 4.76 - 4.44$ nb	$\sigma = 48.7 + 1.3 - 2.5$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
$y^* < 0.5$, $0.3 < m_{jj} < 4.3$ TeV	7	4.5	$\sigma = 48.21 \pm 0.23 + 4.03 - 3.8$ nb	$\sigma = 46.1 + 1.2 - 2.5$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
Dijet R=0.6, $\lvert y\rvert < 3.0$, $y^* < 3.0$	7	4.5	$\sigma = 119 \pm 0.4 + 10.9 - 10.3$ nb	$\sigma = 113.3 + 3.1 - 6.1$ nb (NLOJet++, CT10)	JHEP 05 (2014) 059
$2.5 < \lvert y\rvert < 3.0$, $p_T > 100$ GeV	13	3.2	$\sigma = 157 \pm 1 \pm 13$ nb	$\sigma = 165 + 13 - 18$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$2.5 < \lvert y\rvert < 3.0$, $p_T > 100$ GeV	8	20.2	$\sigma = 45.3 \pm 0.3 + 3.9 - 3.8$ nb	$\sigma = 50.4 + 4.1 - 7.4$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$2.5 < \lvert y\rvert < 3.0$, $p_T > 100$ GeV	7	4.5	$\sigma = 29.13 \pm 0.31 + 7.5 - 6.38$ nb	$\sigma = 32.6 + 2.1 - 3.2$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$2.0 < \lvert y\rvert < 2.5$, $p_T > 100$ GeV	13	3.2	$\sigma = 222 \pm 1 + 19 - 20$ nb	$\sigma = 241 + 19 - 26$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$2.0 < \lvert y\rvert < 2.5$, $p_T > 100$ GeV	8	20.2	$\sigma = 79.8 \pm 0.4 \pm 5.4$ nb	$\sigma = 88.8 + 6.4 - 11.7$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$2.0 < \lvert y\rvert < 2.5$, $p_T > 100$ GeV	7	4.5	$\sigma = 57.1 \pm 0.4 + 10.4 - 9.1$ nb	$\sigma = 60.7 + 3.3 - 5.1$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$1.5 < \lvert y\rvert < 2.0$, $p_T > 100$ GeV	13	3.2	$\sigma = 288 \pm 1 \pm 21$ nb	$\sigma = 317 + 24 - 33$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$1.5 < \lvert y\rvert < 2.0$, $p_T > 100$ GeV	8	20.2	$\sigma = 111 \pm 0.4 + 6.9 - 6.8$ nb	$\sigma = 124.7 + 9.5 - 15.4$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$1.5 < \lvert y\rvert < 2.0$, $p_T > 100$ GeV	7	4.5	$\sigma = 83.5 \pm 0.6 + 11.1 - 9.7$ nb	$\sigma = 88.3 + 4.7 - 7.1$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$1.0 < \lvert y\rvert < 1.5$, $p_T > 100$ GeV	13	3.2	$\sigma = 350 \pm 2 \pm 24$ nb	$\sigma = 383 + 28 - 38$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$1.0 < \lvert y\rvert < 1.5$, $p_T > 100$ GeV	8	20.2	$\sigma = 145.4 \pm 0.5 + 8.9 - 8.6$ nb	$\sigma = 157 + 12 - 19$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$1.0 < \lvert y\rvert < 1.5$, $p_T > 100$ GeV	7	4.5	$\sigma = 112.2 \pm 0.7 + 11 - 10.2$ nb	$\sigma = 113.1 + 5.8 - 9.2$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$0.5 < \lvert y\rvert < 1.0$, $p_T > 100$ GeV	13	3.2	$\sigma = 401 \pm 2 \pm 24$ nb	$\sigma = 431 + 33 - 44$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$0.5 < \lvert y\rvert < 1.0$, $p_T > 100$ GeV	8	20.2	$\sigma = 167.9 \pm 0.5 + 9.6 - 9.4$ nb	$\sigma = 182 + 14 - 23$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$0.5 < \lvert y\rvert < 1.0$, $p_T > 100$ GeV	7	4.5	$\sigma = 136.9 \pm 0.8 + 10.9 - 10.5$ nb	$\sigma = 132.3 + 6.9 - 10.7$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$\lvert y\rvert < 0.5$, $p_T > 100$ GeV	13	3.2	$\sigma = 427 \pm 2 \pm 24$ nb	$\sigma = 459 + 35 - 49$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
$\lvert y\rvert < 0.5$, $p_T > 100$ GeV	8	20.2	$\sigma = 177 \pm 0.5 + 9.6 - 9.4$ nb	$\sigma = 196 + 14 - 25$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$\lvert y\rvert < 0.5$, $p_T > 100$ GeV	7	4.5	$\sigma = 145.1 \pm 0.8 + 10.7 - 10.6$ nb	$\sigma = 142.7 + 7.4 - 11.5$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
Incl. jet R=0.4, $\lvert y\rvert < 3.0$	13	3.2	$\sigma = 1845 \pm 4 + 119 - 120$ nb	$\sigma = 1997 + 152 - 208$ nb (NLOJet++, CT14)	JHEP 05 (2018) 195
Incl. jet R=0.4, $\lvert y\rvert < 3.0$	8	20.2	$\sigma = 726.4 \pm 1.1 + 42.7 - 41.8$ nb	$\sigma = 800 + 59 - 100$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
Incl. jet R=0.4, $\lvert y\rvert < 3.0$	7	4.5	$\sigma = 563.9 \pm 1.5 + 55.4 - 51.4$ nb	$\sigma = 569.8 + 29.5 - 46.3$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$2.5 < \lvert y\rvert < 3.0$, $p_T > 100$ GeV	8	20.2	$\sigma = 58.6 \pm 0.8 + 5.8 - 5.6$ nb	$\sigma = 59.8 + 4.1 - 6.6$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$2.5 < \lvert y\rvert < 3.0$, $p_T > 100$ GeV	7	4.5	$\sigma = 37.5 \pm 0.4 + 9.4 - 8.4$ nb	$\sigma = 36.9 + 3.7 - 2.4$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$2.0 < \lvert y\rvert < 2.5$, $p_T > 100$ GeV	8	20.2	$\sigma = 100.5 \pm 1.1 + 8.6 - 8.3$ nb	$\sigma = 105.8 + 6.6 - 11.4$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$2.0 < \lvert y\rvert < 2.5$, $p_T > 100$ GeV	7	4.5	$\sigma = 69.7 \pm 0.6 + 13.5 - 12.7$ nb	$\sigma = 68.6 + 6.3 - 4.3$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$1.5 < \lvert y\rvert < 2.0$, $p_T > 100$ GeV	8	20.2	$\sigma = 140.3 \pm 1.2 + 11.1 - 10.8$ nb	$\sigma = 149.8 + 9.4 - 14.6$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$1.5 < \lvert y\rvert < 2.0$, $p_T > 100$ GeV	7	4.5	$\sigma = 105.5 \pm 0.7 + 16 - 15.2$ nb	$\sigma = 100.2 + 9.2 - 5.9$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$1.0 < \lvert y\rvert < 1.5$, $p_T > 100$ GeV	8	20.2	$\sigma = 190.7 \pm 1.4 + 15 - 14.6$ nb	$\sigma = 189 + 11 - 18$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$1.0 < \lvert y\rvert < 1.5$, $p_T > 100$ GeV	7	4.5	$\sigma = 139.8 \pm 0.9 + 16.5 - 16.2$ nb	$\sigma = 128.8 + 11.7 - 7.4$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$0.5 < \lvert y\rvert < 1.0$, $p_T > 100$ GeV	8	20.2	$\sigma = 221.6 \pm 1.5 + 16.5 - 15.8$ nb	$\sigma = 220 + 13 - 21$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$0.5 < \lvert y\rvert < 1.0$, $p_T > 100$ GeV	7	4.5	$\sigma = 172.7 \pm 0.9 + 15.9 - 14.3$ nb	$\sigma = 151 + 13.8 - 8.6$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
$\lvert y\rvert < 0.5$, $p_T > 100$ GeV	8	20.2	$\sigma = 239.3 \pm 1.6 + 16.5 - 15.9$ nb	$\sigma = 237 + 14 - 24$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
$\lvert y\rvert < 0.5$, $p_T > 100$ GeV	7	4.5	$\sigma = 187 \pm 0.9 + 15.1 - 15$ nb	$\sigma = 162.9 + 15 - 9.2$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153
Incl. jet R=0.6, $\lvert y\rvert < 3.0$	8	20.2	$\sigma = 951 \pm 3 + 72 - 70$ nb	$\sigma = 961 + 58 - 95$ nb (NLOJet++, CT14)	JHEP 09 (2017) 020
Incl. jet R=0.6, $\lvert y\rvert < 3.0$	7	4.5	$\sigma = 712.3 \pm 1.9 + 79.9 - 76$ nb	$\sigma = 648.3 + 58.96 - 37.1$ nb (NLOJet++, CT10)	JHEP 02 (2015) 153

Figure B.2: A summary of the cross-section measurements and the theoretical predictions for QCD jet production in proton-proton collisions at the LHC [4].



Standard Model Production Cross Section Measurements I

Status: June 2024

ATLAS Preliminary

$\sqrt{s}$ = 7, 8, 13, 13.6 TeV

Model	E_{CM} [TeV]	$\int \mathcal{L}\,dt [fb^{-1}]$	Measurement	Theory	Reference
H	13.6	31.4	$\sigma = 58.2 \pm 7.5 \pm 4.5$ pb	$\sigma = 59.9 \pm 2.6$ pb (LHC-HXSWG YR4)	EPJC 84 (2024) 78
H	13	139	$\sigma = 55.5 \pm 3.2 + 2.4 - 2.2$ pb	$\sigma = 55.6 \pm 2.5$ pb (LHC-HXSWG YR4)	JHEP 05 (2023) 028
H	8	20.3	$\sigma = 27.7 \pm 3 + 2.3 - 1.9$ pb	$\sigma = 24.5 + 1.3 - 1.8$ pb (LHC-HXSWG YR4)	EPJC 76 (2016) 6
H	7	4.5	$\sigma = 22.1 + 6.7 - 5.3 + 3.3 - 2.7$ pb	$\sigma = 19.2 + 1 - 1.4$ pb (LHC-HXSWG YR4)	EPJC 76 (2016) 6
H ggF, $|y_H| < 2.5$	13	139	$\sigma = 45.7 + 1.7 - 1.8 + 2.2 - 2.7$ pb	$\sigma = 44.8 \pm 2.6$ pb (LHC-HXSWG)	ATLAS-CONF-2021-053
H ggF	8	20.3	$\sigma = 23.9 \pm 3.1 + 2.1 - 1.9$ pb	$\sigma = 21.4 + 1.2 - 1.6$ pb (LHC-HXSWG YR4)	EPJC 76, 6 (2016)
H VBF, $|y_H| < 2.5$	13	139	$\sigma = 3.86 + 0.29 - 0.28 + 0.37 - 0.33$ pb	$\sigma = 3.5 \pm 0.07$ pb (LHC-HXSWG)	Nature 607, pages 52-59 (2022)
H VBF	8	20.3	$\sigma = 2.43 + 0.5 - 0.49 + 0.29 - 0.31$ pb	$\sigma = 1.6 \pm 0.04$ pb (LHC-HXSWG YR4)	EPJC 76 (2016) 6
WH, $|y_H| < 2.5$	13	139	$\sigma = 1.41 + 0.21 - 0.2 + 0.18 - 0.17$ pb	$\sigma = 1.22 \pm 0.024$ pb (Powheg Box NLO QCD)	Nature 607, pages 52-59 (2022)
ZH, $|y_H| < 2.5$	13	139	$\sigma = 0.77 \pm 0.13 + 0.12 - 0.1$ pb	$\sigma = 0.8 \pm 0.03$ pb (Powheg Box NLO QCD)	Nature 607, pages 52-59 (2022)
VH	13	36.1	$\sigma = 2719 + 947 - 810$ fb	$\sigma = 2255 \pm 44$ fb (NNLO QCD + NLO EW)	JHEP 12 (2017) 024
VH	8	20.3	$\sigma = 1.03 + 0.37 - 0.36 \pm 0.4$ pb	$\sigma = 1.12 \pm 0.03$ pb (NNLO QCD + NLO EW)	EPJC 76 (2016) 6
VH($b\bar{b}$), $|y_H| < 2.5$	13	139	$\sigma = 1190 \pm 130 + 160 - 140$ fb	$\sigma = 1162 + 31 - 29$ fb (Powheg Box NLO QCD)	ATLAS-CONF-2020-027
VH($\gamma\gamma$), $|y_H| < 2.5$	13	139	$\sigma = 6 + 1.4 - 1.3 + 0.5 - 0.4$ fb	$\sigma = 4.54 + 0.13 - 0.12$ fb (Powheg Box NLO QCD)	ATLAS-CONF-2020-027
$t\bar{t}H$	13	139	$\sigma = 560 \pm 80 + 70 - 80$ fb	$\sigma = 580 \pm 50$ fb (LHCHXSWG NLO QCD + NLO EW)	Nature 607, pages 52-59 (2022)
$t\bar{t}H$	8	20.3	$\sigma = 220 \pm 100 \pm 70$ fb	$\sigma = 133 + 8 - 13$ fb (LHCHXSWG NLO QCD + NLO EW)	PLB 784 (2018) 173
$\sigma^{fid}(H \to \gamma\gamma)$	13.6	31.4	$\sigma = 76 \pm 11 + 9 - 7$ fb	$\sigma = 67.6 \pm 3.7$ fb (LHC-HXSWG)	EPJC 84 (2024) 78
$\sigma^{fid}(H \to \gamma\gamma)$	13	139	$\sigma = 67 \pm 5 \pm 4$ fb	$\sigma = 64.2 \pm 3.4$ fb (LHC-HXSWG)	JHEP 08 (2022) 027
$\sigma^{fid}(H \to \gamma\gamma)$	8	20.3	$\sigma = 42.5 \pm 9.8 + 3.1 - 3$ fb	$\sigma = 31 \pm 3.2$ fb (LHC-HXSWG)	ATLAS-CONF-2015-060
$\sigma^{fid}(H \to \gamma\gamma)$	7	4.5	$\sigma = 49 \pm 17 \pm 6$ fb	$\sigma = 24.7 \pm 2.6$ fb (LHC-HXSWG)	ATLAS-CONF-2015-060
VBF $H \to WW^*$	13	139	$\sigma = 0.79 \pm 0.11 + 0.15 - 0.12$ pb	$\sigma = 0.81 \pm 0.02$ pb (NNLO QCD and NLO EW)	PRD 108 (2023) 032005
VBF $H \to WW^*$	8	20.3	$\sigma = 0.51 + 0.17 - 0.15 + 0.13 - 0.08$ pb	$\sigma = 0.35 \pm 0.02$ pb (LHC-HXSWG)	PRD 92 (2015) 012006
VBF $H \to ZZ^*$, $|y_H| < 2.5$	13	139	$\sigma = 123 + 47 - 39 + 10 - 7$ fb	$\sigma = 92.4 + 2.4 - 2.3$ fb (NNLO QCD and NLO EW)	Nature 607, pages 52-59 (2022)
VBF $H \to \tau\tau$, $|y_H| < 2.5$	13	139	$\sigma = 197 \pm 28 + 32 - 26$ fb	$\sigma = 220 \pm 5$ fb (NNLO QCD and NLO EW)	JHEP 08 (2022) 175
VBF $H \to \gamma\gamma$, $|y_H| < 2.5$	13	139	$\sigma = 10.8 \pm 1.6 + 1.7 - 1.4$ fb	$\sigma = 7.94 + 0.22 - 0.21$ fb (NNLO QCD and NLO EW)	Nature 607, pages 52-59 (2022)
$gg \to H \to WW^*$	13	139	$\sigma = 12.4 \pm 0.6 \pm 1.3$ pb	$\sigma = 10.4 \pm 0.5$ pb (N3LO (LHC-HXSWG))	PRD 108 (2023) 032005
$gg \to H \to WW^*$	8	20.3	$\sigma = 4.6 \pm 0.9 + 0.8 - 0.7$ pb	$\sigma = 4.2 \pm 0.5$ pb (LHC-HXSWG)	PRD 92 (2015) 012006
$gg \to H \to WW^*$	7	4.5	$\sigma = 2 \pm 1.7 + 1.2 - 1.1$ pb	$\sigma = 3.3 \pm 0.4$ pb (LHC-HXSWG)	PRD 92 (2015) 012006
$\sigma^{fid}(H \to ZZ \to 4\ell)$	13.6	29.0	$\sigma = 2.8 \pm 0.7 \pm 0.21$ fb	$\sigma = 3.67 \pm 0.19$ fb (N3LO)	EPJC 84 (2024) 78
$\sigma^{fid}(H \to ZZ \to 4\ell)$	13	139	$\sigma = 3.28 \pm 0.3 \pm 0.11$ fb	$\sigma = 3.41 \pm 0.18$ fb (N3LO)	EPJC 80 (2020) 941
$\sigma^{fid}(H \to ZZ \to 4\ell)$	8	20.3	$\sigma = 2.11 + 0.53 - 0.47 \pm 0.1$ fb	$\sigma = 1.3 \pm 0.13$ fb (LHC-HXSWG)	PLB 738 (2014) 234-253
$\sigma^{fid}(H \to ZZ \to 4\ell)$	7	4.5	$\sigma = 1.9 + 1.2 - 0.9 \pm 0.1$ fb	$\sigma = 1.03 \pm 0.11$ fb (LHC-HXSWG)	JHEP 10 (2017) 132
$\sigma^{fid}(H \to \tau\tau)$	13	139	$\sigma = 2.94 \pm 0.21 + 0.37 - 0.32$ pb	$\sigma = 3.17 \pm 0.09$ pb (LHCHiggsXSWG)	JHEP 08 (2022) 175
$\sigma^{fid}(H \to \tau\tau)$	8	20.3	$\sigma = 2.1 \pm 0.4 + 0.5 - 0.4$ pb	$\sigma = 1.39 \pm 0.14$ pb (LHC-HXSWG)	JHEP 04 117 (2015)
$\sigma^{fid}(H \to \tau\tau)$	7	4.5	$\sigma = 1 + 0.9 - 0.8 + 0.9 - 0.8$ pb	$\sigma = 1.09 \pm 0.11$ pb (LHC-HXSWG)	JHEP 04 117 (2015)

Figure B.3: A summary of the cross-section measurements and the theoretical predictions for Higgs boson processes in proton-proton collisions at the LHC [4].

C The trigger selection

C.1 The trigger menu

The detailed trigger selections for taking the $VH, H \rightarrow b\bar{b}$ and $VH, H \rightarrow c\bar{c}$ candidate events can be found in Table C.1.

Table C.1: The different triggers and their respective thresholds during the Run2 (2015 to 2018) data collection period.

Trigger	Trigger Name	Period	L1 Threshold	HLT Threshold
$E_{\mathrm{T}}^{\mathrm{miss}}$	HLT_xe70_mht_L1XE50	2015	50 GeV	70 GeV
	HLT_xe90_mht_L1XE50	2016	50 GeV	90 GeV
	HLT_xe110_mht_L1XE50	2016	50 GeV	110 GeV
	HLT_xe110_pufit_L1XE55	2017	55 GeV	110 GeV
	HLT_xe110_pufit_xe70_L1XE50	2018	50 GeV	110 GeV
Electron	HLT_e24_lhmedium_L1EM20VH	2015	20 GeV	24 GeV
	HLT_e60_lhmedium	2015	22 GeV	60 GeV
	HLT_e120_lhloose	2015	22 GeV	120 GeV
	HLT_e26_lhtight_nod0_ivarloose	2016 - 2018	22 GeV	26 GeV
	HLT_e60_lhmedium	2016 - 2018	22 GeV	60 GeV
	HLT_e140_lhloose	2016 - 2018	22 GeV	140 GeV
	HLT_e300_etcut	2018	22 GeV	300 GeV
Muon	HLT_mu20_iloose_L1MU15	2015	15 GeV	20 GeV
	HLT_mu50	2015 - 2018	20 GeV	50 GeV
	HLT_mu26_ivarmedium	2016 - 2018	20 GeV	26 GeV

C.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger

The $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers deployed at L1 has limited availability to detector information, and needs to provide fast performance with a good background rejection. Therefore, the *online* $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithms used for the trigger are different from the *offline* $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithms used for reconstruction as described in Section 4.8. Since the trigger selections need to be fast (in an order of 100 ms), the use of inner detector tracking information is computationally expensive (in an order of 1-5 s). Therefore, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger uses only the information from the calorimeter [86].

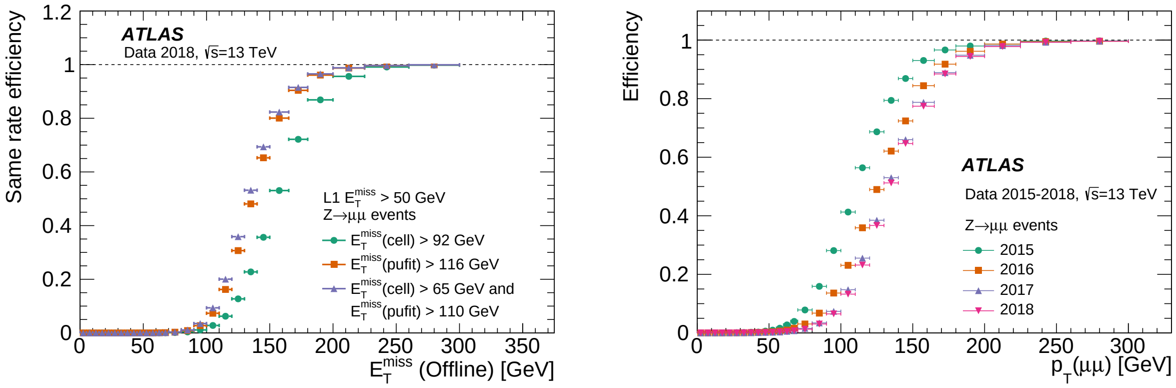
The $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the momentum imbalance in the calorimeter, which is calculated using a set of energy deposits in the calorimeter, generally referred to as *elements*. Elements could be all of the calorimeter cells, or reconstructed jets. The two components of the $\overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}}$ vector is calculated as,

$$E_x^{\mathrm{miss}} = \sum_{i=1}^{|\mathrm{Elements}|} E_i \sin\theta_i \cos\theta_i, \tag{C.1}$$

$$E_y^{\text{miss}} = \sum_{i=1}^{|\text{Elements}|} E_i \sin\theta_i \sin\theta_i, \tag{C.2}$$

and the magnitude is obtained as $E_{\mathrm{T}}^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}$. Muons are treated similar to neutrinos in the $E_{\mathrm{T}}^{\text{miss}}$ trigger calculations since they are mostly invisible in the calorimeter. Therefore, events having muons with high p_T are collected using the $E_{\mathrm{T}}^{\text{miss}}$ trigger for the one and two charged-lepton channels.

At L1, the $E_{\mathrm{T}}^{\text{miss}}$ is calculated based on digitized signals from calorimeter cells forming projective towers of granularity of 0.1×0.1 in (η, ϕ). The digitization results in counts that nominally corresponds to 1 GeV in E_{T}. A fixed threshold, depending on η is applied per tower, and the energy of any tower below this threshold is set to zero. Since the peak luminosity increased by a factor of four times to 2.1×10^{34} cm^{2}s^{-1} (in 2017-2018) from 0.5×10^{34} cm^{2}s^{-1} (in 2015), higher thresholds were applied to provide fixed occupancies. After the thresholds are applied, the towers are summed into larger projective towers of granularity 0.2×0.2 in (η, ϕ), referred to as jet elements. The $E_{\mathrm{T}}^{\text{miss}}$ is calculated using these jet elements using Equation C.1 and C.2.



(**a**) The `pufit` and `mht` trigger efficiencies.

(**b**) The efficiency of $E_{\mathrm{T}}^{\text{miss}}$ triggers from 2015 to 2018.

Figure C.1: The efficiency for the $E_{\mathrm{T}}^{\text{miss}}$ triggers measured using $Z \to \mu\mu$ events, as a function of the transverse momentum of the two muons $p_{\mathrm{T}}(\mu\mu)$. Since muons have little interaction with the calorimeter, $p_{\mathrm{T}}(\mu\mu)$ is a good estimate of the $E_{\mathrm{T}}^{\text{miss}}$ [86].

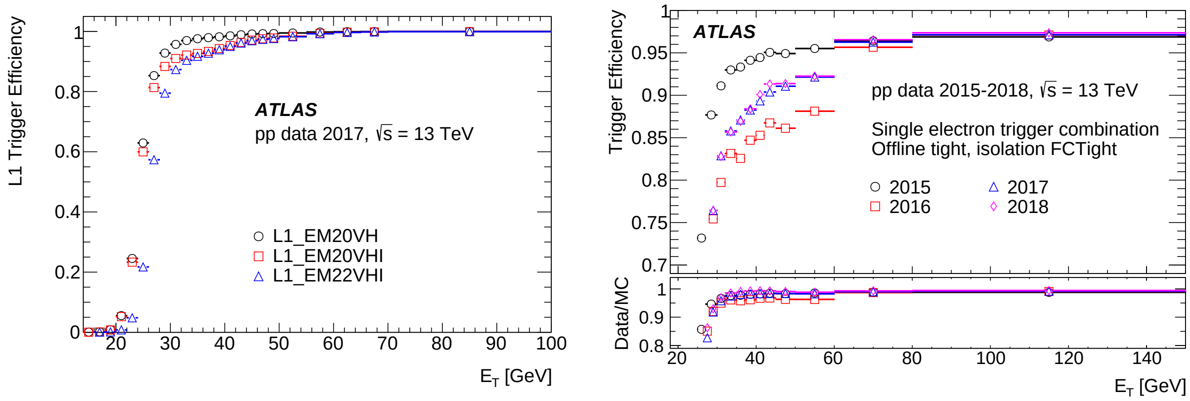
At the HLT, the $E_{\mathrm{T}}^{\text{miss}}$ is recalculated for events accepted by the L1. There are several HLT algorithms used for calculating the $E_{\mathrm{T}}^{\text{miss}}$, but the ones relevant for the data taken in this analysis is the `cell`, `mht`, and `pufit` algorithms. `cell` is the most basic algorithm, and determines the $E_{\mathrm{T}}^{\text{miss}}$ from a sum over the 188,000 calorimeter cells. The `mht` algorithm uses only calibrated jets for the $E_{\mathrm{T}}^{\text{miss}}$ reconstruction, since hadronic jets tend to dominate the visible momentum in most events of interest. Finally the `pufit` algorithm corrects for pile-up effects on the high-E_{T} calorimeter signals contributing to $E_{\mathrm{T}}^{\text{miss}}$, and the pile-up subtracted $\eta - \phi$ regions are used to determine $E_{\mathrm{T}}^{\text{miss}}$. The `pufit` algorithm is observed to have good performance even at higher pile-up.

The L1 trigger rate was reduced by adjusting calorimeter noise thresholds to cope with higher pile-up in from 2017, but mostly had constant a threshold of $E_{\mathrm{T}}^{\text{miss}} > 50$ GeV throughout the Run 2 data taking period. The HLT $E_{\mathrm{T}}^{\text{miss}}$ trigger was the `mht` trigger in the years 2015-2016, and a combination of the `cell` trigger and the `pufit` trigger in the years 2017-2018 that has better performance than using the individual triggers as seen in Figure C.1(a). The overall (L1 & HLT) $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger chains have different thresholds from 2015 to 2018, varying between 50 GeV to 110 GeV. However, events used in the analysis are required to have passed a trigger threshold of $E_{\mathrm{T}}^{\mathrm{miss}} = 150$ GeV to ensure a good trigger efficiency as seen from Figure C.1(b).

Leptonically decaying τ-leptons end up in the electron and muon channels. Hadronically decaying τ-leptons are collected using the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers, with the $E_{\mathrm{T}}^{\mathrm{miss}}$ accounting for the neutrino in the τ-lepton decay.

C.3 Single electron trigger

The L1 trigger for electrons uses the calorimeter information in the $|\eta| < 2.5$ region to build regions-of-interest (RoIs) to be investigated by the HLT. A sliding window approach, where a window contains 4×4 trigger towers with granularity 0.1×0.1 in (η, ϕ), is used for the electromagnetic transverse energy E_{T} reconstruction (detailed in [87]). The reconstructed E_{T} at L1 is required to satisfy a nominal threshold of $E_{\mathrm{T}} > 20$ GeV in 2015, and $E_{\mathrm{T}} > 22$ GeV in 2016 to 2018, and changes from -2 GeV to +3 GeV with respect to the nominal depending on η. A selection to reject hadron activity is also applied, where the sum of E_{T} in the hadronic towers matched to the central region of the cluster is required to be greater than 1 GeV and $E_{\mathrm{T}}/23.0 - 0.2$ GeV. In 2016 to 2018, a further isolation criteria was required; candidate electrons are rejected if the sum of E_{T} in the 12 towers surrounding the central cluster region is at least 2 GeV, and exceeds $E_{\mathrm{T}}/8.0 - 1.8$ GeV. The hadron activity selection and the isolation criteria are not applied at L1 for reconstructed $E_{\mathrm{T}} > 50$ GeV.



(a) L1 trigger efficiency.

(b) Combined trigger efficiency.

Figure C.2: The L1 and combined single-electron trigger efficiency as a function of E_{T} [87]. The shaper efficiency turn on for 2015 is due to looser electron ID requirements (`lhmedium` vs `lhtight` from 2016) and lower E_{T} threshold (24 GeV in 2015, and 26 GeV from 2016).

The HLT electron reconstruction is based on two steps; step one is a fast calorimeter reconstruction and selection step, and step two is a precision reconstruction step using similar reconstruction and calibration steps used in offline analysis (described in more detail in Section 4.2). In the fast reconstruction step, electrons are required to have calorimeter clusters are matched to tracks reconstructed from a fast tracking algorithm inside the RoI only. In the precision reconstruction step, precision tracks within the RoI are extrapolated to the second layer of the electromagnetic calorimeter, and are required to match clusters within $\Delta\eta < 0.05$ and $\Delta\phi < 0.05$ radians. Then electron candidates are selected using a multivariate likelihood discriminant giving four operating points, with varying signal to background rejection criteria similar to the offline criteria discussed in Section 4.2.

The electrons that are selected from the full (L1 & HLT) trigger chain are required to have a E_{T} threshold of 24 GeV for 2015, and 26 GeV for 2016 to 2018. The trigger efficiencies are typically around 70% at the E_{T} thresholds, and reach more than[10] 90% for $E_{\mathrm{T}} > 50$ GeV.

C.4 Single muon trigger

In the L1 muon trigger, the p_{T} of the muon is estimated from the degree of deviation of hit patterns in the muon chambers compared to a muon of infinite momentum, based on six programmable p_{T} thresholds [78]. The number of muons passing each threshold is used in the conditions for the global L1 trigger. For the events which are accepted by L1, the p_{T} thresholds and the corresponding ROIs are sent to the HLT. The ROIs have a typical dimension of 0.1×0.1 (0.03×0.03) in $\Delta\eta \times \Delta\phi$ for the RPCs (TGCs).

The geometrical coverage of the L1 trigger is almost 99% in the end-cap region, but around 80% in the barrel region is due to gaps around the $\eta = 0$ region. The L1 trigger decision in the barrel region is based on the coincidence of hits from three (two) concentric RPC stations, for the three high (low) p_{T} thresholds. However, during the Run 2 data collection period, only the three high-p_{T} triggers were used for collecting single-muon signatures. The L1 trigger decision on the end-cap region is based on the coincidence of hits in the TGC stations in the middle layer (i.e the Big Wheel). Since the main source of trigger background in the L1 muon end-cap system are from low momentum charged particles emerging from the end-cap toroid magnets and beam shielding, a coincidence with the TGC-FI was added from 2015.

The muon HLT selects events in two stages similar to the electron HLT; fast reconstruction algorithms are executed in the first stage, followed by muon algorithms similar to the offline reconstruction. The RoIs identified by the L1 trigger enables the fast algorithms to select regions of the detector where interesting features reside. The muon standard-alone (SA) algorithms refines the L1 candidate by constructing the muon track using MDT hit information within the RoI. Following this step, the muon track (which is based only on the muon spectrometer so far), is back-extrapolated to the interaction point, and statically combined with tracks reconstructed from the inner detector to form a combined (CB) muon candidate. If no CB muons are formed, muon candidates are searched for by extrapolating ID tracks to the muon detectors. If there are corresponding track segments, CB muons are formed. Additionally, an isolation criteria as defined by a cut on the ratio of the summed transverse momentum of the nearby additional tracks, to the p_{T} of the muon (Equation C.3) is used,

$$\sum_{\Delta z < 6(2)\ \mathrm{mm}} \frac{p_{\mathrm{T}}^{\mathrm{trk}}}{p_{\mathrm{T}}^{\mu}} < \mathrm{cut}, \tag{C.3}$$

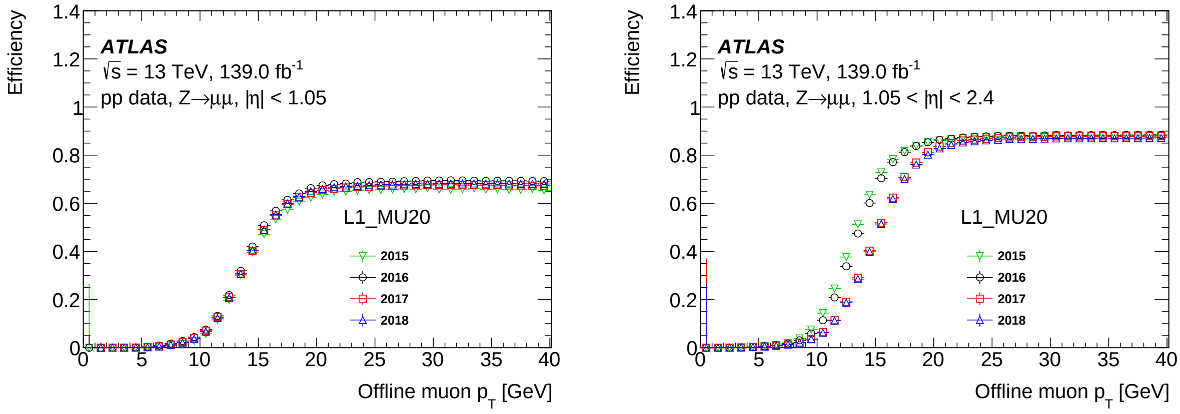
with the cut value changing based on the working point; for the Medium working point for triggers in the current analysis, the cut value was 6% in 2015, and 7% from 2016.

---

[10] 2016 is an exception, where the trigger efficiencies are slightly lower.

The default L1 single-muon triggers used to collect the data used in the analysis were unprescaled, and had a p_{T} threshold of 20 GeV. At the HLT, these triggers require at least one CB muon with $p_{\mathrm{T}} \geq 26$ GeV, or $p_{\mathrm{T}} \geq 50$ GeV. For the trigger requiring a threshold of 26 GeV, a medium isolation was also required. However, in 2015 another trigger with a L1 p_{T} threshold of 15 GeV, and HLT p_{T} threshold of 20 GeV was also used to collect data.

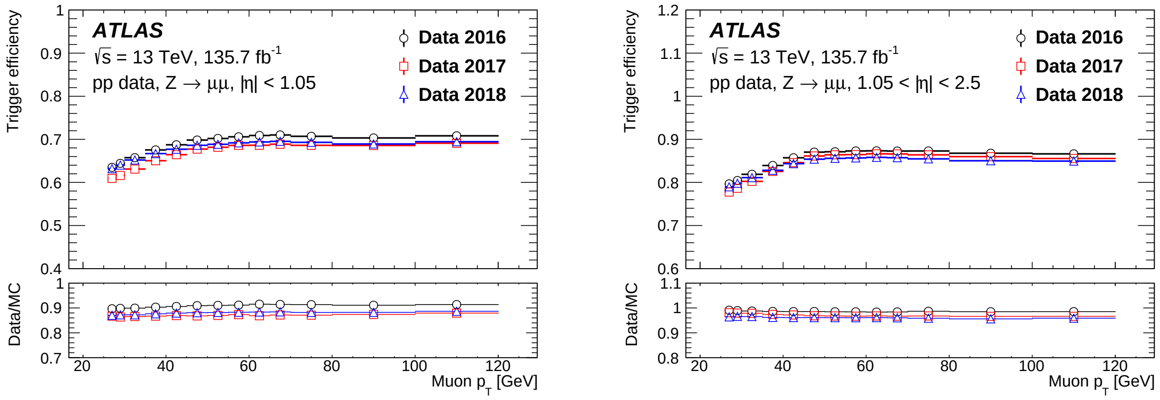
The performance of the ATLAS muon trigger is evaluated primarily using $Z \rightarrow \mu\mu$ events, using the tag-and-probe method. The L1 trigger efficiency in the barrel and end-cap regions are shown in Figure C.3, and the combined trigger (L1 & HLT) efficiency for the barrel and end-cap is shown in Figure C.4.



(**a**) L1 muon trigger efficiency: Barrel.

(**b**) L1 muon trigger efficiency: End-caps.

Figure C.3: The L1 single-muon trigger efficiency as a function of the muon p_{T} [78]. Errors bars show only the statistical uncertainties.



(**a**) L1 muon trigger efficiency: Barrel.

(**b**) L1 muon trigger efficiency: End-caps.

Figure C.4: The combined single-muon trigger efficiency as a function of the muon p_{T} [78]. Errors bars show only the statistical uncertainties.

D Object reconstruction

D.1 Electrons

The reconstruction of electrons involves information from the inner detector, the electromagnetic calorimeter (ECal), and the hadron calorimeter (HCal). The energy of the electromagnetic clusters are reconstructed from energies in the ECal cells. The electromagnetic *topo-clusters* are clusters of energy deposits in topologically connected ECal and HCal cells, used in the reconstruction of electrons and photons. Topo-clusters are required to have an electromagnetic energy above 400 MeV, and more than 50% of the total cluster energy to be the electromagnetic energy.

Electron reconstruction

The electron reconstruction algorithm first prepares tracks and topo-clusters it will use for reconstruction. The topo-clusters are matched to one or several inner detector tracks, and the algorithm refits the track-cluster matching accounting for bremsstrahlung. The matched tracks are required to satisfy a $|d_0|/\sigma_{d_0} < 5$, and $|z_0|/\sin\theta < 0.5$ mm, where σ_{d_0} is the uncertainty in d_0.

Topo-clusters are used to form dynamic, variable-size clusters of energy deposits in the ECal, known as *superclusters*. Dynamic clusters are useful over fixed-size clusters (used in ATLAS before 2017), since dynamic clusters can change in size to recover energy from bremsstrahlung photons or from electron-photon conversions. Further, the energy calibration for the dynamic clustering algorithm can still achieve a similar linear energy response to fixed-sized clustering algorithms. More details on electron reconstruction can be found in Appendix D.1, and [114].

The basis of superclusters are formed by *seed cluster candidates*, and topo-clusters near the seed cluster candidates are identified as *satellite cluster candidates*. Satellite cluster candidates may emerge from bremsstrahlung or topo-cluster splitting. Figure D.1 shows an illustration of the superclustering which is explained below, and more details can be found at [114].

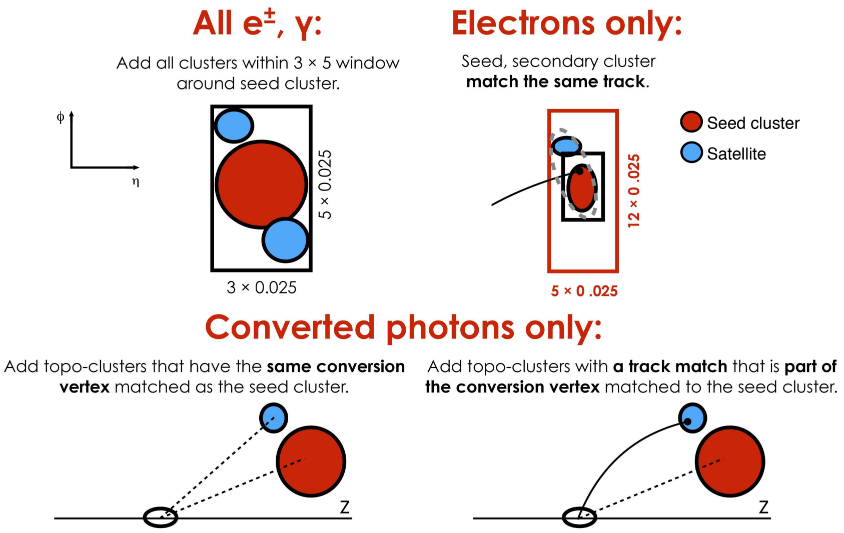


Figure D.1: The diagram illustrates the superclustering algorithm for electrons and photons [114].

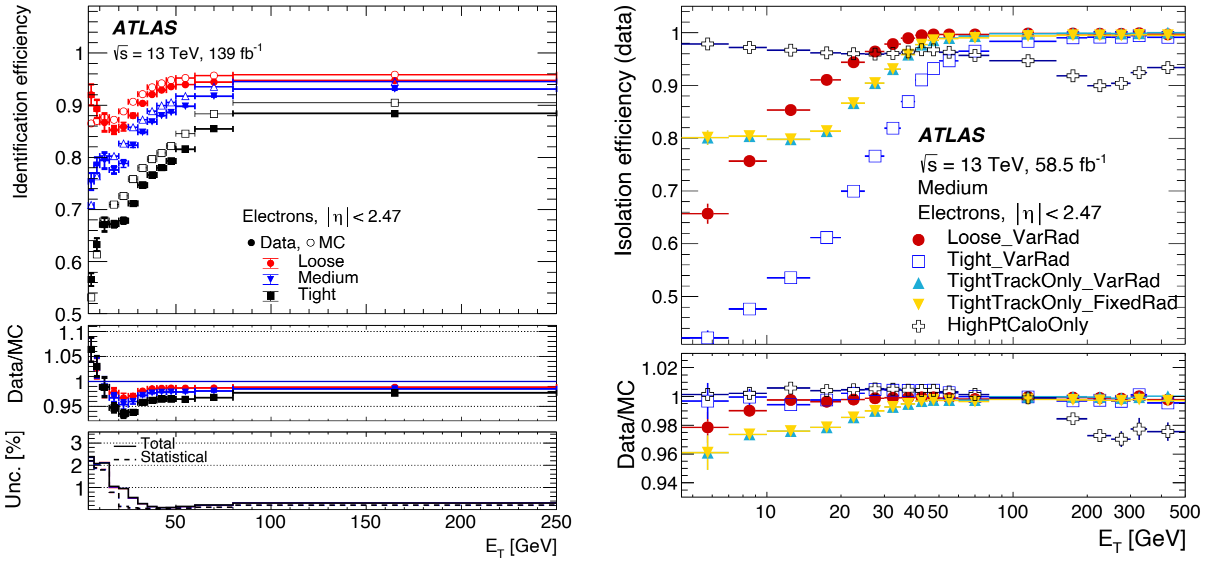
To select the seed cluster candidate, the initial set of selected topo-clusters are sorted descending in the electron transverse energy E_T (calculated using the electromagnetic energy). Then each topo-cluster is tested in the sorted order one by one to see if it can be used as a seed cluster. For a topo-cluster to become electron seed cluster, it should have a E_T >1 GeV, and must be matched to a track with at least four hits in the pixel+SCT detectors.

For both electrons and photons, topo-clusters within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycentre are identified as satellite cluster candidates. For electrons, a cluster is also considered a satellite if it falls within a window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ around the seed cluster barycentre. Satellite clusters are added to the seed cluster candidate if they satisfy the necessary selection criteria,

Final step in superclustering is to assign calorimeter cells to the supercluster. Only the cells from the pre-sampler, and the first three LAr calorimeter layers are considered in this step. However, in the transition region $1.4 < \eta < 1.6$ between the ECal barrel and endcap, the energy measured in the scintillator between calorimeter cryostats is also added.

Electron identification and isolation

The reconstructed electron candidates are passed through several quality criteria known as *identification* selections, in order to improve the purity of the prompt electrons coming from the primary vertex. Prompt electrons are identified using a likelihood based method, which uses quantities measured in the inner detector, the calorimeter, and the combined inner detector and calorimeter. Such quantities are chosen according to their ability to discriminate prompt isolated electrons from fake electrons; hadronic jets, converted photons, or electrons from heavy flavour hadron decays.



(a) Identification efficiency as a function of E_T.

(b) Isolation efficiency as a function of E_T.

Figure D.2: The identification efficiency and isolation efficiency for electron candidates. The identification and isolation efficiencies are measured using a large sample of $Z \rightarrow ee$ events taken using single-electron and di-electron triggers, with the Loose or Very Loose identification criterion [118].

In ATLAS, three different identification operating points are defined to be used for analyses;

Tight, Medium, and Loose. The identification efficiencies in the energy range of 20 GeV< E_T <50 GeV are 93%, 87%, and 79% for the Loose, Medium and Tight operating points, and gradually increase from low to high E_T, as seen in Figure D.2(a). Despite the reduced efficiency for Medium and Tight operating points compared to the Loose operating point, they have a better rejection of background processes; a factor of 1.9 (3.1) better background rejection for Medium (Tight) operating points with respect to the Loose operating point [118].

Non-prompt electrons, which are not from the primary vertex, are emitted by other physics objects like semi-leptonically decaying heavy hadrons in jets. Non-prompt electrons are not rejected by the electron identification selections, and hence, an *isolation* selection is applied to separate prompt electrons from and non-prompt electrons. The isolation can be applied based on the energy deposits in the calorimeter (known as *calorimeter isolation*), or using tracks of nearby charged particles (known as *track isolation*), or from a combination of both the calorimeter energy deposits and tracks. For calorimeter isolation, the energy around the electromagnetic shower is determined using a cone of $\Delta R = 0.2$ around the electron cluster barycenter (i.e the energy-weighted central point), and labeled as E_T^{cone20}. For track isolation, a variable size cone is defined with a given ΔR around the electron track, which contains the scalar sum of the transverse momenta of all tracks (excluding the electron track) associated with the primary vertex. Depending on the isolation criteria, the $\Delta R = 0.2$ (labeled p_T^{cone20}), or $\Delta R = \min(10\ \mathrm{GeV}/p_T, 0.3)$ (labeled $p_T^{\mathrm{varcone30}}$), is computed from tracks with $p_T > 1$ GeV and $|\eta| < 2.5$.

The ATLAS experiment defines the isolation criteria listed in Table D.1, with fixed requirements on the calorimeter and track isolation variables [118]. The isolation efficiency as a function of E_T is as seen in Figure D.2(b).

Table D.1: The electron isolation criteria defined by ATLAS. There are three more categories; *Gradient*, *TightTrackOnly* and *TightTrackOnly_FixedRad*, which are not shown in the table but more information can be found in [119].

Criteria	**Calorimeter isolation**	**Track isolation**
HighPtCaloOnly	$E_T^{\mathrm{cone20}} < \max(0.015 \times p_T, 3.5)$ GeV	-
Tight_VarRad	$E_T^{\mathrm{cone20}}/p_T < 0.06$	$p_T^{\mathrm{varcone30}}/p_T < 0.06$
Loose_VarRad	$E_T^{\mathrm{cone20}}/p_T < 0.2$	$p_T^{\mathrm{varcone30}}/p_T < 0.15$

Electron calibration

The energy of an electron candidate is built by combining the energy of the cells in the presampler and the first three electromagnetic calorimeter layers[11] assigned to the supercluster. The electron energy is calibrated as described in [119] using a boosted decision tree regression algorithm.

The difference in *energy scale* between the data and simulation is defined as α_i, and the mismodelling in *energy resolution* is parameterised as constant term c_i, where i corresponds to differ-

---

[11] However, in the transition region $1.4 < \eta < 1.6$ between the ECal barrel and endcap, the energy measured in the scintillator between calorimeter cryostats is also added.

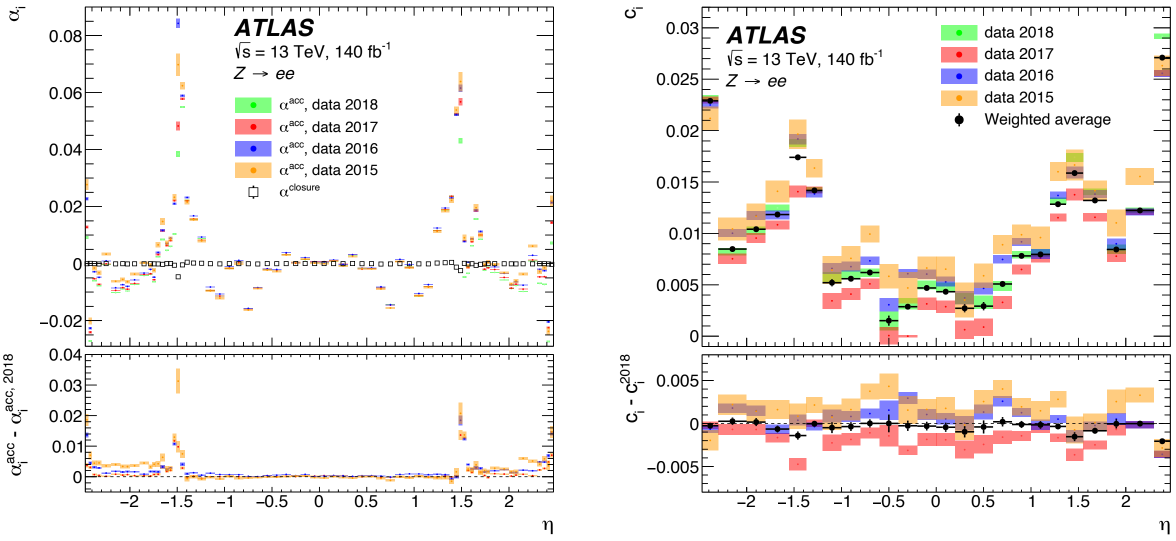
ent η regions. Electron energy scale corrections are applied to data, while the energy resolution corrections are applied to the simulated samples as given in Equation D.1 and Equation D.2.

$$E^{\text{data,corr}} = E^{\text{data}}/(1+\alpha_i) \tag{D.1}$$

$$\left(\frac{\sigma_E}{E}\right)^{\text{MC,corr}} = \left(\frac{\sigma_E}{E}\right)^{\text{MC}} \oplus c_i \tag{D.2}$$

with $\oplus$ denoting the sum in quadrature. To provide a correct description of the ECal response as a function of E_T, first the longitudinal layers are calibrated separately since the layers are segmented in depth. After this calibration step, a simulation-based calibration is applied identically to cluster energies reconstructed from data, and simulated events. Then a set of additional corrections are applied to data to account for energy response variations in specific regions (like non-optimal high voltage in regions, etc.) not included in the simulation. A final correction for the calorimeter response is derived from $Z \to ee$ events, such that the peak of the Z resonance reconstructed from data coincides with that of the simulation.

Figure D.3 shows the energy scale and energy resolution corrections as a function of η.



(a) Energy scale factors as a function of η. (b) Energy resolution term c_i as a function of η.

Figure D.3: The energy scale calibration factors α_i, and the energy resolution term c_i [119].

D.2 Photons

In ATLAS, there are two types of photon objects that are reconstructed; converted photons and unconverted photons. Photons which convert to electrons in the inner detector, around 20% at low $|\eta|$ to about 65% at $|\eta| \approx 2.3$, are labeled as converted photons. Similar to electrons, superclusters in the ATLAS electromagnetic calorimeter are used in the reconstruction of photons. Converted photon objects are superclusters matched to a conversion vertex, and unconverted photons are superclusters matched neither to an electron track nor a conversion vertex. The building of the photon superclusters proceeds independently, but in similar steps to building electron superclusters.

Photons are selected using a set of cuts on the calorimeter variables, in the range $|\eta| < 2.37$, with the transition region $1.37 < |\eta| < 1.52$ between the ECal barrel and the end-cap excluded.

D.3 Muons

Muon candidates are reconstructed primarily based on information from the inner detector and the muon spectrometer (MS) tracking detectors. Information from the calorimeters is also used in determination of track parameters, to account for cases of large energy loss in calorimeters, and for muon spectrometer independent selection of muon candidate inner detector tracks [120].

Muon reconstruction

Muon reconstruction starts with the short straight-line[12] local track segments reconstructed from hits in individual muon spectrometer stations[120]. Segments in the different stations are combined into preliminary track candidates, and then information from precision measurements in the bending plane (from MS tracking detectors) are combined with a second coordinate (from MS trigger detectors) to obtain a 3D track candidates. Finally, a global χ^2 fits the muon trajectory through the magnetic field, taking into account the muon interactions with the detector material and effect of possible misalignment between detector chambers.

Using the muon trajectory from the global χ^2 fit, outlier hits are removed and hits along the trajectory (originally not assigned to the trajectory) are added. The χ^2 fit is repeated using the the updated hit information. For tracks that share a large fraction of hits with each other, only the lower-quality track is rejected. The final set of tracks are again re-fitted with a loose interaction point constraint, taking into account also the energy loss in the calorimeters, and the transverse momentum p_T of the muon candidate at the interaction point is extrapolated from the fit.

The reconstruction of muons using the full detector information is done next, following five main reconstruction strategies that lead to different *types* of muons.

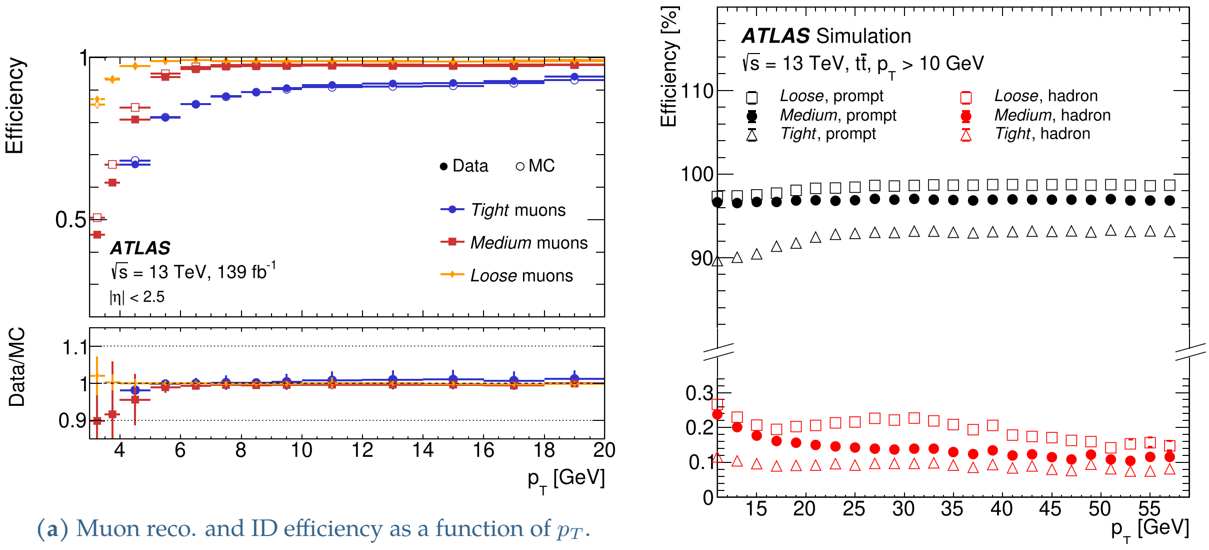
- **Combined muons** are reconstructed by matching MS tracks to inner detector tracks, and then performing a combined track fit based on the inner detector and MS hits, while also accounting for the energy loss in the calorimeters.
- **Inside-out muons** are reconstructed using an algorithm which extrapolates inner detector tracks to the MS, and searches for ≥ 3 loosely aligned MS hits. The inner detector hits, the energy loss in the calorimeters, and the MS hits are then used in a combined track fit.
- **MS extrapolated muons** are reconstructed when a MS track cannot be matched to an inner detector track, and the muon candidate parameters are extrapolated to the beam line. Such muons are used to extend the acceptance to $|\eta| < 2.7$, beyond the inner detectors acceptance of $|\eta| < 2.5$.

---

[12] Muons are straight lines in the $r - \phi$ plane, since the toroid magnet bends muons in the $r - z$ plane.

- **Segment tagged muons** are reconstructed by extrapolating an inner detector track to the MS, and requiring the a tight angular matching to at least one reconstructed MS segment. A successfully matched inner detector track is identified as a muon candidate, and the muon parameters are taken directly from the inner detector track fit.
- **Calorimeter tagged muons** are reconstructed by extrapolating inner detector tracks to calorimeter energy deposits consistent with a minimum-ionizing particle. Such energy deposits are used to tag the inner detector track as a muon, and the muon parameters are taken directly from the inner detector track fit.

While muon reconstruction algorithms use inner detector tracks with p_T down to 2 GeV, a p_T threshold of 5 GeV is required for *calorimeter tagged muons* to suppress a large background contamination at low p_T. Figure D.4(a) shows the reconstruction efficiency for muons identified in the isolation criteria described in Table D.2.



(a) Muon reco. and ID efficiency as a function of p_T.

(b) Muon ID efficiency as a function of p_T.

Figure D.4: The reconstruction and identification efficiency for muon candidates measured using $J/\psi \to \mu\mu$ events, and the simulated ID efficiency for prompt muons and hadrons as a function of p_T [120].

Muon identification and isolation

Similar to electrons, muons are also required to pass a set of quality criteria referred to as $identification$ criteria to increase the purity of high-quality muon candidates. The muon identification is done through a set of requirements, known as *selection working points*, on the number of hits in the different inner detectors and MS stations, on the track fit properties, and on variables testing the compatibility of the individual measurements in the two detectors systems. Several working points, as described in Table D.2 are defined to be used by analyses having different requirements on the prompt muon[13] identification and rejection of the non-prompt muon background. Further selection requirements on the vertex association for muon candidates are done to reject the non-prompt muon background; $|d_0|/\sigma(d_0) < 3$ and $|z_0| \sin\theta < 0.5$ mm.

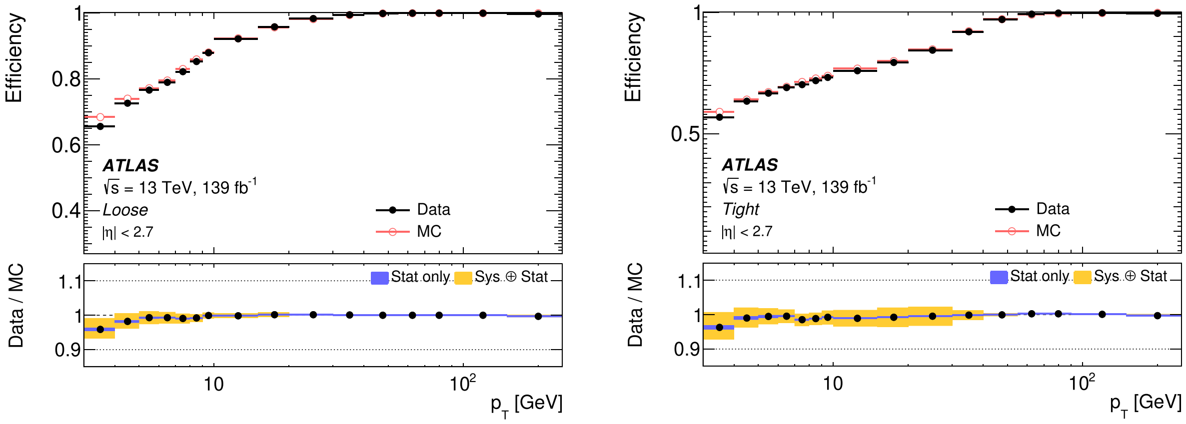
---

[13] Similar to prompt electrons, prompt muons are muons coming from the primary vertex, and non-prompt muons are from hadron decays in flight, pile-up interactions, and cosmic muons.

Table D.2: The muon identification efficiencies on the different selection working points [120].

Selection WP	$3 < p_T(\text{GeV}) < 5$		$5 < p_T(\text{GeV}) < 20$		$20 < p_T(\text{GeV}) < 100$		$p_T(\text{GeV}) > 100$	
	$\epsilon_\mu(\%)$	$\epsilon_{had}(\%)$	$\epsilon_\mu(\%)$	$\epsilon_{had}(\%)$	$\epsilon_\mu(\%)$	$\epsilon_{had}(\%)$	$\epsilon_\mu(\%)$	$\epsilon_{had}(\%)$
Loose	90	1.17	98	1.06	99	0.25	98	0.12
Medium	70	0.63	97	0.85	97	0.17	97	0.07
Tight	36	0.15	90	0.38	93	0.12	93	0.04
Low p_T (CB)	86	0.82	95	0.71	97	0.17	97	0.07
Low p_T (MV)	88	0.73	96	0.66	97	0.17	97	0.07
High p_T	45	0.34	79	0.60	80	0.13	80	0.05

The muon isolation criteria, similar to electrons, aims to discriminate prompt muons from non-prompt muons coming from hadronic sources by measuring the amount of hadronic activity in the vicinity of a muon. The muon isolation is defined by the transverse energy (or transverse momentum if only tracks are considered) reconstructed in a cone around a muon, divided by the muon p_T. The isolation can be defined using only the inner detector tracks, or using only the calorimeter, or using both detector systems with particle flow (explained in Section D.6). Several isolation working points shown in Table D.3 are defined in ATLAS for different use cases in analyses. Figure D.5 shows the p_T trend in the isolation efficiency for two of the working points.



(a) Muon isolation efficiency for the Loose WP. (b) Muon isolation efficiency for the Tight WP.

Figure D.5: The isolation efficiency for muon candidates measured using $Z \to \mu\mu$ events as a function in p_T [120].

Muon track based isolation is similar to electrons, where a variable size cone is defined with a given ΔR around the muon, which contains the scalar sum of the transverse momenta of all tracks (excluding the muon track) associated with the primary vertex. Calorimeter based isolation (labeled $E_T^{\text{topoetcone20}}$) is also considered for some criteria, and is defined as the sum if the traverse energy of the topo-clusters in a cone of size $\Delta R = 0.2$ around the muon, after subtracting the contribution of the muon itself.

The efficiency for prompt muons (ϵ_μ) and the efficiency for muons from bottom/charm hadron

Table D.3: The muon isolation criteria defined by ATLAS. There are four more categories *PflowLoose*, *PflowTight*, *PLBDTLoose*, and *PLBDTTight* which are not shown in the table, but described in [120].

Criteria	**Isolation requirement**	**Track p_T requirement**
Loose	$p_T^{\text{varcone30}} < 0.15 \cdot p_T^{\mu}$ $E_T^{\text{topoetcone20}} < 0.3 \cdot p_T^{\mu}$	$p_T > 1$ GeV
Tight	$p_T^{\text{varcone30}} < 0.04 \cdot p_T^{\mu}$ $E_T^{\text{topoetcone20}} < 0.15 \cdot p_T^{\mu}$	
HighPtTrackOnly	$p_T^{\text{cone20}} < 1.25$ GeV	$p_T > 1$ GeV
TightTrackOnly	$p_T^{\text{varcone30}} < 0.06 \cdot p_T^{\mu}$	

semi-leptonic decays (ϵ_{HF}) in the corresponding isolation working points of Table D.3 are given in Table D.4.

Table D.4: The muon isolation efficiencies of the different isolation working points [120].

Selection WP	$3 < p_T(\text{GeV}) < 5$		$5 < p_T(\text{GeV}) < 20$		$20 < p_T(\text{GeV}) < 100$		$p_T(\text{GeV}) > 100$	
	$\epsilon_\mu(\%)$	$\epsilon_{\text{HF}}(\%)$	$\epsilon_\mu(\%)$	$\epsilon_{\text{HF}}(\%)$	$\epsilon_\mu(\%)$	$\epsilon_{\text{HF}}(\%)$	$\epsilon_\mu(\%)$	$\epsilon_{\text{HF}}(\%)$
Loose	63	14.3	86	7.2	97	6.1	99	12.7
Tight	53	11.9	70	4.2	89	1.0	98	1.6
PflowLoose	62	12.9	86	6.8	97	5.0	99	9.1
PflowTight	45	8.5	63	3.1	87	0.9	97	0.8
HighPtTrackOnly	92	35.9	92	17.2	92	4.5	92	0.6
TightTrackOnly	80	19.9	81	7.0	94	3.2	99	3.3

D.4 Muon calibration

The first step in calibrating the muon momentum is to measure charge-dependant bias in the scale of muon momentum measurement, which is caused by the imperfect knowledge on the real detector geometry. The charge-dependant biases can be approximated as,

$$\frac{q}{\hat{p}} = \frac{q}{p} + q \cdot \delta_s, \tag{D.3}$$

where $q = \pm 1$ is the charge of the muon, p is the corrected momentum of the muon, $\hat{p}$ is the uncorrected momentum of the muon, and δ_s is the strength of the bias. The bias is parameterized in 24×24 equal $\eta - \phi$ regions as $\delta_s(\eta, \phi)$, and measured using $Z \to \mu\mu$ events. Then a dedicated correction to the bias is applied to correct the muon momentum following the equation,

$$p_T = \frac{\hat{p}_T}{1 - q\hat{\delta}_s(\eta, \phi)\hat{p}_T}. \tag{D.4}$$

Next, $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events are used to measure the muon momentum resolution and scale, and compared with that predicted by the simulation. Then a calibration procedure is applied to the simulated events in order to improve the agreement between the simulation and data. The validation of the calibration procedure is done using $J/\psi \to \mu\mu$ and $Z \to \mu\mu$, while also utilizing an independent $\Upsilon \to \mu\mu$ sample [121].

The muon momentum calibration is performed according to Equation D.5,

$$p_T^{\mathrm{Cor}} = \frac{p_T^{MC} + \sum_{n=0}^{1} s_n(\eta, \phi) \left(p_T^{MC}\right)^n}{1 + \sum_{m=0}^{2} \Delta r_m(\eta, \phi) \left(p_T^{MC}\right)^{m-1} g_m}, \tag{D.5}$$

where p_T^{MC} is the uncorrected muon transverse momentum, g_m are normally distributed random variables, and terms $s_n(\eta, \phi)$ and $\Delta r_m(\eta, \phi)$ describe the muon momentum scale and momentum resolution corrections. The term $s_0(\eta, \phi)$ corrects the energy loss effects in the detector, and $s_1(\eta, \phi)$ corrects the inaccuracy in the magnetic field description. The denominator of Equation D.5 accounts for momentum resolution smearing on the relative muon p_T, which can be parameterised as,

$$\frac{\sigma(p_T)}{p_T} = \frac{r_0}{p_T} \oplus r_1 \oplus r_2 \cdot p_T, \tag{D.6}$$

with $\oplus$ denoting the sum in quadrature. r_0 accounts for the fluctuations of energy loss in traversed material, r_1 accounts for the effects such as multiple scattering or local magnetic field modeling, and r_2 accounts for effects such as spatial resolution of hits or detector misalignments. The corrections are derived in 18 pseudo-rapidity η regions as described in [121].

D.5 Taus

τ leptons decay either leptonically as $\tau \to \nu_\tau \nu_l l$, or hadronically as $\tau \to \nu_\tau + hadrons$ after a decay length of 87 um. Hence, due to the short decay length, most τ leptons decay before reaching the innermost layer of the pixel detector system, and only the decay products of the τ leptons can be observed. Leptonic decays will produce isolated electrons or muons, with a missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) due to neutrinos. Hadronic decays, accounting to 65 % of the τ decays, will mostly include one charged pion (one prong) or three charged pions (three prong). Therefore, a hadronic jet from a τ lepton decay will have a signature corresponding to that of a narrow jet with one or three tracks in the inner detector. Since the neutrino from the hadronic τ lepton decay cannot be reconstructed, the combination of all the visible τ lepton decay products is referred to as $\tau_{\mathrm{had-vis}}$.

$\tau_{\mathrm{had-vis}}$ candidates are reconstructed as jets formed using the anti-k_t algorithm, with the distance parameter $R = 0.4$. 3D *topo-clusters* calibrated using a local hadronic calibration serves as an input to the anti-k_t algorithm [128]. The jet axis is defined as the energy-weighted barycenter of all topo-clusters, while the vertex candidate with the highest p_T-weighted fraction of all tracks with $p_T > 500$ MeV within a cone of $R = 0.2$ around the $\tau_{\mathrm{had-vis}}$ jet axis is taken as the

$\tau_{\mathrm{had-vis}}$ vertex. A set of boosted decision trees (BDTs)[14] are used to classify the tracks within $R = 0.4$ around the τ axis into *core* and *isolation* tracks, with the core tracks used to define the number of prongs. Finally, $\tau_{\mathrm{had-vis}}$ candidates with $p_T > 10\,\mathrm{GeV}$, and within the $|\eta| < 2.5$ region are selected, with candidates in the $1.37 < |\eta| < 1.52$ region rejected due to the transition region between the barrel and end-cap calorimeters.

The energy calibration of $\tau_{\mathrm{had-vis}}$ candidates follow the procedure described in [149]. Initially a pile-up subtraction and an energy response correction is applied to the sum of energy of the *topo-clusters* within the $R = 0.2$ cone around the $\tau_{\mathrm{had-vis}}$ jet axis. This energy, together with information form the τ particle flow [150], is used to calculate the final energy of the $\tau_{\mathrm{had-vis}}$ candidate using a boosted regression tree.

Table D.5: The τ lepton identification efficiencies on the different selection working points [149].

Working point	**Signal efficiency**		**Background rejection**	
	1-prong	**3-prong**	**1-prong**	**3-prong**
Tight	60%	45%	70	700
Medium	75%	60%	35	240
Loose	85%	75%	21	90
Very Loose	95%	95%	9.9	16

After reconstruction and calibration, a recurrent neural network (RNN) based on track and calorimeter information is used to identify the true $\tau_{\mathrm{had-vis}}$ from fake $\tau_{\mathrm{had-vis}}$ that could originate from quark or gluon initiated jets [128]. Similar to electrons and muons, several working points are defined for the $\tau_{\mathrm{had-vis}}$ identification, and Figure D.6 shows the trend in the $\tau_{\mathrm{had-vis}}$ efficiency vs the fake $\tau_{\mathrm{had-vis}}$ rejection, while Table D.5 details the actual values.

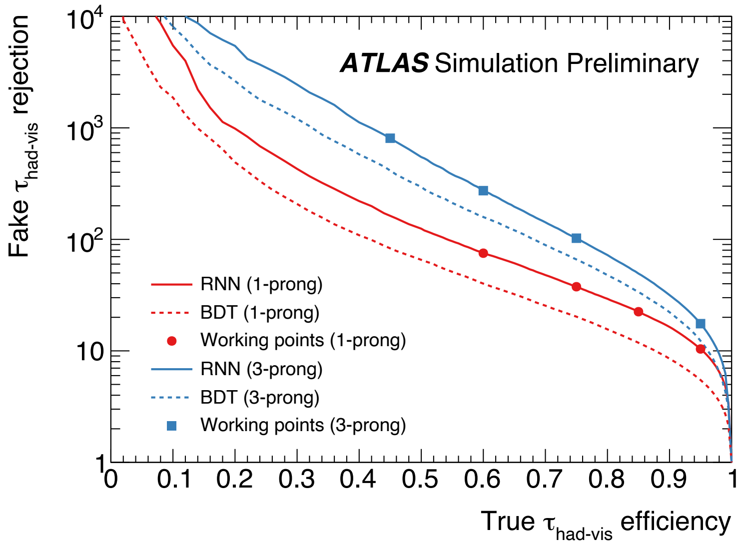


Figure D.6: The true $\tau_{\mathrm{had-vis}}$ identification efficiency vs the fake $\tau_{\mathrm{had-vis}}$ rejection using the RNN-based method (current analysis) and the BDT based method (previous analysis) [128].

---

[14] BDTs are explained in detail in MVA Section 7.1.

D.6 Jets

Object definition using the particle flow algorithm

The particle flow algorithm combines the capabilities of the inner detector in reconstructing charged particles, and the calorimeter's ability to reconstruct both charged and neutral particles. However, since the inner detector has an acceptance for $|\eta| < 2.5$, jets reconstructed in the forward region $(2.5 < |\eta| < 4.5)$ uses topo-clusters from the calorimeter to reconstruct particle flow jets.

There are several advantages of integrating the tracking and calorimeter information to reconstruct the hadronic jets. Since the calorimeter energy resolution for a single charged pion follows the Equation 2.6, while the inverse transverse momentum resolution for the tracker follows Equation 2.5, the tracker has significantly better momentum resolution for low energy particles compared to the calorimeter, while the calorimeter's energy resolution is superior at higher energies. The angular resolution for single charged particles reconstructed using inner detector tracks is also much better compared to the calorimeter.

Further, the acceptance is extended for softer (i.e low energy) particles which are essential for computation of $E_{\mathrm{T}}^{\mathrm{miss}}$, as tracks are reconstructed for charged particles with a minimum p_{T} of 500 MeV[15]. Also, such low p_{T} charged particles originating within a hadronic jet could get swept out of the jet cone due to the magnetic field, but using the tracks azimuthal coordinates at the perigee can cluster these particle into jets. In addition to the discussed benefits of using tracks from the inner detector, the ability determine whether a track originates from the primary vertex aids in rejecting signals from pile-up vertices as well.

Particle flow requires a special procedure; to avoid double counting the energy of a particle during reconstruction, the correct signal in the calorimeter needs to be identified for the tracks used for reconstruction. Therefore, for particle tracks that are determined to be used for reconstruction, the corresponding energy deposits in the calorimeter are subtracted. For all tracks up to $p_{\mathrm{T}}^{\mathrm{trk}} = 100$ GeV, if the measured cluster energy E^{clus} in a cone of size $\Delta R = 0.15$ around the extrapolated particles satisfies,

$$\frac{E^{\mathrm{clus}} - \langle E_{\mathrm{dep}} \rangle}{\sigma(E_{\mathrm{dep}})} < 33.2 \times \log_{10}\left(\frac{40\ \mathrm{GeV}}{p_{\mathrm{T}}^{\mathrm{trk}}}\right), \tag{D.7}$$

then the corresponding energy deposit in the calorimeter is subtracted. Here $\langle E_{\mathrm{dep}} \rangle$ is the mean of the energy deposited by pions. With this parameterization, the subtraction is performed at low track momenta unless the measured calorimeter activity E^{clus} is very high, such as in a very dense environment where the accuracy of the subtraction is degraded. Above $p_{\mathrm{T}}^{\mathrm{trk}} = 100$ GeV no track information is used, and the the particle flow algorithm relies on the good calorimeter performance at high energies.

After subtraction, scaling factors are applied to account for the difference in response between the measured and true particle energies; one is for topo-clusters at the EM scale, and one is for tracks for which the energy scale is closer to the true particle energy. Finally jet reconstruction

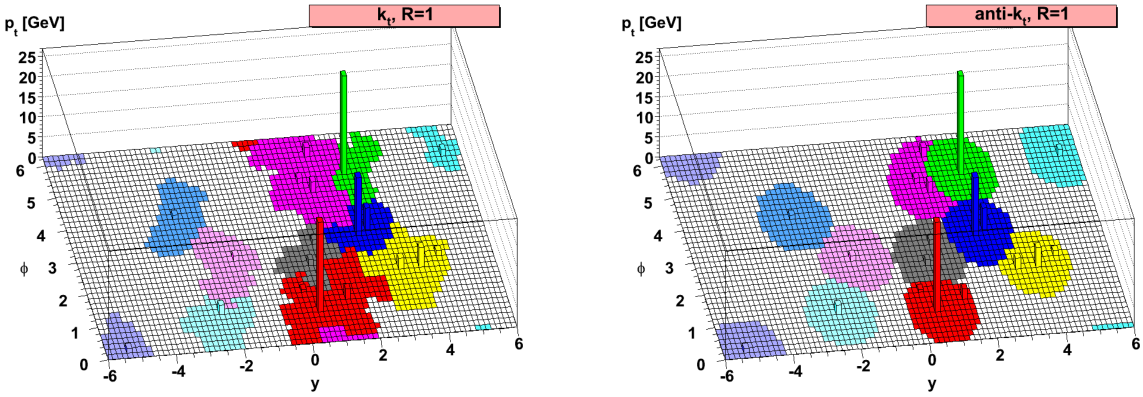
---

[15] To suppress noise, the calorimeter topo-clusters have a much higher threshold for energy deposits

is done using an ensemble of particle flow objects consisting the remaining calorimeter energy deposits and the tracks that are matched to the hard interaction.

Jet clustering using the anti-k_t algorithm

Jet reconstruction is done using the anti-k_t algorithm [122], using the four-vector objects reconstructed with the Particle Flow algorithm. The anti-k_t algorithm is an algorithm used to iteratively merge topo-clusters, which is *infrared safe* and *collinear safe*; an algorithm is deemed infrared safe if the reconstructed jet properties, such as the four-momentum, are stable in the case of emission of additional soft (i.e low energy) particles such as gluons, and collinear safe implies that the reconstructed jet properties are stable with respect to the number of emitted collinear particles.



(**a**) Example of clustering by the k_t algorithm.

(**b**) Example of clustering by the anti-k_t algorithm

Figure D.7: Example showing the comparison between the k_t algorithm (with parameter $p = 1$) and the anti-k_t algorithm (with parameter $p = -1$) [122].

The association of clusters to a jet is performed in the following procedure. For each topo-cluster i, first the distance d_{ij} to topo-cluster j is calculated using Equation D.8, and the distance d_{iB} to the beam B is calculated using Equation D.9. Then the smallets distance $\min(d_{ij})$ between clusters is found, If $\min(d_{ij}) < d_{iB}$, then the cluster j is merged into a single new cluster. If $\min(d_{ij}) > d_{iB}$, then the cluster i is considered as a jet, and it is removed from the list of clusters to be considered. The procedure is repeated with the updated list of objects until no objects are remaining.

$$d_{ij} = \min\left(k_t^{2p}(i), k_t^{2p}(j)\right) \frac{\Delta R(i,j)^2}{R^2}. \tag{D.8}$$

$$d_{ij} = k_t^{2p}(i). \tag{D.9}$$

Here k_t is simply the transverse momentum p_{T}, R is the radius parameter chosen to determine the final size of jets, $\Delta R(i,j)$ is the angular distance between the clusters, and p is a parameter that governs the relative power of the energy versus geometrical $\Delta R(i,j)$ scales. The p parameter can take various values, and $p = -1$ which translates to $k_t^{-2} \rightarrow 1/p_{\mathrm{T}}^2$ for the terms in Equation D.8 and D.9 is referred to as the **anti-k_t algorithm**.

The iterative process and the term min $\left(k_t^{-2}(i), k_t^{-2}(j)\right)$ ensures that the topo-clusters are associated to hard (high-energy) objects compared to soft objects. Figure D.7 shows an example comparing the jet clustering between the k_t algorithm (with parameter $p = 1$) and the anti-k_t algorithm, where anti-k_t is seen to have a better shape for hard jets compared to the k_t algorithm (for example, the pair of jets near $\phi = 5, y = 2$).

Jet selections to suppress pile-up and fake jets

The **Jet Vertex Tagger** (JVT) algorithm is used to suppress jets arising from pile-up interactions, while correctly identifying jets attributed to the hard-scattering vertex [123]. The JVT procedure builds a multivariate discriminant for each jet within $|\eta| < 2.4$ based on tracks *ghost-associated* with the jet; the ghost association algorithm collect tracks within a geometrical area of a jet, taking into account the possible overlap of the area between jets. In simulation, the efficiency for picking jets arriving from the primary vertex is estimated to be 92%. However, since the rate of pile-up jets with $p_{\mathrm{T}} \geq 60$ GeV is sufficiently small, the JVT requirement is not applied for jets with $p_{\mathrm{T}} \geq 60$ GeV.

Due to the lack of tracking information in the forward region (i.e $|\eta| > 2.5$), the JVT algorithm cannot be used for forward jets. Hence, another algorithm known as the **forward Jet Vertex Tagger** (fJVT) is applied for forward jets with $p_{\mathrm{T}} < 120$ GeV.

The **Jet Cleaning Tool** [151] is a tool used in ATLAS to remove fake jets (also known as unclean jets) that are due to,

- Beam induced background from proton losses upstream of the interaction point, that can induce secondary cascades of muons reaching the ATLAS detector. The energy depositions created by such muons can be reconstructed as fake jets.
- Cosmic muons reaching the ATLAS detector during collisions.
- Calorimeter noise from large scale coherent noise or isolated pathological cells. Calorimeter cells which are constantly or periodically very noisy are masked prior to the jet p_{T} or $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, and events with a large noise are rejected during data quality inspections. However, a small fraction of noisy events can still be left which needs to be cleaned using the Jet Cleaning Tool.

The jet cleaning tool defines two working points for clean jet selections as in Table D.6.

Table D.6: The efficiency in the selection working points of the jet cleaning tool used in ATLAS. The efficiencies are measured using a tag-and-probe method using di-jet events that are back-to-back in the transverse plane [151].

Working point	Selection efficiency	p_{T} range
Loose	99.5%	$p_{\mathrm{T}} > 20$ GeV
	99.9%	$p_{\mathrm{T}} > 100$ GeV
Tight	95%	$p_{\mathrm{T}} > 20$ GeV
	99.5%	$p_{\mathrm{T}} > 100$ GeV

D.7 Jet flavor tagging

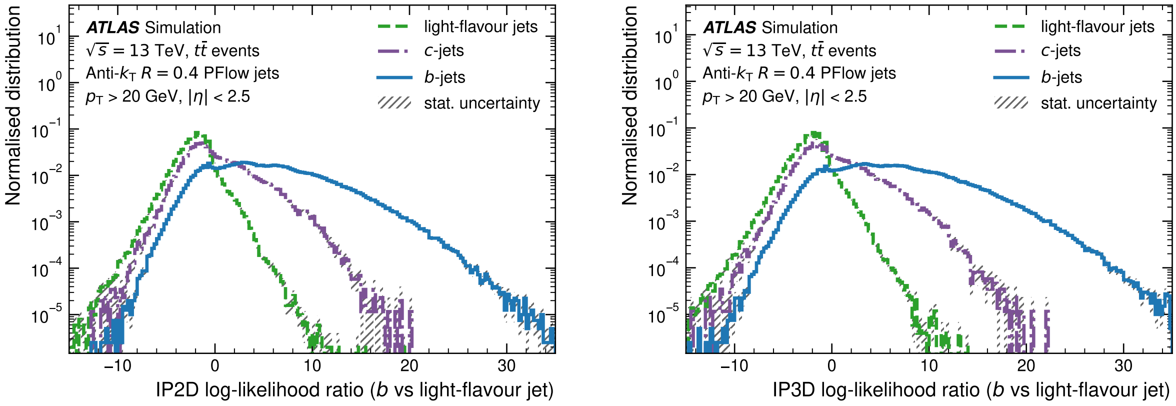
Low level flavor tagging algorithms fall into two broad categories. The first category consists of the **IP2D** and **IP3D** algorithms, and are based on exploiting the large impact parameters of the tracks from b-hadron and c-hadron decays. The **RNNIP** algorithm, a new algorithm used in the $VH, H \to b\bar{b}/c\bar{c}$ analysis, uses a recurrent neural network that learns track impact parameter correlations to further improve the jet flavor discrimination. The second category consists of the **SV1**, that attempts to reconstruct a secondary vertex inside a jet, and the JetFitter algorithms, which aims to reconstruct the full b-hadron and c-hadron decay chains.

The `IP2D`, `IP3D` and `RNNIP` algorithms

The `IP2D` tagger uses the *signed transverse impact parameter* significance (d_0/σ_{d_0}) of tracks to create a discriminant that can be used to tag b-jets. The `IP3D` tagger uses both the signed transverse impact parameter significance (d_0/σ_{d_0}) of tracks, and the longitudinal impact parameter significance (z_0/σ_{z_0}) in a 2D template to account for their correlation. Probability density functions (PDF), obtained from reference MC simulated histograms based on the signed transverse and longitudinal impact parameter significance of tracks associated with b-jet, c-jet and light-jets, are used to calculate the ratio of b-jet, c-jet and light-jet fractions on a per-track basis. Using the PDF's, log likelihood ratio (LLR) discriminants are defined as the sum of the probability ratios for each jet-flavor hypotheses. Figure D.8 shows an example of the LLR discriminant for the b-jet and light-jet hypotheses is defined as,

$$\sum_{i=1}^{N} \ln\left(p_b(S_i)/p_{\text{light}}(S_i)\right) \tag{D.10}$$

where N is the number of tracks associated with the jet, and p_b (p_{light}) is the PDF for the b-jet (light-jet) hypotheses, and S_i is the impact parameter significance of the track i. Similarly the LLR discriminants are defined for the b-jet and c-jet hypotheses, and the c-jet and light-jet hypotheses.



(a) LLR for the IP2D algorithm.

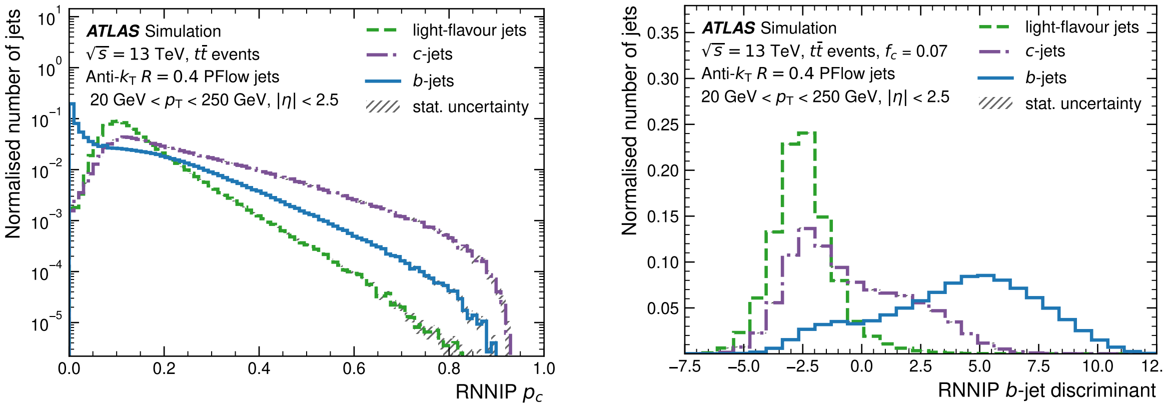
(b) LLR for the IP3D algorithm.

Figure D.8: The log likelihood ratio (LLR) discriminant for the ratio of the b-jet to light-flavour jet hypothesis probabilities using the IP2D and IP3D algorithms [36].

In heavy hadron decays, charged particles emerge from secondary or tertiary vertices with large impact parameters. Since the impact parameters of such charged particle tracks are intrinsically correlated[16], the `RNNIP` algorithm is trained to exploit the correlation in order to better discriminate between b-jets, c-jets, and light-jets [36]. The output of the RNN are the probabilities p_b, p_c, p_{light}, with the jet being a b-jet, c-jets, or light-jets, which is combined into the b-jet tagging discriminant function in Equation D.11.

$$D_{\text{RNNIP}} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c)p_{\text{light}}}\right) \tag{D.11}$$

f_c denotes the c-jet fraction, which is optimized to 0.07 for b-jet tagging by ATLAS. Figure D.9 shows the output of the RNN, and the `RNNIP` discriminant at $f_c = 0.07$. For the `RNNIP` discriminant in Figure D.9(b), it can be seen at values $D_{\text{RNNIP}} > 2.5$ a majority of the true b-jets can be tagged, while having a good light-jet rejection and moderate c-jet contamination.



(a) Output of the RNNIP algorithm for c-jets.

(b) The RNNIP discriminant D_{RNNIP} at $f_c = 0.07$.

Figure D.9: The output probabilities from the RNNIP algorithm [36].

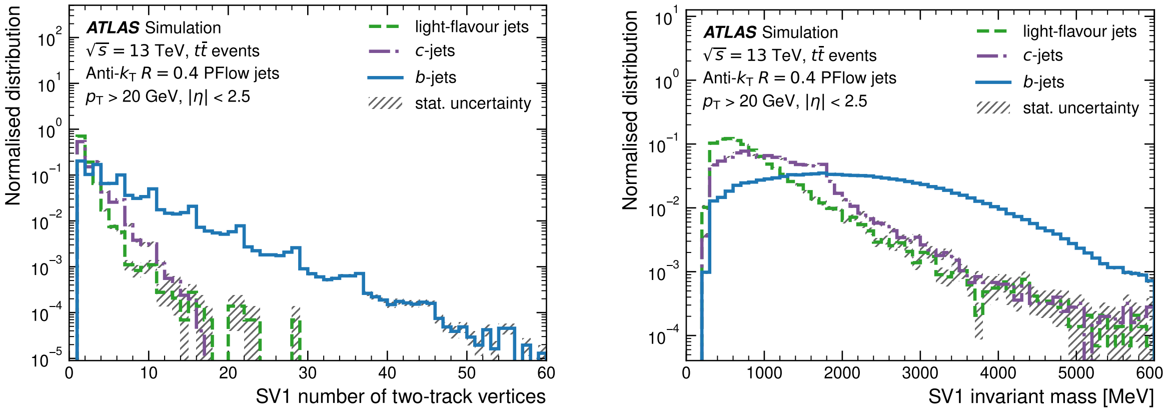
The `SV1` and JetFitter algorithms

The secondary vertex tagging algorithm (`SV1`) reconstructs a single displaced vertex in a jet. First, possible two-track vertices that can be built from tracks associated with the jet are identified, while rejecting two-track vertices compatible with K^0_S or λ meson decays, photon conversion to e^+e^-, and hadron interaction with the detector material. The `SV1` algorithm runs iteratively on all tracks of the identified two track vertices, combining each track using a χ^2 approach to fit one secondary vertex, with an invariant mass less than 6 GeV. This approach typically assigns the decay products of b-hadrons and c-hadrons to a single common secondary vertex. Finally, eight discriminating variables associated with the `SV1` algorithm, as listed in Table D.7, will be used as inputs to the high level flavor tagging algorithms, and Figure D.10 shows some example distributions of these variables.

The JetFitter, also known as the topological multi-vertex finding algorithm, exploits the topological structure of b-hadrons and c-hadrons (referred to as heavy hadrons) decays inside the

---

[16] In heavy hadron decays, if one track with large impact parameter is found, it is likely to find another track with a large impact parameter as well.

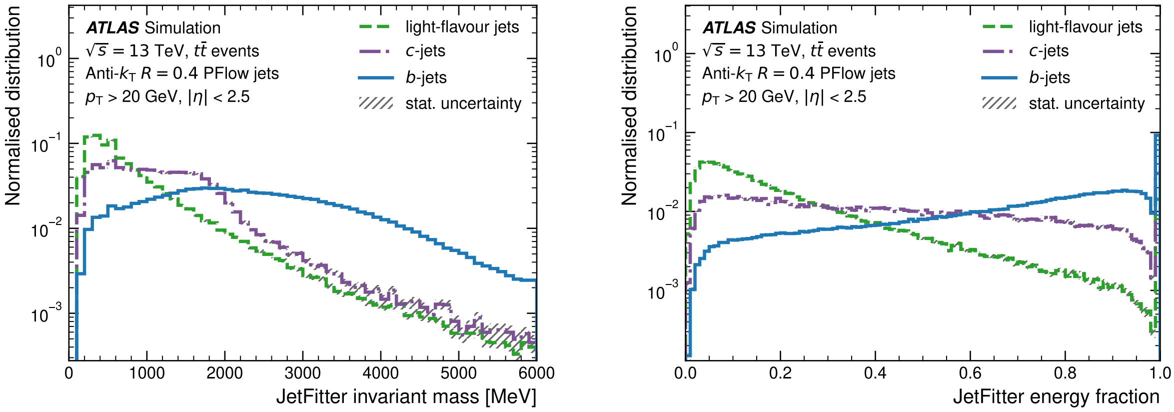


(**a**) Number of two track vertices.

(**b**) Invariant mass of the secondary vertex.

Figure D.10: Two out of the eight discriminating variables associated with the SV1 secondary vertex tagging algorithm [36].

jet to reconstruct the full decay chain [36]. A modified Kalman filter is used to find a common line on which the heavy hadron decay vertices lie, considering the heavy hadron flight path as well as the vertex positions. Through this method, it is possible to resolve the heavy hadron decay vertex even if only a single track is attributed to it. Similar to the `SV1` algorithm, there are eight b-jet discriminating variables associated with the JetFitter. However, by further exploiting c-hadron decay properties such as only a single reconstructed secondary vertex, intermediate charged decay multiplicity and moderate decay length with respect to b-hadron decays, nine c-jet discriminating variables associated with the JetFitter are defined. Altogether, these 17 variables as listed in Table D.7, serve as inputs to the high level flavor tagging algorithms. Figure D.11 shows some example distributions of some JetFitter variables.



(**a**) Number of two track vertices.

(**b**) Invariant mass of the secondary vertex.

Figure D.11: Two out of the 17 discriminating variables associated with the JetFitter topological multi-vertex finding algorithm [36].

High level flavor tagging algorithms

High level flavor tagging algorithms are based on machine learning based classifiers like boosted decision trees and neural networks. In the previous $VH, H \to c\bar{c}$ analysis, the MV2 high level flavor tagging algorithm, based on a boosted decision tree was used for jet flavor tagging [30]. The high level flavor tagging algorithm used in the current analysis comes from the $DL1$ series of algorithms, which is based on a deep feed-forward neural network. The input variables to $DL1r$ algorithm (the algorithm used in the current analysis) consists of the general variables denoted in Table D.7; the use of the RNNIP algorithm is the specialty of the $DL1r$ algorithm, where r stands for the RNNIP algorithm.

Table D.7: The input variables used by the DL1 series of b-tagging algorithms [125]. PV is an abbreviation for primary vertex, SV is an abbreviation for secondary vertex, and TV is an abbreviation for tertiary vertex. Y_{trk} is the rapidity of tracks.

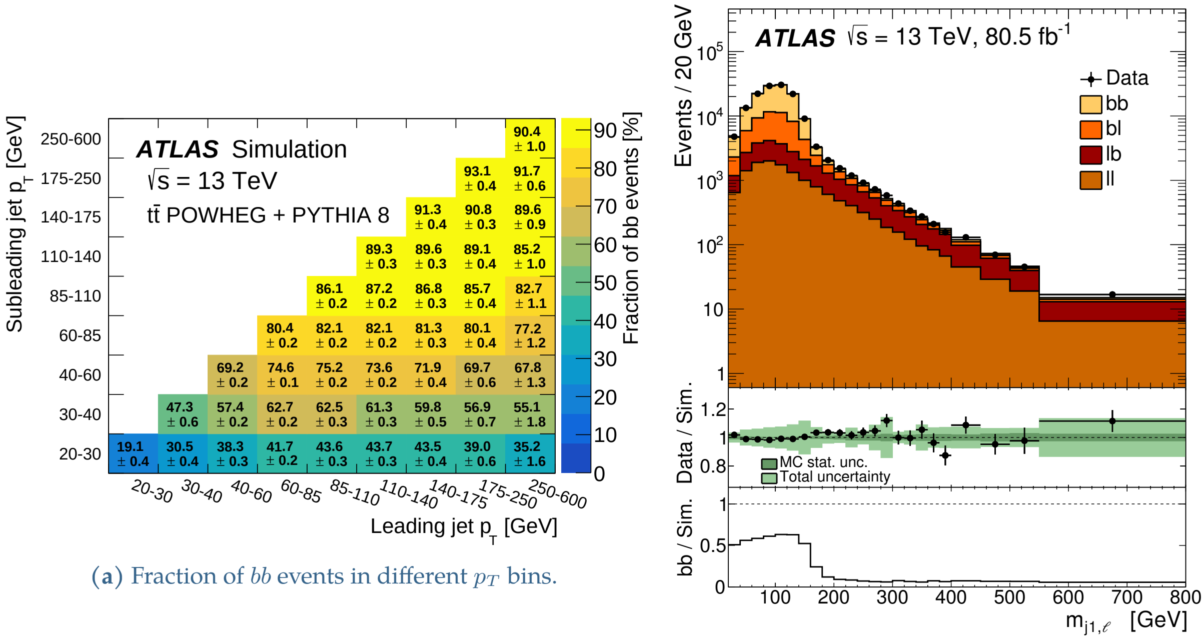
Input type	**Variable**	**Description**
Kinematics	p_T	Jet transverse momentum
	η	Jet pseudo-rapidity
IP2D, IP3D	$\log(P_b/P_{light})$	LLR between b-jet and light-jet hypothesis
	$\log(P_b/P_c)$	LLR between b-jet and c-jet hypothesis
	$\log(P_c/P_{light})$	LLR between c-jet and light-jet hypothesis
RNNIP	$\log(P_b)$	b-jet probability
	$\log(P_c)$	c-jet probability
	$\log(P_{light})$	light-flavor jet probability
SV1	$m(\text{SV})$	Invariant mass of tracks at the SV, assuming pion mass
	$f_E(\text{SV})$	Energy fraction of tracks associated with the SV
	$N_{\text{TrkAtVtx}}(\text{SV})$	Number of tracks associated with the SV
	$N_{\text{2TrkVtx}}(\text{SV})$	Number of two track vertex candidates
	$L_{xy}(\text{SV})$	Transverse distance between PV and SV
	$L_{xyz}(\text{SV})$	Distance between PV and SV
	$S_{xyz}(\text{SV})$	Distance between PV and SV, divided by uncertainty
	$\Delta R(\vec{p}_{\text{jet}}, \vec{p}_{\text{vtx}})(\text{SV})$	ΔR between jet axis & direction of SV relative to PV
JetFitter	$m(\text{JF})$	Invariant mass of tracks from displaced vertices
	$f_E(\text{JF})$	Energy fraction of tracks associated with displaced vertices
	$N_{\text{TrkAtVtx}}(\text{JF})$	Number of tracks from multi-prong displaced vertices
	$N_{\text{2TrkVtx}}(\text{JF})$	Number of two track vertices prior to decay chain fit
	$N_{1-\text{trk vertices}}(\text{JF})$	Number of single prong displaced vertices
	$N_{\geq 2-\text{trk vertices}}(\text{JF})$	Number of multi-prong displaced vertices
	$S_{xyz}(\text{JF})$	Distance between PV and SV, divided by the uncertainty
	$\Delta R(\vec{p}_{\text{jet}}, \vec{p}_{\text{vtx}})(\text{JF})$	ΔR between jet axis & direction of SV relative to PV
JetFitter c-tagging	$L_{xy}(\text{SV},\ \text{TV})(\text{JF})$	Transverse distance to the SV or TV from PV
	$L_{xyz}(\text{SV},\ \text{TV})(\text{JF})$	Distance to the SV or TV from PV
	$m(\text{SV},\ \text{TV})(\text{JF})$	Invariant mass of tracks associated with SV or TV
	$E_{\text{Trk}}(\text{SV},\ \text{TV})(\text{JF})$	Energy fraction of tracks associated with SV or TV
	$f_E(\text{SV},\ \text{TV})(\text{JF})$	Fraction of charged jet energy associated with SV or TV
	$N_{\text{TrkAtVtx}}(\text{SV},\ \text{TV})(\text{JF})$	Number of tracks associated with SV or TV
	$Y_{\text{trk}}^{\text{min}}, Y_{\text{trk}}^{\text{max}}, Y_{\text{trk}}^{\text{avg}}(\text{SV},\ \text{TV})(\text{JF})$	Min, max, and average Y_{trk} associated with SV or TV

Calibration of the high-level flavor tagging algorithms

The imperfect description of the detector response and physics modelling effects in MC simulations requires the measurement of the flavor tagging performance with real collision data. However, due to varying levels of statistics for the jet flavors b, c and light in the regions optimized for flavor tagging, each jet flavor needs to be calibrated in dedicated procedures.

b-jet tagging efficiency calibration

Given the abundance of top pair production ($t\bar{t}$) in proton-proton collisions at the LHC, and the $t \to Wb$ branching ratio is almost 100%, an abundant source of b-jets can be obtained by selecting the $t\bar{t}$ process, which then can be used to extract the b-jet tagging efficiency (ϵ_b). A very pure sample of $t\bar{t}$ is obtained through requiring the W-bosons to decay leptonically to electrons or muons plus a neutrino, referred to as the di-leptonic $t\bar{t}$ process.



Subleading jet p_T [GeV]	20-30	30-40	40-60	60-85	85-110	110-140	140-175	175-250	250-600
250-600									90.4 ± 1.0
175-250								93.1 ± 0.4	91.7 ± 0.6
140-175							91.3 ± 0.4	90.8 ± 0.3	89.6 ± 0.9
110-140						89.3 ± 0.3	89.6 ± 0.3	89.1 ± 0.4	85.2 ± 1.0
85-110					86.1 ± 0.2	87.2 ± 0.2	86.8 ± 0.3	85.7 ± 0.4	82.7 ± 1.1
60-85				80.4 ± 0.2	82.1 ± 0.2	82.1 ± 0.2	81.3 ± 0.3	80.1 ± 0.4	77.2 ± 1.2
40-60			69.2 ± 0.2	74.6 ± 0.1	75.2 ± 0.2	73.6 ± 0.2	71.9 ± 0.4	69.7 ± 0.6	67.8 ± 1.3
30-40		47.3 ± 0.6	57.4 ± 0.2	62.7 ± 0.2	62.5 ± 0.3	61.3 ± 0.3	59.8 ± 0.5	56.9 ± 0.7	55.1 ± 1.8
20-30	19.1 ± 0.4	30.5 ± 0.4	38.3 ± 0.3	41.7 ± 0.2	43.6 ± 0.3	43.7 ± 0.3	43.5 ± 0.4	39.0 ± 0.6	35.2 ± 1.6

(**a**) Fraction of bb events in different p_T bins.

(**b**) The invariant mass $m(j_1, l_1)$.

Figure D.12: The left plots show the fraction of bb events in different p_T bins of the leading and sub-leading jets. The right plot shows the invariant mass $m(j_1, l_1)$ for the leading jet in the selected events. The $m(j_2, l_2)$ for the sub-leading jet also has a similar distribution [125].

After several events selections to reduced the non-$t\bar{t}$ background, the $t\bar{t}$ events are first binned into *nine bins* according to the p_T of the leading jet (i.e jet with the highest p_T, labeled as j_1), and the p_T sub-leading jet (i.e jet with the second highest p_T, labeled as j_2), as seen in Figure D.12(a). Then in each p_T bin, events are assigned to a signal region (SR) and control regions (CR) based on a cut $m(j_1, l_1) < 175$ GeV and $m(j_2, l_2) < 175$ GeV, where $m(j_i, l_i)$ is the invariant mass of jet j_i and its associated lepton l_i; since $t\bar{t}$ events with two b-jets (labeled bb) have an upper limit around the top mass $m_t = 175$ GeV (as seen in Figure D.12(b)), the $m(j_i, l_i)$ cuts will increase the purity of b-jets in the signal region.

The combinations with non b-jets from the top decay, labeled l, are bl, lb, ll, are constrained from the CRs, which have a larger non-bb sample fraction compared to the SR. Then finally the SR events are classified into the different b-jet tagging working points of the $DL1r$ algorithm [125], to measure the ϵ_b in each $DL1r$ working point as a function of p_T through a combinatorial likelihood approach. This process is illustrated in Figure D.13.

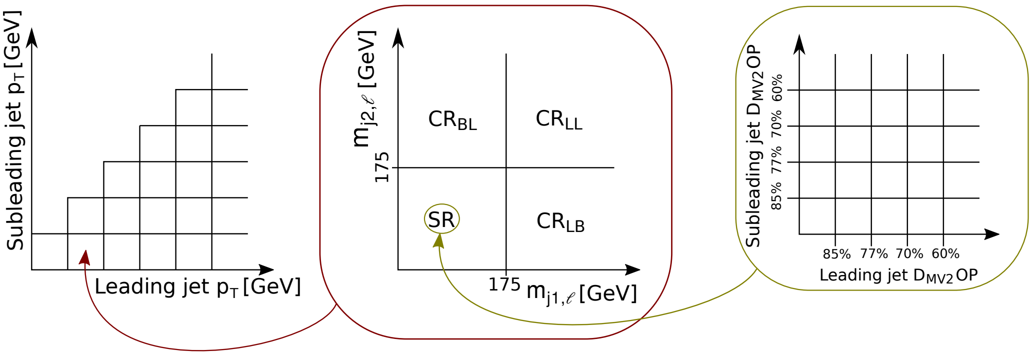


Figure D.13: An illustration showing the event categorization for measuring ϵ_b [125]

In ATLAS there are four default regions used for b-jet tagging with different ϵ_b; ϵ_b is 60%, 70%, 77%, and 85% where 70% (for example) implies that 70% of the true b-jets are estimated to be correctly tagged as b-jets in this working point.. Regions with less b-jet tagging efficiency have better c-jet and light-flavor jet rejection, making these regions ideal for obtaining pure b-jet samples. Figure 4.9 shows the data to MC correction (also known as scaling factors) for the measured b-jet tagging efficiencies ϵ_b for the four different of the $DL1$ algorithm, and the associated uncertainties.

c-jet tagging efficiency calibration

Similar to calibrating b-jet tagging efficiency, c-jet tagging efficiency calibration also uses $t\bar{t}$ events, but one of the W-bosons is required to decay hadronically, while the other W-boson is required to decay leptonically to electrons or muons plus a neutrino (referred to as the semi-leptonic $t\bar{t}$ process) [126].

First, event selections are applied to reduce the non-semi-leptonic $t\bar{t}$ processes;

- A lepton selection require a $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV (to account for the neutrino), and the transverse mass of the W-boson $m_T = \sqrt{2\, p_T^l\, E_T^{miss}(1 - \cos\Delta\phi)} > 40$ GeV.
- A baseline jet selection, requires four jets in an event to account for two b-jets from the top-decay, and two more jets from the W-boson decay. All jets are required to have a $p_T > 40$ GeV. It should be noted that there are two more jet selections for enhancing measurements in specific regions.

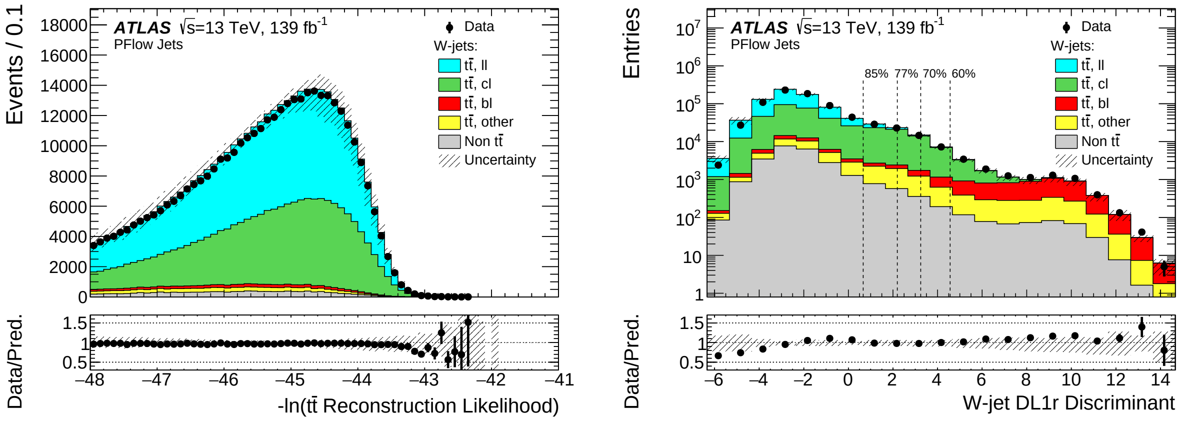
Then an algorithm, using a likelihood function, assigns the four jets to the $t\bar{t}$ decay topology; either the b-jets from the top decay, or the jets from the W-decay. Further selections on the negative likelihood value to be greater than -48, and requiring the b-jets from the top decay to be in the very tight 60% b-tagging working point[17].

After the selections, a sample that is 99.8% pure in true b-jets coming from a top-decay, with

---

[17] The working point has a $\epsilon_b = 58\%$, with very good rejection on c-jets and light-flavor jets: $\epsilon_c = 2.65\%$, $\epsilon_{\text{light}} = 0.051\%$

a good reduction in the non-$t\bar{t}$ background for the jets coming from the W-boson decay is obtained as seen in Figure D.14(a).

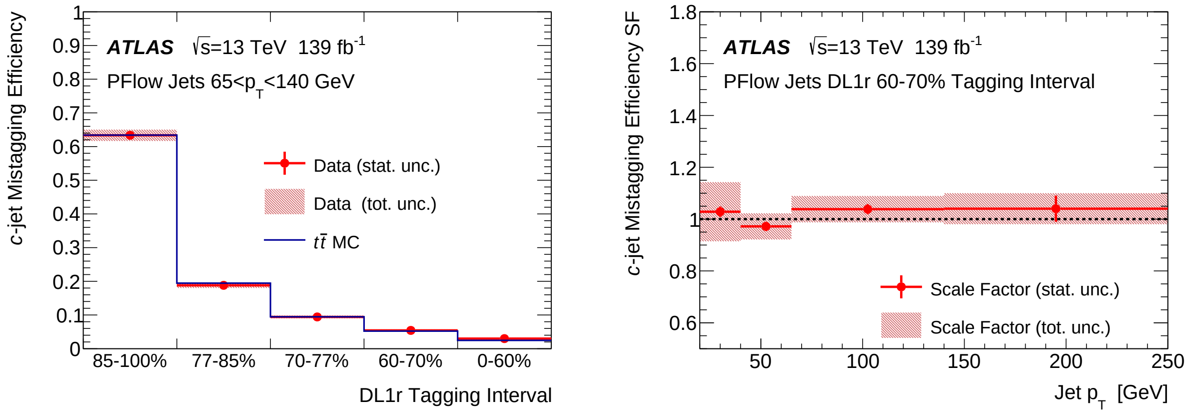


(a) Jets from the W-boson decay in $t\bar{t}$ reco. likelihood.

(b) Jets from the W-boson decay binned in D_{DL1r}.

Figure D.14: The data and MC distribution of the jets from the W boson decay as a function of the $t\bar{t}$ reco. likelihood and the D_{DL1r} discriminant [126].

The jets from the W-boson decay are then classified into four p_T regions (20-40 GeV, 40-65 GeV, 65-140 GeV, and 140-250 GeV), and working points defined using the $DL1r$ discriminant. Figure D.14(b) shows the jets from the W-boson decay in the default four $DL1r$ working points used in ATLAS. Next, the ϵ_b (ϵ_{light}) scale factors are applied to the true b-jets (light-flavor jets), and the remaining c-jets are corrected to real collision data through a χ^2 minimization fit to obtain the c-jet scaling factors. Figure D.15 shows two examples of the calibration in specific b-jet tagging working points.



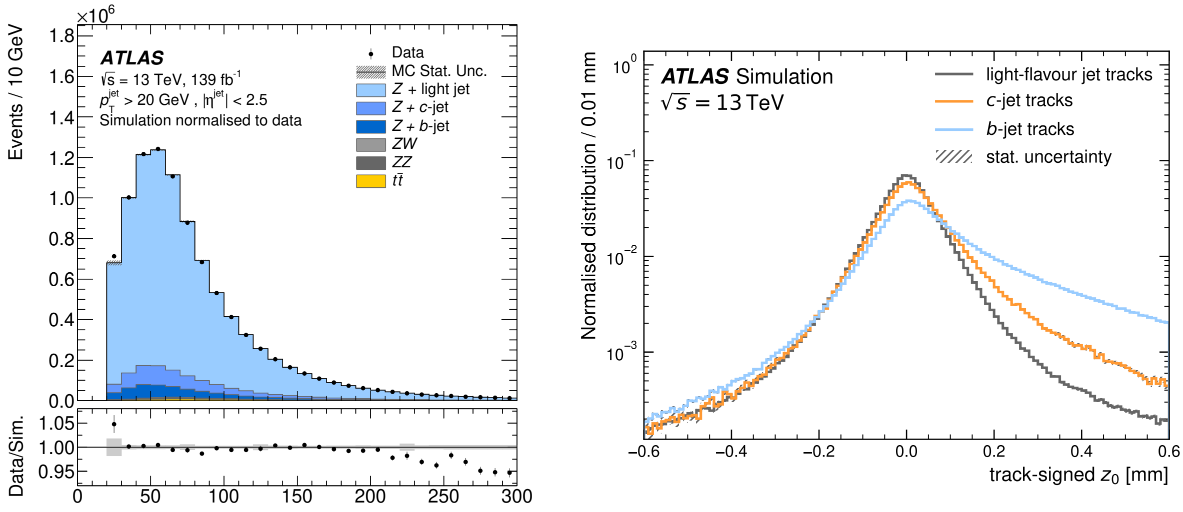
(a) The ϵ_c in different working points.

(b) The ϵ_c scale factors in the 60-70% interval.

Figure D.15: The calibrated c-jet tagging efficiency and scale factors in specific b-jet tagging working points[126].

Light-flavor jet tagging efficiency calibration

Calibrating the light-flavor jet tagging efficiencies ϵ_{light} (also referred to as mis-tagging efficiencies) is challenging as a consequence of the good rejection of light-jets in the flavor tagging working points; there is limited statistics of light-flavor jets to determine ϵ_{light} accurately. Therefore, a sample enriched in light-flavor jets is constructed using a modified $DL1r$ algorithm, known as the $DL1r^{\text{Flip}}$ algorithm, that achieves a lower ϵ_b and ϵ_c with a moderate ϵ_{light}, to enhance the light-flavor jet ratio in the sample.



(**a**) The leading jet p_T of the selected samples.

(**b**) The signed longitudinal impact parameter z_0 .

Figure D.16: The left plots shows the p_T of the leading jet of the selected Z+jets samples. The right plot shows the signed-longitudinal impact parameter z_0 [127].

The modified $DL1r$ tagger method, commonly referred to as the *negative tag method*, first determines the negative tag rate $\epsilon^{\text{data}}_{\text{neg,light}}$ in a data sample enriched in light-flavor jets, and then extracts the real mis-tagging rate ϵ_{light} in the relevant flavor tagging working points [127]. For extracting $\epsilon^{\text{data}}_{\text{neg,light}}$, a Z+jets sample with the Z-boson decaying to two leptons is used. To reduce non-Z+jets background, the invariant mass of the leptons is required to have $81 < m_{ll} <$ 101 GeV, and the event must also contain at least one jet with $p_T > 20$ GeV. Further, the p_T of the Z-boson is required to have $p_T^Z > 50$ GeV to avoid the mis-modelling in simulation. Figure D.16(a) shows the normalized MC simulation to data for the leading jet p_T, where it is seen the selected sample is enriched in light-flavor jets.

The negative tag method is based on the assumption that light-flavor jets are mis-tagged as b-jets due to the finite resolution of the reconstructed inner detector track trajectories and impact parameters. Under this assumption, signed impact parameters d_0 and z_0 are expected to be almost symmetric around zero for light-flavor jets, but not for b-hadrons or c-hadrons due to their longer decay lengths as seen in Figure D.16(b). Hence, the probability for light-flavor jets to be mis-tagged remains almost the same after flipping the sign of the impact parameters, and a good approximation of the light-flavor jet mis-tag rates can be obtained by running the IP2D and IP3D algorithms with the flipped signs.

Similar features are expected for secondary vertices and signed decay length significance (which are seeded from tracks), and hence, the negative tag method can be extended to SV1 and JETFITTER algorithms. It should be noted that the flavor tagging algorithms are not retrained, and uses the same input types but the modified values that are determined from the flipped version of the low level flavor tagging algorithms[18].

Similar to ϵ_b and ϵ_c calibration, the ϵ_{light} calibrations are performed independently in different p_T bins for the different working points of the $DL1r$ algorithm, but utilizing the $DL1r^{\text{Flip}}$ discriminant templates. The $\epsilon^{\text{data}}_{\text{neg},b}$, $\epsilon^{\text{data}}_{\text{neg,light}}$, and the ϵ_{light} scale factors are extracted from a binned fit to the m_{SV} templates, while $\epsilon^{\text{data}}_{\text{neg},c}$ and the corresponding scale factor is fixed to the MC simulation value with a 30% uncertainty due to limited sensitivity. The extracted ϵ_{light} scale factors are used in the end, with an extrapolation uncertainty from $DL1r^{\text{Flip}}$ to $DL1r$.

---

[18] The flipped versions of the low level flavor tagging algorithms are referred to as IP2DNeg, IP3DNeg, SV1Flip, and JETFITTER-Flip

E The event selection and categorization

E.1 The flavor tag selection on the additional jets

Figure E.1 shows the case for events with more than two jets, when the additional jets are classified in the different flavor tag bins. In the zero and one charged lepton channels, if the additional jets are classified in the C_T bin, the event is re-classified to the $\boldsymbol{BC_T}$ region due to possibility that such events are originating from top quark processes.

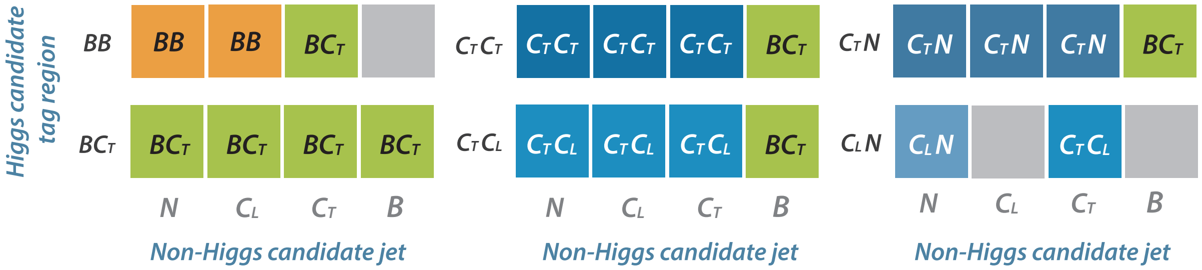


Figure E.1: The flavor tag region classification for events with more than two jets, when the additional non-Higgs candidate jet is tagged in the respective flavor tag bins.

E.2 The ΔR selection

As described in Section 5.3.2, the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cuts of Table 5.6 are derived from studies in the 2-lepton channel of the resolved $VH, H \to b\bar{b}$ analysis, and the same cuts are then harmonized between other channels, as well as with the resolved $VH, H \to c\bar{c}$ analysis. This is to be able to correlate the SR-CR extrapolation uncertainties and normalisation factors between different tag regions.

Table E.1 details the percentage of signal ending below the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cuts separating the signal region and the high ΔR control region (i.e the cuts in Table 5.6), for each lepton channel in the $VH, H \to b\bar{b}$ and $VH, H \to c\bar{c}$ analyses working points.

Table E.1: The percentage of signal under the CR cuts of Table 5.6 in various tag regions and channels. BB is the 2 b-tag working point in the $VH, H \to b\bar{b}$ analysis, and C_TC_T, C_TC_L and C_TN are the c-tag working points in the $VH, H \to c\bar{c}$ analysis.

Lepton channel	n-jets	VHbb in BB	VHcc in C_TC_T	VHcc in C_TC_L	VHcc in C_TN
0-lepton channel	2 jets	95.812 ± 0.174 %	96.681 ± 0.506 %	93.154 ± 0.374 %	88.122 ± 0.212 %
	3 jets	85.490 ± 0.528 %	86.505 ± 0.492 %	79.302 ± 0.336 %	77.476 ± 0.212 %
	4 jets	92.306 ± 1.751 %	90.277 ± 0.654 %	83.496 ± 0.441 %	80.559 ± 0.293 %
1-lepton channel	2 jets	95.992 ± 0.185 %	96.503 ± 0.544 %	93.696 ± 0.402 %	90.794 ± 0.229 %
	3 jets	83.424 ± 0.523 %	84.159 ± 0.501 %	78.289 ± 0.345 %	78.963 ± 0.222 %
2-lepton channel	2 jets	95.354 ± 0.216 %	96.400 ± 0.698 %	93.929 ± 0.519 %	89.937 ± 0.297 %
	3 jets	83.108 ± 0.551 %	84.100 ± 0.644 %	78.124 ± 0.443 %	76.282 ± 0.284 %
	4 jets	87.645 ± 1.674 %	84.434 ± 0.825 %	77.878 ± 0.552 %	75.729 ± 0.386 %
	≥ 5 jets	90.488 ± 4.983 %	84.205 ± 0.903 %	77.184 ± 0.582 %	73.073 ± 0.454 %

It is clearly seen that the harmonization of the high $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cut to other channels in the

$VH, H \to b\bar{b}$ analysis yields also the same percentage of signal events under the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ cut. This also holds true for the $C_T C_T$ region in the $VH, H \to c\bar{c}$ analysis. However, for the $C_T C_L$ region and $C_T N$ regions in the $VH, H \to c\bar{c}$ analysis, the percentage is slightly lower, especially in the ≥ 3 jet regions. This is because it is more likely to tag one random light-jet (from ISR or FSR or pile-up) causing the $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ to be higher in the $C_T C_L$ and $C_T N$ tag regions. In contrast, for BB and $C_T C_T$ tag regions, it is more likely to catch, in most of the cases, the actual b/c-jet from the Higgs decay.

Table E.2 and Table E.3 shows the percentage under the CR cut for $C_T C_L$ and $C_T N$ regions respectively, split in the jet flavor combination. Here, it is clearly seen that the lower signal percentage for $C_T C_L$ and $C_T N$ regions (as seen in Table E.1) is arising mainly from the mixed-flavor component.

Table E.2: The percentage of signal under the CR cuts, in the $C_T C_L$ tag region and split in heavy, mixed and light flavor combinations.

Lepton channel	nJet region	Heavy flavor	Mixed flavor	Light flavor
0-lepton channel	2 jets	94.447 ± 0.383 %	64.891 ± 1.564 %	62.188 ± 47.917 %
	3 jets	84.698 ± 0.376 %	47.144 ± 0.668 %	31.825 ± 12.247 %
	4 jets	90.688 ± 0.518 %	57.489 ± 0.782 %	54.687 ± 10.955 %
1-lepton channel	2 jets	94.479 ± 0.411 %	77.519 ± 1.861 %	30.264 ± 32.539 %
	3 jets	82.220 ± 0.379 %	54.868 ± 0.796 %	33.584 ± 10.483 %
2-lepton channel	2 jets	94.794 ± 0.534 %	78.228 ± 2.182 %	76.334 ± 53.218 %
	3 jets	82.041 ± 0.487 %	56.103 ± 1.007 %	54.083 ± 22.429 %
	4 jets	83.764 ± 0.641 %	57.810 ± 1.038 %	66.948 ± 14.938 %
	≥ 5 jets	84.582 ± 0.722 %	60.912 ± 0.952 %	59.696 ± 8.151 %

Table E.3: The percentage of signal under the CR cuts, in the $C_T N$ tag region and split in heavy, mixed and light flavor combinations.

Lepton channel	nJet region	Heavy flavor	Mixed flavor	Light flavor
0-lepton channel	2 jets	93.998 ± 0.247 %	68.304 ± 0.399 %	69.579 ± 16.594 %
	3 jets	90.049 ± 0.291 %	56.590 ± 0.290 %	48.255 ± 5.459 %
	4 jets	95.175 ± 0.443 %	64.777 ± 0.378 %	57.257 ± 4.727 %
1-lepton channel	2 jets	93.949 ± 0.266 %	78.930 ± 0.466 %	75.595 ± 13.910 %
	3 jets	87.508 ± 0.299 %	62.958 ± 0.332 %	61.192 ± 6.754 %
2-lepton channel	2 jets	93.958 ± 0.348 %	78.110 ± 0.560 %	82.093 ± 18.048 %
	3 jets	87.022 ± 0.390 %	60.808 ± 0.400 %	63.880 ± 8.377 %
	4 jets	90.425 ± 0.601 %	62.389 ± 0.488 %	59.151 ± 6.538 %
	≥ 5 jets	90.425 ± 0.815 %	62.748 ± 0.534 %	62.201 ± 5.008 %

E.3 The percentage of events in the analysis regions

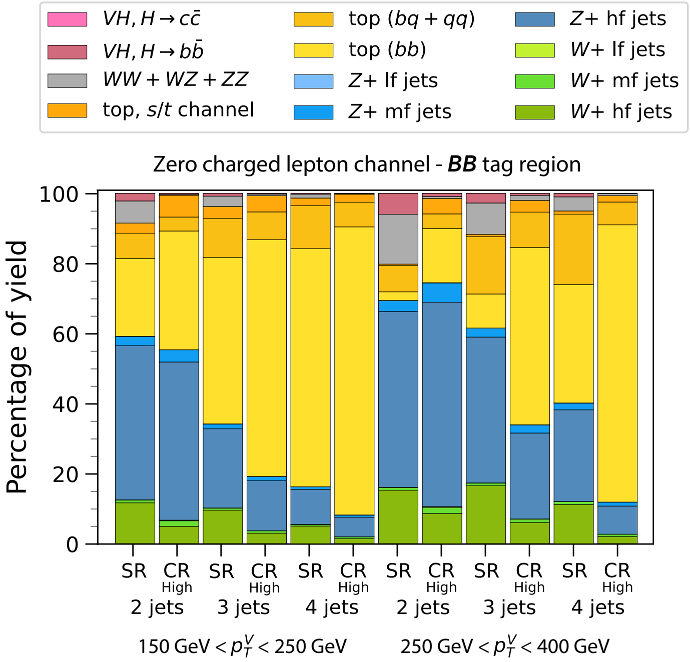


Figure E.2: Contribution of signal and background in the analysis regions of the BB-tag region in the zero charged lepton channel.

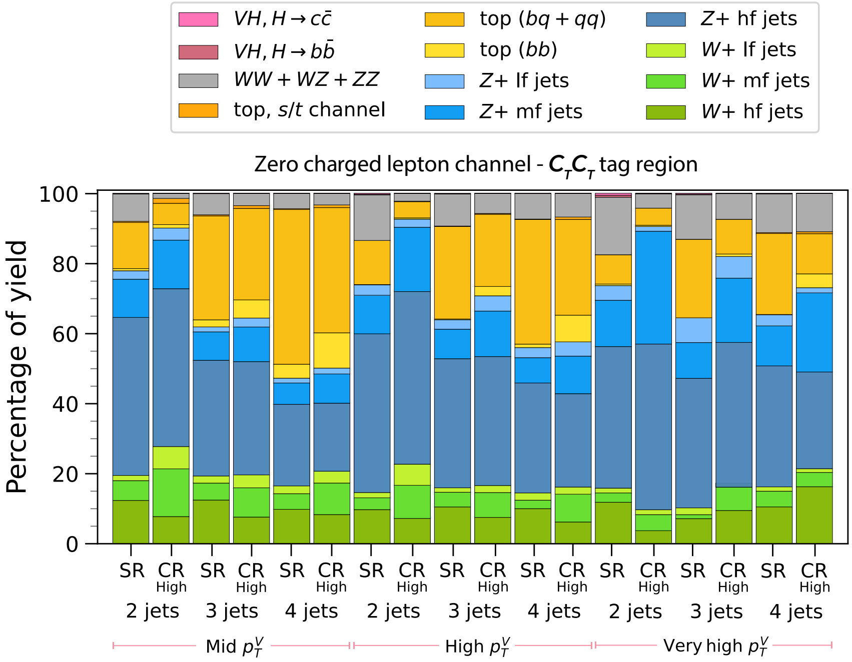


Figure E.3: Contribution of signal and background in the analysis regions of the $C_T C_T$-tag region in the zero charged lepton channel. `Mid` p_{T}^{V}, `High` p_{T}^{V}, and `Very high` p_{T}^{V} stand for the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV, 250 GeV $< p_{\mathrm{T}}^{V} <$ 400 GeV, and $p_{\mathrm{T}}^{V} >$ 400 GeV regions.

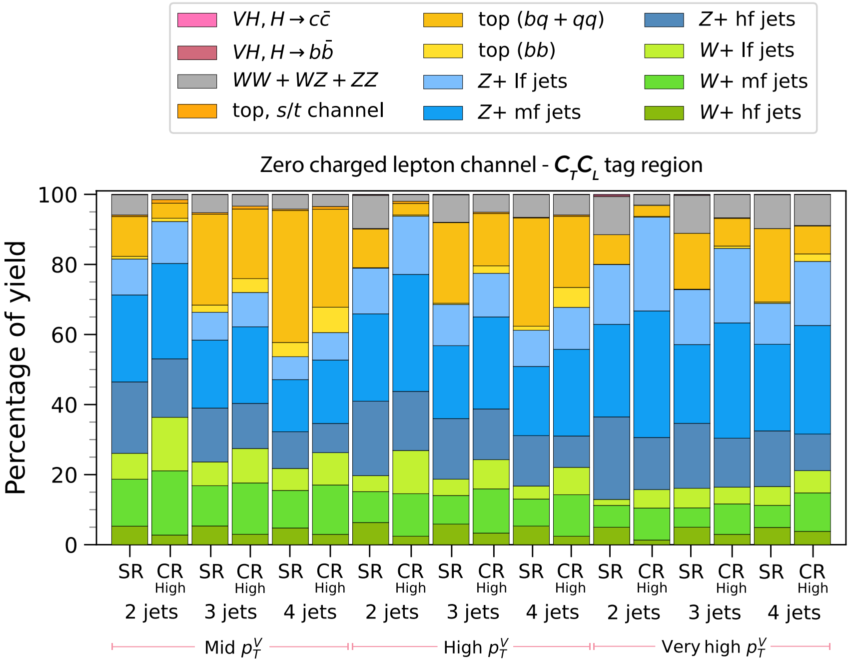


Figure E.4: Contribution of signal and background in the analysis regions of the $C_T C_L$-tag region in the zero charged lepton channel. `Mid` p_{T}^{V}, `High` p_{T}^{V}, and `Very high` p_{T}^{V} stand for the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV, 250 GeV $< p_{\mathrm{T}}^{V} <$ 400 GeV, and $p_{\mathrm{T}}^{V} >$ 400 GeV regions.

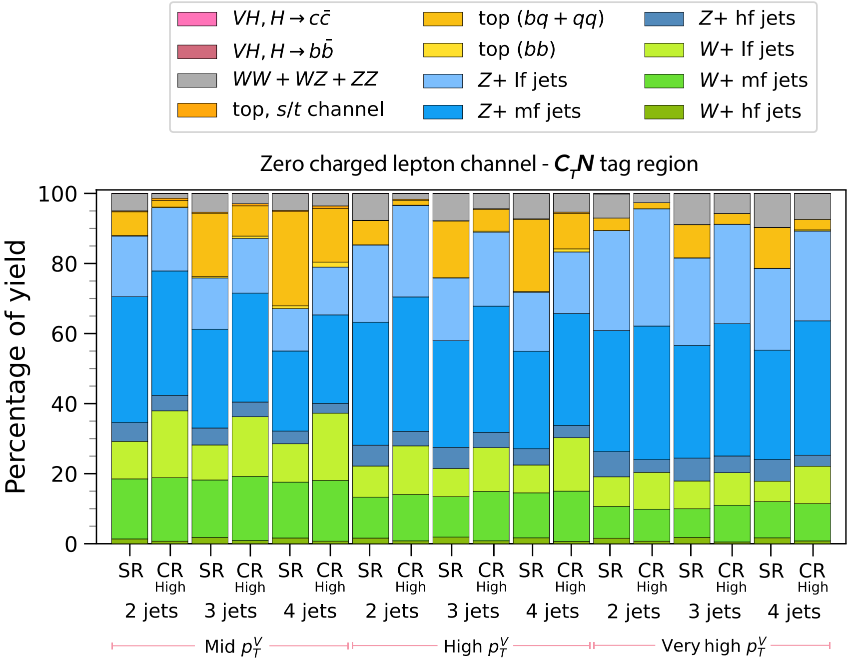


Figure E.5: Contribution of signal and background in the analysis regions of the $C_T N$-tag region in the zero charged lepton channel. `Mid` p_{T}^{V}, `High` p_{T}^{V}, and `Very high` p_{T}^{V} stand for the 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV, 250 GeV $< p_{\mathrm{T}}^{V} <$ 400 GeV, and $p_{\mathrm{T}}^{V} >$ 400 GeV regions.

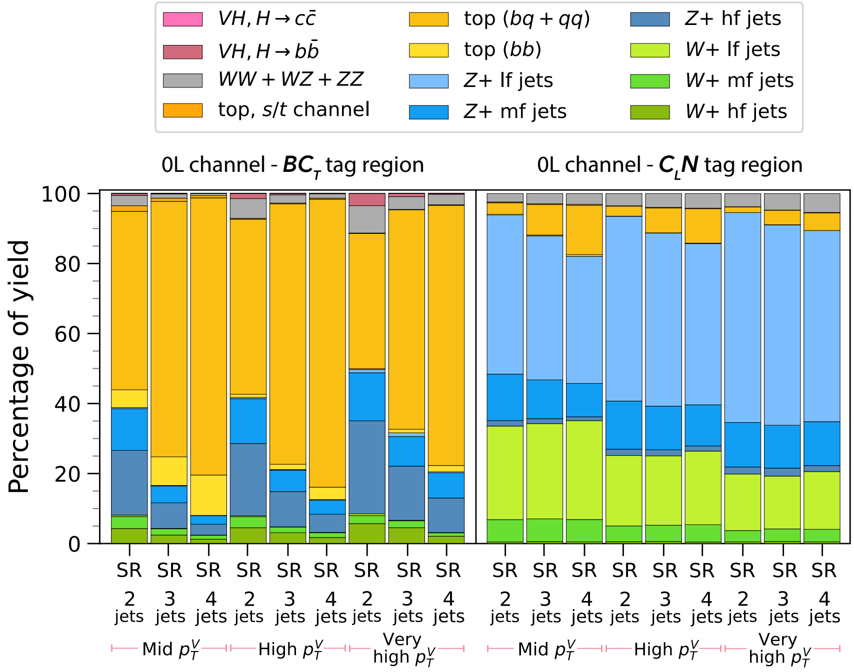


Figure E.6: Contribution of signal and background in the analysis regions of the BC_T-tag and C_LN regions in the zero charged lepton channel. `Mid` p_T^V, `High` p_T^V, and `Very high` p_T^V stand for the 150 GeV $< p_T^V <$ 250 GeV, 250 GeV $< p_T^V <$ 400 GeV, and $p_T^V >$ 400 GeV regions.

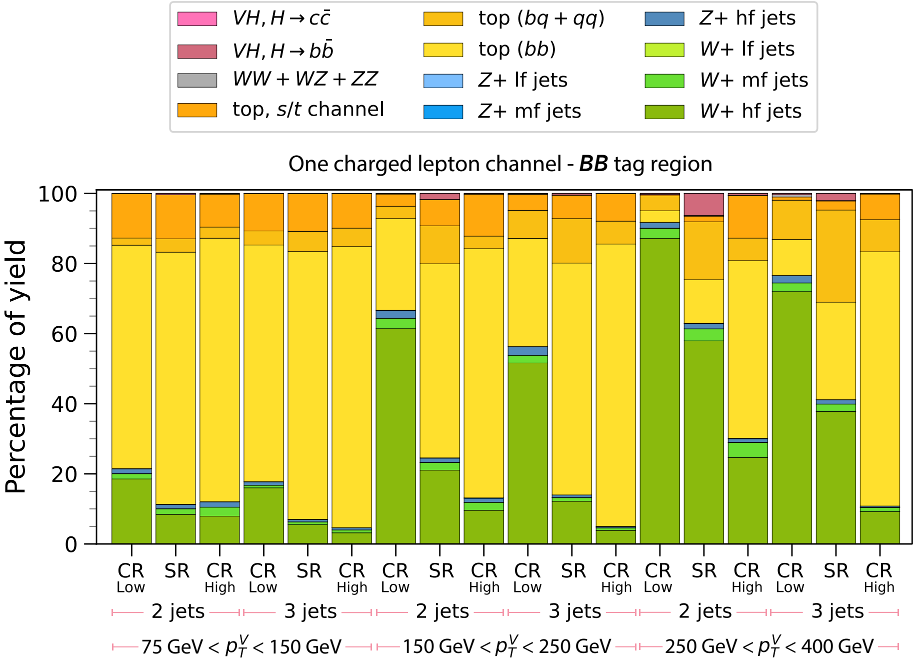


Figure E.7: Contribution of signal and background in the analysis regions of the BB-tag region in the one charged lepton channel.

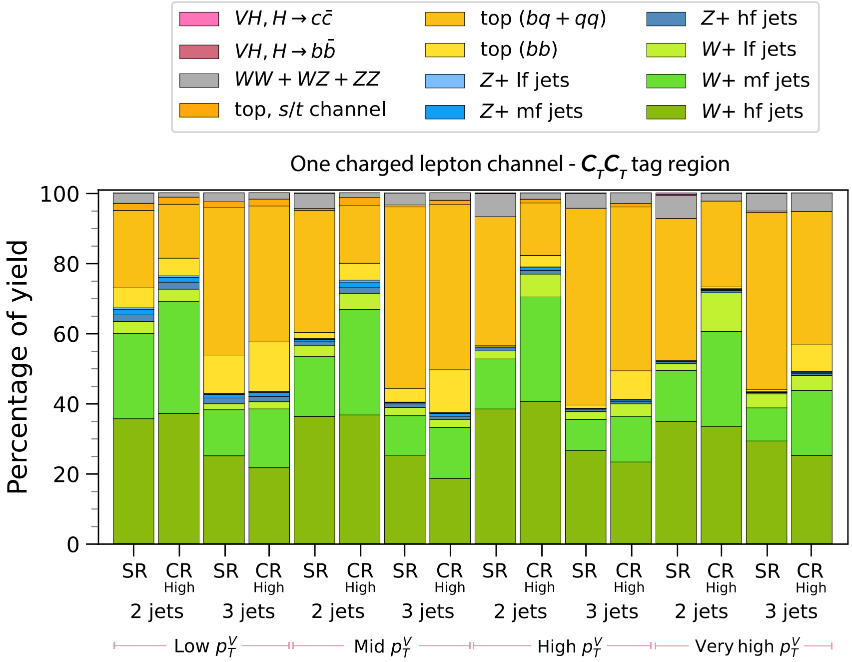


Figure E.8: Contribution of signal and background in the analysis regions of the C_TC_T-tag region in the one charged lepton channel. `Low` p_{T}^V, `Mid` p_{T}^V, `High` p_{T}^V, and `Very high` p_{T}^V stand for the 75 GeV $< p_{\mathrm{T}}^V <$ 150 GeV, 150 GeV $< p_{\mathrm{T}}^V <$ 250 GeV, 250 GeV $< p_{\mathrm{T}}^V <$ 400 GeV, and $p_{\mathrm{T}}^V >$ 400 GeV regions.

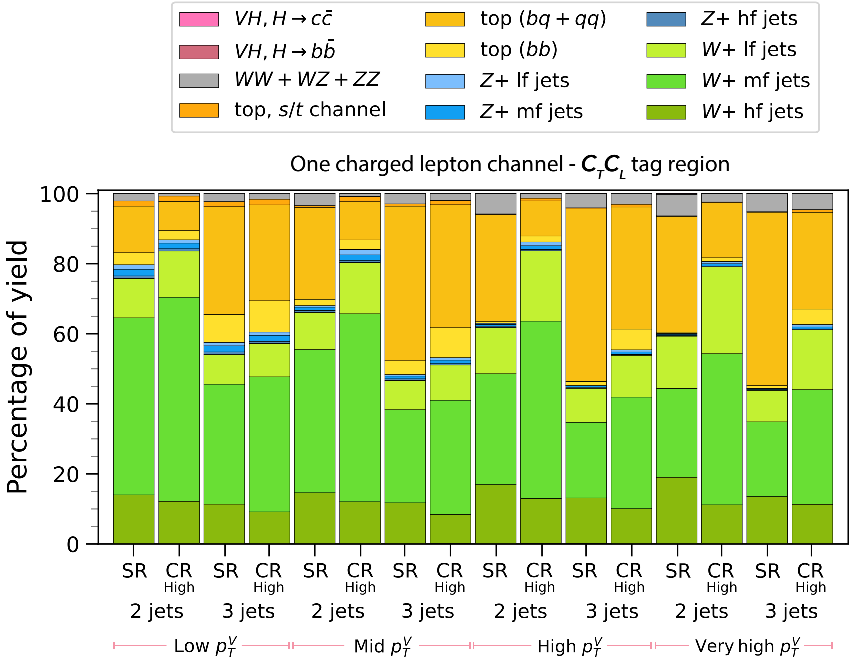


Figure E.9: Contribution of signal and background in the analysis regions of the C_TC_L-tag region in the one charged lepton channel. `Low` p_{T}^V, `Mid` p_{T}^V, `High` p_{T}^V, and `Very high` p_{T}^V stand for the 75 GeV $< p_{\mathrm{T}}^V <$ 150 GeV, 150 GeV $< p_{\mathrm{T}}^V <$ 250 GeV, 250 GeV $< p_{\mathrm{T}}^V <$ 400 GeV, and $p_{\mathrm{T}}^V >$ 400 GeV regions.

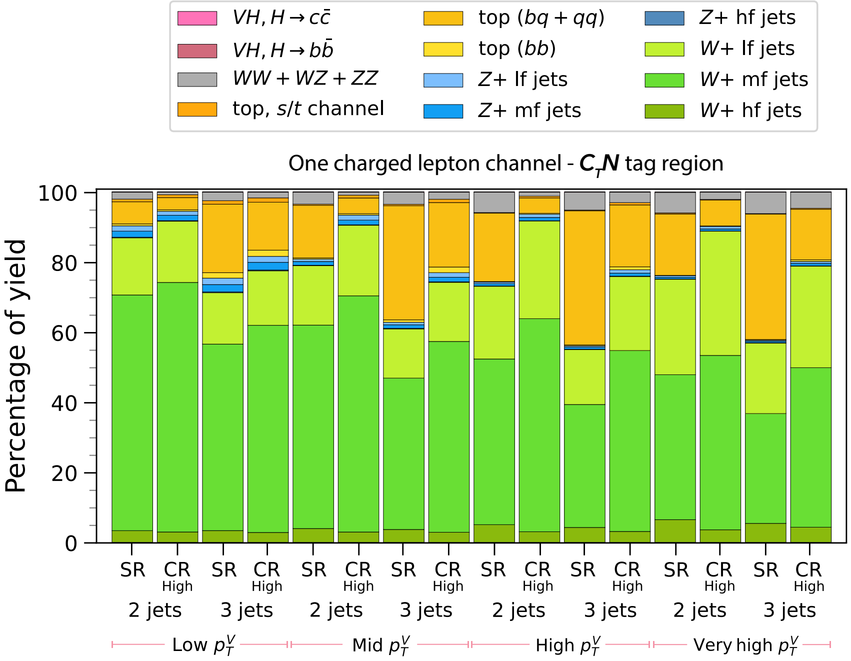


Figure E.10: Contribution of signal and background in the analysis regions of the C_TN-tag region in the one charged lepton channel. `Low` p_{T}^V, `Mid` p_{T}^V, `High` p_{T}^V, and `Very high` p_{T}^V stand for the 75 GeV $< p_{\mathrm{T}}^V <$ 150 GeV, 150 GeV $< p_{\mathrm{T}}^V <$ 250 GeV, 250 GeV $< p_{\mathrm{T}}^V <$ 400 GeV, and $p_{\mathrm{T}}^V >$ 400 GeV regions.

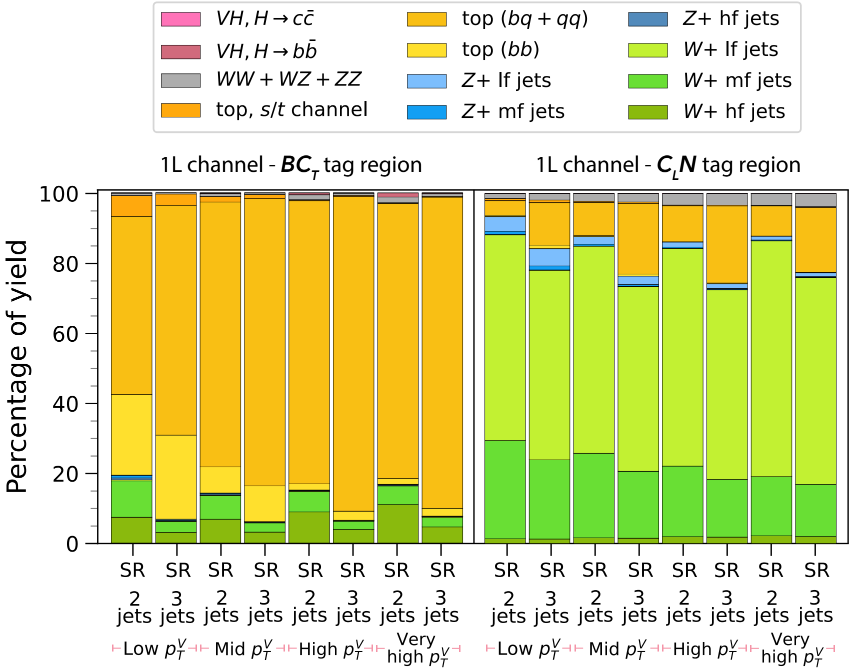


Figure E.11: Contribution of signal and background in the analysis regions of the BC_T-tag and C_LN regions in the one charged lepton channel. `Low` p_{T}^V, `Mid` p_{T}^V, `High` p_{T}^V, and `Very high` p_{T}^V stand for the 75 GeV $< p_{\mathrm{T}}^V <$ 150 GeV, 150 GeV $< p_{\mathrm{T}}^V <$ 250 GeV, 250 GeV $< p_{\mathrm{T}}^V <$ 400 GeV, and $p_{\mathrm{T}}^V >$ 400 GeV regions.

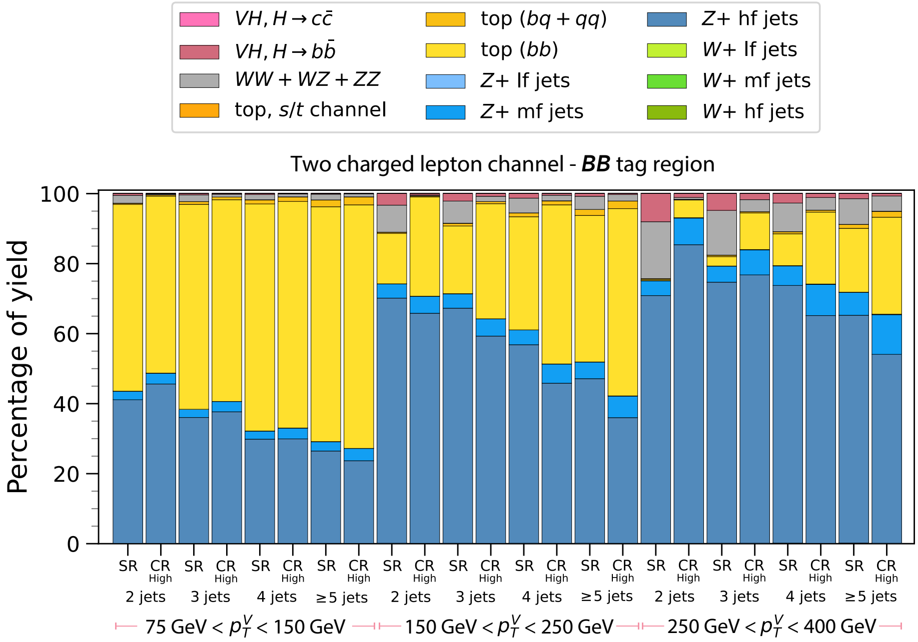


Figure E.12: Contribution of signal and background in the analysis regions of the BB-tag region in the two charged lepton channel.

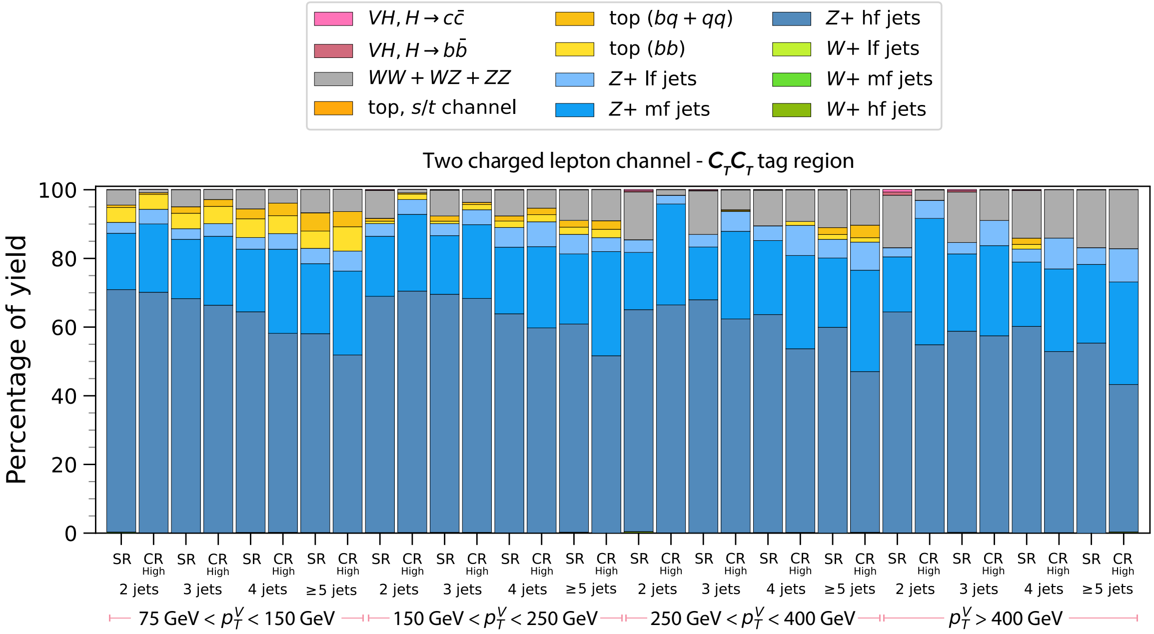


Figure E.13: Contribution of signal and background in the analysis regions of the C_TC_T-tag region in the two charged lepton channel. `Low` p_T^V, `Mid` p_T^V, `High` p_T^V, and `Very high` p_T^V stand for the 75 GeV $< p_T^V <$ 150 GeV, 150 GeV $< p_T^V <$ 250 GeV, 250 GeV $< p_T^V <$ 400 GeV, and $p_T^V >$ 400 GeV regions.

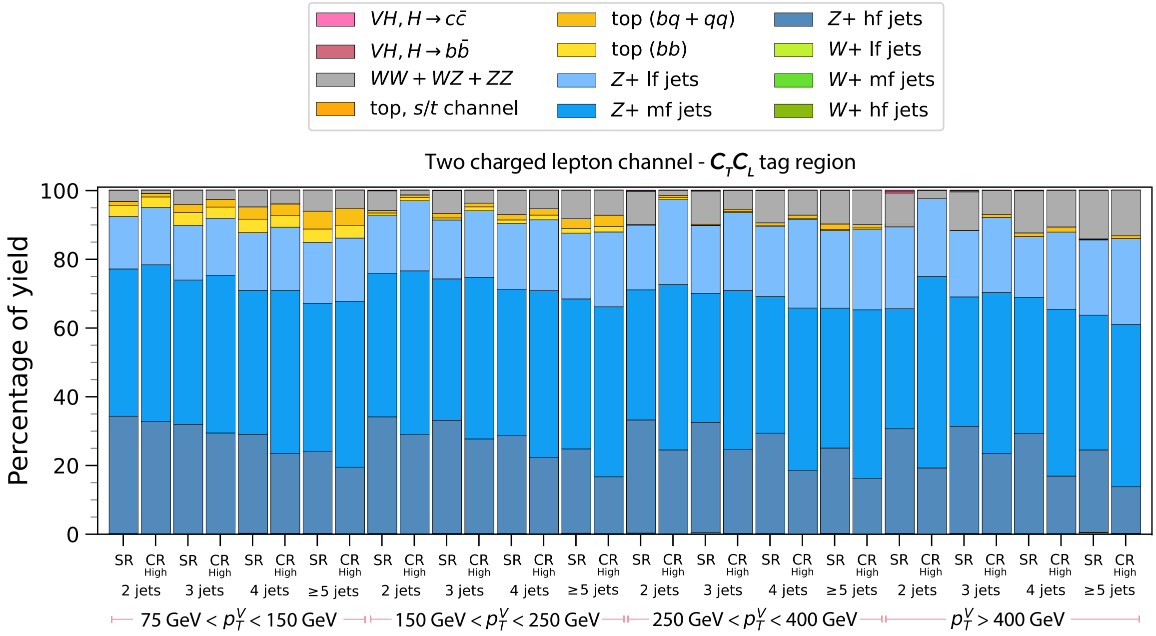


Figure E.14: Contribution of signal and background in the analysis regions of the $C_T C_L$-tag region in the two charged lepton channel. `Low` p_T^V, `Mid` p_T^V, `High` p_T^V, and `Very high` p_T^V stand for the 75 GeV $< p_T^V <$ 150 GeV, 150 GeV $< p_T^V <$ 250 GeV, 250 GeV $< p_T^V <$ 400 GeV, and $p_T^V >$ 400 GeV regions.

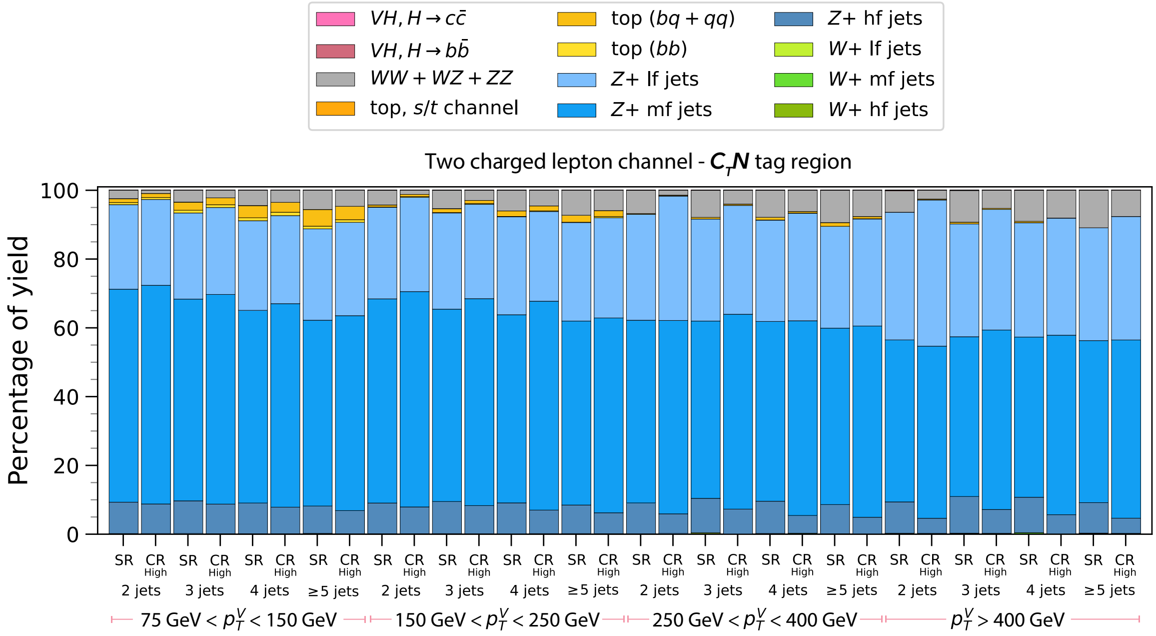


Figure E.15: Contribution of signal and background in the analysis regions of the $C_T N$-tag region in the two charged lepton channel. `Low` p_T^V, `Mid` p_T^V, `High` p_T^V, and `Very high` p_T^V stand for the 75 GeV $< p_T^V <$ 150 GeV, 150 GeV $< p_T^V <$ 250 GeV, 250 GeV $< p_T^V <$ 400 GeV, and $p_T^V >$ 400 GeV regions.

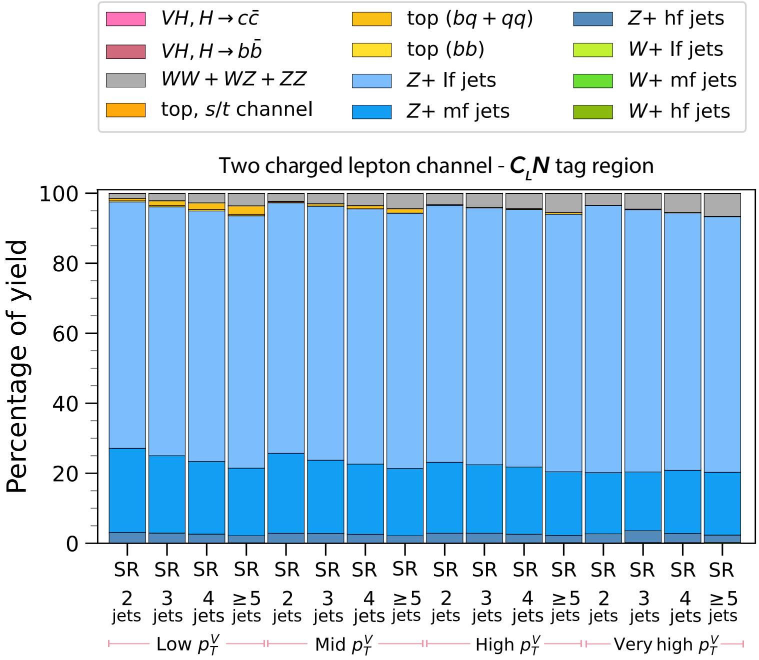


Figure E.16: Contribution of signal and background in the analysis regions of the BC_T-tag and C_LN regions in the two charged lepton channel. `Low` p_{T}^V, `Mid` p_{T}^V, `High` p_{T}^V, and `Very high` p_{T}^V stand for the 75 GeV $< p_{\mathrm{T}}^V <$ 150 GeV, 150 GeV $< p_{\mathrm{T}}^V <$ 250 GeV, 250 GeV $< p_{\mathrm{T}}^V <$ 400 GeV, and $p_{\mathrm{T}}^V >$ 400 GeV regions.

F Truth flavor tagging

This appendix details some in depth explanation on truth flavor tagging, and also shows some additional performance plots not detailed in the main section.

F.1 The GNN model architecture

As discussed in Section 6.2.2, the inputs to the GNN are fully connected graphs, with each node representing a jet in the event with nine input features. The input graph passes through several message passing networks that update the representation of the nodes, while also taking into account the information of the neighboring nodes linked by the edges. Each linear layer in the network is followed by a rectified linear activation function. Then a second fully connected network (referred to as the *jet efficiency network*) updates the input node representations, and the last layer (which is followed by a normalised exponential activation) outputs the flavour tagging efficiency for each jet, per each flavor tag bin. An illustration of the GNN model is shown in the Figure F.1, and a detailed explanation can be found in references [130, 137].

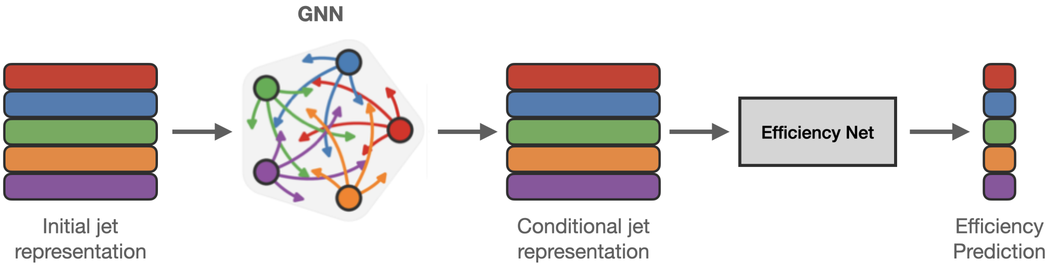


Figure F.1: The network architecture of the GNN model used for parameterising the flavor tagging efficiencies in truth flavor tagging [130].

F.2 Exclusive truth tagging vs inclusive truth tagging

Exclusive truth tagging is chosen when a veto is applied on the additional jets (i.e - the non-Higgs candidate jets). This is the case for VHbb, where we have a veto on additional b-tagged jets. Hence, an example of the computed weight (when picking 2 out of 3 jets) is,

$$\begin{aligned} P(\text{2tag3jet}|\theta) =& \epsilon_{j_1}(\mathbf{x}|\theta)\cdot\epsilon_{j_2}(\mathbf{x}|\theta)\cdot(1-\epsilon_{j_3}(\mathbf{x}|\theta)) \\ &+\epsilon_{j_1}(\mathbf{x}|\theta)\cdot(1-\epsilon_{j_2}(\mathbf{x}|\theta))\cdot\epsilon_{j_3}(\mathbf{x}|\theta)+(1-\epsilon_{j_1}(\mathbf{x}|\theta))\cdot\epsilon_{j_2}(\mathbf{x}|\theta)\cdot\epsilon_{j_3}(\mathbf{x}|\theta) \end{aligned} \tag{F.1}$$

If we apply exclusive truth tagging in VHcc, the probabilities will be as shown in figure 6.3. However, in VHcc we do not require a veto on additional jets that are c-tagged, and this is reflected on the calculation of the weights. This approach is known as **inclusive truth tagging**, and the re-calculated weight can be shown like below (note that there is still a b-veto on the 3rd jet).

$$P(\text{2tag3jet}|\theta) = \epsilon^L_{j1}\epsilon^L_{j2}(1-\epsilon^{b70}_{j3}) + \epsilon^L_{j1}\epsilon^T_{j2}(1-\epsilon^{b70}_{j3}) + \epsilon^T_{j1}\epsilon^L_{j2}(1-\epsilon^{b70}_{j3}) + \epsilon^T_{j1}\epsilon^T_{j2}(1-\epsilon^{b70}_{j3}) + \ldots.. \tag{F.2}$$

F.3 Truth tagging strategy in the $VH, H \to b\bar{b}/c\bar{c}$ analysis

Table F.1 shows the final strategy adopted for truth flavor tagging in the $VH, H \to b\bar{b}/c\bar{c}$ analysis, while Figure F.2 shows an illustration of the truth flavor tagging in the BB and C_TC_T tag regions.

Table F.1: The different strategies used for truth tagging. The abbreviations DT and TT used above stand for direct tag and truth tag respectively. WP stands for working point.

	VHbb resolved	**VHbb boosted**	**VHcc**
Hybrid tagging	Yes (b-jets are DT'd)	No (fully TT'd)	No (fully TT'd)
Truth tag WP	70% b & 70% b	85% b & 85% b	c-tight & c-tight
MC stat. % for TT regions	100%	100%	8%
V+jets	hybrid tagged	truth tagged	truth tagged
single-top s/t	hybrid tagged	truth tagged	truth tagged
single-top Wt	direct tagged	truth tagged	direct tagged
$t\bar{t}$	direct tagged	truth tagged	direct tagged
diboson	direct tagged	direct tagged	direct tagged
signal	direct tagged	direct tagged	direct tagged

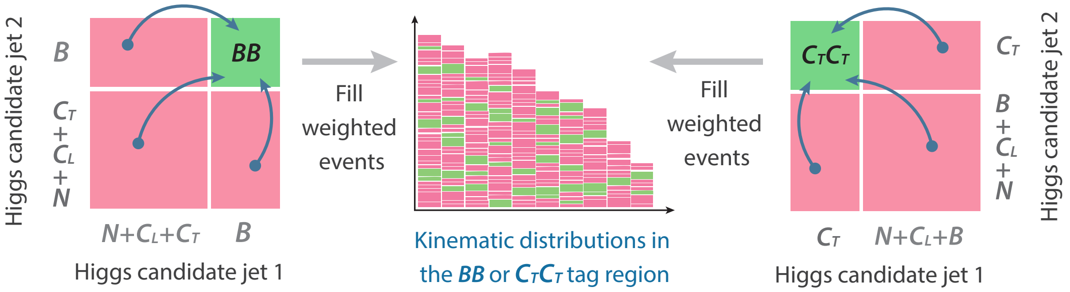


Figure F.2: An illustration showing truth flavor tagging in the BB and C_TC_T tag regions. The green events in the kinematic distribution shows events that are originally enter the BB or C_TC_T tag regions, while red events are originally classified in other tag regions through direct flavor tagging. As the figure illustrates, all green and red events are used to populate the kinematic distributions in truth flavor tagging.

F.4 Strategy in the $VH, H \to c\bar{c}$ analysis

Due to technical reasons, the choice of the $DL1r$ working point should be the same for all the jets, which limits the application of truth flavor tagging in the $VH, H \to c\bar{c}$ analysis regions. Due to this limitation, the truth flavor tagging can only be implemented in a limited number of choices in the $VH, H \to c\bar{c}$ analysis regions. Figure F.3 shows the three choices, or *strategies*, that were considered in the analysis.

When considering the three strategies, priority is given to include the C_TC_T tag region, since it is the most sensitive region to the $VH, H \to c\bar{c}$ signal; reduction of the MC statistical uncertainties

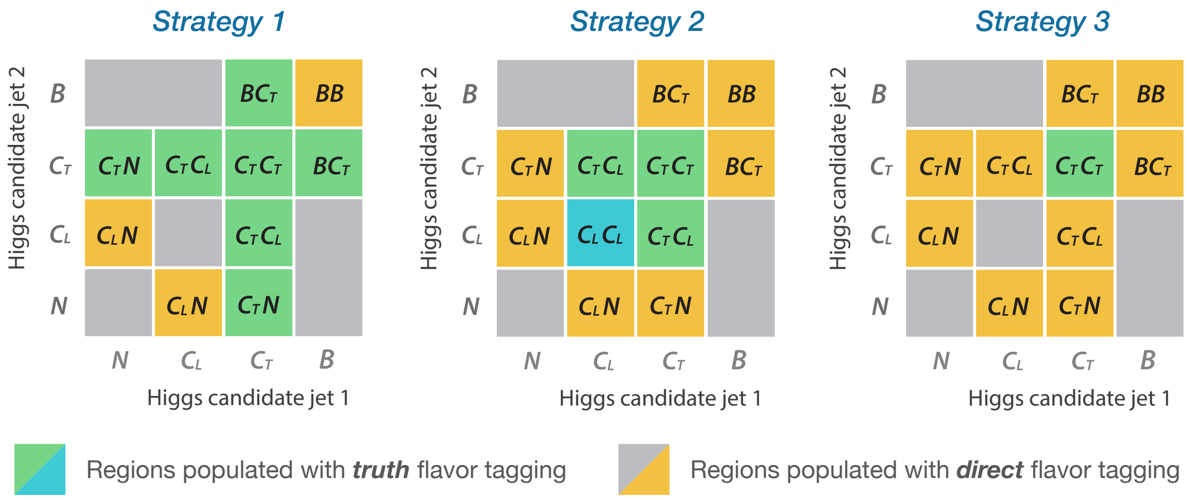


Figure F.3: The three different strategies investigated for truth flavor tagging. The *green* and *blue* regions are populated by truth tagging, and the *grey* and *yellow* regions are populated with direct flavor tagging. Green and yellow regions are the regions that are used in the analysis.

are crucial in the C_TC_T region through truth flavor tagging. The BB tag region is ignored in the considerations, since it is populated independently.

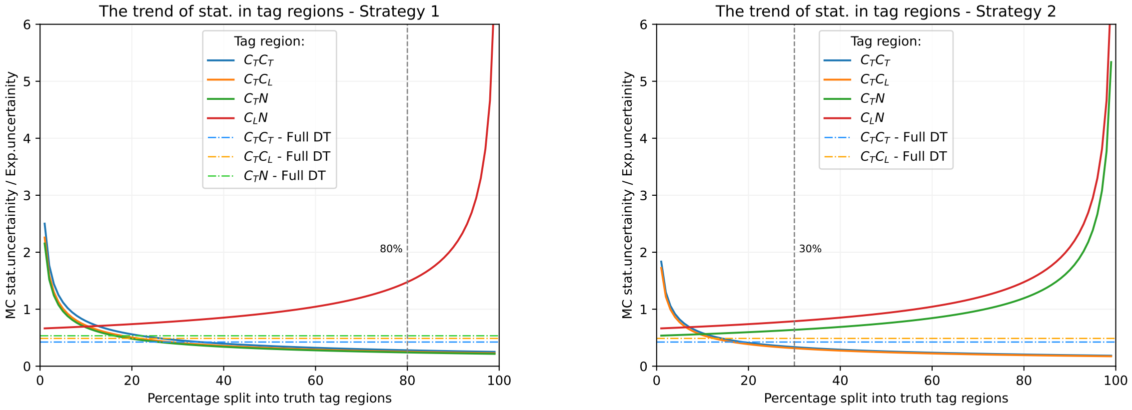
- Strategy 1 gives the signal regions C_TC_T, C_TC_L, C_TN as well as the top quark bq processes control region BC_T through truth flavor tagging. However, the C_LN control region for V+light flavor jets needs to be populated using direct flavor tagging[19].
- Strategy 2 gives the signal regions C_TC_T, C_TC_L through truth flavor tagging, as well as the region C_LC_L that is not used in the analysis. Other regions C_TN, BC_T and C_LN requires to be populated through direct flavor tagging.
- Strategy 3 gives only the signal region C_TC_T through truth flavor tagging. All other regions C_TN, BC_T and C_LN requires to be populated through direct flavor tagging.

It is important to remind that none of the strategies can produce all required regions through truth flavor tagging, and some regions need to be produce through direct flavor tagging. Hence, as a consequence there will be an overlap of MC events between the truth flavor tagged regions, and direct flavor tagged regions. To avoid this issue, and to reduce the burden on the number of times the flavor tagging selection is repeated, a percentage of events will be split for truth flavor tagging, and the remaining events are used for direct flavor tagging.

The percentage to split the events can be estimated by comparing the trend of the ratio between the actual MC statistical uncertainty, and the expected data statistical uncertainty, which is shown in Figure F.4. The ideal percentage lies in the region between the intersection of the solid curve with the dotted curves, and where the solid curves reach 1. The former condition is to ensure there is a merit of using truth flavor tagging (compared to direct flavor tagging) to populate the regions, while the latter condition is to ensure other regions have adequate MC statistics through direct flavor tagging.

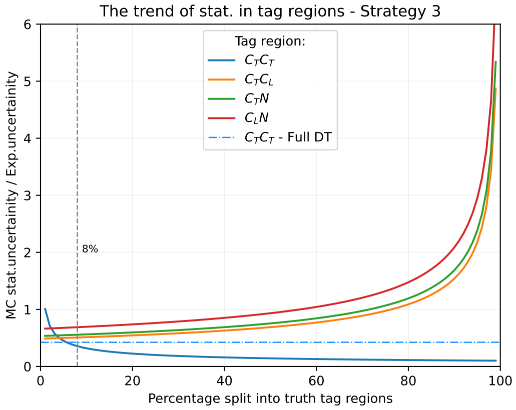
---

[19] To populate the C_LN also with truth flavor tagging, one needs to include C_LC_L and BC_L regions not used in the analysis, which is not optimum.



(**a**) Strategy 1.

(**b**) Strategy 2.



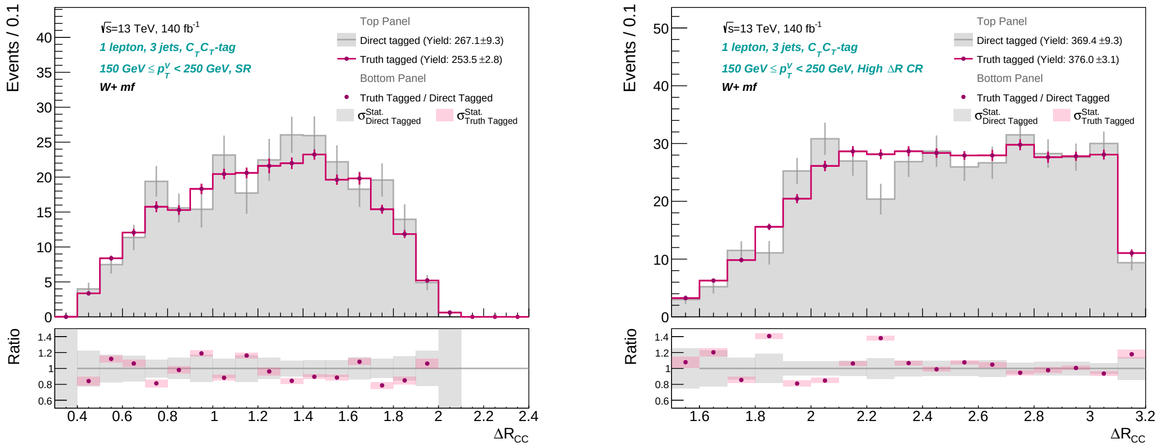
(**c**) Strategy 3.

Figure F.4: The ratio between the MC statistical uncertainty vs the expected data statistical uncertainty, at different percentages of splitting the simulated samples for truth flavor tagging and direct flavor tagging.

The percentage to split the events was decided at 80% for strategy 1, 30% for strategy 2, and 8% for strategy 3. Finally, strategy 3 was adopted in the analysis.

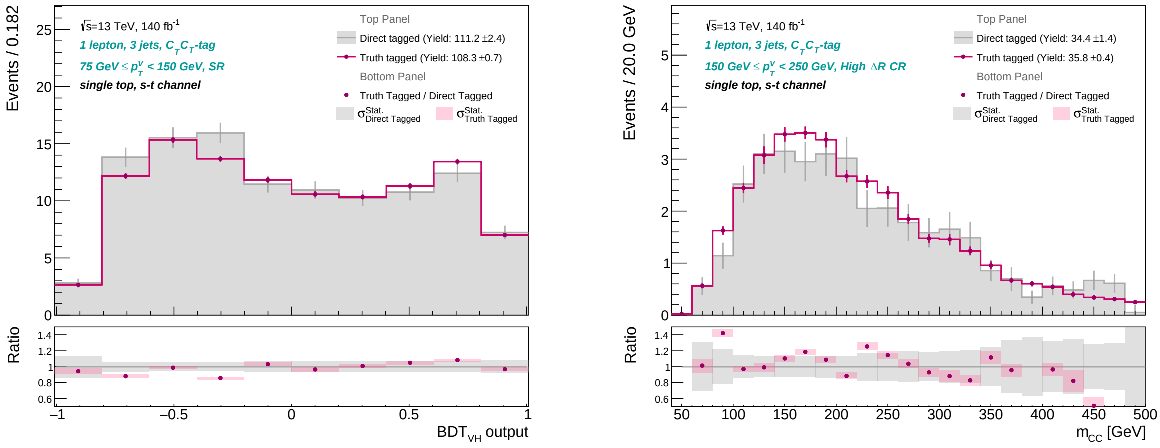
F.5 Additional closure plots in the $VH, H \to c\bar{c}$ regions

This section summarizes some additional plots showing the closure between direct flavor tagging, and the GNN based truth flavor tagging in the $C_T C_T$ region (i.e the $VH, H \to c\bar{c}$ signal region).



(**a**) $W + mf$ jets in the 1L $C_T C_T$ signal region. (**b**) $W + mf$ jets in the 1L $C_T C_T$ high ΔR control region.

Figure F.5: The closure between direct flavor tagging and the GNN based truth flavor tagging for the angular separation $\Delta R(\boldsymbol{j_1}, \boldsymbol{j_2})$ between the Higgs candidate jets, in the $C_T C_T$ working point. mf and *1L* are abbreviations for *mixed flavor* and the *one charged lepton* channel.

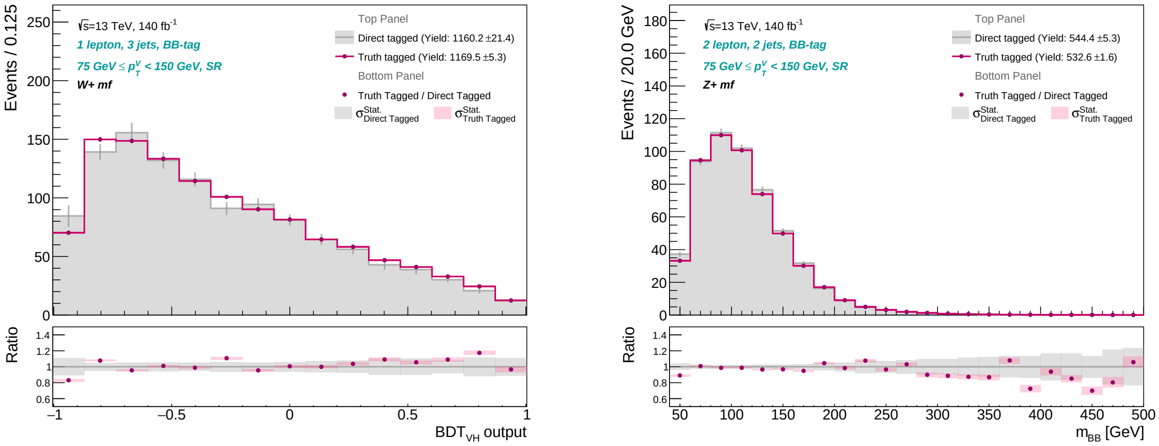


(**a**) single top in the 1L $C_T C_T$ signal region. (**b**) single top in the 1L $C_T C_T$ high ΔR control region.

Figure F.6: The closure between direct flavor tagging and the GNN based truth flavor tagging for $m_{j_1 j_2}$ and BDT in the $C_T C_T$ working point. *1L* is an abbreviation for the *one charged lepton* channels.

F.6 Additional closure plots in the $VH, H \to b\bar{b}$ regions

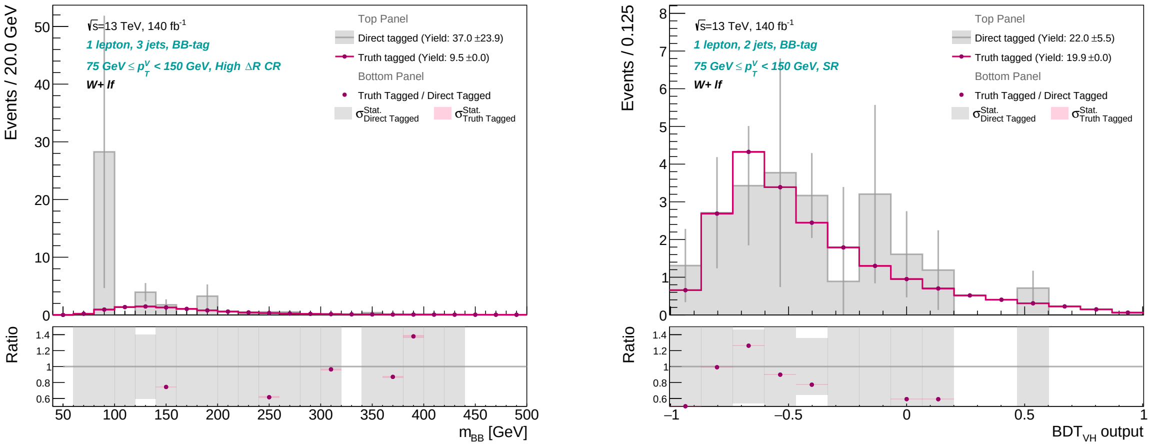
This section summarizes some additional plots showing the closure between direct flavor tagging, and the GNN based truth flavor tagging in the BB region (i.e the $VH, H \to b\bar{b}$ signal region).



(a) $W + mf$ jets in the 1L BB signal region. (b) $Z + mf$ jets in the 2L BB signal region.

Figure F.7: The closure between direct flavor tagging and the GNN based truth flavor tagging for $m_{j_1 j_2}$ and BDT output, in the BB region. mf is an abbreviations for *mixed flavor*, while *1L* and *2L* are abbreviations for the *one charged lepton* and *two charged lepton* channels.

The two plots below show the modeling of the W+light flavor background, where the direct tagged distributions are seen to be very depleted in MC statistics. In such distributions, an event with large MC weights can appear as seen in Figure F.8(a). However, truth flavor tagging managed to model these distributions much smoother.



(a) $W + lf$ jets in the 1L BB high ΔR control region. (b) $W + lf$ jets in the 1L BB signal region.

Figure F.8: The closure between direct flavor tagging and the GNN based truth flavor tagging for $m_{j_1 j_2}$ and BDT output, in the BB region. lf is an abbreviation for *light flavor*, while *1L* is an abbreviation for the *one charged lepton* channel.

G The multivariate analysis

G.1 The $VH, H \rightarrow b\bar{b}$ MVA

In the $VH, H \rightarrow b\bar{b}$ analysis uses gradient boosting for the resolved regime ($p_{\mathrm{T}}^{V} < 400$ GeV), while AdaBoost is used for the boosted regime ($p_{\mathrm{T}}^{V} > 400$ GeV).

G.2 Transformation D

The merging (or re-mapping) using transformation D can be described using the formula in Equation G.1,

$$Z_i(j,k) \equiv z_s \frac{n_s(j,k)}{N_s} + z_b \frac{n_b(j,k)}{N_b} \ , \tag{G.1}$$

where, $Z_i(j,k)$ is the i_{th} bin of the transformed distribution, obtained from merging all the bins in the range $\geq j$ and $\leq k$. $n_s(j,k)$ and $n_b(j,k)$ are the original number of signal and background present in the same range. N_s and N_b are the total number of signal and background events, respectively. Then, z_s and z_b are parameters used to tune the merging.

The bins are merged from the right-most bin (referred to k_{last}), by merging bins to the left of k_{last} one after the other. Then, at each step the below conditions are checked.

- If the computed value of $Z_i(j, k_{\mathrm{last}})$ is > 1.
- If the statistical uncertainty (of the merged bin) is $< 20\%$.
- If the yield (signal + background) is ≥ 3 events.

Once all the above three conditions are met at a particular bin j, all the bins between j and k_{last}, and including j and k_{last}, are merged into a single bin. Then this process is repeated by setting $k_{\mathrm{last}} = j - 1$, and it stops when the first bin (or the left most bin in the original distribution) is reached. After Transformation D, the bins will have different widths, but for better visualization, the bins are shown equidistant in the new BDT distributions.

The parameters z_s and z_b are tuned such that the number of bins have enough statistics while ensuring good sensitivity to the signal. The final choices of these parameters are detailed in the Table G.1.

Table G.1: The values of the parameters z_s and z_b used in Transformation D, in different $VH, H \rightarrow c\bar{c}$ phase spaces.

Tag region	75 GeV < p_{T}^{V} < 150 GeV		150 GeV < p_{T}^{V} < 250 GeV				p_{T}^{V} > 250 GeV	
			0L/1L		2L			
	z_s	z_b	z_s	z_b	z_s	z_b	z_s	z_b
TT	5	3	5	3	2	2	2	2
LT/XT	10	5	10	5	5	5	5	3
Other	10	5	10	5	10	5	10	5

H Fit framework and results

H.1 Treatment of nuisance parameters

The nuisance parameters that have shape effects are introduced as alternate templates to the fit model. However, since statistical fluctuations on templates could create artificially large uncertainties in certain bins, this may lead to over constraining nuisance parameters when fitting. Over constraining a nuisance parameter will not give the required systematic variation in regions where the signal strength is measured, and may lead to a biased measurement. To prevent this, a **smoothing** of the alternate template is done; the bins in the template are iteratively re-binned until the statistical uncertainty is less than 5% in the set of bins.

The fit model used in the analysis has a large number of nuisance parameters as discussed in the previous sections, but some nuisance parameters have negligible systematic variation on the signal and background templates used in the fit. Excluding such nuisance parameters, which is known as **pruning**, helps with the computation time as well as the convergence on the fit. In the fit model, nuisance parameters that are having systematic effects like below are pruned when fitting[20].

- Nuisance parameters having normalization effects smaller than 0.5%.
- Nuisance parameters having shape effects where none of the bins have a variation above 0.5% with respect to the nominal template[21].
- In regions where the signal to background contribution is more than 2% (sensitive regions), nuisance parameters for background processes having less than 2% of the signal yield.
- In non-sensitive regions, nuisance parameters that have a variation of less than 0.5% of the total background.

H.2 Different fit diagnostics

The effect of the nuisance parameters on the fit is determined through dedicated measurements.

Pulls and Constraints

The deviation of the post-fit value $\hat{\theta}$ of a nuisance parameter from its nominal value θ_0 in terms of the standard deviation σ_{θ_0} is known as the **pull**,

$$\text{pull} = \frac{\hat{\theta} - \theta_0}{\sigma_{\theta_0}} = \hat{\theta} \, , \tag{H.1}$$

where by definition $\theta_0 = 0$ and $\sigma_{\theta_0} = 1$. For normalization factors the pre-fit nominal value $\tau_0 = 1$, and since normalization factors $\vec{\tau}$ are not associated to any prior, the fitted values are given directly.

---

[20] The pruning is done separate for each nuisance parameter, separately for different processes, and separate in the regions.

[21] Nominal templates are the default signal and background template introduced to the fit model.

The decrease in the variance of the nuisance parameter after performing the fit is known as the **constraint**, and indicates the certainty with which the NP is known. The uncertainty on the pull is calculated using a covariance matrix, which is given by the inverse Hesse matrix about the point of maximal likelihood. If the fit model is able to extract more information on a given nuisance parameter, then the resulting uncertainty will be smaller than unity and the fit will **constrain** the specific nuisance parameter.

Ranking

To determine the impact that a nuisance parameter θ_i has on the signal strength μ, a likelihood scan is performed by fixing the value of the nuisance parameter θ_i to the scanning point and refitting all the other nuisance parameters $\theta_{j\neq i}$. The scan is continued until the negative log-likelihood difference with respect to the nominal fit reaches $\frac{1}{2}$. This point marks the $+1\sigma$ impact of this nuisance parameter on the signal strength, and the scan is repeated starting from the nominal position into the other direction to get -1σ impact. Finally the impact of the nuisance parameters on the signal strength μ is **ranked** to determine the most important nuisance parameters in the analysis.

Correlation matrix

Correlation matrices show the correlation between a pair of nuisance parameters or POI's, where only the nuisance parameters that have at least one correlation larger than a certain threshold are shown in the matrix. Note that uncertainties not shown in the below plots are pruned.

H.3 Pull plots for nuisance parameters and normalisation factors

The pulls on the normalisation factors and nuisance parameters, defined in Appendix H.2, are shown in Figures H.1 to H.6.

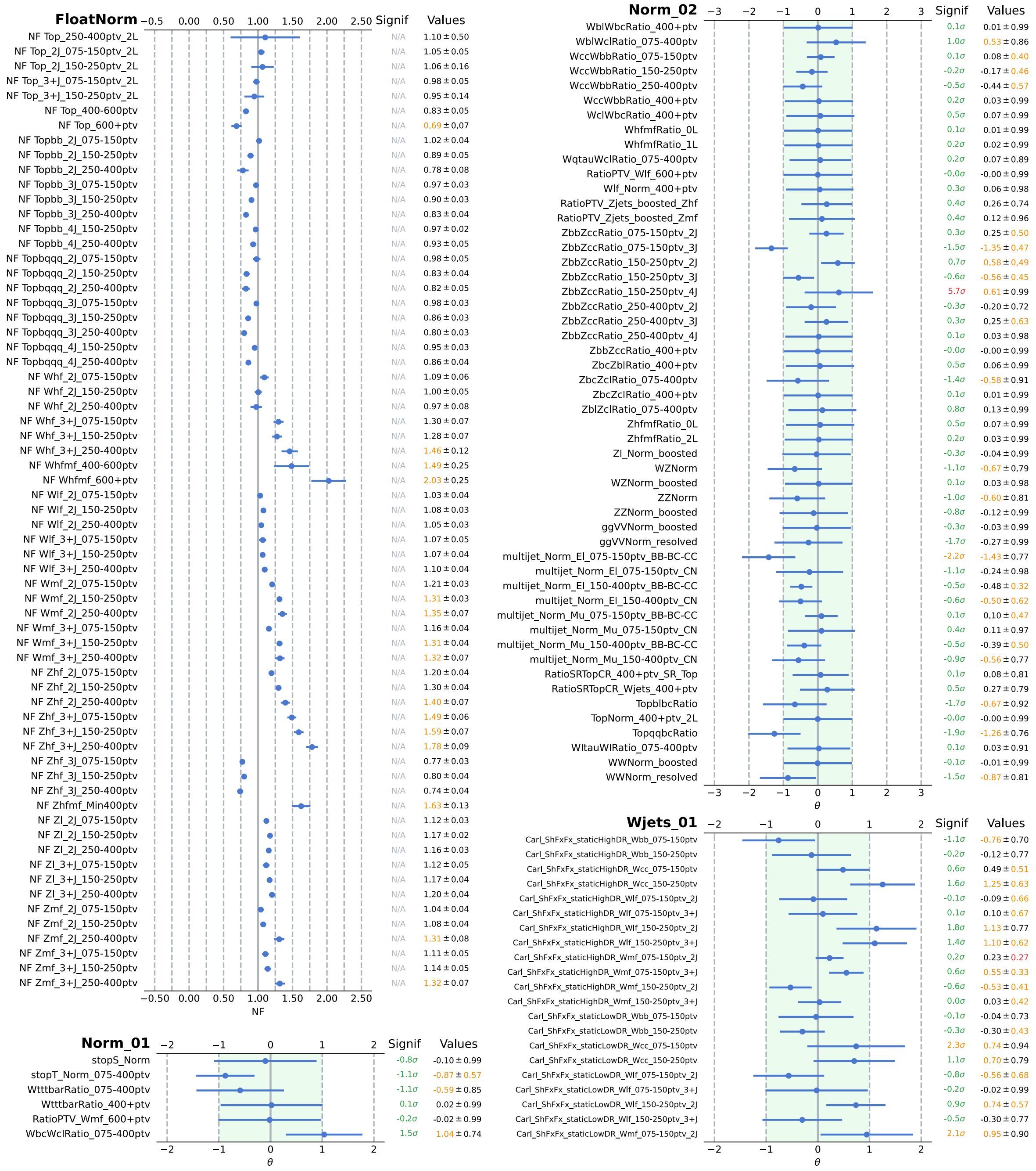


Figure H.1: The pulls on the normalisation factors and W+jets modelling uncertainties.

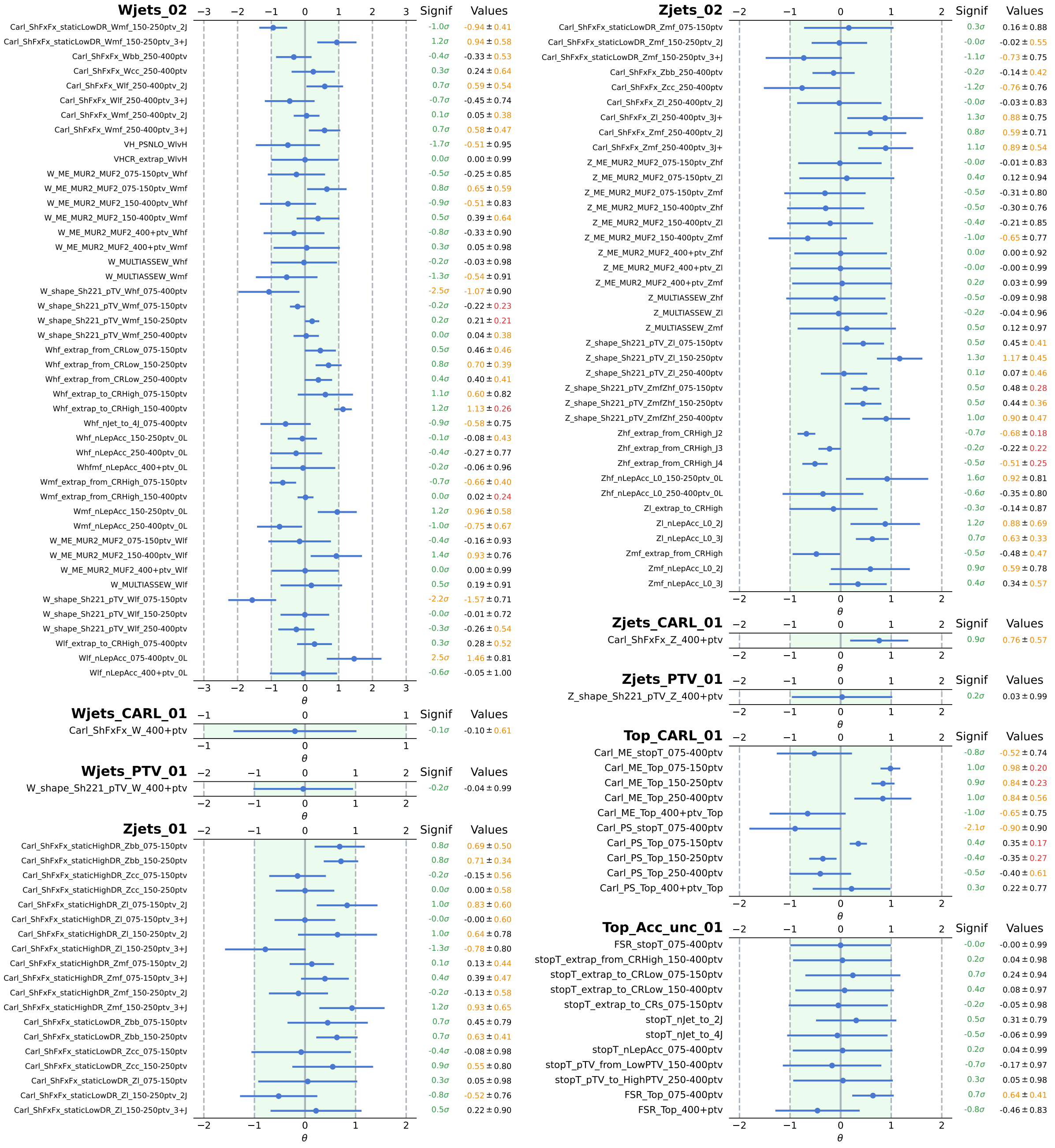


Figure H.2: The pulls on the W+jets, Z+jets and top-quark process modelling uncertainties.

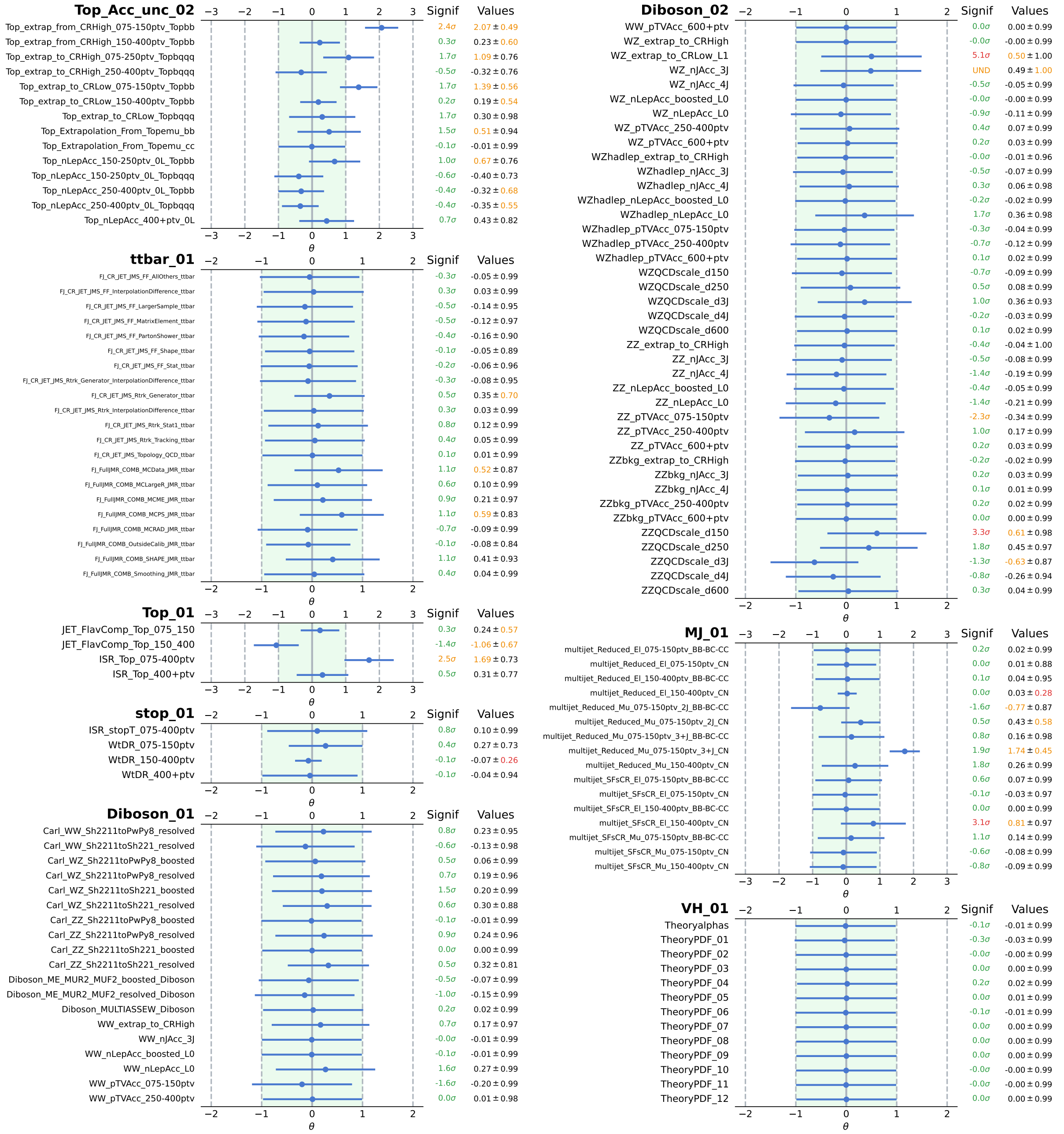


Figure H.3: The pulls on the top-quark process, diboson, multi-jet and VH signal modelling uncertainties.

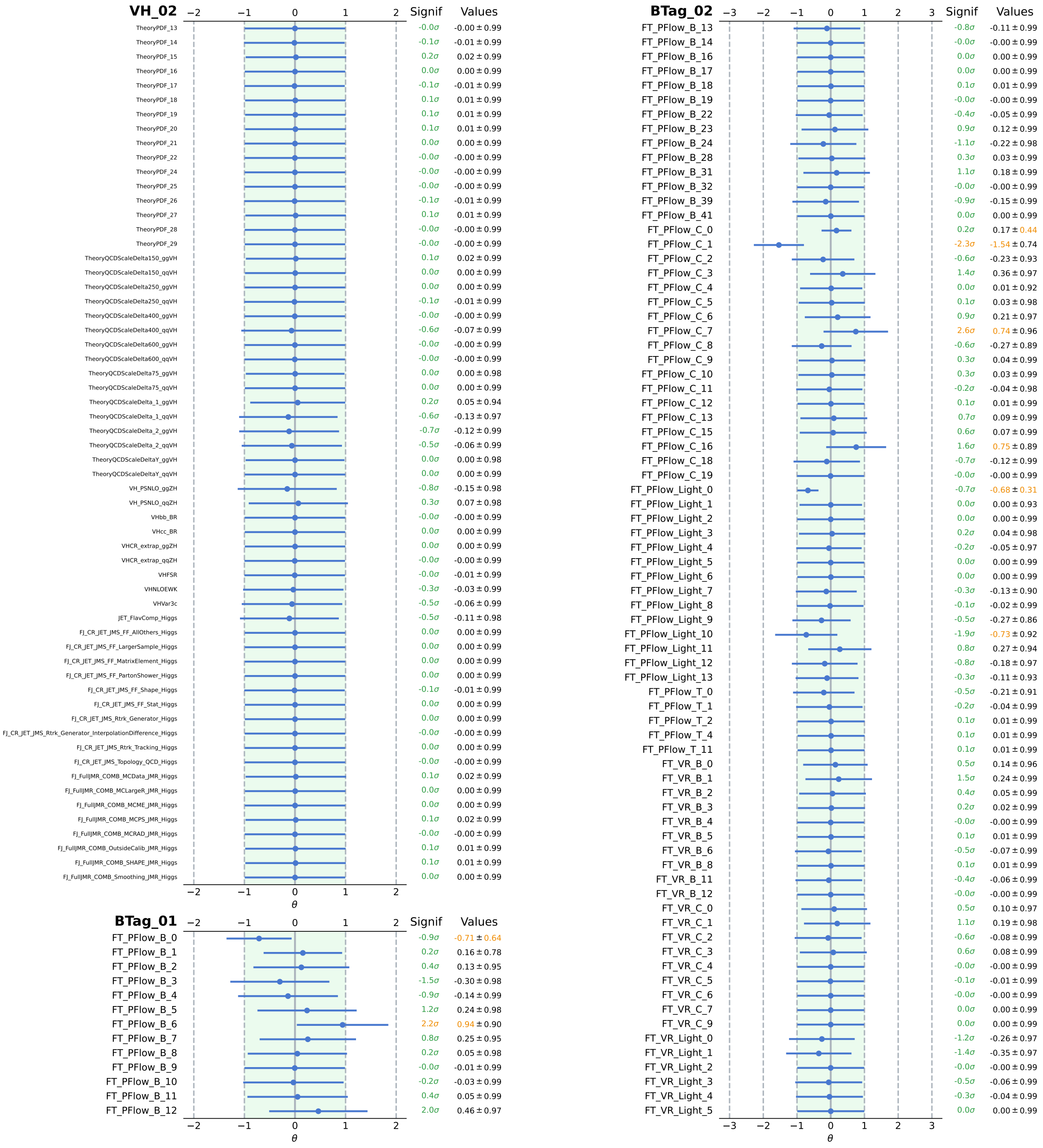


Figure H.4: The pulls on the VH signal modelling and jet-flavor tagging uncertainties.

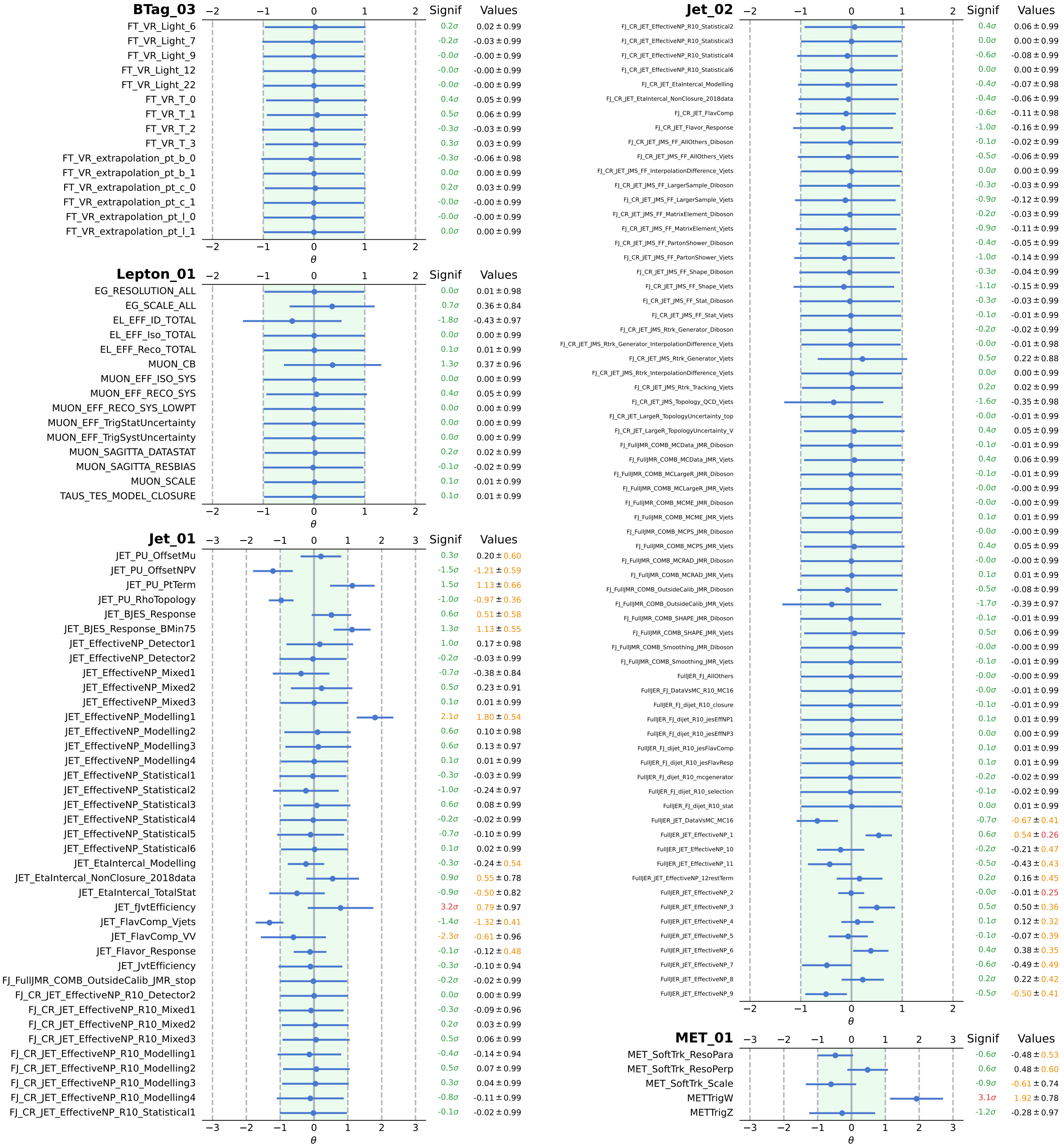


Figure H.5: The pulls on the jet-flavor tagging, lepton, jet, and $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties.

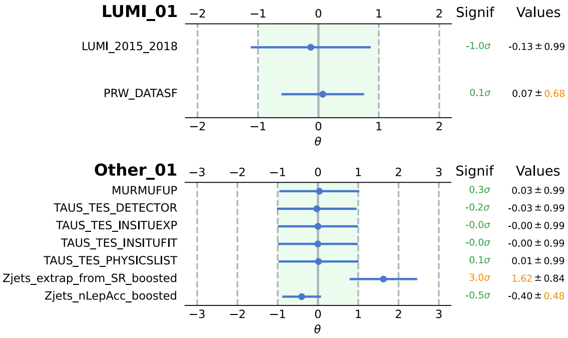


Figure H.6: The pulls on the luminosity and other uncertainties.

H.4 The ranking of nuisance parameters

Figure H.7 shows the ranking of the top 20 nuisance parameters and normalisation factors affecting the measurement of $\mu_{VH,H\to c\bar{c}}$.

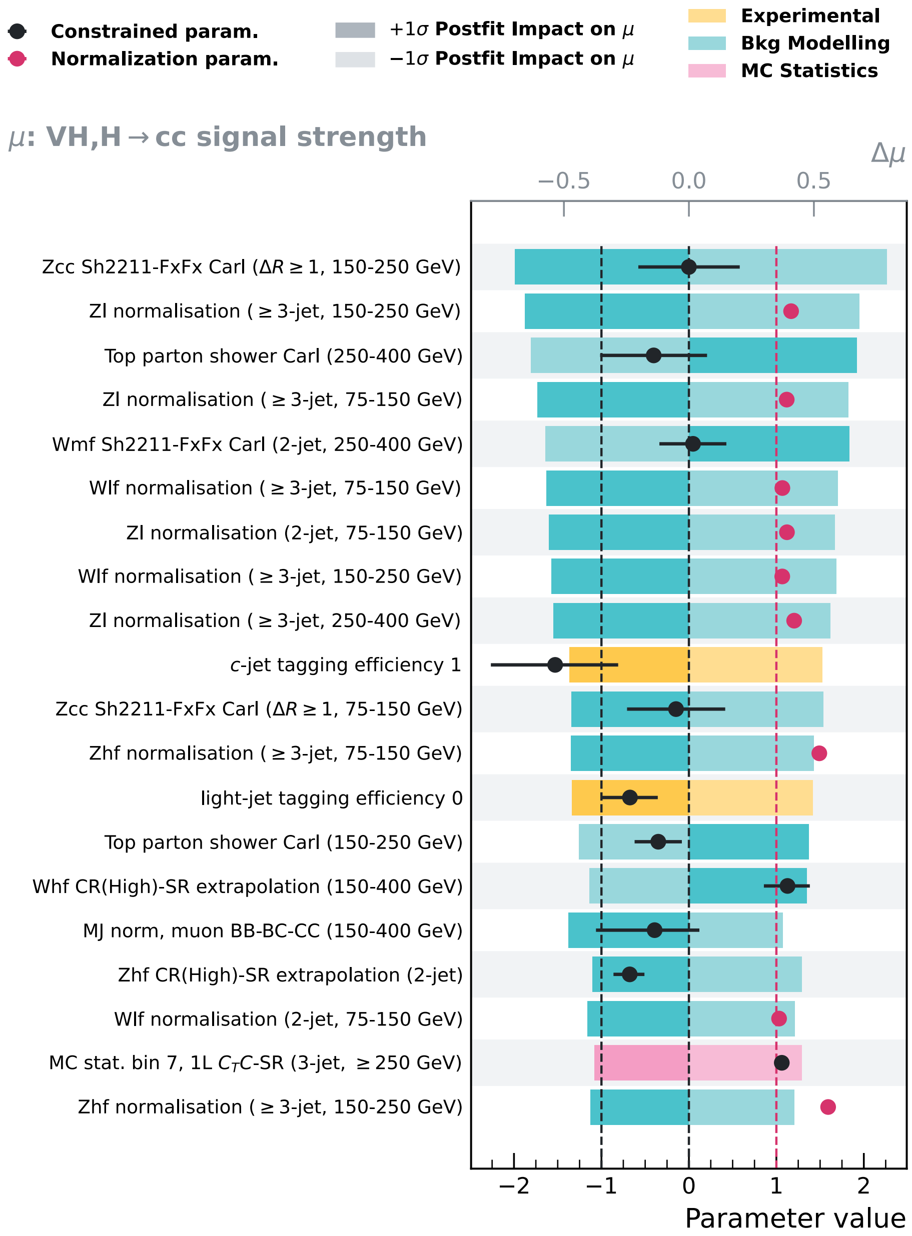


Figure H.7: The ranking of the top 20 nuisance parameters and normalisation factors affecting $\mu_{VH,H\to c\bar{c}}$.

H.5 The results of the 6POI fit

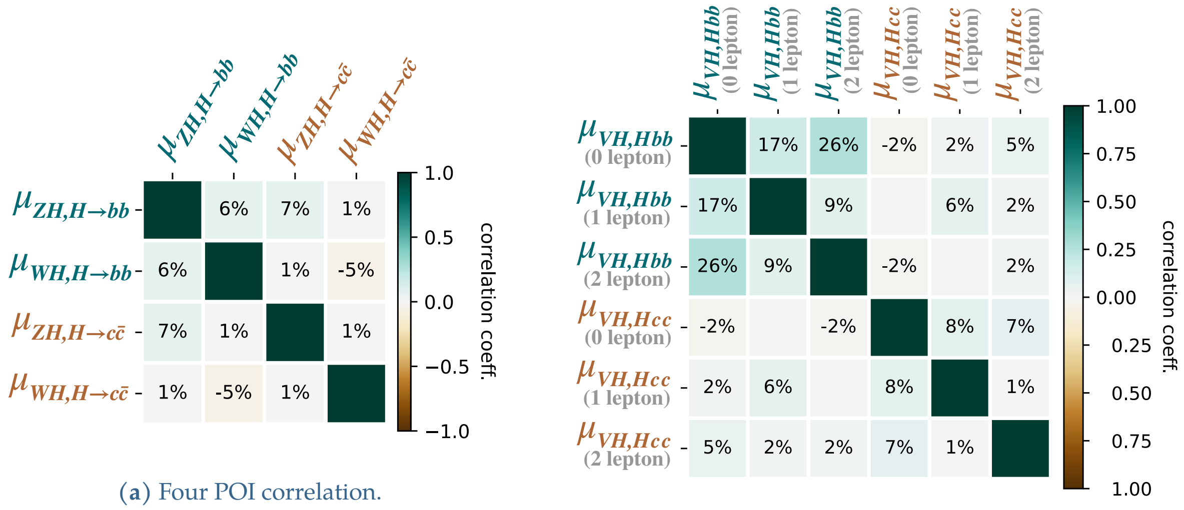
For the six POI fit, the signal strengths are obtained as,

$$
\begin{aligned}
\mu_{VH,H\to b\bar{b}}\ (\text{0 lepton}) &= 0.84^{+0.25}_{-0.23} = 0.84\ ^{+0.18}_{-0.17}\ (\text{statistical})\ ^{+0.18}_{-0.15}\ (\text{systematic}),\\
\mu_{VH,H\to b\bar{b}}\ (\text{1 lepton}) &= 0.95^{+0.21}_{-0.20} = 0.95\ ^{+0.14}_{-0.14}\ (\text{statistical})\ ^{+0.16}_{-0.14}\ (\text{systematic}),\\
\mu_{VH,H\to b\bar{b}}\ (\text{2 leptons}) &= 0.92^{+0.28}_{-0.25} = 0.92\ ^{+0.21}_{-0.21}\ (\text{statistical})\ ^{+0.18}_{-0.14}\ (\text{systematic}),\\
\mu_{VH,H\to c\bar{c}}\ (\text{0 lepton}) &= -5.7^{+8.6}_{-8.6} = -5.7\ ^{+6.8}_{-6.6}\ (\text{statistical})\ ^{+5.3}_{-5.5}\ (\text{systematic}),\\
\mu_{VH,H\to c\bar{c}}\ (\text{1 lepton}) &= 3.8^{+9.0}_{-8.6} = 3.8\ ^{+6.2}_{-6.1}\ (\text{statistical})\ ^{+6.5}_{-6.1}\ (\text{systematic}),\\
\mu_{VH,H\to c\bar{c}}\ (\text{2 leptons}) &= 4.3^{+9.2}_{-8.6} = 4.3\ ^{+7.6}_{-7.3}\ (\text{statistical})\ ^{+5.2}_{-4.5}\ (\text{systematic}).
\end{aligned}
$$

which is consistent to the standard model prediction of $\mu = 1$, within uncertainties.

H.6 The correlation matrices

The correlation matrices for the four POI signal strengths is shown in Figure H.8(a), where The correlation between the $ZH, H \to c\bar{c}$ and $WH, H \to c\bar{c}$ signal strengths are at 1%. The correlation matrices for the six POI signal strengths is shown in Figure H.8(b), where the large correlations between the zero charged lepton channel, with the one and two charged lepton channels, could be attributed to the extrapolations of th different background normalisation from the one and two charged lepton channels to the zero charged lepton channels[22].



	$\mu_{ZH,H\to bb}$	$\mu_{WH,H\to bb}$	$\mu_{ZH,H\to c\bar{c}}$	$\mu_{WH,H\to c\bar{c}}$
$\mu_{ZH,H\to bb}$		6%	7%	1%
$\mu_{WH,H\to bb}$	6%		1%	-5%
$\mu_{ZH,H\to c\bar{c}}$	7%	1%		1%
$\mu_{WH,H\to c\bar{c}}$	1%	-5%	1%	

(a) Four POI correlation.

	$\mu_{VH,Hbb}$ (0 lepton)	$\mu_{VH,Hbb}$ (1 lepton)	$\mu_{VH,Hbb}$ (2 lepton)	$\mu_{VH,Hcc}$ (0 lepton)	$\mu_{VH,Hcc}$ (1 lepton)	$\mu_{VH,Hcc}$ (2 lepton)
$\mu_{VH,Hbb}$ (0 lepton)		17%	26%	-2%	2%	5%
$\mu_{VH,Hbb}$ (1 lepton)	17%		9%		6%	2%
$\mu_{VH,Hbb}$ (2 lepton)	26%	9%		-2%		2%
$\mu_{VH,Hcc}$ (0 lepton)	-2%		-2%		8%	7%
$\mu_{VH,Hcc}$ (1 lepton)	2%	6%		8%		1%
$\mu_{VH,Hcc}$ (2 lepton)	5%	2%	2%	7%	1%	

(b) Six POI correlation.

Figure H.8: The correlation plots for the signal strengths between $\mu_{VH,H\to c\bar{c}}$ and $\mu_{VH,H\to b\bar{b}}$. Four POI corresponds to a fit where the two POI's are de-correlated for the ZH and WH processes, and six POI corresponds to a fit where the two POI's are de-correlated for the charged lepton channels.

---

[22] For example, if the $Z+$ heavy flavor background normalisation constrained in the two charged lepton channel gets smaller, signal strengths in both the zero and two charged lepton channels will be increased to increase the signal yield.

Figures H.9 to H.11 shows the correlation matrices between the signal strengths and several nuisance parameters which have large correlations with the signal strengths, for the two POI, four POI, and six P

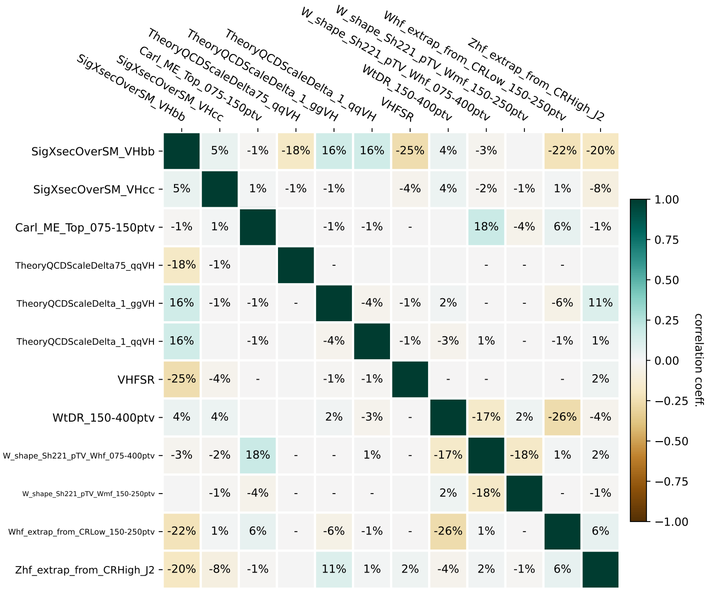


Figure H.9: The correlation of signal strengths and systematic uncertainties in the two POI fit.

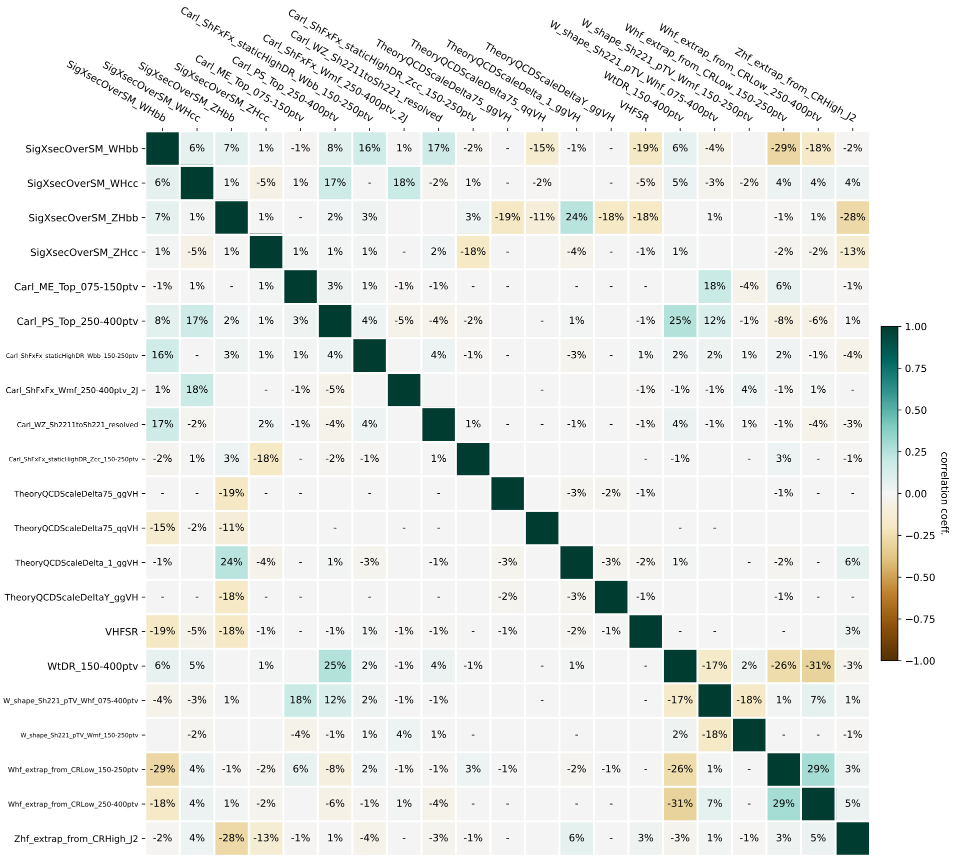
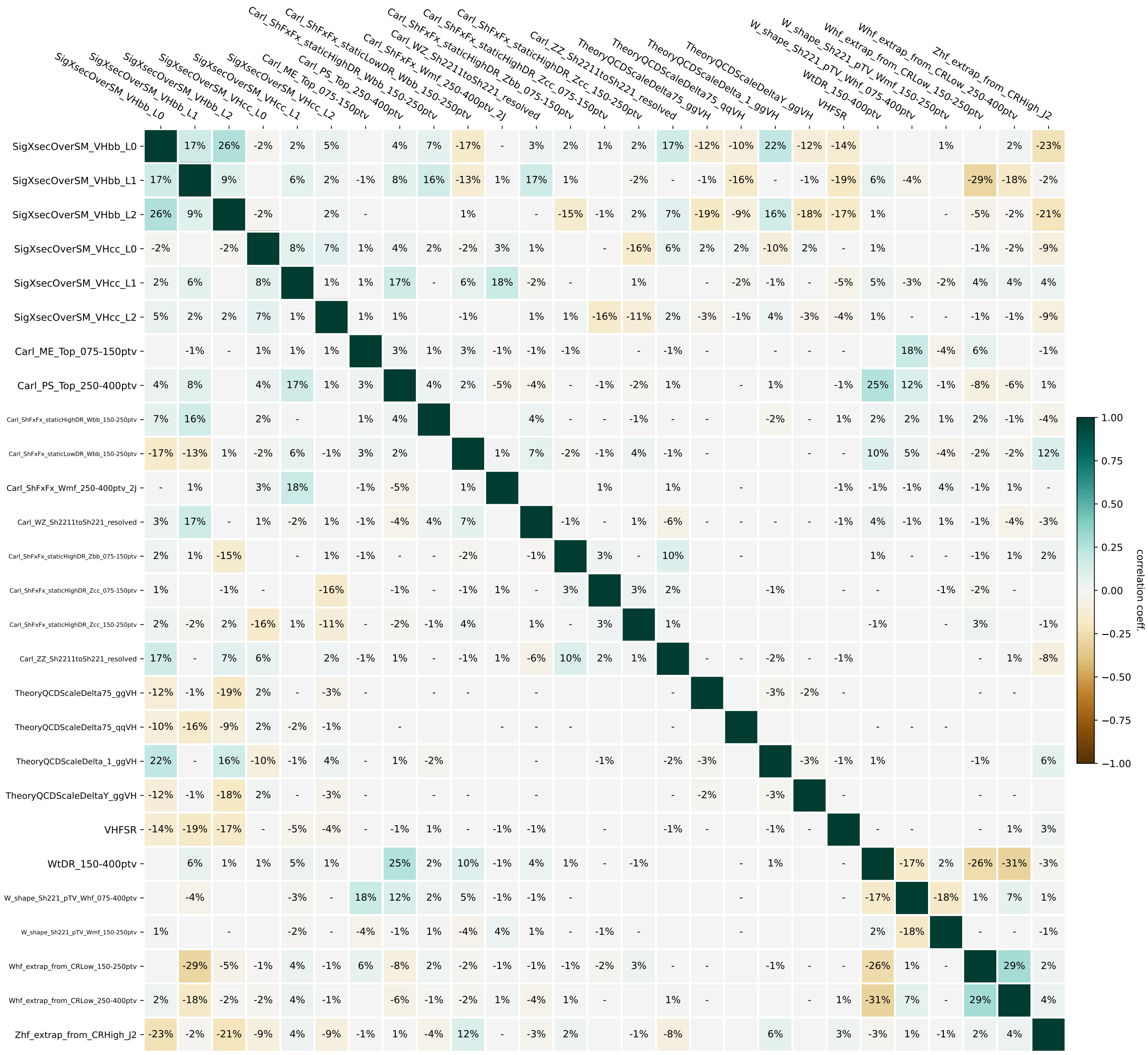


Figure H.10: The correlation of signal strengths and systematic uncertainties in the four POI fit.



	SigXsecOverSM_VHbb_L0	SigXsecOverSM_VHbb_L1	SigXsecOverSM_VHbb_L2	SigXsecOverSM_VHcc_L0	SigXsecOverSM_VHcc_L1	SigXsecOverSM_VHcc_L2	Carl_ME_Top_075-150ptv	Carl_PS_Top_250-400ptv	Carl_ShFxFx_staticHighDR_Wbb_150-250ptv	Carl_ShFxFx_staticLowDR_Wbb_150-250ptv	Carl_ShFxFx_Wmf_250-400ptv_2J	Carl_WZ_Sh2211toSh221_resolved	Carl_ShFxFx_staticHighDR_Zbb_075-150ptv	Carl_ShFxFx_staticHighDR_Zcc_075-150ptv	Carl_ShFxFx_staticHighDR_Zcc_150-250ptv	Carl_ZZ_Sh2211toSh221_resolved	TheoryQCDScaleDelta75_ggVH	TheoryQCDScaleDelta75_qqVH	TheoryQCDScaleDelta_1_ggVH	TheoryQCDScaleDeltaY_ggVH	VHFSR	WtDR_150-400ptv	W_shape_Sh221_pTV_Whf_075-400ptv	W_shape_Sh221_pTV_Wmf_150-250ptv	Whf_extrap_from_CRLow_150-250ptv	Whf_extrap_from_CRLow_250-400ptv	Zhf_extrap_from_CRHigh_J2
SigXsecOverSM_VHbb_L0		17%	26%	-2%	2%	5%		4%	7%	-17%	-	3%	2%	1%	2%	17%	-12%	-10%	22%	-12%	-14%			1%		2%	-23%
SigXsecOverSM_VHbb_L1	17%		9%		6%	2%	-1%	8%	16%	-13%	1%	17%	1%		-2%	-	-1%	-16%	-	-1%	-19%	6%	-4%		-29%	-18%	-2%
SigXsecOverSM_VHbb_L2	26%	9%		-2%		2%	-			1%		-	-15%	-1%	2%	7%	-19%	-9%	16%	-18%	-17%	1%		-	-5%	-2%	-21%
SigXsecOverSM_VHcc_L0	-2%		-2%		8%	7%	1%	4%	2%	-2%	3%	1%		-	-16%	6%	2%	2%	-10%	2%	-	1%			-1%	-2%	-9%
SigXsecOverSM_VHcc_L1	2%	6%		8%		1%	1%	17%	-	6%	18%	-2%	-		1%		-	-2%	-1%	-	-5%	5%	-3%	-2%	4%	4%	4%
SigXsecOverSM_VHcc_L2	5%	2%	2%	7%	1%		1%	1%		-1%		1%	1%	-16%	-11%	2%	-3%	-1%	4%	-3%	-4%	1%	-	-	-1%	-1%	-9%
Carl_ME_Top_075-150ptv		-1%	-	1%	1%	1%		3%	1%	3%	-1%	-1%	-1%		-	-1%	-		-	-	-		18%	-4%	6%		-1%
Carl_PS_Top_250-400ptv	4%	8%		4%	17%	1%	3%		4%	2%	-5%	-4%	-	-1%	-2%	1%		-	1%		-1%	25%	12%	-1%	-8%	-6%	1%
Carl_ShFxFx_staticHighDR_Wbb_150-250ptv	7%	16%		2%	-		1%	4%				4%	-	-	-1%	-	-		-2%	-	1%	2%	2%	1%	2%	-1%	-4%
Carl_ShFxFx_staticLowDR_Wbb_150-250ptv	-17%	-13%	1%	-2%	6%	-1%	3%	2%			1%	7%	-2%	-1%	4%	-1%	-			-	-	10%	5%	-4%	-2%	-2%	12%
Carl_ShFxFx_Wmf_250-400ptv_2J	-	1%		3%	18%		-1%	-5%		1%				1%		1%		-			-1%	-1%	-1%	4%	-1%	1%	-
Carl_WZ_Sh2211toSh221_resolved	3%	17%	-	1%	-2%	1%	-1%	-4%	4%	7%			-1%	-	1%	-6%	-	-	-	-	-1%	4%	-1%	1%	-1%	-4%	-3%
Carl_ShFxFx_staticHighDR_Zbb_075-150ptv	2%	1%	-15%		-	1%	-1%	-	-	-2%		-1%		3%	-	10%		-				1%	-	-	-1%	1%	2%
Carl_ShFxFx_staticHighDR_Zcc_075-150ptv	1%		-1%	-		-16%		-1%	-	-1%	1%	-	3%		3%	2%			-1%		-	-		-1%	-2%	-	
Carl_ShFxFx_staticHighDR_Zcc_150-250ptv	2%	-2%	2%	-16%	1%	-11%	-	-2%	-1%	4%		1%	-	3%		1%						-1%		-	3%		-1%
Carl_ZZ_Sh2211toSh221_resolved	17%	-	7%	6%		2%	-1%	1%	-	-1%	1%	-6%	10%	2%	1%		-	-	-2%	-	-1%				-	1%	-8%
TheoryQCDScaleDelta75_ggVH	-12%	-1%	-19%	2%	-	-3%	-		-	-		-				-			-3%	-2%	-				-	-	
TheoryQCDScaleDelta75_qqVH	-10%	-16%	-9%	2%	-2%	-1%		-			-	-	-			-						-	-	-			
TheoryQCDScaleDelta_1_ggVH	22%	-	16%	-10%	-1%	4%	-	1%	-2%			-		-1%		-2%	-3%			-3%	-1%	1%			-1%		6%
TheoryQCDScaleDeltaY_ggVH	-12%	-1%	-18%	2%	-	-3%	-		-	-		-				-	-2%		-3%		-				-	-	
VHFSR	-14%	-19%	-17%	-	-5%	-4%	-	-1%	1%	-	-1%	-1%		-		-1%	-		-1%	-		-	-		-	1%	3%
WtDR_150-400ptv		6%	1%	1%	5%	1%		25%	2%	10%	-1%	4%	1%	-	-1%			-	1%		-		-17%	2%	-26%	-31%	-3%
W_shape_Sh221_pTV_Whf_075-400ptv		-4%			-3%	-	18%	12%	2%	5%	-1%	-1%	-					-			-	-17%		-18%	1%	7%	1%
W_shape_Sh221_pTV_Wmf_150-250ptv	1%		-		-2%	-	-4%	-1%	1%	-4%	4%	1%	-	-1%	-			-				2%	-18%		-	-	-1%
Whf_extrap_from_CRLow_150-250ptv		-29%	-5%	-1%	4%	-1%	6%	-8%	2%	-2%	-1%	-1%	-1%	-2%	3%	-	-		-1%	-	-	-26%	1%	-		29%	2%
Whf_extrap_from_CRLow_250-400ptv	2%	-18%	-2%	-2%	4%	-1%		-6%	-1%	-2%	1%	-4%	1%	-		1%	-			-	1%	-31%	7%	-	29%		4%
Zhf_extrap_from_CRHigh_J2	-23%	-2%	-21%	-9%	4%	-9%	-1%	1%	-4%	12%	-	-3%	2%		-1%	-8%			6%		3%	-3%	1%	-1%	2%	4%	

Figure H.11: The correlation of signal strengths and systematic uncertainties in the six POI fit.

H.7 Data and MC distributions for Δ R control regions

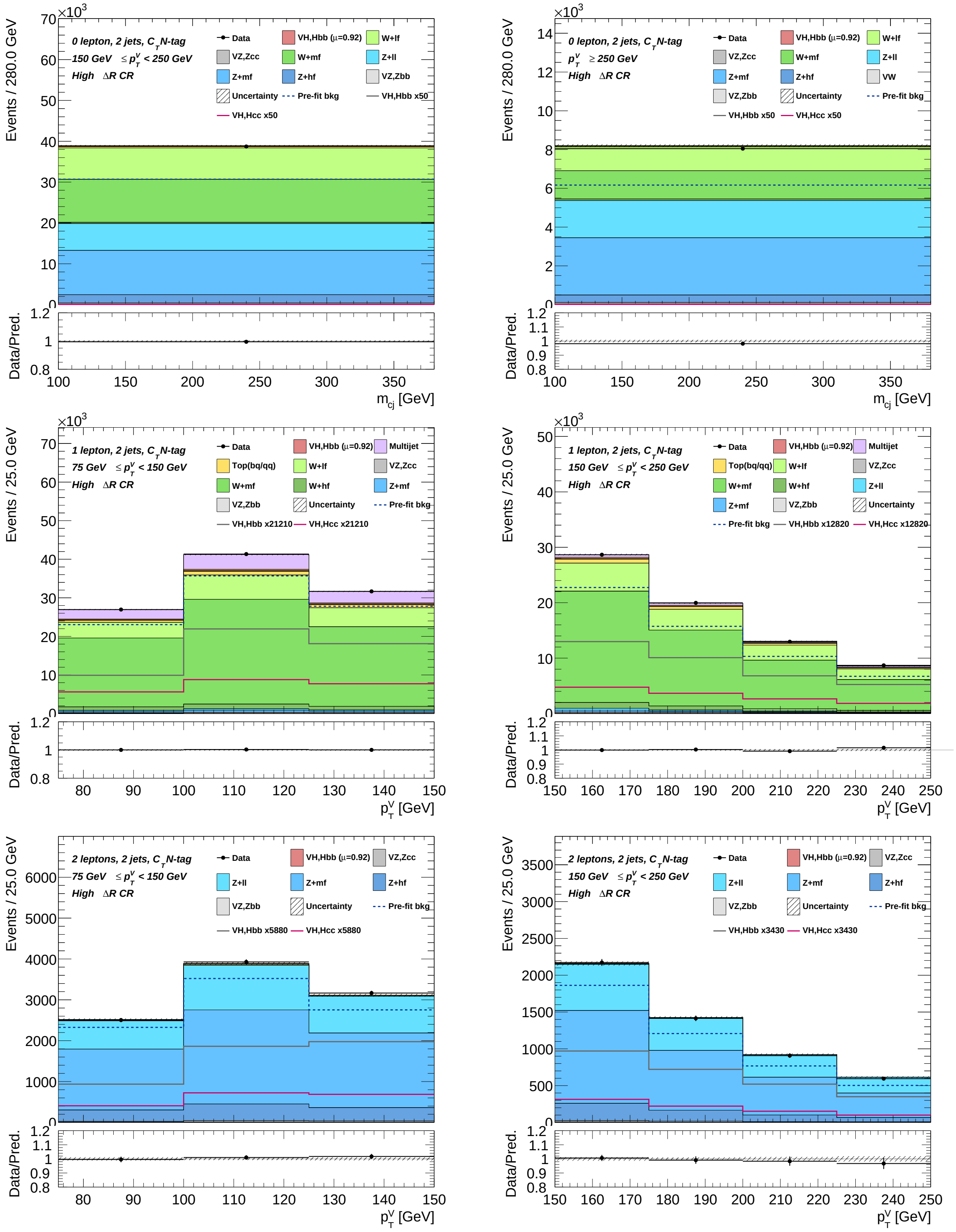


Figure H.12: Post-fit distributions for the high ΔR control region of the $C_T N$-tag region in the respective charged lepton channels.

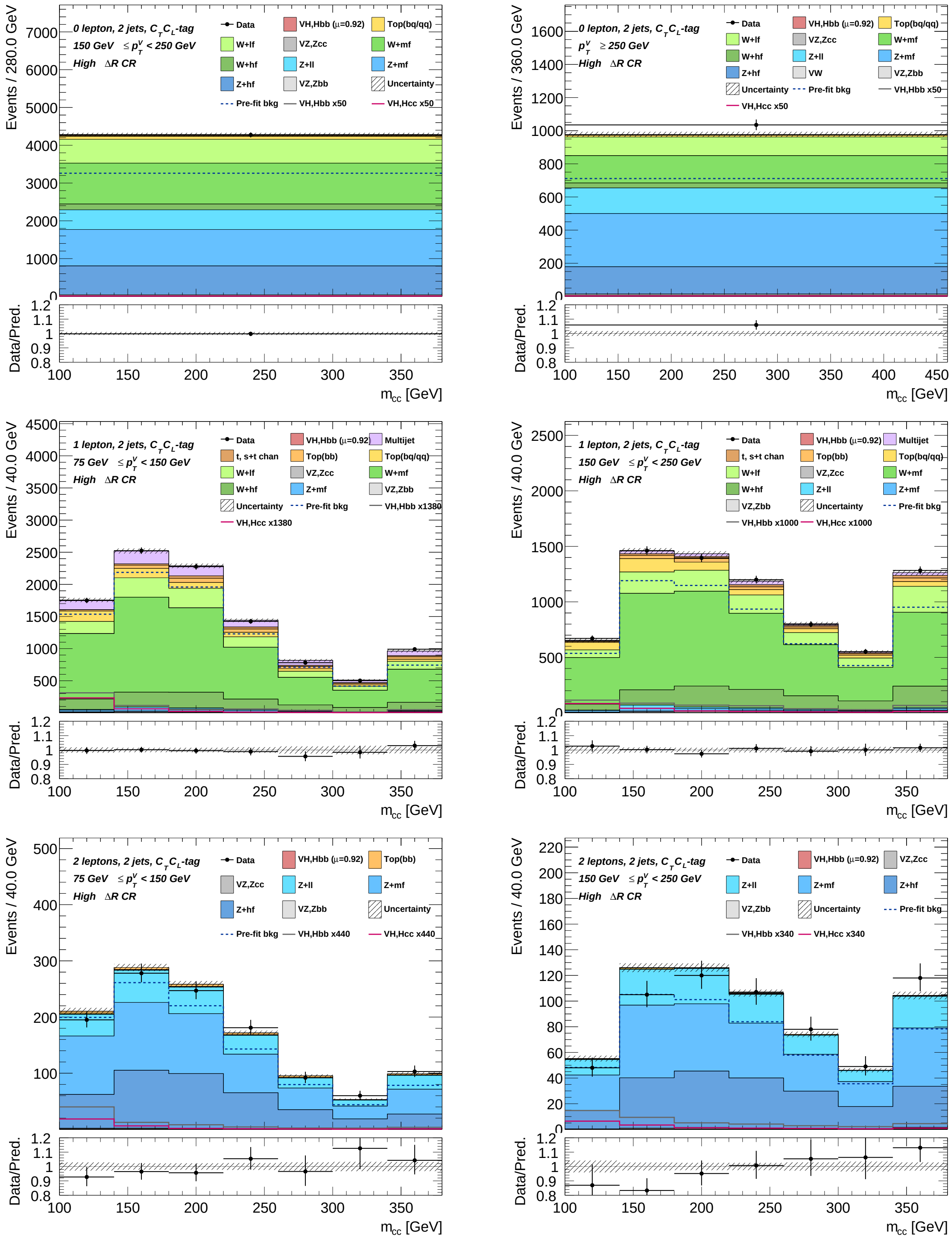


Figure H.13: Post-fit distributions for the high ΔR control region of the $C_T C_L$-tag region in the respective charged lepton channels.

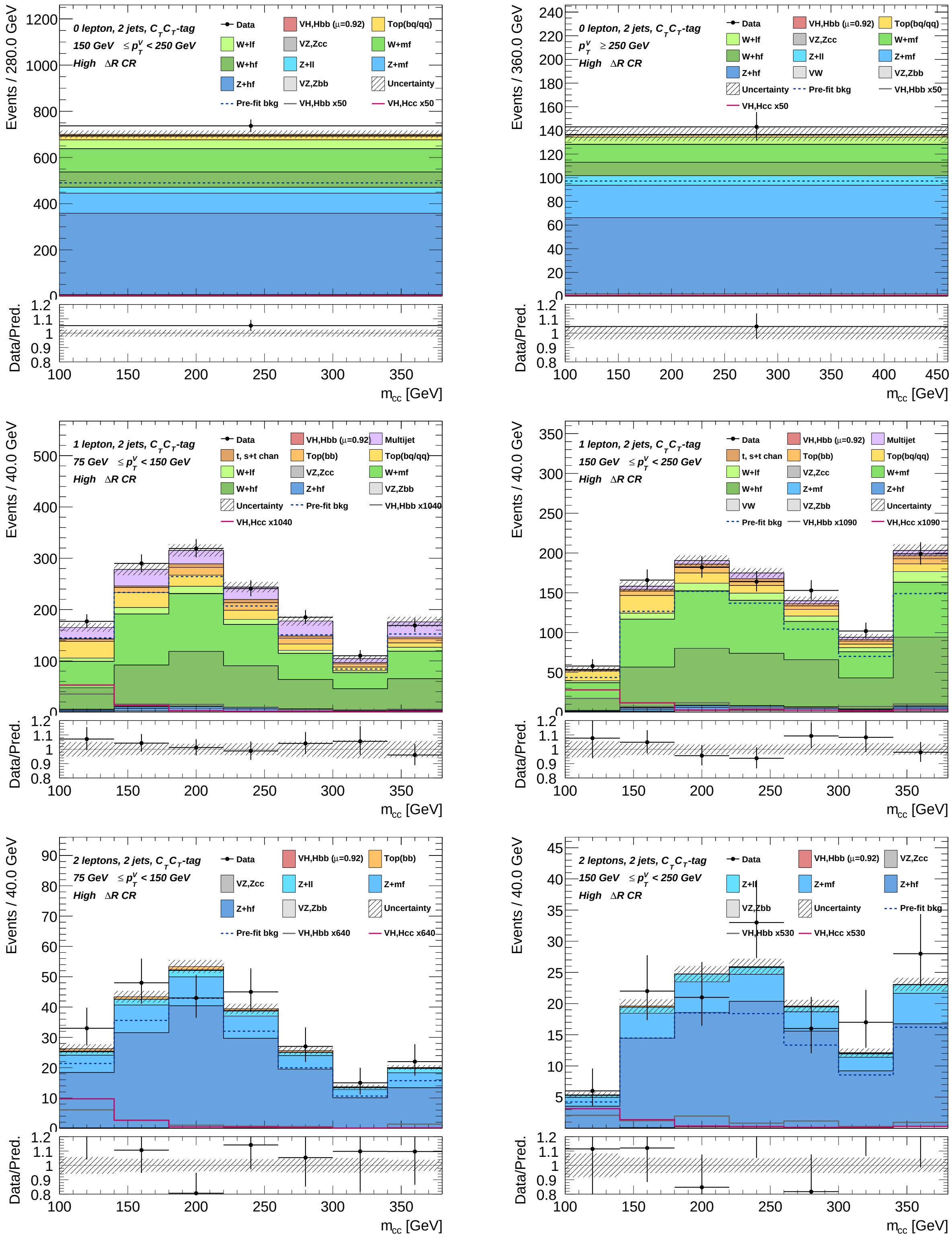


Figure H.14: Post-fit distributions for the high ΔR control region of the $C_T C_T$-tag region in the one and two charged lepton channels.

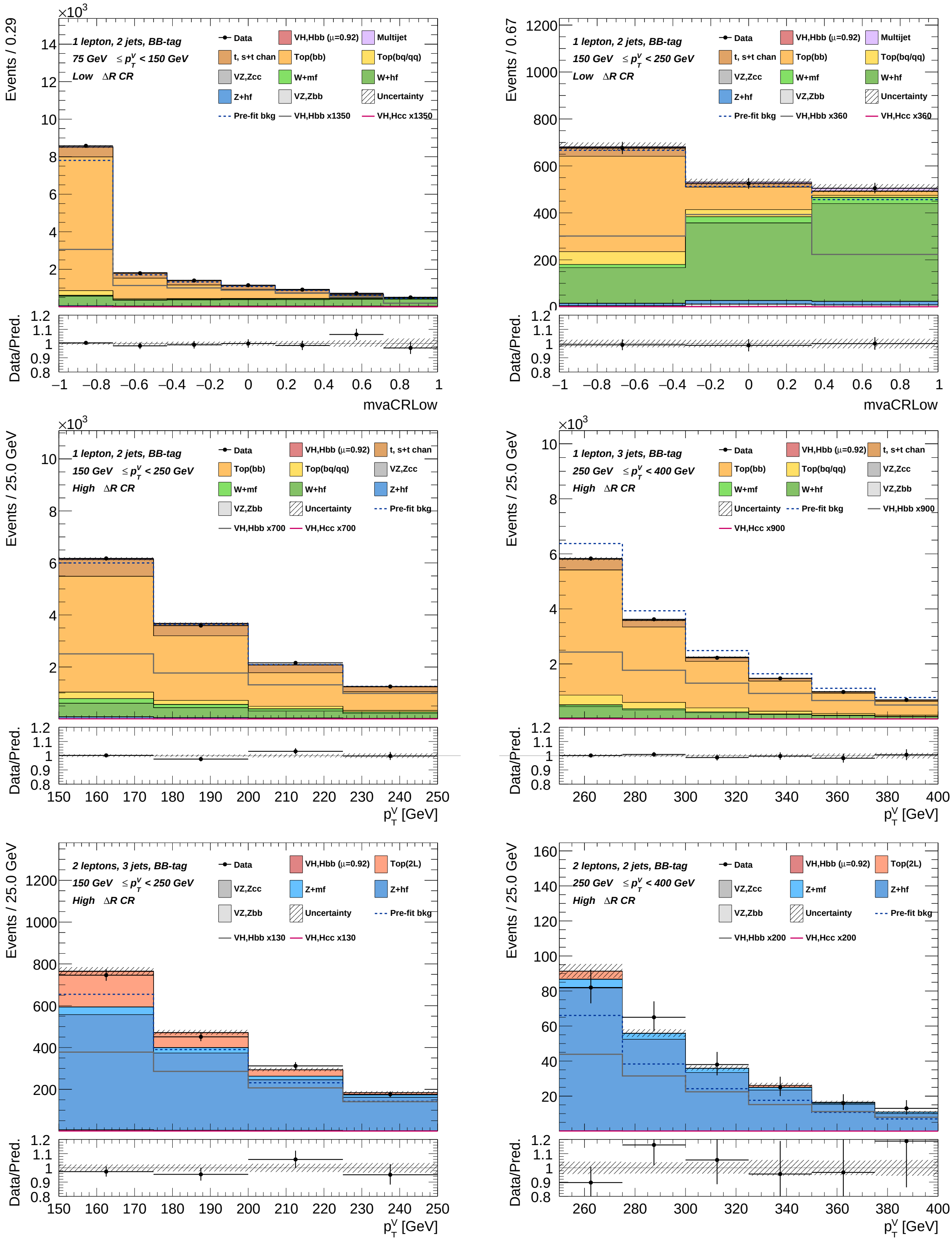


Figure H.15: Post-fit distributions for the low ΔR and high ΔR control region of the BB-tag region in the one and two charged lepton channels.

H.8 Data and MC distributions for top process control regions

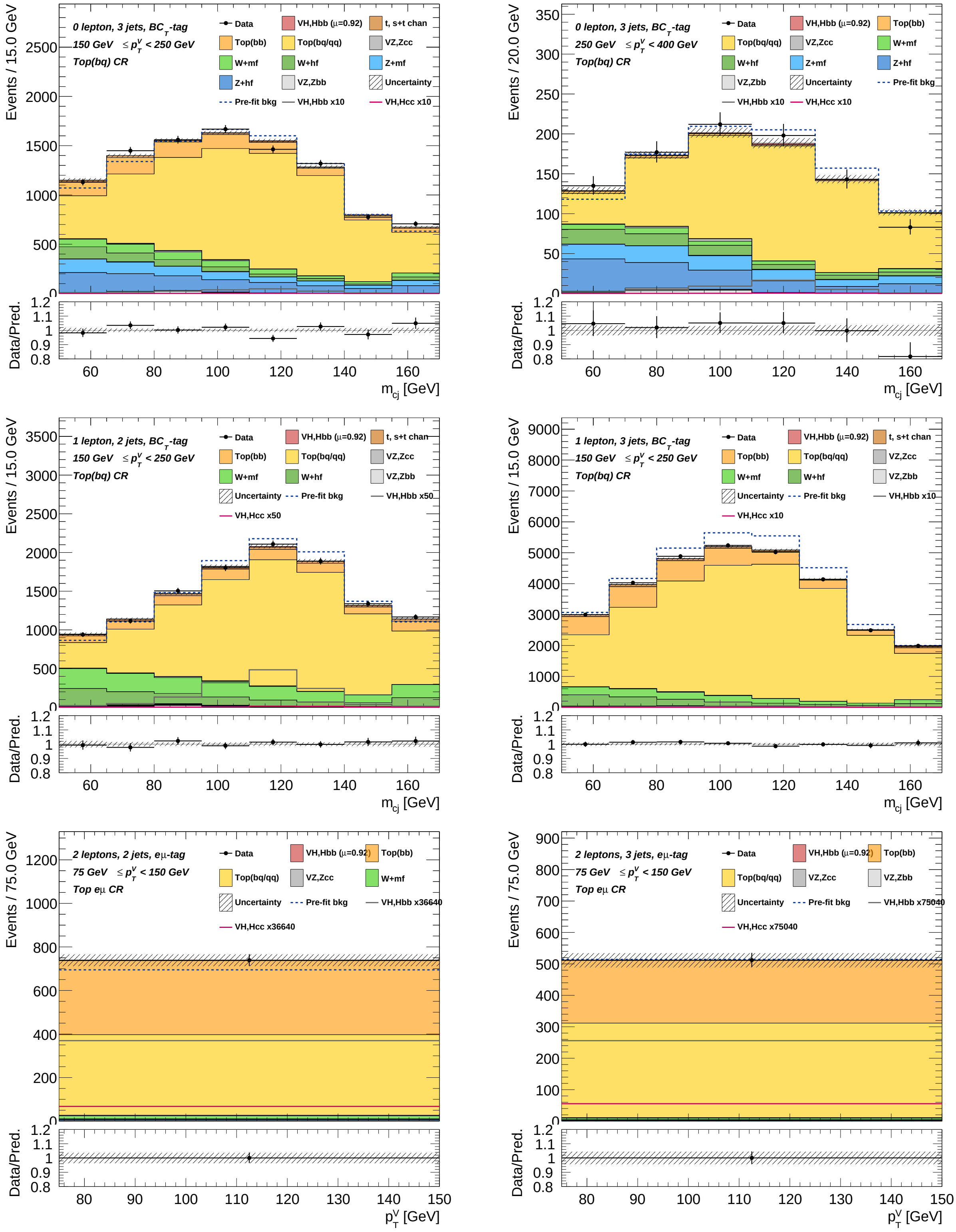


Figure H.16: Post-fit distributions for the top bq control region (i.e the BC_T-tag region) and the top $e\mu$ control regions in the respective charged lepton channels.

H.9 Data and MC distributions for V+light jet control regions

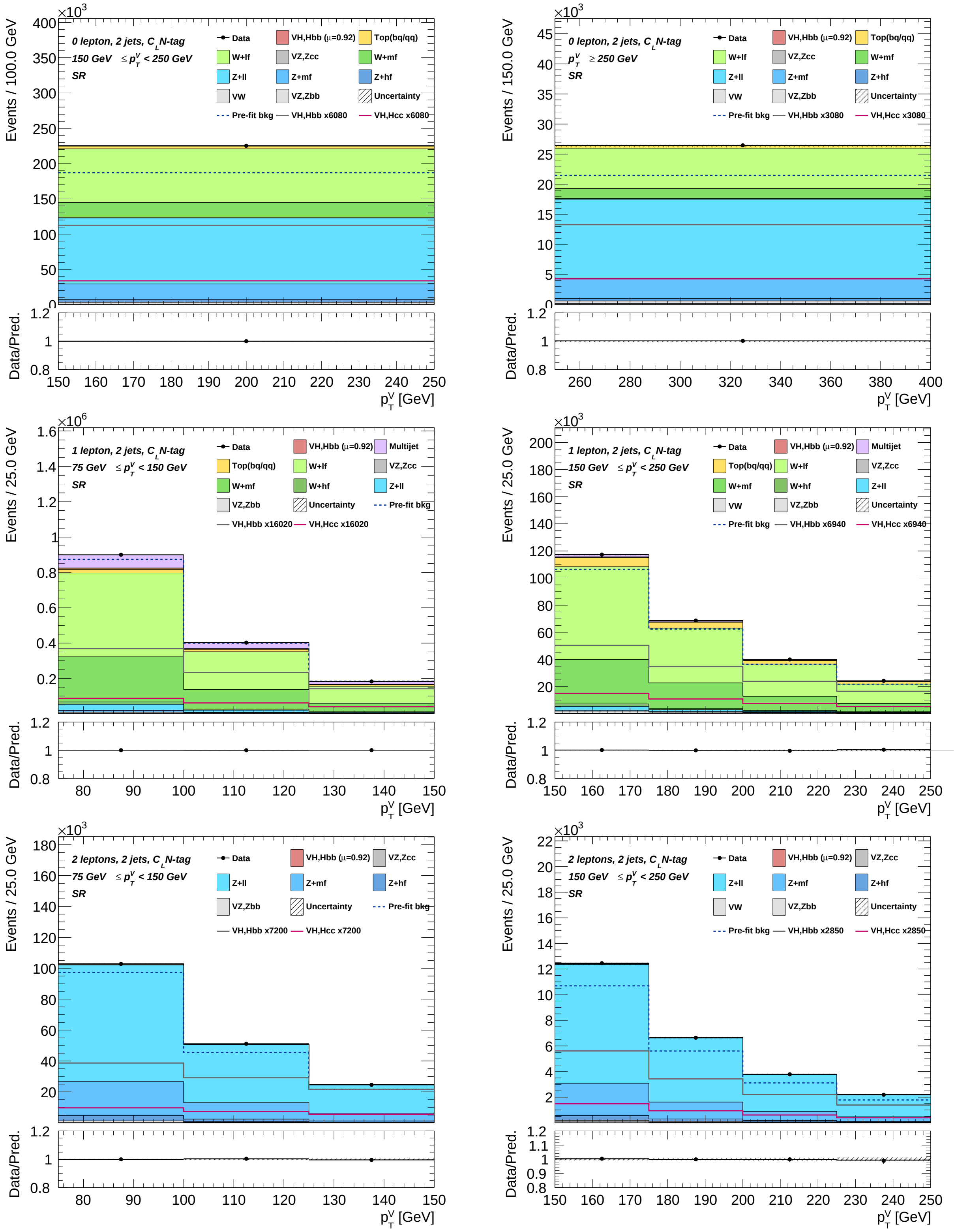


Figure H.17: Post-fit distributions for the V+light jet control region (i.e the $C_L N$-tag region) in the respective charged lepton channels.

References

[1] 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].

[2] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 2013.

[3] Ahmed Abokhalil. "The Higgs Mechanism and Higgs Boson: Unveiling the Symmetry of the Universe". In: (2023). URL: https://arxiv.org/abs/2306.01019.

[4] *Standard Model Summary Plots June 2024*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2024-011. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2903866.

[5] ATLAS Collaboration. "Measurement of the total cross section and ρ-parameter from elastic scattering in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (2022). arXiv: 2207.12246 [hep-ex].

[6] ATLAS Collaboration. "Measurement of inclusive jet and dijet cross-sections in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 05 (2018), p. 195. DOI: 10.1007/JHEP05(2018)195. arXiv: 1711.02692 [hep-ex].

[7] ATLAS Collaboration. "Precise measurements of W- and Z-boson transverse momentum spectra with the ATLAS detector using pp collisions at $\sqrt{s} = 5.02$ TeV and 13 TeV". In: *The European Physical Journal C* 84.10 (Oct. 2024). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-024-13414-0. URL: http://dx.doi.org/10.1140/epjc/s10052-024-13414-0.

[8] ATLAS Collaboration. "Inclusive and differential cross-sections for dilepton $t\bar{t}$ production measured in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector". In: *Journal of High Energy Physics* 2023.7 (July 2023). ISSN: 1029-8479. DOI: 10.1007/jhep07(2023)141. URL: http://dx.doi.org/10.1007/JHEP07(2023)141.

[9] ATLAS Collaboration. "Measurement of t-channel production of single top quarks and antiquarks in pp collisions at 13 TeV using the full ATLAS Run 2 data sample". In: *Journal of High Energy Physics* 2024.5 (May 2024). ISSN: 1029-8479. DOI: 10.1007/jhep05(2024)305. URL: http://dx.doi.org/10.1007/JHEP05(2024)305.

[10] ATLAS Collaboration. "Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with ATLAS". In: *JHEP* 01 (2018), p. 063. DOI: 10.1007/JHEP01(2018)063. arXiv: 1612.07231 [hep-ex].

[11] ATLAS Collaboration. "Measurement of single top-quark production in the s-channel in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (2022). arXiv: 2209.08990 [hep-ex].

[12] ATLAS Collaboration. "Measurement of fiducial and differential W^+W^- production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 79 (2019), p. 884. DOI: 10.1140/epjc/s10052-019-7371-6. arXiv: 1905.04242 [hep-ex].

[13] ATLAS Collaboration. "Measurement of $W^\pm Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 79 (2019), p. 535. DOI: 10.1140/epjc/s10052-019-7027-6. arXiv: 1902.05759 [hep-ex].

[14] ATLAS Collaboration. "$ZZ \to \ell^+\ell^-\ell'^+\ell'^-$ cross-section measurements and search for anomalous triple gauge couplings in 13 TeV pp collisions with the ATLAS detector". In: *Phys. Rev. D* 97 (2018), p. 032005. DOI: 10.1103/PhysRevD.97.032005. arXiv: 1709.07703 [hep-ex].

[15] 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: *JHEP* 05 (2023), p. 028. DOI: 10.1007/JHEP05(2023)028. arXiv: 2207.08615 [hep-ex].

[16] LHC Higgs Cross Section Working Group. *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*. en. 2017. DOI: 10.23731/CYRM-2017-002. URL: https://e-publishing.cern.ch/index.php/CYRM/issue/view/32.

[17] Particle Data Group et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2022.8 (Aug. 2022), p. 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptac097. eprint: https://academic.oup.com/ptep/article-pdf/2022/8/083C01/49175539/ptac097.pdf. URL: https://doi.org/10.1093/ptep/ptac097.

[18] LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*. en. 2013. DOI: 10.5170/CERN-2013-004. URL: http://cds.cern.ch/record/1559921.

[19] A. Bazavov et al. "Up-, down-, strange-, charm-, and bottom-quark masses from four-flavor lattice QCD". In: *Phys. Rev. D* 98 (5 Sept. 2018), p. 054517. DOI: 10.1103/PhysRevD.98.054517. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.054517.

[20] ATLAS Collaboration. "A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery". In: *Nature* 607 (2022), pp. 52–59. DOI: 10.1038/s41586-022-04893-w. arXiv: 2207.00092 [hep-ex].

[21] 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].

[22] 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].

[23] ATLAS Collaboration. "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector". In: *Eur. Phys. J. C* 75 (2015), p. 476. DOI: 10.1140/epjc/s10052-015-3685-1. arXiv: 1506.05669 [hep-ex]. Erratum: in: *Eur. Phys. J. C* 76 (2016), p. 152. DOI: 10.1140/epjc/s10052-016-3934-y.

[24] 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].

[25] ATLAS Collaboration. *Search for the associated production of charm quarks and a Higgs boson decaying into a photon pair with the ATLAS detector*. 2024. arXiv: 2407.15550 [hep-ex]. URL: https://arxiv.org/abs/2407.15550.

[26] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[27] CMS Collaboration. “Search for Higgs boson decays into Z and J/ψ and for Higgs and Z boson decays into J/ψ or Y pairs in pp collisions at s=13 TeV”. In: *Physics Letters B* 842 (2023), p. 137534. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2022.137534. URL: https://www.sciencedirect.com/science/article/pii/S0370269322006682.

[28] *Search for Higgs boson production in association with a charm quark in the diphoton decay channel*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2905239.

[29] CMS Collaboration. “Search for Higgs Boson and Observation of Z Boson through Their Decay into a Charm Quark-Antiquark Pair in Boosted Topologies in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV”. In: *Phys. Rev. Lett.* 131 (4 July 2023), p. 041801. DOI: 10.1103/PhysRevLett.131.041801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.131.041801.

[30] 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].

[31] CMS Collaboration. “Search for Higgs Boson Decay to a Charm Quark-Antiquark Pair in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV”. In: *Phys. Rev. Lett.* 131 (6 Aug. 2023), p. 061801. DOI: 10.1103/PhysRevLett.131.061801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.131.061801.

[32] 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].

[33] Fady Bishara et al. “Constraining Light-Quark Yukawa Couplings from Higgs Distributions”. In: *Phys. Rev. Lett.* 118 (12 Mar. 2017), p. 121801. DOI: 10.1103/PhysRevLett.118.121801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.118.121801.

[34] 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].

[35] ATLAS Collaboration. “Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Lett. B* 816 (2021), p. 136204. DOI: 10.1016/j.physletb.2021.136204. arXiv: 2008.02508 [hep-ex].

[36] ATLAS Collaboration. “ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset”. In: (2022). arXiv: 2211.16345 [physics.data-an].

[37] Gilles Louppe, Kyle Cranmer, and Juan Pavez. *carl: a likelihood-free inference toolbox*. Mar. 2016. DOI: 10.5281/zenodo.47798. URL: http://dx.doi.org/10.5281/zenodo.47798.

[38] Philippe Mouche. “Overall view of the LHC. Vue d’ensemble du LHC”. In: (2014). General Photo. URL: https://cds.cern.ch/record/1708847.

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

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

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

[42] The CMS Collaboration. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004. DOI: `10.1088/1748-0221/3/08/S08004`. URL: `https://dx.doi.org/10.1088/1748-0221/3/08/S08004`.

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

[44] Mike Lamont. *LHC report: Run 1 - the final flurry*. `https://home.cern/news/news/accelerators/lhc-report-run-1-final-flurry`. Accessed: 2024-11-25.

[45] Eckhard Elsen and Frédérick Bordry. *A successful conclusion to Run 2*. `https://home.cern/news/opinion/accelerators/successful-conclusion-run-2`. Accessed: 2024-11-25.

[46] Rende Steerenberg. *Accelerator Report: LHC Run 3 achieves record-breaking integrated luminosity*. `https://home.cern/news/news/accelerators/accelerator-report-lhc-run-3-achieves-record-breaking-integrated-luminosity`. Accessed: 2024-11-25.

[47] CERN. *New schedule for CERN's accelerators*. `https://home.cern/news/news/accelerators/new-schedule-cerns-accelerators`. Accessed: 2024-11-25.

[48] Ewa Lopienska. "The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022". In: (2022). General Photo. URL: `https://cds.cern.ch/record/2800984`.

[49] O. Brüning, H. Burkhardt, and S. Myers. "The large hadron collider". In: *Progress in Particle and Nuclear Physics* 67.3 (2012), pp. 705–734. ISSN: 0146-6410. DOI: `https://doi.org/10.1016/j.ppnp.2012.03.001`. URL: `https://www.sciencedirect.com/science/article/pii/S0146641012000695`.

[50] Daniel Boussard and Trevor Paul R Linnecar. *The LHC Superconducting RF System*. Tech. rep. Geneva: CERN, 1999. URL: `https://cds.cern.ch/record/410377`.

[51] CERN. *Accelerating: Radiofrequency cavities*. `https://home.cern/science/engineering/accelerating-radiofrequency-cavities`. Accessed: 2024-11-25.

[52] Ana Lopes and Melissa Loyse Perrey. "FAQ-LHC The guide". 2022. URL: `https://cds.cern.ch/record/2809109`.

[53] CERN. *Facts and figures about the LHC*. `https://home.cern/resources/faqs/facts-and-figures-about-lhc`. Accessed: 2024-11-23.

[54] Karsten Eggert, K Honkavaara, and Andreas Morsch. *Luminosity considerations for the LHC*. Tech. rep. Geneva: CERN, 1994. URL: `https://cds.cern.ch/record/260711`.

[55] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC". In: (2022). arXiv: `2212.09379 [hep-ex]`.

[56] 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]`.

[57] Joao Pequenao. "Computer generated image of the whole ATLAS detector". 2008. URL: `https://cds.cern.ch/record/1095924`.

[58] Joao Pequenao. "Computer generated image of the ATLAS inner detector". 2008. URL: https://cds.cern.ch/record/1095926.

[59] ATLAS Collaboration. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *Eur. Phys. J. C* 77 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[60] ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report, Volume 2*. ATLAS-TDR-5, CERN-LHCC-97-017. 1997. URL: https://cds.cern.ch/record/331064.

[61] ATLAS Collaboration. *ATLAS Pixel Detector: Technical Design Report*. ATLAS-TDR-11; CERN-LHCC-98-013. 1998. URL: https://cds.cern.ch/record/381263.

[62] ATLAS Collaboration. "Alignment of the ATLAS Inner Detector in Run-2". In: *Eur. Phys. J. C* 80 (2020), p. 1194. DOI: 10.1140/epjc/s10052-020-08700-6. arXiv: 2007.07624 [hep-ex].

[63] ATLAS Collaboration. *ATLAS Insertable B-Layer: Technical Design Report*. ATLAS-TDR-19; CERN-LHCC-2010-013. 2010. URL: https://cds.cern.ch/record/1291633. Addendum: ATLAS-TDR-19-ADD-1; CERN-LHCC-2012-009. 2012. URL: https://cds.cern.ch/record/1451888.

[64] B. Abbott et al. "Production and integration of the ATLAS Insertable B-Layer". In: *JINST* 13 (2018), T05008. DOI: 10.1088/1748-0221/13/05/T05008. arXiv: 1803.00844 [physics.ins-det].

[65] ATLAS Collaboration. "Operation and performance of the ATLAS semiconductor tracker in LHC Run 2". In: *JINST* 17 (2021), P01013. DOI: 10.1088/1748-0221/17/01/P01013. arXiv: 2109.02591 [physics.ins-det].

[66] A. Abdesselam et al. "The barrel modules of the ATLAS semiconductor tracker". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 568.2 (2006), pp. 642–671. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2006.08.036. URL: https://www.sciencedirect.com/science/article/pii/S016890020601388X.

[67] A. Abdesselam et al. "The ATLAS semiconductor tracker end-cap module". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 575.3 (2007), pp. 353–389. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2007.02.019. URL: https://www.sciencedirect.com/science/article/pii/S0168900207003270.

[68] F. Campabadal et al. "Design and performance of the ABCD3TA ASIC for readout of silicon strip detectors in the ATLAS semiconductor tracker". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 552.3 (2005), pp. 292–328. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2005.07.002. URL: https://www.sciencedirect.com/science/article/pii/S0168900205013926.

[69] ATLAS Collaboration. "Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties". In: *JINST* 12 (2017), P05002. DOI: 10.1088/1748-0221/12/05/P05002. arXiv: 1702.06473 [hep-ex].

[70] John Jackson. *Classical Electrodynamics*. John Wiley and Sons, 1999.

[71] Joao Pequenao. "Computer Generated image of the ATLAS calorimeter". 2008. URL: https://cds.cern.ch/record/1095927.

[72] 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].

[73] Joao Pequenao. “Computer generated image of the ATLAS Muons subsystem”. 2008. URL: https://cds.cern.ch/record/1095929.

[74] ATLAS Collaboration. *ATLAS New Small Wheel: Technical Design Report*. ATLAS-TDR-020; CERN-LHCC-2013-006. 2013. URL: https://cds.cern.ch/record/1552862.

[75] ATLAS Collaboration. “Commissioning of the ATLAS Muon Spectrometer with cosmic rays”. In: *Eur. Phys. J. C* 70 (2010), p. 875. DOI: 10.1140/epjc/s10052-010-1415-2. arXiv: 1006.4384 [hep-ex].

[76] ATLAS Collaboration. *ATLAS Muon Spectrometer: Technical Design Report*. ATLAS-TDR-10; CERN-LHCC-97-022. CERN, 1997. URL: https://cds.cern.ch/record/331068.

[77] ATLAS Collaboration. “Resolution of the ATLAS muon spectrometer monitored drift tubes in LHC Run 2”. In: *JINST* 14 (2019), P09011. DOI: 10.1088/1748-0221/14/09/P09011. arXiv: 1906.12226 [hep-ex].

[78] 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].

[79] ATLAS Collaboration. “Performance of the ATLAS RPC detector and Level-1 muon barrel trigger at $\sqrt{s} = 13$ TeV”. In: *JINST* 16 (2021), P07029. DOI: 10.1088/1748-0221/16/07/P07029. arXiv: 2103.01029 [hep-ex].

[80] ATLAS Collaboration. *ATLAS*. WORLD SCIENTIFIC, 2019. DOI: 10.1142/11030. eprint: https://www.worldscientific.com/doi/pdf/10.1142/11030. URL: https://www.worldscientific.com/doi/abs/10.1142/11030.

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

[82] 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].

[83] Rosa Simoniello. “The ATLAS Level-1 Topological Processor: from design to routine usage in Run-2”. In: (2018), pp. 1–4. DOI: 10.1109/NSSMIC.2018.8824280.

[84] A. Sidoti. “Minimum Bias Trigger Scintillators in ATLAS Run II”. In: *Journal of Instrumentation* 9.10 (Oct. 2014), p. C10020. DOI: 10.1088/1748-0221/9/10/C10020. URL: https://dx.doi.org/10.1088/1748-0221/9/10/C10020.

[85] ATLAS Collaboration. *Athena*. Version 22.0.1. Apr. 2019. DOI: 10.5281/zenodo.2641997. URL: https://doi.org/10.5281/zenodo.2641997.

[86] 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].

[87] 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].

[88] ATLAS Collaboration. *Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays to charm quarks with the ATLAS detector*. ATLAS-CONF-2021-021. 2021. URL: https://cds.cern.ch/record/2771724.

[89] 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].

[90] P.J. Clark. "The ATLAS Detector Simulation". In: *Nuclear Physics B - Proceedings Supplements* 215.1 (2011). Proceedings of the 12th Topical Seminar on Innovative Particle and Radiation Detectors (IPRD10), pp. 85–88. ISSN: 0920-5632. DOI: `https://doi.org/10.1016/j.nuclphysbps.2011.03.142`. URL: `https://www.sciencedirect.com/science/article/pii/S092056321100212X`.

[91] A. Schälicke et al. "An event generator for particle production in high-energy collisions". In: *Progress in Particle and Nuclear Physics* 53.1 (2004). Heavy Ion Reaction from Nuclear to Quark Matter, pp. 329–338. ISSN: 0146-6410. DOI: `https://doi.org/10.1016/j.ppnp.2004.02.031`. URL: `https://www.sciencedirect.com/science/article/pii/S014664100400047X`.

[92] Zachary Marshall and the ATLAS Collaboration. "Simulation of Pile-up in the ATLAS Experiment". In: *Journal of Physics: Conference Series* 513.2 (June 2014), p. 022024. DOI: `10.1088/1742-6596/513/2/022024`. URL: `https://dx.doi.org/10.1088/1742-6596/513/2/022024`.

[93] B. Andersson et al. "Parton fragmentation and string dynamics". In: *Physics Reports* 97.2 (1983), pp. 31–145. ISSN: 0370-1573. DOI: `https://doi.org/10.1016/0370-1573(83)90080-7`. URL: `https://www.sciencedirect.com/science/article/pii/0370157383900807`.

[94] G. Marchesini and B.R. Webber. "Simulation of QCD jets including soft gluon interference". In: *Nuclear Physics B* 238.1 (1984), pp. 1–29. ISSN: 0550-3213. DOI: `https://doi.org/10.1016/0550-3213(84)90463-2`. URL: `https://www.sciencedirect.com/science/article/pii/0550321384904632`.

[95] ATLAS Collaboration. "Measurement of distributions sensitive to the underlying event in inclusive Z boson production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 79 (2019), p. 666. DOI: `10.1140/epjc/s10052-019-7162-0`. arXiv: `1905.09752 [hep-ex]`.

[96] Jacob J. Ethier and Emanuele R. Nocera. "Parton Distributions in Nucleons and Nuclei". In: *Annual Review of Nuclear and Particle Science* 70.Volume 70, 2020 (2020), pp. 43–76. ISSN: 1545-4134. DOI: `https://doi.org/10.1146/annurev-nucl-011720-042725`. URL: `https://www.annualreviews.org/content/journals/10.1146/annurev-nucl-011720-042725`.

[97] G. 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: *Journal of High Energy Physics* 2013.10 (Oct. 2013), p. 083. ISSN: 1029-8479. DOI: `10.1007/JHEP10(2013)0833`. URL: `https://doi.org/10.1007/JHEP10(2013)083`.

[98] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. "A brief introduction to PYTHIA 8.1". In: *Computer Physics Communications* 178.11 (2008), pp. 852–867. ISSN: 0010-4655. DOI: `https://doi.org/10.1016/j.cpc.2008.01.036`. URL: `https://www.sciencedirect.com/science/article/pii/S0010465508000441`.

[99] The NNPDF collaboration. "Parton distributions for the LHC run II". In: *Journal of High Energy Physics* 2015.4 (Apr. 2015), p. 040. ISSN: 1029-8479. DOI: `10.1007/JHEP04(2015)040`. URL: `https://doi.org/10.1007/JHEP04(2015)040`.

[100] 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]`.

[101] John Campbell, Tobias Neumann, and Zack Sullivan. "Single-top-quark production in the t-channel at NNLO". In: *JHEP* 02 (2021), p. 040. DOI: 10.1007/JHEP02(2021)040. arXiv: 2012.01574 [hep-ph].

[102] Nikolaos Kidonakis and Nodoka Yamanaka. "Higher-order corrections for tW production at high-energy hadron colliders". In: *JHEP* 05 (2021), p. 278. DOI: 10.1007/JHEP05(2021)278. arXiv: 2102.11300 [hep-ph].

[103] LHC Top Working Group. *NLO single-top channel cross sections*. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec. Accessed: 2024-11-13. Nov. 2022.

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

[105] Stefano Frixione et al. "Single-top hadroproduction in association with a W boson". In: *Journal of High Energy Physics* 2008.07 (July 2008), p. 029. DOI: 10.1088/1126-6708/2008/07/029. URL: https://dx.doi.org/10.1088/1126-6708/2008/07/029.

[106] ATLAS Collaboration. "Modelling and computational improvements to the simulation of single vector-boson plus jet processes for the ATLAS experiment". In: *JHEP* 08 (2021), p. 089. DOI: 10.1007/JHEP08(2022)089. arXiv: 2112.09588 [hep-ex].

[107] Enrico Bothmann et al. "Event generation with Sherpa 2.2". In: *SciPost Phys.* 7 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. URL: https://scipost.org/10.21468/SciPostPhys.7.3.034.

[108] Frank Krauss and Davide Napoletano. "Towards a fully massive five-flavor scheme". In: *Phys. Rev. D* 98 (9 Nov. 2018), p. 096002. DOI: 10.1103/PhysRevD.98.096002. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.096002.

[109] S. Schumann and F. Krauss. "A parton shower algorithm based on Catani-Seymour dipole factorisation". In: *Journal of High Energy Physics* 2008.03 (Mar. 2008), p. 038. DOI: 10.1088/1126-6708/2008/03/038. URL: https://dx.doi.org/10.1088/1126-6708/2008/03/038.

[110] J.C Winter, F. Krauss, and G. Soff. "A modified cluster-hadronisation model". In: *The European Physical Journal C - Particles and Fields* 36.3 (Aug. 2008). ISSN: 1434-6052. DOI: 10.1140/epjc/s2004-01960-8. URL: https://doi.org/10.1140/epjc/s2004-01960-8.

[111] ATLAS Collaboration. *Multi-Boson Simulation for 13 TeV ATLAS Analyses*. ATL-PHYS-PUB-2017-005. 2017. URL: https://cds.cern.ch/record/2261933.

[112] Fabrizio Caola et al. "QCD corrections to ZZ production in gluon fusion at the LHC". In: *Phys. Rev. D* 92 (9 Nov. 2015), p. 094028. DOI: 10.1103/PhysRevD.92.094028. URL: https://link.aps.org/doi/10.1103/PhysRevD.92.094028.

[113] 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.

[114] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data". In: *JINST* 14 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[115] ATLAS Collaboration. "Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2". In: *Eur. Phys. J. C* 77 (2017), p. 673. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex].

[116] G Borissov et al. "ATLAS strategy for primary vertex reconstruction during Run-2 of the LHC". In: *Journal of Physics: Conference Series* 664.7 (Dec. 2015), p. 072041. DOI: 10.1088/

1742-6596/664/7/072041. URL: https://dx.doi.org/10.1088/1742-6596/664/7/072041.

[117] Wolfgang Waltenberger, Rudolf Frühwirth, and Pascal Vanlaer. "Adaptive vertex fitting". In: *Journal of Physics G: Nuclear and Particle Physics* 34.12 (Nov. 2007), N343. DOI: 10.1088/0954-3899/34/12/N01. URL: https://dx.doi.org/10.1088/0954-3899/34/12/N01.

[118] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data". In: *JINST* 19 (2023), P02009. DOI: 10.1088/1748-0221/19/02/P02009. arXiv: 2309.05471 [hep-ex].

[119] ATLAS Collaboration. "Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment". In: *JHEP* 05 (2024), p. 162. DOI: 10.1007/JHEP05(2024)162. arXiv: 2308.13362 [hep-ex].

[120] ATLAS Collaboration. "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 81 (2021), p. 578. DOI: 10.1140/epjc/s10052-021-09233-2. arXiv: 2012.00578 [hep-ex].

[121] ATLAS Collaboration. "Studies of the muon momentum calibration and performance of the ATLAS detector with pp collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 83 (2023), p. 686. DOI: 10.1140/epjc/s10052-023-11584-x. arXiv: 2212.07338 [hep-ex].

[122] 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].

[123] ATLAS Collaboration. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ 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].

[124] ATLAS Collaboration. *Forward jet vertex tagging using the particle flow algorithm*. ATL-PHYS-PUB-2019-026. 2019. URL: https://cds.cern.ch/record/2683100.

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

[126] ATLAS Collaboration. "Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector". In: *Eur. Phys. J. C* 82 (2021), p. 95. DOI: 10.1140/epjc/s10052-021-09843-w. arXiv: 2109.10627 [hep-ex].

[127] ATLAS Collaboration. "Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using $139\,\text{fb}^{-1}$ of ATLAS proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 83 (2023), p. 728. DOI: 10.1140/epjc/s10052-023-11736-z. arXiv: 2301.06319 [hep-ex].

[128] ATLAS Collaboration. *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*. ATL-PHYS-PUB-2019-033. 2019. URL: https://cds.cern.ch/record/2688062.

[129] ATLAS Collaboration. "The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using $140\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ pp collisions". In: (2024). arXiv: 2402.05858 [hep-ex].

[130] ATLAS Collaboration. *Flavor Tagging Efficiency Parametrisations with Graph Neural Networks*. ATL-PHYS-PUB-2022-041. 2022. URL: https://cds.cern.ch/record/2825433.

[131] *What Are Graph Neural Networks?* https://blogs.nvidia.com/blog/what-are-graph-neural-networks/. Accessed: 2024-10-17.

[132] *A Gentle Introduction to Graph Neural Networks*. `https://distill.pub/2021/gnn-intro/`. Accessed: 2024-10-17.

[133] Huilin Qu and Loukas Gouskos. "Jet tagging via particle clouds". In: *Phys. Rev. D* 101 (5 Mar. 2020), p. 056019. DOI: `10.1103/PhysRevD.101.056019`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.101.056019`.

[134] ATLAS Collaboration. *Graph Neural Network Jet Flavour Tagging with the ATLAS Detector*. ATL-PHYS-PUB-2022-027. 2022. URL: `https://cds.cern.ch/record/2811135`.

[135] *Deep Graph Library*. `https://www.dgl.ai`. Accessed: 2024-10-17.

[136] *PyTorch*. `https://pytorch.org`. Accessed: 2024-10-17.

[137] F. A. Di Bello et al. *Efficiency Parameterization with Neural Networks*. May 2021. DOI: `10.1007/s41781-021-00059-x`. URL: `https://doi.org/10.1007/s41781-021-00059-x`.

[138] *Open Neural Network Exchange*. `https://onnx.ai`. Accessed: 2024-10-17.

[139] A. Hoecker et al. "TMVA - Toolkit for Multivariate Data Analysis". In: (2009). arXiv: `physics/0703039 [physics.data-an]`.

[140] *What is a decision tree?* `https://www.ibm.com/topics/decision-trees`. Accessed: 2024-10-02.

[141] *TMVA::GiniIndex Class Reference*. `https://root.cern.ch/doc/master/classTMVA_1_1GiniIndex.html`. Accessed: 2024-09-30.

[142] Dorival Gon çalves and Junya Nakamura. "Role of the Z polarization in the $H \to b\bar{b}$ measurement". In: *Phys. Rev. D* 98 (9 Nov. 2018), p. 093005. DOI: `10.1103/PhysRevD.98.093005`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.98.093005`.

[143] Dorival Gon çalves and Junya Nakamura. "Boosting the $H \to$ invisibles searches with Z boson polarization". In: *Phys. Rev. D* 99 (5 Mar. 2019), p. 055021. DOI: `10.1103/PhysRevD.99.055021`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.99.055021`.

[144] Kyle Cranmer, Juan Pavez, and Gilles Louppe. *Approximating Likelihood Ratios with Calibrated Discriminative Classifiers*. 2016. arXiv: `1506.02169 [stat.AP]`. URL: `https://arxiv.org/abs/1506.02169`.

[145] ATLAS Collaboration. *Evaluation of theoretical uncertainties for simplified template cross section measurements of V-associated production of the Higgs boson*. ATL-PHYS-PUB-2018-035. 2018. URL: `https://cds.cern.ch/record/2649241`.

[146] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *J. Phys. G* 43 (2016), p. 023001. DOI: `10.1088/0954-3899/43/2/023001`. arXiv: `1510.03865 [hep-ph]`.

[147] ATLAS Collaboration. *Jet Flavour Tagging With GN1 and DL1d. Generator dependence, Run 2 and Run 3 data agreement studies*. `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2023-01/`. Accessed: 2025-01-14. Feb. 2023.

[148] C. N. Yang and R. L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". In: *Phys. Rev.* 96 (1 Oct. 1954), pp. 191–195. DOI: `10.1103/PhysRev.96.191`. URL: `https://link.aps.org/doi/10.1103/PhysRev.96.191`.

[149] ATLAS Collaboration. *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*. ATLAS-CONF-2017-029. 2017. URL: `https://cds.cern.ch/record/2261772`.

[150] ATLAS Collaboration. "Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment". In: *Eur. Phys. J. C* 76 (2016), p. 295. DOI: `10.1140/epjc/s10052-016-4110-0`. arXiv: `1512.05955 [hep-ex]`.

[151] ATLAS Collaboration. *Selection of jets produced in* 13 *TeV proton–proton collisions with the ATLAS detector*. ATLAS-CONF-2015-029. 2015. URL: https://cds.cern.ch/record/2037702.