Measurements of the associated production of a Higgs boson decaying into two beauty quarks with a weak vector boson

Luca Ambroz
University College Oxford
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity Term 2020

Questa tesi è dedicata
ai miei nonni,
alla mia famiglia
e a Lucia.

Abstract

For more than four decades, the Higgs boson has been the most sought-after particle in Particle Physics. Its observation, by the ATLAS and CMS collaborations in 2012, has confirmed the existence of the Higgs field, which is responsible for the spontaneous symmetry breaking mechanism in the electroweak sector and the mass generation of fermions through the Yukawa interaction. Since then, many measurements have been performed, with increasing accuracy, to check whether its properties matched the Standard Model predictions.

The measurement of the Higgs boson decay to beauty quarks is the first test of the Yukawa coupling to bottom-type quarks. Although this decay channel has the largest branching ratio, it has been difficult to observe it at the LHC due to the presence of the large multi-jet background. In this regard, the associated production of a Higgs boson with a leptonically decaying vector boson offers an excellent leverage to reject this background. Furthermore, this production mode is particularly sensitive to Beyond Standard Model physics in the form of Effective Field Theories operators.

This thesis presents inclusive and differential measurements of the associated production of a Higgs boson decaying into two beauty quarks with a weak vector boson using the full Run 2 ATLAS dataset. The inclusive measured signal strength is $\mu_{VH}^{b\bar{b}} = 1.02^{+0.18}_{-0.17} = 1.02^{+0.12}_{-0.11}(\text{stat.})^{+0.14}_{-0.13}(\text{syst.})$, which corresponds to a rejection of the background-only hypothesis with a significance of 6.7 standard deviations, when 6.7 standard deviations are also expected. The differential measurements are reported in five bins of Simplified Template Cross Sections, defined with respect to the transverse momentum of the vector boson. The total uncertainties on each bin range from 85% to 30%. These analyses represent some of the first steps towards the era of precision measurements in Higgs physics, which will characterize the future physics programme of the LHC and the high-luminosity LHC.

Contents

List of Figures

List of Tables

Preface

The work presented in this thesis has been produced in collaboration with other members of the ATLAS Collaboration. In a large collaboration, almost no work can be described as completely individual: this thesis is no exception. Through my DPhil, I have worked on different projects and below I will highlight my main contributions.

- *Flavour-tagging*. I completed my ATLAS Qualification Task in the Flavour-tagging group for the in-data calibration of the flavour-tagging efficiencies for high-transverse-momentum jets using multi-jet events. I have developed and maintained the analysis software framework, implementing systematics prescriptions, producing fit inputs and conducting general studies. I have presented the results of this effort at the *ATLAS Joint Flavour Tagging and* $H \rightarrow b\bar{b}$ *Workshop 2017* at Stony Brook University, New York, USA.

- *Development of novel silicon detectors*. I have characterized monolithic pixel detectors fabricated by LFoundry. I have contributed to Edge-TCT (Transient Current Technique) measurements on the CCPD_LF sensor and Front-TCT measurements on the LF-CPIX sensor. I was also involved with the irradiation campaign that took place at the Birmingham MC40 Cyclotron. The results have also contributed to *Study of prototypes of LFoundry active CMOS pixels sensors for the ATLAS detector* [1].

- *Detector operations*. I monitored the ATLAS Inner Detector operations for about ~ 100 hours at the Inner Detector Desk during data-taking periods.

- *Higgs physics*. I have been an active member of the *Hbb* group since October 2017. During this period, I have contributed to three publications [2, 3, 4]. The data delivered by the LHC between 2015 and 2017 was used for the first two publications while the third one also included the data collected in 2018. For the first publication, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector* [2], I have developed a dedicated analysis framework, VHBBTRUTHFRAMEWORK, for the 0-Lepton channel and conducted the truth-level studies. I have also contributed implemented and validated $E_{\mathrm{T}}^{\mathrm{miss}}$ $t\bar{t}$ samples that helped to reach the observation milestone. Additionally, I was responsible for the 0-Lepton multivariate VH and VZ fits. For the second publication, *Measurement of $VH, H \to b\bar{b}$ production as a function of the vector-boson transverse momentum in 13 TeV pp collisions with the ATLAS detector* [3], I have evaluated the effects of the residual Simplified Template Cross Sections acceptance for electroweak corrections, scale variations and alternative parton distribution functions. Moreover, I performed fit studies on different STXS configurations. I have presented the results of the two previous publications at the *IOP Joint Astroparticle Physics and High Energy Particle Physics Annual Conference* at Imperial College London and the *XXIX International Symposium on Lepton Photon Interactions at High Energies* hosted by the University of Toronto, Canada [5]. The third publication, *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* [4], will be the main focus of this thesis. For this result, I was the main 0-Lepton analyser and I co-authored the internal documentation. Furthermore, I presented the analysis strategy at its first Higgs approval meeting, one of the most important steps necessary to bring a result to publication in ATLAS. I have contributed to multiple studies of the signal and backgrounds composition. In addition, I proposed the introduction of the track-based soft $E_{\mathrm{T}}^{\mathrm{miss}}$ term in the 0-Lepton channel. Lastly, I initiated novel studies of Final-State Radiation recovery and Multivariate classification of the main physics processes.

- *Supervision*. I have co-supervised a four-months project of an MPhys student. The project aim was to investigate new selections for the multi-jet background in the $VH, H \to b\bar{b}$ analysis using dedicated simulated samples and then validating the findings with a data-driven method. His evaluation of the current multi-jet selection with the full Run-2 dataset is reported in Appendix C.

During my studies, I was also awarded three academic prizes. I received one of three Graduate Symposium Prizes for the "most effective presentation" in 2017 from the Oxford Sub-Department of Particle Physics. From the same institution, I have also received the Donald H. Perkins Prize for "outstanding performance in the first year of postgraduate study for the year 2016/2017". Lastly, I received the 2018 Gilberto Bernardini Young Scientist Award from the Italian Physical Society.

The thesis is structured as follows. Chapter 1 provides an introduction to some fundamental concepts in Particle Physics. Chapter 2 describes the LHC accelerator complex and the ATLAS detector. Chapter 3 presents the physics objects used in the $VH, H \to b\bar{b}$ analysis. Chapter 4 and Chapter 5 contain my original work. The former chapter describes the $VH, H \to b\bar{b}$ analysis with the full Run 2 dataset with particular focus on the 0-Lepton channel. The latter contains studies detailing possible analysis improvements for the next iterations of the analysis. I have produced all the plots in these last two chapters, except the ones marked with the ATLAS label, which have been produced in collaboration with the *Hbb* group and have been published in [4].

Chapter 1

Theoretical framework

The Standard Model of Particle Physics is the current best theoretical framework that describes the interactions of the elementary constituents of the universe. A fundamental ingredient of this theory, the Higgs boson, has been discovered only recently. The nature of this particle has to be yet fully understood through experimental measurements. The Higgs boson is a quantum excitation of the Higgs field, which is responsible for a mechanism that generates the mass of elementary particles. In this chapter, an introduction to the Standard Model, with a particular focus on Higgs physics, will be presented.

1.1 Particles of the Standard Model

The *Standard Model* is a comprehensive theory, which describes the electromagnetic, weak and strong interactions between elementary particles through the formalism of relativistic quantum fields and local gauge symmetries of the groups $SU(3)_C \times SU(2)_L \times U(1)_Y$ [6, 7]. The gravitational force, for which a quantum-mechanical description is not available, is not included in the Standard Model. In this framework, *fermions* are the fundamental blocks of matter and *bosons* are the force-carrying particles. A summary of the main properties of the particles in the Standard Model can be found in Figure 1.1.

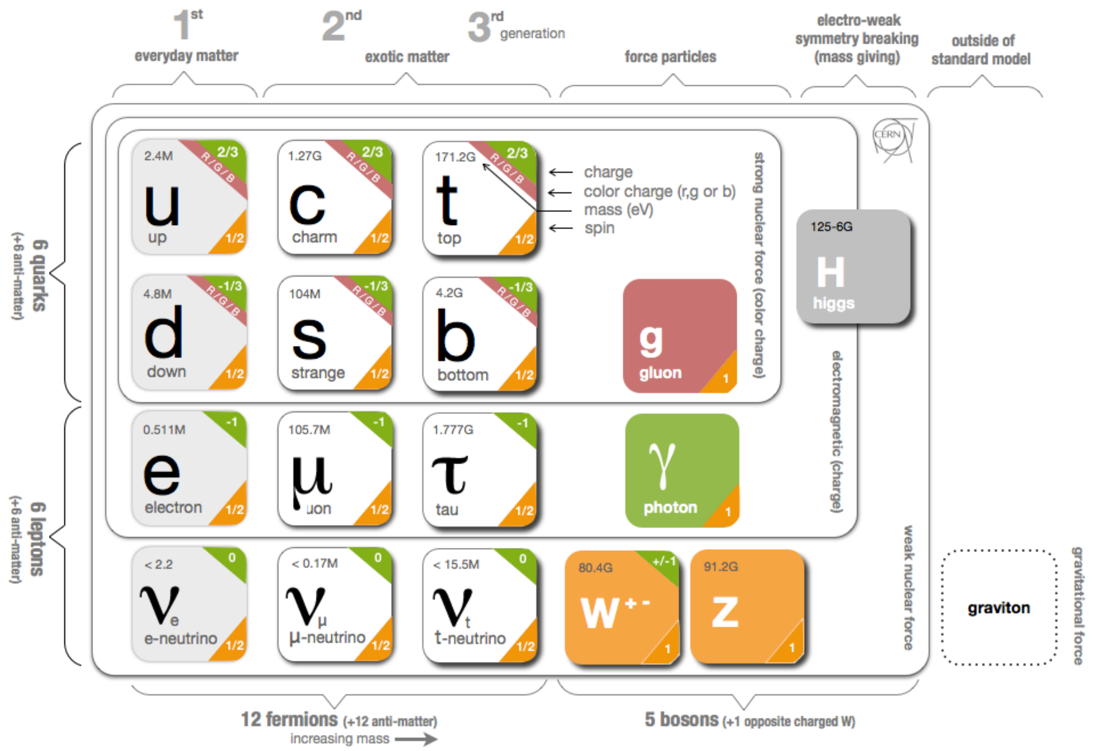


Figure 1.1: Summary of the Standard Model particles plus the graviton, an elementary particle predicted by many quantum gravity field theories and not yet observed [8].

The Standard Model has been extensively and successfully tested at many experiments and colliders [9, 10]. Nonetheless, many questions remain to be answered, such as the inclusion of a theory of gravity, the nature of neutrinos' masses, the hierarchy problem and the asymmetry between matter and anti-matter [6].

Fermions are half-integer spin particles that can be divided into *leptons* and *quarks*, based on whether they can interact via the strong force or not. Furthermore, each fermion is grouped into *generations* according to their flavour quantum number and mass.

Leptons can be further divided into electrically charged leptons (e^-, μ^-, τ^-) and the flavour-associated neutrinos (ν_e, ν_μ, ν_τ), which do not carry the electric charge and can only interact weakly. Conversely, quarks are characterised by a fractional electric charge and by the strong colour charge. Every fermion in the Standard Model is associated with an anti-particle characterised by the same mass and opposite quantum numbers.

Gluons (g) are the massless mediators of the strong force [11]. They confine quarks in hadrons, which are composite states of quarks [12]. A *meson* is a composite state of a quark and anti-quark, while a *baryon* is a composite state of three quarks or anti-quarks; these states are predicted according to the theory proposed by Gell-Mann and Zweig [13, 14]. Gluons and quarks carry *colour charge*, the strong-interaction quantum number, which comes in three colours: red, green and blue. The $W^{\pm}$ and Z^0 (often simply referred to W and Z) are the mediators of the weak force [15], and the photon (γ) is the mediator of the electromagnetic interaction. The Higgs mechanism (Section 1.3) generates masses for the former particles but leaves the latter massless.

1.2 The fundamental interactions

Interactions in the Standard Model arise as invariances of the Lagrangian under a *local gauge transformation* of a symmetry group [16]. For a given Dirac spinor ψ, the free Lagrangian is defined as

$$\mathcal{L}_{free} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi \tag{1.1}$$

where $\bar{\psi} \equiv \psi^{\dagger}\gamma^0$. Given a *global gauge transformation* $\psi \to \psi' = e^{iq\alpha}\psi$ characterised by a charge q, the Lagrangian transforms as

$$\bar{\psi} \to \bar{\psi}' = e^{-iq\alpha}\psi \tag{1.2}$$

$$\mathcal{L}_{free} \to \mathcal{L}'_{free} = \mathcal{L}_{free} \tag{1.3}$$

This implies through Noether's theorem that the charge Q is conserved

$$j^{\mu} = q\bar{\psi}\gamma\psi \qquad \partial_{\mu}j^{\mu} = 0 \tag{1.4}$$

$$Q = \int j^0 d^3x \qquad \frac{dQ}{dt} = 0 \tag{1.5}$$

For a local interaction $\alpha(x)$ in the Abelian group $U(1)$, Equation 1.3 does not hold since

$$\partial_{\mu}\psi \to \partial_{\mu}\psi' = \partial_{\mu}(e^{iq\alpha(x)}\psi(x)) = e^{iq\alpha(x)}[\partial_{\mu}\psi(x) + iq(\partial_{\mu}\alpha(x))\psi(x)] \neq e^{iq\alpha(x)}\partial_{\mu}\psi \tag{1.6}$$

The problem originates from the fact that $\partial_\mu \psi$ does not transform in the same way as ψ under the gauge transformation. To address this issue the *covariant derivative* is defined as

$$D_\mu = \partial_\mu - iqA_\mu \tag{1.7}$$

where A_μ is the electromagnetic four-potential. The covariant derivative under gauge transformation becomes

$$D_\mu \psi \to (D_\mu \psi)' = e^{iq\alpha(x)} D_\mu \psi = e^{iq\alpha(x)} (\partial_\mu \psi - iqA_\mu \psi) \tag{1.8}$$

$$(D_\mu \psi)' = \partial_\mu \psi' - iqA'_\mu \psi' = e^{iq\alpha(x)} [\partial_\mu \psi(x) + iq(\partial_\mu \alpha(x))\psi(x) - iqA'_\mu \psi(x)] \tag{1.9}$$

By comparing Equation 1.8 and Equation 1.9, it is possible to see that the transformation of the electromagnetic four-potential is

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \alpha(x) \tag{1.10}$$

In conclusion, the local gauge transformation of the free Lagrangian is

$$\mathcal{L}_{free} \xrightarrow{\partial_\mu \to D_\mu} \bar{\psi}(i\gamma^\mu D_\mu - m)\psi = \mathcal{L}_{free} + q\bar{\psi}\gamma^\mu A_\mu \psi = \mathcal{L}_{free} + j^\mu A_\mu = \mathcal{L}_{free} + \mathcal{L}_{int} \tag{1.11}$$

It is important to notice that Equation 1.11 does not contain derivatives of A_μ and thus does not describes its temporal evolution. Therefore, the kinetic term $\mathcal{L}_k = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ of A_μ, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, can be added to Equation 1.11 to obtain the complete *Quantum Electrodynamics* (QED) Lagrangian.

In analogy with the reasoning detailed for QED, the symmetry group $SU(2)$ has been introduced to describe the weak interaction and the group $SU(3)_C$ for the strong interaction, as part of *Quantum Chromodynamics* (QCD). The electromagnetic and weak forces can be unified under a single electroweak group $SU(2)_L \times U(1)_Y$. The local gauge $SU(2)_L$ transformation with generators $\boldsymbol{\sigma}$ (Pauli matrices) is defined as

$$\psi \to \psi' = e^{i\boldsymbol{\alpha}(x)\cdot\frac{\boldsymbol{\sigma}}{2}} \psi \tag{1.12}$$

To preserve the symmetry of the Lagrangian, it is necessary to add three additional gauge fields: W_1^μ, W_2^μ, W_3^μ. Therefore, the covariant derivative for the electroweak group is

$$D_\mu = (\partial_\mu + i\frac{g}{2}\sigma W_\mu + iY\frac{g'}{2}B_\mu) \tag{1.13}$$

where I is the weak hypercharge, g is the $SU(2)$ gauge coupling constant, g' is the $U(1)$ gauge coupling constant and B_μ is the massless mediator of the group $U(1)$.

1.3 The Higgs mechanism

With the description of $SU(2)_L \times U(1)_Y$ presented in the previous section, the electroweak bosons arising from the theory are massless. To introduce the mass terms for the boson $W^\pm$ and Z^0 while keeping the photon massless, Brout, Englert, Higgs, Kibble, Hagen, and Guralnik [17, 18, 19] introduced in the Lagrangian a new complex scalar doublet:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{1.14}$$

The Higgs Lagrangian is defined as

$$\mathcal{L}_{Higgs} = |D_\mu\phi|^2 - V(\phi) = |D_\mu\phi|^2 - \mu^2|\phi|^2 - \lambda|\phi|^4 \tag{1.15}$$

and by computing the derivative of the potential, it is possible to obtain its maxima and minima

$$\frac{dV(\phi)}{d\phi}\bigg|_{\phi=\phi_0} = 2\mu^2\phi_0 + 4\lambda\phi_0^3 = 0 \tag{1.16}$$

The solutions depend on the sign of μ and λ; if they are both real and positive there is the trivial solution $\phi_0 = 0$. Conversely, if $\mu^2 < 0$ and $\lambda > 0$ then

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \sqrt{\frac{-\mu^2}{\lambda}} \end{pmatrix} \equiv \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.17}$$

where we have arbitrarily set $\phi^+ = 0$. However, since the potential $V(\phi)$ is invariant under rotations in the (ϕ^+, ϕ^0) plane, other choices of values satisfying Equation 1.16 would have been allowed. In conclusion, the favoured shape of the Higgs potential can be seen in Figure 1.2.

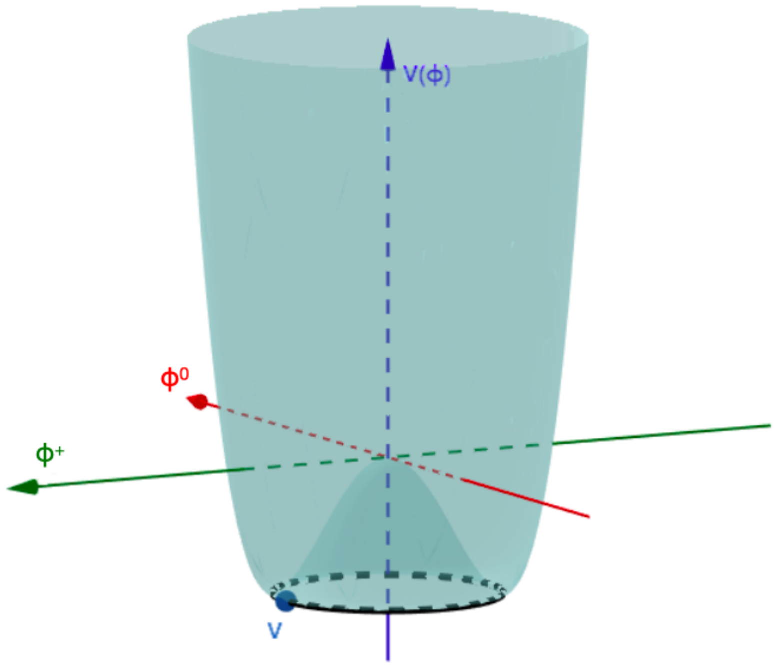


Figure 1.2: A visual representation of the Standard Model Higgs boson potential.

The Higgs Lagrangian in Equation 1.15 can be expanded around its minimum ϕ_0 by a small quantity H along the ϕ^0 axis, where the covariant derivative is defined in Equation 1.13. The final result can be written as

$$\begin{aligned}\mathcal{L}_{Higgs} &= \frac{1}{2}(\partial H)^2 + \frac{g^2}{4}(v+H)^2(W_\mu^+ W^{\mu-}) \\ &+ \frac{1}{8}(g^2+g'^2)Z_\mu Z^\mu (v+H) \\ &+ \frac{\mu^2}{2}(v+H)^2 - \frac{\lambda}{4}(v+H)^4\end{aligned} \tag{1.18}$$

From Equation 1.18, it follows that the $W^\pm$ and Z^0 are massive and a new boson of mass $m_H = \sqrt{-2\mu^2}$ has appeared. The procedure described above, known as *electroweak symmetry breaking*, generates only the masses of the bosons in the electroweak sectors. In order to generate the mass terms of the fermions, it is necessary to include the *Yukawa coupling terms* in the Lagrangian

$$\mathcal{L}_{Yukawa} = -G_f(\bar{\psi}_L \phi \psi_R + \bar{\psi}_R \phi \psi_L) \tag{1.19}$$

where $\psi_{L(R)}$ is the left (right) handed fermion isospin doublet (singlet) and G_f is the coupling constant for each fermion, which can be experimentally measured.

1.4 Higgs phenomenology

In 2012, the ATLAS and CMS experiments claimed the observation of a new particle compatible with the production and decay of the Standard Model Higgs boson [20, 21]. The decays channels primarily used for this discovery were $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^*$ and, $H \rightarrow WW^*$ in ATLAS; a fit to the signals extracted from the first two channels allowed a first mass measurement, m_H, near 125 GeV. In the next sections, the Higgs boson production and decay modes will be described.

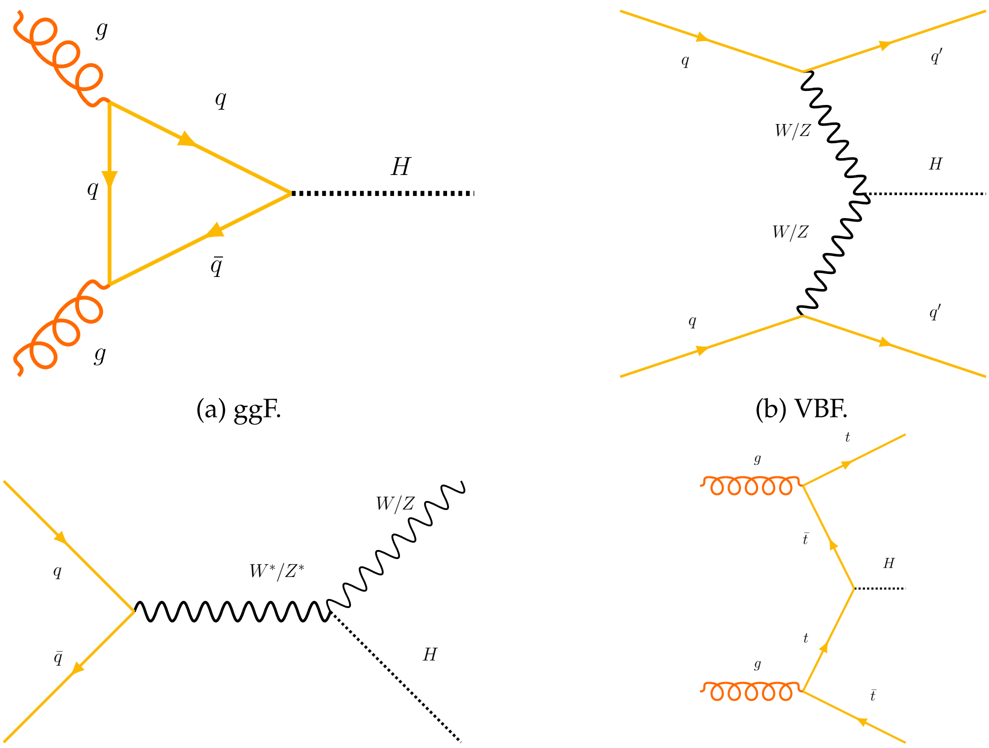
1.4.1 Higgs production mechanisms at the LHC

At the LHC collider (Section 2.1), the main Higgs production mechanisms are the gluon-gluon fusion (ggF), vector boson fusion (VBF), associated production with a vector boson (VH) and associated production with a pair of heavy quarks ($t\bar{t}$H and $b\bar{b}$H). The Feynman diagrams for these processes are shown in Figure 1.3. The Higgs production cross-sections as a function of the centre-of-mass energy, $\sqrt{s}$, for proton-proton collisions can be seen in Figure 1.4(a); these processes are initiated by either the valence quarks in the nucleus or by the sea quarks (and anti-quarks) or by the gluons [22].

At the centre-of-mass energy of 13 TeV, the dominant production mode is the gluon-gluon fusion (σ_{ggF} = 48.6 pb [22]) due to the contribution from the gluon parton distributions [23]. The top quark, given its large mass, contributes the most in the ggF loop.

The VBF production (σ_{VBF} = 3.78 pb [22]) is characterised by the merging of two off-shell vector bosons that have been radiated from the incoming quarks in the protons. In a VBF event, forward jets from the incoming protons allow such events to be distinguished from QCD events [24].

In the VH production mechanism, the Higgs is produced in association with a vector-boson W or Z. The vector boson leptonic decays offer excellent signatures for trigger-



(a) ggF. (b) VBF. (c) VH. (d) ttH.

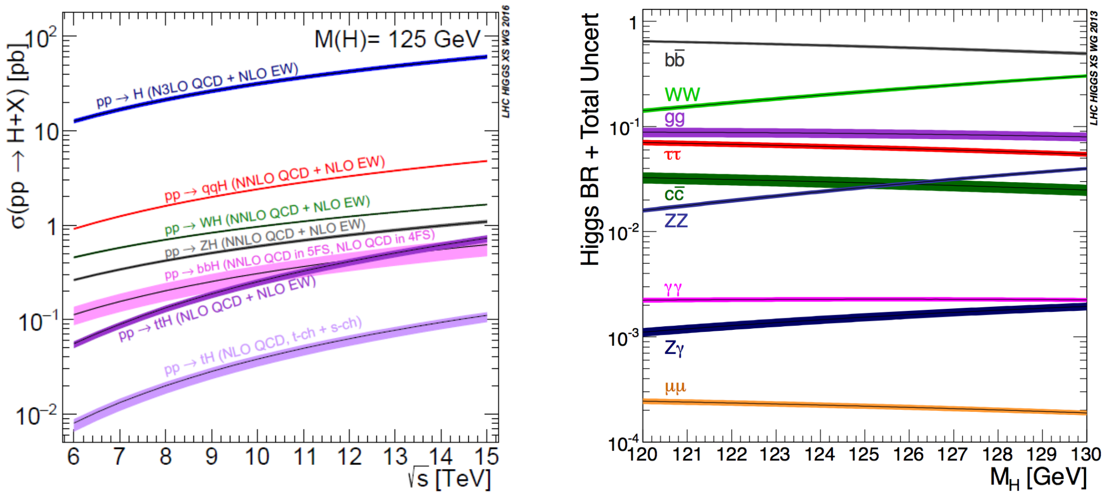
Figure 1.3: Dominant production modes for the Higgs boson in proton-proton collisions. The quark loop in (a) is dominated by the top quark.

ing and rejecting the QCD background. This production mode is the most sensitive for searching the $H \to b\bar{b}$ decays, as it will be discussed in Chapter 4.

Lastly, the $t\bar{t}$H production mode (σ_{ttH} = 0.51 pb [22]) allows the direct coupling of the top quark to the Higgs boson from a tree-level diagram to be probed. Similar to the $t\bar{t}$H production, the $b\bar{b}$H production is characterised by a slightly smaller cross-section at 13 TeV (σ_{bbH} = 0.49 pb [22]).

1.4.2 Higgs decay modes

The lifetime of the SM Higgs boson with a mass of 125 GeV is $\mathcal{O}(10^{-22})$s for a *Total Width* (Γ_H) predicted from the Standard Model of 4.03 MeV [25]. Therefore, the Higgs boson can only be detected through its decay products. At tree level, the Higgs boson can decay into pairs of quarks, leptons and electroweak gauge bosons. Furthermore, loop diagrams allow the Higgs boson to decay to massless particles such as photons and gluons. The main contributors of these loops are $t(\bar{t})$ quarks and $W^{\pm}$ bosons. The



(a) Cross-sections production. (b) Branching ratios.

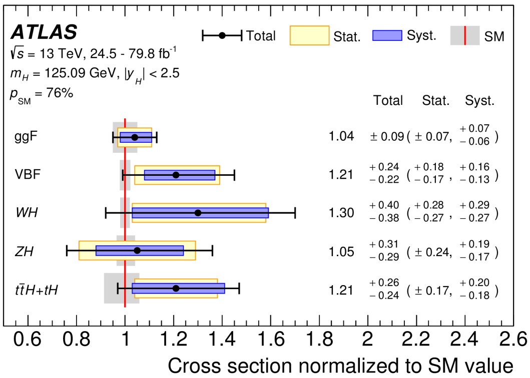
Figure 1.4: On the left plot, the Higgs boson production cross-sections at LHC as a function of the centre-of-mass energy [22]. The gluon-gluon-fusion production mode is referred as $pp \to H$ and the vector-boson-fusion production mode is referred as $pp \to qqH$. On the right plot, the Higgs boson branching ratios as a function of its mass [22].

Branching Ratio (BR) is defined as the fraction of an individual decay mode with respect to the total Higgs boson width:

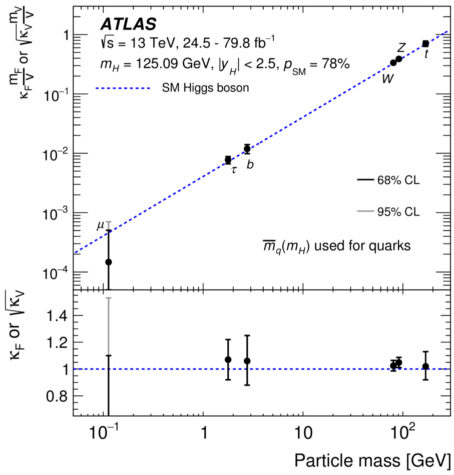
$$BR(H \to X_i) = \frac{\Gamma(H \to X_i)}{\sum_j \Gamma(H \to X_j)} \tag{1.20}$$

The BR predictions for the Higgs boson change according to its mass, as can be seen in Figure 1.4(b). The largest BR, for $m_H = 125$ GeV, is the decay into a $b\bar{b}$ pair, which will be the main focus of this thesis; this decay is enhanced by the relatively large mass of the b-quarks and the coupling constant for fermions. However, a Higgs boson of $m_H = 125$ GeV cannot decay in a $t\bar{t}$ pair because it is not allowed by kinematics. The decays into W^+W^- and ZZ have the second and fourth-largest BRs; this type of decay is suppressed by the presence of an off-shell vector boson. The decay of the Higgs into a gg is experimentally challenging at the LHC and has not been directly probed yet due to the large multi-jet background. The decay to $\tau^-\tau^+$ was the first measurement of the Higgs boson coupling to fermions. The decay to the second generation of fermions ($c\bar{c}$ and $\mu^+\mu^-$) and $Z\gamma$ are very challenging due to the small signal-to-background ratios. Lastly, the decay into two photons is one of the most sensitive decay modes because of the excellent resolution of the electromagnetic calorimeter [26] and the possibility of reconstructing the complete Higgs-boson kinematics.

All the current measurements of production cross-sections and the Higgs coupling values are compatible with the Standard Model as can be seen in Figure 1.5 [27].



(a) Measured production cross-sections.



(b) Measured coupling-strength modifiers.

Figure 1.5: (a) Measured cross-sections for Higgs production mechanisms at the LHC with the assumption of Standard Model BRs. (b) Standard Model reduced coupling-strength modifiers (K_F) for particles coupling to the Higgs boson [27].

1.5 Standard Model Effective Field Theories

Effective Field Theories (EFTs) are general models that describe the physics at an energy scale E by approximating the underlying complete physical theory, which is characterised by energy scale $\Lambda \gg E$ [28].

The Fermi theory of beta decay is an example of effective field theory. In this theory, the neutron decay $n \to p + e^- + \bar{\nu}_e$ is interpreted as a point-like interaction among the four fermions characterised by an $E \sim \mathcal{O}(MeV)$, which is significantly smaller than the mass of the W boson $m_W \sim 80$ GeV that defines the energy scale of the electroweak interaction.

At particle colliders, EFTs could shed light on physics processes whose energy scale Λ is much higher than energies currently available, $E \sim \mathcal{O}(TeV)$. Furthermore, EFTs are built based on few very general assumptions that can probe a great variety of underlying physical theories.

EFTs can be postulated as an expansion of the Standard Model (SMEFT), in particular of the Higgs sector given the central role that the Higgs boson plays in many Beyond Standard Model (BSM) theories. The Standard Model Lagrangian has a mass dimension $d = 4$. Therefore, new terms in the Lagrangian will have to keep $d = 4$ and the SMEFT Lagrangian can be written as

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i^{(D)}}{\Lambda^{D-4}} \mathcal{O}_i^{(D)} \tag{1.21}$$

where $O_i^{(D)}$ are dimension-D operators, and $c_i^{(D)}$ are the Wilson coefficients, which parametrise the strength of the new interactions. The operators $O_i^{(D)}$ are defined according to the symmetries that they respect, to their power D and to the fields to which they couple. In most of the SMEFTs only operators with $D = 6$ are considered given that operators with $D = 5$ or $D = 7$ violate either the conservation of lepton or baryon number, while operators with $D > 7$ are further suppressed by the energy scale Λ. Given the previous requirements, a complete basis for $O_i^{(D=6)}$ operators is made of 59 independent operators [29].

From an experimental point of view, one of the effects of the $O_i^{(D)}$ operators is the change in the Higgs production cross-sections at high transverse momentum. Therefore, the advent of differential measurements in p_{T} in the Higgs sector can put stringent constraints on SMEFT Wilson coefficients.

Chapter 2

The LHC and the ATLAS detector

The ATLAS experiment is one of the two multi-purpose detectors operating at the LHC. These detectors, combined with the accelerator complex, provide unique data capable of expanding our knowledge of Particle Physics. In this chapter, Section 2.1 gives a brief description of the LHC. Then, in Section 2.2, the ATLAS detector and its main sub-detectors are outlined.

2.1 The Large Hadron Collider

The Large Hadron Collider [30] is a proton and heavy-ion collider. It is located at the European Organization for Nuclear Research (CERN), beneath the Franco-Swiss border near Geneva.

Protons at the LHC are accelerated in five stages to reach their designed final energy (see Figure 2.1). The first step of the acceleration chain is the ionization of hydrogen atoms to produce protons. Then, the protons are injected into LINAC2, a linear accelerator, which will increase their energy up to 50 MeV. The next accelerator stage after the LINAC2 is the Booster, the first circular accelerator in the chain, where the protons' energy reaches 1.4 GeV. From the Booster, protons enter the Proton Synchrotron (PS) gaining energy of 25 GeV. The penultimate stage of the accelerator process is the Super

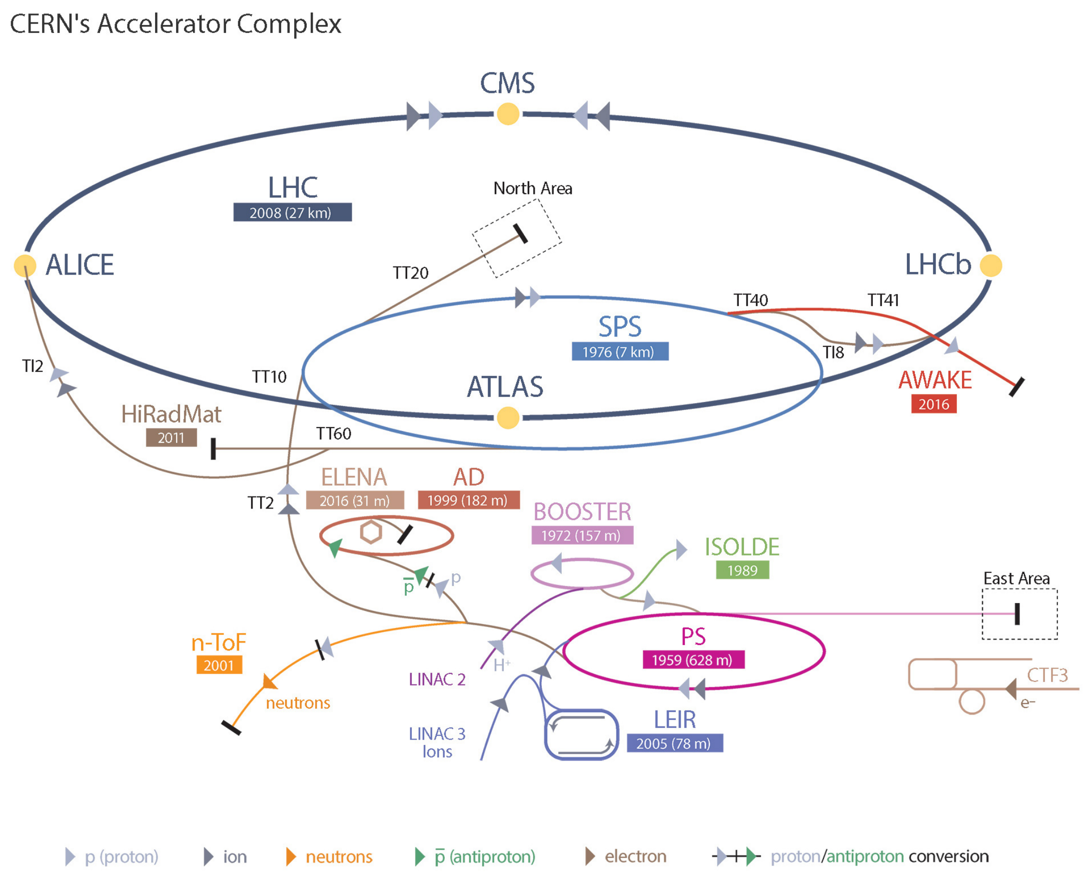


Figure 2.1: CERN's acceleration complex [32].

Proton Synchrotron (SPS), an accelerator ring with a circumference of 7 km that pushes the protons to 450 GeV. Lastly, the particles are injected into the LHC, where they reach their final energy. The LHC has two pipes in which particles travel in opposite directions, allowing for collisions at four Interaction Points (IPs) where the ALICE, ATLAS, CMS and LHCb detectors are located.

The LHC is designed to operate at a proton-proton centre-of-mass energy of $\sqrt{s} = 14$ TeV. However, this designed centre-of-mass energy has not yet been reached: in 2010 and 2011 the energy was $\sqrt{s} = 7$ TeV, in 2012 $\sqrt{s} = 8$ TeV and between 2015 and 2018 $\sqrt{s} = 13$ TeV [31].

At the LHC, protons circulate in 2808 bunches per beam each containing a maximum of $\sim 1.2 \times 10^{11}$ particles. Two consecutive bunches are separated by 25 ns. Radiofrequency cavities (RF), operating at a temperature of 4.5 K, in a superconducting state to remove electrical resistance, and a potential of 2 MV, accelerate particles through oscillating electromagnetic fields. Superconducting dipole (NbTi) magnets, operating at a temperature of 1.9 K and a maximum current of 11700 A, bend the trajectories of the particles to keep them in the ring. Quadrupole and higher-order magnets are displaced throughout the collider to stabilise and focus the beams. A vacuum system removes the majority of air molecules in the beam pipes that would interfere with the protons, keeping the pressure of $\sim 10^{-11}$ mbar.

2.1.1 Luminosity

At a particle collider, the number of expected events N for a given process of cross-section σ over a fixed time is [33]

$$N = \sigma \cdot L \tag{2.1}$$

where L is the *integrated luminosity*, a parameter that is entirely defined by the accelerator characteristics. Since the accelerator performance varies over time, the integrated luminosity is computed as the integral over time of the *instantaneous luminosity* $\mathcal{L}$:

$$L = \int \mathcal{L}(t')dt' \tag{2.2}$$

$$\mathcal{L} = \frac{f_{rev} n_b N_b^2 \gamma_r}{4\pi \epsilon_n \beta^*} F \tag{2.3}$$

in which f_{rev} is the revolution frequency, n_b is the number of bunches per beam, N_b is the number of particles in each colliding beam, γ_r is a relativistic factor, ϵ_n is the normalized transverse beam emittance, β^* is the beta function at the collision point and F the geometric luminosity reduction factor that takes into account the angle between the colliding beams [33]. ϵ_n and β^* measure the transverse-spatial and momentum-space spreads of the beams, respectively. The integrated luminosity is commonly measured in powers of inverse barns $[L] = \mathrm{b}^{-1}$, where 1 barn = 10^{-28} m^2.

At the LHC, the luminosity decays exponentially after the start of collisions due to the degradation of intensities and emittances of the circulating beams. For this reason, once the protons are accelerated to their final energy they only circulate and collide for about half a day. The peak luminosity recorded by ATLAS was 2×10^{34} cm^{-2}s^{-1}, which is a factor of two higher than the LHC design luminosity. The total integrated luminosity delivered to the ATLAS detector between 2011 and 2018 is displayed in Figure 2.2(a). The luminosity is measured with multiple sub-detectors and algorithms in ATLAS; in particular, the measurement used in physics analyses was performed by the LUCID2 Cherenkov detector [34]. Data is required to pass good quality requirements in order to guarantee the correct operation of the detector during data-taking [35].

The data-taking period between 2011 and 2012 is called Run 1, whereas the period between 2015 and 2018 is called Run 2. The total Run 2 integrated luminosity is $L = 139.0 \pm 2.4\ \text{fb}^{-1}$ [36]. The LHC programme is also scheduled to run a third period of data-taking, Run 3, between 2021 and 2024 with similar beam conditions to Run 2 [37].

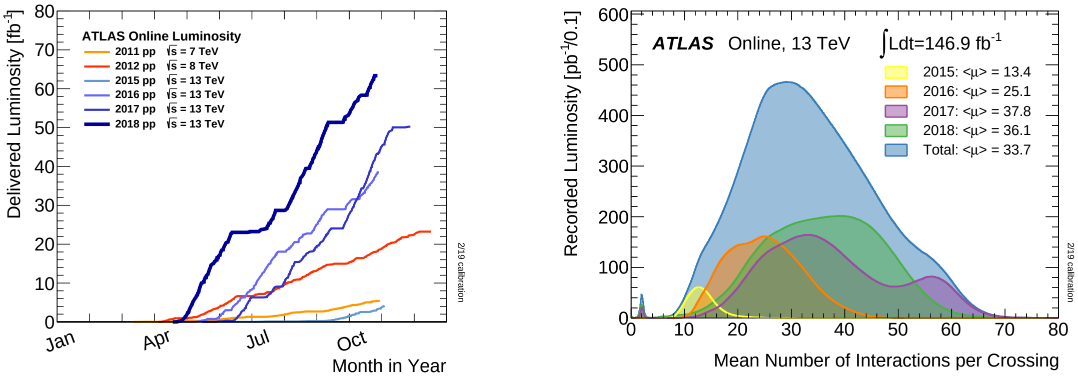
2.1.2 Pile-up

The majority of the events of interest for the detectors at the LHC come from the high-energy hard-scattering collisions. However, with the increase of instantaneous luminosity, other lower-transverse-momentum collisions might happen at the same time. Thus, *in-time pile-up* (often simply referred as pile-up) is caused by multiple soft collisions within the same bunch crossing. There is also a second type of pile-up, *out-of-time pile-up*, which originates from collisions from interactions in neighbouring bunch crossings. Those pile-up collisions overlay with the hard-scattering one, effectively degrading its reconstruction quality. The number of interactions per bunch crossing μ (see Figure 2.2(b)) can be calculated using the inelastic proton-proton cross-section and some accelerator parameters defined in Section 2.1.1:

$$\mu = \frac{L\sigma_{inel}}{n_b f_{rev}} \tag{2.4}$$

2.2 The ATLAS detector

The ATLAS experiment [38] is a general-purpose detector designed to test the current understanding of Particle Physics and probe new theories beyond the scope of the



(a) Delivered luminosity. (b) Number of interactions per crossing.

Figure 2.2: (a) Delivered luminosity to the ATLAS detector [31]. (b) Luminosity-weighted distribution of the mean number of interactions per crossing for the Run 2 (2015-2018) proton-proton collision data at $\sqrt{s} = 13$ TeV [31].

Standard Model. The detector has a cylindrical shape, centred around the beam axis, that provides a near-complete 4π coverage. In Figure 2.3, a cutaway scheme of the ATLAS detector and its main sub-detectors is shown. The detector measures 45 m in length and 25 m in diameter, weighing more than 7000 tons.

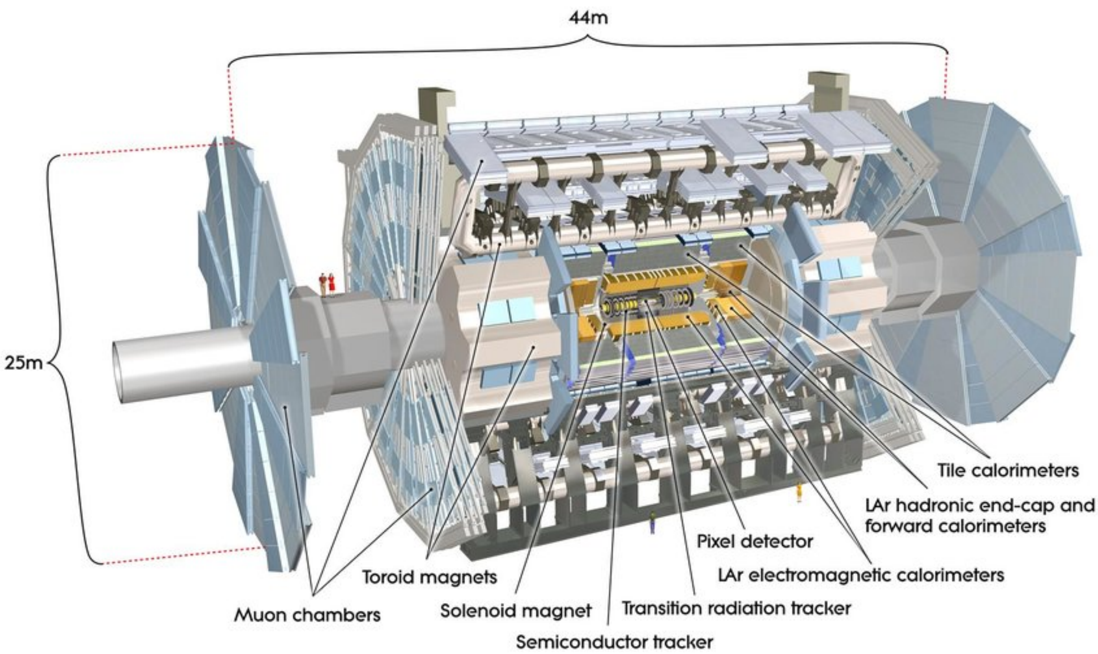


Figure 2.3: Cutaway scheme of ATLAS detector [38].

In ATLAS, a right-handed coordinate system with reference to the LHC ring is used. The origin of the coordinate system is the nominal interaction point, at the centre of the detector. The beam-line defines the z-axis, where the positive direction is anti-clockwise. The positive x-axis points towards the centre of the LHC ring, and the positive y-axis points upwards. A cylindrical coordinate system is defined in ATLAS based on the azimuthal angle ϕ around the z-axis in the xy-plane, and the polar angle θ is defined as the angular separation from the zy-plane. For a particle of four-momentum vector $(E, \boldsymbol{p})$, its *transverse momentum* is defined as $p_T = |\boldsymbol{p}| \sin(\theta)$ and *rapidity* is defined as $y = \frac{1}{2} \ln[(E + p_z)/(E - p_z)]$. For an ultra-relativistic particle, $E \gg m$, the *pseudo-rapidity* $\eta = \ln \cot(\theta/2)$ approximates the rapidity. Differences in pseudo-rapidity $\Delta\eta$ are invariant under Lorentz boosts along the z-axis. Moreover, angular distances are measured in terms of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, also referred to as dR.

Lastly, the ATLAS detector is divided into three main regions according to η: the central barrel ($|\eta| < 2.5$) and the two forward end-caps ($2.5 \leq |\eta| < 5$).

2.2.1 Magnet system

The ATLAS magnet system [39] is composed of four sets of superconducting magnets: a *Central Solenoid*, a *Barrel Toroid*, and two *End-Cap Toroids*. This configuration restricts the magnetic field to the Inner Detector and thus minimizes its presence in the calorimeters, which would degrade their performances.

The central solenoid surrounds the Inner Detector (see Section 2.2.2). It generates a magnetic field of 2 T parallel to the z-axis, which deflects charged particles' trajectories along the ϕ coordinate. The strength of the magnetic field is almost constant along the radial direction, whereas it decreases along the z-axis due to the finite length of the solenoid.

The Barrel Toroid and End-Cap Toroids each consist of eight coils arranged in the radial direction with respect to the z-axis. The former produces a magnetic field up to 3.9 T, whereas the latter a magnetic field up to 4.1 T. The toroids' design aims at bending the muons' trajectories in the η direction to allow precise measurement of their momenta.

2.2.2 Inner Detector

The ATLAS Inner Detector (ID) [40, 41] is designed to reconstruct and measure the tracks of charged particles. The solenoid bends the trajectories of the charged particles, thus allowing their transverse momenta to be measured. The ID is the innermost sub-detector system and consists of four sub-detectors: the Insertable B-Layer (IBL), the Pixel detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). A schematic of the ID is shown in Figure 2.4. The ID is characterised by a *penetration depth* that ranges between 0.4 X_0 in the barrel and 1.5 X_0 at about $|\eta|$=2 [40] where X_0 is the *radiation length*. The radiation length is the average distance over which an electron reduces its energy to $E_{final} = E_{initial}/e$ by emitting bremsstrahlung radiation, where e is Napier's constant.

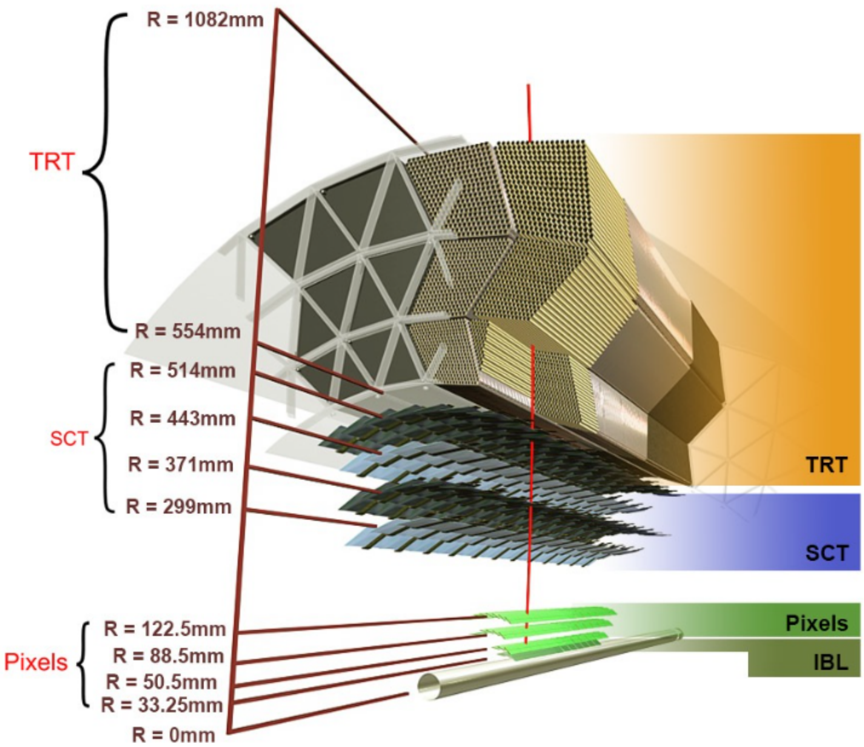


Figure 2.4: Schematic of the ATLAS Inner Detector [42].

The IBL is the innermost layer of the ID. It was installed at the beginning of Run 2, to improve tracking performances, in particular the ones related to b-tagging (see Section 3.2.4.3) [41]. It consists of a layer of silicon pixels placed at $r = 33.25$ mm and $|z| < 332$ mm. The Pixel detector makes up the second part of the ID from $r = 50.5$ mm to

$r = 122.5$ mm [43]. It consists of a central barrel, made of three layers of pixels, and two end-caps, each composed of three disks of pixels. The SCT surrounds the Pixel detector with over 6 million silicon strips organised in four concentric barrel layers and nine disks on each end-cap to cover up to $|\eta| < 2.5$ [44]. Lastly, the TRT is the outermost component of the ID [45]. It consists of drift straws arranged parallel to the z-axis in the barrel, and radially in the end-caps. Each straw tube has a radius of 2 mm, it contains a tungsten anode wire and it is filled with a gas mixture of Xe, and a mixture of quenching gases. Polythylene, a plastic material, is used to fill the space between the straw tubes to induce transition radiation for charged particles, whose intensity is proportional to the Lorenz boost γ ($\propto E/m$). In particular, the transition radiation provides good discrimination between electron and pions. Xenon is an excellent gas to detect X-ray photons produced in the transition radiation due to its high atomic number.

2.2.3 Calorimetry

Located outside the solenoid magnet, the ATLAS calorimeters are designed to accurately measure particle energies. The calorimeters are hermetic along the ϕ coordinate and they cover in pseudo-rapidity up to $|\eta| < 4.9$. They are of two types: an inner *electromagnetic calorimeter* and an outer *hadronic calorimeter*, as shown in Figure 2.6. Both calorimeters are *sampling calorimeters* consisting of multiple layers of *passive* and *active* materials. In the passive layers, dense materials with a high atomic number are used to generate a cascade of particles, called *shower*, from an incoming particle. The active layers measure, through ionisation and scintillation, the energy deposited by the shower. This measure allows the energy of the incoming particle to be assessed. For this reason, calorimeters require large volumes to contain the whole particle shower.

2.2.3.1 Electromagnetic calorimeter

The ATLAS electromagnetic calorimeter (ECAL) is built with lead as the passive material and liquid argon as the active material [26]. High-energy photons ($E_\gamma \gg 1$ MeV)

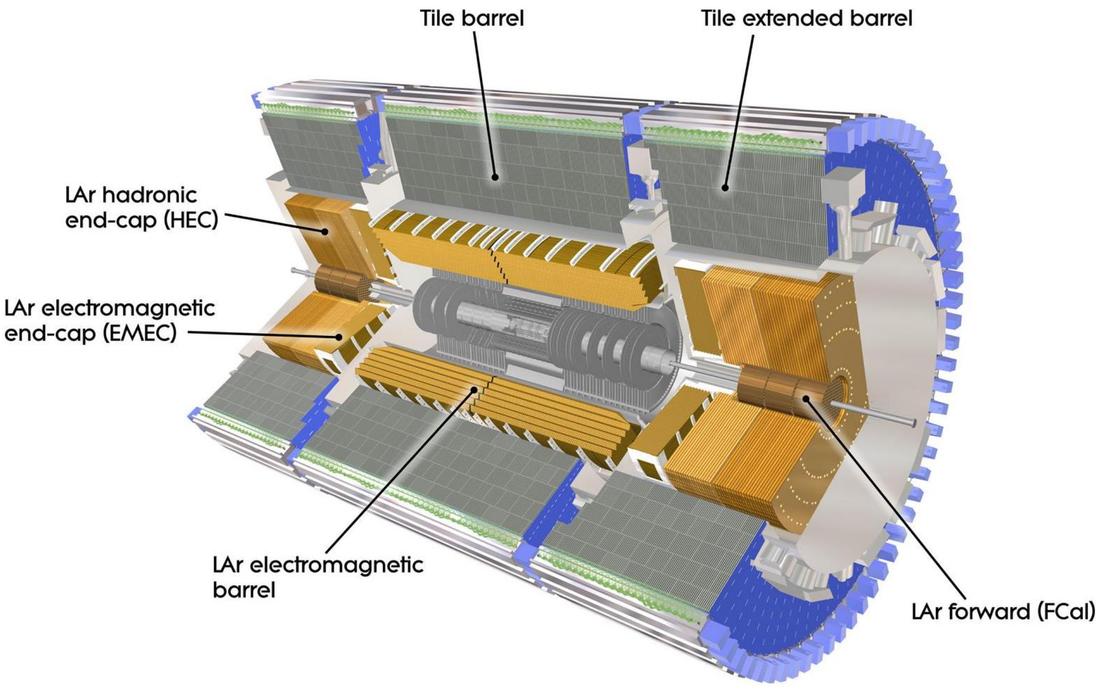


Figure 2.5: A representation of the electromagnetic and hadronic calorimeters in the ATLAS detector [38].

interact with matter mainly through *pair production* and high-energy electrons through *bremsstrahlung* [46]. The mean length to reduce the energy of an electron by $1/e$ is the radiation length X_0, while the mean length for a photon to produce an electron-positron pair is λ_{pair}. In general, since $\lambda_{pair} \simeq \frac{9}{7}X_0$, the electromagnetic shower can be approximated by assuming that every X_0 in depth, either an electron emits a photon through bremsstrahlung or a photon creates a positron-electron pair. Since the radiation length can be approximated as $X_0 \sim 180A/Z^2$ [g/cm^2], ATLAS chose lead ($X_0^{Pb} = 6.37$ g/cm^2 [46]) as a passive material for developing the shower in a compact space. Liquid argon is used to drift the ionization produced by the shower on copper electrodes arranged in a honeycomb structure. Liquid argon is characterised by chemical stability, radiation hardness, and good energy resolution ($\lesssim 10\%\sqrt{E[GeV]}$) for the calorimeter [47]. The ECAL has a central barrel covering $|\eta| < 1.475$, and two end-caps covering $1.375 < |\eta| < 3.2$. A single layer of liquid argon, called the pre-sampler, is placed before the ECAL barrel to correct for the energy loss in the ID and the magnet. The overall ECAL depth ranges from 22 X_0 in the barrel to more than 24 X_0 in the end-caps.

2.2.3.2 Hadronic calorimeter

The ATLAS hadronic calorimeter (HCAL) is designed to reconstruct hadronic activity [48]. Hadrons deposit their energies through collisions mediated by the strong interaction, creating a hadronic shower. The charged hadrons produced in the pp collisions deposit their energies through excitations and ionisation. The strong interaction creates a lot of mesons, among which the most abundant are π^0 and $\pi^{\pm}$. Neutral π^0 mesons decay almost immediately into a pair of photons, which deposit their energies in a limited space through an electromagnetic shower. The charged $\pi^{\pm}$ mesons decay most of the time into a muon and a neutrino: the muon deposits a minimum ionising signal in the calorimeter, whereas the neutrino leaves the calorimeter without interacting. Overall, a hadronic shower can be described in terms of *nuclear interaction length* λ_{int}, which is the average distance that a hadron travels before undergoing a strong reaction. For a given material $\lambda_{int} \gg X_0$, thus hadronic calorimeters are generally larger than electromagnetic calorimeters. The HCAL consists of a central tile barrel calorimeter, two hadronic end-cap calorimeters (HEC) and two forward calorimeters (FCAL). The tile calorimeter is constructed from alternating tiles of steel absorber and plastic scintillator; it is composed of a central barrel that extends to $|\eta| < 1.0$ and two additional barrels between $0.8 < |\eta| < 1.7$ on each side. The scintillators are read out by wavelength-shifting fibre optics that are connected to photo-multiplier tubes. The tile calorimeter is segmented in modules of η, ϕ and r segments. These cells are arranged in three layers: the first two layers 0.1×0.1 $(\Delta\eta \times \Delta\phi)$ while the most external one 0.2×0.1 $(\Delta\eta \times \Delta\phi)$. The HEC covers between $1.5 < |\eta| < 3.2$, while the FCAL between $3.1 < |\eta| < 4.9$. Both sub-systems use liquid argon as an active medium; however, the HEC uses copper as passive material, whereas the FCAL is made of tungsten. The overall HCAL depth ranges from $\sim 7\ \lambda_{int}$ in the central part of the tile barrel calorimeter to $\sim 10\ \lambda_{int}$ in the FCAL.

2.2.4 Muon Spectrometer

The Muon Spectrometer is the outermost sub-system of the ATLAS detector [49]. It is designed to trigger upon and measure the trajectories of the muons. Muons, while travelling through the inner detector and the calorimeters, lose only a small amount of energy; this is due to their large mass compared to that of electrons ($m_\mu \sim 200 \cdot m_e$) that suppresses the probability of bremsstrahlung ($\sigma_{brem} \propto 1/m^2$). A muon, travelling through the detector, typically leaves a single track, which is bent in the magnetic field. The muon spectrometer, covering up to $|\eta| < 2.7$, consists of four types of sub-systems: the Monitored Drift Tubes (MDT), the Cathode Strip Chambers (CSC), the Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC). The MDT and CSC provide precision tracking of the muons. They are gaseous sub-detectors operating with a mixture of Ar and CO_2. Instead, RPC and TGC are used for fast triggering on the muons. They are characterised by high efficiency and low dead time.

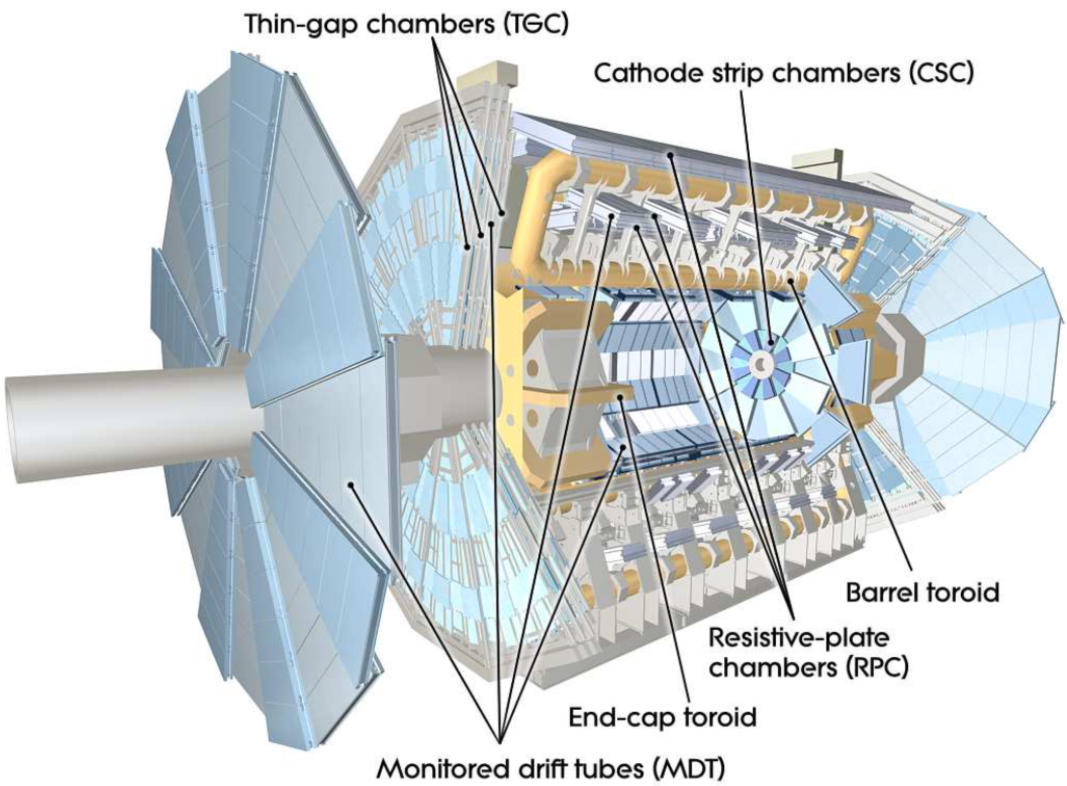


Figure 2.6: An overview of the ATLAS muon spectrometer [38].

2.2.5 Trigger and data acquisition

The ATLAS detector cannot record all the collisions that occur every 25 ns (40 MHz). For this reason, it uses a Trigger and Data Acquisition System (TDAQ) to record only the collisions that are the most interesting for physics analyses. The TDAQ is a multi-level system designed to reject non-interesting physics collisions as soon as it is possible to make a decision. The two main levels of the TDAQ are: the Level 1 Trigger (L1) and the High Level Trigger (HLT); the former reduces the rate of recorded events to $\lesssim$ 100 kHz, while the latter allows the rate to be reduced to $\sim$ 1.2 kHz [50].

The L1 Trigger is a hardware-based trigger that combines inputs from the calorimeter (L1Calo) and the muon triggers (L1Muon). The L1 selects events that contain potential leptons, photons, high-transverse-momentum jets or that are characterised by a large imbalance in the transverse momentum. For the L1Calo, the decision is based upon the reconstruction of Regions of Interest (RoI) defined as clusters of deposited transverse energy. The RoI are then transferred to the HLT to speed up the reconstruction. For the L1Muon to trigger a muon, the RPC or the TGC require two or three hits according to the muon transverse momentum. For Run 2, the L1Topo, a new L1 Trigger, was included in the ATLAS TDAQ. L1Topo combines information from the L1Calo and the L1Muon and performs kinematic selections to identify interesting events [51].

The HLT is a software-based trigger that combines the RoI with information from all the sub-detectors to perform a first reconstruction of the events. At this level, more complex signatures can be determined, such as events that contain b-jets, hadronic taus and many other combinations of physics objects. The ensemble of triggers in ATLAS is called the *Trigger Menu*. Lastly, if an event is selected, it is saved in the Tier-0 CERN data centre for full off-line reconstruction, calibration and analysis.

2.3 Simulations

In ATLAS, simulations are used to describe physical processes as well as the particles' interactions with the detector. Simulating a process in most of the cases require calculating multi-dimensional integrals. The most efficient way to compute these integrals is through a family of random sampling techniques known as *Monte Carlo* methods (MC) [52]. These techniques are so pervasive that often MC simulations are simply referred to as MC.

At LHC, the production cross-section for a proton-proton collision producing a final-state X is

$$\sigma_{pp\to X} = \sum_{a,b} \int_0^1 dx_1 dx_2 f_a(x_a, \mu_F) f_b(x_b, \mu_F) \hat{\sigma}_{a,b\to X}(s, \mu_R, \mu_F) \tag{2.5}$$

where a, b are the partons extracted from proton, x_1 and x_2 are the proton momentum fractions carried by partons, $f_i(x_i, \mu_F)$ are the *parton density functions* (PDFs), s is the square of the centre-of-mass-energy, μ_F is the *factorisation scale*, μ_R is the *renormalisation scale* and $\hat{\sigma}_{a,b\to X}$ is the *partonic cross-section* between the two partons involved in the hard-scattering interaction, which contains the *matrix element* (ME) [53]. PDFs provide the probability to find *partons* (quarks and gluons) in a proton as a function of the fraction of its momentum carried by the parton [54]. PDF values depend on the energy scale of the collisions, which are parametrised by the *factorisation scale* μ_F and *renormalisation scale* μ_R [55]. The factorisation scale can be interpreted as the scale that separates the long- and short-distance physics, while the renormalisation scale is used to treat ultraviolet divergences. Furthermore, *factorisation theorems* allow the non-perturbative QCD effects to be decoupled from the partonic cross-section. Then, when quarks and gluons are created after the hard-scattering collision, they are too energetic to form bound states and they produce a cascade of particles through the radiation of a gluon from a quark and via gluon splitting; this process is modelled with *parton shower* (PS) generators. Once the evolution of the system reaches an energy in which the perturbative QCD regime can no longer be applied, phenomenological *hadronisation* models are used to simulate the formation of hadrons [56]. The *underlying event*, which is the additional soft activity originating from the primary vertex, is added to the hadronistion models. Then, the simulation of pile-up collisions is added on top of the event using soft QCD processes. Lastly, the interactions of the final-state particles with the ATLAS detector are simulated using the GEANT4 software package [57] and then the standard ATLAS object-reconstruction software is run.

Chapter 3

Object reconstruction

To perform physics analyses, the electronic signals produced by the detector need to be reconstructed as physics objects. Physics objects are the final-state particles and observables measured by ATLAS. The reconstruction procedure is divided into several subsequent steps. The first one is the reconstruction of low-level objects in each subdetector, known as *detector objects*, such as primary vertices, particle tracks and energy clusters deposited in the calorimeters. Then, the detector objects are combined to create *high-level physics objects*, such as electrons, photons and muons.

3.1 Detector objects

In this section, the key ingredients for building high-level physics objects will be discussed.

3.1.1 Tracks and primary vertex

Charged particles deposit energy through excitation and ionization when they interact with matter. In the Inner Detector, the energy deposits are reconstructed in 3D points called *hits*. Then, an algorithm combines hits into *tracks* using a three-dimensional

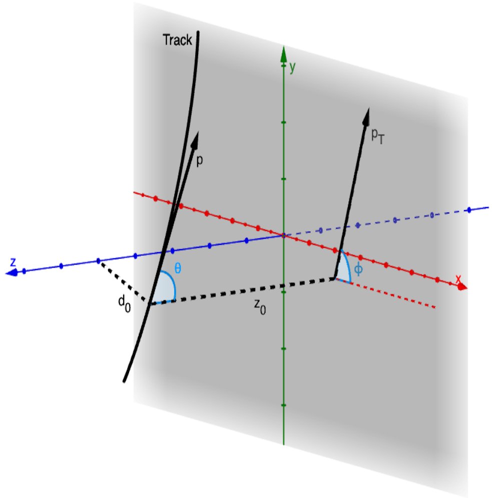
χ^2 fit [58]. Tracks are created starting from seven hits in the silicon layer of the ID, the *track seed*, and then developing outwards by adding additional hits compatible with the seed. With this method, it is possible to assign single hits to multiple tracks, thus ambiguities might arise. To solve these ambiguities, tracks are ranked by quality according to multiple factors such as number of associated hits, number of shared hits, number of missing hits in the silicon layers (*holes*) and value of the χ^2 of the fit.

A charged particle immersed in a solenoid magnetic field moves along a helix trajectory. The trajectory is fully specified by five parameters, as shown in Figure 3.1(a): $(d_0, z_0, q/p_{\mathrm{T}}, \theta, \phi)$. The transverse and longitudinal impact parameters are respectively d_0 and z_0; they represent the minimum distance to the centre of the coordinate system in the $x - y$ plane and along the longitudinal direction. The quantity q/p_{T}, where q is the charge of the particle, is related to the radius of curvature of a charged particle since $p_{\mathrm{T}}[GeV] = 0.3q[Z]B[T]R[m]$. The θ and ϕ angles indicate the direction of the track. Once tracks have been computed, vertices are calculated as the intersection of two or more tracks plus the beam spot as an additional constraint. Due to the high pile-up conditions during the Run 2 data taking, see Figure 2.2(b), multiple vertices are reconstructed for a single bunch crossing. The *primary vertex* is defined as the vertex with the largest $\sum p_{\mathrm{T}}^2$ associated with its tracks. The primary vertex is the vertex in which most likely the hard-scattering interaction has occurred. Thus, tracks not associated with the primary vertex are regarded as potential sources of pile-up. A collision with 66 reconstructed vertices is displayed in Figure 3.1(b) as an example.

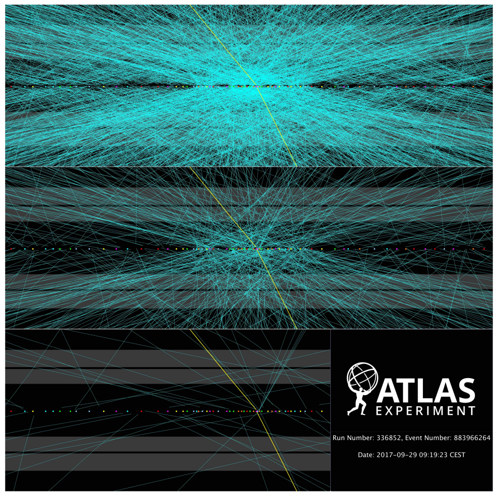
The track reconstruction efficiency depends on the p_{T} and η of the track, reaching a maximum efficiency of $\sim 90\%$, for $p_{\mathrm{T}} > 5$ GeV and Run 2 data-taking conditions [59].

3.1.2 Topological clusters

The *topological clustering algorithm* allows the three-dimensional particle showers to be reconstructed in the calorimeters [61]. This method has been used as the key ingredient to reconstruct jets since Run 1 [62] and lately has also been used to reconstruct electrons and photons due to its ability to better measure bremsstrahlung photons and electrons



(a) Track parameters.



(b) Reconstructed vertices.

Figure 3.1: (a) Schematic of the track parameters for a charged particle moving inside the Inner Detector. (b) An event display of a $Z \to \mu^+\mu^-$ candidate plus 65 additional reconstructed vertices from pile-up interactions with different selections on the minimum p_T of the tracks, from top to bottom: 100 MeV, 1 GeV, and 5 GeV [60].

from photon conversions [63].

The topological clustering algorithm consists of three main steps. The first step is dedicated to finding a *proto-cluster*; these cells are characterised by measured energy exceeding by a factor of four the expected noise. The expected noise is calculated as a quadrature sum of electronic noise and a pile-up contribution, which depends on the instantaneous luminosity of the LHC. At this point, proto-clusters are ordered in descending order according to their signal-to-noise ratio. The second step consists in looking for neighbouring cells near each seed with energy larger than twice the noise. If such a neighbour is found, it is added to the proto-cluster. If two proto-clusters are adjacent they are merged. Once there are no more cells or proto-clusters to merge according to previous criteria, all the remaining neighbouring cells with signal above noise are added to their respective proto-clusters. Then, a splitting step is applied to separate local energy maxima in each single proto-cluster. Lastly, the remaining proto-clusters are renamed topo-clusters.

3.2 High-level physics objects

In the next sections, the high-level physics objects relevant for the $VH, H \to b\bar{b}$ will be described.

3.2.1 Electrons

Electrons are reconstructed by matching candidate tracks to topological clusters from the ECAL [63, 64]. If multiple tracks are associated with the same topological cluster, they are ranked according to their respective quality and angular separation from the topological cluster. The highest-ranking track, known as *primary track*, is also required to satisfy $d_0/\sigma_{d_0} < 5$ and $|\Delta z_0 \sin\theta| < 0.5$ mm, where d_0 and Δz_0 are calculated with respect to the primary vertex and σ_{d_0} is the resolution of d_0. Once a candidate electron has been reconstructed, it is important to distinguish whether an electron has been produced during the hard-scattering collision (*prompt*) or not. Non-prompt leptons arise, for example, from semi-leptonic decays of heavy-flavour jets or photon conversion. To reject non-prompt electrons and fake electrons (such as mis-reconstructed pions or photons), two criteria have been developed: *identification* and *isolation*.

A likelihood-based method (LH) is used to identify electrons in ATLAS during Run 2 [63]. The combination of shower profiles and tracking variables allows a discriminant for identifying electron candidates to be created. Three working points are defined based on increasingly tighter cuts on the discriminant: `LH-Loose`, `LH-Medium`, and `LH-Tight` [65]. A *working point* is a selection criterion that corresponds to a fixed performance on a given representative reference sample. The working points have been optimised in bins of electron transverse energy E_{T} and pseudo-rapidity. For example, for $E_{\mathrm{T}} > 25$ GeV, the `LH-Loose` working point provides an identification efficiency greater than 90% for real electrons, with an efficiency of objects mistakenly reconstructed as electrons smaller than 0.8%.

Isolation criteria are defined to reject electrons that have a significant amount of hadronic activity close-by and are therefore likely to be non-prompt. Two variables are defined for this purpose [65]: $E_{\mathrm{T}}^{cone0.2}$ to quantify calorimeter isolation and $p_{\mathrm{T}}^{varcone0.2}$ to quantify track isolation. The former is defined as the sum of transverse energies of the topological clusters within $\Delta R < 0.2$ from the candidate electron. The latter is defined as the sum of transverse momenta of all the tracks ($\sum_i^{tracks} p_{\mathrm{T},i}$) within $\Delta R = min(0.2, 10\,\mathrm{GeV}/p_{\mathrm{T}})$ from the candidate electron track. For the $VH, H \to b\bar{b}$ analysis, two working points are used for the event selection: `FixedCutHighPtCaloOnly` and `FCLoose`: the first one is defined by $E_{\mathrm{T}}^{cone0.2} < 3.5$ GeV, whereas the second one is characterised by an approximate constant efficiency of 99% for real electrons.

The efficiencies to reconstruct prompt electrons are measured in data using the J/ψ and Z resonances [65]. From these measurements, correction scale factors are calculated to adjust the simulations to the values found in data.

3.2.2 Muons

In ATLAS, muons are firstly reconstructed independently in the Inner Detector and the Muon Spectrometer [66]. In a second step, the information of various sub-detectors is combined to obtain the best performance in different regions of the detector. In this way, four types of muons are reconstructed: *combined muons*, *segment-tagged muons*, *calorimeter-tagged muons* and *extrapolated muons*. A combined muon is built from a global fit of the hits in the Inner Detector and the Muon Spectrometer. To recover low-p_{T} muons, a track in the Inner Detector can be classified as a segment-tagged muon if there are few hits in the muon chambers consistent with the track. In the central region of the detector ($|\eta| < 0.1$) the muon chambers offer limited coverage due to the presence of cabling and supporting structures; in this region, a track in the Inner Detector matched to an energy deposit in the calorimeter system consistent with a minimum-ionizing particle can be reconstructed as a calorimeter-tagged muon. In the forward regions ($2.5 < |\eta| < 2.7$), where the Inner Detector does not provide coverage, extrapolated muons are reconstructed using only the Muon Spectrometer. More than 97% of the muons in the $VH, H \to b\bar{b}$ analysis are combined muons. Moreover, the track associated with the muon needs to satisfy quality criteria: $d_0/\sigma_{d_0} < 3$ and $|\Delta z_0 \sin\theta| < 0.5$ mm. The identification and isolation criteria are applied to muons as described below.

For muon identification, four working points are defined: *ID-Loose*, *ID-Medium*, *ID-Tight* and *ID-High-*p_{T}. The first three are designed to be increasingly tighter, while the fourth one is optimized for high transverse momentum muons. The working points are defined based on several discriminating variables, among which being the normalised χ^2 of the combined track fit. For the $VH, H \rightarrow b\bar{b}$ analysis, two working points are used: the *ID-Loose* in the 2-Lepton channel and the *ID-Medium* in the 1-Lepton channel (Section 4.1).

To reject non-prompt muons, three isolation variables are defined: $E_{\mathrm{T}}^{cone0.2}$, $p_{\mathrm{T}}^{varcone0.3}$ and $p_{\mathrm{T}}^{cone0.2}$. The first two have similar definitions to the ones described for electron in Section 3.2.1, whereas the third one is defined as the sum of the transverse momenta of all tracks ($\sum_i^{tracks} p_{\mathrm{T},i}$) within $\Delta R < 0.2$ around the candidate muon. For the $VH, H \rightarrow b\bar{b}$ analysis, two working points, based on the previous variables, are considered: `FixedCutLoose`, which keeps an efficiency for real muons $\sim$ 99%, and `FixedCutHighPtTrackOnly` defined by $p_{\mathrm{T}}^{cone0.2} < 1.25$ GeV.

Similarly to electrons, muon-efficiency scale factors are measured in data using the J/ψ and Z resonances [66].

3.2.3 Hadronic taus

The tau lepton has an average decay length of 87 μm so it typically decays before reaching the Inner Detector [46]. It can decay into hadrons, τ_{had}, or leptons ($\tau \rightarrow e\nu_e\nu_\tau$ and $\tau \rightarrow \mu\nu_\mu\nu_\tau$). Taus decaying leptonically are considered as electrons and muons, plus $E_{\mathrm{T}}^{\mathrm{miss}}$, in the $VH, H \rightarrow b\bar{b}$ analysis. The branching ratio for the combined hadronic channels is $\sim$ 65%. A tau decaying hadronically is characterised by the presence of a single charged pion in 75% (1-prong) of the cases and three charged pions in about 22% (3-prong) of the decays. The charged pions are classified as a jet of particles. Neutral pions are additional decay products in 78% of all tau hadronic decays. A Boosted Decision Tree algorithm [67] has been developed to identify jets as originating from hadronic tau decays and calibrated identification working points have been provided: *Loose*, *Medium* and *Tight* [68]. In the $VH, H \rightarrow b\bar{b}$ analysis, hadronic taus are selected using the *Medium* ID working point, which has a tau identification efficiency of 75% for 1-prong events and 60% for 3-prong events and corresponding mis-identifications rates for hadronic jets as hadronic taus of 20 and 150, respectively.

3.2.4 Jets

Immediately after quarks and gluons have been produced in a hard-scattering collision, they fragment and hadronise, forming a spray of collimated particles, known as *jet* [69]. Multiple algorithms can be used to cluster particles or calorimeter deposits into jets. Jets can be assembled with two types of procedure: *top-down* and *bottom-up*. The former one is based on the principle that energy, in QCD, flows into a cone. The latter approach tends to combine particles around a high-transverse-momentum seed according to a given metric. Many criteria define a good jet reconstruction algorithm, among which there is infrared and collinear (IRC) safety [69]. If the number of jets in an event is not modified by a collinear splitting of a high-momentum parton or the addition of a soft emission, the jet algorithm is said to be IRC safe. The *anti-k_t* jet clustering algorithm is IRC safe and provides jets with circular shapes in the $\eta - \phi$ plane [70].

The anti-k_t algorithm is the main method for reconstructing jets in ATLAS [62]. For *EMTopo* jets, topological clusters are used as inputs of the anti-k_t algorithm. EM, in EMTopo, refers to the measured energy calibrated at the electromagnetic energy scale, which does not take into account corrections for the signal loss originating from hadrons for the non-compensating nature of the ATLAS calorimeters [62].

In the anti-k_t algorithm, each pair of topological clusters ij is associated with the following metric:

$$d_{ij} = \min\left(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2}\right)\frac{\Delta_{ij}^2}{R^2} \tag{3.1}$$

$$d_{iB} = \frac{1}{k_{t,i}^2} \tag{3.2}$$

where $k_{t,i}$ is the transverse momentum of the cluster i, $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ measures the distance between the two clusters and R is a fixed parameter that determines approximately the jet radius. For the $VH, H \to b\bar{b}$ analysis and the majority of the ATLAS analyses that use calorimeter jets, R is set to 0.4. Lastly, d_{iB} can be interpreted as the distance between the particle i and the beam. The clustering procedure consists of the following steps:

1. Calculate d_{ij} and d_{iB} for all clusters.
2. Find the minimum distance among d_{ij} and d_{iB}.
   (a) If the minimum distance is d_{ij}, the momenta of the clusters i and j are combined.
   (b) If the minimum distance is d_{iB}, the cluster i is classified as jet and removed from the list of the clusters.
3. Repeat from step 1. until all clusters are combined into jets.

An alternative approach to reconstruct jets combing information from the tracker and the calorimeters is the *particle-flow* reconstruction [71]. In this procedure, at first, the energy deposits in the calorimeters associated with charged particles are removed. Then, jets are reconstructed using tracks matched to the primary vertex and the remaining energy deposits in the calorimeters from the neutral particles. These particle-flow objects are subsequently reconstructed using the anti-k_t algorithm. There are many benefits associated with using particle-flow jets:

- Tracks' transverse momentum from charged particles can be reconstructed down to $p_{\mathrm{T}}^{trk} > 400$ MeV with good precision.
- Tracks' angular resolution is significantly better than the resolution provided by the calorimeters.
- Particle-flow jets are more resilient to pile-up jets compared to EMTopo jets since it is possible to match the calorimeters' energy deposits with the primary vertex.

EMTopo jets will be used in the $VH, H \to b\bar{b}$ analysis described in Chapter 4, while particle-flow jets will be tested in Section 5.1.

3.2.4.1 Jet Energy-Scale calibration

The *Jet Energy-Scale* calibration (JES) aims at correcting the measured jet energy deposits to that of truth jets reconstructed at the particle-level energy scale [72]. *Truth jets* are jets reconstructed from simulations using stable final-state particles as inputs, before they go through the detector simulation and reconstruction. Therefore, truth jets do not suffer from resolution and efficiency degradations. Muons, neutrinos and particles from pile-up are excluded from the reconstruction of truth jets. Reconstructed calorimeter jets, from simulations, are matched to truth jets by requiring $\Delta R_{reco,truth} < 0.3$. The JES calibration procedure is applied in the following order:

- *Origin correction*. When jets are reconstructed for the first time, they point to the centre of the detector. This first correction recomputes the jets' four-momentum to point to the hard-scatter primary vertex, improving the jet η resolution up to a factor 20 for jets with $p_{\rm T} > 200$ GeV ($\eta_{res} \sim 0.03 \to \eta_{res} \sim 0.006$) [72].
- *Jet area-based pile-up correction*. This correction is an area-based method that removes the per-event pile-up contribution to the transverse momentum of each jet according to its area [73].
- *Residual pile-up correction*. A second pile-up correction is applied which depends on the number of reconstructed vertices in the event.
- *Absolute MC-based calibration*. The four momenta of the reconstructed jets are corrected to the particle-level energy scale.
- *Global Sequential calibration* (GSC): The GSC is a calibration that aims at correcting for the different calorimeter responses to quark-initiated and gluon-initiated jets.
- *Residual in situ calibration* (only to data). This is the last step in the JES calibration procedure and does not affect simulated events. To further improve Data /

MC agreement, jets in data are calibrated using well understood and precisely measured physics processes. The $Z \to l^+l^-$+*jets* and γ+*jets balance methods* rely on the excellent resolution for photons and leptons to tightly constrain the jet p_{T} using the conservation of energy.

The total uncertainty associated with the complete Jet Energy-Scale calibration is $\sim$ 4.5% at 20 GeV, $\sim$ 1% at 200 GeV and $\sim$ 2% at 2 TeV for an inclusive di-jet sample [72].

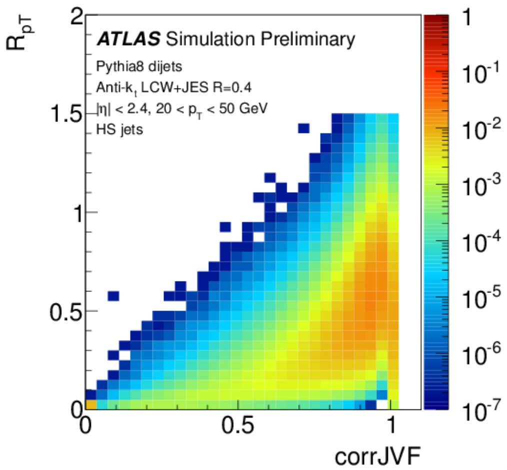
3.2.4.2 Suppression of pile-up jets

Suppression of jets originating from pile-up vertices is critical in Run 2 due to the average number of pile-up verteces being larger than 30. The jet-vertex-tagger (JVT) is designed to discriminate hard-scattering jets from pile-up [74]. It is constructed by combining two variables using a k-nearest-neighbour algorithm, R_{pT} and corrJVF, which are independent of the number of reconstructed primary vertices; R_{pT} is defined as the scalar p_{T} sum of the tracks that are associated with the jet and the hard-scattering vertex divided by the calibrated jet p_{T}; corrJVF is defined as the scalar p_{T} sum of the tracks that are associated with the jet and the hard-scattering vertex divided by a weighted average of all the tracks associated to the jet. R_{pT} and corrJVF offer complementary discrimination as can be seen in Figure 3.2(a) and 3.2(b), which is further enhanced in the JVT shown in Figure 3.2(c).

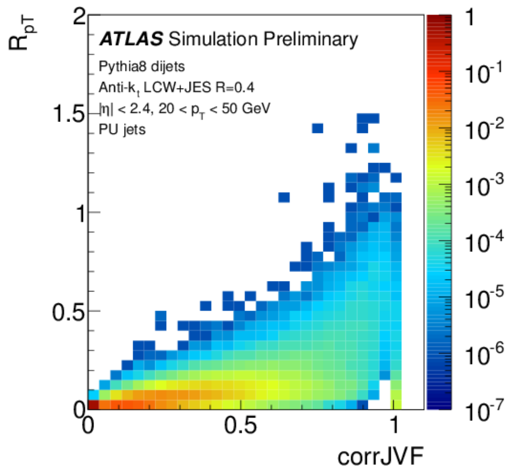
Three working points for the hard-scatter efficiency have been defined in the kinematic range 20 GeV $< p_{\mathrm{T}} <$ 120 GeV and $|\eta| < 2.5$; the working points for EMTopo jets are shown in Table 3.1.

Working points	JVT	Efficiency EMTopo	Pileup fake rates
Loose	> 0.11	97%	3%
Medium	> 0.59	92%	1.0%
Tight	> 0.91	85%	0.4%

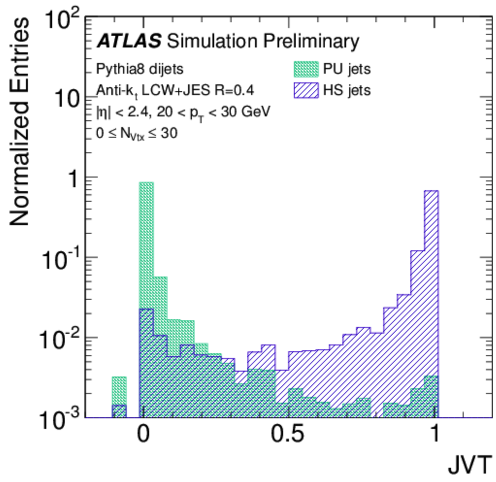
Table 3.1: Working points and selection thresholds for the JVT hard-scatter jet efficiencies and pileup fake rates in EMTopo jets.



(a) Hard-scatter jets.



(b) Pile-up jets.



(c) JVT.

Figure 3.2: Correlations of corrJVF and R_{pT} for hard-scatter (a) and pile-up jets (b) [74]. JVT output score for hard-scatter (a) and pile-up jets (c) [74].

3.2.4.3 *b*-tagging

The identification of jets originating from beauty quark production, *b-tagging*, is of great importance for many physics programs in ATLAS, such as $H \to b\bar{b}$ measurements. Particles containing at least one b-quark, b-hadrons, are characterized by many properties that allow their identification. They have a relatively long lifetime of about 1.5 ps. Therefore, a b-hadron with $p_T = 50$ GeV will travel on average 3 mm in the transverse plane before decaying ($L_{xy} = \beta\gamma c\tau$, see Figure 3.3). The Inner Detector allows the identification of primary vertices, where the proton-proton collision occurs, and secondary vertices, which are vertices displaced from the z-axis where long-lived hadrons decay into lighter particles. A second property that characterizes b-hadrons is the high mass of their decay products compared to light-flavour hadrons. This feature results in a large number of decay products, a phenomenon called *higher multiplicity*. Moreover, b-hadron decay products are characterized by higher transverse momenta compared to the decay products of light and charm quarks. A third property which is typical of b-hadrons is their large branching ratio for semileptonic decays. The processes $B \to X_c + l + \nu$ and $B \to X_u + l + \nu$, where l could be either an electron or a muon and X_c and X_u are decay products containing charm and up quarks, have a branching ratio of $\sim 10\%$ and $\sim 9\%$, respectively [46].

A jet containing at least one b-hadron is conventionally called a b-jet. A c-jet is defined to contain at least one c-hadron but no b-hadrons. Lastly, a *light*-jet (or l-jet) does not contain any b/c-hadrons.

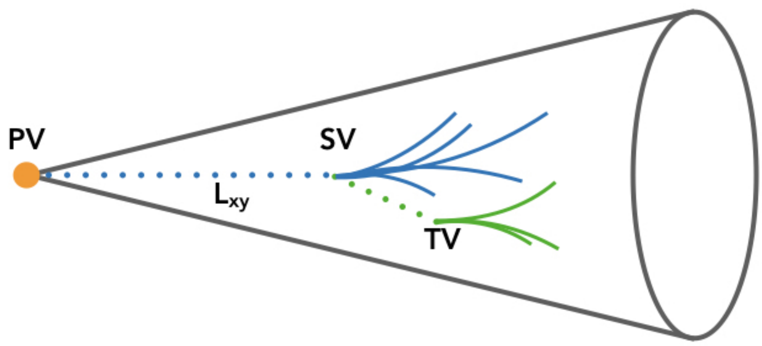
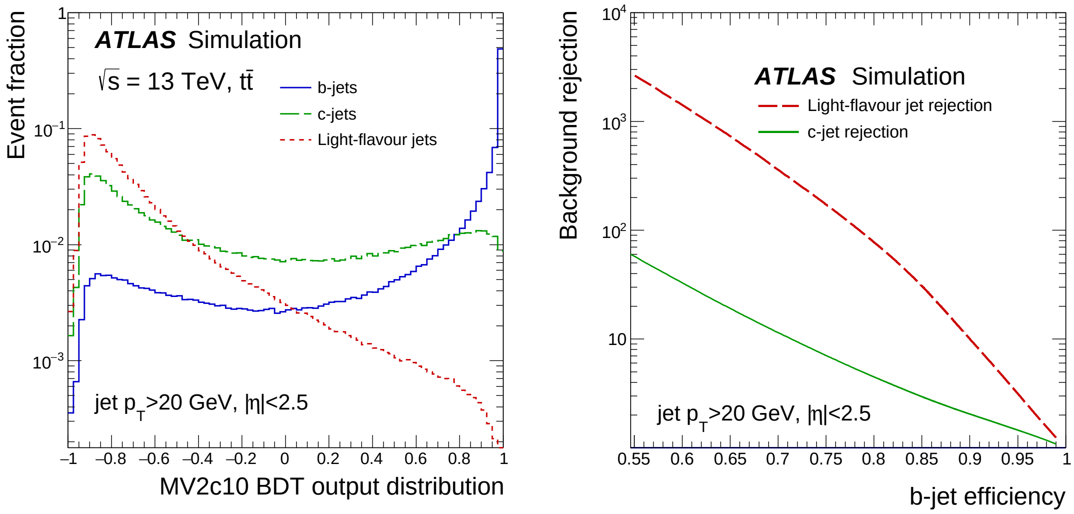


Figure 3.3: Pictorial representation of a b-jet, in which the b-hadron decays into a c-hadron. Primary (PV), secondary (SV) and tertiary (TV) vertices are highlighted.

B-tagging algorithms are designed to discriminate b-jets from c-jets and l-jets using the properties mentioned above. In ATLAS, there are two types of b-tagging algorithms [75, 76]: low-level algorithms, which directly use features of b-jets, and high-level algorithms, which combines the outputs of the low-level algorithms in a single discriminant. For the $VH, H \to b\bar{b}$ analysis, the b-tagging high-level algorithm of interest is the MV2c10, whose inputs originate from three low-level taggers: IP3D, SV1 and JetFitter. The IP3D combines the transverse and longitudinal impact parameter significance (d_0/σ_{d_0} and $z_0 \sin\theta/\sigma_{z_0 \sin\theta}$) of the tracks in the jet into a single log-likelihood-ratio discriminant. The secondary vertex finding algorithm, SV1, attempts the reconstruction of the displaced secondary vertex; this algorithm aims at rejecting secondary vertices in jets originating from non-b-hadrons, such as from K_S and Λ, by calculating the invariant mass of the tracks associated with the secondary vertex. The JetFitter algorithm searches for the full $b \to c$ decay chain inside a jet by looking for the secondary and tertiary vertices. The MV2c10 algorithm is a BDT [67] trained on simulated top-quark pair production ($t\bar{t}$) events; where b-jets are considered as the signal, and a mixture of c-jets (10%) and *light*-jets (90%) are considered background for the training [75]. The output score of the MV2c10 can be seen in Figure 3.4(a) and the b-tagging efficiency as a function of the c/l-rejection rates is shown in Figure 3.4(b).



(a) MV2c10 output score.

(b) B-tagging efficiency vs c-/l-rejections.

Figure 3.4: MV2c10 output scores from the $t\bar{t}$ simulated sample (a). B-tagging efficiency as a function of the c/l-rejection rates (b) [75].

For the $VH, H \to b\bar{b}$ analysis, central jets ($|\eta| < 2.5$) are tagged as b-jets by requiring the MV2c10 to be above a given threshold; the chosen threshold corresponds to a 70% efficiency for identifying a b-jet and a rejection rate of 12 for c-jet and 381 for l-jet.

The Data/Simulations scale factors (shown in Figure 3.5 for b-jets) for the MV2c10 are measured in data by using tagged jets to correct for the different tagger performance on simulated events compared to data events. The scale factors and their respective uncertainties for b-jets, c-jets and l-jets are calculated in five MV2c10 bins and are measured in samples enriched with $t\bar{t}$ events for jet p_T between 20 GeV and 600 GeV [76]. For jet p_T greater than 600 GeV the uncertainties are extrapolated from simulations. The

c-tagging scale factors are measured in samples enriched with semi-leptonic $t\bar{t}$ events [77], while the scale factors for the mistagging rates of l-jets are measured in multi-jet events [78]. Lastly, MC-MC scale factors are included to account for the discrepancies in the efficiency between different MC generators.

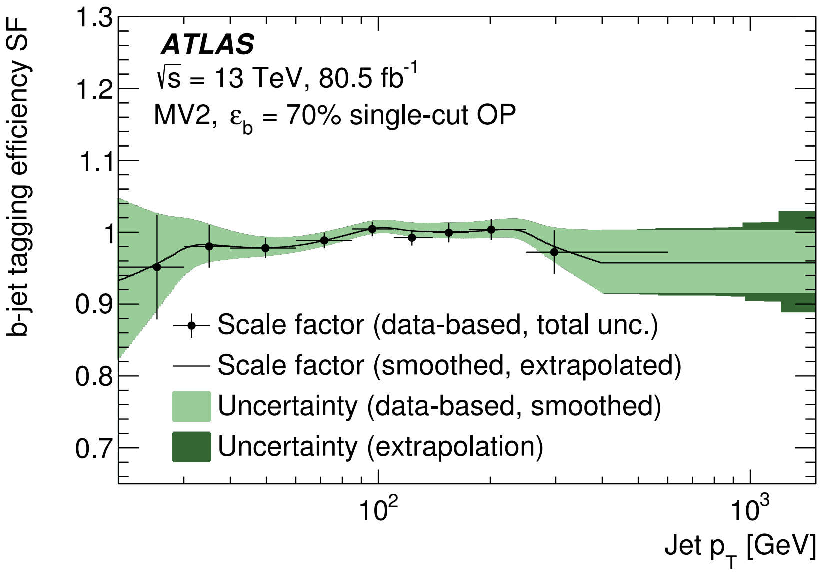


Figure 3.5: B-tagging efficiency scale factors for the MV2c10 70% working point [76].

3.2.5 Missing Transverse Energy

Particles not interacting electromagnetically or strongly do not leave any energy deposit in the ATLAS detector. However, it is possible to infer their production thanks to the laws of conservation of energy and momentum. At the LHC, the initial momenta along the z-axis of the colliding partons are unknown; however, their initial momenta in the transverse plain is approximately zero. Therefore, any missing transverse momentum, measured after a collision, can be attributed to non-interacting particles such as neutrinos leaving the event undetected. The missing transverse energy is calculated by the negative vectorial sum of two components: a *hard term* and a *soft term*. The hard missing transverse energy corresponds to the sum of the momenta in the transverse plane of the reconstructed and calibrated physics objects. The soft missing transverse energy, also known as the *track-based soft term* (TST), is calculated from the vectorial sum of transverse momenta of tracks associated with the primary vertex but not to any physics objects [79].

$$\begin{aligned}
\mathbf{E}_{\mathbf{T}}^{\mathbf{miss}} &= -\sum_{\text{hard}} \mathbf{p}_{\mathbf{T}} - \sum_{\text{soft}} \mathbf{p}_{\mathbf{T}} \\
&= -\sum_{e} \mathbf{p}_{\mathbf{T}} - \sum_{\mu} \mathbf{p}_{\mathbf{T}} - \sum_{\tau} \mathbf{p}_{\mathbf{T}} - \sum_{\gamma} \mathbf{p}_{\mathbf{T}} - \sum_{jets} \mathbf{p}_{\mathbf{T}} - \sum_{\substack{unused\\tracks}} \mathbf{p}_{\mathbf{T}}
\end{aligned} \tag{3.3}$$

Since $\mathbf{E}_{\mathbf{T}}^{\mathbf{miss}}$ is a vector in the transverse xy-plane, it is possible to calculate its magnitude and direction in the ϕ coordinate:

$$E_T^{\text{miss}} = |\mathbf{E}_{\mathbf{T}}^{\mathbf{miss}}| = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2} \tag{3.4}$$

$$\phi^{\text{miss}} = \arctan\left(\frac{E_y^{\text{miss}}}{E_x^{\text{miss}}}\right). \tag{3.5}$$

The track-based soft term brings advantages and disadvantages. On one hand, it is quite resilient to pile-up, since tracks are associated with the primary vertex. On the other hand, it does not include neutral particles and particles outside the acceptance of the Inner Detector ($|\eta| > 2.5$) [80].

The performance of E_T^{miss} reconstruction is evaluated for scale (linearity), resolution, and sensitivity to the data-taking conditions. It is measured in events with no expected E_T^{miss}, such as $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$, and events where the E_T^{miss} is expected from the presence of at least a neutrino, such as W decays and top-quark pair production [79]. The E_T^{miss} resolution depends on the detector acceptance, reconstruction efficiency and resolutions of all physics objects; in particular, the dominating contribution originates from jet resolution.

3.2.6 Overlap Removal procedure

After leptons and jets have been reconstructed, some physics objects may share tracks or calorimeter clusters. The *Overlap Removal Procedure* is designed to remove ambiguities and double-counting between reconstructed electrons, muons, jets and hadronically decaying tau leptons [81]. For example, hadronic taus are almost always reconstructed also as jets. Two physics objects are considered to be the same based on criteria such as shared tracks or proximity in space. To resolve the ambiguity, the Physics Object is only kept in one of the reconstructed object categories and is thus unambiguously defined. The overlap removal affects the following pairs of physics objects:

- τ_{had} - e: if $\Delta R(\tau_{had}, e) < 0.2$, the τ_{had} is removed.
- τ_{had} - μ: if $\Delta R(\tau_{had}, \mu) < 0.2$, the τ_{had} is removed.
- e - μ: if a combined μ shares a track with an e, the e is removed.
- e - jet: if $\Delta R(e, jet) < 0.2$, the jet is removed.
- μ - jet: if $\Delta R(\mu, jet) < 0.2$ and $N^{jet}_{traks} < 3$, the jet is removed.
- τ_{had} - jet: if $\Delta R(\tau_{had}, jet) < 0.2$, the jet is removed

Chapter 4

Measurements of $\mathbf{H \to b\bar{b}}$ decays in the VH production mode

The Higgs boson was introduced in 1964 [18] to explain electroweak symmetry breaking. The first phenomenological studies of its production and decay rates dated back to the early 70s [82]. Since then, this boson has been searched for in many experiments, before being finally discovered at the LHC by ATLAS and CMS [20, 21]. The search for this particle has been particularly challenging because the theoretical predictions did not provide any clear indication of what m_H should have been [83]. First experimental evidence in the mid-70s from nuclear physics allowed to set a lower limit $m_H \gtrsim \mathcal{O}(10)$ MeV [84]. Conversely, theoretical calculations suggested an upper mass limit of $m_H \lesssim \mathcal{O}(1)$ TeV [85]. In this mass region, the $H \to b\bar{b}$ decays played an important role in the search. As soon as the Higgs decay into two beauty quarks becomes kinematically allowed at $m_H \sim \mathcal{O}(10)$ GeV, the $H \to b\bar{b}$ decays become the leading branching ratio up until $m_H \sim \mathcal{O}(150)$ GeV, where the WW decay mode becomes dominant. For this reason, the $H \to b\bar{b}$ decay mode was the primary search channel at LEP [86] and the Tevatron [82]. In 2012, at the Tevatron $p\bar{p}$ collider, the CDF and D0 experiments observed an excess of events in the data compared with the background predictions in the mass range $[120, 135]$ GeV [87]. In general, at hadron colliders, targeting the VH ($V = W, Z$) production mode provides the highest sensitivity in the $H \to b\bar{b}$ decay channel because the leptonic signatures from V decays allow the vast multi-jet background to be rejected and provide effective triggers [88].

At the LHC, $H \to b\bar{b}$ decays were observed in 2018 by ATLAS and CMS [2, 89]. For ATLAS, the observation of this decay channel has been possible thanks to the combination of the available production channels and its Run 1 measurements. However, the leading contributor to the observation of $H \to b\bar{b}$ decays was the $VH, H \to b\bar{b}$ Run 2 analysis, with a dataset of $L = 79.8$ fb^{-1}.

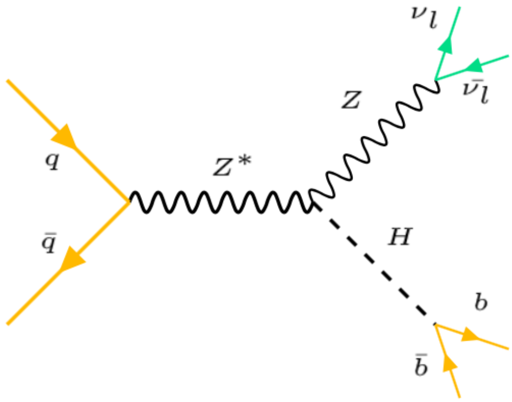
The $VH, H \to b\bar{b}$ analysis was further improved using the full Run 2 data sample of $L = 139$ fb^{-1} as described in the next sections with particular focus on the 0-Lepton channel (see Section 4.1) [4].

4.1 Analysis overview

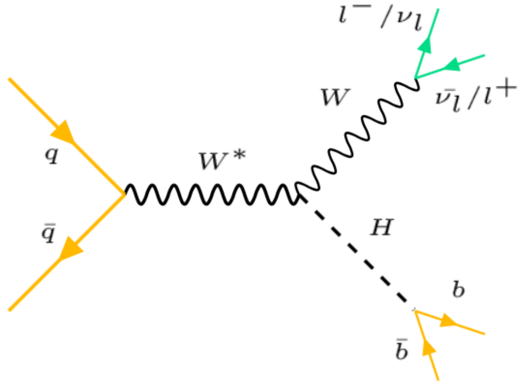
In the $VH, H \to b\bar{b}$ analysis , three lepton channels are defined according to the number of charged leptons (electrons, muons or taus decaying leptonically) present in each event: 0-Lepton, 1-Lepton and 2-Lepton to target the $ZH \to \nu\nu b\bar{b}$, $WH \to l\nu b\bar{b}$ and $ZH \to llb\bar{b}$ signatures. For each channel, events are required to have exactly two jets identified as b-jets using the b-tagging requirement detailed in Section 3.2.4.3 and additional non-b-tagged jets are allowed. The signal and background signatures will be discussed in Section 4.2 and Section 4.3. In Section 4.4, the simulated samples for the signal and background processes are detailed. Based on the kinematic properties of signal and backgrounds, the complete event selection, described in Section 4.5, will be applied to the events. The definition of signal and control regions will be discussed in Section 4.6. In Section 4.7, the Simplified Template Cross-Section framework will be introduced. Strategies to reduce the statistical errors associated with the simulated samples will be discussed in Section 4.9. Three techniques to improve the resolution of the invariant mass of the two b-jets system, m_{bb}, will be presented in Section 4.8. To better separate the signal from the background processes, multivariate techniques will be introduced in Section 4.10. The statistical analysis procedure will be discussed in Section 4.11. The main multivariate analysis will be accompanied with two supporting results: a multivariate analysis aimed at measuring the $VZ, Z \to b\bar{b}$ diboson signal strength and a non-multivariate analysis aimed at measuring the $VH, H \to b\bar{b}$ processes using the invariant mass of the two b-jets system. The latter will be referred to as *di-jet analysis*. The full description of the analysis systematic uncertainties and the fit's result will be shown in Section 4.12 and Section 4.13.

4.2 Signal

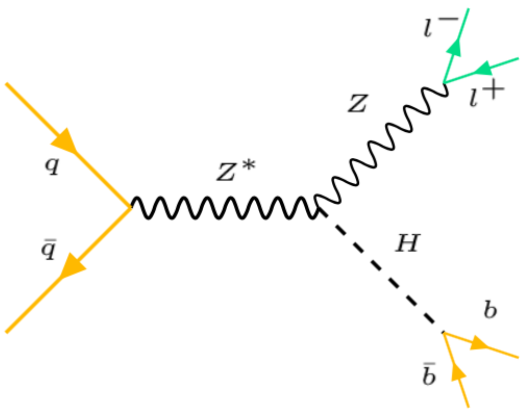
At the LHC, the VH production mode can occur in two ways: through the annihilation of a quark-antiquark pair or a loop initiated by two gluons; the leading-order Feynman diagrams for these two processes are shown in Figure 4.1 and 4.2.



(a) 0-Lepton.



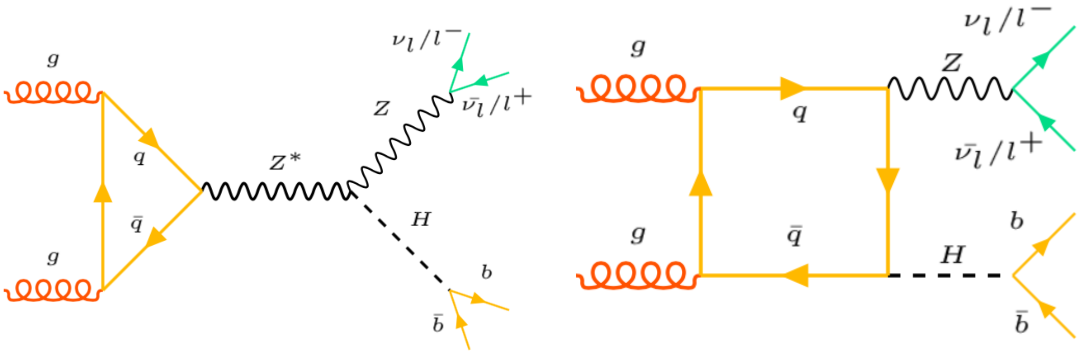
(b) 1-Lepton.



(c) 2-Lepton.

Figure 4.1: Leading-order Feynman diagrams for quark initiated $VH, H \to b\bar{b}$ processes for the 0-Lepton (a), 1-Lepton (b) and 2-Lepton (c) channels.

The signal-over-background ratio of $VH, H \to b\bar{b}$ is more than one order of magnitude smaller than for $H \to ZZ^*$. Particularly, the VH cross-section for transverse momenta of the vector boson greater than 120 GeV is only $\sim$ 5% of the total VH cross-section



(a) Triangle diagram. (b) Box diagram.

Figure 4.2: Leading-order Feynman diagrams for gluon initiated $VH, H \to b\bar{b}$ processes for the 0-Lepton and 2-Lepton channels.

[88]. However, at these high values of p_{T}^{V}, the VH production becomes the most sensitive mode to search for $H \to b\bar{b}$ decays. In this regime, the Higgs boson and the vector boson are mostly produced back-to-back and in the central region of the detector, where leptons from the vector-boson decay can be easily used for triggering. In Figure 4.3, the $E_{\mathrm{T}}^{\mathrm{miss}}$ shapes of the signal and main backgrounds for the 0-Lepton channel show that the $VH, H \to b\bar{b}$ signal is harder. To enhance the analysis sensitivity, signal and control regions will also be divided according to p_{T}^{V}, as described in Section 4.6.

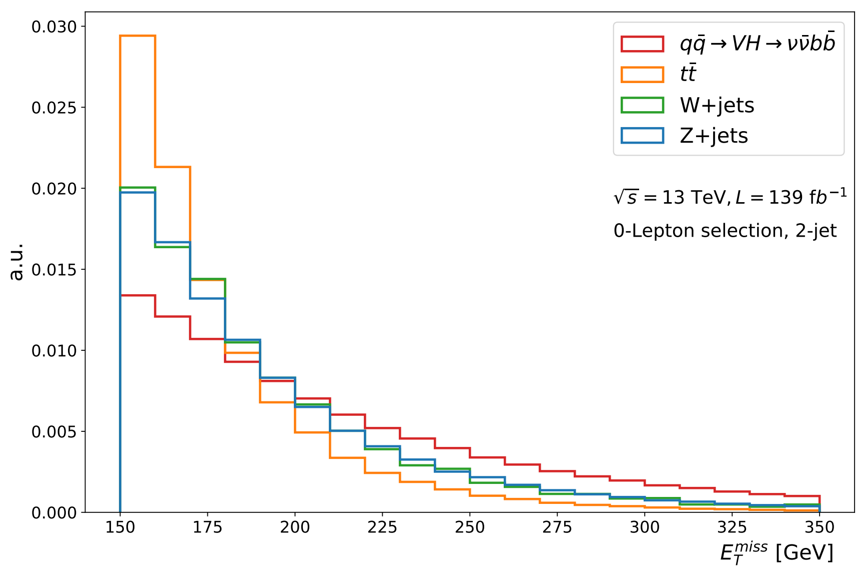


Figure 4.3: $E_{\mathrm{T}}^{\mathrm{miss}}$ normalised distributions for the VH 0-Lepton signal and main backgrounds for events with 2 jets.

As the transverse momentum of the vector boson increases, the b-jets originating from the Higgs tend to get closer to one another. Under the assumption of $m_{b_1}, m_{b_2} \ll m_H$ (m_{b_1} and m_{b_2} are ordered according to their p_{T}), the angular separation of the two b-jets can be approximated as [90]:

$$\Delta R(b_1, b_2) \sim 2 \cdot \frac{m_H}{p_{\mathrm{T}}^H} \tag{4.1}$$

Figure 4.4 shows the distribution of $\Delta R(b_1, b_2)$ against Higgs transverse momentum for the signal $q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}$; this behaviour will be exploited in Section 4.6 to define signal and control regions.

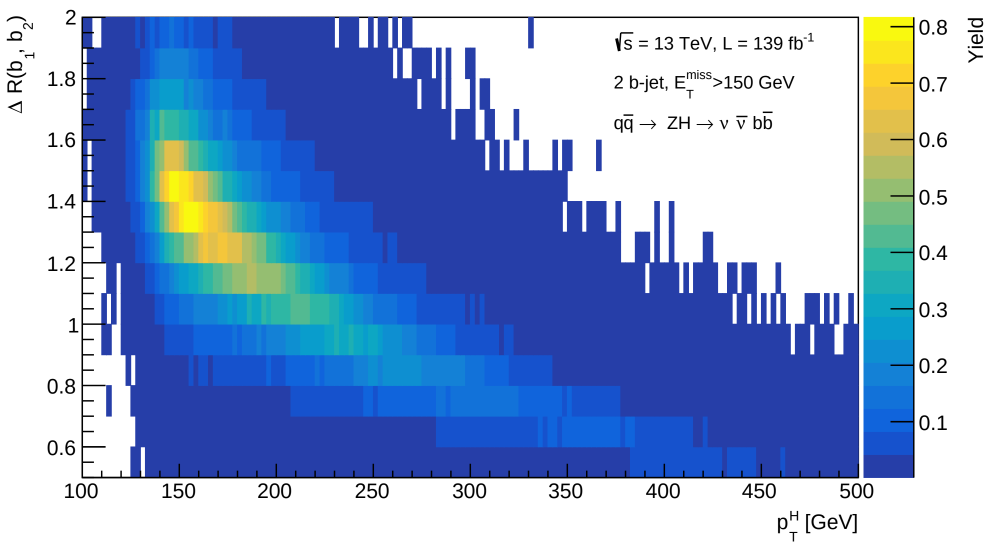


Figure 4.4: The signal $q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}$ $\Delta R(b_1, b_2)$ distribution as a function of the Higgs traverse momentum for events with two jets and $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV.

The most powerful variable to discriminate the $VH, H \to b\bar{b}$ signal against all the backgrounds is the invariant mass of the two b-jets, m_{bb}, which peaks at 125 GeV for the signal and is non-resonant for the most important backgrounds except for diboson production (see Section 4.3.3). This variable is the most highly ranked in the main multivariate analysis (see Section 4.10) and it is fitted directly in the di-jet analysis. The resolution on m_{bb} is mainly limited by the Jet Energy Resolution; however, there are some techniques specific to the $VH, H \to b\bar{b}$ analysis that further improve upon the GSC jet energy calibration (see Section 3.2.4.1), as it will be shown in Section 4.8.

It is quite common for a b-jet to emit a hard gluon, which gets reconstructed as an additional jet. This process is known as Final-State Radiation (FSR). Similarly, partons inside the protons can emit Initial-State Radiation (ISR) before the hard-scattering collision. Therefore, in the $VH, H \rightarrow b\bar{b}$ analysis, events are categorised according to the total number of jets in the event, requiring always exactly two b-jets and allowing one more additional non-b-tagged jet for the 0-Lepton channel. Events with four or more jets are not considered in the 0-Lepton channel because in that category too many $t\bar{t}$ background events would enter. A study of the Final-State Radiation will be presented in Section 4.2.1.

Although the 0-Lepton channel is dominated by the $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ processes, it has also a sizeable component from $WH \rightarrow l\nu b\bar{b}$. In fact, the WH component entering in 0-Lepton can be divided into two parts: W decaying into charged leptons lost outside the acceptance of the detector and W decaying into τ that decays hadronically. The latter case will be discussed in Section 5.4.

4.2.1 Final-state radiation

The 3-jet category in the $VH, H \rightarrow b\bar{b}$ analysis contains several events originating from Final-State Radiation. These events are characterised by peculiar kinematics that is interesting to study particularly to improve the jet energy resolution. In the available signal MC simulations, it is possible to classify the third jet, which is non-b-tagged, into four groups. The first three groups have distinct physical origins, while the last one has been introduced due to the impossibility of labelling jets as ISR or FSR for $p_{\text{T}} < 25$ GeV:

- *Pile-up (PU)*: the third jet does not originate from the hard-scatter vertex.
- *ISR*: the third jet has a $p_{\text{T}} > 25$ GeV and originates from ISR.
- *FSR*: the third jet has a $p_{\text{T}} > 25$ GeV and originates from FSR.
- *Low* p_{T}: the third jet has a $p_{\text{T}} < 25$ GeV and comes either from ISR or FSR.

FSR events are 14.8% of the total 0-Lepton 3-jet signal events, while PU contributes

16.3%, ISR 48.5% and Low p_T 20.4%. Figure 4.5 shows m_{bb} distributions for various categories of 3-jet events. FSR events are characterised by an m_{bb} distribution that peaks at values lower than 100 GeV, whereas the m_{bb} distribution for PU and ISR events peaks around the the Higgs boson mass of 125 GeV.

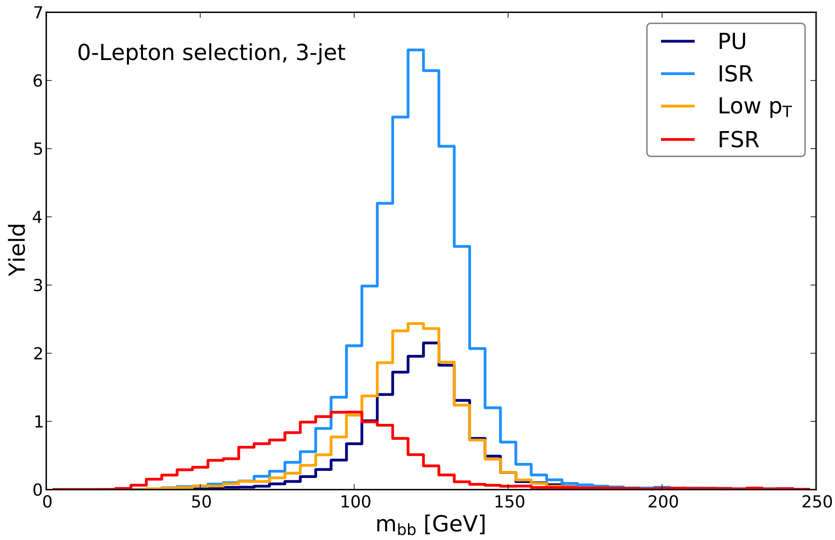


Figure 4.5: Simulated m_{bb} distributions for the 0-Lepton 3-jet category in signal events. The colour scheme encodes the origin of the third jet. The event yield on the y-axis represents the number of simulated weighted events in each bin.

In Figure 4.6, the invariant mass of the three-jet system, m_{bbj}, for FSR events peaks near 125 GeV while PU and ISR peak at higher values. It can also be seen that for low p_T events, the m_{bbj} resembles a mixture of the FSR and ISR distributions.

It is possible to visualize graphically the position of the FSR jet with respect to the two b-jets using a simplified model and few assumptions. For every event, the leading b-jet is translated to the origin of the $\eta - \phi$ plane. This transformation can then also be applied to the sub-leading b-jet and the third jet

$$\begin{cases} \eta'_{b_2} = \eta_{b_2} - \eta_{b_1} \\ \phi'_{b_2} = \phi_{b_2} - \phi_{b_1} \end{cases} \qquad \begin{cases} \eta'_{j_3} = \eta_{j_3} - \eta_{b_1} \\ \phi'_{j_3} = \phi_{j_3} - \phi_{b_1} \end{cases} \tag{4.2}$$

Then, the sub-leading b-jet can be rotated on the diagonal on the first quadrant of the $\eta - \phi$ plane

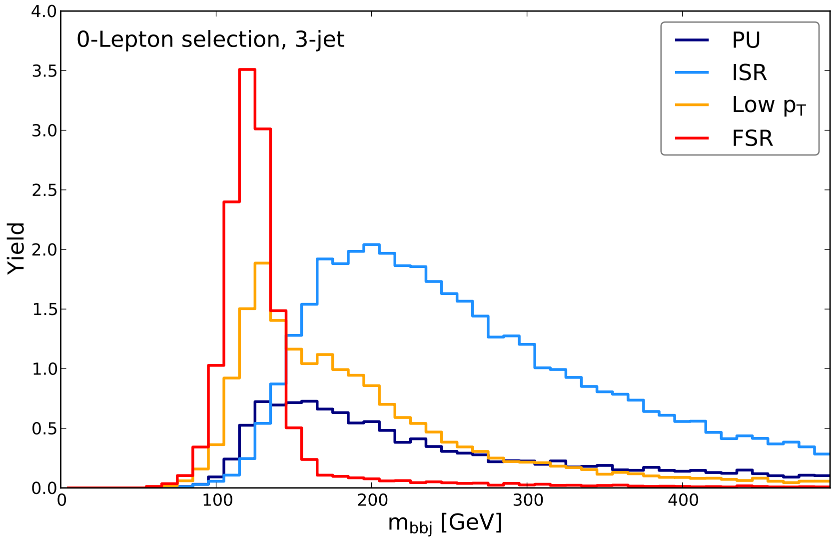


Figure 4.6: Simulated m_{bbj} distributions for the 0-Lepton 3-jet category in signal events. The colour scheme encodes the origin of the third jet. The event yield on the y-axis represents the number of simulated weighted events in each bin.

$$\eta''_{b_2} = \phi''_{b_2} > 0$$

$$\begin{cases} \eta''_{b_2} = \eta'_{b_2} \cdot \cos\alpha_{rot} - \phi'_{b_2} \cdot \sin\alpha_{rot} \\ \phi''_{b_2} = \eta'_{b_2} \cdot \sin\alpha_{rot} + \phi'_{b_2} \cdot \cos\alpha_{rot} \end{cases} \tag{4.3}$$

$$\alpha_{rot} = \arctan\left(\frac{\eta'_{b_2} - \phi'_{b_2}}{\eta'_{b_2} + \phi'_{b_2}}\right)$$

If $\eta''_{b_2} < 0$ (rotation in the third quadrant), then a factor of π is added to α_{rot}. Analogously, the third jet can be rotated using the same transformation

$$\begin{cases} \eta''_{j_3} = \eta'_{j_3} \cdot \cos\alpha_{rot} - \phi'_{j_3} \cdot \sin\alpha_{rot} \\ \phi''_{j_3} = \eta'_{j_3} \cdot \sin\alpha_{rot} + \phi'_{j_3} \cdot \cos\alpha_{rot} \end{cases} . \tag{4.4}$$

At this point, rather than showing the position of each sub-leading b-jet along the diagonal, it is possible to show an "average" position using the relation in Equation 4.1 and assuming $p_{\mathrm{T}}^{H} = p_{\mathrm{T}}^{V}$. After considering the previous assumptions and the coordinates transformation, the result is shown in Figure 4.7.

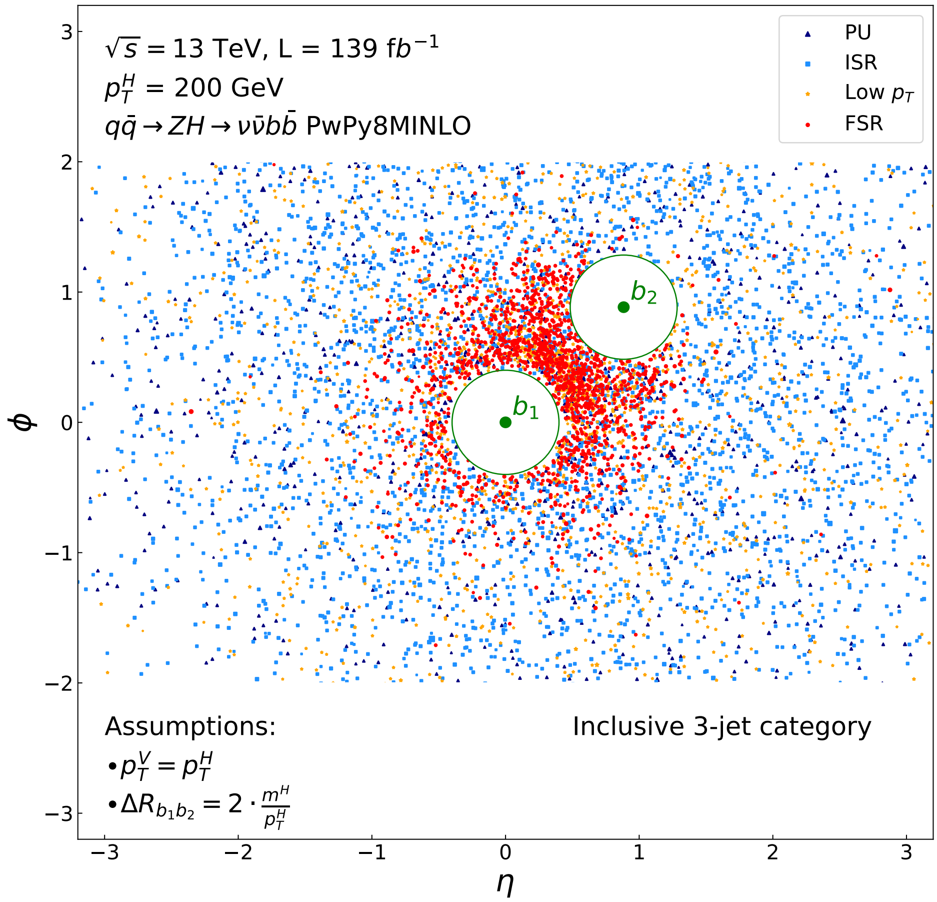


Figure 4.7: $\eta - \phi$ distribution of the third jet in the 0-Lepton 3-jet category for events with $p_{\mathrm{T}}^{H} \in [195, 205]$ GeV. A simplified model is used to assume the position of the second b-jets in the event. The size of the marker is proportional to the respective event weight. Analogous plots, with different values of p_{T}^{H}, can be found in Appendix B. The colour scheme encodes the origin of the third jet.

Figure 4.7 indicates that FSR jets tend to form mostly between the two b-jets. Thus, it is possible to construct the following variable:

$$\Delta R(b_1, j_3) + \Delta R(b_2, j_3) \tag{4.5}$$

which defines an ellipse in which the b-jets are the two foci. The variable defined in Equation 4.5 offers excellent discrimination for FSR jets using only topological properties, as can be seen in Figure 4.8. This property will be used to define a Final-State

Radiation Recovery strategy in Section 5.2 to improve the energy resolution of b-jets.

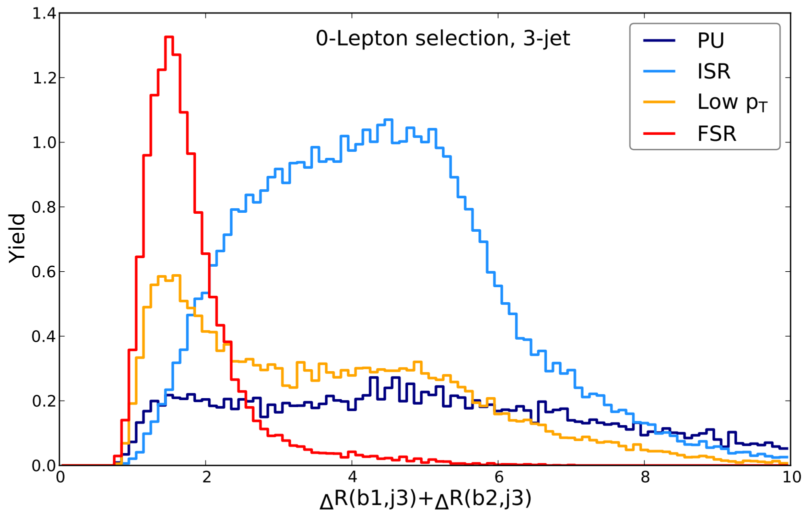


Figure 4.8: $\Delta R(b_1, j_3) + \Delta R(b_2, j_3)$ distribution for the third jet in the 0-Lepton 3-jet category in signal events. The colour scheme encodes the origin of the third jet. The event yield on the y-axis represents the number of simulated weighted events in each bin.

4.3 Backgrounds

Many backgrounds enter the phase space of $VH, H \to b\bar{b}$ analysis. In the next sections, the main characteristics of each background will be presented in order of decreasing importance.

4.3.1 V+jets production

The leading-order Feynman diagrams for the W+jets and Z+jets productions are shown in Figure 4.9. When the gluon in Figure 4.9(b) splits into two b-quarks, it creates the same final-state objects as the signal sought in all the 2-jet categories. For this reason, the V+jets background is one of the most challenging of the $VH, H \to b\bar{b}$ analysis.

The Z+jets is the largest background entering the 0-Lepton phase space due to its large cross-section, compared to the signal, and its branching ratio to two neutrinos $BR(Z \to \nu\bar{\nu}) = 20\%$ [46]. Other decays of the Z boson are negligible in this lepton channel. The normalisation of the Z+jets background is extrapolated from the 2-Lepton channel as

(a) $V + q$. (b) $V + g$.

Figure 4.9: Leading-order Feynman diagrams for W+jets and Z+jets production.

will be discussed in Section 4.12. The Z+jets, with the Z boson decaying into charged leptons, is the largest background entering the 2-Lepton channel.

In Table 4.1, the simulated W+jets decays entering in the 0-Lepton phase space are reported. About half of the events originate from tau hadronic decays. For the others W leptonic decays, the e or μ could be outside the detector acceptance: because they have large $|\eta^{lep}|$ or low p_{T}^{lep}. Lastly, in a very limited number of events, the lepton is reconstructed as the third jet in the 3-jet category, as can be seen in the last column of Table 4.1(b).

Kinematically, the W+jets and Z+jets in the 0-Lepton channel are very similar to each other. One key difference between the two is the distribution of $\Delta R(b_1, b_2)$ (also referred as $dR(b_1, b_2)$), shown in Figure 4.10.

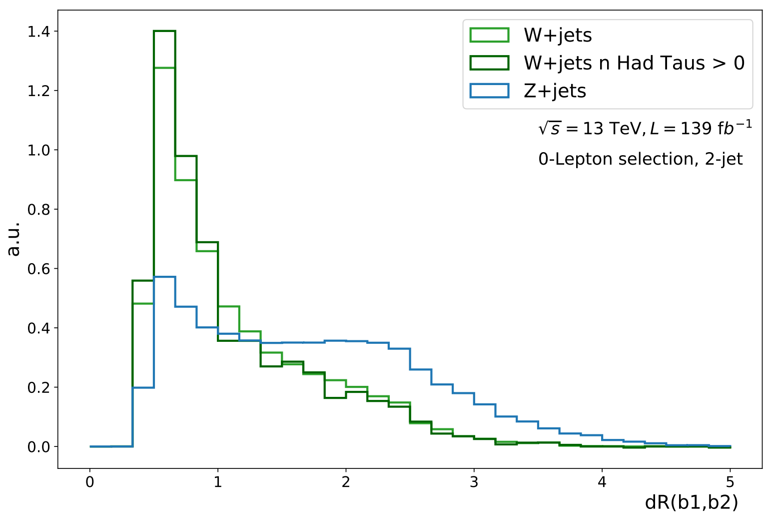


Figure 4.10: $dR(b1, b2)$ normalised shapes for V+jets processes.

W decay	Events	Events in acceptance
$W \to e\nu_e$	85(7.5%)	16(18.3%)
$W \to \mu\nu_\mu$	259(23.0%)	30(11.8%)
$W \to \tau\nu_\tau, \tau \to e\nu_e$	96(8.6%)	14(14.2%)
$W \to \tau\nu_\tau, \tau \to \mu\nu_\mu$	134(11.9%)	53(39.5%)
$W \to \tau\nu_\tau, \tau \to had$	551(49.0%)	314(57.0%)
$W \to had$	-	-

(a) 2-jet inclusive category.

W decay	Events	Events in acceptance	Events $\Delta R(l, j_3) < 0.3$
$W \to e\nu_e$	251(10.1%)	69(27.4%)	61(23.9%)
$W \to \mu\nu_\mu$	568(22.7%)	61(10.8%)	13(2.2%)
$W \to \tau\nu_\tau, \tau \to e\nu_e$	199(8.0%)	26(13.1%)	7(3.3%)
$W \to \tau\nu_\tau, \tau \to \mu\nu_\mu$	312(12.5%)	107(34.2%)	3(0.1%)
$W \to \tau\nu_\tau, \tau \to had$	1168(46.7%)	662(56.7%)	92(7.8%)
$W \to had$	-	-	-

(b) 3-jet inclusive category.

Table 4.1: Simulated W+jets decays in the 0-Lepton channel. The decays are categorised according to the truth lepton in the event, which is missed by the reconstruction procedure. Table (a) and (b) show respectively the inclusive 2-jet and 3-jet categories. The *Events* column contains the total number of weighted events passing the 0-Lepton selection. The percentages in the Event column are calculated with respect of the total W+jets events, i.e. column-wise. The *Events in acceptance* column shows the numbers of events in which the truth lepton is within the detector acceptance. The acceptance region considered is: $|\eta^{lep,\tau_{had}}| < 2.5$ and $p_{\mathrm{T}}^{lep} > 10$ GeV or $p_{\mathrm{T}}^{\tau_{had}} > 20$ GeV. The *Events* $\Delta R(l, j_3) < 0.3$ column contains the numbers of events in which the truth lepton is within an angular distance of 0.3 from the reconstructed third jet; events passing this criterion indicate that the lepton as been reconstructed as a jet. The percentages in the last columns are calculated with respect to the W decay, i.e. row-wise.

For W+jets, $dR(b_1, b_2)$ is particularly peaked at low values. This result can be interpreted with the help of Table 4.1: since in $W \to \tau \to had$ decays there is only one neutrino and part of the energy of the W-system gets absorbed by the τ, the W-system has also to be highly boosted to reach the threshold $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV. Therefore, the recoiling gluon has to be also highly boosted and thus the resulting jets are very close to each other. For what concerns the invariant mass of the b-jet system, the V+jets processes peak at lower m_{bb} than the signal, as can be seen in Figure 4.11.

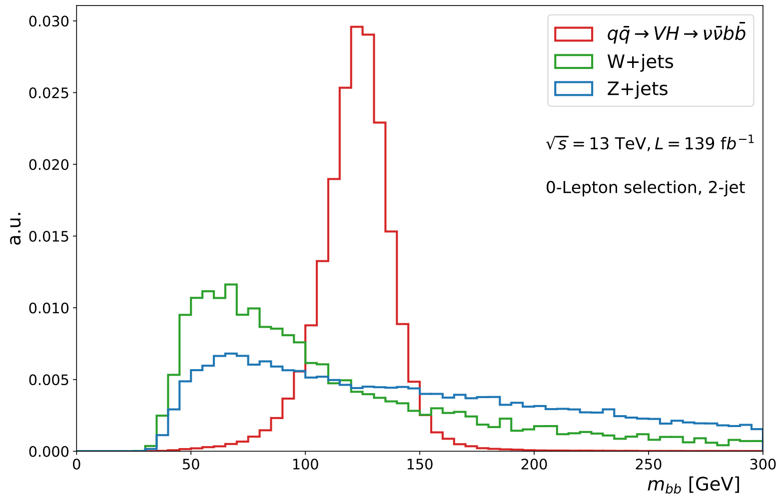


Figure 4.11: Normalised m_{bb} distributions for the signal and the V+jets backgrounds.

4.3.2 Top pair and single top productions

The production of top-quark pairs ($t\bar{t}$) constitutes an important background for the $VH, H \to b\bar{b}$ analysis. This background is the largest that enters the 1-Lepton channel. A simplified schematic of the $t\bar{t}$ decay is reported in Figure 4.12. The production of $t\bar{t}$ gives rise to many leptons and b-jets, which reduces the number of events entering the 0-Lepton channel. Nonetheless, many events still enter the 0-Lepton phase space, mostly due to lepton and b-jet misidentification.

Fully hadronic $t\bar{t}$ decays on average produce six jets, from the b-quarks and the W hadronic decays, with no additional neutrinos produced from the W bosons. Due to the small $E_{\mathrm{T}}^{\mathrm{miss}}$ and large jet multiplicity, fully hadronic $t\bar{t}$ decays are negligible in the 0-Lepton channel.

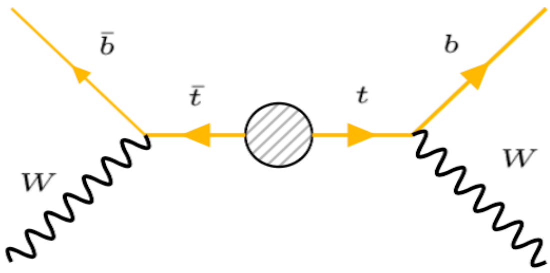


Figure 4.12: Simplified Feynmann diagram $t\bar{t}$ production.

Semi-leptonic and di-leptonic $t\bar{t}$ decays entering the 0-Lepton channel are, in most of the cases, characterised by the presence of τ leptons, as it can be seen in Figure 4.13 and Figure 4.14. Naturally, for 3-jet events, the fraction of semi-leptonic decays is much higher. Furthermore, the event selection threshold can be reached by additional $E_{\mathrm{T}}^{\mathrm{miss}}$ from the neutrinos originating from semi-leptonic b-jet decays.

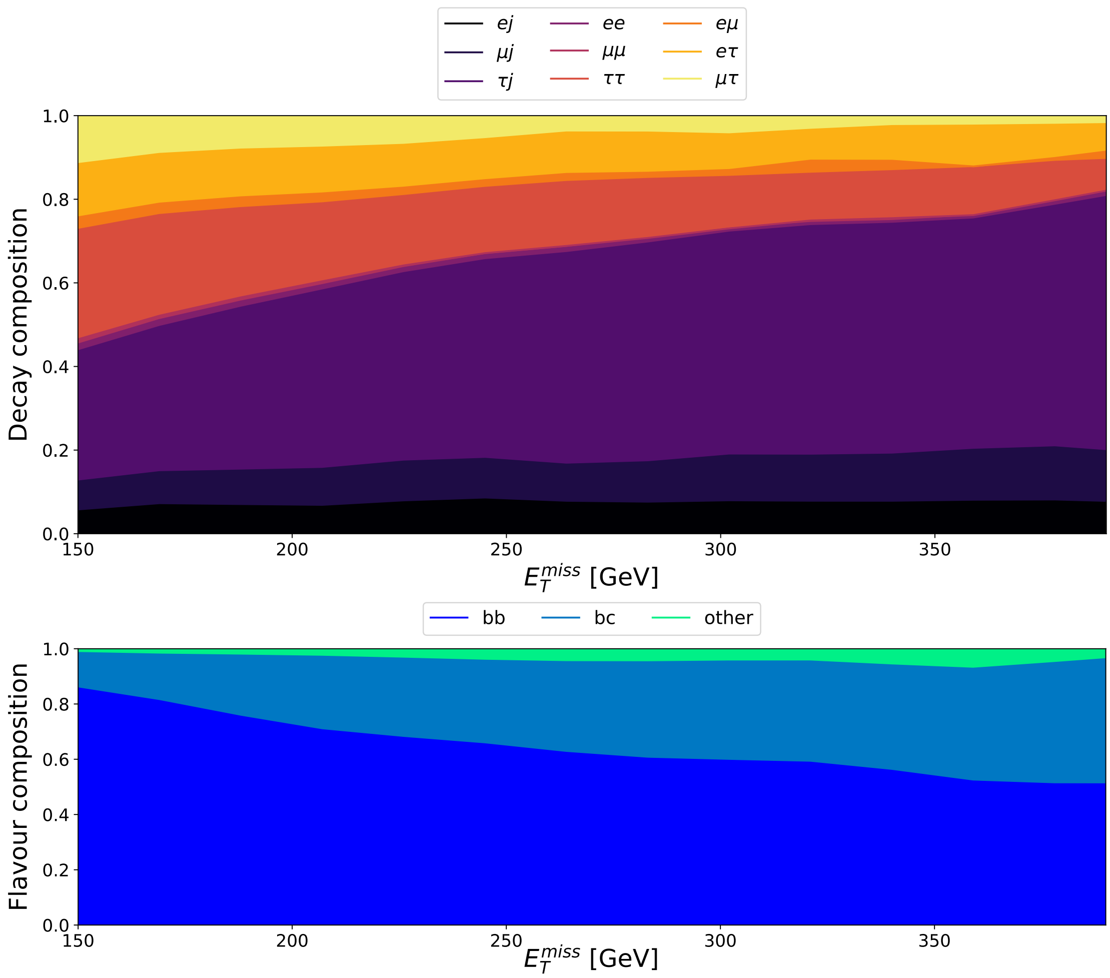


Figure 4.13: Decay and flavour composition for $t\bar{t}$ events entering the 2-jet 0-Lepton channel as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$.

It is interesting to notice that at low $E_{\mathrm{T}}^{\mathrm{miss}}$, there is a good fraction of di-leptonic $t\bar{t}$ events and the two b-jets reconstructed mainly come from the bottom quarks originating from the top quarks.

For high values of $E_{\mathrm{T}}^{\mathrm{miss}}$, the two reconstructed b-jets tend to originate more often from a b and a c quark. This is because, at high $E_{\mathrm{T}}^{\mathrm{miss}}$, a W decaying into cs quarks can be reconstructed as a b-jet. This process is more common at high $E_{\mathrm{T}}^{\mathrm{miss}}$ since the W and the b-quark tend to be more close to each other. In this scenario, the second b-quark from the top decay is not identified.

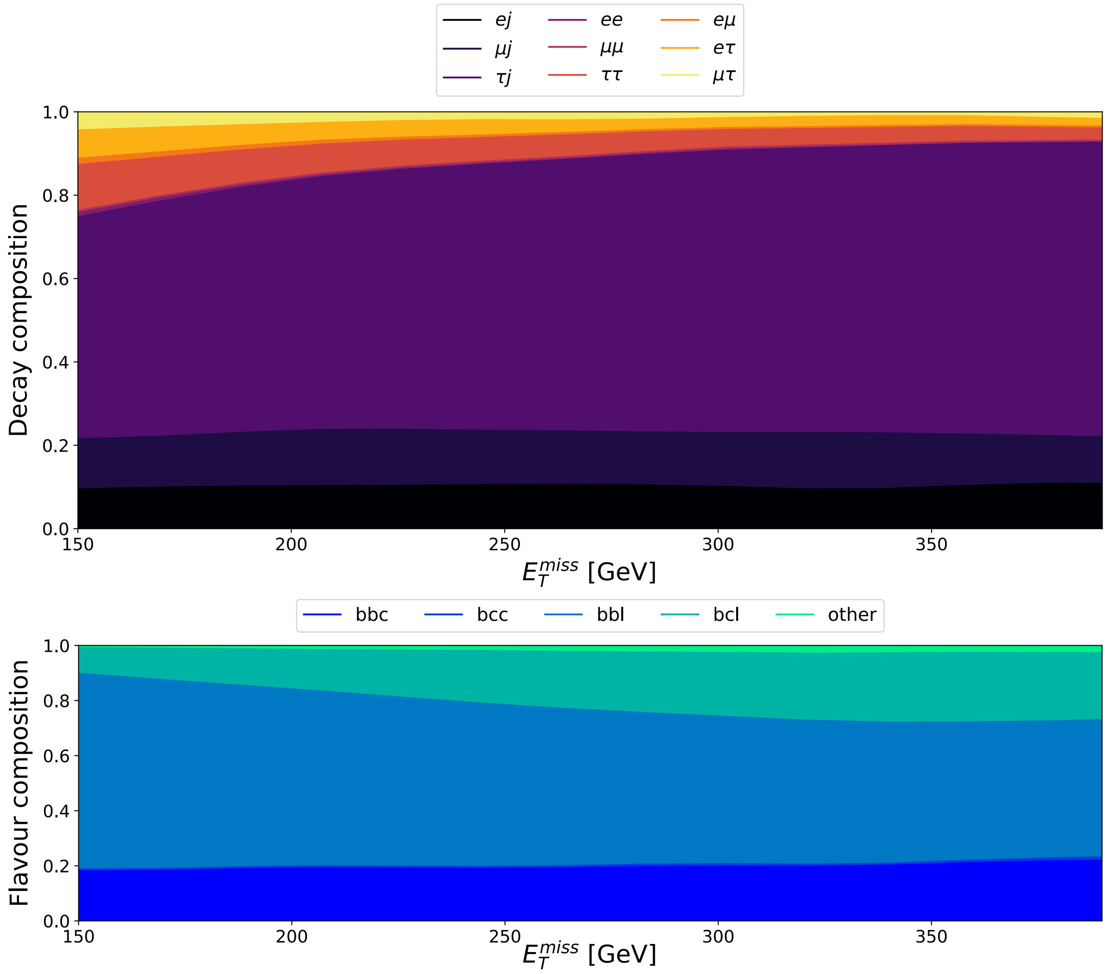
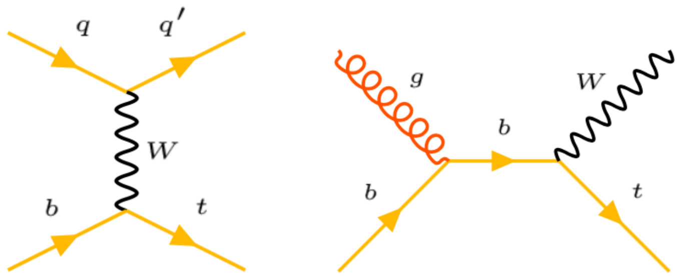


Figure 4.14: Decay and flavour composition for $t\bar{t}$ events entering the 3-jet 0-Lepton channel as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$.

Leading-order diagrams for the three single-top production channels are illustrated in Figure 4.15: $t-$channel, $Wt-$channel and $s-$channel. The Wt production has been generated with a diagram-removal procedure to account for the interference with $t\bar{t}$ since many beyond-leading-order diagrams for Wt production can also be interpreted as the production of $t\bar{t}$ at leading order [91]. The single top production plays a minor role in the 0-Lepton channel because of its relatively low cross-section and the fact that is not often characterised by high values of $E_{\mathrm{T}}^{\mathrm{miss}}$.

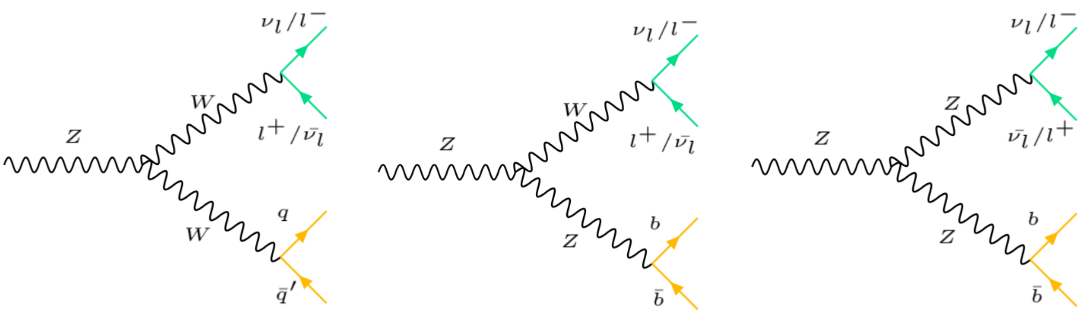


(a) $t-$channel. (b) $Wt-$channel. (c) $s-$channel.

Figure 4.15: Leading-order Feynman diagrams for single-top production in the $t-$channel (a), $Wt-$channel (b) and the $s-$channel (c).

4.3.3 Diboson production

The diboson processes entering the $VH, H \rightarrow b\bar{b}$ phase space are shown in Figure 4.16. The ZZ process gives the highest contribution to the 0-Lepton and 2-Lepton channels, while the WZ is largest in the 1-Lepton channel. The ZZ is also further enhanced due to combinatorics since both Z bosons can decay to $b-$quark pairs and leptons pairs, while in the WZ case, only the Z can decay to $b-$quark pairs. The WW process gives the smallest contribution, mostly in the 1-Lepton channel in the high p_{T}^{V} regime.



(a) WW. (b) WZ. (c) ZZ.

Figure 4.16: Examples of diboson Feynman diagrams entering the $VH, H \rightarrow b\bar{b}$ phase space.

The signal strength of the $WZ, Z \rightarrow b\bar{b}$ and $ZZ, Z \rightarrow b\bar{b}$ processes, where the second vector boson decays leptonically, is measured as a cross-check of the main $VH, H \rightarrow b\bar{b}$ result.

4.3.4 Multi-jet production

Multi-jet production is one of the most abundant processes at the LHC. A schematic for this process is shown in Figure 4.17. Multi-jet enter the $VH, H \rightarrow b\bar{b}$ analysis phase space due to $E_{\mathrm{T}}^{\mathrm{miss}}$ mis-measurements and jets mis-identified as prompt leptons.

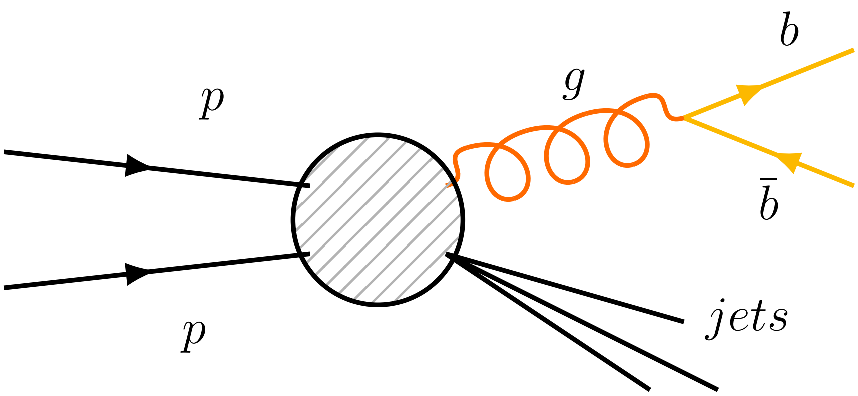


Figure 4.17: Simplified diagram for multi-jet production.

For the 0-Lepton channel, the multi-jet background is negligible thanks to four topological selections, also referred to as *anti-QCD cuts*:

1. $\Delta\phi(b_1, b_2) < 140°$
2. $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, bb) > 120°$
3. $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{miss}) < 90°$
4. $\min\left[\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, jets)\right] > 20°$ for 2-jet events and $> 30°$ for 3-jet events

The first selection is motivated by Equation 4.1, since at high p_{T}^{H} the two b-jets tend to be closer to each other. On the contrary, the multi-jet $\Delta\phi(b_1, b_2)$ distribution tends to peak at 180°, especially for 2-jet events. For the second cut, at high p_{T}^{H} the vector boson tends to be produced back-to-back with respect to the Higgs boson. The third selection assures that the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector measured in the calorimeters is compatible with the p_{T}^{miss} vector measured in the tracker within 90°. The fourth selection removes the remaining multi-jet contribution. Details on the estimation of the remaining multi-jet contribution in the 0-Lepton channel are provided in Appendix C.

In the 1-Lepton channel, the multi-jet contribution cannot be reduced to a negligible level without unacceptably reducing the signal sensitivity. The multi-jet contribution is estimated separately for the 2-jet and 3-jet categories, and in the electron and muon sub-channels. The transverse mass of the W-boson candidate, m_{T}^{W}, offers good discrimination between the multi-jet and the electroweak backgrounds, such as $t\bar{t}$ and W+jets; in fact, leptonic decays of the W boson exhibits a characteristic Jacobian peak in the m_{T}^{W} distribution. For this reason, m_{T}^{W} is fitted in multi-jet-enriched regions orthogonal to the signal and control regions in order to determine the contribution of this background. The multi-jet-enriched regions are defined by applying the 1-Lepton event selection (see Section 4.5), requiring exactly one b-jet rather than two, considering inclusively signal and control regions and inverting (logical `not`) the lepton isolation requirement. The normalisations of the multi-jet-enriched regions are calculated as a difference between the measured data and the simulated electroweak backgrounds after fitting the m_{T}^{W} distribution. Then, the shapes of all the variables can be extrapolated from the multi-jet enriched regions to the signal and control regions; this is achieved by extrapolating $dR(b_1, b_2)$ and p_{T}^{V}.

For the 2-Lepton channel, the multi-jet background is reduced to a negligible level after applying the event selection (Section 4.5). This reduction stems from the requirement of two isolated leptons with invariant mass consistent with the Z boson. The remaining multi-jet contribution is evaluated using the side-bands of the dilepton mass distribution, m_{ll}, in a background-enriched region assuming an exponentially falling distribution for the multi-jet contribution [92]. The background-enriched region is defined by requiring same-sign leptons and removing the Z-boson mass window (81 GeV $< m_{ll} <$ 101 GeV).

4.4 Simulated samples

Simulated MC events are used to model the signal and all the backgrounds for the $VH, H \to b\bar{b}$ analysis, except for the multi-jet production. All events are simulated at least to Next-to-Leading-Order (NLO) accuracy for QCD calculations, except for the gluon induced ZH and ZZ productions which are calculated at Leading-Order (LO) accuracy. An additional electroweak NLO correction is applied to the $qq \to VH$ processes as a function of the transverse momentum of the vector boson using the HAWK package [93, 94]. The default matrix element generators used for this analysis are: POWHEG-BOX V2 [95], for signal interfaced with GOSAM [96] + MINLO [97, 98], and SHERPA [99]. The list of the generated samples is reported in Table 4.2. Moreover, alternative generators are used to derive systematic uncertainties as described in Section 4.12.2. Signal and background events have been simulated using the PDF set provided by the NNPDF3.0NLO collaboration [100]. For the relevant generators, the Higgs mass and $H \to b\bar{b}$ BR have been set to 125 GeV and 58%, respectively, while the top-quark mass to 172.5 GeV. Pile-up has been simulated using PYTHIA 8.186 [101] with the A2 set of tuned parameters and MSTW2008LO PDFs [102].

Filtered samples are used to enhance the available amount of simulated events in the corners of phase space where most of the events are removed by the event selection. These samples are designed to reduce the statistical error of the simulated data set. Filtered samples are generated in bins of *truth-level variables* using stable final-state particles as inputs, instead of the reconstructed physics objects. In the 0-Lepton channel $t\bar{t}$ filtered samples have been generated in order to populate regions with high values of $E_{\mathrm{T}}^{\mathrm{miss}}$; details on these samples will be provided in Section 4.9.1. The main $V + jets$ samples have been generated in bins of $max(H_{\mathrm{T}}, p_{\mathrm{T}}^{V})$ where H_{T} is the scalar p_{T} sum of all the jets in the event and p_{T}^{V} is the transverse momentum of the vector boson. The $V + jets$ $max(H_{\mathrm{T}}, p_{\mathrm{T}}^{V})$ samples are then combined with samples filtered by p_{T}^{V} and m_{jj}, simply referred to as p_{T}^{V}-m_{jj} samples (details in Section 4.9.2). Lastly, the inclusive diboson $VZ, Z \to q\bar{q}$ sample has been combined with a flavour-filtered $VZ, Z \to b\bar{b}$ sample.

Process	Matrix Element generator	Parton Shower	Cross-section order	$\sigma \times$ BR [pb]
Signal				
$qq \to ZH \to llb\bar{b}$	POWHEG-BOX v2 + GOSAM + MINLO	PYTHIA 8.212	NNLO(QCD)+NLO(EW)	$8.91 \cdot 10^{-2}$
$qq \to W^+H \to l^+\nu b\bar{b}$	POWHEG-BOX v2 + GOSAM + MINLO	PYTHIA 8.212	NNLO(QCD)+NLO(EW)	$1.64 \cdot 10^{-1}$
$qq \to W^-H \to l^-\bar{\nu}b\bar{b}$	POWHEG-BOX v2 + GOSAM + MINLO	PYTHIA 8.212	NNLO(QCD)+NLO(EW)	$1.04 \cdot 10^{-1}$
$qq \to ZH \to \nu\bar{\nu}b\bar{b}$	POWHEG-BOX v2 + GOSAM + MINLO	PYTHIA 8.212	NNLO(QCD)+NLO(EW)	$2.99 \cdot 10^{-2}$
$gg \to ZH \to \nu\bar{\nu}b\bar{b}$	POWHEG-BOX v2 + GOSAM + MINLO	PYTHIA 8.212	NLO+NLL [103]	$1.43 \cdot 10^{-2}$
$gg \to ZH \to llb\bar{b}$	POWHEG-BOX v2 + GOSAM + MINLO	PYTHIA 8.212	NLO+NLL	$4.80 \cdot 10^{-3}$
Top pair and single top productions				
$t\bar{t}$	POWHEG-BOX v2	PYTHIA 8.230	NNLO+NNLL [104]	832
s-channel s-top	POWHEG-BOX v2	PYTHIA 8.230	NLO [105]	3.35
t-channel s-top	POWHEG-BOX v2	PYTHIA 8.230	NLO [106]	70.4
Wt-channel s-top	POWHEG-BOX v2	PYTHIA 8.230	Approximate NNLO [107]	71.7
Vector boson + jets				
$W \to \ell\nu$	SHERPA 2.2.1	SHERPA 2.2.1	NNLO [108]	$6.02 \cdot 10^4$
$Z/\gamma^* \to \ell\ell$	SHERPA 2.2.1	SHERPA 2.2.1	NNLO	$6.30 \cdot 10^3$
$Z \to \nu\nu$	SHERPA 2.2.1	SHERPA 2.2.1	NNLO	$1.07 \cdot 10^4$
Diboson				
$qq \to WW$	SHERPA 2.2.1	SHERPA 2.2.1	NLO	45.7
$qq \to WZ$	SHERPA 2.2.1	SHERPA 2.2.1	NLO	21.7
$qq \to ZZ$	SHERPA 2.2.1	SHERPA 2.2.1	NLO	6.53
$gg \to VV$	SHERPA 2.2.2	SHERPA 2.2.2	NLO	3.09

Table 4.2: Nominal simulated samples used in $VH, H \to b\bar{b}$ analysis and corresponding cross-sections at $\sqrt{s} = 13$ TeV [4].

4.5 Event selection

The complete event selection for the three lepton channels in the $VH, H \to b\bar{b}$ analysis is reported in Table 4.3. All the events are required to have exactly two b-jets, forming the Higgs-boson candidate. One or more additional non-b-tagged jets are allowed, depending on the lepton channel. Central jets must have $p_{\mathrm{T}} > 20$ GeV while forward jets $p_{\mathrm{T}} > 30$ GeV. The highest p_{T} b-jet, also referred to as *leading b-jet*, must have a transverse momentum greater than $p_{\mathrm{T}}^{b_1} > 45$ GeV, given the high p_{T}^{V} of the vector boson against which it recoils. In the next sections, the specific event selection for each channel is described. The signal acceptances for the three channels after the complete event selection are reported in Table 4.4.

4.5.1 0-Lepton selection

Events with no loose leptons (definition in Section 3.2.1 and Section 3.2.2) are accepted in the 0-Lepton channel. The multi-jet background is rejected using four selections described in Section 4.3.4. The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is used to select events during on-line data acquisition. Different trigger thresholds are deployed for the various years of data-taking to account for different beam conditions. $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger thresholds range between 70 GeV and 110 GeV. An off-line selection on the fully-calibrated $E_{\mathrm{T}}^{\mathrm{miss}}$ is placed at 150 GeV. The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency is measured in the 1-Lepton channel using single-muon triggered data in V+jets and $t\bar{t}$ events. The trigger efficiency reaches a plateau at $E_{\mathrm{T}}^{\mathrm{miss}} \sim 180$ GeV, and before that correction scale factors, with values greater than 0.9, are applied to the simulations to correct for residual turn-on effects for differences in data and simulations. Lastly, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger shows a small dependence on the jet multiplicity in the event, which is removed by applying a selection on the scalar sum of the transverse momenta of the jets: $H_{\mathrm{T}} > 120$ GeV and $H_{\mathrm{T}} > 150$ GeV for 2-jet and 3-jet events, respectively

4.5.2 1-Lepton selection

In the 1-Lepton channel, events with exactly one tight lepton (definition in Section 3.2.1 and Section 3.2.2) are accepted (see Sections 3.2.1 and 3.2.2). Single-electron triggers are used in the electron sub-channel and $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are used in the muon sub-channel since these triggers perform better at high p_{T}^{V} than muon triggers. An off-line p_{T} requirement is applied to leptons: $p_{\mathrm{T}}^{e} > 27$ GeV and $p_{\mathrm{T}}^{\mu} > 25$ GeV. An additional selection $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV is applied to improve the rejection of the multi-jet background in the 1-Lepton channel.

4.5.3 2-Lepton selection

Events with exactly two same-flavour loose leptons are accepted in the 2-Lepton channel. Single-lepton triggers are used both for the electron and muon sub-channels. The leading lepton p_{T} has to be greater than 27 GeV. Lastly, the invariant mass of the two lepton system, m_{ll}, has to be compatible with the Z boson mass (81 GeV $< m_{ll} <$ 101 GeV).

Number of jets	2 or 3 jets (3 or more for 2-Lepton).
b-jets	Exactly 2 b-jets
$p_{\mathrm{T}}^{b_1}$	> 45 GeV
Central jets p_{T}	> 20 GeV
Forward jets p_{T}	> 30 GeV
p_{T}^{V} regions	$[75, 150]$ GeV (only 2-Lepton), $[150, 250]$ GeV, $[250, \infty]$ GeV
Signal regions	$\Delta R(b_1, b_2)$ signal selection
Control regions	High and low $\Delta R(b_1, b_2)$ side-bands

(a) Common selections and categorisation

Trigger	$E_{\mathrm{T}}^{\mathrm{miss}}$
Leptons	0 loose leptons
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 150 GeV
H_{T}	> 120 GeV (2 jets), >150 GeV (3 jets)
$\min[\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jets})]$	$> 20°$ (2 jets), $> 30°$ (3 jets)
$\Delta\phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, bb)$	$> 120°$
$\Delta\phi(b_1, b_2)$	$< 140°$
$\Delta\phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})$	$< 90°$

(b) 0-Lepton

Sub-channels	electron sub-channel and muon sub-channel
Trigger	single lepton (e sub-channel)
	$E_{\mathrm{T}}^{\mathrm{miss}}$ (μ sub-channel)
Leptons	1 tight electron, $p_{\mathrm{T}}^{e} > 27$ GeV (e sub-channel)
	1 tight muon, $p_{\mathrm{T}}^{\mu} > 25$ GeV (μ sub-channel)
	0 additional loose leptons
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 30 GeV (e sub-channel)

(c) 1-Lepton

Trigger	single lepton
Leptons	2 loose lepton, $p_{\mathrm{T}}^{l} > 27$ GeV
Lepton flavours	same
Lepton sign charges	opposite only for muons
m_{ll}	$81 < m_{ll} < 101$ GeV

(d) 2-Lepton

Table 4.3: Event selection and categorisation for the $VH, H \to b\bar{b}$ analysis [4].

Process	Acceptance [%]		
	0-Lepton	1-Lepton	2-Lepton
$qq \to ZH \to \nu\nu b\bar{b}$	1.9	–	–
$gg \to ZH \to \nu\nu b\bar{b}$	3.5	–	–
$qq \to WH \to l\nu b\bar{b}$	0.2	1.1	–
$qq \to ZH \to llb\bar{b}$	< 0.1	0.2	6.4
$gg \to ZH \to llb\bar{b}$	< 0.1	0.3	14.5

Table 4.4: Signal acceptances obtained from simulated signal samples for the three lepton channels. The acceptances are computed as the fraction of events remaining in the inclusive signal and control regions after the event selection has been applied [4].

4.6 Event Categorisation

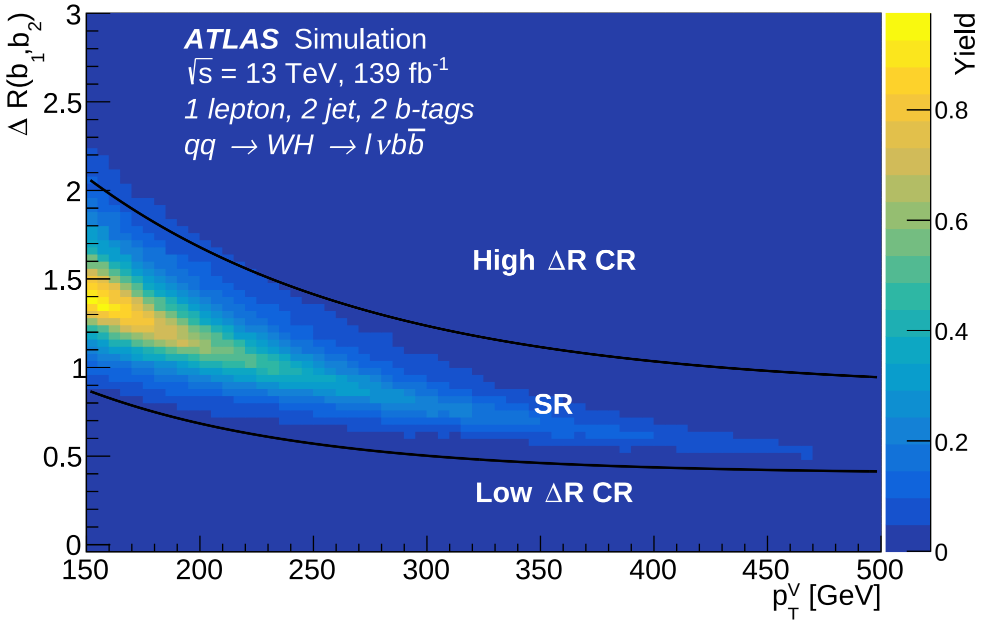
Events are categorised according to multiple criteria in signal and control regions. The categorisation criteria are the number of charged leptons, the transverse momentum of the reconstructed vector boson (or $E_{\mathrm{T}}^{\mathrm{miss}}$ for the 0-Lepton channel) and a continuous selection based on two p_{T}^{V} dependent $\Delta R(b_1, b_2)$ criteria. The continuous selection is reported in Equation 4.6 and Equation 4.7. The criteria have been designed for the 1-Lepton channel and extended to the two other lepton channels. The selections in Equation 4.6 define a control region for high values of $\Delta R(b_1, b_2)$, referred to as High-ΔR-CR, which is enriched in $t\bar{t}$ events in the 0-Lepton and 1-Lepton channels. The specific values of the continuous selections have been designed to keep 95% and 85% of the signal events in the signal region (SR). Conversely, the selections in Equation 4.7 define the Low-ΔR-CR, which retains 90% of the diboson events in the SR, which is sufficient to perform the diboson validation analysis (see Section 4.13.3). In Figure 4.18, the SRs and CRs selections for the 1-Lepton channel are shown.

$$\Delta R_{\text{2-jet category}}(b_1, b_2) > 0.87 + e^{1.38 - 0.00795 \cdot p_T^V}$$
$$\Delta R_{\text{3-jet category}}(b_1, b_2) > 0.76 + e^{1.33 - 0.00730 \cdot p_T^V} \tag{4.6}$$

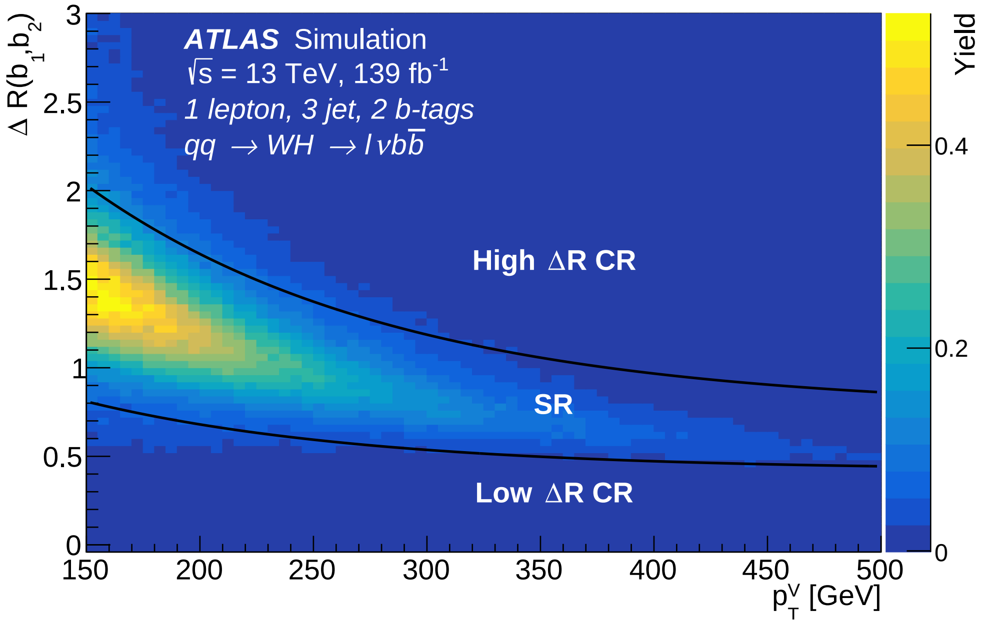
$$\Delta R_{\text{2-jet category}}(b_1, b_2) < 0.40 + e^{0.788 - 0.01023 \cdot p_T^V}$$
$$\Delta R_{\text{3-jet category}}(b_1, b_2) < 0.42 + e^{0.268 - 0.00809 \cdot p_T^V} \tag{4.7}$$

In Figure 4.19, the pre-fit simulated background composition is reported for the 0-Lepton analysis regions.

In addition to the CRs and SRs previously described, in the 2-Lepton channel, six very pure control regions ($e\mu$-CRs) are defined. In the $e\mu$-CRs, the same-flavour lepton required by the 2-Lepton event section is replaced by requiring exactly one electron and one muon. The $e\mu$-CRs are defined for each of the 2 jet categories and 3 p_T^V categories in the 2-Lepton channel. These control regions are more than 99% pure in $t\bar{t}$ and single-top Wt events. The ratio between the simulated and the measured data yield agrees within 1% in all the $e\mu$-CRs. Therefore, shapes and normalisations are extrapolated from data in the $e\mu$-CRs to the analysis SRs and CRs. The advantage of this method is that the only modelling uncertainty associated with the extrapolation is the data statistics in the $e\mu$-CRs.



(a) 2-jet.



(b) 3-jet.

Figure 4.18: Bi-dimensional $\Delta R(b_1, b_2)$ distribution as a function of p_T^V for 2-jet (a) and 3-jet (b) categories in the 1-Lepton channel. The black lines separate the signal and control regions defined in Equation 4.6 and Equation 4.7 [4].

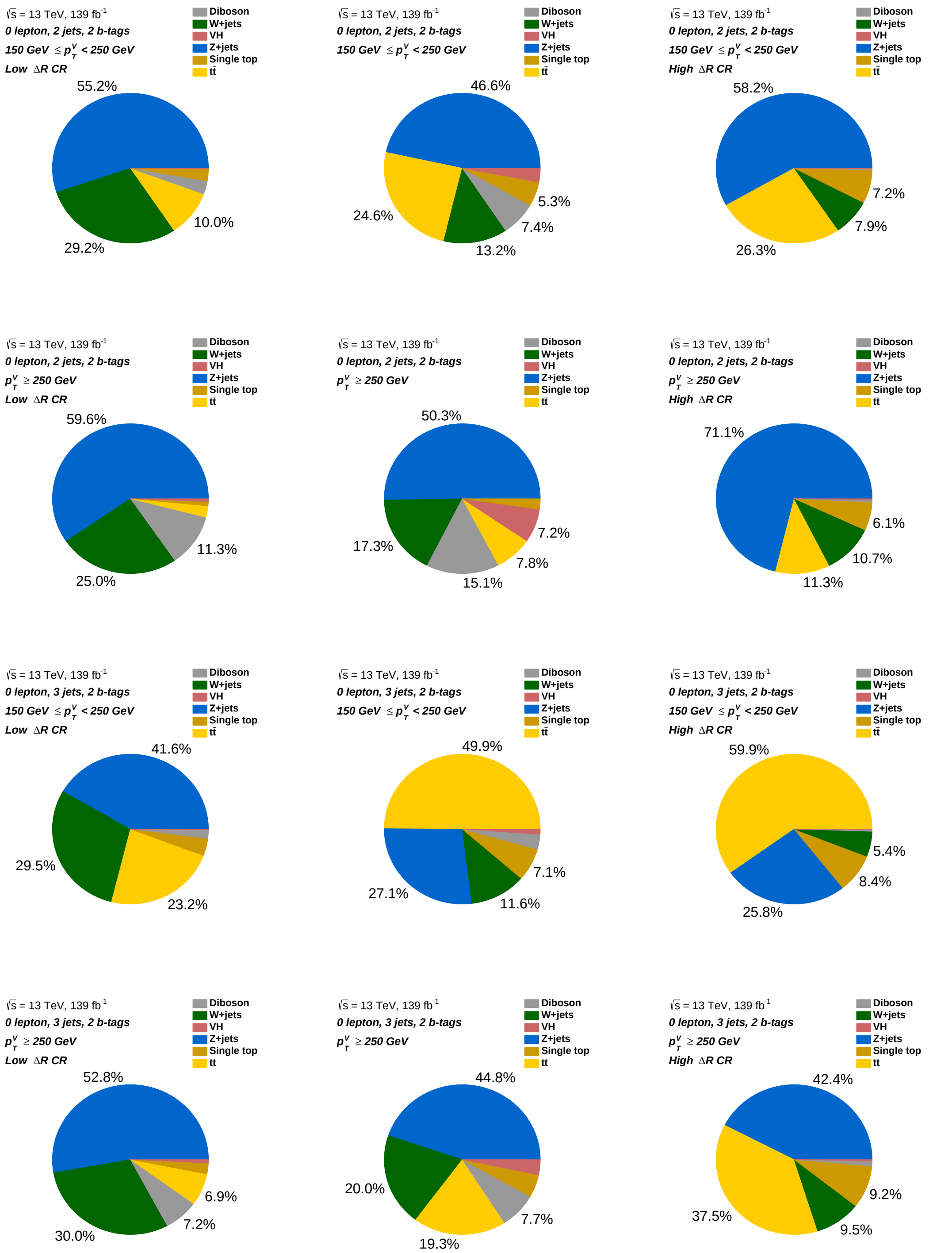


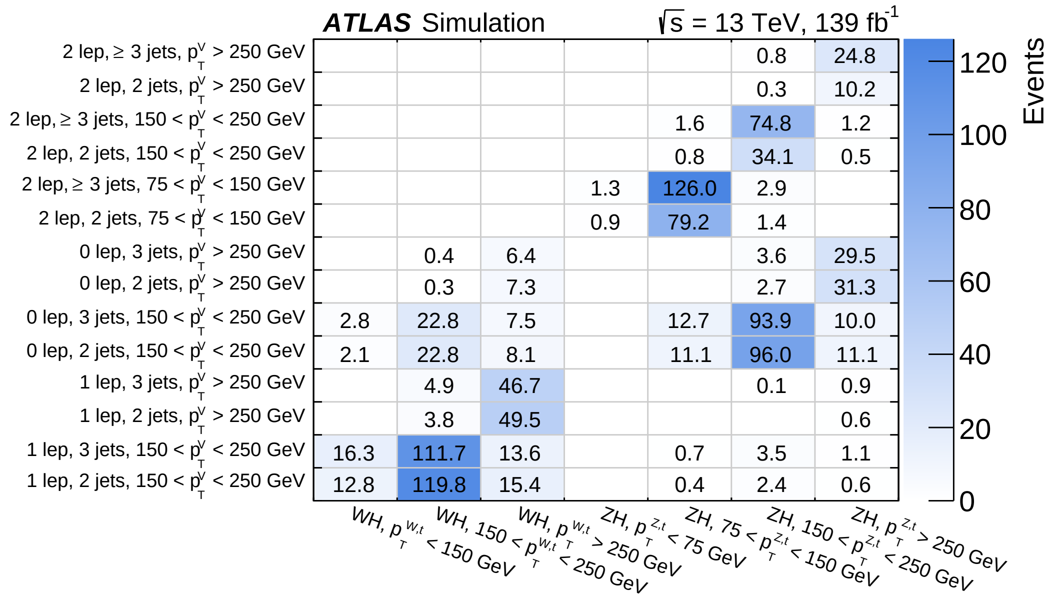
Figure 4.19: Pre-fit 0-Lepton background composition in signal and control regions. From left to right: Low-ΔR-CR, SR and High-ΔR-CR. From top to bottom: 2-jet 150 GeV $< E_T^{miss} <$ 250 GeV category, 2-jet 250 GeV $< E_T^{miss}$ category, 3-jet 150 GeV $< E_T^{miss} <$ 250 GeV category and 3-jet 250 GeV $< E_T^{miss}$ category. Values below 5% are not reported on the pie charts.

4.7 Simplified Template Cross Sections

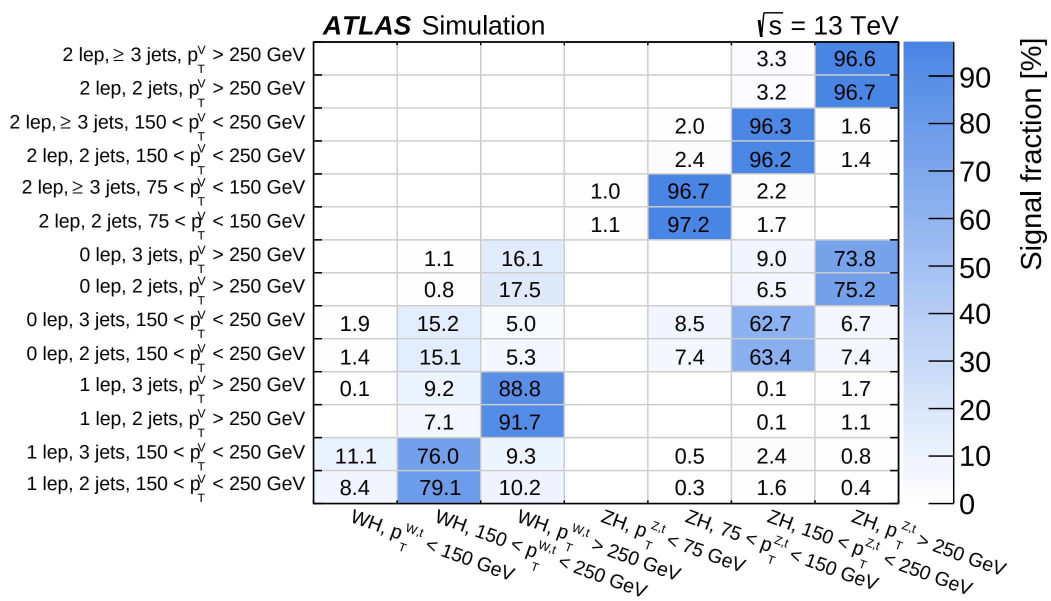
One of the main goals of the analysis is to perform cross-sections measurements in five *Simplified Template Cross Sections* (STXS) bins.

STXSs are fiducial cross-sections defined in simplified fiducial regions [109]. A *fiducial region* is a portion of the truth-level phase-space, which is populated by particles with a relatively long lifetime $> \mathcal{O}(10^{-8})$ s. At the detector-level, particles are reconstructed in physics objects with a certain efficiency within the detector acceptance. The aim of a *fiducial measurement* is to define a selection as close as possible between the truth-level and the detector-level in order to minimize the extrapolation between the two. The simple cuts that are applied in fiducial measurements cannot be integrated easily with a MVA technique. In fact, MVAs, which increase most of the current analysis sensitivities, are not easily reproducible at the truth level. A compromise between performing a fiducial measurement and using advanced multivariate techniques can be reached by introducing the STXS framework [110, 111]. Rather than considering fully differential distributions, such as p_{T}^{V}, in the STXS framework the cross-sections are divided into sub-cross sections, known as *bins*. Moreover, the Branching Ratio is factorised out of the measurement to reduce further theory dependent factors originating from combining measurements.

For the $VH, H \rightarrow b\bar{b}$ analysis, the chosen STXS bins are shown in Table 4.5. The signal distributions for each STXS bins have been estimated using generator truth information, denoted with a superscript t. With this notation p_{T}^{V} is the traverse momentum of the vector boson reconstructed from the physics objects measured by ATLAS, while $p_{\mathrm{T}}^{V,\,\mathrm{t}}$ is the truth-level boson transverse momentum. As illustrated in Table 4.5, for each category at truth-level there is a corresponding category at the reconstructed level. This specular definition of truth-level and reconstructed categories allows the correlation among the STXS bins to be reduced, as it can be seen from Figure 4.20(a) and Figure 4.20(b). Lastly, since the $VH, H \rightarrow b\bar{b}$ analysis has very low sensitivity below $p_{\mathrm{T}}^{V} < 150$ GeV for WH production and $p_{\mathrm{T}}^{V} < 75$ GeV for ZH production, there are no corresponding STXS bins for those regions.



(a) Signal event yields.



(b) Signal acceptances.

Figure 4.20: Signal event yields (a) and signal acceptances (b) for STXS and reconstructed categories [4]. Only yield contributions above 0.1 and acceptance contributions above 0.1% are reported.

STXS region		Corresponding reconstructed analysis regions		
Process	$p_{\mathrm{T}}^{V,\,\mathrm{t}}$ interval	Number of charged leptons	p_{T}^{V} interval	Number of jets
WH	150–250 GeV	1	150–250 GeV	2, 3
WH	$>$ 250 GeV	1	$>$ 250 GeV	2, 3
ZH	75–150 GeV	2	75–150 GeV	2, $\geq$ 3
ZH	150–250 GeV	0	150–250 GeV	2, 3
		2	150–250 GeV	2, $\geq$3
ZH	$>$ 250 GeV	0	$>$ 250 GeV	2, 3
		2	$>$ 250 GeV	2, $\geq$3

Table 4.5: The simplified template cross-section regions and the corresponding reconstructed analysis regions [4].

4.8 Improvements on invariant mass resolution

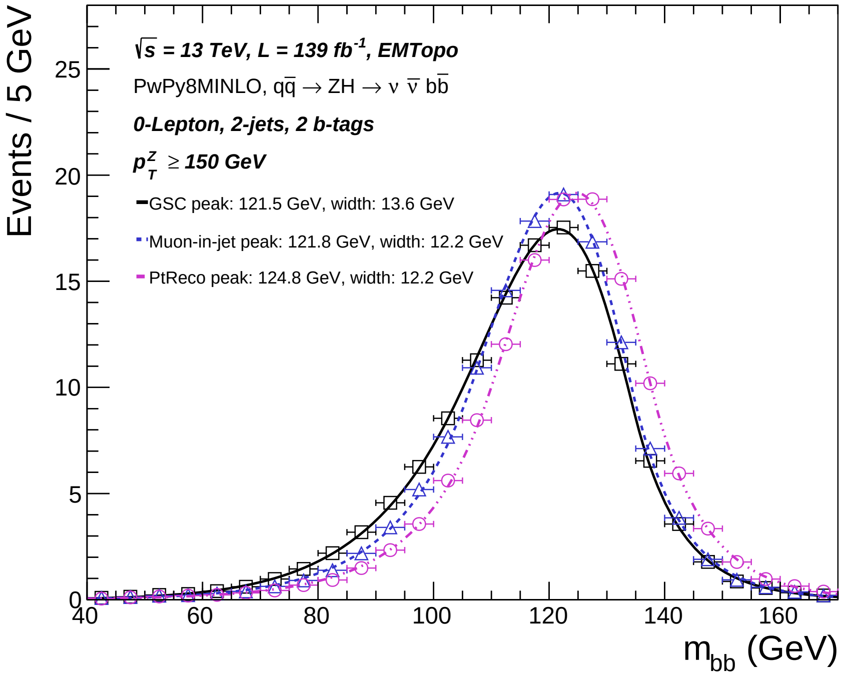
The invariant mass of the two b-jets is the most important variable to discriminate between the $VH, H \to b\bar{b}$ signal and the background processes. Several corrections can be applied to the m_{bb}, which exploit either the particular nature of b-jets or unique features of the $VH, H \to b\bar{b}$ topology. The m_{bb} corrections are:

- *Muon-in-jet correction*. B-hadrons during their decay chain can produce a muon. In fact, 12% of the b-jets passing the event selection contain at least one muon inside their jet cone. In the calorimeters, a muon deposits only a few GeV of energy loss. To recover the correct jet energy response, the muon deposited energy in the calorimeters is removed and the muon reconstructed four-momentum is added to the b-jet. This correction is applied only to muons with $p_{\mathrm{T}}^{\mu} > 4$ GeV and $|\eta^{\mu}| < 2.7$. Moreover, the muon has to be within $\Delta R(\mu, b - \mathrm{jet}) < min(0.4, 0.04 + 10/p_{\mathrm{T}}^{\mu})$ from the b-jet to better reject prompt muons. If more than one muon is found in the jet cone, the closest one in ΔR to the jet is chosen. Systematic uncertainties on the momentum correction and efficiency scale factors were found to be negligible. This correction is applied to all the three lepton channels.

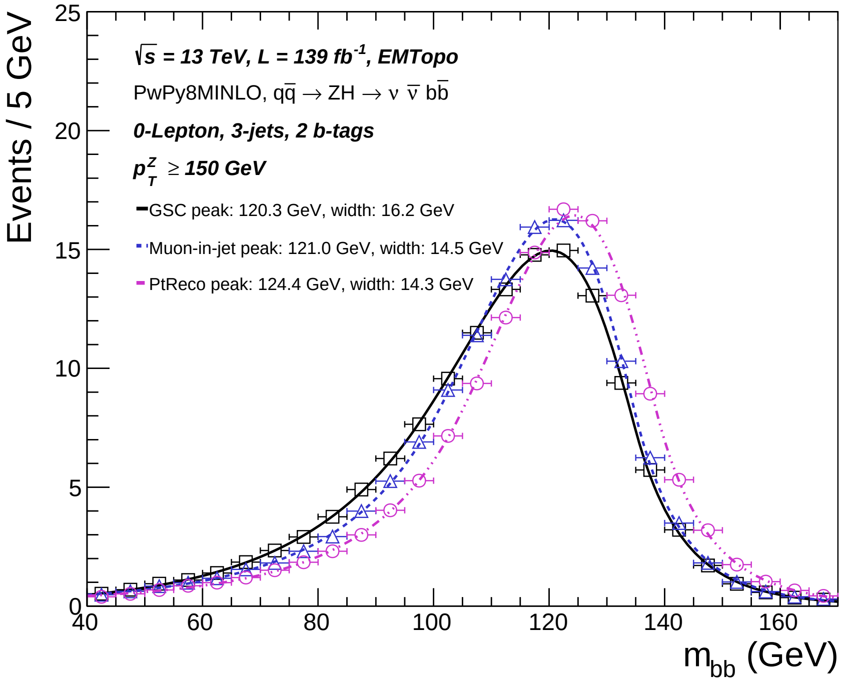
- *PtReco correction*. After the muon-in-jet correction, the PtReco correction is applied. This correction aims at recovering the energy loss associated with possible neutrinos inside the *b*-jet and particles escaping the reconstruction cone of the *b*-jet. Correction scale factors were calculated in simulated $ZH \rightarrow l^+l^-b\bar{b}$ events by taking the ratio of the p_{T} distributions of truth jets and the corresponding reconstructed ones. The correction scale factors were computed separately for events in which the *b*-jet decays semi-leptonically or hadronically. The correction factors are derived in the nominal analysis and propagated to all systematic variations. For this reason, no additional uncertainty is applied on top of the standard jet uncertainties. This correction is applied only to the 0-Lepton and 1-Lepton channels.

- *Kinematic Fit*. In the 2-Lepton channel, after the muon-in-jet correction, a Kinematic Fit is performed. In this channel, $E_{\mathrm{T}}^{\mathrm{miss}}$ originates only from the inherent transverse momentum of the incoming partons and from the neutrinos in the semi-leptonic decays of the two *b*-jets. Except for these two effects, the $ZH \rightarrow l^+l^-b\bar{b}$ events should be perfectly balanced. Moreover, the energy resolution of leptons is significantly better than that of *b*-jets. These properties allow a fit to be performed in which the energies of the particles are left floating within the respective resolutions, calibrating the energy of the *b*-jets using the precisely measured leptons.

The *b*-jet energy corrections for the 0-Lepton channel can be seen in Figure 4.21.



(a) 2-jet.



(b) 3-jet.

Figure 4.21: Comparison of the m_{bb} distributions for the b-jet energy corrections in the 0-Lepton channel for 2-jet and 3-jet categories. The jet reconstruction used in the analysis is EMTopo (see Section 3.2.4). The distributions have been fitted with a Bukin function to account for the asymmetric shape of the peak [112].

4.9 Improvements on the available number of simulated events

Having a large number of simulated events in the analysis phase space is of primary importance for every physics analysis. In this section, two techniques are discussed to increase the number of available events: simulated filtered samples (introduced in Section 4.4) and statistical b-tagging.

4.9.1 Filtered $t\bar{t}$ truth $E_{\mathrm{T}}^{\mathrm{miss}}$ samples

The $t\bar{t}$ production is one of the largest backgrounds for the $VH, H \rightarrow b\bar{b}$ analysis. To properly describe this background, it is necessary to have a sufficient number of generated Monte Carlo events. In the 0-Lepton channel, the $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV selection (Section 4.5) removes a large number of events. Therefore, generating events with high values of $E_{\mathrm{T}}^{\mathrm{miss}}$ will greatly increase the probability of events passing the final analysis selection.

Filtered $t\bar{t}$ truth $E_{\mathrm{T}}^{\mathrm{miss}}$ samples have first been introduced in the $VH, H \rightarrow b\bar{b}$ analysis with the 2015 - 2017 dataset [2]. Therefore, the plots and tables in this section have been done with $L = 79.8$ fb^{-1}, improving the global expected sensitivity of the analysis by 2% and contributing to reach the observation of $H \rightarrow b\bar{b}$ decays. For the Full Run 2 analysis with $L = 139$ fb^{-1}, the number of generated events of each sample has been increased by a factor ~ 2.

Table 4.6 summarises the additional samples (also referred as slices) generated and the filters applied. For the filtered samples, truth $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined at truth-level as the vectorial sum of all neutrinos originating from the hard-scatter process. The events have been generated with POWHEG-BOX V2 with the same cross-section but different filter efficiencies. All the events generated are either di-leptonic or semi-leptonic given that fully hadronic $t\bar{t}$ events do not enter in the 0-Lepton phase space (see Section 4.3.2).

Figure 4.22 shows the truth $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the four $E_{\mathrm{T}}^{\mathrm{miss}}$ filtered samples, compared to the inclusive $t\bar{t}$ sample in the 0-Lepton 2-jet and 3-jet categories. The top panel

Process	Truth selection	ϵ_{filter}	Simulated events
$t\bar{t}$	-	0.54386	235 M
$t\bar{t}$	$100 <$ truth $E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV	0.0952	22 M
$t\bar{t}$	$200 <$ truth $E_{\mathrm{T}}^{\mathrm{miss}} < 300$ GeV	0.0080	25.1 M
$t\bar{t}$	$300 <$ truth $E_{\mathrm{T}}^{\mathrm{miss}} < 400$ GeV	0.0011	6.6 M
$t\bar{t}$	truth $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV	0.0003	5 M

Table 4.6: $t\bar{t}$ $E_{\mathrm{T}}^{\mathrm{miss}}$ filtered samples used in the 0-Lepton channel during the 2015 - 2017 analysis [2]. For the full Run 2 analysis, the statistics of each sample has been increased by almost a factor ~ 2.

presents the comparison of the MC statistical uncertainty in each bin, while the second panel shows the ratio of the inclusive sample to the filtered ones.

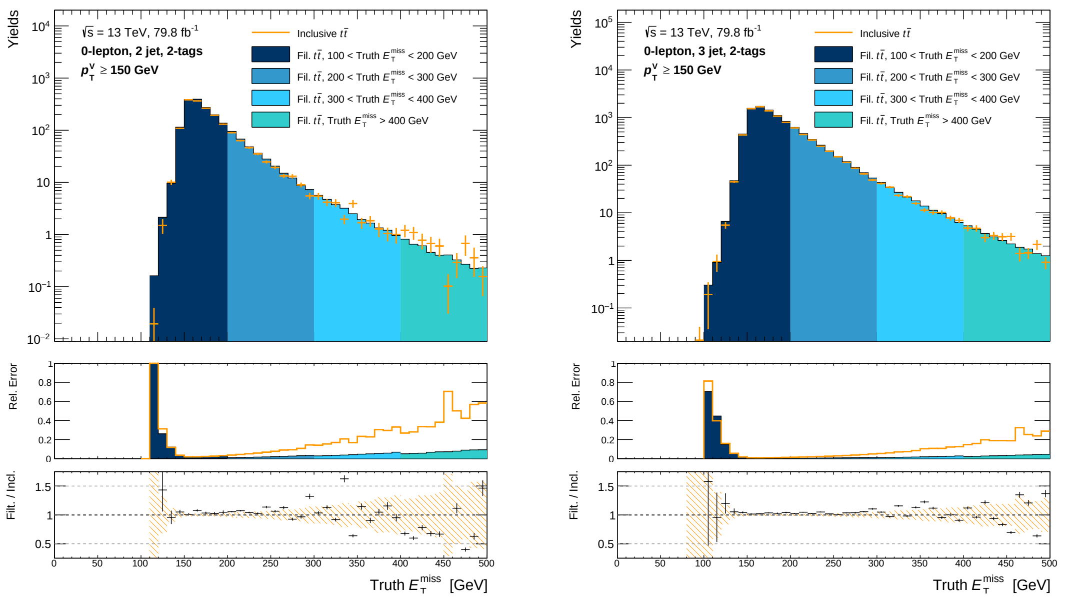


Figure 4.22: The truth $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the four $E_{\mathrm{T}}^{\mathrm{miss}}$ filtered samples, compared to the inclusive $t\bar{t}$ sample in the 0-Lepton 2-jet (left) and 3-jet (right) categories. The middle panel presents the comparison of the MC statistical uncertainty in each bin while the bottom panel shows the ratio of the inclusive to the filtered MC samples.

Due to the method by which the samples were generated, events from slices with truth $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV can in principle overlap with the inclusive sample, while for the 100 GeV $<$ truth $E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV slice, there is no overlap of events. Therefore only events in the filtered samples were included for truth $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV. The complete prescription to combine these samples is as follows:

- Select events only from the E_T^{miss} filtered samples if truth $E_T^{miss} > 200$ GeV.
- Select events from both the nominal and E_T^{miss} filtered samples but multiplying their respective MC weight by a factor 0.5 if 100 GeV $<$ truth $E_T^{miss} < 200$ GeV
- Select events from only the inclusive $t\bar{t}$ sample if truth $E_T^{miss} < 100$ GeV

The factor of 0.5 for the 100 GeV $<$ truth $E_T^{miss} < 200$ GeV slice was chosen because the inclusive sample and the corresponding filtered sample have comparable statistics in that region; more details are available in Appendix A.

Figure 4.23 presents the E_T^{miss} distribution for the combination of the four E_T^{miss} filtered samples and the inclusive sample, compared to the stand-alone $t\bar{t}$ sample by itself. As can be seen for $E_T^{miss} > 200$ GeV, the MC statistical uncertainty is reduced by a factor of 5. In addition, for the $100 < E_T^{miss} < 200$ GeV slice, the MC statistical uncertainty is reduced by a factor $\sqrt{2}$.

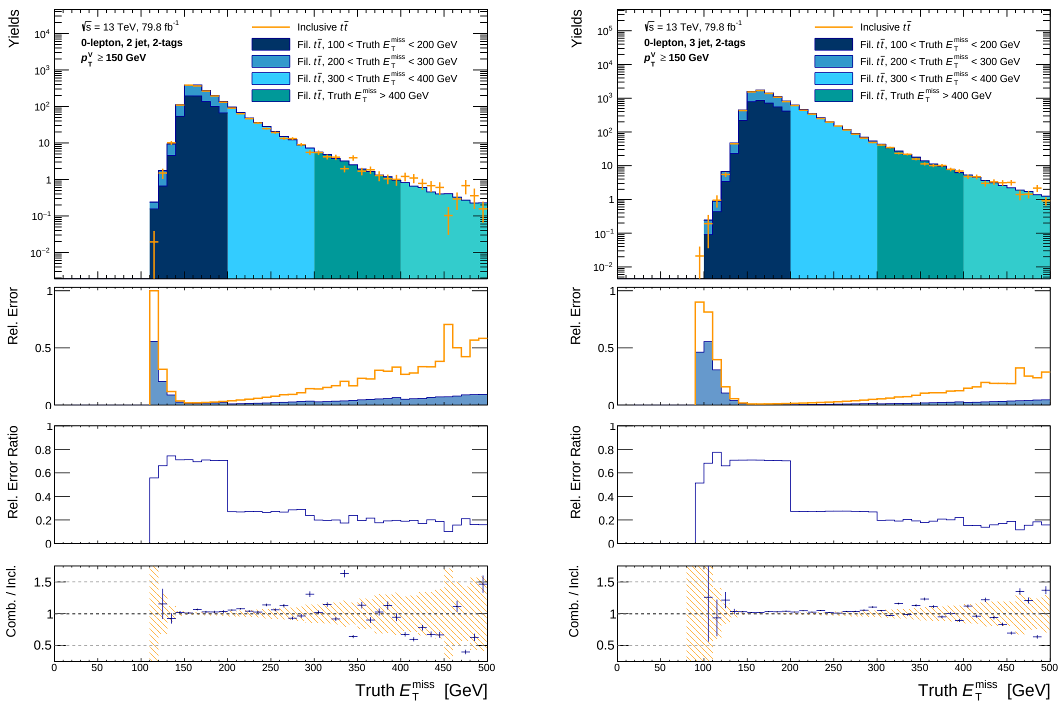


Figure 4.23: The truth E_T^{miss} distribution of the combination of the four E_T^{miss} filtered samples and the inclusive sample, compared to the inclusive $t\bar{t}$ sample in the 0-Lepton 2-jet (left) and 3-jet (right) categories. The second panel from the top presents the comparison of the MC statistical uncertainty in each bin, the third panel from the top presents the relative reduction in the MC statistical uncertainty and the bottom panel shows the ratio of the inclusive to the filtered MC samples.

The BDT$_{VH}$ discriminant is the fitted variable in the MVA measurement; details on the MVA analysis strategy will be provided in Section 4.10. Figure 4.24 presents the comparison of the 0-Lepton BDT$_{VH}$ discriminant distribution in the 2-jet and 3-jet signal regions, with and without the combination of the filtered $t\bar{t}$ truth $E_{\mathrm{T}}^{\mathrm{miss}}$ samples. Good agreement between the inclusive $t\bar{t}$ sample and the combination of the filtered samples is observed. The method achieves a 40% reduction in the MC statistical uncertainty across the entire BDT$_{VH}$ distribution. This increases to a 75% reduction in the statistical uncertainty in the most sensitive bin.

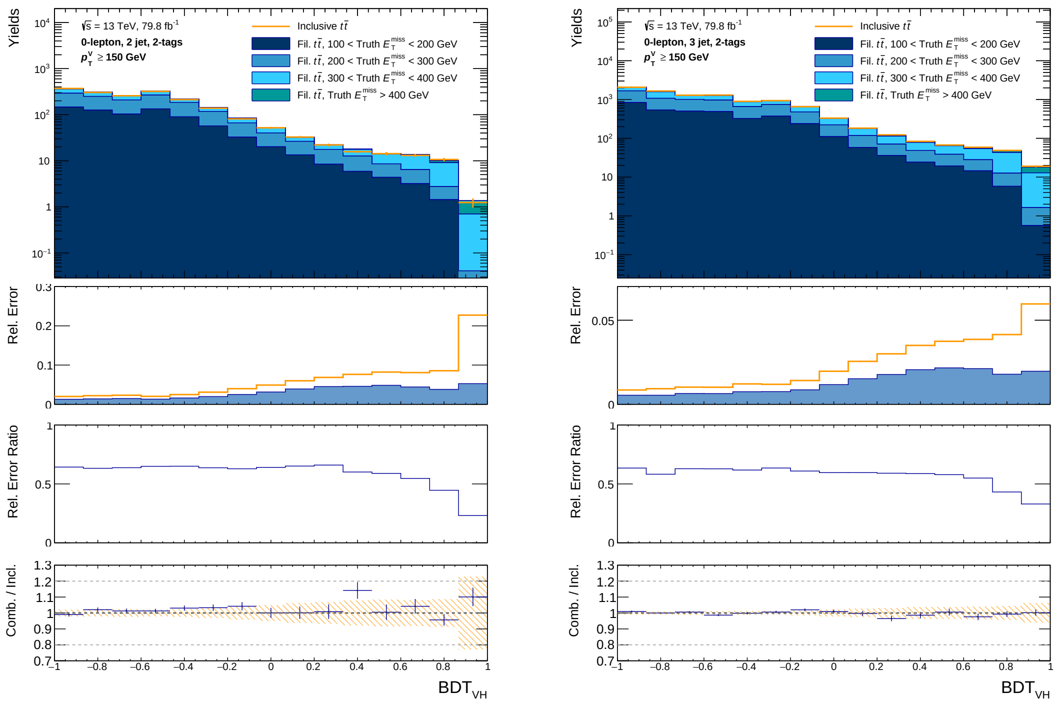
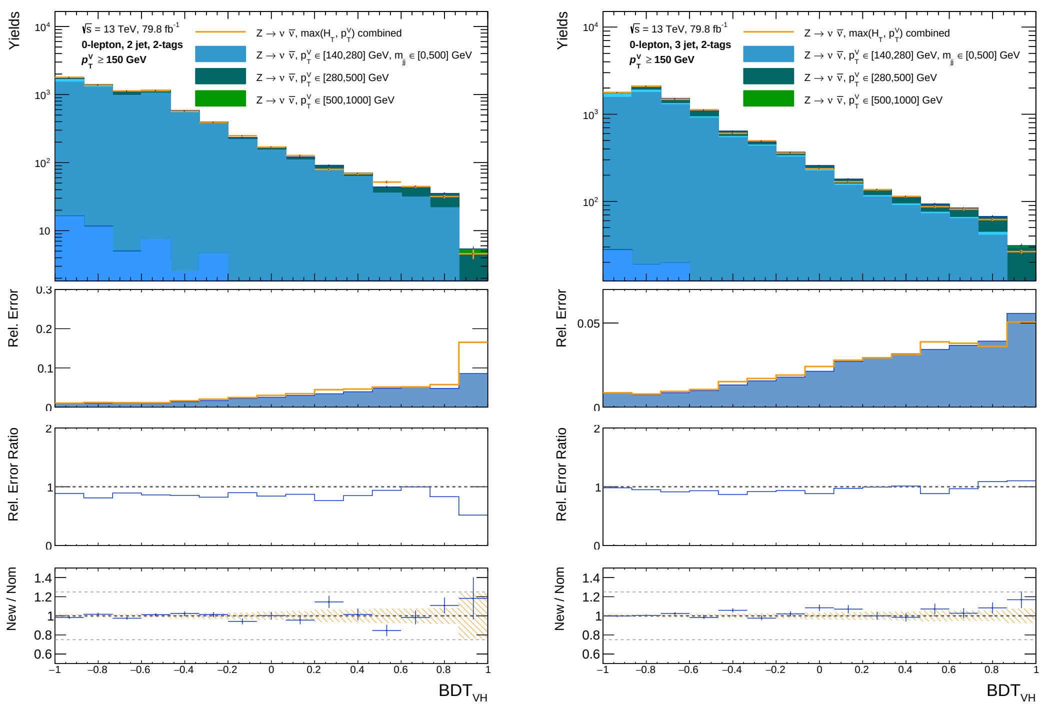


Figure 4.24: The 0-Lepton BDT_{VH} discriminant distribution of the combination of the four $E_{\mathrm{T}}^{\mathrm{miss}}$ filtered samples and the inclusive sample, compared to the inclusive $t\bar{t}$ sample by itself in the 0-Lepton (left) 2-jet and (right) 3-jet categories. The second panel from the top presents the comparison of the MC statistical uncertainty in each bin, the third panel from the top presents the relative reduction in the MC statistical uncertainty and the bottom panel shows the ratio of the inclusive to the filtered MC samples.

4.9.2 Filtered samples for Z+jets

A new filter being developed by a team studying $H \rightarrow invisible$ decays [113] for the $Z \rightarrow \nu\bar{\nu}$ background has been tested during the observation analysis [2]. This filter was charactered by being sliced in two quantities: p_{T}^{V} and m_{jj}. The former quantity would populate events with high $E_{\mathrm{T}}^{\mathrm{miss}}$ while the latter, defined as the invariant mass of the two largest p_{T} jets in the event, would populate events with high m_{jj}, which is characteristic of the VBF topology. In Table 4.7, the $Z \rightarrow \nu\bar{\nu}$ filtered samples are shown.

In Figure 4.25, the BDT_{VH} discriminant distribution for the $max(H_{\mathrm{T}}, p_{\mathrm{T}}^{V})$ combined samples and the p_{T}^{V}-m_{jj} samples are reported for the 0-Lepton 2-jet (left) and 3-jet categories (right). Given the similar statistical error of the two types of sample, it was decided to combine them, each weighted by a factor 0.5, in a similar way as the $100 <$ truth $E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV slice described in Section 4.9.1. The two slices that contribute most in the 0-Lepton phase space are: $140 < p_{\mathrm{T}}^{V} < 280$ GeV with $0 < m_{jj} < 500$ GeV and $280 < p_{\mathrm{T}}^{V} < 500$ GeV. Unfortunately, due to a delay, these p_{T}^{V} samples were not included in the observation analysis [2]. However, they have been included in the full Run 2 analysis.



(a) 2-jet.　　(b) 3-jet.

Figure 4.25: The 0-Lepton 2-jet (left) and 3-jet (right) BDT_{VH} discriminant distribution for the $max(H_{\mathrm{T}}, p_{\mathrm{T}}^{V})$ combined samples and the p_{T}^{V} samples. In the legend, only the samples contributing the most in the 0-Lepton phase space have been highlighted. The second panel from the top presents the comparison of the MC statistical uncertainty in each bin for the two types of sample, the third panel from the top reports the ratio of the relative errors and lastly the bottom panel shows the ratio of the $max(H_{\mathrm{T}}, p_{\mathrm{T}}^{V})$ to the filtered p_{T}^{V} samples.

Truth selection	σ [pb]	Simulated events
$70 < p_T^V < 100$ GeV	323.649	13.5 M
$100 < p_T^V < 140$ GeV, $0 < m_{jj} < 500$ GeV	111.267	9 M
$100 < p_T^V < 140$ GeV, $0.5 < m_{jj} < 1$ TeV	4.39739	585 k
$100 < p_T^V < 140$ GeV, $1 < m_{jj} < 13$ TeV	1.18657	180 k
$140 < p_T^V < 280$ GeV, $0 < m_{jj} < 500$ GeV	56.4852	6.75 M
$140 < p_T^V < 280$ GeV, $0.5 < m_{jj} < 1$ TeV	4.32468	720 k
$140 < p_T^V < 280$ GeV, $1 < m_{jj} < 13$ TeV	1.42901	247.5 k
$280 < p_T^V < 500$ GeV	4.41812	900 k
$0.5 < p_T^V < 1$ TeV	0.3044	500 k
$1 < p_T^V < 13$ TeV	$5.852 \cdot 10^{-3}$	200 k

Table 4.7: $Z \to \nu\bar{\nu}$ + jets additional samples filtered in p_T^V and m_{jj} relevant for the $VH, H \to b\bar{b}$ analysis.

4.9.3 Statistical b-tagging

The MV2c10 algorithm described in Section 3.2.4.3 is characterised by a high rejection rate for c-jets and l-jets. Nevertheless, a sizeable amount of events containing these flavours of jets enters the $VH, H \to b\bar{b}$ analysis phase space. By removing all the events in which the non-b-flavour jets do not meet the b-tagging value threshold (*direct-tagging*), the statistics of these samples would be highly depleted; this depletion creates irregular templates for physics processes with non-b-flavour jets, which suffer from statistical fluctuations.

An alternative approach to direct tagging is *truth-tagging* (also referred as *statistical b-tagging*): to each MC event it is possible to assign a weight based on the probability to have n b-tagged jets in the event, rather than removing the event completely when the b-tagging requirement is not met. In this way, all the events are kept and simply scaled down by a weight factor. As an example, for events with three jets, in which only two jets are required to be b-tagged, the probability weight is calculated as $w_{event} = \epsilon_1^{f_1}\epsilon_2^{f_2}(1-\epsilon_3^{f_3}) + \epsilon_1^{f_1}(1-\epsilon_2^{f_2})\epsilon_3^{f_3} + (1-\epsilon_1^{f_1})\epsilon_2^{f_2}\epsilon_3^{f_3}$, where ϵ_i is the b-tagging efficiency for a jet of a given flavour f_i. Although the truth-tagging approach increases the available statistics, it is characterised by a slightly worse agreement between data and simulation compared to direct-tagging, especially when b-jets are truth tagged. For this reason, a compromised approach has been chosen, called *hybrid-tagging*. In hybrid-tagging, b-jets are direct-tagged while non-b-jets are truth-tagged; the event weight is a product of the b-jet flavour tagging scale factors (in Figure 3.5) for the b-jets, and the truth-tagging weight for the non-b-jets.

4.10 Multivariate analysis strategy

Boosted Decision Trees (BDTs) have been trained to maximize the analysis' sensitivity [67]. BDTs are a multivariate technique commonly used in High Energy Physics to discriminate a signal process from its backgrounds [114]. A TMVA implementation of the Gradient Boosting algorithm has been chosen as training method for the BDTs [115].

In the main multivariate analysis, a set of BDTs has been trained in order to separate the VH signal from all the other processes. Similarly, in the diboson cross-check analysis, a second set of BDTs has been trained to distinguish the VZ processes from all the others, including the VH production. The trainings have been performed using all the simulated samples reported in Table 4.2. Moreover, separate BDTs have been trained for each lepton channel and jet category in the inclusive phase space. For each separate lepton channel and jet category, the inclusive phase space is defined by merging the different SRs, CRs and p_{T}^{V} bins (except for 75 GeV $< p_{\mathrm{T}}^{V} <$ 150 GeV in the 2-Lepton channel). The training dataset has been divided into two subsets (50% $-$ 50%): the first sub-set has been used for the actual training of the BDTs, while the second sub-set has been used to test for *over-training*. Over-training occurs when the machine learning model has learnt the training dataset in too much detail and is not able to provide good predictions on a new dataset. To check for over-training, the BDT output scores evaluated on the two sub-sets have been compared; since very similar distributions have been found between the training and testing sub-sets, it has been concluded that no over-training occurred. Similarly, also the Receiver Operating Characteristic (ROC) [116] curves for the training and testing samples have been compared to provide a second cross-check. Lastly, the training-testing procedure has been then repeated with inverted sub-sets to exploit the full available statistics.

Input variables (also known as features) for the BDT_{VH} and the BDT_{VZ} are reported in Table 4.8 for each lepton channel.

Variable	0-Lepton	1-Lepton	2-Lepton
m_{bb}	$\times$	$\times$	$\times$
$\Delta R(b_1, b_2)$	$\times$	$\times$	$\times$
$p_{\mathrm{T}}^{b_1}$	$\times$	$\times$	$\times$
$p_{\mathrm{T}}^{b_2}$	$\times$	$\times$	$\times$
p_{T}^{V}	$\equiv E_{\mathrm{T}}^{\mathrm{miss}}$	$\times$	$\times$
$\Delta\phi(V, bb)$	$\times$	$\times$	$\times$
MV2(b_1)	$\times$	$\times$	
MV2(b_2)	$\times$	$\times$	
$\|\Delta\eta(b_1, b_2)\|$	$\times$		
m_{eff}	$\times$		
$p_{\mathrm{T}}^{\mathrm{miss,st}}$	$\times$		
$E_{\mathrm{T}}^{\mathrm{miss}}$	$\times$	$\times$	
$\min[\Delta\phi(\ell, b)]$		$\times$	
m_{T}^{W}		$\times$	
$\|\Delta y(V, bb)\|$		$\times$	
m_{top}		$\times$	
$\|\Delta\eta(V, bb)\|$			$\times$
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{S_{\mathrm{T}}}$			$\times$
$m_{\ell\ell}$			$\times$
$\cos\theta(\ell^-, Z)$			$\times$
		Only in 3-jet events	
$p_{\mathrm{T}}^{\mathrm{jet}_3}$	$\times$	$\times$	$\times$
m_{bbj}	$\times$	$\times$	$\times$

Table 4.8: Variables used for the training and evaluation of the BDTs in each lepton channel.

The invariant mass of the di-jet system, the leading and sub-leading jet p_{T}, the angular separation in ΔR between the two b-jets, the transverse momentum of vector boson and the angular separation in $\Delta\phi$ between the di-jet system and the vector boson are the common variables used in all the three lepton channels. The p_{T} of the third jet and the invariant mass of the three-jet system are used in all the lepton channels only in the 3-jet category. In the 0-Lepton and 1-Lepton channels also the binned b-tagging scores (quantiles) for the b-jets have been used to discriminate against the $t\bar{t}$ background. In 0-Lepton channel, also to reject the $t\bar{t}$ background, the magnitude of the track-based missing transverse energy, $p_{\mathrm{T}}^{\mathrm{miss,st}}$, has been added (details in Section 4.10.1). The other variables specific to the 0-Lepton channel are the m_{eff}, which is the scalar sum of the p_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event, and $|\Delta\eta(b_1, b_2)|$. In the 1-Lepton channel, m_{top} aims at reconstructing the mass of the leptonically decaying top quark, while m_{T}^{W} is the transverse mass of the W boson candidate; min$[\Delta\phi(\ell, b)]$ and $|\Delta y(V, bb)|$ also provide good discrimination against $t\bar{t}$. In the 2-Lepton channel, most of the variables are selected to discriminate against $Z + jets$, except for $E_{\mathrm{T}}^{\mathrm{miss}} / \sqrt{S_{\mathrm{T}}}$, which is effective at suppressing the di-leptonic $t\bar{t}$ decays, where S_{T} is the scalar sum of transverse momenta of the physic objects in the event.

The BDTs output scores are fitted in histograms, whose binning has been optimised in order to maximise the expected sensitivity while keeping the statistical error in each bin below a certain threshold. Details on the procedure can be found in Reference [2]. For the main VH analysis, the SR histograms in regions with $p_{\mathrm{T}}^{V} < 250$ GeV have 15 bins, while SR histograms in regions with $p_{\mathrm{T}}^{V} > 250$ GeV have 8 bins due to the smaller number of events in these regions.

4.10.1 New input features: *b*-tagging scores and $p_{\mathrm{T}}^{\mathrm{miss,st}}$

Three new variables have been added to the 0-Lepton channel for the full Run 2 $VH, H \to b\bar{b}$ analysis with respect to the observation analysis [2]: the two *b*-tagging quantile scores of the leading and sub-leading *b*-jets (MV2) and the soft $E_{\mathrm{T}}^{\mathrm{miss}}$ ($p_{\mathrm{T}}^{\mathrm{miss,st}}$). The main target for both sets of variables is the $t\bar{t}$ background. In order to assess the potential of each variable, several metrics have been carefully evaluated:

1. *Shapes*. The normalised variable shapes for the signal VH have been compared with the ones of backgrounds to assess the discriminating potential of each variable.

2. *Correlations*. The linear correlation, $\rho_{x,y} = cov(x, y)/(\sigma_x \sigma_y)$ between each training variable has been tested. A new variable highly correlated with one already present might indicate that the discriminating power of the new variable is already partially included in the old one. The correlation matrices for the signal and main background processes for the 2-jet category in the 0-Lepton channel are reported in Appendix D.

3. *MVA retraining and evaluation of the ROC*. An XGBoost [117] based implementation of the MVA discriminant has been used to test the new variables, in parallel to the implementation in TMVA. Details on the XGBoost implementation is provided in Section 5.3. For the retraining, each new candidate variable has been added, one at a time in the MVA training. Then, the Receiver Operating Characteristic curve area has been used as metric to assess the effectivness of the new variable [116].

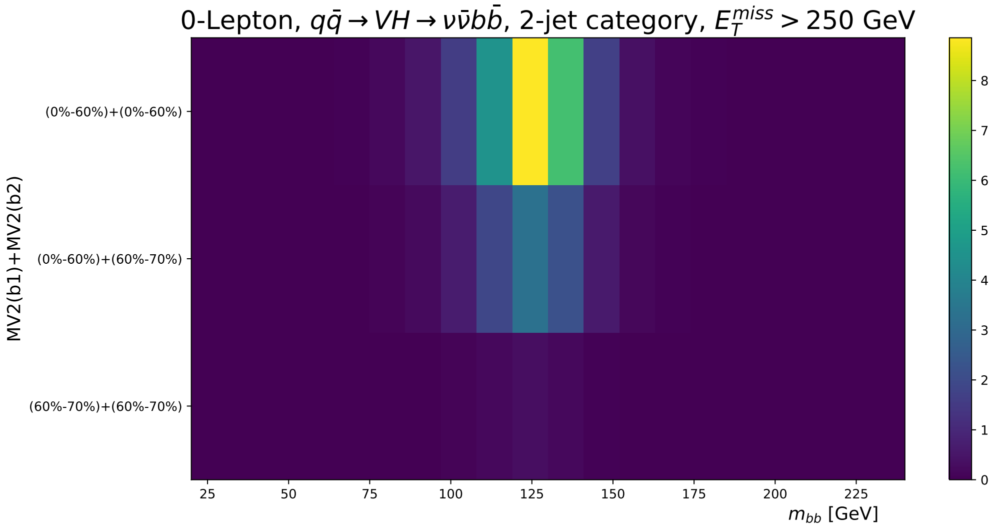
4. *Simplified significance*. A simplified formula [118] can be used to estimate the analysis significance without performing the full fit:

$$S = \sqrt{\sum_i [2(s_i + b_i)\ln(1 + \frac{s_i}{b_i}) - s_i]} \tag{4.8}$$

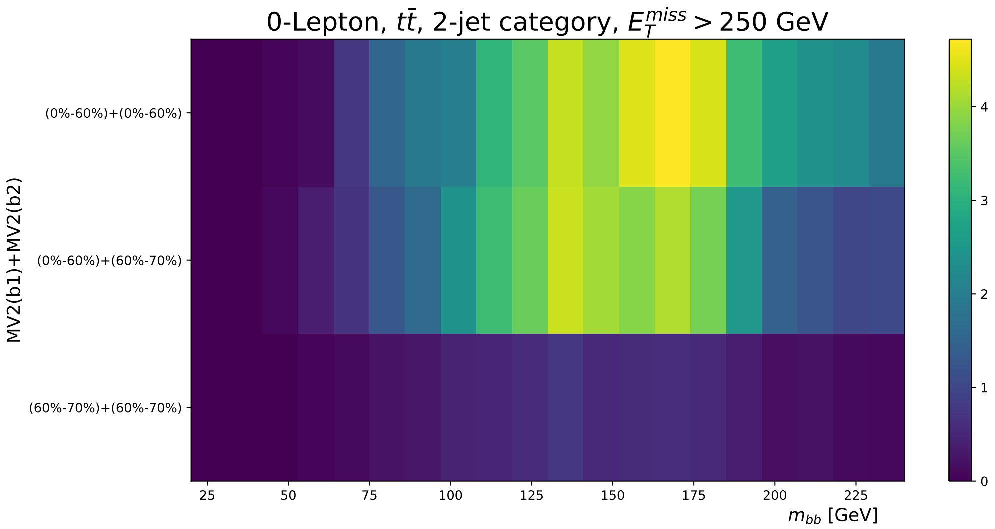
   where the summations run over the MVA bins for signal and background events.

5. *Significance*. The analysis significance has also been determined using the full Asimov fit, described in Section 4.11.

MV2(b_1) and MV2(b_2) quantiles are very effective at recognizing the $t\bar{t} - bc$ background in the high p_T^V regions. Since these variables are assigned only to b-jets that have already been tagged at the 70% working point, they can assume only two values: one for a b-tagging score in the range $(60\% - 70\%)$ and another one for $(0\% - 60\%)$. The sum of MV2(b_1) and MV2(b_2) for the signal and $t\bar{t}$ 2-jets events in the $p_T^V > 250$ GeV region is shown in Figure 4.26. The comparison clearly indicates that a substantial fraction of $t\bar{t}$ events is not as pure in bb content as the signal. The inclusion of MV2(b_1) and MV2(b_2) in the MVA improves the significance of Asimov fit by $\sim 7\%$.



(a) $q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}$.



(b) $t\bar{t}$.

Figure 4.26: Sum of MV2(b_1) and MV2(b_2) quantiles for 0-Lepton 2-jets events for $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV. The top plot shows the main 0-Lepton signal, while the bottom plot shows the $t\bar{t}$ background.

Another variable that was found to improve the sensitivity is the track-based soft missing energy $p_{\mathrm{T}}^{\mathrm{miss,st}}$, described in Section 3.2.5. This variable evaluates the amount of non-reconstructed objects in the event, which is particularly useful especially for recognizing $t\bar{t}$ events that are characterised by mis-reconstructed leptons and jets. The shape of $p_{\mathrm{T}}^{\mathrm{miss,st}}$ is displayed in Figure 4.27; as it can be noticed, the shapes of the signal and $t\bar{t}$ are different. Moreover, $p_{\mathrm{T}}^{\mathrm{miss,st}}$ is not particularly correlated with the other variables, such as $E_{\mathrm{T}}^{\mathrm{miss}}$, as it can be seen in Appendix D.

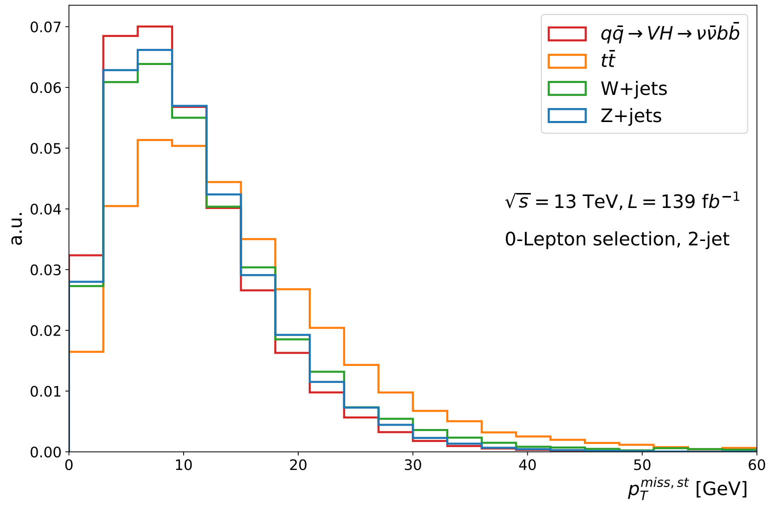


Figure 4.27: Normalised shapes of $p_{\mathrm{T}}^{\mathrm{miss,st}}$ for the main signal and background processes in the 0-Lepton channel 2-jet category.

One important concern for adding this variable to the MVA was its modelling in different MC generators due the difficulty of accurately modelling mis-identification of leptons and jets. The comparison of $p_{\mathrm{T}}^{\mathrm{miss,st}}$ for the nominal and an alternative MC (POWHEG+HERWIG7) generator in $t\bar{t}$ events shows a good agreement as noted in Figure 4.28. Adding this variable to the MVA improves the Asimov significance by $\sim 2\%$.

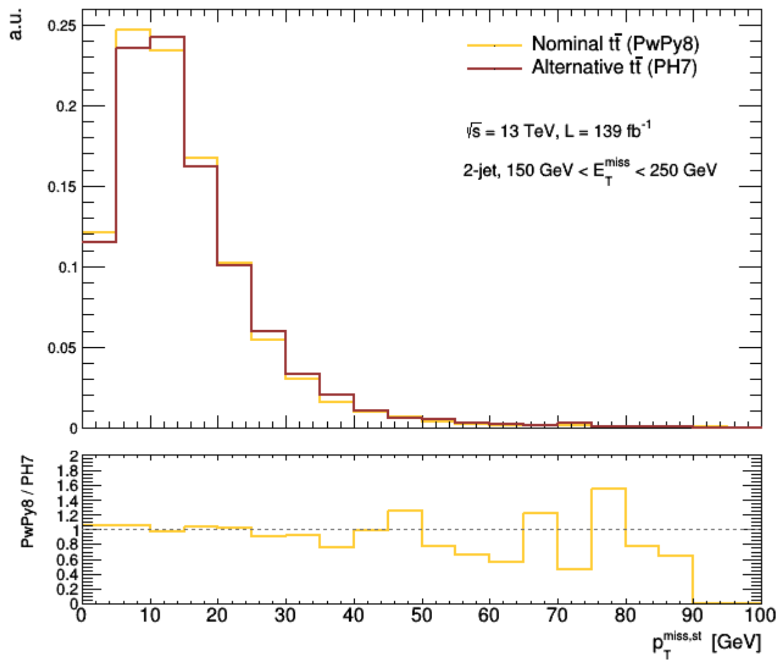


Figure 4.28: Comparison of $p_{\mathrm{T}}^{\mathrm{miss,st}}$ for the nominal and alternative MC generators in $t\bar{t}$ 2-jet events with 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.

4.11 Statistical analysis

A simultaneous fit procedure is used in the $VH, H \to b\bar{b}$ analysis to measure the parameters of interest (POI) by maximising a global likelihood function [118]. The POI are of two types: signal strengths μ and cross-section times branching ratio.

4.11.1 The likelihood function

The *likelihood function* $\mathcal{L}$ assesses the goodness of fit between two datasets: the first one composed by the measured data and the second one obtained by MC simulations for a given hypothesis to test. For this analysis, the hypotheses to test are the POI with respect to the values predicted by the Standard Model. The signal strength is defined as

$$\mu = \frac{\sigma \cdot BR}{\sigma_{SM} \cdot BR_{SM}} \tag{4.9}$$

assuming a Poisson probability distribution in each bin. The likelihood for the signal strength can be written as

$$\mathcal{L}(\mu)_{Poisson} = \prod_{j}^{reg} \prod_{i}^{bins} \frac{(\mu \cdot s_{ij} + b_{ij})^{n_{ij}}}{n_{ij}!} e^{-(\mu \cdot s_{ij} + b_{ij})} \tag{4.10}$$

where n_{ij} are the measured number of data events for a given bin i in a region j, while s_{ij} and b_{ij} are the predicted numbers of signal and backgrounds events for the same bin. The definition of the likelihood in Equation 4.11 covers only statistical fluctuations originating from the statistical nature of the measurement. To include uncertainties originating from the limited resolution of the experimental apparatus or theoretical prediction, the likelihood can be extended by adding additional degrees of freedom in the form of *nuisance parameters*. The description of the systematic uncertainties entering the analysis will be provided in Section 4.12. The set of nuisance parameters $\boldsymbol{\theta}$ enters in the likelihood as Gaussian multiplicative terms

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \mathcal{L}(\mu)_{Poisson} \cdot \prod_{k}^{N_{sys}} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(\bar{\theta_k} - \theta_k)^2}{2\sigma_k^2}} \tag{4.11}$$

where $\bar{\theta_k}$ is defined to be equal to $\bar{\theta_k} = 1$ for normalisation uncertainties and to $\bar{\theta_k} = 0$ for the others, and σ_k^2 represents the uncertainty associated with the nuisance parameter θ_k. Lastly, one additional factor is included in the likelihood to account for a finite amount of simulated events, according to the Barlow and Beeston approach [119], where a new nuisance parameter per bin of the fitted distributions provides an additional degree of freedom for the total MC prediction within its statistical uncertainty.

For the STXS measurement, the signal strength in Equation 4.9 and the likelihood in 4.11 can be extended to the cross-section times branching ratio POI in the following way:

$$\mu_{POI,l} = \frac{\sigma_l \cdot BR(H \to b\bar{b}) \cdot BR(V \to lep)}{\sigma_{SM,l} \cdot BR_{SM}(H \to b\bar{b}) \cdot BR_{SM}(V \to lep)} \tag{4.12}$$

$$\mathcal{L}(\mu)_{Poisson} = \prod_j^{reg} \prod_i^{bins} \frac{(\sum_l^{STXS} \mu_{POI,l} \cdot s_{ijl} + b_{ij})^{n_{ij}}}{n_{ij}!} e^{-(\sum_l^{STXS} \mu_{POI,l} \cdot s_{ijl} + b_{ij})} \tag{4.13}$$

where $STXS$ are the five STXS bins.

4.11.2 The likelihood ratio

From the likelihood function, it is possible to define a statistical test to quantify the compatibility between two hypotheses:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{4.14}$$

known as *likelihood ratio*. In Equation 4.14, $\hat{\hat{\boldsymbol{\theta}}}$ represents the value of the set of nuisance parameters $\boldsymbol{\theta}$ that maximises the likelihood for a given value of μ, whereas $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ represent the values that maximise the likelihood when both $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are allowed to vary simultaneously. A small value of $\lambda(\mu)$ indicates a bad agreement between the measured data and the initial hypothesis for μ. Discovering a signal can be thought of as rejecting the background-only hypothesis; with this purpose, the test statistics q_0

can be defined

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{4.15}$$

The condition in the second line of Equation 4.15 creates a statistical test that considers a signal process that does not interfere destructively with the background processes [118]. To quantify the rejection of the hypothesis of only background, it is useful to define the *p-value* as

$$p_{0,obs} = \int_{q_{0,obs}}^{\infty} f(q_0|\mu = 0)dq_0 \tag{4.16}$$

where $q_{0,obs}$ is the measured test statistics in data and $f(q_0|\mu = 0)$ is the probability distribution function of the test statistics, which can be approximated analytically thanks to the Wald approximation [118]. In conclusion, the p-value quantifies the probability of obtaining the measured results assuming only a fluctuation of the backgrounds. A common way to quote the p-value is in terms of the standard deviations of a standard Gaussian distribution using its inverse cumulative distribution, Φ^{-1}:

$$Z = \Phi^{-1}(p_{0,obs} - 1) \tag{4.17}$$

In the particle physics community, $Z = 3$ is considered the threshold for the *evidence* of a new phenomenon, $Z = 4$ is regarded as *strong evidence*, while $Z = 5$ is required to claim the *observation*.

The uncertainties $\hat{\sigma}$ associated with the observed $\hat{\mu}$ are the values that satisfy $\ln(\mathcal{L}(\hat{\mu} \pm \hat{\sigma})) \approx \ln(\mathcal{L}_{max}) - 1/2$, which is obtained by Taylor-expanding the log-likelihood around its maximum. In order to evaluate the contribution of the nuisance parameters to the measured μ, two methods are used: *the breakdown method* and *the ranking method*.

In the breakdown method, one or more systematic variations belonging to a common group (e.g. jet systematics) are removed from the fit model. Then, the likelihood is re-maximised and the error on μ is recalculated. Lastly, the error contribution of the given systematic error is calculated by subtracting quadratically the total error estimated without the systematic from the total error with all the systematics. This procedure is repeated for all the groups of systematics.

In the ranking method, a group of systematics are fixed to the upward and downward variations corresponding to $\sigma = \pm 1$. Then, the likelihood is re-maximised and the shift $\Delta\mu^{\pm}_{sys} = \mu_{\pm 1\sigma} - \hat{\mu}$ calculated. The shifts are calculated for all the systematics and then ranked from the largest to the smallest.

Before performing the unblinded fit to measure μ, at least two other types of fit are performed: the *Asimov fit* and the *conditional fit to data*. In the Asimov fit, an Asimov dataset is created such that the estimators extracted by the fit correspond by construction to the true parameters [118]. This type of fit allows correlations between nuisance parameters to be evaluated and provides an estimate of the expected sensitivity. Conversely, in a conditional fit to data, μ is fixed to $\mu = 1$ and it is possible to evaluate how the fit model matches the data in bins where few signal events, with respect to background events, are expected.

4.12 Systematic uncertainties

The systematic uncertainties impacting the $VH, H \to b\bar{b}$ analysis can be grouped into two main categories: experimental uncertainties and modelling uncertainties.

4.12.1 Experimental uncertainties

Experimental systematic uncertainties are associated with physics objects and beam conditions measurements. Standardised prescriptions are provided to physics analyses by dedicated working groups based on independent measurements. The experimental systematic uncertainties affecting the $VH, H \to b\bar{b}$ analysis are reported below.

- *Leptons*. The leptons experimental systematic uncertainties take into account the lepton energy / momentum scale and resolution. There are a series of uncertainties covering the efficiencies of lepton trigger, reconstruction, identification, and isolation [65, 66].

- *Jets*. The two main sources of uncertainties affecting jets are the Jet Energy Scale [72] and the Jet Energy Resolution [120]; the former is characterised by 30 independent components while the latter by 8 independent components. A specific uncertainty on the efficiency of the JVT is included.

- *B-tagging*. *B*-tagging scale-factor uncertainties (described in Section 3.2.4.3) are decomposed into uncorrelated components, which are evaluated independently. Eigenvectors [76] causing variations smaller than 1% on the *b*-tagging scale factors in all the *b*-tagging efficiency bins $(0\%, 60\%)$, $(60\%, 70\%)$ and $(70\%, 100\%)$ are removed. Moreover, eigenvectors causing variations bigger than 1% on the *b*-tagging scale factors on some efficiency bins but whose total contribution is less than 10% of the total uncertainty in the bins are also removed. After applying this pruning procedure, the number of independent systematic uncertainties are 29 for *b*-jets, 18 for *c*-jets and 10 for *l*-jets.

- $E_{\text{T}}^{\text{miss}}$. The $E_{\text{T}}^{\text{miss}}$ is recomputed during the evaluation of all the systematics. Furthermore, three additional systematic variations are associated with the trigger $E_{\text{T}}^{\text{miss}}$: the first one is related to the statistical error of the scale factors, the second systematic accounts for different scale factors measured in different physics processes and the last one accounts for a small kinematic dependence of $E_{\text{T}}^{\text{miss}}$ trigger efficiency on H_{T}. Lastly, specific systematics are included to cover resolution and efficiency uncertainties of the tracks used to build the $p_{\text{T}}^{\text{miss,st}}$.

- *Luminosity*. The total combined uncertainty on the integrated luminosity for the Run 2 is 1.7% [36].

- *Pile-up*. Scale-factor variations are introduced in the simulations to cover the data pile-up profiles across the years of data-taking [121].

4.12.2 Modelling uncertainties

MC simulations are one of the cornerstones on which the $VH, H \to b\bar{b}$ analysis is built. However, due to the physics assumptions entering the simulations and limited MC statistics, an accurate estimation of modelling uncertainties is required to provide accurate measurements. For each physics process, the main MC generator is referred to as *nominal*. Alternative settings of the nominal generator, expressed as internal weights or different generators, are compared to the nominal generator in order to compute *modelling uncertainties*. When multiple settings/generators are available for the same variation, the one producing the largest effect is considered. For signal, diboson and single-top processes, given the limited number of fully reconstructed simulated events, the modelling uncertainties are calculated at truth-level, due to the larger available statistics, with a dedicated analysis framework called VHBBTRUTHFRAMEWORK. Instead, for V+jets and $t\bar{t}$ productions, the modelling uncertainties are calculated on the fully reconstructed objects since for these processes the number of fully reconstructed events are sufficient. In the analysis, three types of modelling uncertainties are considered:

- *Normalisation*. The total cross-section uncertainty of a given process is parametrised by its normalisation uncertainty. It is calculated as a ratio between the number of events in the nominal sample and in the alternative sample, as reported in Equation 4.18.

$$\sigma_{norm} = 1 - \frac{N_{alt}}{N_{nom}} \tag{4.18}$$

  This type of uncertainty covers the smallest background entering the $VH, H \to b\bar{b}$ analysis phase space. Conversely, the normalisation of the main backgrounds ($t\bar{t}$, Z+bb and W+bb) are left floating in the fit.

- *Acceptance*. Acceptance uncertainties account for the migration of events from a category A to B and vice-versa [122], as indicated in Equation 4.19.

$$\sigma_{acc} = 1 - \left(\frac{N^A_{alt}}{N^B_{alt}} \bigg/ \frac{N^A_{nom}}{N^B_{nom}} \right) \tag{4.19}$$

  Acceptance effects depend implicitly on the variables used to define the analysis categories such as p_{T}^V, $\Delta R(b_1, b_2)$ and N_{jets}.

- *Shape*. Shape uncertainties keep constant the number of events within an analysis region while varying the distribution of a chosen variable. In principle, the uncertainty originating from the comparison of the nominal and alternative generator in a given distribution for a bin i is reported in Equation 4.20; however, to smooth the statistical fluctuation between adjacent bins, the alternative-over-nominal ratio is fitted with a linear or exponential function.

$$\sigma^i_{shape} = 1 - \frac{N^i_{alt}}{N^i_{nom}}. \tag{4.20}$$

Three types of distributions are considered for shape uncertainties m_{bb}, p_{T}^V and BDT_S. BDT_S is the BDT output score of a BDT trained to distinguish the nominal and alternative generators for the same physics process and event category [123]. The ratio of BDT_S of the nominal and alternative samples is used for re-weighting the nominal sample to match the alternative one (the re-weighting function called BDT_R), providing simultaneous coverage of all input variables entering the BDT. The input training variables of BDT_S are the same as BDT_{VH}, listed in Table 4.8, except for p_{T}^V, which is excluded from the training. Before training BDT_S, the p_{T}^V distribution of the nominal sample is re-weighted to match the alternative sample, thereby factorising away all dependency of the p_T^V differences between the two generator predictions during the BDT_S training.

4.12.2.1 Signal uncertainties

Table 4.9 shows the signal uncertainties. The normalisation uncertainties affect the total cross-section and $H \to b\bar{b}$ branching ratio; their prescriptions are provided by the LHC Higgs Cross Section Working Group [25]. Then, acceptance uncertainties are associated with the migrations between STXS bins with the procedure detailed in [124]. Variations of the factorisation and renormalisation scales, parton shower and underlying event (PS/UE), and PDF sets have been evaluated [124]. For the acceptance uncertainties, the range in Table 4.9 indicates the maximum and minimum values across the various regions. Lastly, shape (on m_{bb} and p_{T}^V) and acceptance uncertainties are also estimated for the analysis-reconstructed categories for the variations described previously, plus a specific uncertainty on the electro-weak correction as a function of the p_{T}^V.

Signal		
Cross-section	Norm	0.7% (qq) 25% (gg)
$H \to b\bar{b}$ branching ratio	Norm	1.7%
Scale variations in STXS bins	Acc	3.0 – 3.9% ($qq \to WH$) 6.7 – 12% ($qq \to ZH$) 37 – 100% ($gg \to ZH$)
PS/UE variations in STXS bins	Acc	1 – 5% ($qq \to VH$) 5 – 20% ($gg \to ZH$)
PDF+α_S variations in STXS bins	Acc	1.8 – 2.2% ($qq \to WH$) 1.4 – 1.7% ($qq \to ZH$) 2.9 – 3.3% ($gg \to ZH$)
m_{bb} from scale variations	Acc+Shape	($qq \to VH$, $gg \to ZH$)
m_{bb} from PS/UE variations	Acc+Shape	
m_{bb} from PDF+α_S variations	Acc+Shape	
p_T^V from NLO EW correction	Acc+Shape	

Table 4.9: Summary of the systematic uncertainties in the signal.

4.12.2.2 Backgrounds uncertainties

Table 4.10 and Table 4.11 show summaries of the background modelling systematics. Details on each background are reported below:

- *V + jets*. For this background, the flavour contributions bb, bc, bl and cc are referred to collectively as V + HF. Acceptance uncertainties are calculated for the parton shower and the underling event variations using alternative generators (MADGRAPH5_AMC@NLO+PYTHIA8). Furthermore, ratios on the yields of different flavours of the jets produced in association with the vector bosons provide additional acceptance uncertainties. Extrapolation uncertainties are also introduced in the 0-Lepton channel (as ratio of yields between different regions) for the W+jets extrapolated from the 1-Lepton channel, and for the Z+jets extrapolated from the 2-Lepton channel. Similarly, extrapolation uncertainties from SR-to-CR are calculated for all the three lepton channels. Lastly, Z+jets shapes are calculated using m_{bb} side-bands in the 2-Lepton channel after all the other backgrounds have been removed. For W+jets, the shapes are obtained with the BDT_R method described in Section 4.12.2.
- $t\bar{t}$. For 0-Lepton and 1-Lepton channels, the alternative generators are MADGRAPH5 AMC@NLO + PYTHIA8 for the matrix-element variation and POWHEG

+ HERWIG 7 for the parton shower generator. Uncertainties on the flavour composition are calculated by comparing the difference in the ratio of the bc_{yield}/bb_{yield} to the alternative matrix-element and parton-shower generator samples. The procedure is repeated substituting bc with the "others", i.e. the sum of all the non-bb and bc components. Lastly, shapes using the BDT_R method are estimated for both the alternative matrix-element and parton shower generators. For the 2-Lepton channel, the $t\bar{t}$ background is estimated using the $e\mu$ CR (see Section 4.6).

- *Single top*. For the Wt and $t-$channels (see Section 4.3.2), uncertainties are calculated for the normalisation, relative acceptance and shapes. The uncertainties are calculated separately for the different truth-flavour of the selected two b-jets. The alternative generators considered are POWHEG + HERWIG++ for the different parton-shower model, POWHEG + PYTHIA 8 using the diagram subtraction scheme as an alternative interference model with $t\bar{t}$ [91, 125]. For the $s-$channel, due to the small contribution of this background, only normalization uncertainties are considered.

- *Multi-jet*. The multi-jet background is negligible in the 0-Lepton and 2-Lepton channels. For the 1-Lepton channel, several shapes are considered accounting for efficiencies of the lepton triggers, the isolation requirements and the fitted electroweak normalisation factors. The largest shape uncertainty in each analysis region is chosen as the shape systematic. For the normalisation uncertainty, two contributions are added in quadrature: the first one originating from relaxing the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut in the electron sub-channel and re-performing the fit and the second one from fitting $\Delta\Phi(l, bb)$ instead of m_{T}^{W}.

- *Diboson*. For the WW channel, which yields a small contribution to the analysis phase space, only the normalization uncertainty is considered. For ZZ and WZ, the alternative generators used for the parton shower and underlying event simulations are POWHEG + PYTHIA 8 and POWHEG + HERWIG++. Normalisation, acceptance and shape (m_{bb} and p_{T}^{V}) uncertainties are used for ZZ and WZ.

	Z + jets	
$Z + ll$	Norm	18%
$Z + cl$	Norm	23%
Z + HF	Norm	Floating (2-jet, 3-jet) × (75 GeV $< p_T^V <$ 150 GeV, $p_T^V >$ 150 GeV)
$Z + bc$-to-$Z + bb$	Acc	30 – 40%
$Z + cc$-to-$Z + bb$	Acc	13 – 16%
$Z + bl$-to-$Z + bb$	Acc	20 – 28%
SR-to-low ΔR CR	Acc	3.8 – 9.9% (75 GeV $< p_T^V <$ 150 GeV, $p_T^V >$ 150 GeV)
SR-to-high ΔR CR	Acc	2.7 – 4.1% (75 GeV $< p_T^V <$ 150 GeV, $p_T^V >$ 150 GeV)
0-to-2 lepton	Acc	7%
p_T^V	Acc+Shape	(75 GeV $< p_T^V <$ 150 GeV, $p_T^V >$ 150 GeV)
m_{bb}	Shape	(75 GeV $< p_T^V <$ 150 GeV, $p_T^V >$ 150 GeV)
	W + jets	
$W + ll$	Norm	32%
$W + cl$	Norm	37%
W + HF	Norm	Floating (2-jet, 3-jet)
$W + bc$-to-$W + bb$	Acc	15% (0-Lepton) and 30% (1-Lepton)
$W + cc$-to-$W + bb$	Acc	10% (0-Lepton) and 30% (1-Lepton)
$W + bl$-to-$W + bb$	Acc	26% (0-Lepton) and 23% (1-Lepton)
SR-to-CR	Acc	3.6-15%
0-to-1 lepton	Acc	5%
p_T^V	Acc+Shape	(2-jet, 3-jet)
R_{BDT}	Shape	
	$t\bar{t}$ (0+1-Lepton channels only)	
$t\bar{t}$	Norm	Floating (2-jet, 3-jet)
0-to-1 lepton	Acc	8%
$t\bar{t}$ bc-to-bb (ME)	Acc	7.6 – 8.2% (0-Lepton), 1.3 – 3.8% (1-Lepton)
$t\bar{t}$ bc-to-bb (PS)	Acc	2.1 – 3.2% (0-Lepton), 1.5 – 7.1% (1-Lepton)
$t\bar{t}$ other-to-bb (ME)	Acc	2.8 – 6.4% (0-Lepton), 3.3 – 5.7% (1-Lepton)
$t\bar{t}$ other-to-bb (PS)	Acc	5.6 – 13% (0-Lepton), 0.3 – 2.1% (1-Lepton)
p_T^V	Acc+Shape	(2-jet, 3-jet)
R_{BDT} ME variation	Acc+Shape	(2-jet, 3-jet)
R_{BDT} PS variation	Acc+Shape	(0-Lepton, 1-Lepton)
	Single top	
Cross-section	Norm	4.6% (s-channel), 4.4% (t-channel), 6.2% (Wt)
Acceptance 2-jet	Acc	17% (t-channel), 55% ($Wt(bb)$), 24% (Wt(other))
Acceptance 3-jet	Acc	20% (t-channel), 51% ($Wt(bb)$), 21% (Wt(other))
m_{bb}	Acc+Shape	(t-channel, $Wt(bb)$, Wt(other))
p_T^V	Acc+Shape	(t-channel, $Wt(bb)$, Wt(other))
	Multi-jet (1-Lepton)	
Cross-section	Norm	30 – 200% (2-jet), 100% (3-jet)
BDT template	Acc+Shape	

Table 4.10: Summary of the modelling uncertainties for the V+jets, $t\bar{t}$, single top and multi-jet backgrounds.

	ZZ	
Cross-section	Norm	20%
0-to-2 lepton	Acc	6%
Scale variations	Acc	10 – 18%
PS/UE variations for 2 or more jets	Acc	6%
PS/UE variations for 3 jets	Acc	7% (0-Lepton), 3% (2-Lepton)
m_{bb} from scale variations	Acc+Shape	correlated with WZ
p_{T}^{V} from scale variations	Acc+Shape	correlated with WZ
m_{bb} from PS/UE variations	Acc+Shape	correlated with WZ
p_{T}^{V} from PS/UE variations	Acc+Shape	correlated with WZ
m_{bb} from matrix-element variations	Acc+Shape	correlated with WZ
	WZ	
Cross-section	Norm	26%
0-to-1 lepton ratio	Acc	11%
Scale variations	Acc	13 – 21%
PS/UE variations for 2 or more jets	Acc	4%
PS/UE variations for 3 jets	Acc	11%
m_{bb} from scale variations	Acc+Shape	correlated with ZZ
p_{T}^{V} from scale variations	Acc+Shape	correlated with ZZ
m_{bb} from PS/UE variations	Acc+Shape	correlated with ZZ
p_{T}^{V} from PS/UE variations	Acc+Shape	correlated with ZZ
m_{bb} from matrix-element variations	Acc+Shape	correlated with ZZ
	WW	
Cross-section	Norm	25%

Table 4.11: Summary of the modelling uncertainties for the diboson background.

4.13 Results

In this section, the fit results of the main MVA analysis, the di-jet analysis and the di-boson analysis will be presented for the Full Run 2 dataset. An example of a selected event characterised by a high score of BDT$_{VH}$ is shown in Figure 4.29.

The summary of the fitted variables in the signal and control regions for the main MVA analysis is provided in Table 4.12. For the di-jet analysis, the fitted variable in the signal region is m_{bb} instead of BDT$_{VH}$, while for the di-boson analysis it is BDT$_{VZ}$.

4.13.1 Multivariate analysis

The results of the MVA analysis will be presented in two steps: first the signal strength measurements in Section 4.13.1.1, then the differential measurements in Section 4.13.1.2.

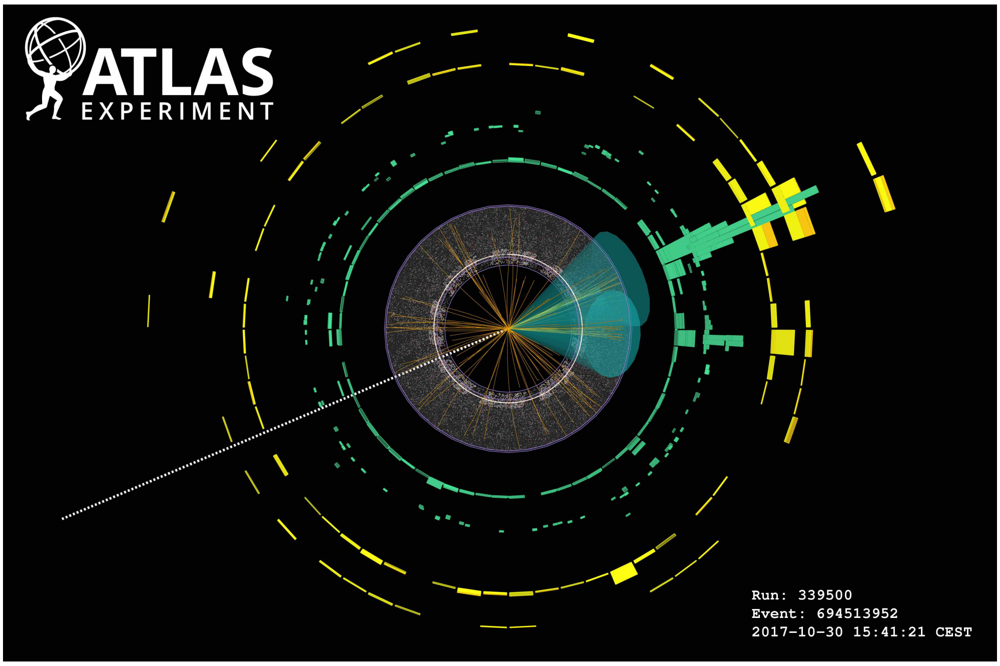


Figure 4.29: Event display of a $ZH \to \nu\bar{\nu}b\bar{b}$ candidate in the 0-Lepton channel interacting with the ATLAS detector [2, 60].

4.13.1.1 Signal strength measurements

For the MVA analysis, after performing the global likelihood fit across the three lepton channels, the measured signal strength for the $VH, H \to b\bar{b}$ is

$$\mu_{VH}^{b\bar{b}} = 1.02^{+0.18}_{-0.17} = 1.02^{+0.12}_{-0.11}(\text{stat.})^{+0.14}_{-0.13}(\text{syst.}) \quad (4.21)$$

This result is compatible with the Standard Model Higgs boson within experimental errors. The measured signal strength corresponds to a rejection of the background-only hypothesis with a significance of 6.7 standard deviations, while the expected significance from the global Asimov fit was also 6.7 standard deviations. The combined fit has also been performed separately on the signal strengths for the WH and ZH productions where the WH and ZH processes are left floating independently:

$$\mu_{WH}^{b\bar{b}} = 0.95^{+0.27}_{-0.25} = 0.95 \pm 0.18(\text{stat.})^{+0.19}_{-0.18}(\text{syst.}) \quad (4.22)$$

$$\mu_{ZH}^{b\bar{b}} = 1.08^{+0.25}_{-0.23} = 1.08 \pm 0.17(\text{stat.})^{+0.18}_{-0.15}(\text{syst.}) \quad (4.23)$$

Channel	Regions	Categories					
		75 GeV< p_{T}^{V} < 150 GeV		150 GeV< p_{T}^{V} < 250 GeV		p_{T}^{V} > 250 GeV	
		2-jets	3-jets	2-jets	3-jets	2-jets	3-jets
0-Lepton	Low-ΔR-CR	–	–	yields	yields	yields	yields
	SR	–	–	BDT_{VH}	BDT_{VH}	BDT_{VH}	BDT_{VH}
	High-ΔR-CR	–	–	yields	yields	yields	yields
1-Lepton	Low-ΔR-CR	– -	– -	yields	yields	yields	yields
	SR	– -	– -	BDT_{VH}	BDT_{VH}	BDT_{VH}	BDT_{VH}
	High-ΔR-CR	– -	– -	yields	yields	yields	yields
2-Lepton	Low-ΔR-CR	yields	yields	yields	yields	yields	yields
	SR	BDT_{VH}	BDT_{VH}	BDT_{VH}	BDT_{VH}	BDT_{VH}	BDT_{VH}
	High-ΔR-CR	yields	yields	yields	yields	yields	yields

Table 4.12: Analysis regions and fitted variables in the three lepton channels for the $VH, H \to b\bar{b}$ multivariate analysis [4]. In each control region, the number of events (yields) is fitted. Instead, in each signal region, the shape of the output discriminant (BDT_{VH}) is fitted.

The observed and expected significances for the WH production are respectively 4.0 and 4.1, which indicate strong evidence for this process, while the observed and expected significance for the ZH production are respectively 5.3 and 5.1, which allows an observation of ZH production mode to be claimed. The likelihood contours for the Higgs boson signal strengths WH and ZH are shown in Figure 4.30.

The normalisation factors of the floating backgrounds are reported in Table 4.13. The $Z +$ HF scale factors deviate substantially from unity; the difference has been attributed to a poor modelling; high values of the $Z +$ HF scale factors have also been observed by the CMS $VH, H \to b\bar{b}$ [89] and by a dedicated ATLAS $Z \to b\bar{b}$ analysis [126]. In Table 4.14 and Table 4.15, the post-fit signal, background and data yields for the three lepton channels are reported.

Figure 4.31 shows the post-fit distributions for the fitted variables in the 0-Lepton 2-jet category, while Figure 4.32 shows the same variables for the 3-jet category. The BDT_{VH} output score is shown with the binning used in the global fit. The post-fit event yields for all the analysis regions and lepton channels grouped in bins of $\log_{10}(S/B)$ are shown in Figure 4.33.

The ranking and breakdown of the systematic uncertainties, described in Section 4.11.2,

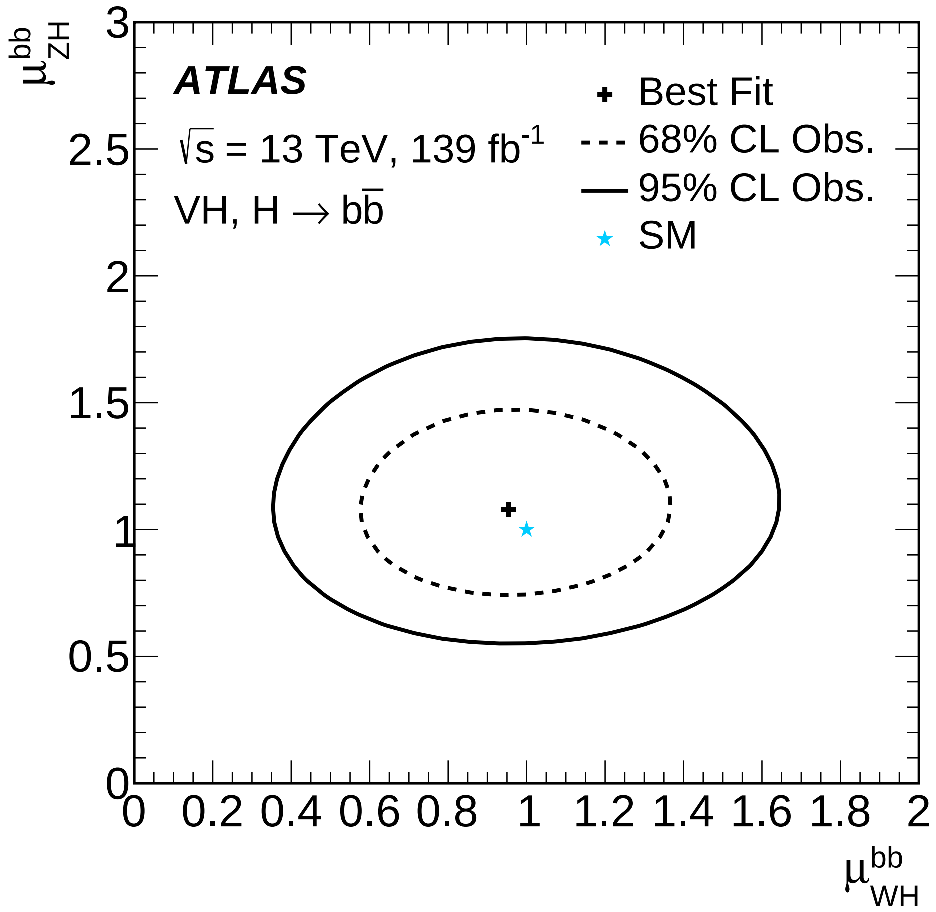


Figure 4.30: The likelihood contours for the WH and ZH production signal strengths.

are shown in Figure 4.34 and Table 4.16, respectively.

In the ranking plot in Figure 4.34, for the experimental systematics, the largest contributions come from the flavour-tagging uncorrelated components and the jet energy scale. For the modelling systematics, the main contributions arise from the shapes of the BDT_R for the W boson and VH theory systematics. Furthermore, the seventh and tenth systematics are related to the diboson modelling; these systematics are ranked so high because they have not been updated with the VHBBTRUTHFRAMEWORK since the previous publication [2] due to lack of time. Uncertainties on $Z+$HF appear in the ranking because of the poor modelling of this background in the MC generators. This mis-modeling is particularly enhanced in the control regions, yielding large extrapolation uncertainties.

From the breakdown of systematic uncertainties (see Table 4.16), it is interesting to

Process and Category	Normalisation factor
$t\bar{t}$ 2-jet	0.98 ± 0.09
$t\bar{t}$ 3-jet	0.93 ± 0.06
$W +$ HF 2-jet	1.06 ± 0.11
$W +$ HF 3-jet	1.15 ± 0.09
$Z +$ HF 2-jet, $75 < p_{\mathrm{T}}^{V} < 150$ GeV	1.28 ± 0.08
$Z +$ HF 3-jet, $75 < p_{\mathrm{T}}^{V} < 150$ GeV	1.17 ± 0.05
$Z +$ HF 2-jet, 150 GeV $< p_{\mathrm{T}}^{V}$	1.16 ± 0.07
$Z +$ HF 3-jet, 150 GeV $< p_{\mathrm{T}}^{V}$	1.09 ± 0.04

Table 4.13: Normalisation scale factors for the main backgrounds of the MVA $VH, H \rightarrow b\bar{b}$ analysis after the global likelihood fit [4].

notice that the MC statistical error is outside the five largest sources of uncertainties thanks to a large number of simulated events, statistical b-tagging and dedicated MC filters. In the ATLAS publication of the evidence of $H \rightarrow b\bar{b}$ decays with $L = 36.1$ fb^{-1} with early Run 2 data, the MC statistics was the second-largest uncertainty [127]; the largest uncertainty was the modelling of the signal, which did not affect the expected p-value since the expected p-value is calculated on the fluctuation of the background-only hypothesis. The MC statistics is also the largest systematic in the CMS $H \rightarrow b\bar{b}$ observation analysis [89].

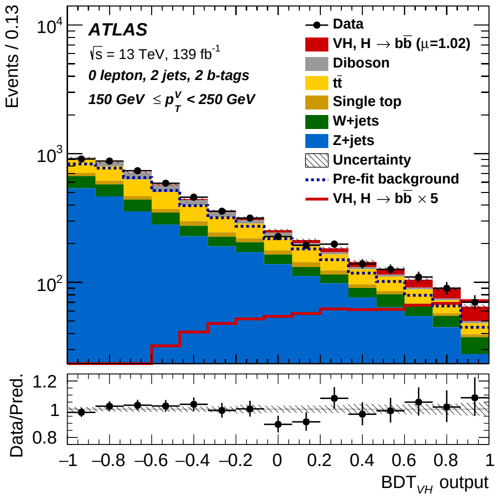
Overall, the $VH, H \rightarrow b\bar{b}$ analysis is already systematically limited, as it can be seen from the top of Table 4.16. Therefore, improving the b-tagging and the jet energy scale, on one side, and the modelling, on the other side, will be crucial to further improve the analysis sensitivity. For this reason, simply adding more data during the next period of data taking (Run 3) would not significantly improve the precision of the measurement. For this reason, some new analysis strategy improvements will be discussed in Chapter 5.

Signal regions	0-Lepton $150\,\text{GeV} < p_T^V < 250\,\text{GeV}$		0-Lepton $p_T^V > 250\,\text{GeV}$		1-Lepton $150\,\text{GeV} < p_T^V < 250\,\text{GeV}$		1-Lepton $p_T^V > 250\,\text{GeV}$	
Sample	2-jet	3-jet	2-jet	3-jet	2-jet	3-jet	2-jet	3-jet
Z + jets	2846 ± 80	3830 ± 160	338 ± 13	533 ± 23	102 ± 5	207 ± 11	10 ± 1	25 ± 2
W + jets	634 ± 63	1500 ± 110	83 ± 9	220 ± 19	1850 ± 160	4080 ± 270	353 ± 33	935 ± 69
Single top	237 ± 35	770 ± 130	9 ± 2	36 ± 8	990 ± 160	3570 ± 600	70 ± 15	313 ± 66
$t\bar{t}$	1157 ± 76	5470 ± 220	39 ± 5	151 ± 16	4600 ± 210	21030 ± 620	188 ± 19	970 ± 58
Diboson	360 ± 55	333 ± 79	86 ± 13	70 ± 17	229 ± 57	264 ± 83	57 ± 15	68 ± 22
Multi-jet	–	–	–	–	130 ± 47	24 ± 14	16 ± 10	7 ± 8
Total bkg.	5234 ± 63	11910 ± 100	554 ± 15	1011 ± 21	7899 ± 82	29170 ± 170	694 ± 23	2318 ± 42
Signal ($\mu = 1.02$)	147 ± 24	130 ± 22	40 ± 6	33 ± 6	148 ± 24	125 ± 21	52 ± 9	44 ± 7
Data	5397	11875	578	1046	8044	29316	727	2378

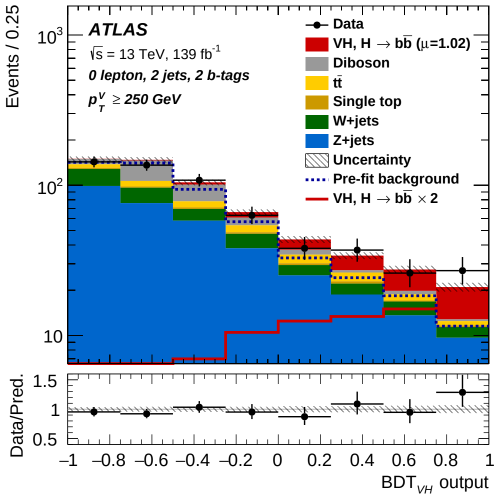
Table 4.14: Post-fit signal, background and data yields for the 0- and 1-Lepton channels [4].

Signal regions	2-Lepton 75 GeV $< p_T^V <$ 150 GeV		2-Lepton 150 GeV $< p_T^V <$ 250 GeV		2-Lepton $p_T^V >$ 250 GeV	
Sample	2-jet	≥3-jet	2-jet	≥3-jet	2-jet	≥3-jet
Z + jets	5900 ± 100	11630 ± 160	716 ± 19	2499 ± 52	84 ± 3	537 ± 16
W + jets	1 ± 0	6 ± 0	< 1	2 ± 0	< 1	< 1
Top	3193 ± 57	8796 ± 87	52 ± 7	389 ± 19	1 ± 1	15 ± 4
Diboson	283 ± 47	443 ± 78	83 ± 14	169 ± 30	20 ± 4	52 ± 10
Total bkg.	9378 ± 86	20880 ± 130	851 ± 19	3058 ± 44	106 ± 4	605 ± 14
Signal ($\mu = 1.02$)	78 ± 14	106 ± 21	34 ± 6	59 ± 12	10 ± 2	18 ± 3
Data	9463	20927	881	3148	123	614

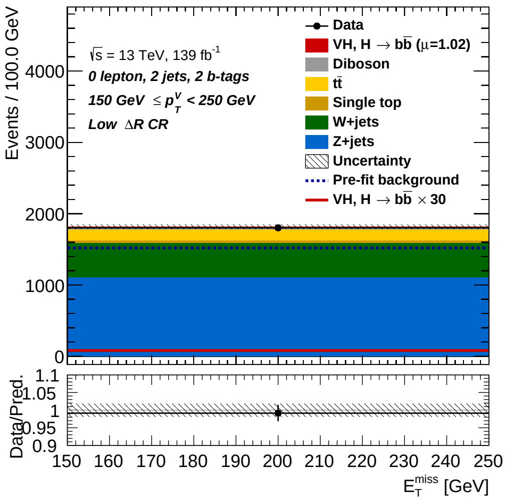
Table 4.15: Post-fit signal, background and data yields for the 2-Lepton channel [4].



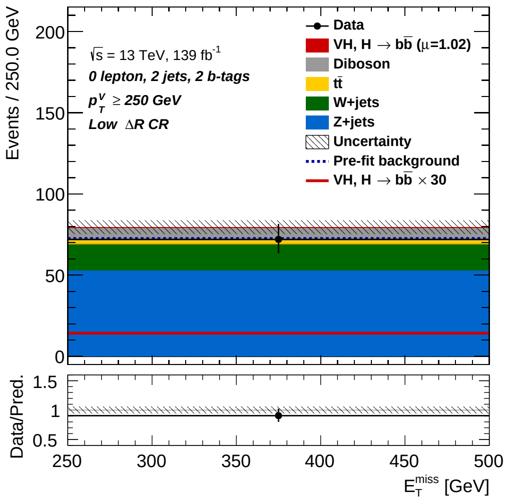
(a) SR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



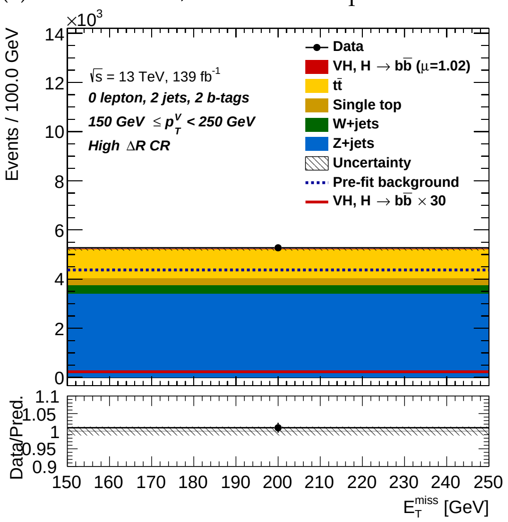
(b) SR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.



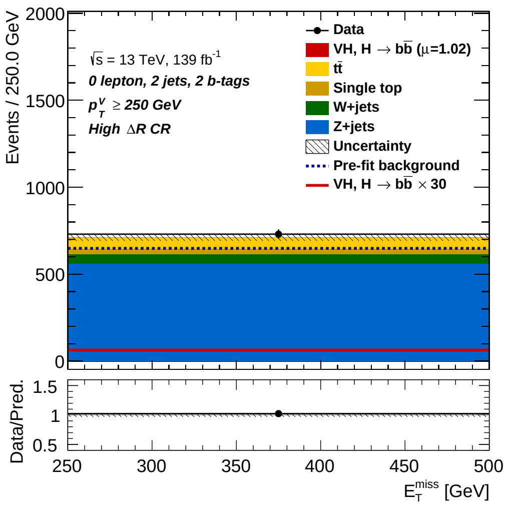
(c) Low-ΔR-CR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



(d) Low-ΔR-CR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

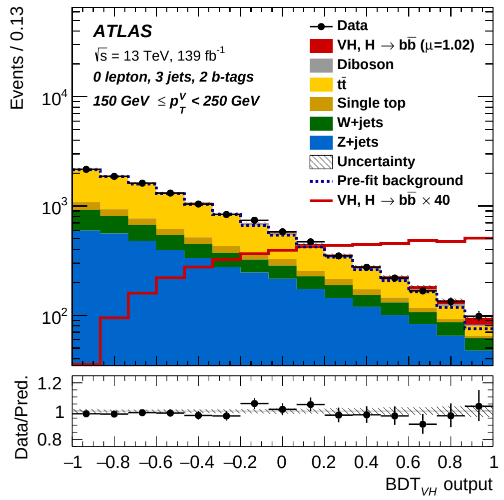


(e) High-ΔR-CR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.

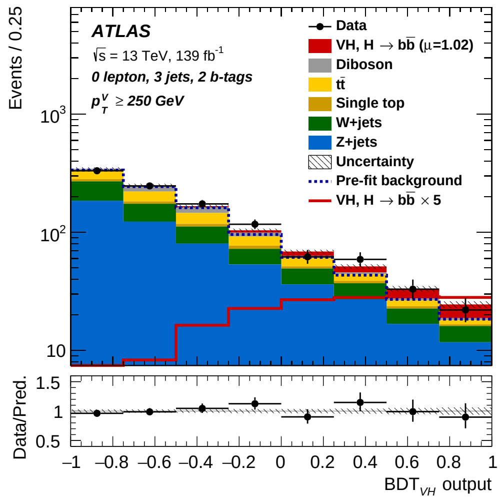


(f) High-ΔR-CR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

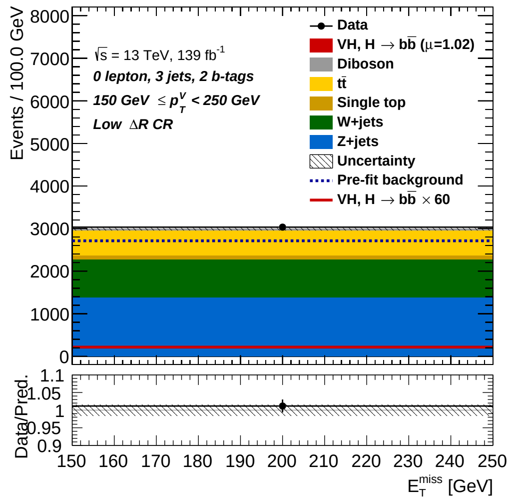
Figure 4.31: The top row shows the BDT$_{VH}$ output post-fit distributions in the SRs while the middle and bottom rows show the post-fit distributions for the CRs. All plots refer to the 0-Lepton channel 2-jet category. The red continuous line shows the signal shape, while the dashed histogram indicates the total pre-fit backgrounds. The displayed errors include both the statistical and systematic errors. The bottom panel shows the ratio between the measured data and the total MC predictions [4].



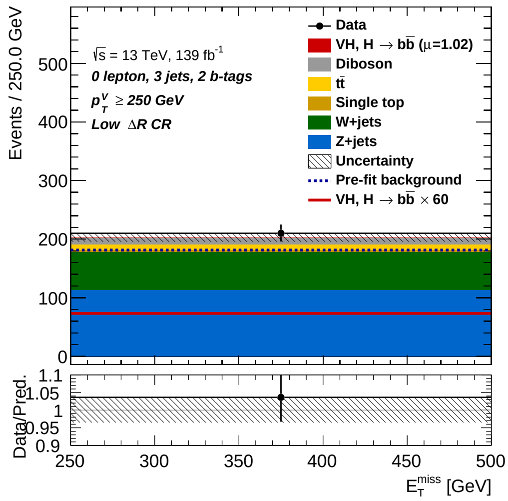
(a) SR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



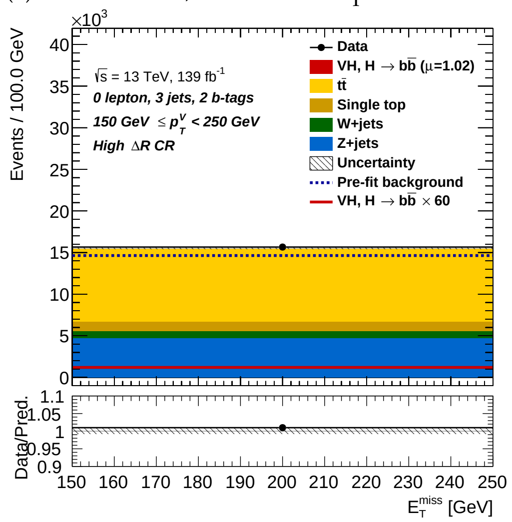
(b) SR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.



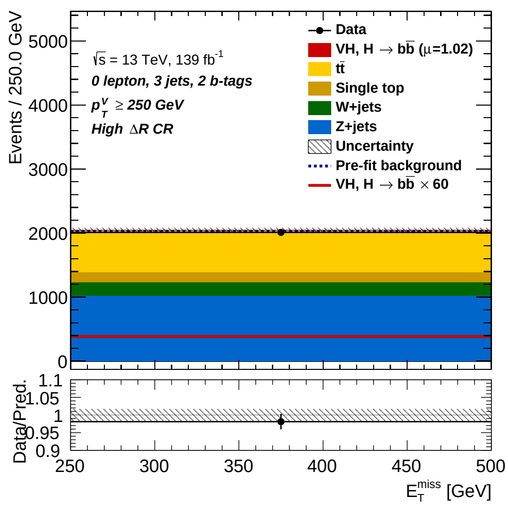
(c) Low-ΔR-CR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



(d) Low-ΔR-CR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.



(e) High-ΔR-CR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



(f) High-ΔR-CR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

Figure 4.32: The top row shows the BDT$_{VH}$ output post-fit distributions in the SRs while the middle and bottom rows show the post-fit distributions for the CRs. All plots refer to the 0-Lepton channel 3-jet category. The red continuous line shows the signal shape, while the dashed histogram indicates the total pre-fit backgrounds. The displayed errors include both the statistical and systematic errors. The bottom panel shows the ratio between the measured data and the total MC predictions [4].

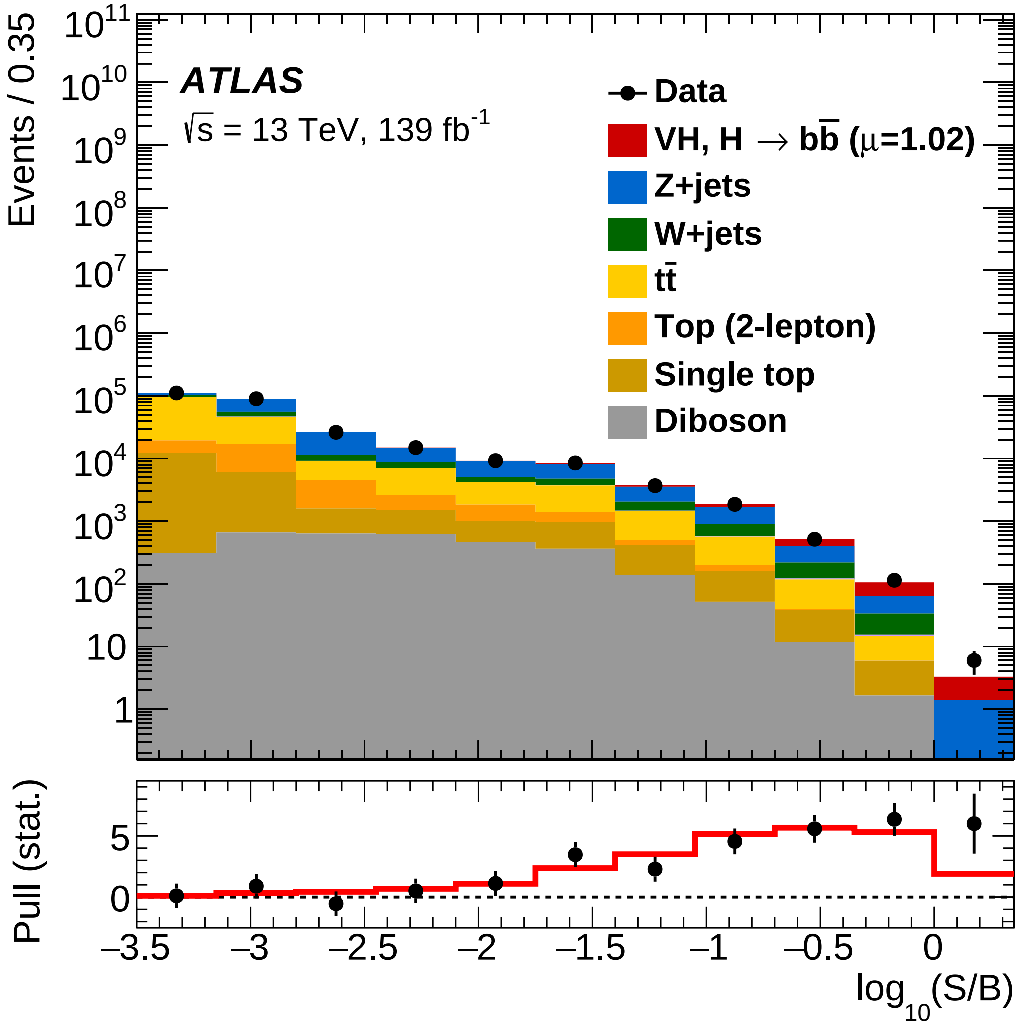


Figure 4.33: Post-fit event yields as a function of $\log_{10}(S/B)$. BDT$_{VH}$ output post-fit bins in all regions and channels have been combined into bins of $\log_{10}(S/B)$. In the bottom pannel, the difference between the data and the total background (*pull*) is shown with only the statistical uncertainties. The red line represents the pull expected from the sum of fitted signal and total background relative to the fitted total background. The dashed line indicates the expected values for the hypothesis of only background [4].

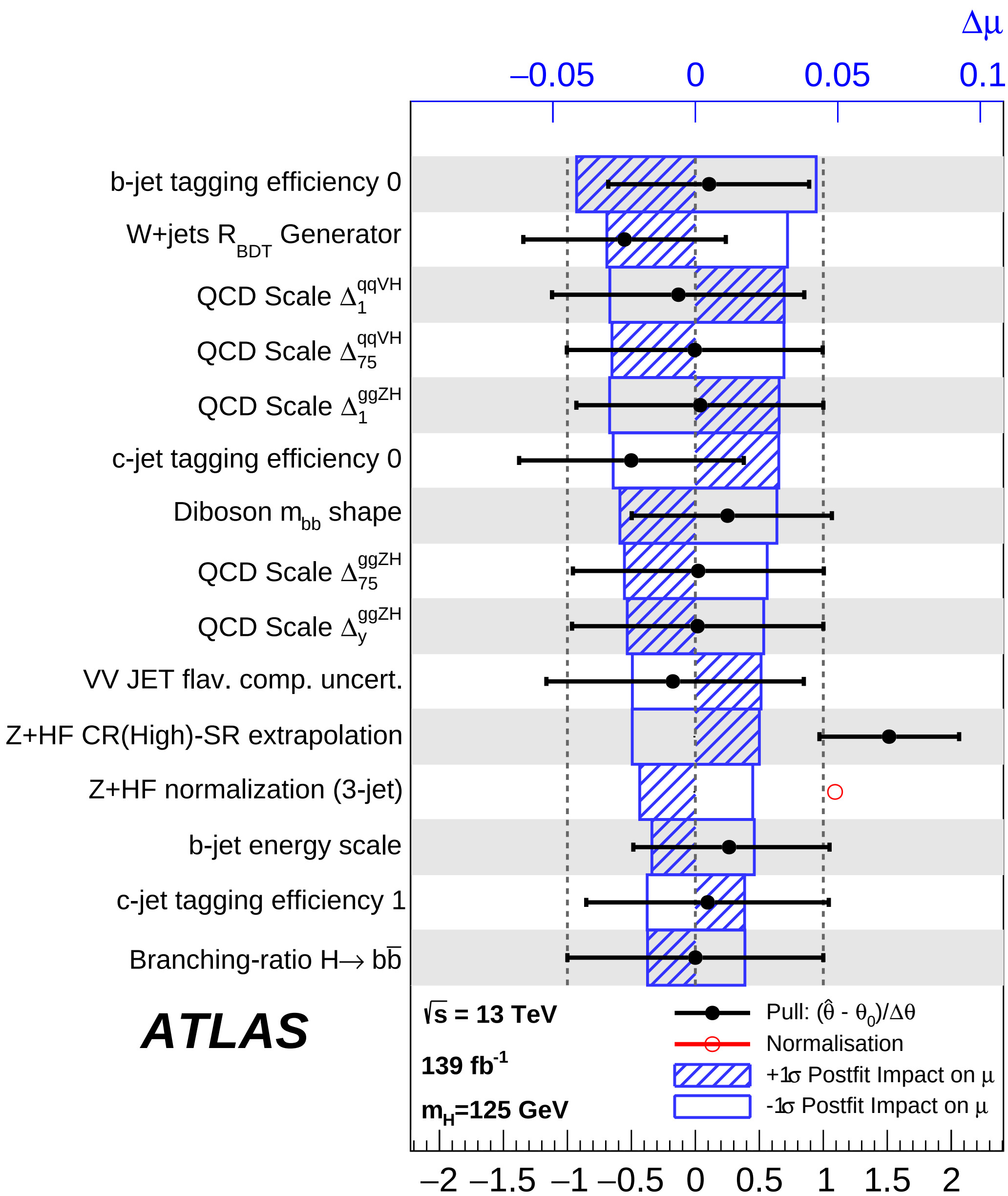


Figure 4.34: Ranking of systematic uncertainties for the fitted Higgs-boson signal strength in the MVA analysis. The top x-axis refers to the blue boxes, which indicate the $\Delta\mu$ variation, while the bottom x-axis refers to the black dots, which indicate the pulls originating from the fit. The hatched boxes indicate upwards variations, whilst the open ones indicate downwards variations. The pull on a nuissance parameter θ_i is the difference between the estimated $\hat{\theta}_i$ from the fit and its nominal value ($pull = (\hat{\theta} - \theta_0)/\Delta\theta$). The pulls are expressed in units of standard deviations. The red dot, which refers to the bottom x-axis, indicates the normalisation scale factors extracted from the fit [4].

Source of uncertainty		σ_μ VH	WH	ZH
Total		0.177	0.260	0.240
Statistical		0.115	0.182	0.171
Systematic		0.134	0.186	0.168
Statistical uncertainties				
Data statistical		0.108	0.171	0.157
$t\bar{t}$ $e\mu$ control region		0.014	0.003	0.026
Floating normalisations		0.034	0.061	0.045
Experimental uncertainties				
Jets		0.043	0.050	0.057
E_T^{miss}		0.015	0.045	0.013
Leptons		0.004	0.015	0.005
b-tagging	b-jets	0.045	0.025	0.064
	c-jets	0.035	0.068	0.010
	light-flavour jets	0.009	0.004	0.014
Pile-up		0.003	0.002	0.007
Luminosity		0.016	0.016	0.016
Theoretical and modelling uncertainties				
Signal		0.072	0.060	0.107
Z + jets		0.032	0.013	0.059
W + jets		0.040	0.079	0.009
$t\bar{t}$		0.021	0.046	0.029
Single top quark		0.019	0.048	0.015
Diboson		0.033	0.033	0.039
Multi-jet		0.005	0.017	0.005
MC statistical		0.031	0.055	0.038

Table 4.16: Breakdown of the contributions to the uncertainty for the Higgs signal-strength measurements [4].

4.13.1.2 STXS measurements

The five measured STXS VH cross-sections, defined in the numerator of Equation 4.12, and their respective Standard Model predictions are shown in Figure 4.35 [4]. All the measurements are consistent with the Standard Model hypothesis. The relative uncertainties on the measurements reach $\sim$ 30% in the highest p_{T}^{V} bins where the statical uncertainty is an important limiting factor. The theoretical uncertainties on the modelling of the VH signal are relatively small, except for the $gg \rightarrow ZH$, originating from the limited precision of the theoretical calculations. Lastly, the background-only hypothesis has been rejected with more than 3σ in three individual STXS bins: ZH 150 GeV $< p_{\mathrm{T}}^{V} <$ 250 GeV, WH $p_{\mathrm{T}}^{V} >$ 250 GeV, and ZH $p_{\mathrm{T}}^{V} >$ 250 GeV.

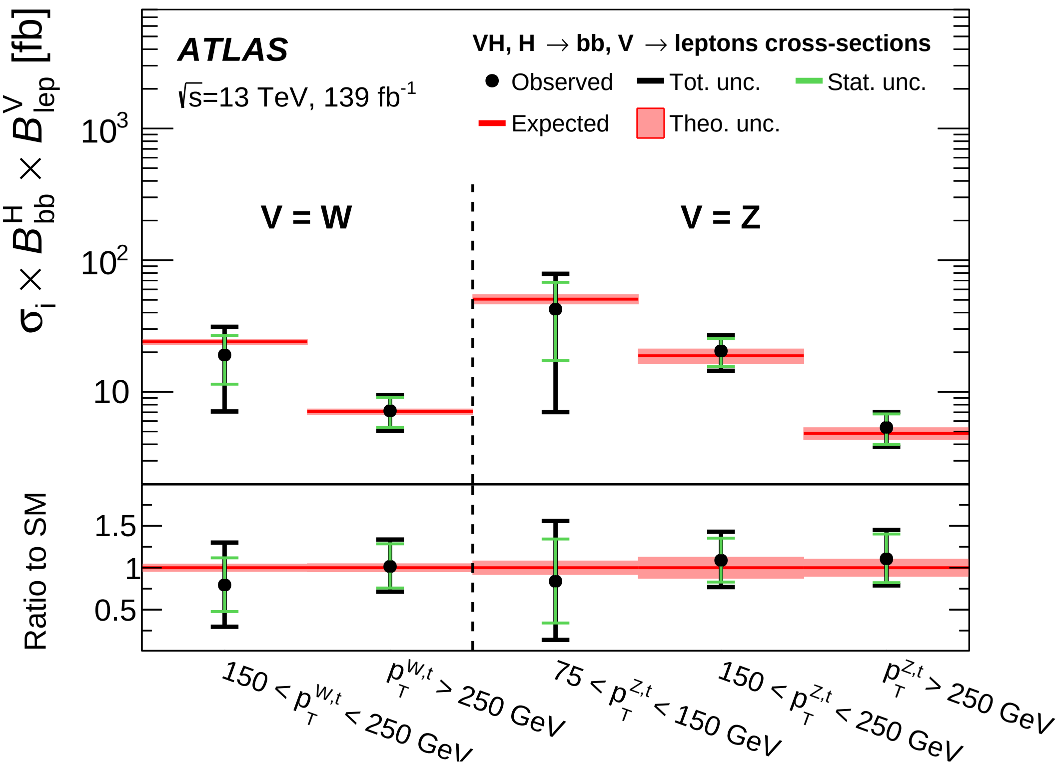


Figure 4.35: Measured cross-sections for the $VH, V \rightarrow lep, H \rightarrow b\bar{b}$ processes in the five STXS bins [4].

The measured correlations between the STXS cross-sections are detailed in Figure 4.36. The correlation matrix is built from the inverse of the Hessian matrix of the likelihood ($H_{ij} = \frac{\partial^2 \mathcal{L}}{\partial\theta_i \partial\theta_j}$). The values on the correlation matrix are almost diagonal due to the reconstructed analysis regions being matched to the STXS bins.

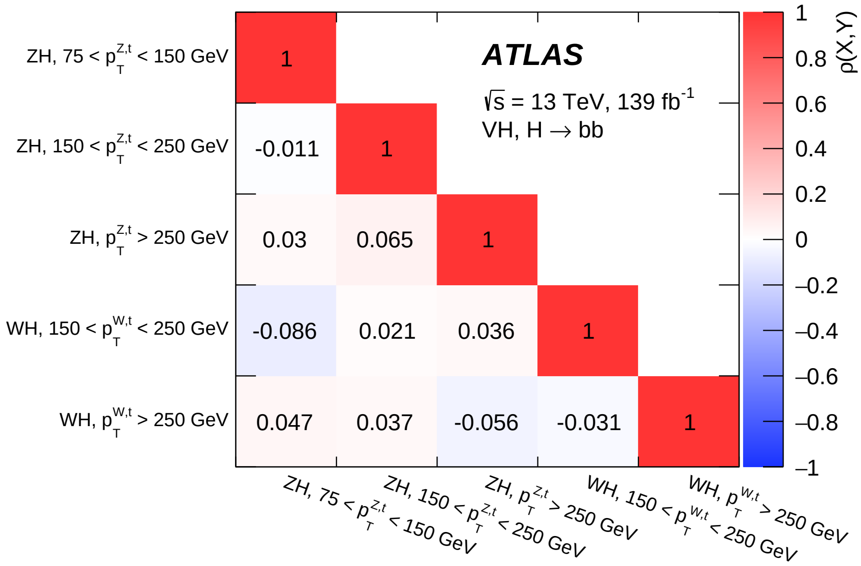


Figure 4.36: Observed correlations between the measured STXS cross-sections [4].

4.13.2 Di-jet analysis

In the di-jet analysis, the m_{bb} is fitted instead of the BDT$_{VH}$ in the SRs in Table 4.12. This analysis aims to measure $\mu_{VH}^{b\bar{b}}$ using the most discriminating variable entering in the BDT to provide a cross-check of the MVA analysis. The measured signal strength for the combined fit is

$$\mu_{VH}^{b\bar{b}} = 1.17^{+0.25}_{-0.23} = 1.17 \pm 0.16(\text{stat.})^{+0.19}_{-0.16}(\text{syst.}) \tag{4.24}$$

the result is in agreement with the Standard Model prediction and with the measurement of the MVA analysis. The observed and expected significances to reject the background-only hypothesis are respectively 5.5 and 4.9. The weighted background-subtracted m_{bb} distribution is shown in Figure 4.37.

4.13.3 Diboson analysis

Before unblinding the VH measurements, the diboson signal strength is measured. This measurement provides a valuable validation on the analysis strategy and the fit model using the Z boson, which was measured with high precision at the LEP collider

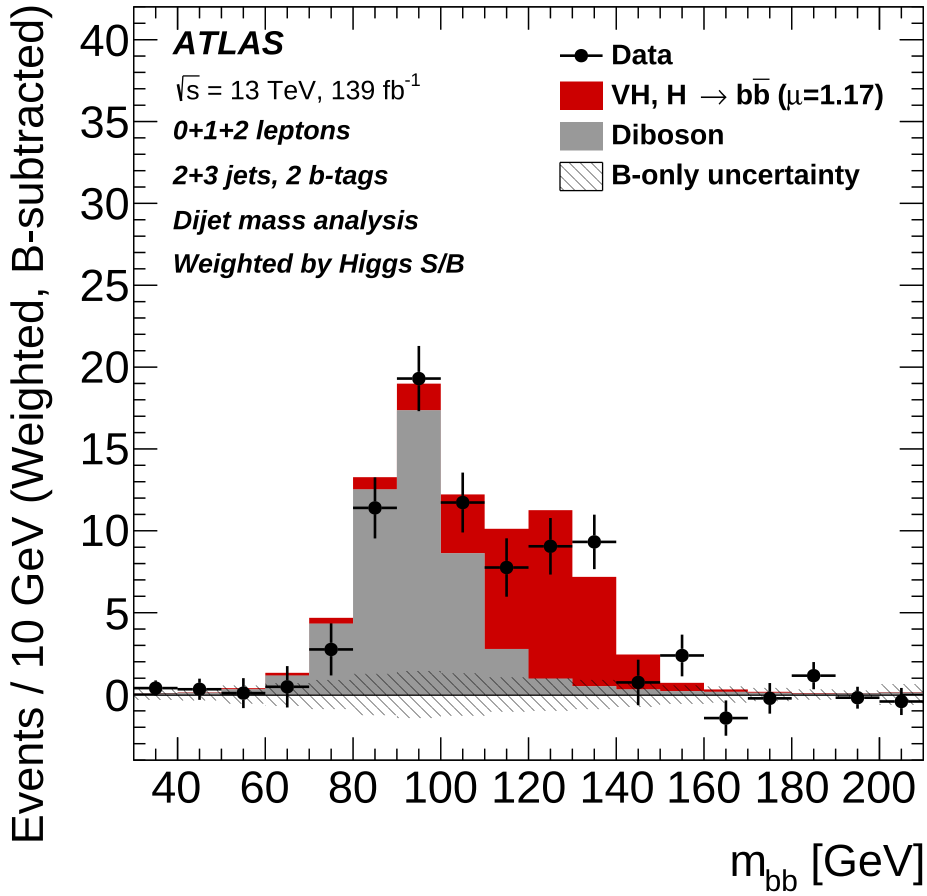
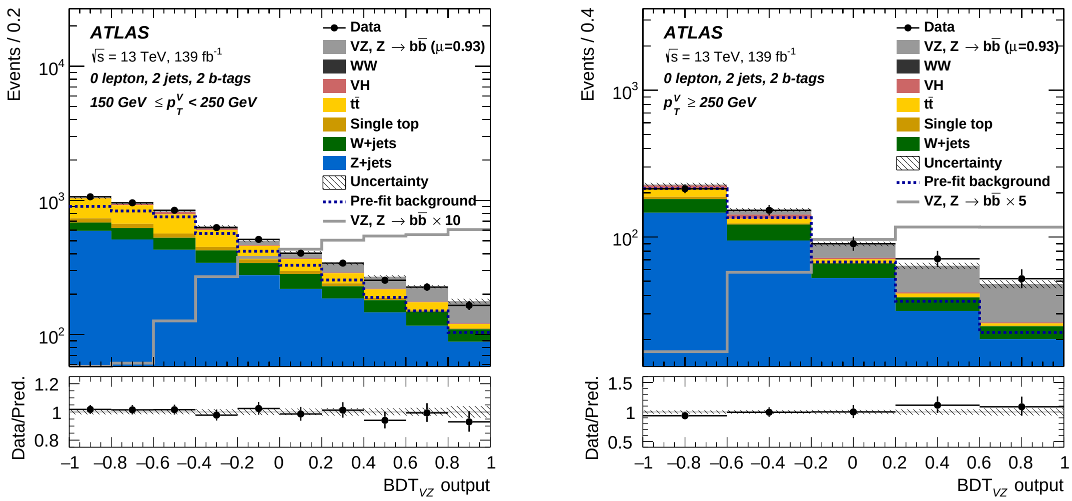


Figure 4.37: The $m_{b\bar{b}}$ distribution in data after the subtraction of all backgrounds except for the diboson processes. The contributions from all the analysis regions are summed and weighted by their respective post-fit S/B. The signal and diboson templates have been scaled by their respective normalisation factors (for the signal $\mu = 1.17$). The size of the total uncertainty for the fitted background is indicated by the hatched band [4].

[128], as a standard candle. For this analysis, BDT$_{VZ}$ is fitted instead of the BDT$_{VH}$ in the SRs in Table 4.12. The BDT$_{VZ}$ distributions in the 0-Lepton SRs for the 2-jet categories are shown in Figure 4.38.

The measured signal strengths, with their respective expected and observed significances, for the combined fit are shown in Table 4.17. All the measurements are consistent with the Standard Model predictions. The high significance of ZZ compared to WZ is due to combinatorics, as explained in Section 4.3.3. The two main limiting factors of this measurement are the modelling of the diboson processes and the limited MC statistics of the diboson samples, which limits the training of the BDT$_{VZ}$.



(a) SR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV. (b) SR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

Figure 4.38: The BDT$_{VZ}$ output post-fit distributions for the 0-Lepton channel 2-jet category in the SRs. The grey continuous line shows the diboson shape, while the dashed histogram indicates the total pre-fit backgrounds. The displayed errors include both the statistical and systematic errors. The bottom panel shows the ratio between the measured data and the total MC predictions [4].

Measured signal strength	Exp. significance	Obs. significance
$\mu^{b\bar{b}}_{WZ} = 0.68^{+0.26}_{-0.24} = 0.68^{+0.15}_{-0.16}(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$	3.9	2.7
$\mu^{b\bar{b}}_{ZZ} = 1.00^{+0.18}_{-0.15} = 1.00^{+0.08}_{-0.08}(\text{stat.})^{+0.16}_{-0.13}(\text{syst.})$	9.3	9.4
$\mu^{b\bar{b}}_{VZ} = 0.93^{+0.15}_{-0.14} = 0.93^{+0.07}_{-0.06}(\text{stat.})^{+0.14}_{-0.12}(\text{syst.})$	10.5	10.3

Table 4.17: Signal strengths, expected and observed significances for the diboson measurements [4].

Chapter 5

Future analysis improvements

The $VH, H \to b\bar{b}$ analysis described in Chapter 4 represents an important milestone towards the understanding of the nature of the Higgs boson. Although all current Higgs measurements are in agreement with the Standard Model, many puzzles are still left in Particle Physics.

The EFT formalism (see Section 1.5) provides a way to probe Beyond Standard Model theories. In particular, EFT Wilson coefficients can be extracted from Simplified Template Cross-Sections measurements. Limits on the Wilson coefficients have been placed using the latest $VH, H \to b\bar{b}$ analysis [4] and there are also efforts to perform global fits combing multiple electroweak measurements taken at LEP and LHC [129]. In the STXS framework, the $VH, H \to b\bar{b}$ analysis plays the leading role in the VH production mode [2]. To improve the sensitivity to the Wilson coefficients, it will be necessary to move to a more differential measurement with more bins than the five currently adopted in Figure 4.35. The LHC Higgs Cross Section Working Group [22] has already outlined the next step, which requires finer STXS binning, for the $VH, H \to b\bar{b}$ analysis as described in Figure 5.1. This will require additional divisions in p_{T}^{V}, the number of additional jets in the event and the production mode ($q\bar{q}$ and gg).

To achieve the precision necessary to measure many STXS bins, new strategies for the $VH, H \to b\bar{b}$ analysis are required; a few will be discussed in this chapter.

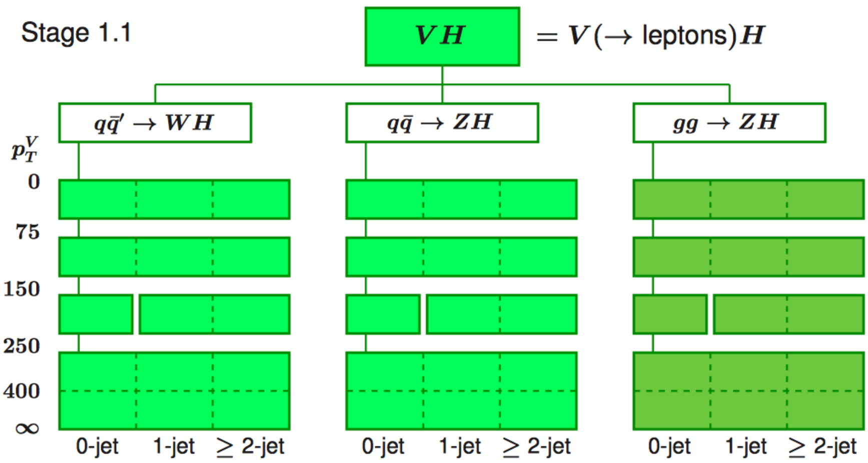
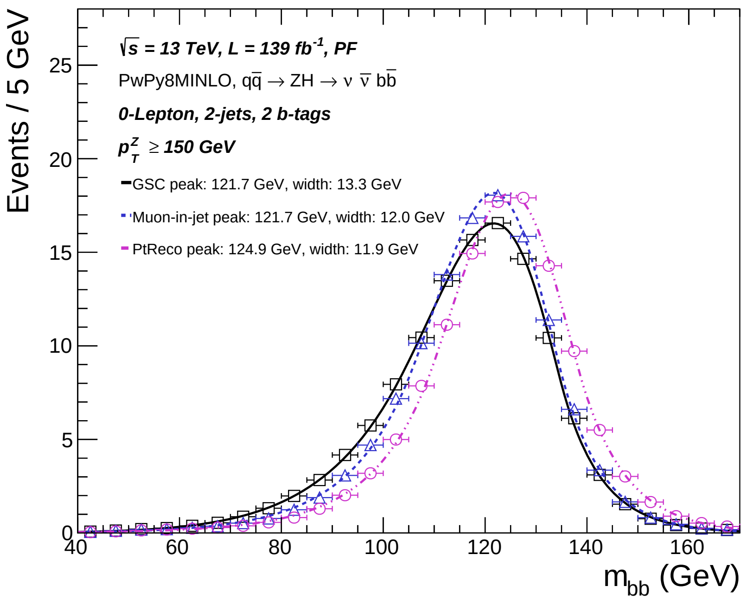


Figure 5.1: Proposed binning for the next iteration of the $VH, H \rightarrow b\bar{b}$ analysis [130].

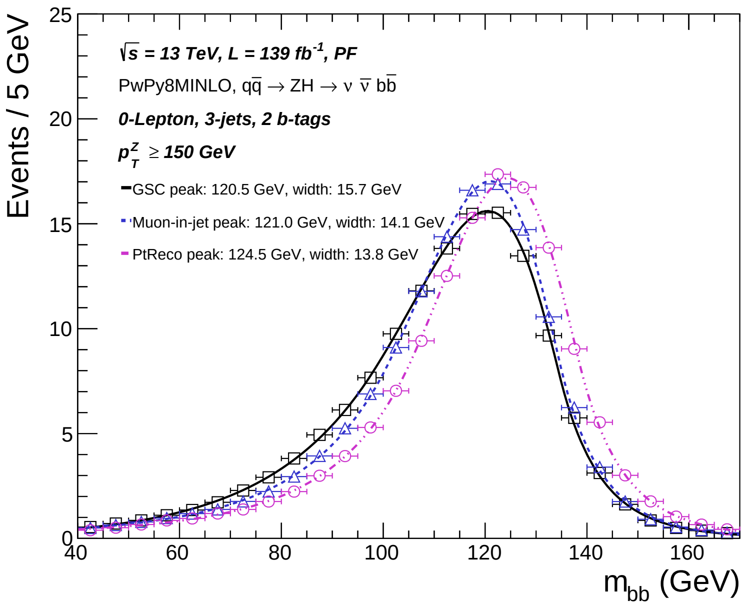
5.1 Particle-flow jets

As described in Section 3.2.4, particle-flow jets [71] promise better energy resolution and pile-up rejections than EMTopo jets. The particle-flow jet collection was tested by ATLAS in January 2019 at the early stages of the Full Run 2 $VH, H \rightarrow b\bar{b}$ analysis. The m_{bb} distribution obtained using particle-flow jets after applying the different energy corrections specific for the $VH, H \rightarrow b\bar{b}$ analysis, which were introduced in Section 4.8, are displayed in Figure 5.2. Compared to the energy resolutions offered by EMTopo jets in Figure 4.21, particle-flow jets offer an improvement of between $\sim 2\%$ to $\sim 4\%$. Furthermore, the resolution for the PtReco correction for particle-flow jets could be slightly better since the scale factors used for this study are not optimized for particle-flow jets but were derived on EMTopo jets.

Unfortunately, a migration of events towards higher jet multiplicity was noticed across three types of sample (VH signal, $t\bar{t}$ and data), mostly outside the analysis phase space. In particular, the signal 2-jet category decreases by 7.2% while the 3-jet category increases by 1.3%. In the case of the $t\bar{t}$ samples, the events in the 2-jet and 3-jet categories decrease by 11.5% and 8.8% respectively. An increase of pile-up jets was the main hypothesis for the jet migration. To validate this idea, the third non-b-tagged jet in the 3-jet category was investigated; this was done to disentangle the effect of the JVT (see Section 3.2.4.2) and the MV2c10 (see Section 3.2.4.3) since the MV2c10 has some antipile-up properties. The ratio between particle-flow and EMTopo jets originating from pile-up for the third jet is shown in Figure 5.3.



(a) 2-jet.



(b) 3-jet.

Figure 5.2: Comparison of the m_{bb} distributions for the b-jet energy corrections in the 0-Lepton channel for 2-jet and 3-jet categories for particle-flow jets (see Section 3.2.4). The distributions have been fitted with a Bukin function to account for the asymmetric shape of the peak [112].

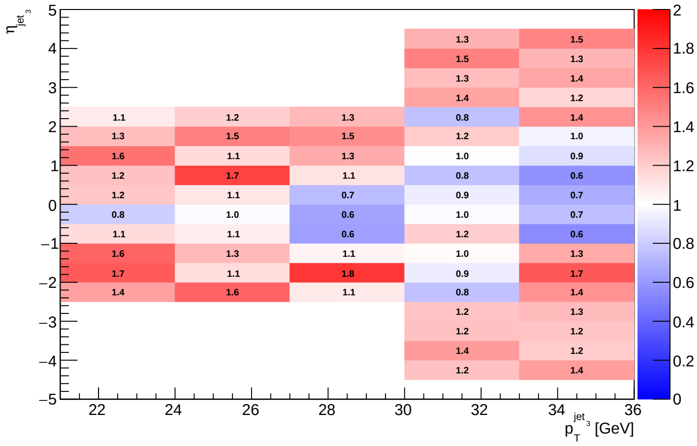


Figure 5.3: The ratio of the number of pile-up jets is displayed for the 0-Lepton 3-jet category between particle-flow and EMTopo jets for the third non-b-tagged jets in simulated signal events; red colours indicate more pile-up jets for particle-flow jets than EMTopo jets.

In Figure 5.4, the average number of jets per events is reported as a function of the number of collisions in the events. These plots corroborated the suspicion that particle-flow jets, with the JVT prescription of January 2019, are less pile-up resistant than EMTopo jets. Therefore, EMTopo jets were chosen for the $VH, H \to b\bar{b}$ analysis described in Chapter 4. For the future, once an updated JVT prescription will be available, particle-flow jets should be used in the analysis for their better energy resolution.

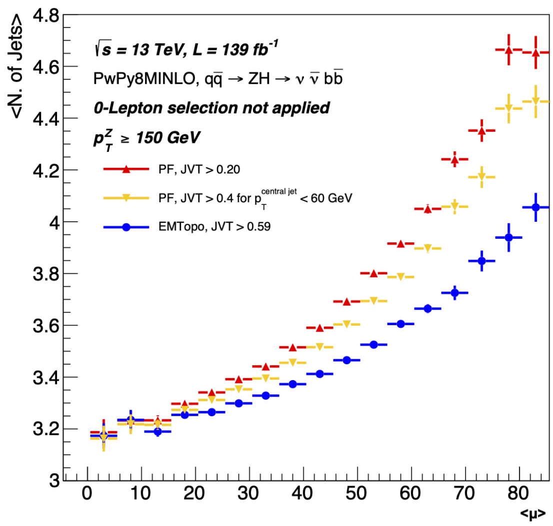


Figure 5.4: Average number of jets as a function of the number of interactions for EM-Topo jets with JVT > 0.59, for particle-flow jets with JVT > 0.20 (January 2019 prescription) and for particle-flow jets with JVT > 0.40 for central jets.

5.2 FSR recovery

B-jets originating from the Higgs boson decay can emit Final-State Radiation (see Section 4.2.1). In the $VH, H \to b\bar{b}$ analysis, most of the events with FSR are categorised in the 3-jet category. However, as shown in Figure 4.5, the m_{bb} resolution for FSR events is substantially degraded. Therefore, a strategy based on topological quantities has been investigated in order to add back the four-momentum of the FSR jet to its closest b-jet.

From Figure 4.7, it can be concluded that the FSR jet is mostly between the two b-jets, indicating that it tends to be mostly collinear with the Higgs momentum. The variable $dR_{b1,j3} + dR_{b2,j3}$ offers great discrimination for FSR from ISR and PU jets (see Section 4.2.1) as shown in Figure 4.8. FSR jets in the 3-jet category are selected by requiring:

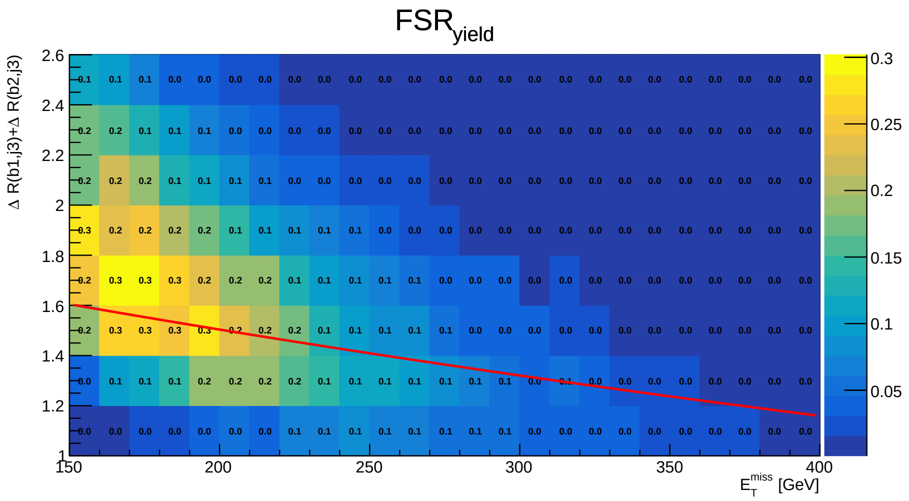
$$dR_{b1,j3} + dR_{b2,j3} < X \tag{5.1}$$

The above selection gives rise to an elliptical shape between the two b-jets and thus it is ideal for recovering FSR jets. The FSR selection can be further refined by exploring its dependence on the transverse momentum of the Higgs boson. Assuming a balanced system with $p_{\mathrm{T}}^{H} \sim p_{T}^{V}$, by plotting $dR_{b1,j3} + dR_{b2,j3}$ as a function of p_{T}^{V} in Figure 5.5, it is possible to deduce an exponential dependence, so the value X of the selection can be parametrised as:

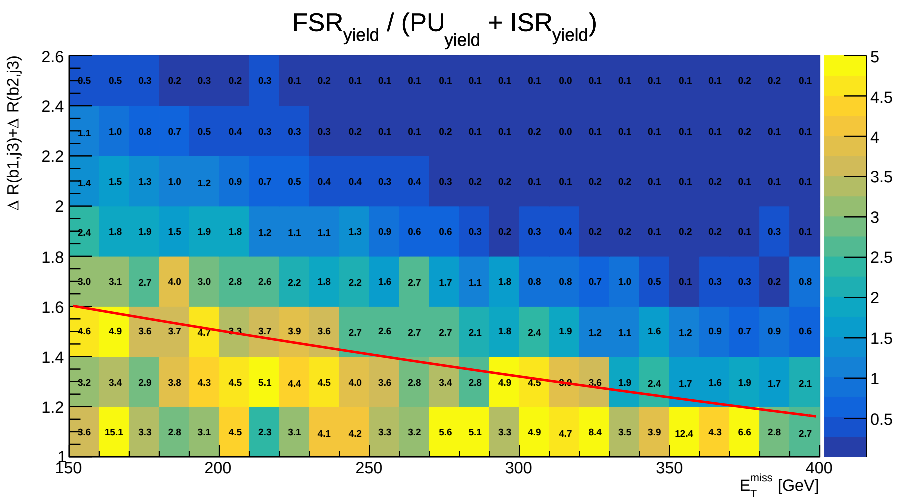
$$X = a + b \cdot e^{-c \cdot p_{T}^{V}} \tag{5.2}$$

where the parameters a, b and c can be tuned for the different lepton channels according to their respective Standard Model backgrounds. In particular, it has been decided to adopt a tighter cut in the 0-Lepton and 1-Lepton channels, due to large the $t\bar{t}$ background. In conclusion, the chosen coefficients for the FSR recovery selection are:

- For the 0-Lepton and 1-Lepton channels: $a = 0$, $b = 1.95$ and $c = 0.0013$;
- For the 2-Lepton channel: $a = 0.03$, $b = 3.55$ and $c = 0.004$;



(a) FSR events as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ for 3-jet events.



(b) Estimated sensitivity as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ for 3-jet events.

Figure 5.5: In the top plot, the $dR_{b1,j3} + dR_{b2,j3}$ distribution is shown as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ for events in the 0-Lepton 3-jet category where the third jet originates from FSR. In the bottom plot, the ratio of the yields between the FSR jets and the non-FSR jets is used to estimate a region with high purity of FSR jets. The red line represents the cut for the FSR selection in the 0-Lepton channel.

To summarize, the proposed FSR recovery strategy consists of migrating 3-jet events satisfying the selection in Equation 5.1 from the 3-jet category to the 2-jet category while adding back the third jet four-momentum to its closest b-jet.

The FSR recovery selection for the 0-Lepton channel is shown in Figure 5.5. Given the third-jet classification described in Section 4.2.1, the FSR recovery mis-classifies PU and ISR jets in $\sim 15\%$ of the cases and in $\sim 30\%$ of the cases the third jet is classified as *Low* p_{T} (mixture of FSR and ISR). This jet composition is fairly constant as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 0-Lepton channel as shown in Figure 5.6.

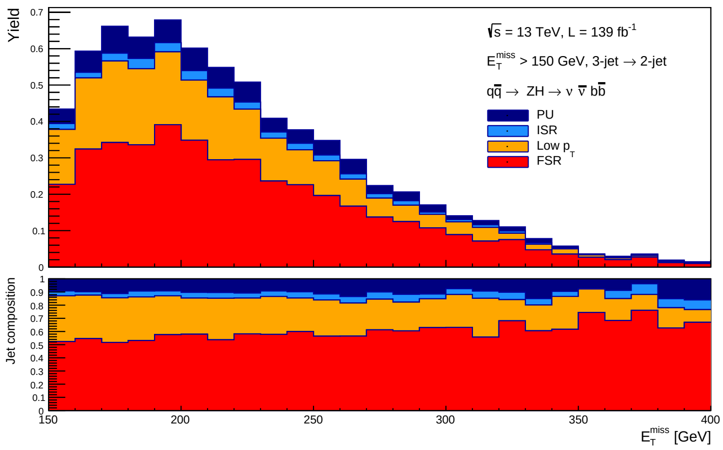


Figure 5.6: Third jet composition as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ for the selected events from FSR recovery in the 0-Lepton channel.

The FSR recovery is applied after the muon-in-jet and PtReco corrections in the 0-Lepton and 1-Lepton channels, and before the Kinematic Fitter in the 2-Lepton channel.

The improvement in the m_{bb} resolution for the events migrating from the 2-jet to the 3-jet category is shown in Figure 5.7. When the third jet is added to its closest b-jet, the standard deviation of the m_{bb} distribution reduces from 18.9 GeV to 14.9 GeV in the 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV category and from 22.7 GeV to 13.1 GeV in the 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$. Moreover, migrated events are more peaked toward the Higgs-boson mass.

The FSR recovery moves events from both signal and background processes from the 3-jet category to the 2-jet category. In Table 5.1, the variations of the event yields of the main processes in the 0-Lepton channel are reported. It is important to notice that the event migration observed for the two main backgrounds in the 0-Lepton channel, Z+jets and $t\bar{t}$, is smaller than for the signal processes.

The FSR recovery has also a positive impact on the analysis sensitivity since low-quality m_{bb} events from the low-significance 3-jet category are raised to the higher

Process	2-jet	2-jet FSR recovery	3-jet	3-jet FSR recovery
$q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}$	95.69	102.32 (+6.9%)	81.97	75.38 (−8.0%)
$gg \to ZH \to \nu\bar{\nu}b\bar{b}$	25.18	26.79 (+6.4%)	37.10	35.49 (−4.3%)
$q\bar{q} \to WH \to l\nu b\bar{b}$	32.83	35.67 (+8.7%)	32.88	30.06 (−8.6%)
$Z + b\bar{b}$	4883.50	5141.41 (+5.3%)	6976.18	6718.27 (−3.7%)
$W + b\bar{b}$	1126.55	1230.21 (+9.2%)	2377.47	2274.49 (−4.3%)
$t\bar{t}$	2482.24	2605.44 (+5.0%)	15122.29	15000.94 (−0.8%)

(a) 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.

Process	2-jet	2-jet FSR recovery	3-jet	3-jet FSR recovery
$q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}$	32.32	34.64 (+7.2%)	27.94	25.64 (−8.3%)
$gg \to ZH \to \nu\bar{\nu}b\bar{b}$	2.94	3.24 (+10.1%)	6.26	5.97 (−4.7%)
$q\bar{q} \to WH \to l\nu b\bar{b}$	7.59	8.13 (+7.2%)	6.84	6.31 (−7.9%)
$Z + b\bar{b}$	666.47	697.25 (+4.6%)	1189.57	1158.79 (−2.6%)
$W + b\bar{b}$	140.31	154.08 (+9.8%)	357.09	343.33 (−3.9%)
$t\bar{t}$	120.27	132.80 (+10.4%)	973.09	960.63 (−1.3%)

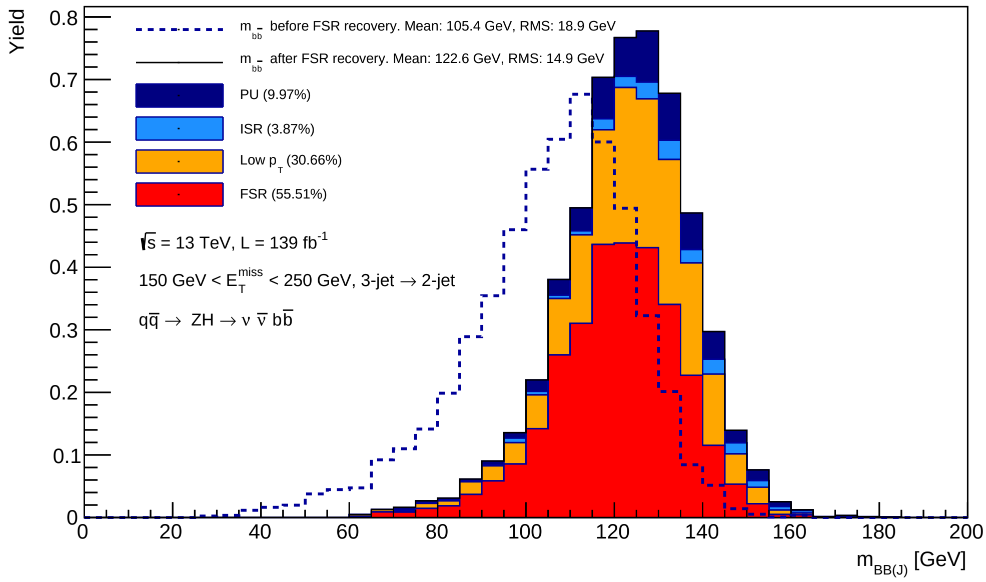
(b) 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

Table 5.1: Weighted event yield variations for the main signal and background processes in the 0-Lepton channel before and after applying the FSR recovery in the inclusive SRs+CRs phase space.

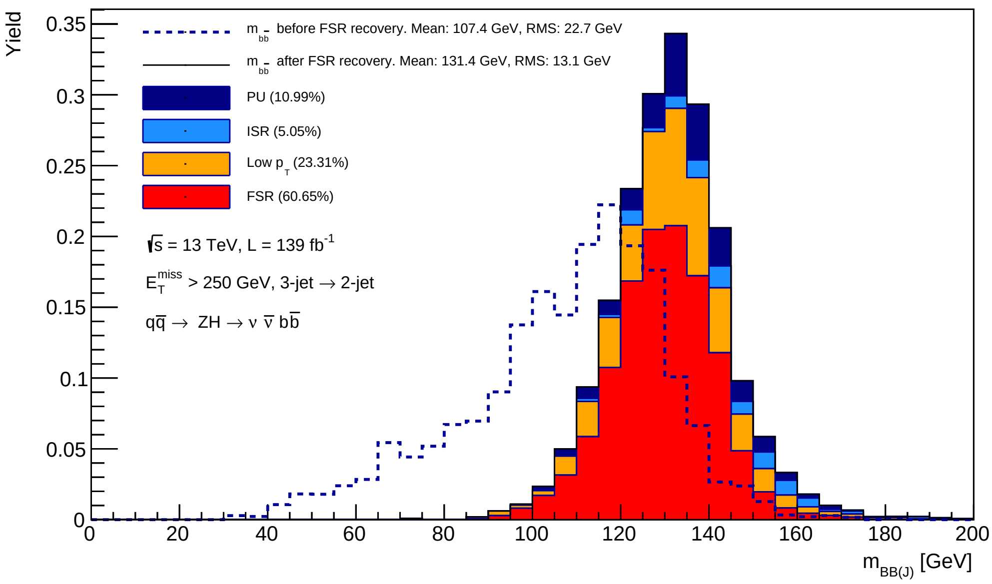
significance 2-jet category with an improved m_{bb}. The impact of FSR recovery has been evaluated using the simplified significance (defined in Equation 4.8) on the m_{bb} distribution with and without this correction. Improvements of 0.85%, 1.63% and 1.1% are observed in 0-Lepton, 1-Lepton and 2-Lepton channels respectively.

Furthermore, the FSR recovery plays an important role in the improvement of the matching of reconstructed and STXS categories when considering a splitting in N_{jets}. In Figure 5.8, signal events are shown with and without applying the FSR recovery. The migration matrix, which describes how events are moved across Reco and STXS categories, becomes more diagonal after the application of the FSR recovery.

In summary, the FSR recovery brings three main advantages: it corrects the m_{bb} toward the Higgs boson mass, improving the resolution, it also slightly improves the expected significance and it contributes to the improvement of the matching of the reconstructed and STXS categories.

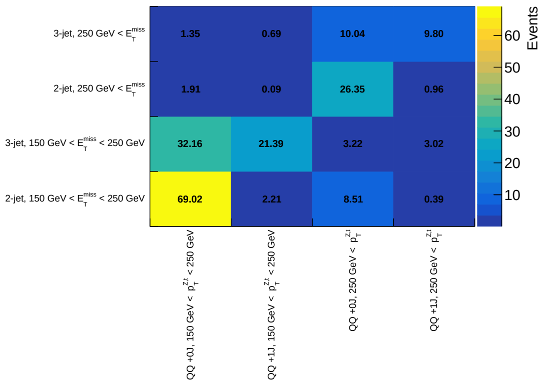


(a) 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.

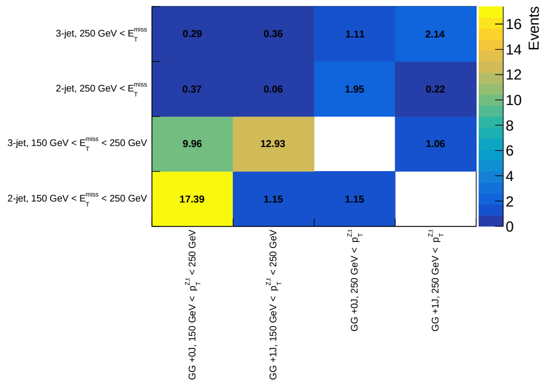


(b) 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

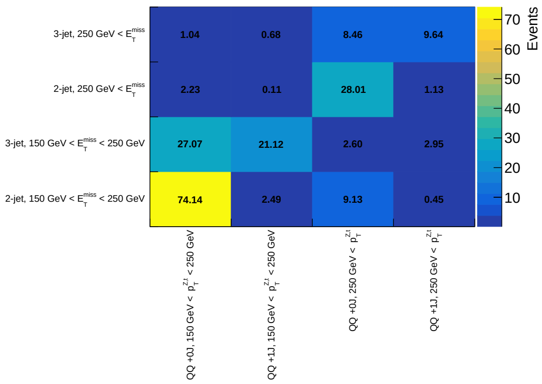
Figure 5.7: Signal events migrated from the 3-jet category to the 2-jet category in the 0-Lepton channel with the FSR recovery. The dashed line indicates the m_{bb} distribution for the uncorrected events while the coloured histogram shows the m_{bb} distribution when the third jet is added to its closest b-jet.



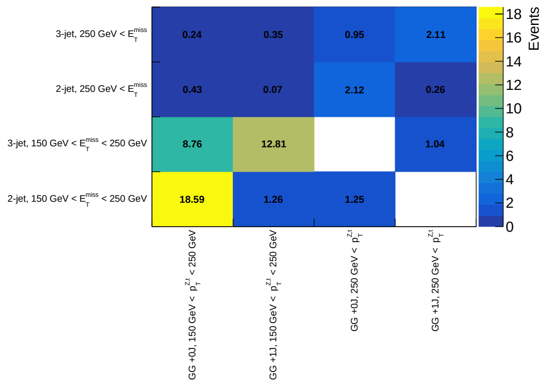
(a) $q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}$. No FSR recovery.



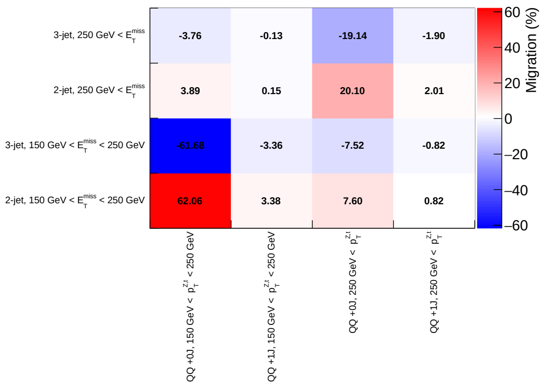
(b) $gg \to ZH \to \nu\bar{\nu}b\bar{b}$. No FSR recovery.



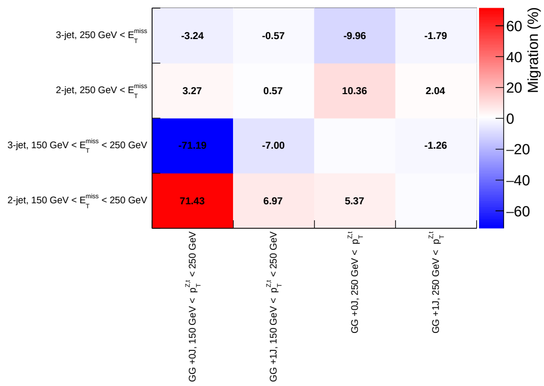
(c) $q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}$. With FSR recovery.



(d) $gg \to ZH \to \nu\bar{\nu}b\bar{b}$. With FSR recovery.



(e) $q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}$. Event migration.



(f) $gg \to ZH \to \nu\bar{\nu}b\bar{b}$. Event migration.

Figure 5.8: FSR recovery event migrations across jet categories for signal $q\bar{q}$ (left) and gg (right) initiated events. In the top plots, the weighted event yields are organised according to the reconstructed (y-axis) and STXS (x-axis) categories are shown without applying the FSR recovery. In the middle plots, the weighted event yields are shown after applying the FSR recovery. In the bottom plots, the event migrations are shown. The event migration for a bin ij is calculated as: $(Events_{ij,FSR} - Events_{ij,NoFSR})$ / $TotEventsMigrated$.

5.3 XGBoost

XGBoost is a powerful supervised machine learning open-source package based on gradient boosted decision trees [117]. This package is widely used by the machine learning community due to its performance and speed.

XGBoost creates a classification model by iteratively adding new decision trees during each training step. The new decision trees are chosen by minimizing an *objective function*

$$obj = L(\boldsymbol{y_i}, \boldsymbol{\hat{y}_i}) + \Omega(\boldsymbol{f}) \tag{5.3}$$

where L is the *loss function* that quantifies the discrepancy between the predicted values by the model during training $\boldsymbol{\hat{y}_i}$ and the true values $\boldsymbol{y_i}$, while Ω is a *regularization function* that depends on the trees, $\boldsymbol{f}$, in the model. The goal of the regularization function is to penalise unnecessarily complex models prone to over training.

In the $VH, H \to b\bar{b}$ analysis XGBoost has been tested as an alternative of the TMVA implementation of the BDT$_{VH}$ and BDT$_{VZ}$ for all the three lepton channels, and as a multi-classifier for the 0-Lepton channel.

XGBoost is characterised by a training procedure that automatically avoids over-training by dividing all the available events of every analysis region in three samples: training, validation and testing. For these studies, the training and validation datasets contain 40% and 10% of the available events, respectively, while the remaining 50% of the events are allocated to the testing dataset. The testing dataset is used to build the MC templates for BDT$_{VH}$ and BDT$_{VZ}$. In order to exploit the full statistics for building the MC templates, a second model, “fold”, is trained for every analysis region by inverting the training and validation samples with the testing samples, as shown in Figure 5.9. The training procedure described so far is analogous to the one adopted in TMVA in Section 4.10 with the exception of the presence of the validation sample.

The validation sample plays the key role in avoiding over-training. During each training round, the loss function is evaluated both for the training and the validation sam-

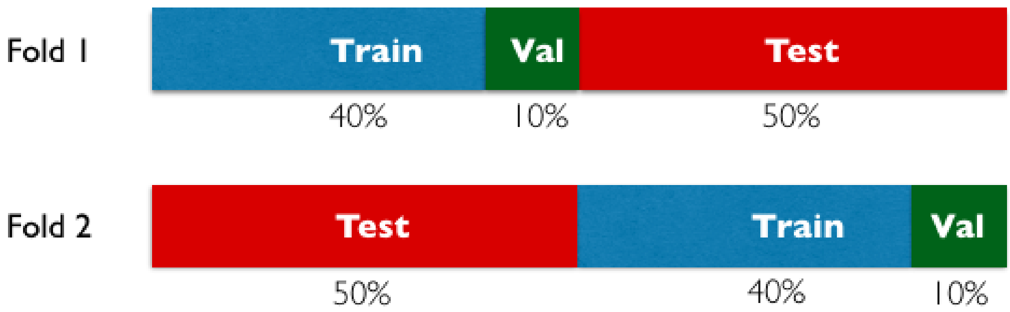


Figure 5.9: Training, validation and testing samples for each analysis region. A two-fold approach is used to exploit the full available statistics.

ples. However, at each stage of training, a new decision tree is added only based on the objective function. The result is shown in Figure 5.10, where by adding new decision trees, the loss function decreases, at first, very fast for both samples before plateauing for the validation sample. At this point, by including additional decision trees, the loss function would continue to decrease for the training sample because the model is learning fine details of that sample; however, the loss function of the validation sample would increase because the model has lost its best ability to predict on a new sample. Therefore, the best model to chose is the one for which the loss function of the validation sample reaches a minimum. XGBoost implements a functionality called *early stopping* that stops the training after N_{rounds} in which the loss function of the validation sample is not decreased; then XGBoost returns the best model in which the last N_{rounds} decision trees have been removed. For these studies, N_{rounds} has been set to 30.

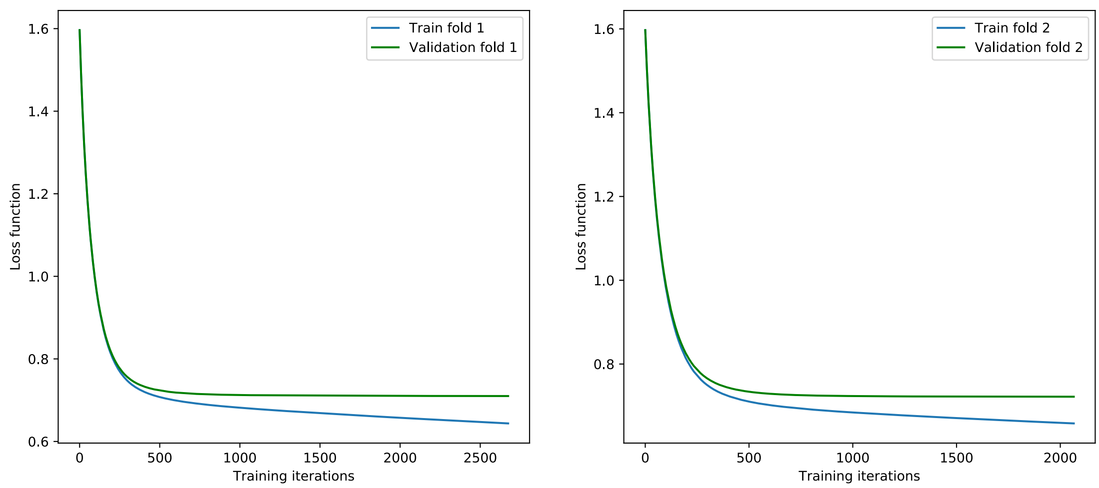


Figure 5.10: XGBoost over-training check for the 2-jet 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV category.

XGBoost also allows *sub-sampling* of the training dataset. Sub-sampling means that XGBoost would randomly sample a portion of the training dataset before growing a new tree. Once the training round has been completed, the training events are put together again before a new sub-sampling occurs. The sub-sampling procedure is also useful to mitigate over-training to avoid that XGBoost would put too much focus on certain events that might be outliers. In these studies, the training sub-sampling fraction has been set to 90%.

Sub-sampling can also be performed on the training features. In this context, before each training round, a fraction of the available input features to build a new tree is removed. Feature sub-sampling is particularly important when one or few features have significantly more discriminating power than others, such as the m_{bb}. In fact, by removing these powerful inputs features, XGBoost can better look into some more subtle correlations that might have been otherwise overshadowed. In these studies, the feature sub-sampling fraction has been set to 90% to remove at least one feature per training round.

XGBoost offers a series of hyper-parameters that can be specifically tuned [131]. In these studies, the ones that have been considered are

- `learning rate`. This parameter reduces the influence of the new added trees during training (default value = 0.3, range: $[0, 1]$).
- `max depth tree`. This consists of the maximum number of splits for a single new tree (default value = 6, range: $[1, \infty]$).
- `gamma`. This is the minimum loss reduction necessary to make a new split in the tree and increase its depth. (default value = 0, range: $[0, \infty]$).
- `min child weight`. This plays a similar role to gamma, with the difference that it bases its decision on the Hessian of the loss function (default value = 1, range: $[0, \infty]$).

For the multi-classifier in the 0-Lepton channel, the optimal values for the hyper-parameters are searched through a grid scan in a region of the parameter phase space near the default values, as shown in Section 5.5.

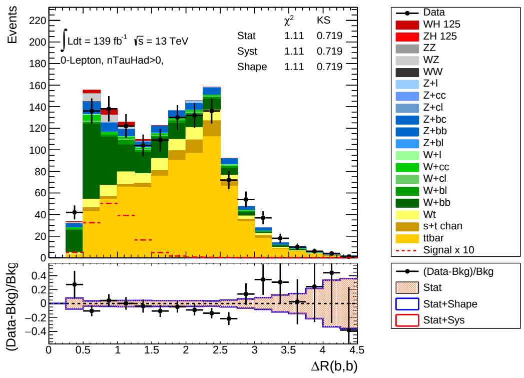
5.4 New variables for the MVA

In order to improve the analysis sensitivity in the 0-Lepton channel, new variables have been tested in the MVA. For these studies, the baseline is an XGBoost signal versus background configuration, analogous to BDT$_{VH}$ in Section 4.10, with all the input features described in Table 4.8, except for $p_{\mathrm{T}}^{\mathrm{miss,st}}$, MV2($b_1$) and MV2($b_2$).

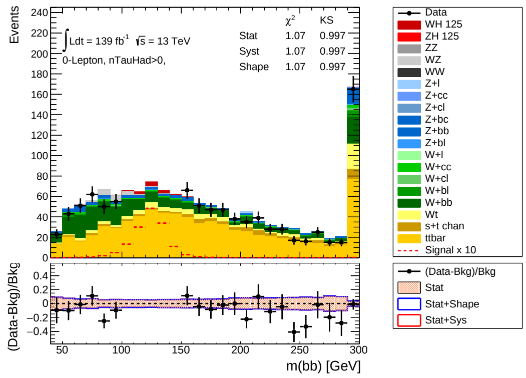
The training settings considered for the four analysis regions in the 0-Lepton channel are:

1. Baseline input variables.
2. Baseline input variables plus $p_{\mathrm{T}}^{\mathrm{miss,st}}$.
3. Baseline input variables plus pseudo-continuous b-tagging quantiles MV2(b_1) and MV2(b_2).
4. Baseline input variables plus number of hadronic taus $nTaus$.
5. Baseline input variables plus $dRBBJ$ in the 3-jet category.
6. Baseline input variables minus $|\Delta\eta(b_1, b_2)|$ and m_{eff}.

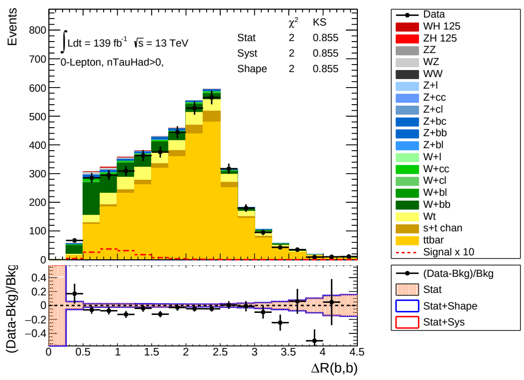
The effectiveness of $p_{\mathrm{T}}^{\mathrm{miss,st}}$, MV2($b_1$) and MV2($b_2$) has already been discussed in Section 4.10.1. Events with at least one hadronic taus allow a very interesting phase space to be investigated, in which the Z+jets background is almost completely absent and the signal originates mostly from $WH \to l\nu b\bar{b}$. The $\Delta R(b_1, b_2)$ and m_{bb} distributions for 0-Lepton events with at least one hadronic tau are shown in Figure 5.11. It is interesting to notice that within this corner of phase-space, it is possible to obtain a good separation for $t\bar{t}$ and W+jets with $\Delta R(b_1, b_2)$. Moreover, the $t\bar{t}$ background in this phase-space is often characterised by a high BDT_{VH}, given that here its m_{bb} peaks near 125 GeV. The $dRBBJ \equiv \Delta R(b_1 + b_2, j_3)$ variable shown in Figure 5.12 is calculated, in the 3-jet category, as the angular separation between the third jet and the bb-system in order to introduce some angular information about the third jet. Lastly, a configuration without the variables $|\Delta\eta(b_1, b_2)|$ and m_{eff} has been tested to check whether XGBoost would particularly benefit from these variables, given that they are very correlated with the others (Appendix D).



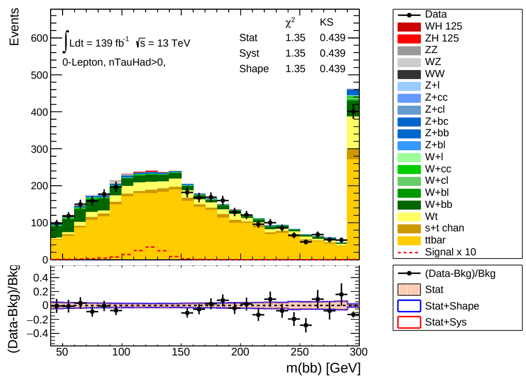
(a) $\Delta R(b_1, b_2)$ 2-jet.



(b) m_{bb} 2-jet



(c) $\Delta R(b_1, b_2)$ 3-jet.



(d) m_{bb} 3-jet.

Figure 5.11: $\Delta R(b_1, b_2)$ (left) and m_{bb} (right) for 0-Lepton 2-jet (top) and 3-jet (bottom) events with at least one hadronic tau for the inclusive SR+CR. The last bin in the plots contains the overflow.

The area under the ROC curves for the six training settings described above are reported in Table 5.2 for every analysis region. Furthermore, it is also reported a training configuration that combines all the six training settings. All new variables bring improvements to the ROC curves in at least one analysis region. The pseudo-continuous b-tagging quantiles are the most effective new variables while the $dRBBJ$ is the least effective. Instead, the removal of $|\Delta\eta(b_1, b_2)|$ and m_{eff} does not visibly degrade the

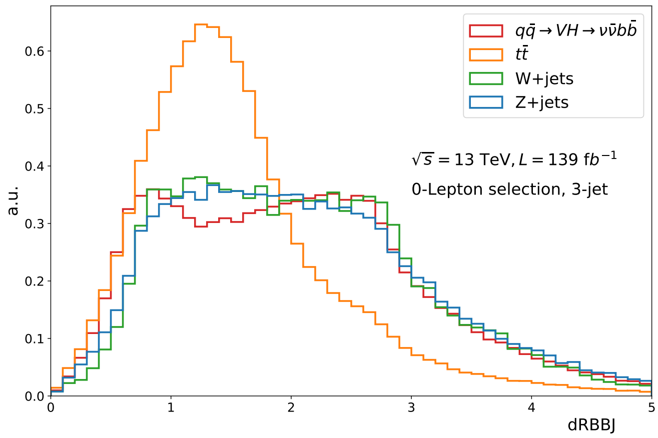


Figure 5.12: dRBBJ shapes for the main signal and background processes in the 0-Lepton 3-jet category.

sensitivity. In conclusion, for future iterations of the analysis, the combined configuration setting could provide the best improvement in the sensitivity. For the analysis described in Chapter 4, only $p_{\mathrm{T}}^{\mathrm{miss,st}}$, MV2($b_1$) and MV2($b_2$) have been included due to time constraints.

Features	2-jet 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV	2-jet 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$
Baseline	0.915	0.930
Baseline plus $p_{\mathrm{T}}^{\mathrm{miss,st}}$	0.915	0.931
Baseline plus MV2(b_1) and MV2(b_1)	0.918	0.932
Baseline plus $nTaus$	0.915	0.931
Baseline minus m_{eff} and $\|\Delta\eta(b_1, b_2)\|$	0.915	0.930
Baseline plus all the above	0.919	0.933

Features	3-jet 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV	3-jet 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$
Baseline	0.892	0.919
Baseline plus $p_{\mathrm{T}}^{\mathrm{miss,st}}$	0.893	0.920
Baseline plus MV2(b_1) and MV2(b_1)	0.895	0.921
Baseline plus $nTaus$	0.895	0.920
Baseline plus $dRBBJ$	0.892	0.920
Baseline minus m_{eff} and $\|\Delta\eta(b_1, b_2)\|$	0.892	0.919
Baseline plus all the above	0.898	0.923

Table 5.2: Area under the receiver operating characteristic curves for the training settings in the 0-Lepton channel analysis regions.

5.5 Multi-classifier

The signal and control regions defined with the continuous cuts in Equation 4.6 and Equation 4.7 allow three analyses in a single phase space to be performed, as shown in Section 4.13. However, these control regions in the 0-Lepton channel are not very pure as can be seen in Figure 4.19. Inspired by the CMS $VH, H \to b\bar{b}$ observation publication [89], a multi-classifier based on the XGBoost package [117] has been tested with the goal of creating very pure control regions without sacrificing analysis sensitivity.

The CMS multi-classifier is a deep neural network that aims at distinguishing five background processes: $V + b\bar{b}$, $V + b$, $t\bar{t}$, $V + light$ and single top production. This is very challenging because the 0-Lepton channel is characterised by few physics objects: two b-jets and one optional additional non-b-jet.

In this section, the possibility of replacing the Low-ΔR and High-ΔR CRs with two control regions defined by a multi-classifier is explored. The first control region would aim at isolating the W+jets background, while the second would target the top processes, both $t\bar{t}$ and single-top production. The top and W+jets backgrounds are respectively the second and third largest backgrounds in the 0-Lepton channel, therefore dedicated control regions should be designed for them. For Z+jets, the largest 0-Lepton background, no control region is needed since the low BDT$_{VH}$ bins in the SR (top row of Figure 4.31 and Figure 4.32) already act as de facto control regions.

The XGBoost multi-classifier has been designed with five outputs to distinguish five processes: signal VH, top, W+jets, Z+jets and diboson. This choice has provided good flexibility during the studies. Individual multi-classifier trainings have been assigned to every combination of jet category and p_{T}^{V} regions. The over-training has been handled using the procedure detailed in Section 5.3. The ntuples used for training were generated applying the FSR recovery strategy described in Section 5.2.

The input variables for the multi-classifier are shown in Table 5.3. Compared to the input variables of BDT$_{VH}$ described in the previous chapter, $|\Delta\eta(b_1, b_2)|$ and m_{eff} have been removed due to the high correlations with other variables (Appendix D) and the non-visible improvement in the area under the ROC curve for XGBoost. Conversely, the number of hadronic taus in the event *nTaus* has been included to provide additional discrimination for W+jets and $t\bar{t}$. Furthermore, for the 2-jet category, a flag called *hasFSR* indicates whether the event has been migrated from the 3-jet category. Lastly, in the 3-jet category, the angular separation between the third jet and the bb-system, *dRBBJ*, has been added.

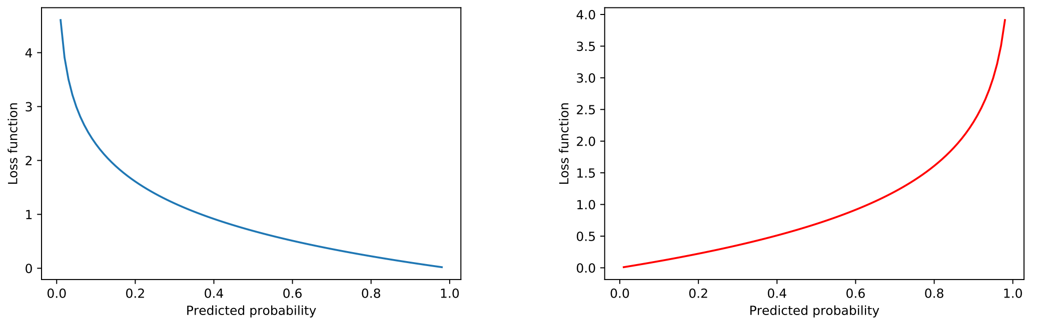
Variable	0-lepton
m_{bb}	×
$\Delta R(b_1, b_2)$	×
$p_T^{b_1}$	×
$p_T^{b_2}$	×
$\Delta\phi(V, bb)$	×
MV2(b_1)	×
MV2(b_2)	×
$p_T^{\text{miss,st}}$	×
E_T^{miss}	×
$nTaus$	×
	Only in 2-jet events
$hasFSR$	×
	Only in 3-jet events
$p_T^{\text{jet}_3}$	×
m_{bbj}	×
$dRBBJ$	×

Table 5.3: Input variables used for the training and evaluation of the multi-classifier in the 0-Lepton channel.

The loss function minimised during the training was the *multidimensional cross-entropy* defined as:

$$MCE = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{i,j} \ln p_{i,j} \tag{5.4}$$

where N is the total number of events for training, M the number of classes for the physics processes, $p_{i,j}$ is the model's probability of assigning label j to the event i, and $y_{i,j}$ is the truth label of the event. For example, assuming that an event i is a W+jets event, $y_{i,j}$ has the values $y_{i,j=W+jet} = 1$ and $y_{i,j\neq W+jet} = 0$. The MCE measures the performance of a classification model whose output is a probability $p_{i,j}$ [132], as shown in Figure 5.13 for the case of only two classes ($M = 2$).



(a) Cross-entropy when true label $y = 1$. (b) Cross-entropy when true label $y = 0$.

Figure 5.13: Cross-entropy as a function of the estimated probability from the classifier.

The MCE is an effective loss function when the output classes have similar importance. However, in the $VH, H \to b\bar{b}$ analysis, correctly recognising the tiny VH signal is far more important than accurately classifying the backgrounds. For this reason, a weight proportional to the MC weight of the event is introduced in the MCE. For a VH signal event, its weight is $w_i = w_{i,MC} / \sum_j^{Signal} w_{j,MC}$ while for a background event its weight is $w_i = w_{i,MC} / \sum_j^{Totbkg} w_{j,MC}$.

Optimal hyper-parameters have been searched through an extensive *grid search*, which is an exhaustive scan through a fixed subset of the hyper-parameters. The considered hyper-parameters are `max depth` (values = 4, 5, 6, 7, 8, 9), `gamma` (values = 0, 0.2, 0.4), `learning rate` (values = 0.01, 0.05, 0.3) and `min child weight` (values = 1, 3, 5). The best set of hyper-parameters was chosen based on the lower value of the MSE obtained during training in the validation sample. The best hyper-parameters for each training are shown in Table 5.4.

Once trained, the five output scores of the multi-classifier are shown in Figure 5.14 for the the 2-jet 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV category. Moreover, the feature importance in the multi-classifier is shown in Figure 5.15, where m_{bb} is observed to be the most important variable as expected.

As previously mentioned, the idea behind this multi-classifier is to construct two control regions: one for W+jets and one for top processes. Therefore, the events are categorized in the following way:

Hyper-parameter	2-jet	3-jet
`max depth`	6	8
`gamma`	0	0.2
`learning rate`	0.01	0.01
`min child weight`	3	1

(a) 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV .

Hyper-parameter	2-jet	3-jet
`max depth`	6	6
`gamma`	0	0.2
`learning rate`	0.01	0.01
`min child weight`	1	3

(b) 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

Table 5.4: Hyper-parameters chosen for the multi-classifier trainings after the grid search.

- If p_W is the largest score of the five multi-classifier's outputs, the event is categorised in the W+jets CR.
- If p_{Top} is the largest score of the five multi-classifier's outputs, the event is categorised in the Top CR.
- In all the other cases, the event is categorised in the SR.

Furthermore, as before, in the CRs the event yield is the variable fitted, whereas in the SRs p_{Signal} is the variable fitted (analogous to BDT$_{VH}$).

For the signal VH, the new SRs and CRs behave in a similarly as the previous regions defined upon the $\Delta R(b_1, b_2)$ continuous cuts, as it is possible to see in Figure 5.16. However, the new regions are characterised by a significantly higher purity, as it is shown in Figure 5.17. Therefore, the new CRs act as "enhanced" Low-ΔR-CRs and High-ΔR-CRs, which would simplify their adoption process into the analysis chain.

To test the goodness of these new regions, a very simple fit model has been used. A 0-Lepton stand-alone fit where all the main backgrounds have been left floating independently between the 2-jet and 3-jet categories has been considered. In this fit

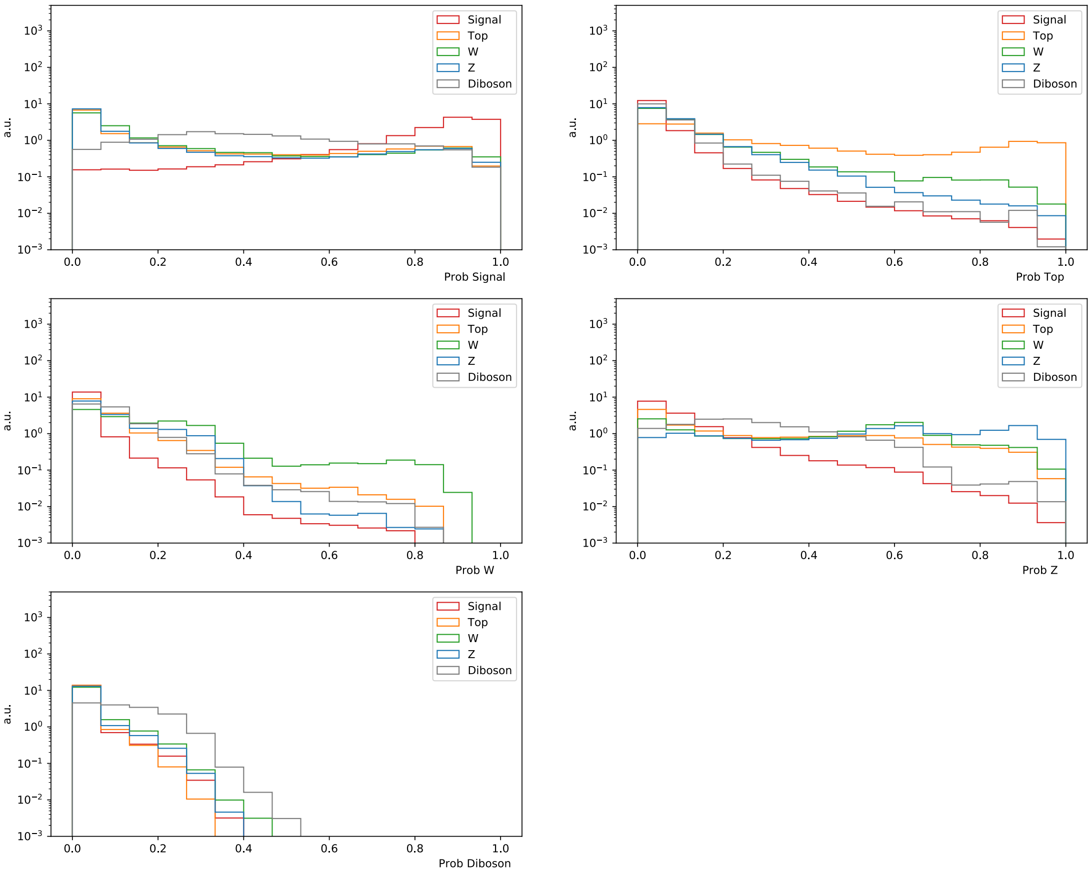
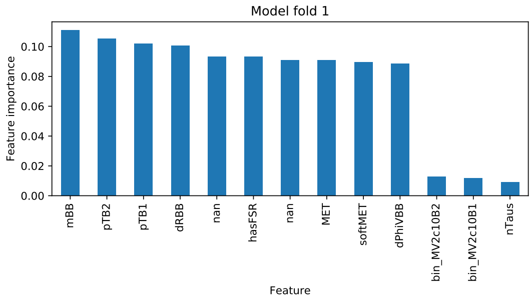


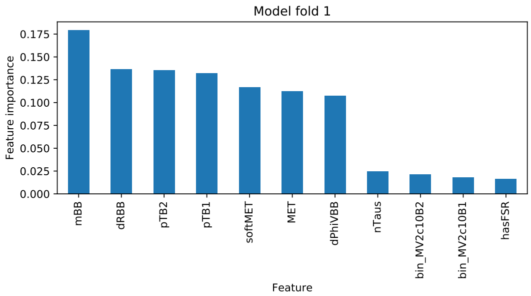
Figure 5.14: Multi-classifier output scores for the five predicted processes for the 2-jet 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV category. The colours of the lines indicate the truth label of the processes.

model, all the experimental and modelling systematics have been neglected, just to focus on the interplay between the various backgrounds. The post-fit distributions for the conditional fit with $\mu = 1$ are shown in Figure 5.18 and Figure 5.19. The pull plot in Figure 5.20 and, in particular, the correlation matrices in Figure 5.21 indicate that, for the same fit model, the new regions are able to significantly reduce the correlations of the normalisation factors compared to the previous ones. Lastly, there is a small improvement in the expected significance of $\sim 2\%$ with the multi-classifier.

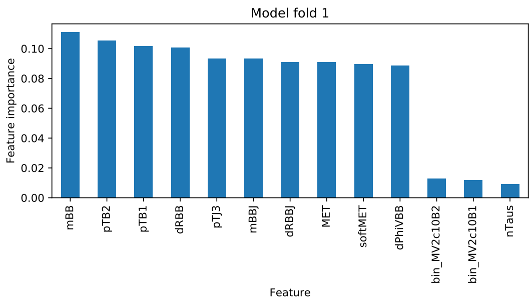
In conclusion, the XGBoost multi-classifiers provides a valuable alternative to the continuous $\Delta R(b_1, b_2)$ SRs and CRs, which is makes further exploration with a more complete fit model worthwhile.



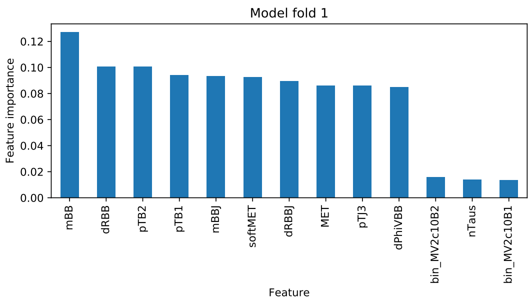
(a) 2-jet.



(b) 2-jet



(c) 3-jet.



(d) 3-jet.

Figure 5.15: Importance of the input variables used in the trainings. The height of each column is proportional to the number of times a variable is used inside the multi-classifier.

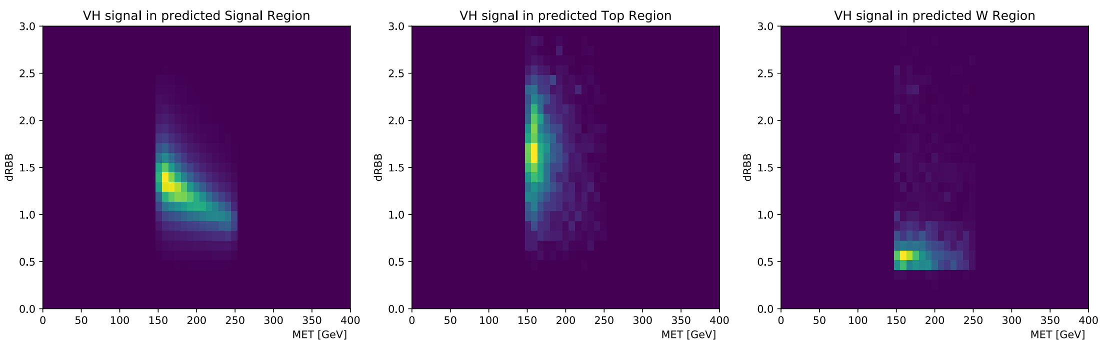


Figure 5.16: $dR(b_1, b_2)$ as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 2-jet 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV category for the VH signal in the Signal, Top and W predicted regions.

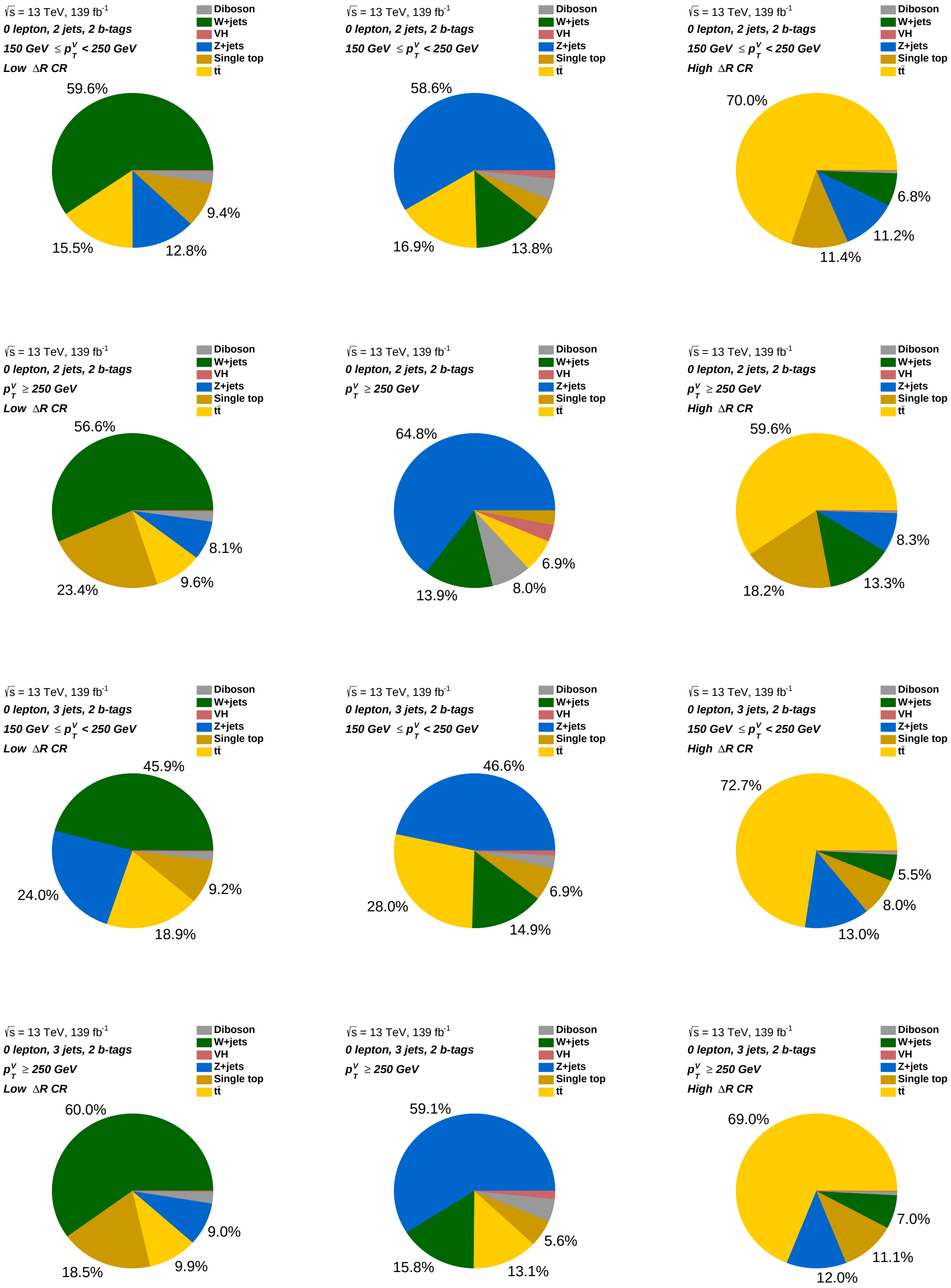
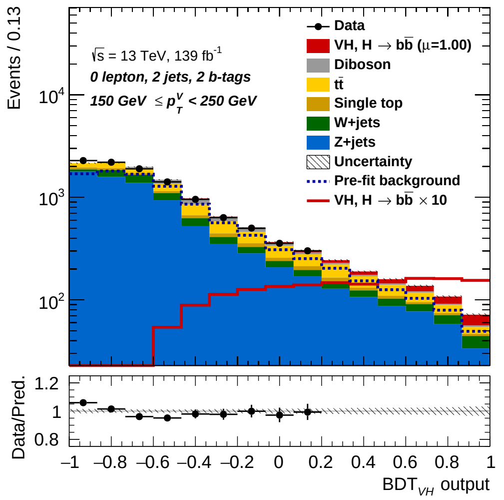
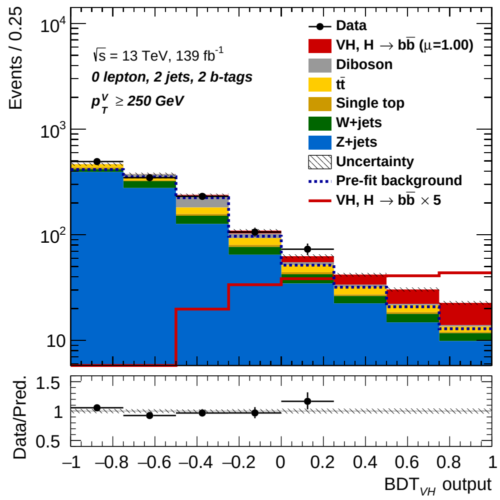


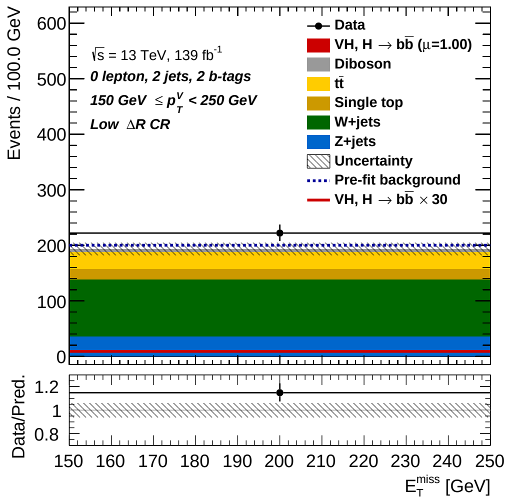
Figure 5.17: Pre-fit 0-Lepton background composition in signal and control regions defined with the XGBoost multi-classifier. From left to right: Low-ΔR-CR, SR and High-ΔR-CR. From top to bottom: 2-jet 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV category, 2-jet 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$ category, 3-jet 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV category and 3-jet 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$ category. Values below 5% are not reported on the pie charts.



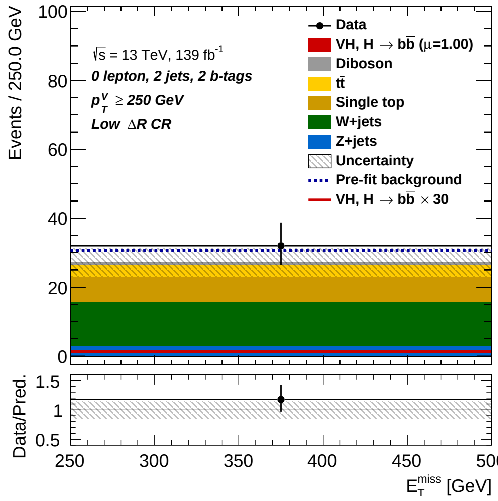
(a) SR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



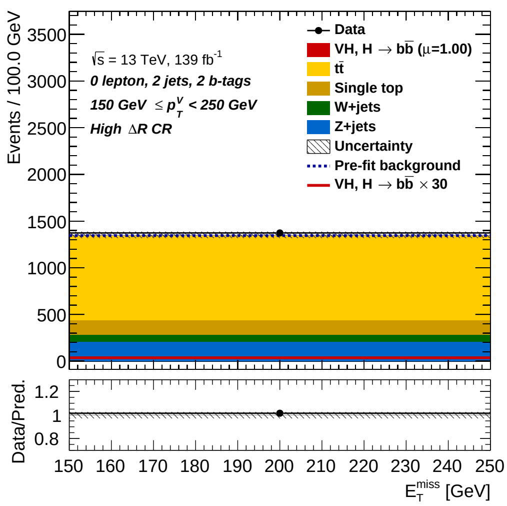
(b) SR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.



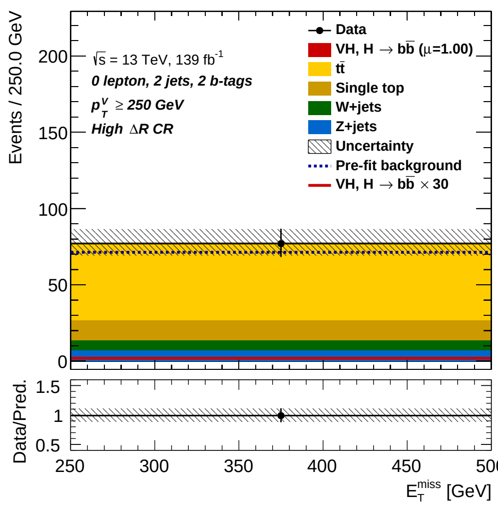
(c) W-CR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



(d) W-CR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

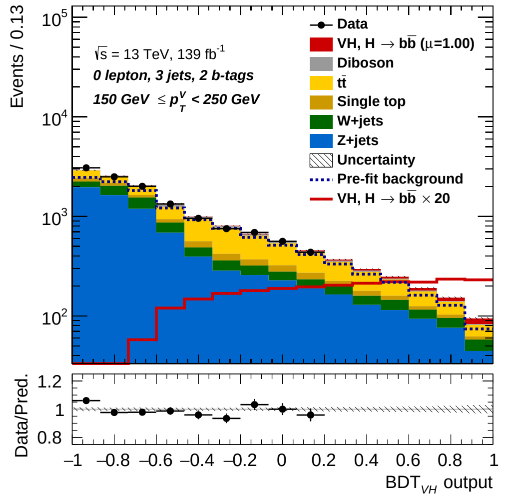


(e) Top-CR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.

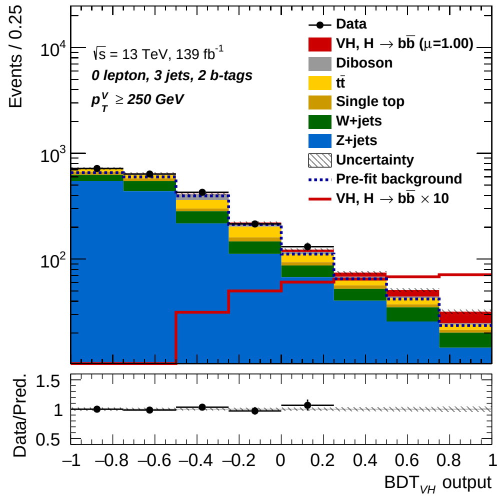


(f) Top-CR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

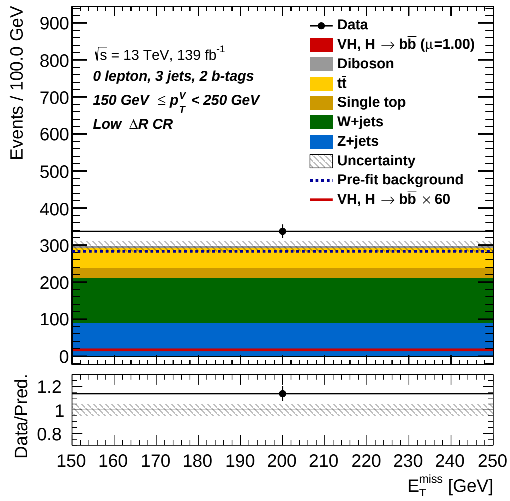
Figure 5.18: The top row shows the conditional post-fit p_{Signal} output distributions of the multi-classifier in the new SRs. The middle and bottom rows show the post-fit yield distributions for the W and Top CRs defined with the multi-classifier. All plots refer to the 0-Lepton channel 2-jet category. The fit model is a standalone simplified 0-Lepton fit with only the MC statistics and the floating normalisations as nuisance parameters.



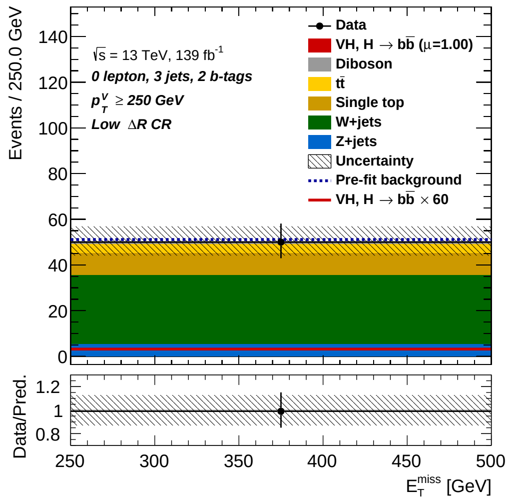
(a) SR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



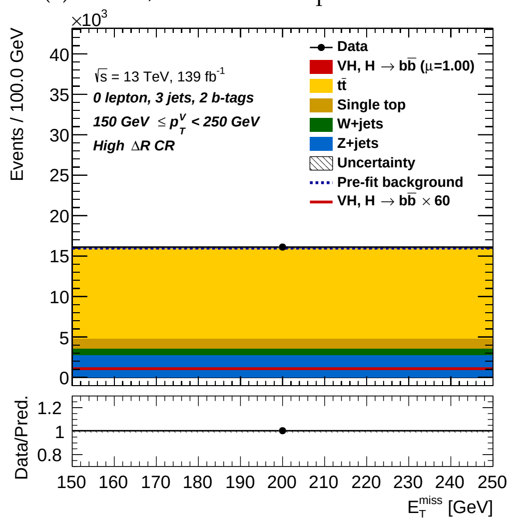
(b) SR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.



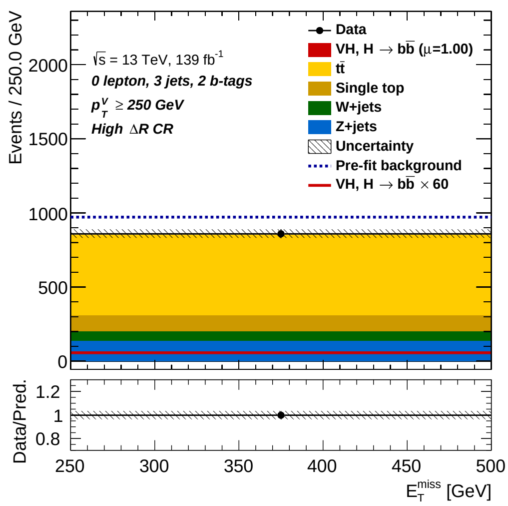
(c) W-CR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



(d) W-CR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.



(e) Top-CR, 150 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 250 GeV.



(f) Top-CR, 250 GeV $< E_{\mathrm{T}}^{\mathrm{miss}}$.

Figure 5.19: The top row shows the conditional post-fit p_{Signal} output distributions of the multi-classifier in the new SRs. The middle and bottom rows show the post-fit yield distributions for the W and Top CRs defined with the multi-classifier. All plots refer to the 0-Lepton channel 3-jet category. The fit model is a standalone simplified 0-Lepton fit with only the MC statistics and the floating normalisations as nuisance parameters.

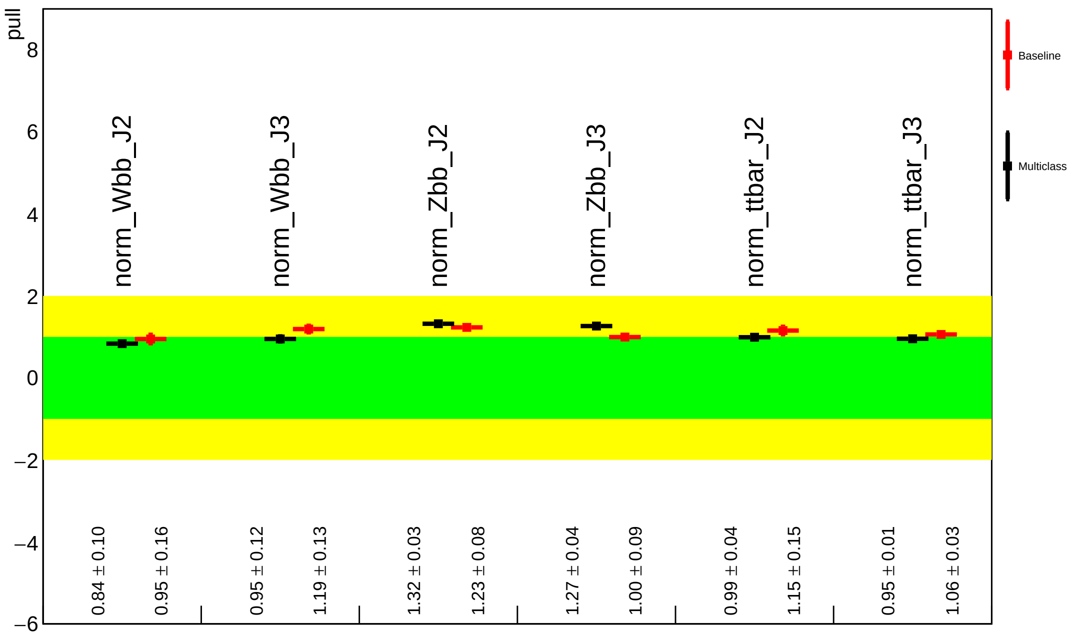
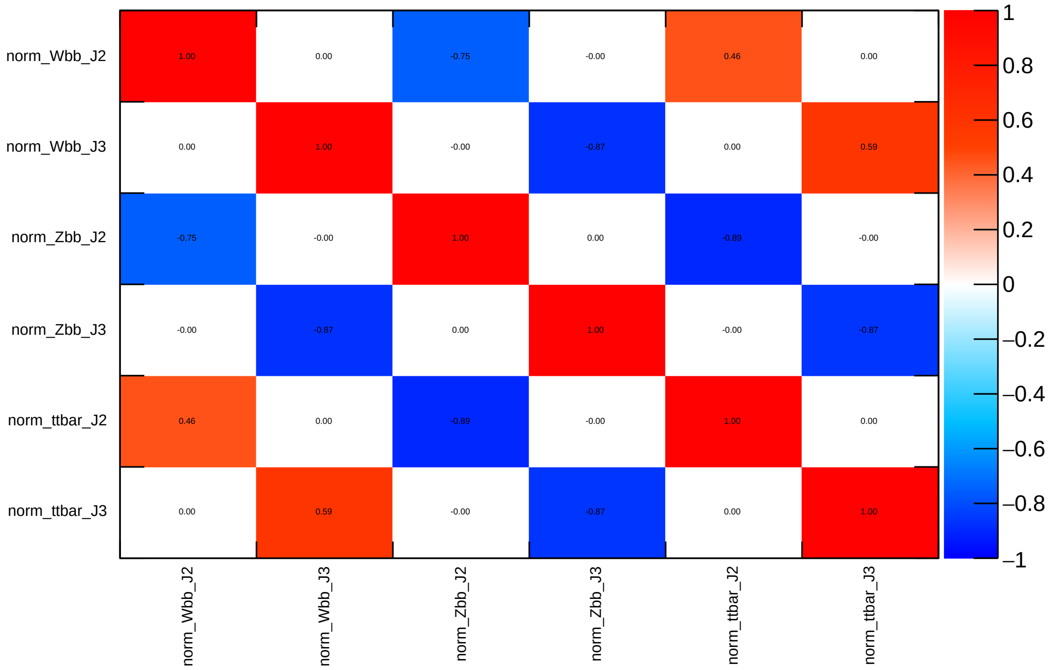
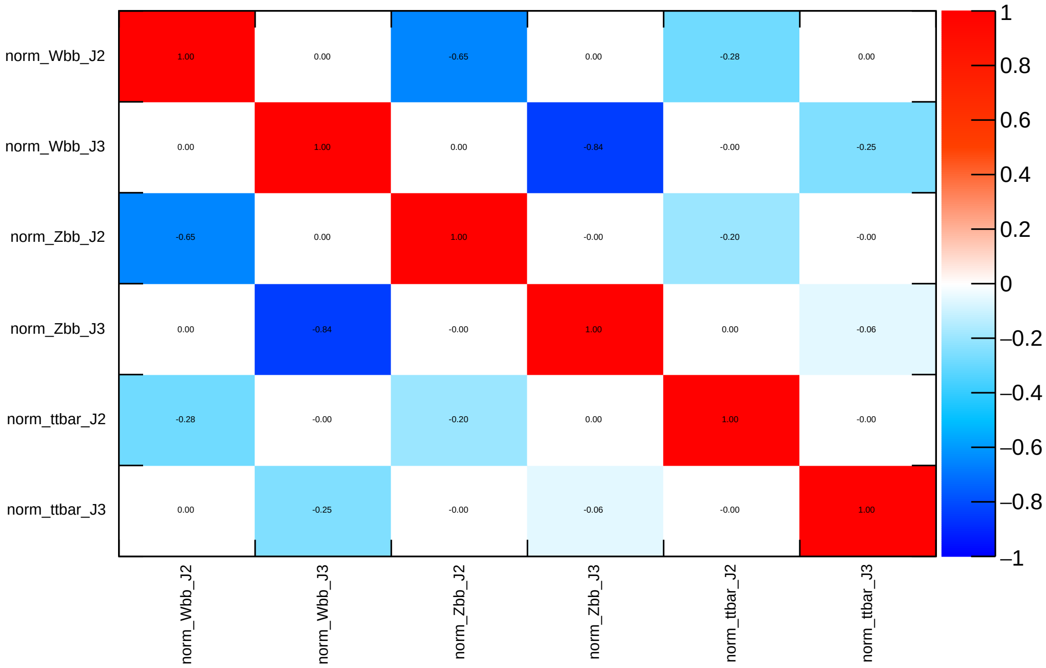


Figure 5.20: Floating normalisations for the simplified conditional 0-Lepton fit for the multi-classifier and the analysis inputs described in Chapter 4.



(a) Baseline 0-Lep.



(b) Multi-classifier.

Figure 5.21: Correlation matrices for the floating normalisations for the simplified conditional 0-Lepton fit for the multi-classifier and the analysis inputs described in Chapter 4.

Chapter 6

Conclusions

Measurements of the Higgs boson decay to beauty quarks play an important role in testing the Standard Model predictions and exploring Beyond the Standard Model theories. In this thesis, the $VH, H \to b\bar{b}$ analysis has been presented using the full Run 2 ATLAS dataset.

The $VH, H \to b\bar{b}$ signal strength has been measured with a global likelihood fit across the three lepton channels: $\mu^{b\bar{b}}_{VH} = 1.02^{+0.18}_{-0.17} = 1.02^{+0.12}_{-0.11}(\text{stat.})^{+0.14}_{-0.13}(\text{syst.})$. This result corresponds to a rejection of the background-only hypothesis with a significance of 6.7 standard deviations, while the expected significance was also 6.7 standard deviations. The $WH, H \to b\bar{b}$ and $ZH, H \to b\bar{b}$ signal strengths have also been separately measured: $\mu^{b\bar{b}}_{WH} = 0.95^{+0.27}_{-0.25} = 0.95 \pm 0.18(\text{stat.})^{+0.19}_{-0.18}(\text{syst.})$ and $\mu^{b\bar{b}}_{ZH} = 1.08^{+0.25}_{-0.23} = 1.08 \pm 0.17(\text{stat.})^{+0.18}_{-0.15}(\text{syst.})$. The former observed and expected significances are respectively 4.0 and 4.1, which indicate strong evidence for the WH production, while the latter significances are respectively 5.3 and 5.1, which allow the observation of ZH production mode to be claimed. Cross-sections of the associated production of a Higgs boson decaying into beauty quarks and a vector boson have been measured as a function of the vector-boson transverse momentum in five bins using the STXS framework. The total uncertainties on each bin ranged from 85% to 30% for the high p_T^V bins. Two cross-check analyses have been performed to validate the $VH, H \to b\bar{b}$ fit strategy. The first one was a $VZ, Z \to b\bar{b}$ multivariate analysis in the same phase-space of the

$VH, H \to b\bar{b}$ analysis, which allowed a careful validation of the fit model. The second one extracted the $VH, H \to b\bar{b}$ signal strength from a fit of the invariant mass of the b-jets. All the previous results are compatible with the Standard Model predictions.

For future iterations of the $VH, H \to b\bar{b}$ measurements, several new analysis techniques have been tested. Particle flow jets offer a better energy resolution, although pile-up rejections prescriptions were not yet ready at the time of the full Run 2 analysis. FSR recovery gives a better m_{bb} resolution and slightly improves the analysis sensitivity; more importantly for future differential measurements in N_{jets}, this technique brings the reconstructed categories closer to the Simplified Template Cross-Sections. Lastly, a multi-classifier in the 0-Lepton channel could significantly improve the purity of the control regions. Overall, these new methods have the potential to make the analysis more sensitive and resilient.

In conclusion, with an upcoming increasing dataset and new methods at our disposal, our understanding of the Higgs sector is only destined to deepen.

Acknowledgements

I owe a debt of gratitude to many people for all their encouragement, advice and support along the journey of my doctorate.

I would like to thank my supervisor Daniela for giving me the opportunity of joining the ATLAS Higgs group, for her guidance and ideas during the DPhil and for entrusting me with intellectual freedom through my later research years. It has been a privilege and an honour to be your student.

Special thanks must go to the ATLAS Oxford group with whom I shared many insightful conversations and from whom I received constructive feedback. This list is by no means exhaustive but does include Ian, Giacomo, Elisabeth, Luigi V, Mariyan, Fikri, Luigi M, Jesse, Cecilia, Miha, Santiago, Maria Giovanna, Ricardo, Maria, Philipp, Yingjie and Siyuan.

My most sincere thanks go ATLAS Hbb group for being an excellent example of collaborative teamwork. I have learnt a lot about the constructive exchange of ideas, setting high goals and delivering timely results. In particular, I owe a huge amount to Valerio that has inspired and guided the majority of the studies in this thesis. He has been an immense source of knowledge and *pain au chocolat* during intense moments.

I would like to thank the CMS Bologna group for introducing me to Particle Physics. A warm thank to my ex-supervisor Daniele who has always been there for me since the beginning of my studies.

I would like to express my gratitude to the University College, the Sub-Department of Particle Physics, the Clarendon Council and the Science and Technology Facilities Council for financially supporting my studies.

Last but not least, none of this would have been possible without the love and help of my family, in Italy and Colorado, and Lucia.

Appendix A

Weights for filtered samples

Combining events in the most effective way from an inclusive sample and a filtered sample can be thought as an optimization problem. In this sense, the goal of the optimization is reducing the overall statistical error. The purpose of this appendix is to show the complete calculation for the $t\bar{t}$ filtered sample coefficients described in Section 4.9.1.

For the slice 100 GeV $<$ truth $E_{\mathrm{T}}^{\mathrm{miss}}$ $<$ 200 GeV, the factor 0.5 was chosen in the following manner. Given an histogram bin i with M_i entries with weights w_j, its error is

$$\sigma_i = \sqrt{\sum_j^M (w_j)^2} \tag{A.1}$$

since the bin content follows a Poisson distribution. When new events are generated in the slice 100 GeV $<$ truth $E_{\mathrm{T}}^{\mathrm{miss}}$ $<$ 200 GeV, N events with weight w_k will enter in bin i. However, in order to keep the bin yield constant y, the weights w_j and w_k need to be scaled by two factors a and b, such that

$$y = a \cdot \sum_j^M w_j + b \cdot \sum_k^N w_k = \sum_j^M w_j = \sum_k^N w_k \tag{A.2}$$

The values a and b, as long as they satisfy Equation A.2, can be chosen arbitrarily. However, they can be optimised to reduce the error in Equation A.1. In fact, Equation

A.1 can be rewritten as

$$\sigma_i(a,b) = \sqrt{a^2 \cdot \sum_j^M (w_j)^2 + b^2 \cdot \sum_k^N (w_k)^2} \tag{A.3}$$

At this point, it is possible to calculate the minimum of the function in Equation A.3 on the constrain of Equation A.2:

$$a \cdot \sum_j^M w_j + b \cdot \sum_k^N w_k = \sum_j^M w_j \tag{A.4}$$

by introducing a curve $\overrightarrow{\gamma}(t) : \mathbb{R}^1 \to \mathbb{R}^2$ that parametrises the constrain:

$$\overrightarrow{\gamma}(t) = \left(t, \frac{\sum_j^M w_j}{\sum_k^N w_k} \cdot (1-t) \right) \tag{A.5}$$

and since a and b are required to be positive, it follows that $0 < t < 1$. Substituting Equation A.5 in Equation A.3 and calculating the stationary points with respect to t gives

$$\frac{d}{dt} \left[\sigma_i(\overrightarrow{\gamma}(t)) \right]_{t=0} = \sqrt{t^2 \cdot \sum_j^M (w_j)^2 + \left[\frac{\sum_j^M w_j}{\sum_k^N w_k} \cdot (1-t) \right]^2 \cdot \sum_k^N (w_k)^2} = 0 \tag{A.6}$$

The result is after solving for t:

$$t = \frac{(\sum_j^M w_j)^2 \cdot \sum_k^N (w_k)^2}{(\sum_k^N w_k)^2 \cdot \sum_j^M (w_j)^2 + (\sum_j^M w_j)^2 + \sum_k^N (w_k)^2} \tag{A.7}$$

which gives back a and b after substituting in the curve $\overrightarrow{\gamma}(t)$:

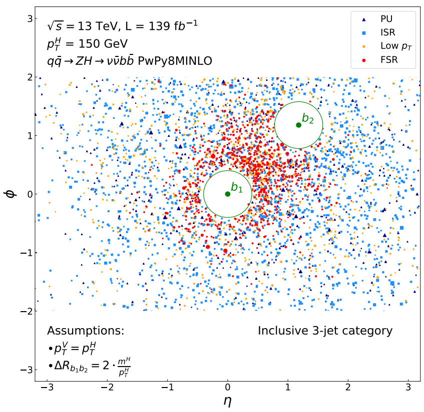
$$a = \frac{(\sum_j^M w_j)^2 \cdot \sum_k^N (w_k)^2}{(\sum_k^N w_k)^2 \cdot \sum_j^M (w_j)^2 + (\sum_j^M w_j)^2 + \sum_k^N (w_k)^2} \tag{A.8}$$

$$b = \frac{\sum_j^M w_j \cdot \sum_k^N w_k \cdot \sum_j^M (w_j)^2}{(\sum_k^N w_k)^2 \cdot \sum_j^M (w_j)^2 + (\sum_j^M w_j)^2 + \sum_k^N (w_k)^2} \tag{A.9}$$

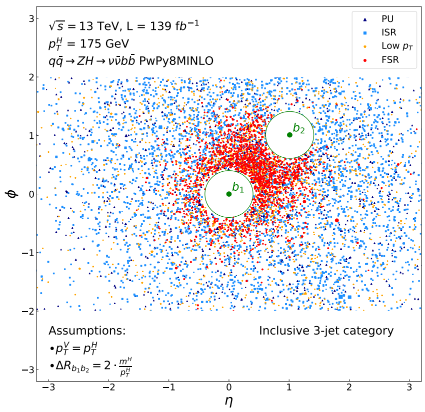
For $M \sim N$, Equations A.8 and A.9 significantly simplify leading to $a = b = 0.5$.

Appendix B

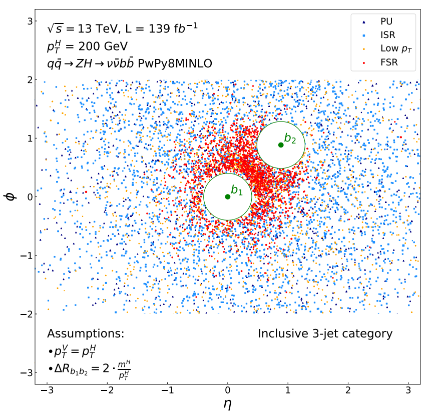
Additional plots for FSR studies



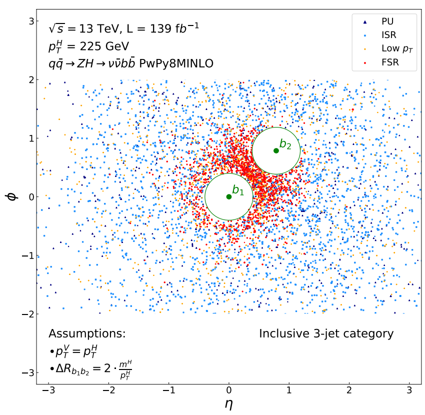
(a) $p_{\mathrm{T}}^{H} \in [150, 155]$ GeV.



(b) $p_{\mathrm{T}}^{H} \in [170, 180]$ GeV.

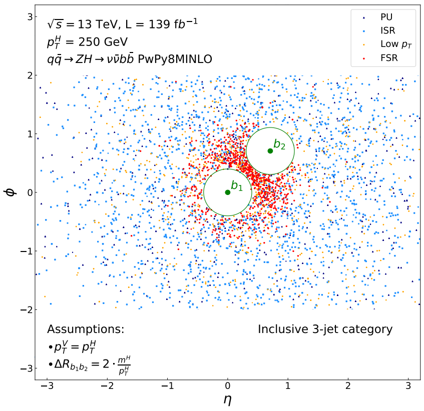


(c) $p_{\mathrm{T}}^{H} \in [195, 205]$ GeV.

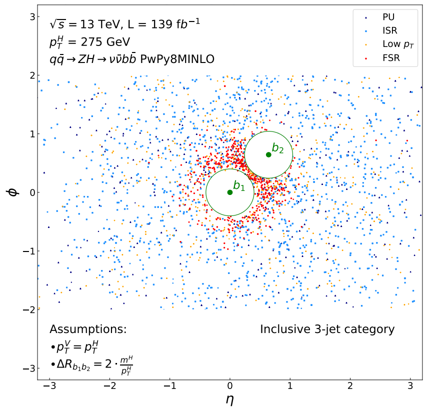


(d) $p_{\mathrm{T}}^{H} \in [220, 230]$ GeV.

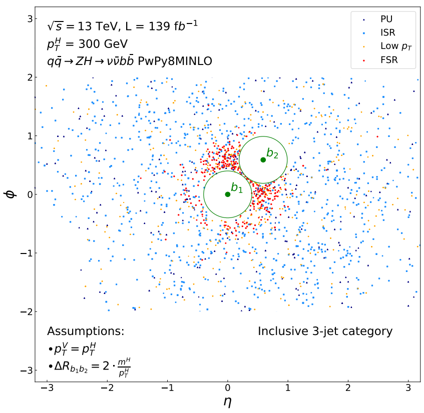
Figure B.1: $\eta - \phi$ distribution of the third jet in the 0-Lepton 3-jet inclusive category for various range of p_{T}^{H}: $150 - 225$ GeV (details in Section 4.2.1).



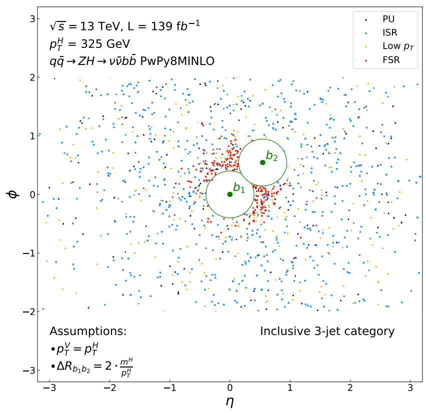
(a) $p_{\mathrm{T}}^{H} \in [245, 255]$ GeV.



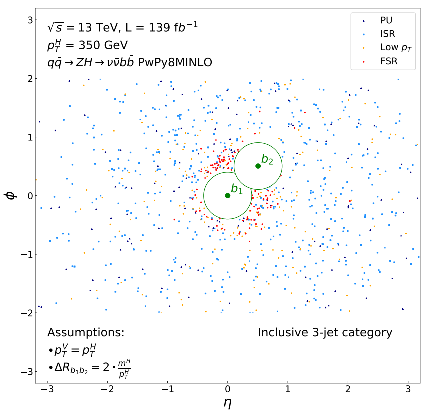
(b) $p_{\mathrm{T}}^{H} \in [270, 280]$ GeV.



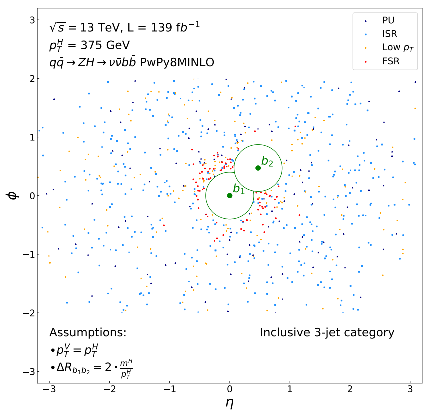
(c) $p_{\mathrm{T}}^{H} \in [295, 305]$ GeV.



(d) $p_{\mathrm{T}}^{H} \in [320, 330]$ GeV.



(e) $p_{\mathrm{T}}^{H} \in [345, 355]$ GeV.



(f) $p_{\mathrm{T}}^{H} \in [370, 380]$ GeV.

Figure B.2: $\eta - \phi$ distribution of the third jet in the 0-Lepton 3-jet inclusive category for various range of p_{T}^{H}: $250 - 375$ GeV (details in Section 4.2.1).

Appendix C

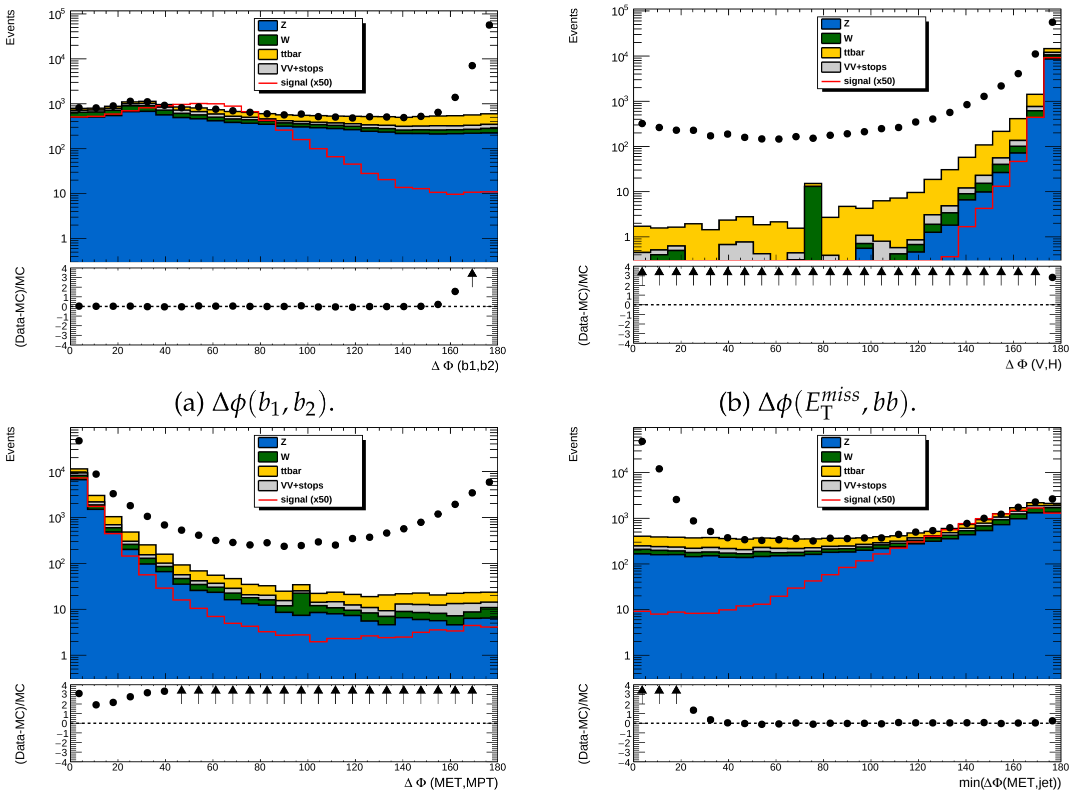
Multi-jet estimation in the 0-Lepton channel

In this appendix, a data-driven method and an MC method will be discussed to estimate the multi-jet background in 0-Lepton.

C.0.0.1 Data-driven estimation

In the 0-Lepton channel, four topological selections have been introduced to suppress the multi-jet background in Section 4.3.4. In Figure C.1 and Figure C.2, the four variable distributions are shown, respectively for the 2-jet and 3-jet categories, before any of the anti-QCD cuts are applied.

The value of the $\min[\Delta\phi(E_{\mathrm{T}}^{miss}, jets)]$ selection is established after performing a template fit. After the first three selections have been applied, the $\min[\Delta\phi(E_{\mathrm{T}}^{miss}, jets)]$ data distribution is fitted using the MC templates for the electroweak processes. The underlying assumption is that the multi-jet template can be extracted from data as a difference between the measured data and the fitted MC templates. The results are shown in Figure C.3. In the fit, the normalisations of the main 0-Lepton backgrounds (Z+jets, W+jets and $t\bar{t}$) are left floating, while the normalisations of single top and di-boson productions are fixed to the MC prediction. Furthermore, the 3-jet category multi-jet shape is fitted with an exponential function of the form $f(x) = A \cdot e^{-x/c}$ in the range $x = \min[\Delta\phi(E_{\mathrm{T}}^{miss}, jets)] \in [0, 50]$.



(a) $\Delta\phi(b_1, b_2)$.

(b) $\Delta\phi(E_{\mathrm{T}}^{miss}, bb)$.

(c) $\Delta\phi(E_{\mathrm{T}}^{miss}, bb)$.

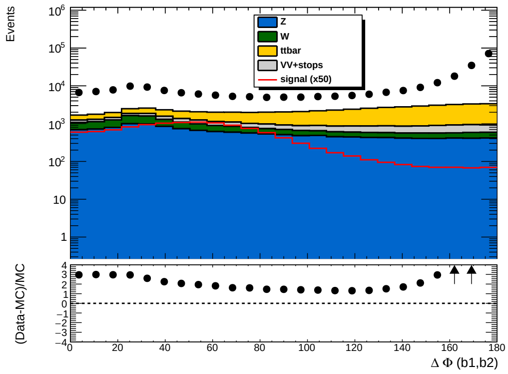
(d) $\min\left[\Delta\phi(E_{\mathrm{T}}^{miss}, jets)\right]$.

Figure C.1: Anti-QCD variable distributions before the anti-QCD selection for the 0-Lepton 2-jet category for $E_{\mathrm{T}}^{miss} > 150$ GeV. The electroweak backgrounds are scaled to the values extracted from the fit.

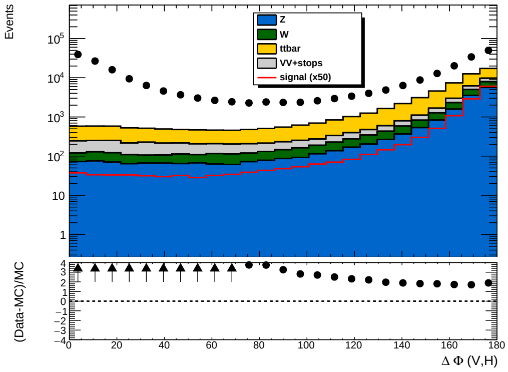
The 20° selection value for 2-jet category eliminates any remaining multi-jet contribution. The 30° selection value for 3-jet category corresponds to keeping less than 10% of multi-jet with respect to the signal yields. The parameters extracted from the exponential fit are: $A = 5409 \pm 168$ and $c = 6.26 \pm 0.18°$.

C.0.0.2 MC based multi-jet estimation

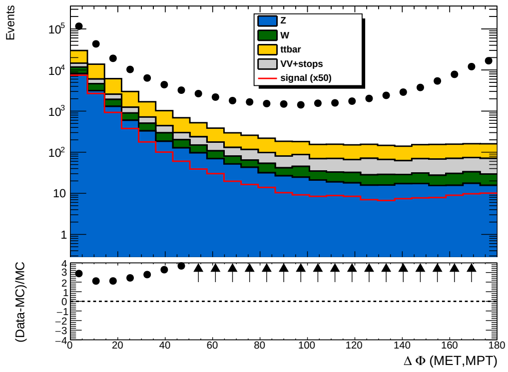
An alternative MC based approach has been tested to complement the data-driven multi-jet estimation, using a simulated di-jet sample [133]. The di-jet sample has been filtered in order to require at least two b-hadrons matched to two separate jets and $p_T^{b1} + p_T^{b2} > 150$ GeV ($\sigma = 609$ nb, $\epsilon_{filter} = 0.018619$). More than 33.9 millions of events have been generated for the corresponding $L = 139$ fb^{-1} dataset. The sample has been generated with the Sherpa generator [99] and using the NNPDF30NNLO pdf



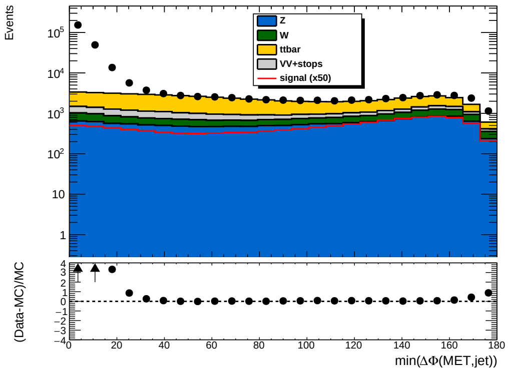
(a) $\Delta\phi(b_1, b_2)$.



(b) $\Delta\phi(E_{\mathrm{T}}^{miss}, bb)$.



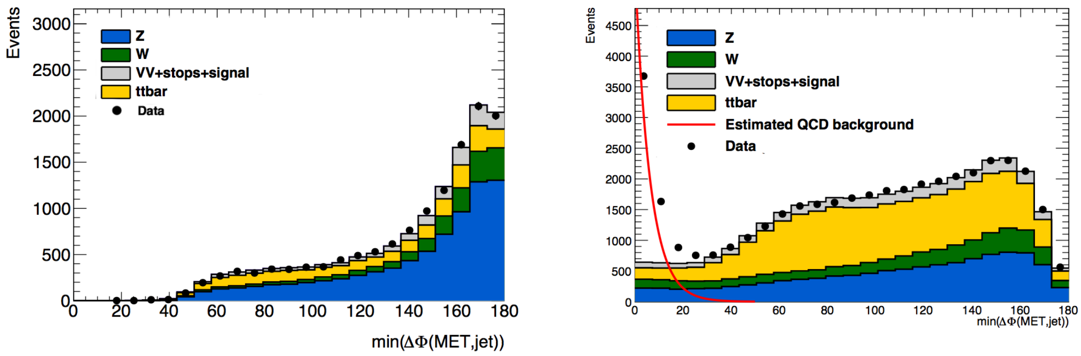
(c) $\Delta\phi(E_{\mathrm{T}}^{miss}, bb)$.



(d) min $\left[\Delta\phi(E_{\mathrm{T}}^{miss}, jets)\right]$.

Figure C.2: Anti-QCD variable distributions before the anti-QCD selection for the 0-Lepton 3-jet category for $E_{\mathrm{T}}^{miss} > 150$ GeV. The electroweak backgrounds are scaled to the values extracted from the fit.

set [134]. After applying the complete 0-Lepton selection, with the anti-QCD cuts, zero events enter in the 2-jet category and less than one hundred enter the 3-jet category, which are reported in Figure C.4. An exponential fit to min$[\Delta\phi(E_{\mathrm{T}}^{miss}, jets)]$, analogous to the one described in Section C.0.0.1, is performed, resulting in $A = 4674 \pm 127$ and $c = 4.72 \pm 0.11°$. These values are compatible with the data-driven approach.



(a) 2-jet.

(b) 3-jet.

Figure C.3: Multi-jet data-driven fit of $\min[\Delta\phi(E_{\mathrm{T}}^{miss}, jets)]$ in the 2-jet (a) and 3-jet (b) categories.

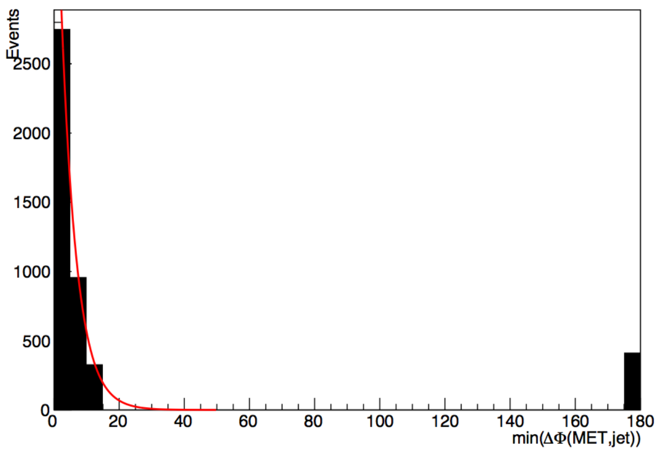


Figure C.4: Multi-jet fit of $\min[\Delta\phi(E_{\mathrm{T}}^{miss}, jets)]$ on the simulated di-jet sample in the 3-jet category.

Appendix D

Correlation matrices for MVA inputs

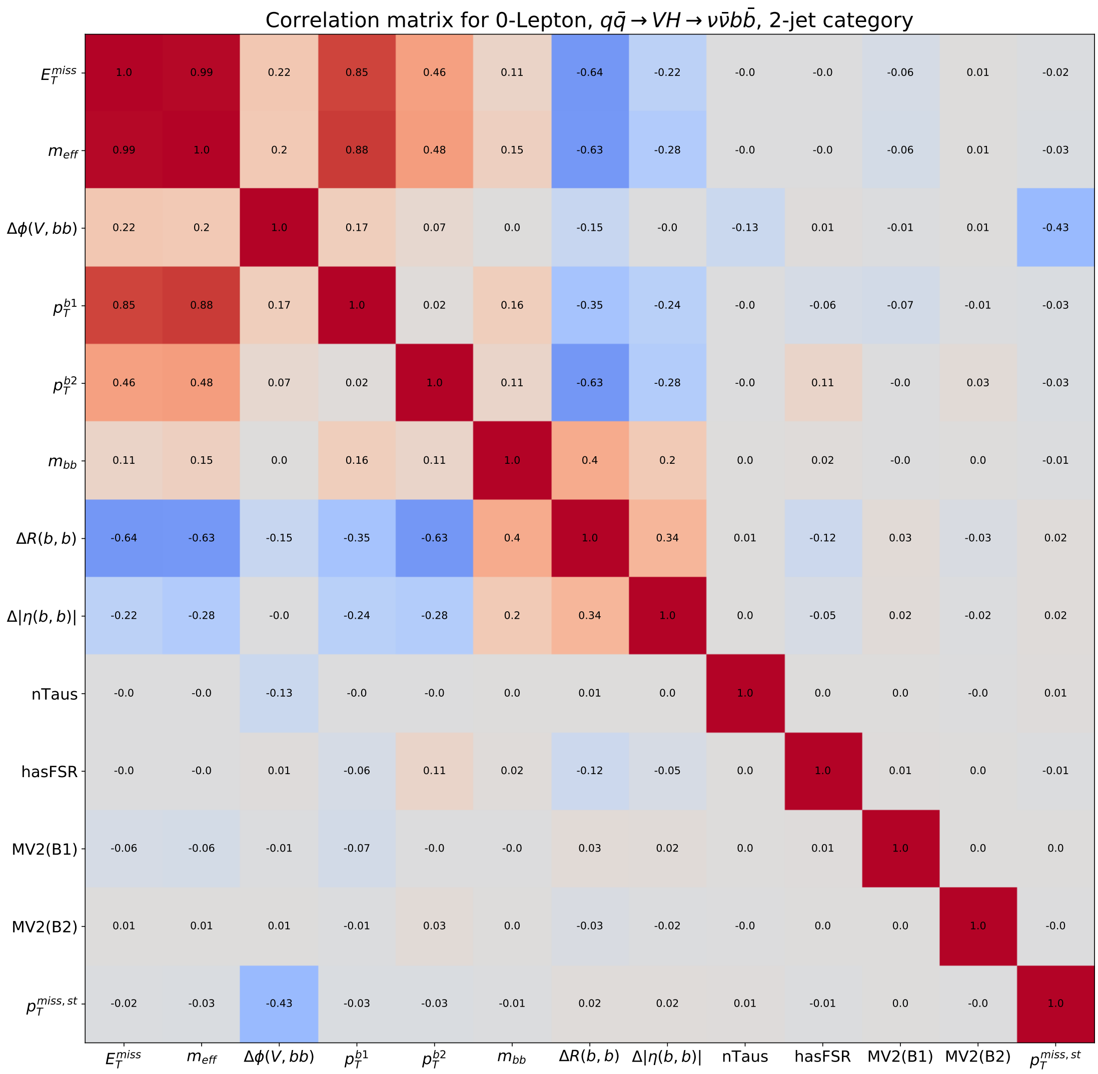


Figure D.1: Linear correlation matrix for the MVA input variables for the inclusive 0-Lepton channel 2-jet category in the $q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}$ simulated sample (details in Section 4.10.1).

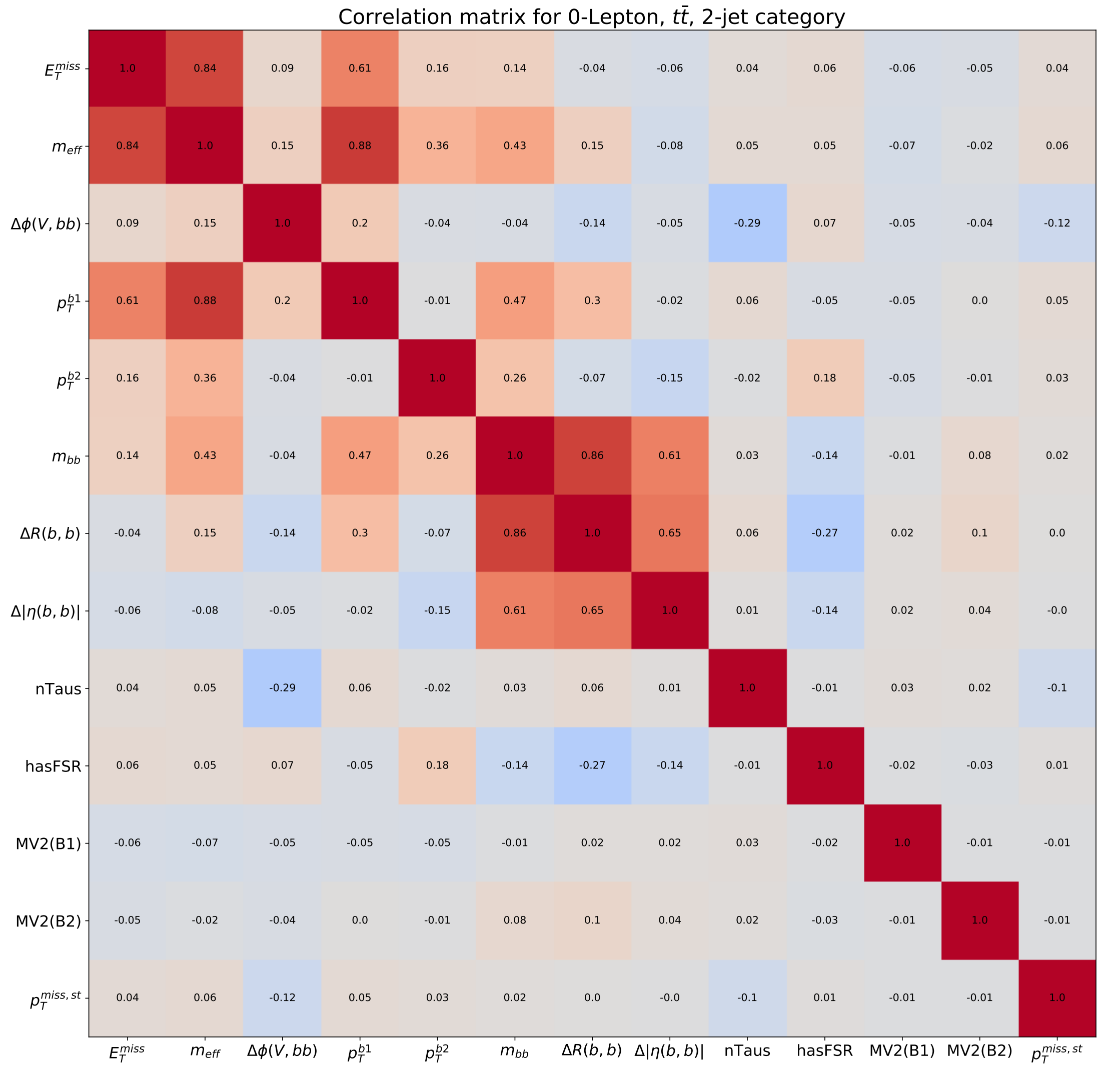


Figure D.2: Linear correlation matrix for the MVA input variables for the inclusive 0-Lepton channel 2-jet category in the $t\bar{t}$ simulated sample (details in Section 4.10.1).

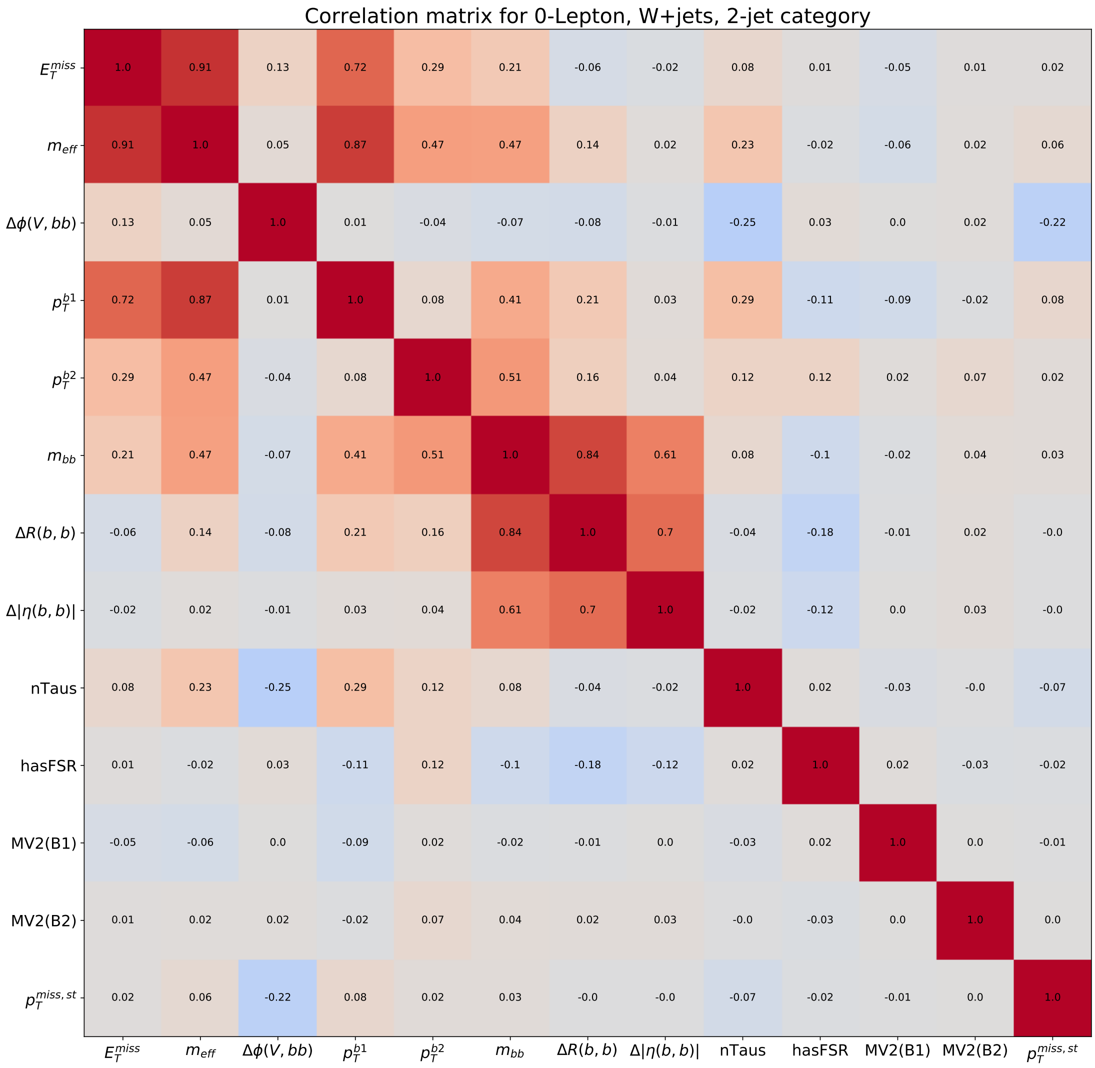


Figure D.3: Linear correlation matrix for the MVA input variables for the inclusive 0-Lepton channel 2-jet category in the W+jets simulated sample (details in Section 4.10.1).

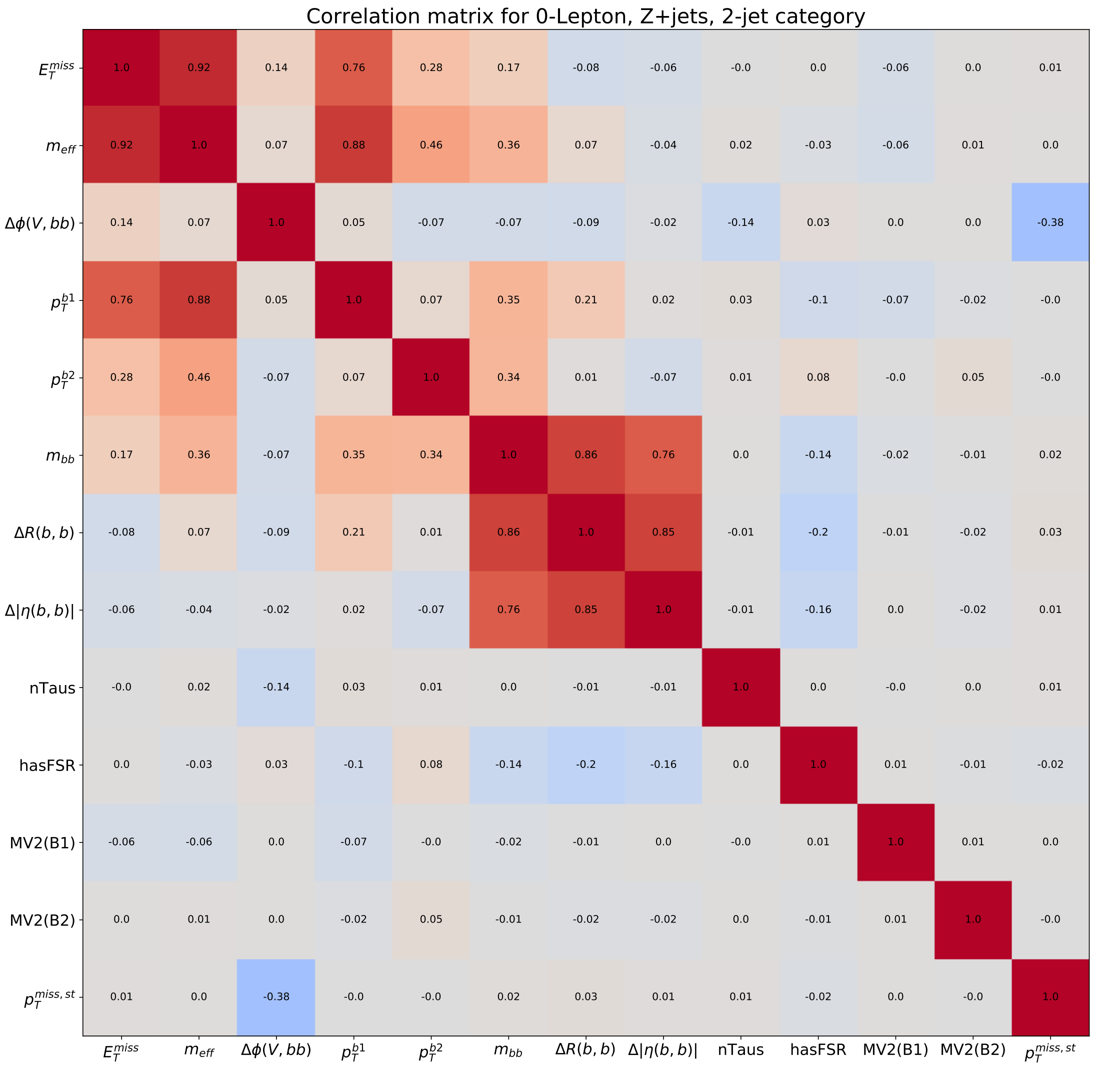


Figure D.4: Linear correlation matrix for the MVA input variables for the inclusive 0-Lepton channel 2-jet category in the Z+jets simulated sample (details in Section 4.10.1).

Bibliography

[1] L. Vigani et al. "Study of prototypes of LFoundry active CMOS pixels sensors for the ATLAS detector". In: *Journal of Instrumentation* 13.02 (2018), pp. C02021–C02021. DOI: 10.1088/1748-0221/13/02/c02021. URL: https://doi.org/10.1088%2F1748-0221%2F13%2F02%2Fc02021.

[2] ATLAS Collaboration. "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector". In: *Phys. Lett.* B786 (2018), pp. 59–86. DOI: 10.1016/j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex].

[3] ATLAS Collaboration. "Measurement of VH, H $\to$ b$\overline{\text{b}}$ production as a function of the vector-boson transverse momentum in 13 TeV pp collisions with the ATLAS detector". In: *JHEP* 05 (2019), p. 141. DOI: 10.1007/JHEP05(2019)141. arXiv: 1903.04618 [hep-ex].

[4] 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*. Tech. rep. ATLAS-CONF-2020-006. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2714885.

[5] Luca Ambroz. "Measurements of the Higgs boson decays to two bottom quarks". In: *PoS* LeptonPhoton2019 (2019), p. 085. DOI: 10.22323/1.367.0085.

[6] Chris Quigg. "Introduction". In: *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions: Second Edition*. STU - Student edition. Princeton University Press, 2013, pp. 1 –24, 430 –431.

[7] Tom W.B. Kibble. "The Standard Model of Particle Physics". In: *European Review* 23.1 (2015), 36–44. DOI: 10.1017/S1062798714000520.

[8] A. Purcell. *Go on a particle quest at the first CERN webfest*. 2012. URL: https://cds.cern.ch/record/1473657.

[9] S. Francesco. "Standard model electroweak measurements at LEP". In: *Proceedings, 41st Rencontres de Moriond on Electroweak Interactions and Unified Theories: La Thuile, Val d'Aoste, Italy, Mar 11-18, 2006*. 2006, pp. 132–139. arXiv: hep-ex/0605093 [hep-ex].

[10] ATLAS Collaboration. "Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C78.2 (2018). [Erratum: Eur. Phys. J.C78,no.11,898(2018)], p. 110. DOI: 10.1140/epjc/s10052-018-6354-3, 10.1140/epjc/s10052-017-5475-4. arXiv: 1701.07240 [hep-ex].

[11] John Ellis. "The Discovery of the Gluon". In: *Int. J. Mod. Phys.* A29.31 (2014). [,189(2015)], p. 1430072. DOI: 10.1142/9789814618113_0012, 10.1142/S0217751X14300725. arXiv: 1409.4232 [hep-ph].

[12] D. J. Gross. "The discovery of asymptotic freedom and the emergence of QCD". In: *Proc. Nat. Acad. Sci.* 102 (2005). [Int. J. Mod. Phys.A20,5717(2005); Rev. Mod. Phys.77,837(2005)], pp. 9099–9108. DOI: 10.1142/S0217751X05029009, 10.1073/pnas.0503831102, 10.1103/RevModPhys.77.837.

[13] M. Gell-Mann. "A schematic model of baryons and mesons". In: *Physics Letters* 8.3 (1964), pp. 214 –215. ISSN: 0031-9163. DOI: https://doi.org/10.1016/S0031-9163(64)92001-3. URL: http://www.sciencedirect.com/science/article/pii/S0031916364920013.

[14] G Zweig. "An SU_3 model for strong interaction symmetry and its breaking; Version 2". In: CERN-TH-412 (1964), 80 p. URL: https://cds.cern.ch/record/570209.

[15] L. Di Lella and C. Rubbia. "The Discovery of the W and Z Particles". In: *Adv. Ser. Dir. High Energy Phys.* 23 (2015), pp. 137–163. DOI: 10.1142/9789814644150_0006. URL: https://cds.cern.ch/record/2103277.

[16] R. Lewis H. "Spontaneous symmetry breaking and the Weinberg-Salam model". In: *Quantum Field Theory*. 2nd ed. Cambridge University Press, 1996, 240–307. DOI: 10.1017/CBO9780511813900.010.

[17] R. Brout F. Englert. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[18] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[19] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (20 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[20] 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.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[21] CMS Collaboration. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[22] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector". In: (2016). DOI: 10.2172/1345634, 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[23] H. M. Georgi et al. "Higgs Bosons from Two-Gluon Annihilation in Proton-Proton Collisions". In: *Phys. Rev. Lett.* 40 (11 1978), pp. 692–694. DOI: 10.1103/PhysRevLett.40.692. URL: https://link.aps.org/doi/10.1103/PhysRevLett.40.692.

[24] Paolo Bolzoni et al. "Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD". In: *Phys. Rev. Lett.* 105 (1 2010), p. 011801. DOI: 10.1103/PhysRevLett.105.011801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.105.011801.

[25] S. Dittmaier et al. "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables". In: (2011). DOI: 10.5170/CERN-2011-002. arXiv: 1101.0593 [hep-ph].

[26] ATLAS Collaboration. *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331061.

[27] ATLAS Collaboration. "Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment". In: *Phys. Rev.* D101.1 (2020), p. 012002. DOI: 10.1103/PhysRevD.101.012002. arXiv: 1909.02845 [hep-ex].

[28] Johann Brehmer. "Higgs Effective Field Theory". In: (2016). URL: https://www.thphys.uni-heidelberg.de/~gk_ppbsm/lib/exe/fetch.php?media=students:lectures:student_lecture_eft.pdf.

[29] B. Grzadkowski et al. "Dimension-Six Terms in the Standard Model Lagrangian". In: *JHEP* 10 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph].

[30] Oliver Sim Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1. URL: https://cds.cern.ch/record/782076.

[31] ATLAS Collaboration. *Luminosity and Pile-up ATLAS Public plots in Run II*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[32] UK Research and Innovation. *CERN Accelerator Complex*. URL: https://stfc.ukri.org/research/particle-physics-and-particle-astrophysics/large-hadron-collider/cern-accelerator-complex/.

[33] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001.

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

[35] ATLAS Collaboration. *Good Run Lists For Analysis Run 2*. 2019. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/GoodRunListsForAnalysisRun2.

[36] ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2019-021. Geneva: CERN, 2019. URL: http://cds.cern.ch/record/2677054.

[37] LHC Commissioning. *Longer term LHC schedule*. URL: https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm.

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

[39] ATLAS Collaboration. *ATLAS magnet system: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/338080.

[40] ATLAS Collaboration. "ATLAS inner detector: Technical design report. Vol. 2". In: (1997).

[41] M Capeans et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. 2010. URL: https://cds.cern.ch/record/1291633.

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

[43] ATLAS Collaboration. "ATLAS pixel detector electronics and sensors". In: *JINST* 3 (2008), P07007. DOI: 10.1088/1748-0221/3/07/P07007.

[44] ATLAS Collaboration. "The barrel modules of the ATLAS semiconductor tracker". In: *Nucl. Instrum. Meth. A* 568 (2006), pp. 642–671. DOI: 10.1016/j.nima.2006.08.036.

[45] ATLAS Collaboration. "The ATLAS TRT barrel detector". In: *JINST* 3 (2008), P02014. DOI: 10.1088/1748-0221/3/02/P02014.

[46] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev.* D98.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[47] Christian W. Fabjan and Fabiola Gianotti. "Calorimetry for particle physics". In: *Rev. Mod. Phys.* 75 (4 2003), pp. 1243–1286. DOI: 10.1103/RevModPhys.75.1243. URL: https://link.aps.org/doi/10.1103/RevModPhys.75.1243.

[48] ATLAS Collaboration. *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331062.

[49] ATLAS Collaboration. *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331068.

[50] ATLAS Collaboration. *Trigger menu in 2018*. Tech. rep. ATL-DAQ-PUB-2019-001. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2693402.

[51] E. Simioni et al. "Upgrade of the ATLAS Level-1 Trigger with event topology information". In: *J. Phys. Conf. Ser.* 664.8 (2015), p. 082052. DOI: 10.1088/1742-6596/664/8/082052.

[52] Florin Pop. "High Performance Numerical Computing for High Energy Physics: A New Challenge for Big Data Science". In: *Adv. High Energy Phys.* 2014 (2014), p. 507690. DOI: 10.1155/2014/507690.

[53] Peter Skands. "Introduction to QCD". In: *Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales*. 2017, pp. 63–124. DOI: 10.1142/9789814525220_0008. arXiv: 1207.2389 [hep-ph].

[54] Davison E. Soper. "Parton distribution functions". In: *Nucl. Phys. Proc. Suppl.* 53 (1997), pp. 69–80. DOI: 10.1016/S0920-5632(96)00600-7. arXiv: hep-lat/9609018 [hep-lat].

[55] John M. Campbell, J. W. Huston, and W. J. Stirling. "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics". In: *Rept. Prog. Phys.* 70 (2007), p. 89. DOI: 10.1088/0034-4885/70/1/R02. arXiv: hep-ph/0611148 [hep-ph].

[56] B. R. Webber. "Fragmentation and hadronization". In: *Int. J. Mod. Phys.* A15S1 (2000). [eConfC990809,577(2000)], pp. 577–606. DOI: 10.1142/S0217751X00005334. arXiv: hep-ph/9912292 [hep-ph].

[57] S. Agostinelli et al. "GEANT4: a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[58] T Cornelissen et al. "The new ATLAS track reconstruction (NEWT)". In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032014. DOI: 10.1088/1742-6596/119/3/032014. URL: https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032014.

[59] ATLAS Collaboration. "Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2". In: *Eur. Phys. J. C* 77.10 (2017), p. 673. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex].

[60] ATLAS Collaboration. *Event Displays from Run 2 physics analyses*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayRun2Physics.

[61] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, 2008. URL: https://cds.cern.ch/record/1099735.

[62] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J.* C77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[63] ATLAS Collaboration. "Performance of electron and photon triggers in ATLAS during LHC Run 2". In: *Eur. Phys. J.* C80.1 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex].

[64] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *JINST* 14.03 (2019), P03017. DOI: 10.1088/1748-0221/14/03/P03017. arXiv: 1812.03848 [hep-ex].

[65] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*. Tech. rep. ATLAS-CONF-2016-024. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2157687.

[66] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[67] Sotiris Kotsiantis, I. Zaharakis, and P. Pintelas. "Machine learning: A review of classification and combining techniques". In: *Artificial Intelligence Review* 26 (Nov. 2006), pp. 159–190. DOI: 10.1007/s10462-007-9052-3.

[68] ATLAS Collaboration. "Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment". In: *Eur. Phys. J.* C76.5 (2016), p. 295. DOI: 10.1140/epjc/s10052-016-4110-0. arXiv: 1512.05955 [hep-ex].

[69] Gavin P. Salam. "Towards Jetography". In: *Eur. Phys. J.* C67 (2010), pp. 637–686. DOI: 10.1140/epjc/s10052-010-1314-6. arXiv: 0906.1833 [hep-ph].

[70] 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].

[71] ATLAS Collaboration. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *Eur. Phys. J.* C77.7 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[72] ATLAS Collaboration. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* D96.7 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[73] Matteo Cacciari and Gavin P. Salam. "Pileup subtraction using jet areas". In: *Phys. Lett.* B659 (2008), pp. 119–126. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph].

[74] ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector*. Tech. rep. ATLAS-CONF-2014-018. Geneva: CERN, 2014. URL: https://cds.cern.ch/record/1700870.

[75] ATLAS Collaboration. "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV". In: *JHEP* 08 (2018), p. 089. DOI: 10.1007/JHEP08(2018)089. arXiv: 1805.01845 [hep-ex].

[76] ATLAS Collaboration. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J.* C79.11 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].

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

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

[79] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J.* C78.11 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[80] ATLAS Collaboration. *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV*. Tech. rep. ATL-PHYS-PUB-2015-023. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2037700.

[81] D Adams et al. *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*. Tech. rep. ATL-COM-PHYS-2014-451. Geneva: CERN, 2014. URL: https://cds.cern.ch/record/1700874.

[82] John Ellis, Mary K. Gaillard, and Dimitri V. Nanopoulos. "A Historical Profile of the Higgs Boson". In: (2012). arXiv: 1201.6045 [hep-ph].

[83] John R. Ellis, Mary K. Gaillard, and Dimitri V. Nanopoulos. "A Phenomenological Profile of the Higgs Boson". In: *Nucl. Phys.* B106 (1976), p. 292. DOI: 10.1016/0550-3213(76)90382-5.

[84] Riccardo Barbieri and Torleif Erik Oskar Ericson. "Evidence Against the Existence of a Low Mass Scalar Boson from Neutron-Nucleus Scattering". In: *Phys. Lett.* 57B (1975), pp. 270–272. DOI: 10.1016/0370-2693(75)90073-8.

[85] Benjamin W. Lee, C. Quigg, and H. B. Thacker. "The Strength of Weak Interactions at Very High-Energies and the Higgs Boson Mass". In: *Phys. Rev. Lett.* 38 (1977), pp. 883–885. DOI: 10.1103/PhysRevLett.38.883.

[86] "Search for the standard model Higgs boson at LEP". In: *Lepton and photon interactions at high energies. Proceedings, 20th International Symposium, LP 2001, Rome, Italy, July 23-28, 2001*. 2001. arXiv: hep-ex/0107029 [hep-ex].

[87] T. Aaltonen et al. "Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron". In: *Phys. Rev. Lett.* 109 (2012), p. 071804. DOI: 10.1103/PhysRevLett.109.071804. arXiv: 1207.6436 [hep-ex].

[88] Jonathan M. Butterworth et al. "Jet substructure as a new Higgs search channel at the LHC". In: *Phys. Rev. Lett.* 100 (2008), p. 242001. DOI: 10.1103/PhysRevLett.100.242001. arXiv: 0802.2470 [hep-ph].

[89] CMS Collaboration. "Observation of Higgs boson decay to bottom quarks". In: *Phys. Rev. Lett.* 121.12 (2018), p. 121801. DOI: 10.1103/PhysRevLett.121.121801. arXiv: 1808.08242 [hep-ex].

[90] Jessie Shelton. "Jet Substructure". In: *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales (TASI 2012): Boulder, Colorado, June 4-29, 2012*. 2013, pp. 303–340. DOI: 10.1142/9789814525220_0007. arXiv: 1302.0260 [hep-ph].

[91] Stefano Frixione et al. "Single-top hadroproduction in association with a W boson". In: *JHEP* 07 (2008), p. 029. DOI: 10.1088/1126-6708/2008/07/029. arXiv: 0805.3067 [hep-ph].

[92] VHbb analysis group. *Object selections for SM Higgs boson produced in association with a vector boson in which $H \to b\bar{b}$ and V decays leptonically with Run-2 data (2015+2016+2017)*. Tech. rep. 2018.

[93] Ansgar Denner et al. "Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK". In: *JHEP* 03 (2012), p. 075. DOI: 10.1007/JHEP03(2012)075. arXiv: 1112.5142 [hep-ph].

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

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

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

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

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

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

[100] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *J. Phys.* G43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[101] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. "A brief introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[102] A. D. Martin et al. "Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections". In: *Eur. Phys. J.* C 64 (2009), p. 653. DOI: 10.1140/epjc/s10052-009-1164-2. arXiv: 0905.3531 [hep-ph].

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

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

[105] Nikolaos Kidonakis. "NNLL resummation for s-channel single top quark production". In: *Phys. Rev. D* 81 (2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028. arXiv: 1001.5034 [hep-ph].

[106] Nikolaos Kidonakis. "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production". In: *Phys. Rev. D* 83 (2011), p. 091503. DOI: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792 [hep-ph].

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

[108] Stefano Catani et al. "Vector Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at Next-to-Next-to-Leading Order". In: *Phys. Rev. Lett.* 103 (2009), p. 082001. DOI: 10.1103/PhysRevLett.103.082001. arXiv: 0903.2120 [hep-ph].

[109] M. Boggia et al. "The HiggsTools handbook: a beginners guide to decoding the Higgs sector". In: *J. Phys.* G45.6 (2018), p. 065004. DOI: 10.1088/1361-6471/aab812. arXiv: 1711.09875 [hep-ph].

[110] J. R. Andersen et al. "Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report". In: *9th Les Houches Workshop on Physics at TeV Colliders (PhysTeV 2015) Les Houches, France, June 1-19, 2015*. 2016. arXiv: 1605.04692 [hep-ph]. URL: http://lss.fnal.gov/archive/2016/conf/fermilab-conf-16-175-ppd-t.pdf.

[111] Nicolas Berger et al. "Simplified Template Cross Sections - Stage 1.1". In: (2019). arXiv: 1906.02754 [hep-ph].

[112] A. D. Bukin. *Fitting function for asymmetric peaks*. 2007. arXiv: 0711.4449v2 [physics.data-an].

[113] ATLAS collaboration. "Search for invisible Higgs boson decays in vector boson fusion at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 793 (2019), pp. 499–519. DOI: 10.1016/j.physletb.2019.04.024. arXiv: 1809.06682 [hep-ex].

[114] Kim Albertsson et al. "Machine Learning in High Energy Physics Community White Paper". In: *J. Phys. Conf. Ser.* 1085.2 (2018), p. 022008. DOI: 10.1088/1742-6596/1085/2/022008. arXiv: 1807.02876 [physics.comp-ph].

[115] Andreas Hoecker et al. *TMVA - Toolkit for Multivariate Data Analysis*. 2007. arXiv: 0703039 [physics.data-an].

[116] Andrew P Bradley. "The use of the area under the ROC curve in the evaluation of machine learning algorithms". In: *Pattern recognition* 30.7 (1997), pp. 1145–1159.

[117] Tianqi Chen and Carlos Guestrin. "Xgboost: A scalable tree boosting system". In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016, pp. 785–794.

[118] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[119] Roger Barlow and Christine Beeston. "Fitting using finite Monte Carlo samples". In: *Computer Physics Communications* 77.2 (1993), pp. 219 –228. ISSN: 0010-4655. DOI: https://doi.org/10.1016/0010-4655(93)90005-W. URL: http://www.sciencedirect.com/science/article/pii/001046559390005W.

[120] ATLAS Collaboration. "Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector". In: *Eur. Phys. J. C* 73.3 (2013), p. 2306. DOI: 10.1140/epjc/s10052-013-2306-0. arXiv: 1210.6210 [hep-ex].

[121] William Buttinger. *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*. Tech. rep. ATL-COM-SOFT-2015-119. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2014726.

[122] VHbb analysis group. *Data/Monte Carlo Modelling Studies for the Standard Model $VH, H \to b\bar{b}$ analysis*. Tech. rep. 2020.

[123] Kyle Cranmer, Juan Pavez, and Gilles Louppe. "Approximating Likelihood Ratios with Calibrated Discriminative Classifiers". In: (June 2015). arXiv: 1506.02169 [stat.AP].

[124] ATLAS Collaboration. *Evaluation of theoretical uncertainties for simplified template cross section measurements of V-associated production of the Higgs boson*. Tech. rep. ATL-PHYS-PUB-2018-035. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2649241.

[125] Chris D. White et al. "Isolating Wt production at the LHC". In: *JHEP* 11 (2009), p. 074. DOI: 10.1088/1126-6708/2009/11/074. arXiv: 0908.0631 [hep-ph].

[126] ATLAS Collaboration. "Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (Mar. 2020). arXiv: 2003.11960 [hep-ex].

[127] ATLAS Collaboration. "Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector". In: *JHEP* 12 (2017), p. 024. DOI: 10.1007/JHEP12(2017)024. arXiv: 1708.03299 [hep-ex].

[128] L3 OPAL SLD LEP Electroweak Working Group SLD Electroweak Group SLD Heavy Flavour Group ALEPH DELPHI. "Precision electroweak measurements on the Z resonance". In: *Phys. Rept.* 427 (2006), pp. 257–454. DOI: 10.1016/j.physrep.2005.12.006. arXiv: hep-ex/0509008.

[129] John Ellis et al. "Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data". In: *JHEP* 06 (2018), p. 146. DOI: 10.1007/JHEP06(2018)146. arXiv: 1803.03252 [hep-ph].

[130] M. Delmastro et al. "Simplified Template Cross Sections – Stage 1.1". In: (2019). 14 pages, 3 figures. arXiv: 1906.02754. URL: http://cds.cern.ch/record/2669925.

[131] Readthedocs. *XGBoost Documentation*. 2020. URL: https://xgboost.readthedocs.io.

[132] Readthedocs. *Loss Functions*. 2017. URL: https://ml-cheatsheet.readthedocs.io/en/latest/loss_functions.html.

[133] Javier Llorente Merino et al. *Multijet simulation for 13 TeV ATLAS Analyses*. Tech. rep. ATL-COM-PHYS-2018-1514. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2646017.

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