Observation of H→bb decays and VH production with the ATLAS detector

Yanhui Ma

NNT : 2019SACLS099

Observation du mode de désintégration $H \to bb$ et de la production associée de VH avec le détecteur ATLAS

Observation of $H \to bb$ decays and VH production with ATLAS detector

Thèse de doctorat de Shandong University et de l'Université Paris-Saclay, préparée à l'Université Paris-Sud

École doctorale n°576 Particules, Hadrons, Énergie, Noyau, Instrumentation, Image, Cosmos et Simulation (PHÉNIICS)
SCHOOL of PHYSICS
Spécialité de doctorat : Physique des particules

Thèse présentée et soutenue à Qingdao, le 28 Mai 2019, par

M. Yanhui Ma

Composition du Jury :

M. Yajun MAO	
Professeur, Peking University	Président
M. Shenjian CHEN	
Professeur, Nanjing University	Rapporteur
M. Matteo CACCIARI	
Professeur, Université Paris Diderot	Rapporteur
M. Lianliang MA	
Professeur, Shandong University	Directeur de thèse
M. Marumi KADO	
Directeur de recherche, Laboratoire de l'Accélérateur Linéaire	Directeur de thèse

Thèse de doctorat

Résumé

Une recherche du boson de Higgs du Modèle Standard produit en association avec un boson W ou Z et se désintégrant en une paire quark-antiquark b a été effectuée avec le détecteur ATLAS. Les données de collisions proton-proton utilisées ont été accumulées durant le Run 2 du Grand Collisionneur de Hadrons du CERN à une énergie dans le centre de masse de 13 TeV, et correspondent à une luminosité intégrée de 79.8 fb^{-1}. Trois canaux contenant zéro, un ou deux leptons chargés (électrons ou muons) sont considérés, correspondant à chacune des désintégrations leptoniques d'un boson W ou Z: $Z \to \nu\nu$, $W \to l\nu$ et $Z \to ll$. Pour un boson de Higgs de masse 125 GeV, un excès d'événements par rapport aux bruits de fonds des autres processus du Modèle Standard est observé avec un niveau de signification statistique de 4.9 déviations standard, à comparer à 4.3 attendues. Le rapport du nombre d'événements observé au nombre attendu est mesuré être $1.16^{+0.27}_{-0.25} = 1.16 \pm 0.16(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$. Ce résultat est combiné avec d'autres d'ATLAS sur la recherche du boson de Higgs se désintégrant dans le mode $b\bar{b}$, utilisant des données du Run 1 et du Run 2. Le niveau de signification mesuré (attendu) pour ce mode de désintégraion est de 5.4 (5.5) déviations standard, ce qui en constitue la première observation directe. De plus, une combinaison des résultats du Run 2 sur la recherche de la production associée du boson de Higgs et d'un boson W ou Z conduit à un niveau de signification observé (attendu) de 5.3 (4.8) déviations standard, et donc à la première observation de ce mode de production.

Most clé: LHC, expérience ATLAS, boson de Higgs, production associée de VH, quark b

Abstract

A search for the Standard Model Higgs boson produced in association with a W or Z boson, and decaying to a $b\bar{b}$ pair has been performed with the ATLAS detector. The data were collected in proton-proton collisions during Run 2 of the Large Hadron Collider at a centre-of-mass energy of 13 TeV, and correspond to an integrated luminosity of 79.8 fb^{-1}. Three channels containing zero, one and two charged leptons (electrons or muons) have been considered to target each of the leptonic decays of the W or Z boson, $Z \to \nu\nu$, $W \to l\nu$ and $Z \to ll$, referred to as as the 0-lepton, 1-lepton and 2-lepton channels, respectively. A data-driven method has been developed to estimate the multijet background in the 1-lepton channel, along with the detailed studies to assign proper uncertainties on the estimation. Extensive studies have been carried out to improve the sensitivity and robustness of the analysis, such as a study of the 1-lepton channel medium p_T^V region and a study of pile-up jets suppression cuts. These have been combined with the works on validating and updating the boosted decision tree training and extensive fit studies, to ensure the robustness of the final results. For a Higgs boson mass of 125 GeV, an excess of events over the expected background from other Standard Model processes is found with an observed significance of 4.9 standard deviations, compared to an expectation of 4.3 standard deviations. The ratio of the measured signal events to the Standard Model expectation equals to $1.16^{+0.27}_{-0.25} = 1.16 \pm 0.16(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$. The result is also combined with the other results from the searches for the Higgs boson in the $b\bar{b}$ decay mode in Run 1 and Run 2, the combination yields an observed (expected) significance of 5.4 (5.5) standard deviations, and therefore provides a direct observation of the Higgs boson decay into a $b\bar{b}$ pair. In addition, a combination of Run 2 results searching for the Higgs boson produced in association with a W or Z boson yields an observed (expected) significance of 5.3 (4.8) standard deviations, and therefore provides a direct observation of Higgs boson being produced in association with a W or Z boson.

Keywords: LHC, ATLAS experiment, Higgs boson, VH associated production, bottom-quark

Synthèse en français

Une recherche du boson de Higgs du Modèle Standard produit en association avec un boson W ou Z et se désintégrant en une paire $b\bar{b}$ a été effectuée au moyen du détecteur ATLAS. Les données ont été acquises en collisions proton-proton au cours du Run 2 du grand collisionneur de hadrons (LHC) du CERN à une énergie dans le centre de masse de 13 TeV et correspondent à une luminosité intégrée de 79.8 fb^{-1}. Trois canaux comportant zéro, un ou deux leptons chargés (électrons ou muons) ont été considérés en vue de chacun des modes de désintégration leptonique du boson W ou Z: $Z \to \nu\nu$, $W \to l\nu$ et $Z \to ll$, dénotés canal 0-lepton, 1-lepton et 2-leptons respectivement. Afin de maximiser la sensibilité à un signal de boson de Higgs, des discriminants multivariés sont construits à partir de variables caractérisant la cinématique des événements sélectionnés. Ces discriminants multivariés sont combinés au moyen d'un ajustement de maximum de vraisemblance, lequel permet d'extraire la force du signal et son niveau de signification. Deux autres analyses sont effectuées pour valider la mthode d'extraction du signal: l'analyse "dijet", où la masse du système des deux jets candidats à provenir de la désintégration d'un boson de Higgs est utilisée comme observable dans l'ajustement pour extraire le taux de production du signal; et l'analyse "dibosons", où l'analyse multivariée nominale est modifiée pour extraire le taux de production du processus $(W/Z)Z$ suivi de la désintégration $Z \to b\bar{b}$.

Dans tous les canaux, les événements doivent comporter exactement deux jets étiquetés comme provenant de la fragmentation d'un quark b, lesquels sont présumés constituer les produits de la désintégration d'un boson de Higgs. Au moins l'un de ces jets étiquetés doit avoir une impulsion transverse p_T supérieure à 45 GeV. L'ensemble des événements est séparé en catégories 3-jets et 2-jets selon qu'un jet additionnel non-étiqueté est ou non présent. Comme le rapport signal à bruit de fond augmente avec l'impulsion transverse du boson de Higgs, les canaux 0- et 1-lepton sont restreints à la région de grand p_T^V ($p_T^V > 150$ GeV), où p_T^V est l'impulsion transverse du boson W ou Z. Dans le canal 2-leptons, la sensibilité est accrue par l'addition de la région de moyen p_T^V (75 GeV $< p_T^V <$ 150 GeV).

Des critères spécifiques à chaque canal sont également appliqués pour sélectionner le boson W ou Z et pour réduire le fond multijet provenant de la production de jets par interaction forte. À l'issue de la sélection des évènements, des arbres de décision boostés (BDT) sont entraînés dans les diverses catégories et régions de signal, dont les variables de sortie sont utilisées par l'ajustement comme discriminants finals. Pour accroitre la sensibilité de l'analyse dijet, un certain nombre de critères supplémentaires sont appliqués aux événements pour réduire la contamination du bruit de fond.

L'analyse statistique repose sur une fonction de vraisemblance $\mathcal{L}(\mu, \boldsymbol{\theta})$ construite comme le produit de probabilités de Poisson sur les éléments des distributions des discriminants finals et de distributions de probabilités pour les paramètres de nuisance $\boldsymbol{\theta}$ représentant les incertitudes systématiques. Le paramètre d'intérêt μ, la force du signal qui multiplie le produit de la section efficace de production associée du boson de Higgs par le rapport d'embranchement de la désintégration $H \to b\bar{b}$ prédit par le Modèle Standard, est extraite par maximisation de la fonction de vraisemblance.

Dans l'analyse multivariée nominale le signal observé dans les données du Run 2, pour une masse du boson de Higgs de 125 GeV et lorsque les trois canaux sont combinés, a un niveau de signification correspondant à 4.9 déviations standard, à comparer à 4.3 attendues. La valeur ajustée de la force du signal est:

$$\mu_{VH}^{bb} = 1.16^{+0.27}_{-0.25} = 1.16 \pm 0.16(\text{stat.})^{+0.21}_{-0.19}(\text{syst.}),$$

où "stat" représente l'incertitude statistique et "syst" celle due aux incertitudes systématques.

Dans l'analyse dibosons, la valeur ajustée de la force du signal est:

$$\mu_{VZ}^{bb} = 1.20^{+0.20}_{-0.18} = 1.20 \pm 0.08(\text{stat.})^{+0.19}_{-0.16}(\text{syst.}),$$

en accord avec la prédiction du Modèle Standard.

Dans l'analyse dijet, le signal de boson de Higgs est observé avec un niveau de signification de 3.6 déviations standard, à comparer à 3.5 attendues, et la valeur ajustée de la force du signal est:

$$\mu_{VH}^{bb} = 1.06^{+0.36}_{-0.33} = 1.06 \pm 020(\text{stat.})^{+0.30}_{-0.26}(\text{syst.}),$$

en accord avec le résultat de l'analyse multivariée nominale.

Le résultat de l'analyse multivariée nominale est combiné avec le résultat correspondant obtenu avec les données du Run 1 et avec les résultats de recherches du boson de Higgs du Modèle Standard se désintégrant en une paire $b\bar{b}$ et produit par fusion de bosons vecteurs (VBF) ou en association avec une paire $t\bar{t}$ ($t\bar{t}H$) à la fois au Run 1 et au Run 2 afin d'augmenter la sensibilité au mode de désintégration $H \to b\bar{b}$. Sous l'hypothèse que les rapports des sections efficaces de production sont ceux prédits par le Modèle Standard pour une masse de boson de Higgs de 125 GeV, la signification statistique obtenue pour ce mode de désintégration est de 5.4 déviations standard, à comparer à 5.5 attendues. Sous l'hypothèse supplémentaire que les sections efficaces de production sont celles prédites par le Modèle Standard, la valeur ajustée de la force du signal de $H \to b\bar{b}$ est:

$$\mu_{H\to b\bar{b}} = 1.01 \pm 0.20 = 1.01 \pm 0.12(\text{stat.})^{+0.16}_{-0.15}(\text{syst.}).$$

Ce résultat constitue une observation directe du mode de désintégration $H \to b\bar{b}$.

Le résultat de l'analyse multivariée nominale des données du Run 2 pour la recherche de la production de $(W/Z)H$ suivie de la désintégration $H \to b\bar{b}$ est également combiné avec ceux d'autres recherches au Run 2 de la production de $(W/Z)H$, où le boson de Higgs se désintègre en deux photons ($H \to \gamma\gamma$) ou en quatre leptons ($H \to ZZ^* \to 4l$). Sous l'hypothèse que les rapports des rapports d'embranchement sont ceux prédits par le Modèle Standard pour une masse de boson de Higgs de 125 GeV, la signification statistique obtenue pour la production associée $(W/Z)H$ est de 5.3 déviations standard, à comparer avec 4.8 attendues. Sous l'hypothèse supplémentaire que les rapports d'embranchement sont ceux prédits par le Modèle Standard, la valeur ajustée de la force du signal de production de $(W/Z)H$ est:

$$\mu_{VH} = 1.13^{+0.24}_{-0.23} = 1.13 \pm 0.0.15(\text{stat.})^{+0.18}_{-0.17}(\text{syst.}).$$

Ce résultat constitue une observation directe du boson de Higgs produit en association avec un boson vecteur.

Acknowledgements

Without the help from countless people, I could not finish the work presented in this thesis. I would like to thank my supervisor from Shandong University, Prof. Lianliang Ma, who took me into the world of particle physics, for his advice, guidance and support during my PhD career. I established some good habits from him that I could benefit from them in all my life. I would like to thank my supervisor from LAL in France, Marumi Kado, for providing me such an excellent chance to work at LAL for two years, I also highly appreciate all the administrative stuff he did for me. I also would like to express my special appreciation and thanks to Jean-François Grivaz, my co-supervisor at LAL. The works I did in this thesis are mainly advised directly by him. We spent quite a lot of time in his or my office for discussing various questions related to my work. I acknowledge such time is one of my most precious memories in my Ph.D career, from where I started to have the idea about what is the right way to think as a researcher. I really enjoyed working with such an experienced, dedicated physicist. I owe so much of what I learned to his experience, enthusiasm and especially his will to share, teach, debate and discuss.

I would also like to thanks to my colleagues and friends in Shandong University and LAL, for the stimulating and beneficial academic discussions, as well as all the fun we have had in the last years. A special thanks to Yongke, not only for all the warm help I got from him, but also for the quite a lot of happy time we spent together in Jinan, in Qingdao, in Orsay and at CERN. Thanks to all the members of the ATLAS Hbb group that I have had the pleasure to work with.

Thanks to my parents and sister, what I am today is a result of your unconditional love and support since the day I was born.

Thanks to Yu Min, for your love and company. Meeting you is the best thing even happened in my life.

Contents

Chapter 1

Introduction

The Higgs boson [1–4] was discovered by the ATLAS and CMS Collaborations [5, 6] in 2012, from the analysis of proton-proton (pp) collision data produced by the Large Hadron Collider (LHC) [7]. Since then, using the Run 1 dataset collected at centre-of-mass energy of 7 TeV and 8 TeV, the properties of the discovered particle have been measured and were found to be compatible with those predicted by the Standard Model (SM) within uncertainties. The Run 2 dataset collected at centre-of-mass energy of 13 TeV provides an excellent opportunity to improve the precision of such measurements, and to challenge theory predictions further. The analyses of Higgs bosons decaying into vector bosons are entering an era of detailed precision measurements [8–14]. For the Higgs boson coupling to the fermions, the decay mode of $H \to \tau\tau$ was first observed from the combination of the ATLAS and CMS analyses [15]. Recently, the Higgs boson coupling to top quarks was directly observed by the ATLAS and CMS Collaborations [16, 17] respectively via the observation of the Higgs boson produced associated with a top-quark pair ($t\bar{t}H$).

The dominant decay of the Higgs boson in SM is into b-quarks pair, with approximately 58% expected branching fraction for a mass of $m_H = 125$ GeV [18]. However, a large amount of background arising from multi-jet production make a search in the dominant gluon-gluon fusion (ggF) production mode extremely challenging at the LHC. The associated production of a Higgs boson and a W or Z bosons [19] are the most sensitive production mode for probing $H \to b\bar{b}$ decays, since the leptonic decay of the W or Z bosons leads to efficient triggering and a significant rejection of the multi-jet backgrounds. This measurement can not only probe the dominant decay of the Higgs boson, and then allows the constraint of the overall Higgs boson decay width [20, 21], but also provide the best sensitivity to the WH and ZH production modes, which are crucial elements in the Higgs

boson measurements interpretation in effective field theories [22].

The $H \to b\bar{b}$ searches in the VH associated production (where V is used to denote W or Z) at the Tevatron by the CDF and D0 Collaborations showed an excess of events with a global significance of 3.1 standard deviations in the mass range of 120 GeV to 135 GeV, and a local significance of 2.8 standard deviations for a Higgs boson with a mass of 125 GeV [23]. With approximately 25 fb^{-1} data from Run 1, ATLAS and CMS reported an excess of events with a significance of 1.4 and 2.1 standard deviations [24, 25], respectively, and the combination of the ATLAS and CMS results yields an excess of events with a significance of 2.6 standard deviations [26]. The $H \to b\bar{b}$ searches have been performed also for the vector-boson fusion (VBF) [27–29] and $t\bar{t}H$ [30–34] production modes, and with high transverse momentum Higgs bosons [35], but with significantly lower sensitivities than for VH production.

This thesis describes a search for the SM Higgs boson decaying into a pair of b-quarks in the VH production mode with the ATLAS detector in Run 2 of the LHC. The pp collision data collected at a centre-of-mass energy of 13 TeV is used in the analysis, with an integrated luminosity of 79.8 fb^{-1}. Events are selected based on the number of charged leptons (electrons or muons) in 0-, 1- and 2-lepton channels, in order to explore the signatures of $ZH \to \nu\nu b\bar{b}$, $WH \to l\nu b\bar{b}$ and $ZH \to llb\bar{b}$, respectively. In order to maximize the sensitivity to the Higgs boson signal, multivariate discriminants are built from variables that describe the kinematics of the selected events. These multivariate discriminants are combined using a binned maximum-likelihood fit (referred to as the global likelihood fit), which allows to extract the signal strength and signal significance. Two other analyses are carried out to validate the signal extraction method: the dijet-mass analysis, where the mass of the dijet system is used as the main fit observable to extract the signal yield; the diboson analysis, where the nominal multivariate analysis is modified to extract the $VZ, Z \to b\bar{b}$ diboson process. In order to maximize the significance of $H \to b\bar{b}$ decay and VH production, the nominal multivariate analysis result is also combined with that of the previously published analysis of Run 1 data [24], with the other searches for $H \to b\bar{b}$ decays and with other searches for the Higgs boson produced in the VH production mode.

The results presented in this thesis are carried out not only by myself but also the other people in the working group, my personal contributions to the analysis are as follows:

- Developing and maintaining the analysis code for the 1-lepton channel.

- Multijet backgrounds estimation in 1-lepton channel, as presented in Section 5.5.2.
- Training and optimization of the multivariate discriminant used in the 1-lepton channel, as presented in Section 5.4.2.1.
- Various optimization of 1-lepton channel analysis, includes adding new analysis sub-channel and region, events selection optimization, etc, as presented in Section 5.6 and Section 5.10.
- Producing the 1-lepton channel inputs for the statistical analysis.
- Extensive fit studies to ensure the robustness of the fit model, provide the final fit results, as presented in Section 5.9.

The structure of the thesis is as follows:

- Chapter 2 covers a brief overview of the theoretical foundations that motivate the research work presented in this thesis.
- Chapter 3 gives a brief overview of the LHC and ATLAS detector.
- Chapter 4 presents the different reconstruction and identification procedures for each type of physics objects used for the research work presented in this thesis.
- Chapter 5 presents an analysis searching for the decay of the Standard Model Higgs boson to a b-quarks pair, in association with the production of a vector boson, using 79.8 fb^{-1} of pp collision data recorded by ATLAS during 2015 to 2017. The combination results with that of the previously published analysis of Run 1 data, with the other searches for $H \to b\bar{b}$ decays and with other searches for the Higgs boson produced in the VH production mode are also presented in this chapter.
- Chapter 6 presents a summary of the work described in this thesis.

Chapter 2

Theoretical Framework

This chapter covers a brief overview of the theoretical foundations that motivate the research work presented in this thesis. A brief overview of the Standard Model of particle physics is given in Section 2.1, then a brief description of Higgs mechanism in Section 2.2. Section 2.3 presents a short description of the physics in hadron collider, following a brief discussion on SM Higgs boson phenomenology in hadron collider in Section 2.4.

2.1 The Standard Model

The SM of particle physics [36, 37] is a Quantum Field Theory (QFT) developed during the second half of the 20^{th} century. The SM is also one of the most thoroughly tested theories of particle physics that has had a great success to explain experimental observations of particle physics. The Higgs boson, observed by the ATLAS and CMS experiments [5, 6] at the LHC in 2012, was the last missing particle predicted by the SM.

The SM of particle physics is a theory that describes the elementary particles and their interactions. Three are three out of the four fundamental interactions described in the SM: the electromagnetic interaction, responsible for the interactions between charged particles; the weak interaction, acting on the nuclear fission and radioactive decays; and the strong interaction, playing an essential role for confining quarks into hadron particles and binding neutrons and protons to create atomic nuclei. The gravitational force is currently not included in the SM.

The elementary particles in the SM can be basically divided in two groups: fermions and bosons. All the elementary particles have their own antiparticles, with same mass and spin. For some particles, like Z boson and photon, their

antiparticles are themselves.

Fermions, which have half-integer spin and are the building blocks of the matter, can be further grouped into two categories, the colourless leptons, and the colour charged quarks. The leptons and quarks can be grouped into three generations, and the first generation is the lightest while the third one is the heaviest.

For leptons, the first generation contains the electron and the electron neutrino, whilst the second and third generations are composed of the muon and the muon neutrino, the tau and the tau neutrino, respectively. Leptons do not undergo strong interaction, neutrinos all carry 0 charge hence do not undergo electromagnetic interaction, but they do interact through the weak interaction, the charged leptons interact through both the electromagnetic interaction and the weak interaction.

Quarks have also six flavours in three generations. The first generation contains the up quark and the down quark. The second generation is composed of the charm quark and the strange quark. And the third generation includes the top quark and the bottom quark . Quarks are the only known elementary particles in the SM whose electric charges are not integer multiples of the elementary charge. Up, charm and top Quarks carry $+\frac{2}{3}$ charge, while down, strange and bottom quarks carry $-\frac{1}{3}$ charge. Quarks interact through all the three fundamental interactions described in the SM, including the strong interaction. Due to the color confinement phenomenon, quarks can not be directly observed in isolation, and must clump together to form hadrons by strong interaction. There are two main types of hadrons, the meson composed of a quark and an antiquark and baryons made of three quarks.

The properties of all the fermions are summarized in Table 2.1.

Table 2.1: Summary of the properties of the half-integer spin fermions of the Standard Model [37].

	Leptons				Quarks			
Generation	Particle		Charge	Mass[MeV]	Particle		Charge	Mass[MeV]
First	electron neutrino	ν_e	0	$< 2.2 \times 10^{-3}$	up	u	$+\frac{2}{3}$	$2.2^{+0.5}_{-0.4}$
	electron	e^-	-1	0.51 ± 0.00	down	d	$-\frac{1}{3}$	$4.7^{+0.5}_{-0.3}$
Second	muon neutrino	ν_μ	0	$< 1.7 \times 10^{-1}$	charm	c	$+\frac{2}{3}$	1275^{+25}_{-35}
	muon	μ^-	-1	105.66 ± 0.00	strange	s	$-\frac{1}{3}$	95^{+9}_{-3}
Third	tau neutrino	ν_τ	0	< 1.55	top	t	$+\frac{2}{3}$	173210 ± 400
	tau	τ^-	-1	1776.86 ± 0.12	bottom	b	$-\frac{1}{3}$	4180^{+40}_{-30}

Bosons, which have integer spin, are the mediators of the three fundamental interactions described in the SM. Bosons can be further grouped into two categories, the gauge bosons with spin equal to one and scalar bosons with zero spin. Gauge bosons are composed with the W boson, the Z boson, the gluon and the photon. The W and Z bosons are known as the mediators of the weak interaction. Gluons act as the exchange particles between quarks in the strong interaction with eight independent types, known as the eight gluon colors. Photons are the mediators of the electromagnetic interaction. Both gluons and photons are massless and carry no charge. Currently, only one scalar boson has been found, the Higgs boson, with a mass around 125GeV. The properties of all the bosons are summarized in Table 2.2.

Table 2.2: Summary of the properties of the integer spin bosons of the Standard Model [37].

Name		J	Mass[GeV]
Photon	γ	1	0
W Boson	$W^{\pm}$	1	80.379 ± 0.012
Z Boson	Z	1	91.1876 ± 0.0021
Gluon	$g(\times 8)$	1	0
Higgs	H	0	125.18 ± 0.16

The SM is based on a gauge symmetry, $SU(3)_c \times SU(2)_L \times U(1)_Y$, in where $SU(3)_c$ is the non-abelian group describing the colour symmetry and strong interactions, $SU(2)_L \times U1(_Y)$ acts on electroweak interactions proposed by Glashow, Salam and Weinberg in 60s [38, 39]. Consider the Dirac Lagrangian density

$$\mathcal{L} = \overline{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi, \tag{2.1}$$

where the $\psi = \psi(x)$ is the Dirac spinor of a spin $\frac{1}{2}$ fermison. Consider the $U(1)$ gauge transformation

$$\psi \rightarrow \psi^{'} = e^{i\alpha(x)}\psi, \tag{2.2}$$

the Dirac Lagrangian then becomes

$$\mathcal{L} \rightarrow \mathcal{L}^{'} = \mathcal{L} - \overline{\psi}\gamma^{\mu}\partial_{\mu}\alpha(x)\psi. \tag{2.3}$$

To conserve the $U(1)$ symmetry of the Lagrangian during this transformation, the partial derivative ∂_{μ} needs to be replaced with the covariant derivative, D_{μ}.

This means, D_μ should satisfy

$$D_\mu\psi(x) \rightarrow D_\mu^{'}\psi(x)^{'} = e^{i\alpha(x)}D_\mu\psi(x), \tag{2.4}$$

in order to conserve the symmetry.

Equation 2.4 is satisifed for

$$D_\mu\psi(x) = \partial_\mu + ieA_m\mu, \tag{2.5}$$

where

$$A_\mu \rightarrow A_\mu^{'} = A_\mu - \frac{1}{e}\partial_\mu\alpha(x). \tag{2.6}$$

In quantum electrodynamics (QED),A_μ can be interpreted as the gauge field for the electromagnetic interaction with interaction strength e. The QED Lagrangian then can be written as

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi, \tag{2.7}$$

where $F_{\mu\nu}$ represents the kinetic energy term of the excitation of the gauge field. To generalise the interaction as an Abelian gauge group, $F_{\mu\nu}$ can be also defined in terms of D_μ,

$$[D_\mu, D_\nu]\psi \equiv ieF_{\mu\nu}\psi, \tag{2.8}$$

one can introduce the strong interaction as the symmetry of the $SU(3)_c$ group via expand this procedure to include non-Abelian gauge groups. By requiring an $SU(2)_L \times U(1)_Y$ symmetry of the SM Lagrangian, the unified electromagnetic and weak interactions can be introduced. Under the $SU(2)_L$ local gauge transformation

$$\psi \rightarrow \psi^{'} = e^{i\alpha(x)\times\frac{\sigma}{2}}\psi, \tag{2.9}$$

to conserve the symmetry of the Lagrangian, an additional 3 gauge fields, W_1^μ, W_2^μ, W_3^μ, are introduced with coupling strength g. In order to explain this in terms of a Lagrangian gauge symmetry, the weak interaction has both a vector and axial-vector ($V-A$) component. By the nature of the $V-A$ interaction, only the left handed (right handed) component of (anti-) particle spinors partake in the charged weak current interaction.

To describe the weak interaction, it is necessary to introduce the weak isospin quantum number, I_W. Left handed fermions are in weak isospin doublets with $I_W = \frac{1}{2}$, whilst right handed fermions are in weak isospin singlets with $I_W = 0$. Particle wave functions coupling to these bosons dependent on the third compo-

nent of the weak isospin charge, I_W^3, with $I_W^3 = \pm\frac{1}{2}$ for the left handed doublet, and $I_W^3 = 0$ for the right handed singlet. The charged flavour changing current is expressed as a linear combination of W_1^μ and W_2^μ,

$$W^{\pm\mu} = \frac{1}{\sqrt{(}2)}(W_1^\mu \pm iW_2^\mu). \tag{2.10}$$

Whilst it seems tempting to associate the Z boson with W_3, experimental observations indicate that the Z boson couples to both left and right handed electrons. Instead, the weak neutral current and photon are expressed as the product of the mixing of W_3^μ and B^μ, the boson of the $U(1)_Y$ symmetry, with coupling $g^{'}$. In the $U(1)_Y$ symmetry, the weak hypercharge is defined as

$$Y = 2Q - 2I_W^3, \tag{2.11}$$

with Q as the charge of the fermion. The mixing of W_3^μ and B^μ is defined in terms of the electroweak mixing angle θ_W,

$$\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \begin{pmatrix} cos\theta_W & sin\theta_W \\ -sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} B^\mu \\ W_3^\mu \end{pmatrix}. \tag{2.12}$$

From equating the $SU(2)_L$ and $U(1)_Y$ currents with the known interaction current of the photon, A_μ, the following relations are obtained

$$e = g^{'} cos\theta_W, \tag{2.13}$$

$$e = gsin\theta_W. \tag{2.14}$$

2.2 The Higgs Mechanism

In the description of SM in Section 2.1, all elementary particles are massless. The Higgs mechanism is introduced to the SM to explain the origin of mass through a process of spontaneous symmetry breaking [1–4]. This section first discuss the coupling between the gauge bosons and the Higgs, along with a discussion on mass for the fermionic sector.

Considering the complex isospin doublet of the Higgs field, with $I_W = \frac{1}{2}$ and $Y = 1$,

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.15}$$

one can introduce four additional degrees of freedom to the Lagrangian,

$$\phi^+ = \frac{\phi_1 + i\phi_2}{\sqrt{2}}, \phi^0 = \frac{\phi_3 + i\phi_4}{\sqrt{2}}, \tag{2.16}$$

and define the $SU(2)_L \times U(1)_L$ covariant derivative as

$$D_\mu = (\partial_\mu + i\frac{g}{2}\sigma W_\mu + iY\frac{g^{'}}{2}B_\mu). \tag{2.17}$$

The Lagrangian for the field can be written as

$$\mathcal{L}_\Phi = (D_\mu\Phi)^+(D^\mu\Phi) - V(\Phi), \tag{2.18}$$

with

$$V(\Phi) = \lambda(\Phi^+\Phi)^2 - \mu^2\Phi^+\Phi, \tag{2.19}$$

where μ and λ are scalar constants. One can see $V(\Phi)$ has minimum through

$$\frac{\partial V}{\partial(\Phi^+\Phi)} = \mu^2 - 2\lambda\Phi^+\Phi, \tag{2.20}$$

$$\frac{\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2}{2} = \frac{\mu^2}{2\lambda}. \tag{2.21}$$

Here we assume $\lambda < 0$ to ensure the potential to have the bounded ground state. There are two possibilities for μ^2. If $\mu^2 > 0$, the Higgs potential is then shown as the dashed line in Figure 2.1, without breaking the symmetry. If $\mu^2 < 0$, the Higgs potential is shown as the solid line in Figure 2.1, with the spontaneous symmetry breaking.

Through a phase rotation, one can set ϕ_1^2, ϕ_2^2, and ϕ_4^2 equal zero, and set $\phi_3 = \upsilon + H(x)$, the Φ can be written as

$$\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \upsilon + H(x) \end{pmatrix}, \tag{2.22}$$

then the Lagrangian can be writeen as

$$\begin{aligned}\mathcal{L}_\Phi = &\frac{1}{2}(\partial_\mu H)(\partial^\mu H) + \frac{g^2}{4}(\upsilon + H)^2(W_\mu^+ W^{\mu-}) \\ &+ \frac{1}{8}(g^2 g^{'2})Z_\mu Z^\mu(\upsilon + H) \\ &+ \frac{\mu^2}{2}(\upsilon + H)^2 - \frac{\lambda}{4}(\upsilon + H)^4.\end{aligned} \tag{2.23}$$

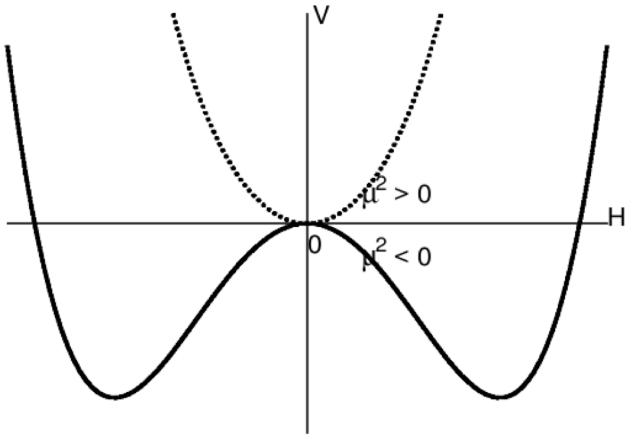


Figure 2.1: The Higgs potential with $\mu^2 > 0$ (dashed line) or $\mu^2 < 0$ (solid line).

Then the mass of the W, Z bosons, photon and Higgs boson can be explicitly written as:

$$m_W = \frac{1}{2} g \upsilon, \tag{2.24}$$

$$m_Z = \frac{1}{2}\sqrt{(g^2 + g'^2)}\upsilon = \frac{m_W}{cos\theta_W}, \tag{2.25}$$

$$m_A = 0, \tag{2.26}$$

$$m_H = \sqrt{-2\lambda\upsilon^2}. \tag{2.27}$$

The Higgs vacuum expectation value υ is related to the the Fermi coupling constant G_F:

$$\upsilon = \sqrt{\frac{1}{\sqrt{2}G_F}}. \tag{2.28}$$

G_F is the coupling constant associated with the weak interaction. The experimental determination of G_F comes from measurements of the muon lifetime [40], which is inversely proportional to the $\sqrt{G_F}$. The measured value of G_F is $1.166 \times 10^{-5}\,\mathrm{GeV}^{-2}$, then one can get the vacuum expectation value $\upsilon = 246.22\,\mathrm{GeV}$, with such a value, based on Equation 2.24 and Equation 2.25, m_W and m_Z can be calculated equal to 81 GeV and 91 GeV, respectively, which are in a very good agreement with the experimental measurements [41, 42].

Then one knows that the W and Z boson can obtain masses via spontaneous

symmetry breaking in the electroweak sector, whilst the photon can remain massless [1, 3]. The same scalar doublet used to generate the masses of the W and Z bosons is also sufficient to generate the masses for fermions, the interaction term between the scalar doublet and the fermion fields can be expressed as:

$$\mathcal{L} = -Y_f(\overline{Q}_L \Phi Q_R + \overline{Q}_R \overline{\Phi} Q_L), \tag{2.29}$$

where $Q_{L,R}$ are the left (right) handed fermion isospin doubletS (singletS) and Φ is the complex scalar Higgs field, and Y_f is called Yukawa coupling constant. This term is not given by the theory and needs to be obtained by experiment for every individual fermions. The Yukawa term is applicable to all fermions. Taking the first generation of leptons as an example, from Equation 2.29, it is found

$$\begin{aligned} \mathcal{L}_e &= -\frac{Y_e}{\sqrt{2}}\left[(\overline{\nu}_e, \overline{e})_L \begin{pmatrix} 0 \\ \upsilon + H(x) \end{pmatrix} e_R + \overline{e}_R\Big(0, \upsilon + H(x)\Big) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L\right] \\ &= -\frac{Y_e(\upsilon + H(x))}{\sqrt{2}}(\overline{e}e). \end{aligned} \tag{2.30}$$

The term $\frac{Y_e \upsilon}{\sqrt{2}}$ then can be interpreted as the electron mass term. Y_e is proportional to the electron mass and needs to be determined from experiment. The above formalism can equally be applied to the second and third generation of fermions, with different Y_f terms. This formalism gives masses only to the "down" type fermions, to include the massed for the "up" type quarks, another term must be added to the Lagrangian

$$\mathcal{L} = -Y_f, up(\overline{Q}_L \widetilde{\Phi}^c Q_R + h.c.), \tag{2.31}$$

$$\widetilde{\Phi}^c = -i\sigma_2 \Phi^* = -\frac{1}{\sqrt{2}} \begin{pmatrix} \upsilon + H(x) \\ 0 \end{pmatrix}, \tag{2.32}$$

where $h.c.$ is the Hermitian conjugate. Then masses are introduced for the "up" type quarks.

2.3 Physics at Hadron Colliders

The colliding particles are not fundamental objects in a proton-proton (pp) machine, the proton can be imagined to be formed from three quarks (uud). Due to quantum fluctuations, virtual pairs of quarks and gluons are created and re-

absorbed continuously and results in a tight interaction among the constituents, these phenomena are dominated by non-perturbative effects due to the low energies involved. A proton-proton interaction can be expressed as the incoherent superposition of the interactions between any two constituents of the two protons, each of them carries a fraction x_1 and x_2 of the incoming momentum of the proton. The formula for the cross section of a process can be written as:

$$\sigma = \sum_{i,j} \int dx_1 dx_2 \cdot f_i(x_1, \mu_F) f_j(x_2, \mu_F) \cdot \hat{\sigma}(\hat{s}, \mu_R, \mu_F), \tag{2.33}$$

where i, j are the different parton types and s is the squared centre-of-mass energy of the collider. $f(x)$ represents the parton distribution function (PDF), defined as the density of parton in the proton to carry a fraction x of the proton momentum. The dependence on a factorization scale μ_F is introduced to renormalize singularities arising from collinear emission of soft gluons and gluon splitting. The proton PDFs can be extracted from the data of deep inelastic scattering experiments (HERA) and from detailed measurements at hadron colliders [43]. According to the approaches and specific inputs used to extract the information from the data, several sets of PDFs are available. PDFs are extracted at a given scale and can be extrapolated to a different energy regime through the DGLAP [44, 45] evolution equations. Such equations describe the evolution of the strong coupling constant αs and the radiation branching properties with energy. Figure 2.2 represents an example of the PDFs extracted from a fit to HERA data [46], as can be seen, at high values of x, quarks carry most of the momentum of the proton and represent the dominant contribution, while at lower values of x, gluons and sea quarks represent the dominant contribution, so that the LHC can be also referred to as a "gluon collider".

$\hat{\sigma}$ is the cross section for the pp interaction occurring at a reduced squared energy $\hat{s}$. $\hat{\sigma}$ can be calculated with the perturbation theory. The theoretical estimations depend on the scales μ_R and μ_F since these scales can be only calculated up a given order in perturbation theory due to practical limitations. One typical way to calculate one of the main sources of theoretical uncertainty is performed by varying these scales by a factor of 0.5 - 2 around the nominal value and evaluating the effect on the result. In general, the higher the order of the calculation, the smaller the effect from scale related uncertainties are expected to be. Next to Leading Order (NLO) calculations which take into account virtual contributions to first order in αs, meaning the emission of extra partons and loop effects, are currently available for most of the processes. An increasing number of the theo-

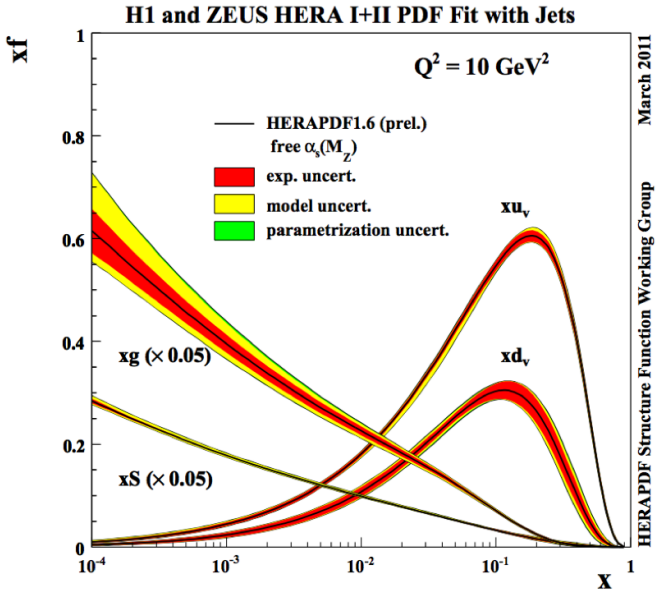


Figure 2.2: An example of the PDFs extracted from a fit to HERA data [46].

retical effort is moving to a more complete set of Next-to-Next to leading order (NNLO) calculation which includes two-loop effects.

Figure 2.3 shows a schematic view of the processes happening in *pp* collisions. Different colors represent different time (energy) scale in the event which are considered as subsequent steps in the calculation and event generation by Monte Carlo (MC) technique (more details about the MC technique are shown in Section 3.5). One parton from each proton can be involved in the main hard scattering collision (upper purple circle). The incoming partons and the partons produced in the hard scattering can undergo a set of radiation emissions or splitting into other partons (red lines). An effective approach in describing the various emissions is called parton shower approach. Radiations from incoming colliding parton are usually defined as initial state radiation (ISR) while emissions from final state partons are normally referred to as final state radiation (FSR). The evolution of the partons continues until the quarks and gluons combine into colourless states (light green circles) and subsequently form hadrons that decay into stable particles (dark green circles). This process is know as hadronization and is responsible for the evolution of the partons into a collimated spray of hadrons called a "jet". The interaction of the two proton remnants is usually defined as underlying events (bottom purple circle).

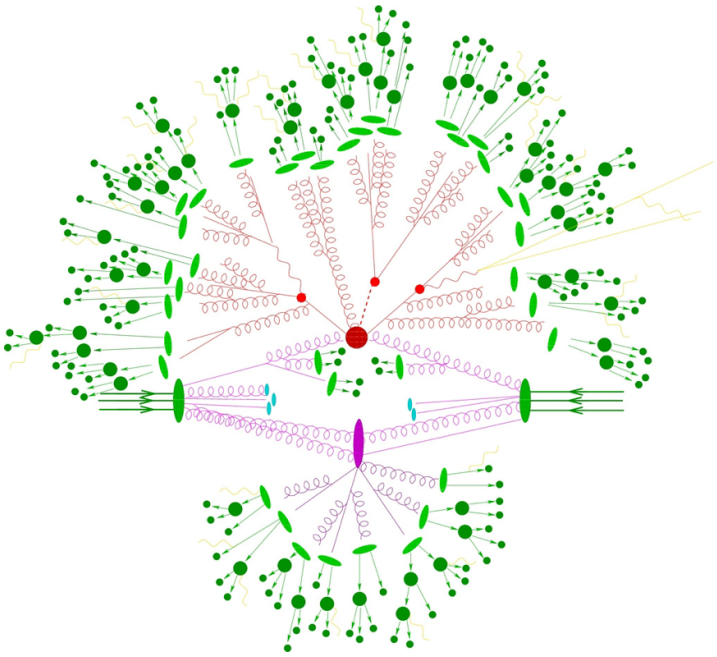


Figure 2.3: Schematic view of the interaction at a pp collider.

2.4 Standard Model Higgs Boson Phenomenology at Hadron Colliders

As discussed in Section 2.2, the Higgs boson plays a key role to give the masses of the bosonic and fermionic particles of the SM. The mass of the Higgs boson is not predicted by theory and has to be measured by the experiment. In this Section, I will discuss the basic phenomenology of SM Higgs boson at the hadron colliders.

2.4.1 Higgs Boson Production Mechanisms

For a SM Higgs boson with mass around 125 GeV, the Feynman diagrams for the main production modes at the LHC at a centre-of-mass energy of 13 TeV are presented in Figure 2.4. The related production cross-sections are presented in Figure 2.5. More detailed descriptions for the main production modes are as follows:

Gluon gluon fusion (ggF): As discussed in Section 2.3, the gluon density is highly dominant in colliding protons, hence ggF is the dominant Higgs boson production mode at the LHC. The production is mediated by a fermion loop, mainly via heavy quark loops (top, bottom) which have large Yukawa

coupling since the Higgs coupling strength is proportional to the mass of the particles.

Vector boson fusion (VBF): Two colliding quarks exchange a virtual W or Z boson, which emits a Higgs boson. This process results in a final state with two hard jets in the forward and backward regions of the detector (a description of the ATLAS detector is given in Chapter 3), which is a very clear and useful signature in the experiment.

Higgs Strahlung (VH): The Higgs boson is produced in association with a W or Z boson. The main contribution is from the quark-initiated process ($pp \to q\bar{q} \to VH$), with a sub-leading contribution from the loop-initiated process ($gg \to ZH$). From an experimental point of view, the presence of electrons or muons from the leptonic decay of the W or Z boson in the final state is an important handle to trigger these events, and provides a strong rejection for the overwhelming multijet backgrounds.

The total production cross sections for $pp \to q\bar{q} \to VH$ at NNLO QCD and NLO EW accuracy [21] are presented in Table 2.3, separately for the production modes of W^+H, W^-H and ZH at centre-of-mass energy of $\sqrt{s}$ = 13 TeV for m_H = 125 GeV. The NNLO QCD calculation is performed with VH@NNLO [47], renormalization and factorization scales are set to $\mu = \mu_R = \mu_F = m_{VH}$, and PDFs are taken from the set of PDF4LHC15_nnlo_mc PDFs. The NLO EW calculation is performed with HAWK [48, 49] and $\mu = \mu_R = \mu_F = m_V + m_H$, using NNPDF2.3QED PDFs which includes the EW corrections.

The cross section for $gg \to ZH$ at NLO QCD accuracy with VH@NNLO, including next-to-leading-log (NLL) effects is also quoted in Table 2.3. The uncertainties in the overall VH production cross-section from missing higher-order terms in the QCD perturbative expansion are obtained by varying the renormalization scale μ_R and factorization scale μ_F independently, from 1/3 to 3 times their original value. The PDF+αs uncertainty in the overall VH production cross-section is calculated from the 68% CL interval using the PDF4LHC15_nnlo_mc PDF set.

The charge asymmetry for the $W^{\pm}H$ cross section at a proton-proton collider as the LHC is due to the different PDFs for quarks and anti-quarks in protons. The much larger scale uncertainties for the ZH production compared to the WH production is mainly due to the contribution from $gg \to ZH$.

Table 2.3: Total WH and ZH cross section at centre-of-mass energy of $\sqrt{s} = 13$ TeV and $m_H = 125$ GeV, with scales and PDF+αs uncertainties. The separated contributions from W^+H, W^-H and $gg \to ZH$ are also presented.

	σ [fb]	Scales (%)	PDF+αs (%)	W^+H [fb]	W^-H [fb]	$gg \to ZH$ [fb]
WH	1373.00	$^{+0.5}_{-0.7}$	± 1.9	840.20	532.50	-
ZH	883.70	$^{+3.8}_{-3.1}$	± 1.6	-	-	123.30

Top fusion ($t\bar{t}H$): The Higgs boson is produced in association with top quarks pair. The rate for this process is very low at LHC, but it is very important to study the direct coupling of the Higgs boson to top quarks via this process (otherwise can only study at loop level in production or decay)

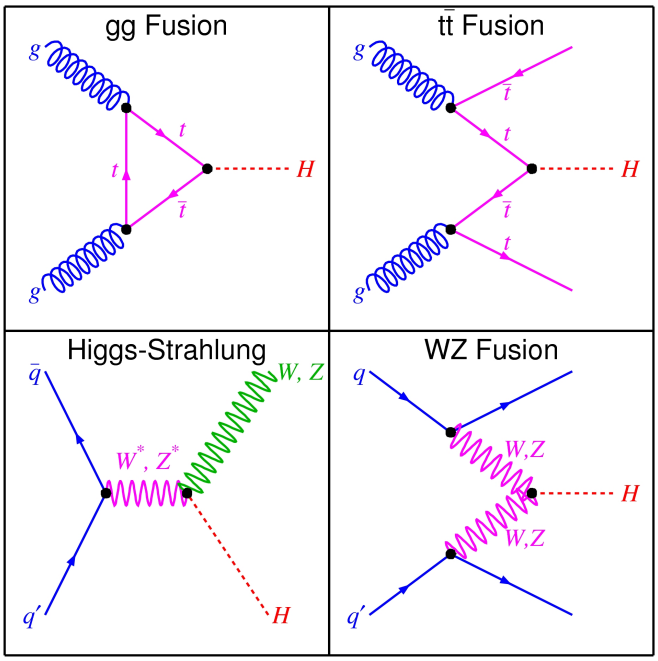


Figure 2.4: Feynman diagrams for the main production modes of the SM Higgs boson in LHC: ggF(top left), $t\bar{t}H$(top right), VH(bottom left) and VBF(bottom right).

2.4.2 Higgs Boson Decay Modes

The Higgs boson has no appreciable lifetime, as already discussed in Section 2.2, the Higgs boson coupling to particles is proportional to their masses,

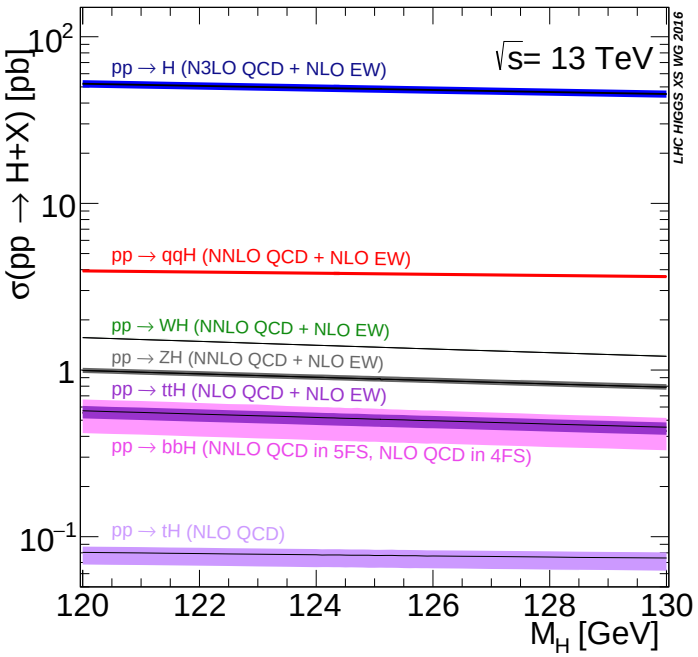


Figure 2.5: SM Higgs boson production cross sections at $\sqrt{s} = 13\,\text{TeV}$ as a function of Higgs boson mass.

therefore the Higgs boson tends to decay to the most massive particles allowed by kinematics. The Higgs boson decay to the massless particles, such as gluons or photons arises from loop corrections involving mainly top quarks and W bosons.

The branching ratio of any Higgs boson decay mode can be expressed as the ratio of the partial width to the total width which comes from the sum of all the possible partial widths:

$$BR(H \to XX) = \frac{\Gamma(H \to XX)}{\sum_i \Gamma(H \to Y_i Y_i)}. \tag{2.34}$$

Figure 2.6 shows the branching ratios for the different Higgs boson decay modes. For a Higgs boson with mass around 125 GeV, the dominant decay mode is to bottom quarks pairs, with a branching ratio about 58%. B-quark is the heaviest quarks in the SM that still form hadrons before undergoing a weak decay, and can be reconstructed as heavy flavour jet in the detector. $H \to b\bar{b}$ decays are accessible in the ATLAS experiment thanks to the experimental ability to identify jets from b-quarks (referred to as b-tagging, more details can be found in Section 4.6). However Several problems still exist for this decay channel:

- First, the presence of b-quark jets in the event is very difficult to be used for the online event selection. As a result, only the production modes with

additional signatures, like $t\bar{t}H$ or VH production mode, resulting in leptonic decays, can be really triggered on efficiently.

- Even for the $t\bar{t}H$ and VH production modes, large backgrounds from events with gluon jet or light quark (u, d, s) make the search for the $H \to b\bar{b}$ decays very challenging. The good performance of the b-jet identification algorithm is a critical factor for rejecting such backgrounds efficiently and keeping a reasonable fraction of events with b-jets at the same time.

- Apart from the background events with gluon or light-jets, the backgrounds events with real b-jets can also be produced copiously at the LHC. Such events can not be rejected by using the b-jet identification algorithm. One of the most important handle against such backgrounds is a good dijet invariant mass resolution.

The second largest decay mode is to W bosons pair, with one of the bosons off-shell. The decay to gluon pair is the third largest modes, but this final state is non-distinguishable from SM background hence is not studied at the LHC. The following decay modes are τ lepton pair production, charm quark pair production, ZZ^* production, and $\gamma\gamma$ production. The ZZ^* decay mode has a very clear experimental signature from the leptonic decays of the Z bosons : $H \to ZZ^* \to 4l$, even with extremely low production rate, this channel provides a distinct opportunity to study the Higgs boson's properties precisely. Similar with $4l$ channel, the $\gamma\gamma$ channel has a low decay rate but a relatively clean final state, and is also a key channel to study precisely the Higgs boson's properties. The latest mass measurement of the Higgs boson with ATLAS detector, combining the $4l$ and $\gamma\gamma$ channels, gives the value at $124.79 \pm 0.37\,\text{GeV}$ [50]. Finally the $\mu\mu$ decay mode, has also a very clean experimental signature, but very challenging due to the extremely suppressed branching fraction, this channel is very important to probe the nature of the Higgs boson coupling with the second generation fermions.

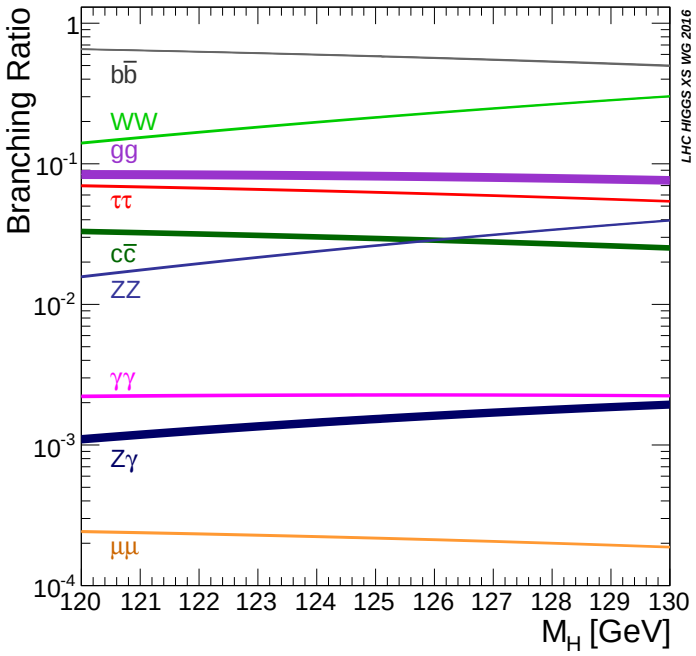


Figure 2.6: SM Higgs boson branching ratios for different decay modes, as a function of the Higgs boson mass.

Chapter 3

The Large Hadron Collider and the ATLAS detector

The research work presented in this thesis is based on the data collected by ATLAS detector at the Large Hadron Collider(LHC). In this chapter, Section 3.1 gives a brief overview of the LHC, Section 3.2 provides an overview of the design and operation of the ATLAS detector. Section 3.3 describes the concept of luminosity and details the size of the recorded dataset used for the analysis presented in this thesis. Section 3.4 presents the pile-up conditions in the corresponding dataset. Finally Section 3.5 gives a short description of the Monte Carlo simulation technique.

3.1 The Large Hadron Collider

The LHC [7] at the European Organisation for Nuclear Research (CERN) is the largest and highest energy particle collider in the world. It is housed in a circular tunnel with 27 km in circumference and 45-175 m in depth underground, which was previously used for the Large Electron-Positron Collider (LEP). The tunnel has four interaction points that are used to host the four main LHC experiments: ATLAS, CMS [51], LHCb [52] and ALICE [53]. There are three smaller experiments located nearby the main interaction points: LHCf [54], MoEDAL [55] and TOTEM [56]. An overview of the LHC complex is presented in Figure 3.1.

The LHC is a two-ring superconducting hadron accelerator and collider, the main physics programme at LHC relies on proton-proton collisions, but the machine is also capable of accelerating lead ions. For the proton-proton collisions, protons are extracted by ionizing the hydrogen gas in an electric field, and then

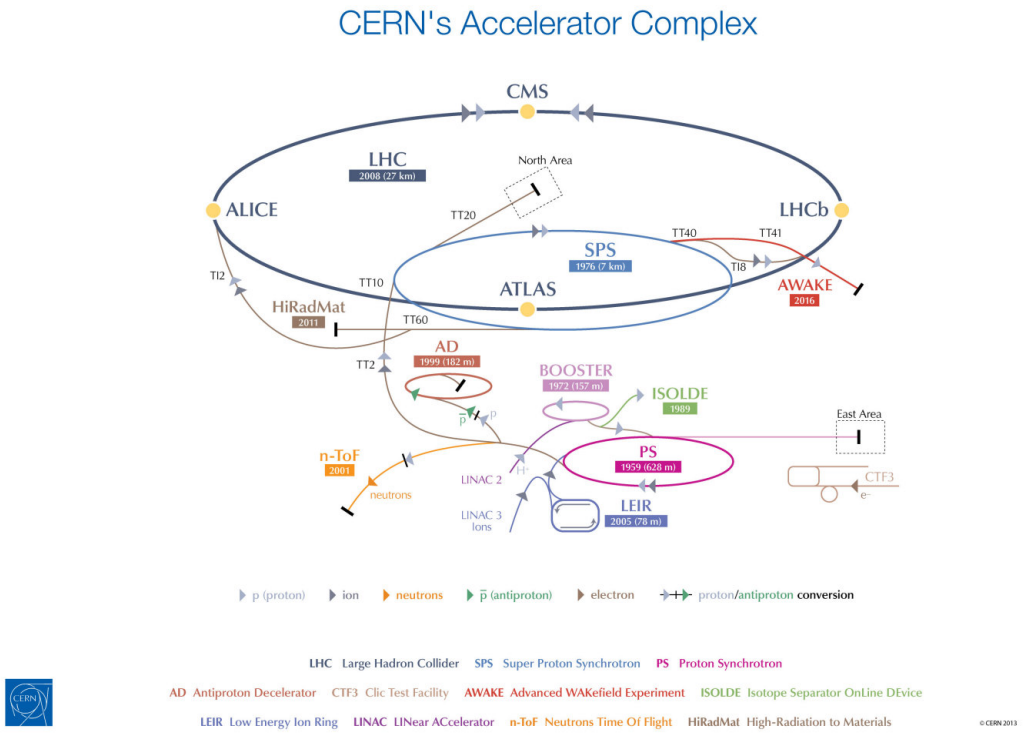


Figure 3.1: Overview of the CERN accelerator complex and the LHC.

first accelerated to an energy of 50 MeV by the linear accelerator Linac 2. The Proton Synchrotron Booster then accelerates the protons to 1.4 GeV, the Proton Synchrotron accelerates the protons to 25GeV afterwards. The beams then pass to the Super Proton Synchrotron (SPS) and reach the energy of 450 GeV before being injected into the LHC ring. In the LHC, the proton beams are finally accelerated to the collision energy. The protons in the LHC are arranged in bunches, each bunch contains approximately 10^{11} protons.

During the Run 1 (2010-2012) of the LHC, the centre of mass energies of proton-proton collisions were 7 TeV and 8 TeV, and the bunch spacing was set to 50 ns. Run 2 started in 2015 at $\sqrt{s} = 13\,\text{TeV}$ with the necessary upgrades of superconducting beampipe magnets during the long shutdown of 2012-2015, the bunch spacing also reduced to 25ns (except a very short period at the beginning of the 2015 data taking). Run 2 was end in December 2018, now the LHC is shutdown again to allow upgrades in preparation for Run 3. Run 3 is scheduled to run from 2021 to 2023 at $\sqrt{s} = 14\,\text{TeV}$. A new physics programme, called High Luminosity LHC (HL-LHC), is scheduled after Run 3, with further upgrades to facilitate instantaneous luminosities seven times larger than the current Run 2 luminosity.

3.2 The ATLAS Detector

The ATLAS [57] (A Toroidal LHC Apparatus) detector is one of the four main physics experiments at the LHC. It is designed to act as a general-purpose experiment, to cover the physics programs for both the precise measurements of SM processes and searches for beyond Standard Model (BSM) physics. In this section, the design and operation of the ATLAS detector are discussed. An overview of the ATLAS detector is presented in Section 3.2.1. Short descriptions of the inner detector, calorimetry system, muon spectrometer, forward detectors, trigger and data acquisition system are presented in Sections 3.2.2 - 3.2.6, respectively.

3.2.1 Overview

The ATLAS detector is the world's largest particle detector with a diameter of 25 m and length of 44 m. It is composed of several components and subsystems as shown in Figure 3.2 , the main subsystems are cylindrically symmetric with respect to the interaction point.

ATLAS detector is designed to be able to identify and reconstruct various

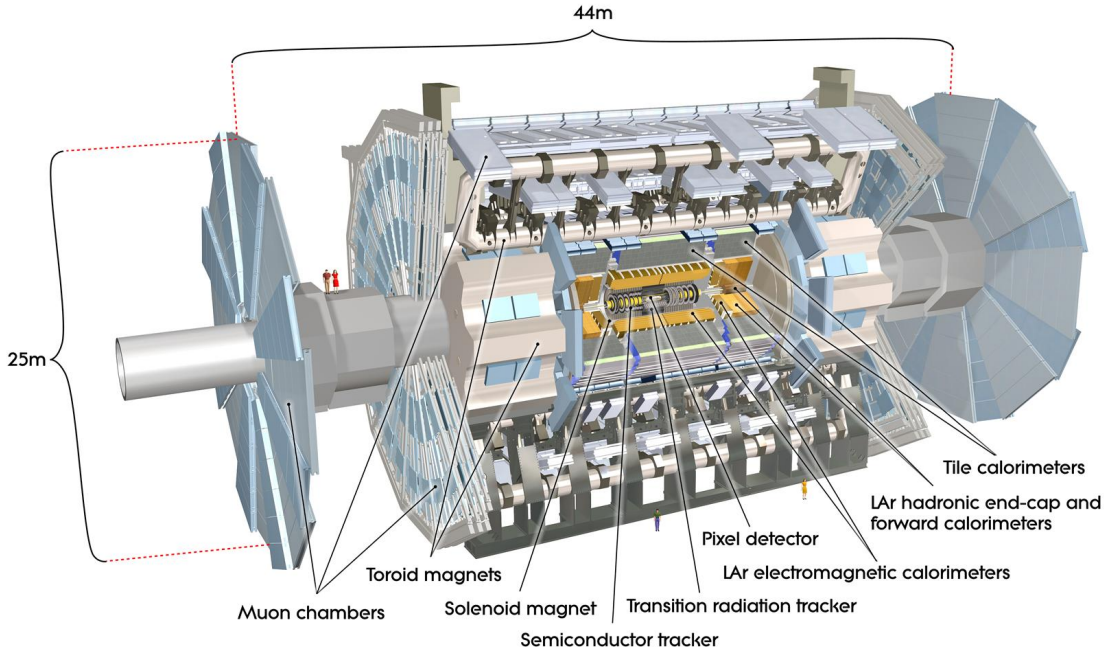


Figure 3.2: Cut-away view of the ATLAS detector.

physical objects, in order to achieve the different physics goals. A schematic view of the interaction of different types of particles within the ATLAS detector is shown in Figure 3.3.

ATLAS uses a right-handed coordinate system, the origin is defined as the interaction point in the centre of the detector. The z-axis coincides with the axis of the beam pipe, while the x-axis points towards the centre of the LHC ring, and the y-axis points upwards. The $x-y$ plane is defined as the transverse plane. The detector can be divided into two parts: A-side for positive values of z, and C-side for negative value of z. The azimuthal angle ϕ is measured around the beam axis, starting from the x-axis, whilst the polar angle θ is defined starting from the beam axis. The transverse momentum and energy in $x-y$ plane, p_T, E_T, are defined as $p_T = psin\theta$ and $E_T = Esin\theta$, respectively. A frequently used angular variable, transformed from the polar angle, the pseudorapidity η, is defined as

$$\eta = -ln(tan\frac{\theta}{2}). \tag{3.1}$$

For a particle in the transverse plane of the detector with $\theta = \frac{\pi}{2}$, the $\eta = 0$, for a particle with $\theta = 0, \pi$, the $\eta = \pm\infty$. The distance between objects in the (η, ϕ) plane is defined as

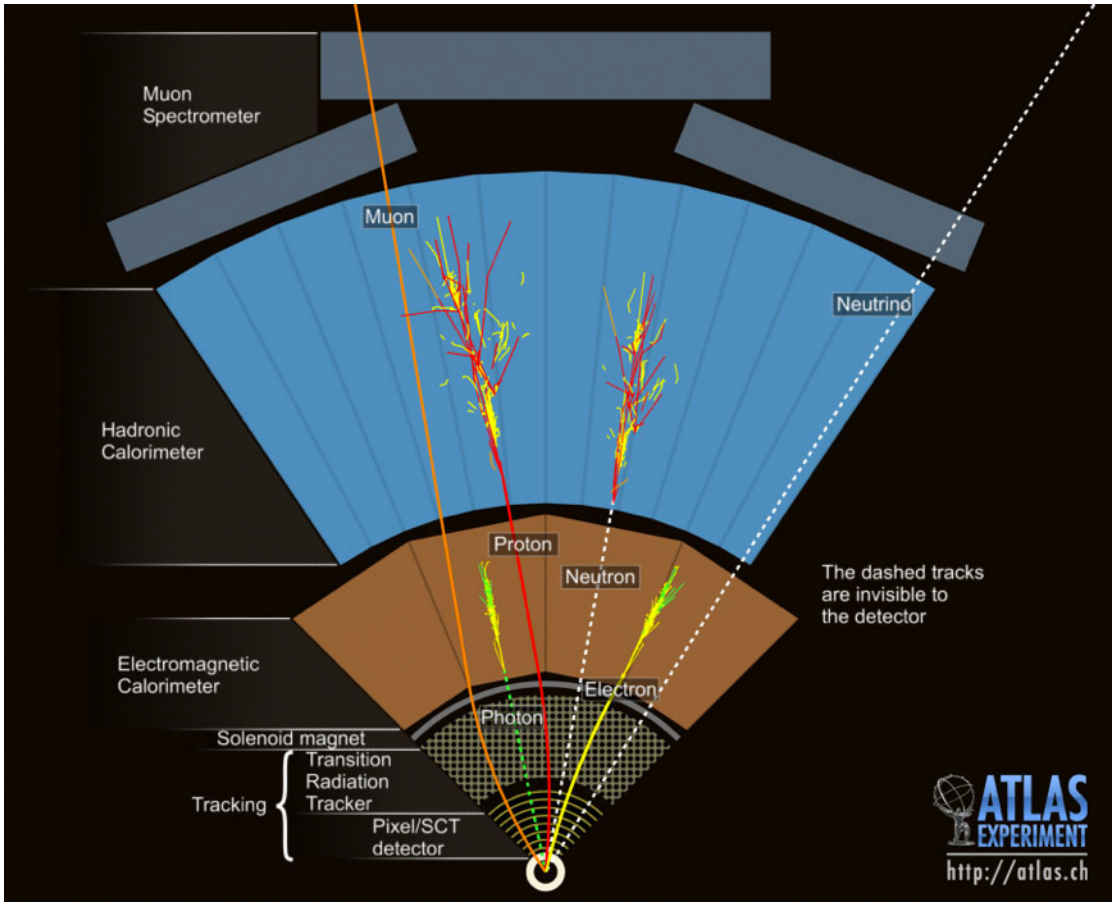


Figure 3.3: Interaction of the different particles in the ATLAS detector.

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \tag{3.2}$$

The magnet system is very important for the detector to provide the different magnetic fields required by the various parts of the apparatus. The ATLAS magnet system contains two magnetic subsystems, one for the inner detector and another one for the muon sectrometer. A solenoid [58] is installed between the electromagnetic calorimeter and the inner detector, and produces the strong magnetic field for the inner detector. Three sets of toroidal magnets [59, 60] are installed just outside the haronic calorimeter and provide the magnetic field for the muon spectrometer. The solenoidal magnet is 5.8 m long and has an inner radius of 1.23 m and an outer radius of 1.28 m. It is a coil of superconducting material with a 8 kA electric current, which provides a magnetic field of 2 T for the inner detector. The toroidal system is 25.3 m long and has an inner radius of 4.7 m and an outer radius of 10.05 m. It includes a barrel toroid (composed of 8 separate coils with an electric current of 20 kA) and two endcaps toroids, which provide a 0.5 T magnetic field in the central region and a 1 T magnetic field in the end-caps.

3.2.2 Inner detector

The Inner Detector [61] (ID) is the central part of the ATLAS detector, immersed in a 2 T magnetic field provided by a solenoid. The acceptance in pseudorapidity is $|\eta| < 2.5$ with full coverage in ϕ. The ID is designed to measure tracks and transverse momentum of charged particles with very good precision. It is also responsible for the reconstruction of the particles's primary and secondary interaction vertices. ID contains three complementary subsystems: a silicon pixel detector (pixel), a silicon micro-strip detector (SCT), a transition-radiation straw-tube tracker (TRT). A cut-away view of the ID is presented in Figure 3.4.

A new pixel-detector layer, insertable B-layer [62] (IBL), was inserted in the ATLAS detector in the first long shutdown period between the LHC Run 1 and Run 2 data taking, at a radius of approximately 30 mm and is the inner-most pixel layer of the ATLAS detector. This new layer was designed to achieve good spatial resolution with special care to be resistant to high radiation. The main improvements from the installation of the IBL are:

- the robustness of the tracking: loss of data in the Pixel B-layer (from for example radiation damage) strongly deteriorates the impact parameter resolution, and then deteriorates the performance of the b-tagging algorithms.

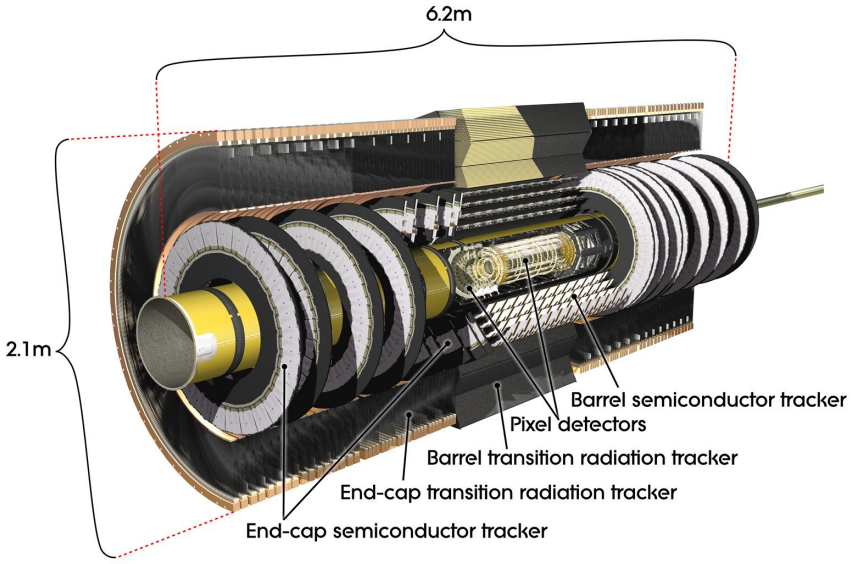


Figure 3.4: Cut-away view of the inner detector.

The b-tagging efficiency can be restored with IBL even in case of complete B-layer failure.

- the luminosity effects: The designed peak luminosity for Run 1 pixel detector has been excessed during LHC Run 2 data taking. This leads to high occupancy from pile up events that will affect the B-layer and limit the b-tagging efficiency. The IBL helps in keeping good tracking performance despite luminosity effects with low occupancy.

- tracking precision: IBL allows to improve the accuracy of track impact parameter reconstruction due to the very short distance to the interaction point, and hence can help to improve the b-tagging performance.

The inclusion of IBL leads an improvement of 10% for the b-tagging algorithm performance in Run 2 when comparing these to Run 1 without IBL included.

The pixel detector has a very high spatial resolution and covers a pseudorapidity region of $|\eta| < 2.5$. In the barrel region, the pixel detector are arranged in three cylindrical layers at a distance of 5.05, 8.85, 12.25 cm from the center of the beam pipe. In the end-caps the pixels are divided into three disks. All the pixels are segmented in $R - \phi$ and z, with a minimum active size 50×400 μm^2, achieving a resolution of 12 μm and 50 μm in $R - \phi$ and z, respectively.

Similarly to the pixel detector, the SCT is also divided into barrel and end-cap regions and covers the pseudorapidity range of $|\eta| < 2.5$. In order to provide 4

additional track points to contribute to the momentum measurement and track reconstruction, in the barrel region, SCT consists of 4 layers approximately at 30, 37, 44, 51 cm from the interaction point, each layer is composed of 2 microstrip sensors. The end-cap parts of the detector are organized in 2×9 disks of microstrips. The microstrip is 6.4 cm long with a pitch of 80 μm, and provides a resolution of 16 μm and 580 μm in $R - \phi$ and z, respectively.

The TRT is outermost-layer of the ID, and covers the pseudorapidity region of $\eta < 2$ and provides information only in the $R - \phi$ plane with a resolution of 130 μm. The TRT has the worse spatial resolution compared to the other ID sub-detectors, but provides up to 36 additional track points which helps a lot to improve the tracks reconstruction, particle identification and momentum measurement, and it also capable to identifying electrons thanks to the transition radiation photons. The TRT consists approximately 350,000 straw tubes, filled with xenon gas. The straws are parallel to the beampipe in the barrel and cover $560 < r < 1080$ mm for $|z| < 720$ mm. In the end-cap region, the straws are perpendicular to the beampipe and cover $617 < r < 1106$ mm for $827 < z < 2774$ mm.

The ID is able to determine a particle's momentum by measuring the curvature of the path of charged particles from hits in the detector. The overall ID track momentum resolution σ_{p_T}, as a function of the track transverse momentum p_T is

$$\frac{\sigma_{p_T}}{p_T} = 0.05\% \cdot p_T \oplus 1\%. \tag{3.3}$$

3.2.3 Calorimetry

The ATLAS calorimeter system is installed outside of the ID, and designed to measure the energy of particles precisely. The system is composed of two subsystems, the electromagnetic and the hadronic calorimeters. The electromagnetic calorimeter is used to measure the energy of electrons and photons, whilst the hadronic calorimeter is designed to measure the energy of hadrons, and limit the punch-through of hadrons into the muon system to make the sure the good performance of the muon chamber. A cut-away view of the calorimeters is shown in Figure 3.5

Electromagnetic Calorimeter The ATLAS electromagnetic calorimeter [63] is composed of two parts, a barrel section that covers pseudorapidity region $|\eta| < 1.475$ and an endcap section that covers $1.375 < |\eta| < 3.2$. It is a sampling calorimeter using Liquid Argon (LAr) as active material and lead as

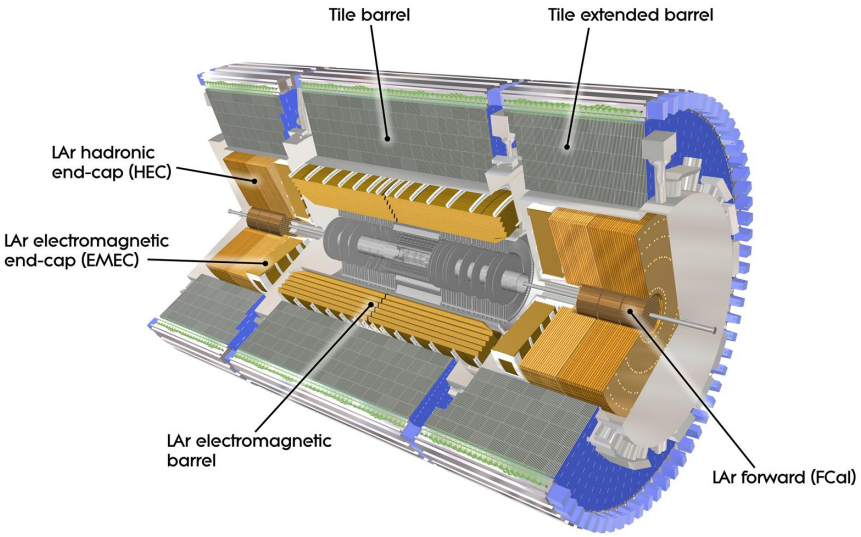


Figure 3.5: Cut-away view of the calorimetry system.

absorber. The barrel has an accordion geometry as shown in Figure 3.6 which provides uniform performances through the detector, with full coverage in ϕ and no crack regions. Electrons and photons will interact with the lead absorbers and build EM showers when entering the calorimeter, then the LAr calorimeter can measure the shower energy by collecting the charge at electrodes. The global collected charge is proportional to the energy of the particle that initiated the shower. The best possible energy resolution provided by electromagnetic calorimeter is

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.3\%. \tag{3.4}$$

Hadronic Calorimeter A hadron can interact by the strong force and also the electromagnetic force in the calorimeter, hence the interaction is fundamentally different compared to the interactions of electrons and photons. The hadron calorimeter [64] use steel as energy-absorbing material, whilst using scintillator tiles to sample the deposited energy. The barrel region of the hadronic calorimeter is composed of three parts, a central part ($|\eta| < 1.0$) and two extended barrels ($0.8 < |\eta| < 1.7$). The hadronic end-cap calorimeter and forward calorimeters provide additional pseudorapidity coverage up to $|\eta| < 4.9$, with both using liquid-argon technology. The hadronic end-cap calorimeter can cover the pseudorapidity region of $1.5 < |\eta| < 3.2$, with using copper as absorber, and the forward calorimeter covers $3.1 < |\eta| < 4.9$

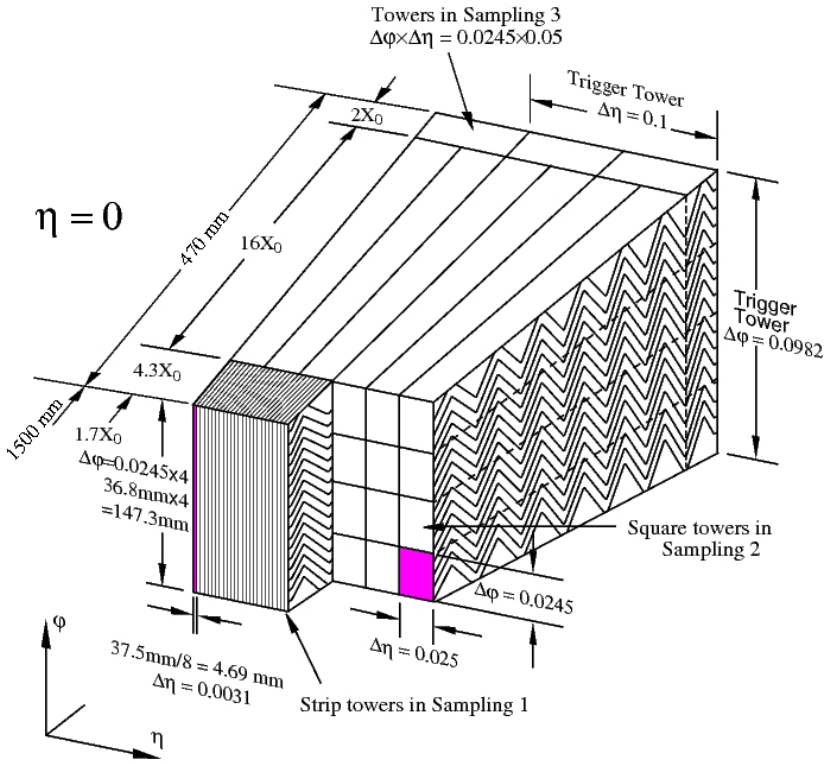


Figure 3.6: Schematic view of the EM barrel modules.

with three layers of absorber (one for copper, two for tungsten). The energy resolution for the hadronic calorimeter (barrel and end-cap) and forward calorimeter are

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%, \tag{3.5}$$

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%, \tag{3.6}$$

respectively.

3.2.4 Muon spectrometer

Muon is the only detectable particle that can pass through the ID and calorimeter without being absorbed. A dedicated muon spectrometer is needed to be able to measure the muon momentum with high precision. The muon spectrometer [65] (MS) is the outermost part of ATLAS and covers the region of $|\eta| < 2.7$. MS has its own trigger system and tracking chambers, the trigger system covers the region

of $|\eta| < 2.5$. A cut-away view of the MS is shown in Figure 3.7.

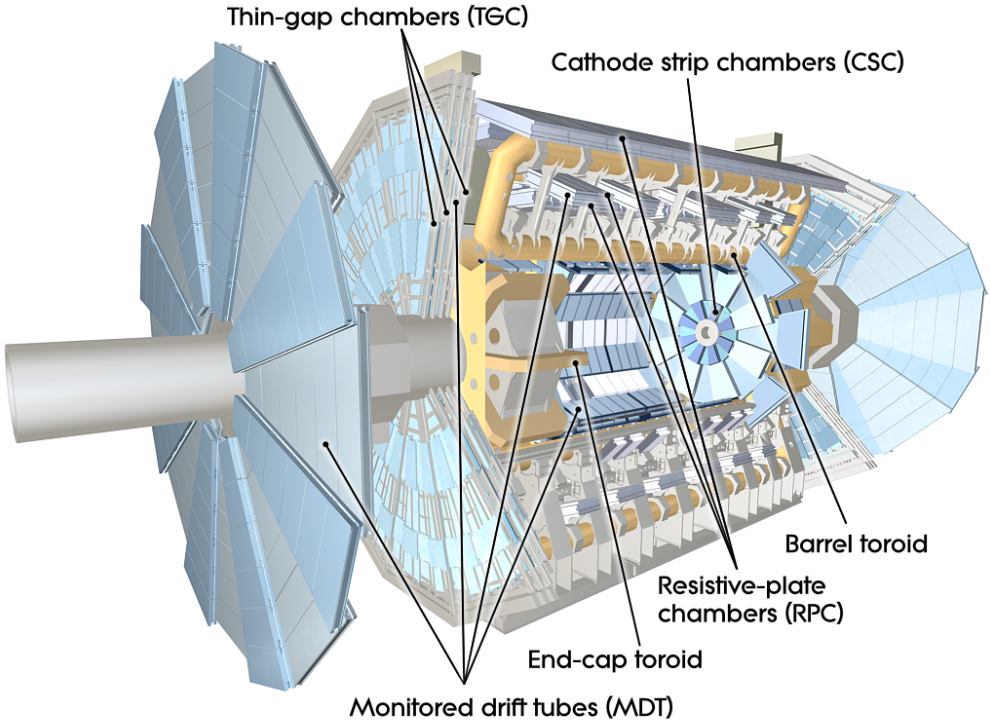


Figure 3.7: Cut-away view of the muon system.

The strong magnetic field, generated by the large superconducting air-core toroidal magnets in the barrel and end-cap regions, provides the capability to measure the muon momentum precisely. There are two types of subdetectors in the MS, one for the precision measurement of the particle momentum, another one for the quick response for the online triggering with coarser resolution. The Monitored Drift Tube (MDT) and the Cathode Strip Chambers (CSC) belong to the first type and cover the pseudorapidity for $|\eta| < 2.7$ and $2 < |\eta| < 2.7$, respectively. The Resistive Plate Chambers (RPC), as well as the Thin Gap Chambers (TGC) are the second type subdetectors, and cover the pseudorapidity for $|\eta| < 1.05$ and $1.0 < |\eta| < 2.4$, respectively.

Typically, the MS can provides momentum measurement with $\frac{\sigma_{p_T}}{p_T} \sim 10\%$ resolution for 1 TeV muons, and $\sim 3\%$ for 100 GeV muons. The muon system information is also able to combined with the information from inner detector to achieve a good efficiency and resolution for low-p_T muons.

3.2.5 Forward detectors

Apart from the main detector systems described above, there are three sets of small detectors designed to provide coverage in the very forward region to study inelastic pp scattering at small angles. From the closest to the farmost distances from the interaction point, the three detectors are : Luminosity measurement Using Cerenkov Integrating Detectors [66] (LUCID), which is the main relative luminosity monitor in ATLAS; Zero-Degree Calorimeter [67] (ZDC), which is designed to detect forward neutrons in heavy-ion collisions and Absolute Luminosity For ATLAS [68] (ALFA).

3.2.6 Trigger and data acquisition system

As discussed in Section 3.1, the bunch spacing for Run 2 data taking is 25 ns, corresponds to a rate of 40 MHz. Such rate is clearly too high for the read-out and storage capabilities allowed by the current ATLAS technology, therefore a dedicated trigger system is used to decide whether a event should be stored or not for offline analyses. The trigger system is designed to reduce event rate to $\sim$ 1 kHz, and providing a first discrimination between hard-scattering events and soft-physics events. The ATLAS trigger system is composed of two main levels, the Level-1 [69] (L1) trigger and high-level trigger [70] (HLT), as shown in Figure 3.8. L1 trigger is a hardware based trigger, designed to finds regions of interest (RoIs) in the calorimeters and muon spectrometer and reduce the event rate to approximately 100 kHz. The decision time for a Level-1 accept is about 2.5 μs. The RoIs are sent to the HLT in which more complicated selection algorithms are used with full granularity detector information in either the RoIs or the whole event. The HLT is able to reduce the event rate from 100 kHz to approximately 1 kHz, with a processing time of about 200 ms. There are two trigger selections are available, one is the "un-prescaled", another one is "prescaled". Prescaled triggers can help to limit the HLT output event rate further and avoid over-burdening the data taking system by retaining only a fraction of the events that passing the HLT. All triggers used in the analysis descripted in this thesis are un-prescaled.

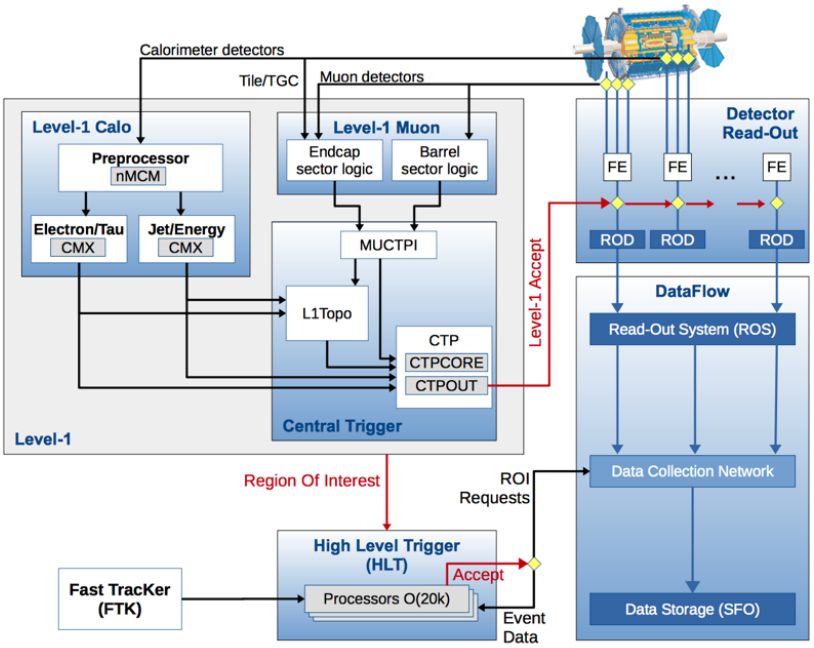


Figure 3.8: Schematic layout of the ATLAS trigger and data acquisition system in Run 2.

3.3 Luminosity

The number of events excepted for a certain process in a given dataset can be expressed as:

$$N = L \cdot \sigma = \sigma \int \mathcal{L} dt, \tag{3.7}$$

where $\mathcal{L}$ is the integrated luminosity over a certain period of data taking, σ is the cross section for the certain process and $\mathcal{L}$ is the instantaneous luminosity. $\mathcal{L}$ can be defined as a function of the beam parameters,

$$\mathcal{L} = \gamma \frac{n_b N^2 f_{rev}}{4\pi\beta^* \epsilon_n} R, \tag{3.8}$$

$$R = \frac{1}{\sqrt{1 + (\frac{\theta_c \sigma_z}{2\sigma_x})^2}}, \tag{3.9}$$

where the definition of the parameters are given in Table 3.1.

The cumulative luminosities delivered by the LHC and recorded by ATLAS for the 2015, 2016 and 2017 data-taking periods at $\sqrt{s} = 13\,\text{TeV}$ are shown in Figure 3.9. The ATLAS data-taking efficiencies are generally above 90%.

Table 3.1: Summary of the beam parameters.

Parameter	Definition
N	Protons per bunch
n_b	Number of bunches per beam
f_{rev}	Revolution frequency
γ	Relativistic γ factor
ϵ_n	Transverse emittance
β^*	β function at interaction point
$\frac{\theta_c}{2}$	Crossing angle at interaction point
σ_z	RMS bunch length
σ_x	RMS transverse beam size

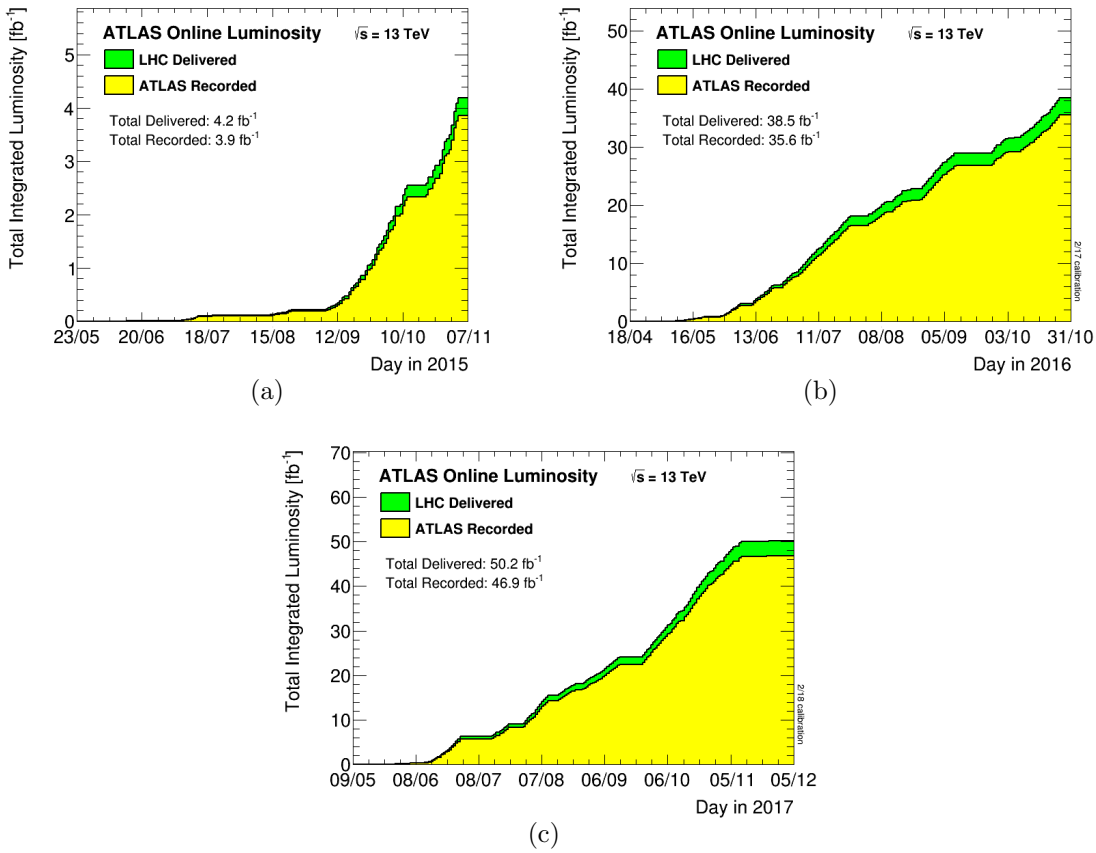


Figure 3.9: Cumulative luminosity versus time delivered by LHC (green) and recorded by ATLAS (yellow) during stable beams for pp collisions at $\sqrt{s} = 13\,\text{TeV}$ for the year of 2015 (a), 2016 (b) and 2017 (c).

3.4 Pile-up

The mean number of inelastic interactions per bunch crossing, referred to pile-up events, is also an important parameter related to the instantaneous luminosity, which can be expressed as

$$\mu = \frac{\mathcal{L}\sigma}{n_b f}, \tag{3.10}$$

where n_b is the number of colliding bunches and f is the bunch crossing frequency, σ is the total inelastic cross section for pp collisions.

Pile-up events are mainly soft interactions which considered as background to the hard interaction interested by the analysis. The level of pile-up effects also the physics objects measurement used in the analysis, the high pile-up worsens the resolution with which we can reconstruct hard-scattering events. The mean number of interactions per bunch crossing, $<\mu>$, for the 2015, 2016 and 2017 datasets are presents in Figure 3.10. The $<\mu>$ in 2015 data-taking was 13.4, and was increased to 25.1 and 37.8 in 2016 and 2017 data-taking due to the increased instantaneous luminosities, the average $<\mu>$ for the three years data-taking is 31.9.

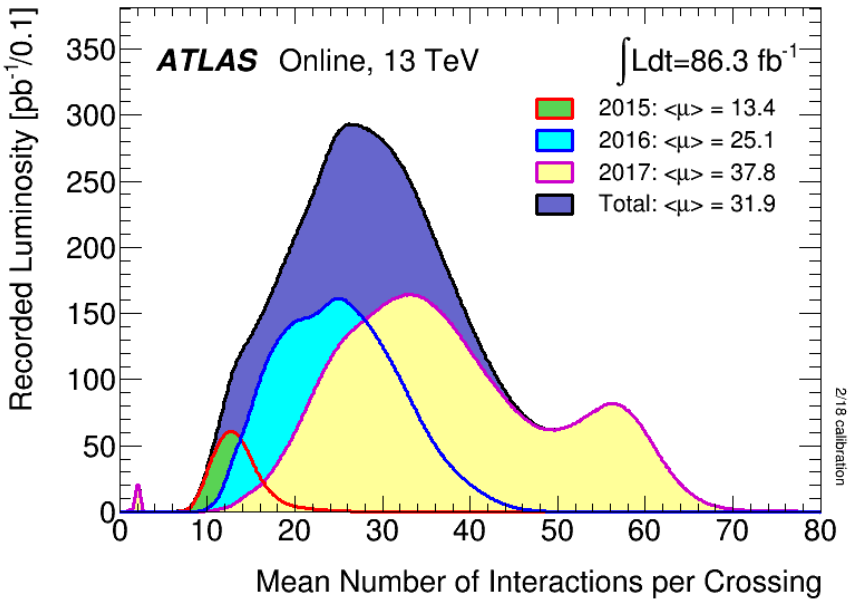


Figure 3.10: Mean number of interactions per bunch crossing for the 2015, 2016 and 2017 ATLAS pp datasets.

3.5 Monte Carlo Simulation

Monte Carlo (MC) simulation is a necessary and important component of experimental particle physics in different phases, such as the design of the detectors, the investigation of the physics reach of detector concepts, the development of data reconstruction software and the physics analysis.

Different Monte Carlo techniques are used to describe the different steps of the collisions as discussed in Section 2.3. The hard scattering processes are calculated in perturbation theory and the first emissions can be also included in the exact fixed order calculation of the scattering matrix element (ME). Parton shower effects are modelled through subsequent branching techniques and help in covering the kinematic range of soft and collinear radiation.

The simulation of MC events follows a series of steps in the ATLAS computing chain [71]. The outcome of the event generation is a list of stable particles stemming from the interaction point, the precise simulation of the interactions between such particles and the detector is performed with the GEANT4 [72] program. In order to reduce the CPU time, a less refined simulation, Atlfast-II [73] (AF2), is also available by applying a parameterized description of the particle showers in the calorimeters. After that, all the Monte Carlo events are also overlaid with additional inelastic events generated in order to simulate the effect from pile-up. A reweighting procedure is then applied to the MC samples so that the distribution of the average number of interactions per bunch crossing matches the corresponding distribution of the data sample.

The events are reconstructed and analyzed after the detector simulation by the same software chains that used also for data in order to convert the signal measured in each sensitive element of the detector to a physical quantity. Finally, in order to improve the description of data, the MC simulation is corrected in the description of the performance of the object reconstruction and identification for any residual disagreement with data.

Chapter 4

Object Reconstruction

A successful and efficient reconstruction and identification of the physics objects is very important for performing physics analyses with the data collected by the ATLAS detector. In this chapter, the different reconstruction and identification procedures are described for each type of physics objects used for the research work presented in this thesis.

4.1 Tracks and Vertices

The charged particles track reconstruction relies on the ID, which is surrounded by a 2 T solenoid magnet, and provides position measurements for charged particles within a range of $|\eta| < 2.5$. The first step for the track reconstruction algorithm [74] is to create the clusters in the Pixel and SCT, and the drift circles in the TRT, next, the clusters and drift circles are transformed into 3D space-points which correspond to a hit, a seed is then formed from 3 hits in either Pixel or SCT. A set of transverse momentum and impact parameter cuts are added on the seeds. For the seeds which pass such cuts, an additional requirement is then applied to them to match a fourth hit which is compatible with the particle's track estimated from the seeds. A combinatorial Kalman filter is then used to build track candidates from the chosen seeds. Track candidates are ranked based on a set of criteria. For the hits which are associated to more than one tracks, they are assigned to the highest ranked track. Track candidates are removed if they have less than 7 hits or $p_T < 400\,\text{MeV}$. Finally, the track candidates are extrapolated to the TRT if there is a valid set of matching drift circles, before the implementation of a track refit with all the information to improve momentum resolution.

As described in Section 3.3, the instantaneous luminosity is usually quite high in the LHC, which means multiple pp interactions can happen in a given bunch crossing. It is then very important to reconstruct the primary vertex where hard scattering process is originated from. To keep a low rate of fake tracks, only the reconstructed tracks which are selected with a set of tighter requirements are used for the primary vertex reconstruction with an iterative vertex finding algorithm [75]. A seed position is selected for the first vertex, then a fit is performed with the tracks and the seed to estimate the best vertex position. The fit is an iterative procedure, less compatible tracks are down-weighted and the vertex position is recomputed in each iteration. Once the vertex position is determined, tracks that are incompatible with the vertex are removed. The same algorithm is then used with the remaining tracks to reconstruct the other vertices. In case more than one vertex are reconstructed in an event, the one with the highest sum of squared track p_T is selected as the primary vertex. The efficiency of primary vertex reconstruction is predicted larger than 99% with a $t\bar{t}$ sample for $\mu = 30$.

4.2 Electrons

4.2.1 Reconstruction

The reconstruction of electron candidates matches the topological clusters (topo-cluster) of energy deposits in the calorimeter to the candidate tracks in the inner detector. The reconstruction proceeds in the following steps:

Topo-cluster reconstruction: The foundation of the electron reconstruction is the topological cell clustering algorithm [76]. Topo-cluster is formed in a way that follows closely spatial signal-significance patterns generated by particle showers. The basic observable, cell signal significance S_{cell}, which controls the cluster formation, is defined as

$$S_{cell} = \frac{E_{cell}}{\sigma_{noise,cell}}, \tag{4.1}$$

where E_{cell} is the energy deposited and $\sigma_{noise,cell}$ is the average noise in the cell from pile up or electronic noise. Both E_{cell} and $\sigma_{noise,cell}$ are measured on the electromagnetic energy scale, in which the energy deposited by electrons and photons are reconstructed correctly but the corrections for the loss of signal for hadrons from the non-compensating character of the ATLAS calorimeters are not included. The algorithm starts by forming proto-clusters

using a set of noise thresholds. The initial cell is required to satisfy $S_{cell} > 4$. The neighbouring cells with $S_{cell} > 2$ is then collected by the proto-cluster. The neighbour cells passing the noise threshold of 2σ is treated as a seed cell in the next iteration and collect each of its neighbors in the proto-cluster. Finally, a set of neighbouring cells with $S_{cell} > 0$ are added to the cluster. These thresholds are known commonly as "4-2-0" topo-cluster reconstruction.

Track reconstruction: Track reconstruction consists of two steps, the pattern recognition and track fit. The pattern recognition uses the pion hypothesis for energy loss from interactions with the detector material. If a p_T greater than 7 GeV track seed can not be extended to a full track with at least seven hits and it falls within one of the EM cluster region of interest, an electron hypothesis is then performed to allow for larger energy loss. The ATLAS Global χ^2 Track Fitter is used to fit the track candidates with either pion hypothesis or electron hypothesis.

Electron specific track fit: The obtained tracks are matched to EM clusters considering the distance in η and ϕ between the position of the track and the cluster barycenter.

Electron candidate reconstruction: one track is chosen as "primary" track based on a specific algorithm if several tracks fullfil the matching condition. The algorithm takes into account the distance of the track and cluster, and also the number of pixel hits and holes.

The track associated with the electron is also required to be compatible with the primary vertex. Two conditions are applied to match this requirement: $d_0/\sigma_{d_0} < 5$ and $\Delta z_0 sin\theta < 0.5$ mm. The impact parameter d_0 is the distance of closest approach of the track to the primary vertex in the r-ϕ projection, σ_{d_0} represents the estimated uncertainty of the d_0 parameter. Δz_0 is the distance along the beam-line between the point where d_0 is measured and the primary vertex, and θ is the polar angle of the track.

4.2.2 Identification

It is possible for a non-prompt physics object (such as electron from photon conversion or semi-leptonic decay of a heavy flavour hadron) being reconstructed as

a prompt electron (from heavy resonance decays, such as $W \to e\nu, Z \to ee$). Electron identification algorithm [77] is used to determine whether the reconstructed electron is signal-like object or such background-like object.

A likelihood-based (LH) method with multivariate analysis (MVA) technique is used as the baseline identification algorithm, to evaluates several properties of the electron candidates simultaneously when making a selection decision. The LH method uses the probability density functions (PDFs) of the discriminating variables for both signal and background, to calculate an overall probability to determine the electron candidate is signal or background. The discriminant $d_{\mathcal{L}}$ is defined as

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \tag{4.2}$$

where

$$\mathcal{L}_{S(B)}(\vec{x}) = \prod_{i=1}^{n} P_{s(b),i}(x_i). \tag{4.3}$$

$\vec{x}$ is the vector of discriminating variable values, $P_{s(b),i}(x_i)$ refers to the signal (background) probability function of the i^{th} variable evaluated at x_i.

The ID algorithm provides three levels of identification operating points, called Loose, Medium and Tight in descending order of signal efficiency. The only difference for these three operation points is the selections used on the LH discriminant, while the variables used to define the LH discriminant are the same. The ID operating points are optimised in several $|\eta|$ and E_T bins. For the electron candidates with $E_T = 25\,\text{GeV}$, the signal (background) efficiencies for these three operating points are in the range from 90 to 78% (0.8 to 0.3%).

4.2.3 Isolation

The isolation requirement is adopted to further reduce the non-prompt electron backgrounds, by using the isolation variables which are capable to quantify the energy of the particles produced around the electron candidate. Two discriminating variables [77] are defined for this purpose :

Calorimeter isolation , $E_T^{cone0.2}$, defined as the sum of transverse energies of the EM clusters, within a cone of $\Delta R = 0.2$ around the candidate electron cluster.

Track isolation , $p_T^{varcone0.2}$, defined as the sum of transverse momenta of all the qualitied tracks, within a cone of $\Delta R = min(0.2, 10\,\text{GeV}/p_T)$ around the candidate electron track. The cone size gets smaller for larger p_T of the electron, to take into account the situation that the other objects can end up very close to the electron in boosted signatures or very busy environnements.

There are basically two types of isolation operating points are defined based on the calorimeter and track isolation variables. First one is the efficiency targeted operating points, in where varying requirements are used in order to obtain a given isolation efficiency. Second one is the fixed requirement operating points, where the upper thresholds on the isolation variables are constant. The definition of the various electron isolation operating points are shown in Tabel 4.1. For same operating points, the ratio of isolation variable and electron p_T is used to improve performance over the full p_T spectrum.

Table 4.1: Summary of the electron isolation operating point definitions.

	Efficiency / Cut value		
Operating point	calorimeter isolation	track isolation	total efficiency
LooseTrackLoose	-	99%	99%
Tight	96%	99%	95%
Gradient	$0.1143\% \times E_T + 92.14\%$	$0.1143\% \times E_T + 92.14\%$	90%/99% at 25/60 GeV
FixedCutTightTrackOnly	-	$P_T^{varcone0.2}/p_T < 0.06$	-
FixedCutHighPtCaloOnly	$E_T^{cone0.2} < 3.5\,\text{GeV}$	-	-

4.2.4 Simulation Correction Factors from Efficiency Measurement

The efficiency to find and select an electron in the ATLAS detector is divided into different components, like reconstruction, identification, isolation, and trigger efficiencies. The total efficiency can be written as:

$$\varepsilon_{total} = \varepsilon_{reconstruction} \times \varepsilon_{identification} \times \varepsilon_{isolation} \times \varepsilon_{trigger}. \tag{4.4}$$

Due to the imperfect modelling of the MC simulation, such as tracking properties or shower shapes in the calorimeters, the efficiencies are needed to be estimated both in data and in simulation, the ratio between data and MC efficiencies is then used as a multiplicative correction factor for MC. The tag-and-probe method [77],

which employs events containing well-known resonance decays to electrons, like $Z \to ee$ and $J/\psi \to ee$, has been used to measure each of these efficiencies.

4.3 Muons

4.3.1 Reconstruction

The reconstruction of muon candidates [78] uses tracks in the ID and MS. For the tracks in MS, a χ^2 fit is used with the hits information. In the ID, muon tracks are reconstructed just like the other charge particles, the combined ID-MS muon reconstruction is then performed with the information from individual subdetectors. Based on which subdetectors are used in the reconstruction, four muon types are defined.

Combined (CB) muon: a global refit is performed with the hits information from both MS and ID to form the combined track.

Segment-tagged (ST) muons: if a track in the ID is associated with at least one local track segment in the MS after the extrapolation, the track is then classified as a muon. The ST muons are mainly used in the situation that muons pass only one layer of MS chambers, due to either their low p_T or they fall in the regions with reduced MS acceptance.

Calorimeter-tagged (CT) muons: track in the ID is classified as a muon if it can be matched to an energy deposit in the calorimeter compatible with a minimum-ionizing particle. This type recovers the acceptance in the region where the MS is only partially instrumented.

Extrapolated (ME) muons: the muon track reconstruction is based on the MS track only with a loose requirement on the compatibility between the track and interaction point. ME muons are used to extend the acceptance into the region $2.5 < |\eta| < 2.7$, which is out of the ID coverage.

Similar with electron reconstruction, the track associated with the muon is also required to be compatible with the primary vertex, by requiring $d_0/\sigma_{d_0} < 3$ and $\Delta z_0 sin\theta < 0.5mm$.

4.3.2 Identification

It is possible for a non-prompt physics object being reconstructed as a prompt muon, such as muon from inflight decay of a hadron (pion and kaon)or semi-leptonic decay of a heavy flavour hadron. Muon identification [78] is designed to suppress background , and keep high efficiency for the prompt muon by applying quality requirement. In the ID, the non-prompt muons originating from in-flight decays of pion and kaon are usually characterized by the presence of a distinctive "kink" topology in the reconstructed track, which leads to a worse fit quality of the combined track, compared to a prompt muon. Several discriminating variables are used to identify the signal muons :

$\frac{q}{p}$ significance: absolute value of the difference in the ratio of the charge and momentum of the muon measured in the ID and MS, divided by their uncertainties.

ρ': absolute value of the difference in the transverse momentum measured in the ID and MS, divided by the p_T of the combined track.

normalized χ^2 of the combined track fit.

Four muon quality operation points are provided using different cuts on these discriminating variables:

ID-Loose all muon types are used to maximize the reconstruction efficiency while providing good-quality muon tracks, the criteria are optimized for the benefit of the $H \to ZZ^* \to 4l$ analysis.

ID-Medium this operation point uses only the CB and ME tracks, and minimizes the systematic uncertainties associated with muon reconstruction and calibration.

ID-Tight this operation working is optimized to maximise the purity of muons by using only the CB muons which satisfy the required selections.

ID-High-p_T this operation points is optimized for the high-mass W'/Z' resonances analysis, and provides better momentum resolution for tracks with p_T above 100 GeV.

For a muon with p_T between 20 GeV and 100 GeV, the efficiency for the prompt and also the non-prompt muon identification provided by these four operation points are shown in Table 4.2.

Table 4.2: Summary of the efficiency of the four muon identification operation points for the prompt signal muon and non-prompt.

Operation Points	efficiency for prompt muon [%]	efficiency for non-prompt muon [%]
Loose	98.1	0.76
Medium	96.1	0.17
Tight	91.8	0.11
High-p_T	80.4	0.13

4.3.3 Isolation

Similar as electron, the isolation requirement is applied to muon to further reduce the non-prompt muon backgrounds. Three discriminaiing variables [78] are defined for this purpose :

Calorimeter isolation: $E_T^{cone0.2}$, defined as the sum of transverse energies of the topological clusters, within a cone of $\Delta R = 0.2$ around the candidate muon.

Track isolation with variable radius: $p_T^{varcone0.3}$, defined as the sum of transverse momenta of all the qualitied tracks, within a cone of $\Delta R = min(0.3, 10\,\text{GeV}/p_T)$ around the candidate muon track.

Track isolation with fixed radius: $p_T^{cone0.2}$, defined as the sum of transverse momenta of all the qualitied tracks, within a cone of $\Delta R = 0.2$ around the candidate muon track.

The definition of the various muon isolation operating points are shown in Tabel 4.3. For same operating points, the ratio of isolation variable and muon p_T is used to improve performance over the full p_T spectrum.

Table 4.3: Summary of the muon isolation operating point definitions.

	Efficiency / Cut value		
Operating point	calorimeter isolation	track isolation	total efficiency
LooseTrackLoose	-	99%	99%
Gradient	$0.1143\% \times E_T + 92.14\%$	$0.1143\% \times E_T + 92.14\%$	90%/99% at 25/60 GeV
FixedCutTight	$E_T^{cone0.2}/p_T < 0.06$	$p_T^{varcone0.3}/p_T < 0.06$	-
FixedCutHighPtTrackOnly	-	$p_T^{cone0.2} < 1.25\,\text{GeV}$	-

4.3.4 Simulation Correction Factors from Efficiency Measurement

Similar with electrons, the efficiency to find and select muons in the ATLAS detector is also divided into different components, like reconstruction, identification, isolation, and trigger efficiencies. The same tag-and-probe method [78] is used to measure each of these efficiencies, by using the $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events. The correction factors are applied to the simulated samples to correct the measured MC efficiencies to the data efficiencies.

4.4 Hadronic Tau

The tau lepton is the only lepton that can decays into hadrons. Tau lepton decays can be basically divided into two modes based on the products of the decays: the leptonic decay that tau lepton decays into tau neutrino, electron (muon) and electron (muon) antineutrino; the hadronic decay that the tau lepton decays into for examples a charged pion, a neutral pion, and a tau neutrino, or three charged pions and a tau neutrino, etc. About 65% tau leptons undergo the hadronic decay (τ_{had}). τ_{had} is reconstructed use the procedure [79] described in Section 4.5. τ_{had} is required to have $p_T > 10\,\text{GeV}$ and $|\eta| < 2.5$ (excluding the transition region corresponding to $1.37 < |\eta| < 1.52$), with exactly 1 or 3 matching charged tracks. The dedicated τ_{had} calibration is developed to correct the energy deposition measured in the calorimeter to the average energy carried by the measured decay products at the generator level. The Boosted Decision Tree (BDT) based τ_{had} identification algorithm is designed to reject backgrounds from hadronic jets. Three τ_{had} identification working points are provided, labelled as loose, medium and tight, and correspond to different τ_{had} identification efficiency. For 1-track case, the target efficiencies are 0.6, 0.55 and 0.45 for loose, medium and tight working points, for 3-track case, the corresponding efficiencies are 0.5, 0.4 and 0.3.

4.5 Jets

In the hadron collider, the quarks and gluons always fragment and hadronize immediately after the production, the only observable object in the detector for these particles is a spray of hadrons, which is called as jet. There are two types of jets reconstructed in ATLAS, calorimeter jets and track jets, using the same anti-k_t

algorithm but with different distance parameters R. For the analysis presented in this thesis, the calorimeter jets with $R = 0.4$ is used. Details for the reconstruction and calibration for this type of jet are discussed in this section.

4.5.1 Reconstruction

Jet reconstruction [80] starts with topological clusters which built from calorimeter cell with more details have already been given in Section 4.2

The anti-k_t algorithm [81] is then used to reconstruct the calorimeter jets by clustering topological clusters. Two distance measures are defined as

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2}, \tag{4.5}$$

$$d_{iB} = k_{ti}^{2p}, \tag{4.6}$$

here $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and k_{ti},y_i and ϕ_i are the transverse momentum, rapidity and azimuth of particle i, respectively. p is a parameter to govern the relative power of the energy versus geometrical scales and equals to -1 in case of anti-k_t algorithm. R is the usual radius parameter and related to the radius of the jet. $R = 0.4$ is used for the studies presented in this thesis. d_{ij} can be introduced as the distance between cluster i ans j, whilst d_{iB} can be introduced as the distance between cluster i and the beam (B). The anti-k_t algorithm is an iterative procedure that starts from computing all distances of d_{ij} and d_{iB}. If the smallest distance is d_{ij}, the four moment of i and j are then combined, else if the smallest distance is d_{iB}, cluster i is removed and is called as a "jet". The procedure is repeated until all the topological clusters are clustered into jets.

4.5.2 Calibration

There are two main purposes for the jet energy calibration. First, due to the energy scale of reconstructed jets does not correspond to the truth-particle jet energy scale (JES), a dedicated jet energy calibration is needed to calibrate the reconstructed jet energy to the corresponding truth-particle jet. Second, the jet energy calibration has to account for the differences of the energy scale of jets between data and MC. A few steps are included in the jet calibration [82]. First, the jet direction is corrected to point back to the primary vertex, then the pile-up effect is removed using an area-based subtraction procedure, next, the jet energy is calibrated by applying the corrections derived from the MC simulation, finally,

an additional correction is applied to the jets in data, to calibrate their energy to the correct value based on in situ studies. This correction is derived from well understood processes, like γ/Z + jets events, using the balance between the energy of the recoiling jet and the well understood decay of γ/Z. The Z + jets events are used for jets with $20\,\text{GeV} < p_T < 500\,\text{GeV}$, whilst γ + jets events are used for jets with $36\,\text{GeV} < p_T < 950\,\text{GeV}$, a system of low-$p_T$ jets is used for high p_T jets with $950\,\text{GeV} < p_T < 2\,\text{TeV}$.

4.5.3 Jet Cleaning and Pile-up Jets Suppression

Reconstructed jets in the ATLAS detector can originate not only from a hard scatter proton collision but also from a non-collision background process or noise in the calorimeters. The development and implementation of a set of selection criteria to distinguish the jets from the different originations is known as jet cleaning [83]. Variables used for the jet cleaning can be basically divided into three categories: variables built from signal pulse shape in the LAr calorimeters, which can help to reduce fake jets due to coherent or sporadic noise in the LAr calorimeters; variables based on jet energy ratio, such as the ratio of jet energy deposited in the electromagnetic calorimeter to the jet total energy; track based variables, such as the ratio of the scalar sum of the p_T of the tracks coming from the primary vertex associated to the jet to the total p_T of the jet. Dedicated selections are formed based on such variables and applied to the jets. The events containing the jets failed the selections are removed. Apart from the jet cleaning, a multivariate combination of two track-based variables, called jet-vertex-tagger (JVT) [84], is developed to further suppress pile-up jets and provides stable hard-scatter jet efficiency in term of number of reconstructed primary vertices. Three working points have been derived for jets with $|\eta| < 2.4$ and $20\,\text{GeV} < p_T < 60\,\text{GeV}$, the cut values and average efficiencies for hard scatter jets are shown in Table 4.4.

Table 4.4: Summary of the cut values and average efficiencies for hard scatter jets for the JVT working points.

Working Points	JVT Cut	Hard scatter jets average efficiency
Loose	> 0.11	97%
Medium	> 0.59	92%
Tight	> 0.91	85%

A limitation of the JVT technique is that it can only be used for jets within

the coverage of the tracking detector, while , jets are reconstructed in the ATLAS detector in the range of $|\eta| < 4.5$. A novel technique, forward jet-vertex-tagger (FJVT) [85], is developed to allow identification and rejection of pile-up jets in the range of $2.5 < |\eta| < 4.5$ by exploiting the correlation between central and forward jets originating from pileup interactions. Two working points have been derived for jets with $2.5 < |\eta| < 4.5$ and $20\,\text{GeV} < p_T < 50\,\text{GeV}$, the cut values, average efficiencies for hard scatter and pile-up jets are shown in Table 4.5

Table 4.5: Summary of the cut values and average efficiencies for hard scatter and pile-up jets for the FJVT working points.

Working Points	FJVT Cut	Hard scatter jets average efficiency	Pile-up jets average efficiency
Loose	< 0.5	92%	60%
Tight	< 0.4	85%	50%

4.6 B-jet Tagging

The identification of jets containing b hadrons, commonly referred as b-tagging, is an important tool used in a number of physics analyses, such as Higgs boson studies, top quark sector precision measurements and new physics searches. The main property used for b-tagging identification algorithms is the relatively longer lifetime ($\sim$ 1.5 ps) of hadrons containing a b-quark compared to other hadrons. A b-hadron with $p_T = 50\,\text{GeV}$ will have an average flight length of $\sim$ 3 mm before its decay, and therefore making at least one vertex displaced from the point where the hard-scatter collision occurred. There are also some other discriminating properties can be used for the b-tagging algorithms, such as the large mass of b-hadrons, large decay multiplicity of b-hadrons and the large momentum fraction carried by b-hadrons. A Schematic view of a b-hadron decay inside a jet is shown in Figure 4.1.

Three baseline algorithms are developed in ATLAS based on the above properties : the impact parameter based algorithms (IP2D, IP3D), inclusive secondary vertex reconstruction algorithms (SV), and the decay chain reconstruction algorithms (JetFitter). Each of the baseline algorithms provides some capacities to separate b-jets from c and light jets. To further reject the c and light jets, a multivariate algorithm [86], MV2c, constructed from combing the outputs of each of the these baseline algorithms, is developed. A boosted decision tree (BDT) is trained

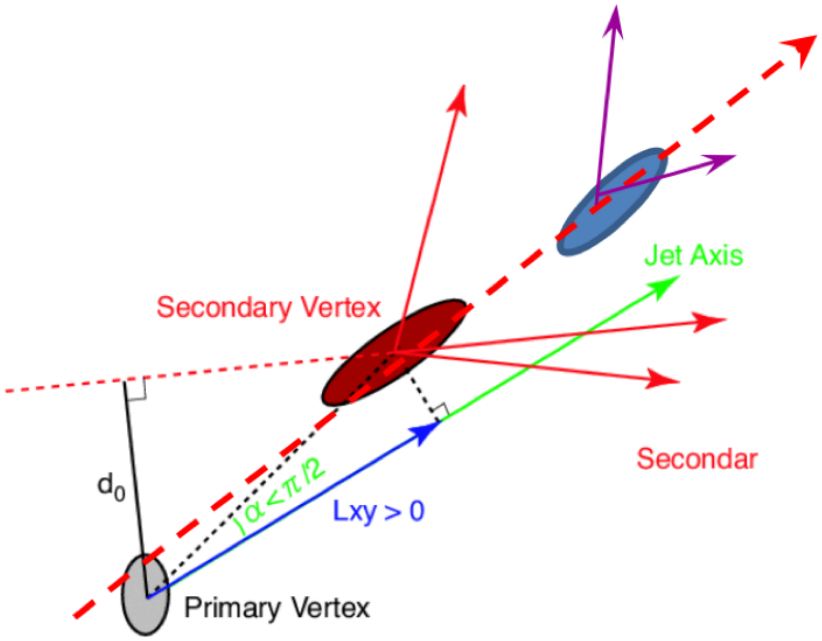
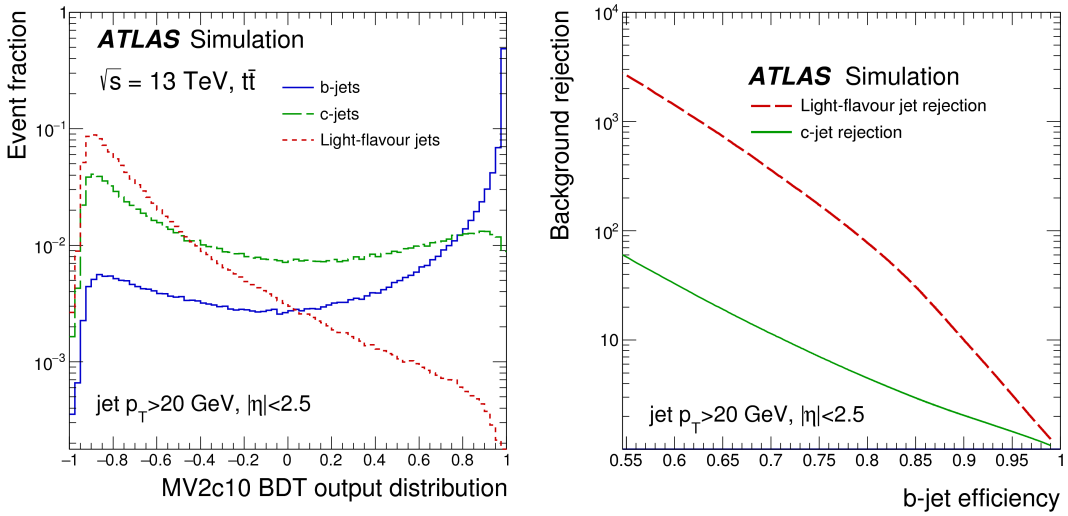


Figure 4.1: Schematic view of a b-hadron decay inside a jet.

with $t\bar{t}$ events, with considering b-jets as signal and c- and light-jets as background. The ratio of c-jets to the total backgrounds can be optimized to improve the c-jets rejection. For the analysis presented in this thesis, the MV2c10 algorithm is used, which means the training sample contains 10% c-jet background and 90% light-jets background. The MV2c10 output for b-jets, c-jets and light-jets in a $t\bar{t}$ sample is shown in Figure 4.2(a), along with the corresponding light-jet and c-jet rejection factors as a function of the b-jet tagging efficiency shown in Figure 4.2(b). The rejection factors for light-jets and c-jets are defined as the inverse of the efficiency for tagging a light-jet or a c-jet as a b-jet, respectively. Four working points for MV2c10 are provided with different b-jet efficiency, and summarized in Tabel 4.6. 70% working point is used for the analysis presented in this thesis, with a rejection rate of 12.2 and 383.3 for c-jet and light-jets, respectively.

Table 4.6: Working point definitions for the 2016 configuration of the MV2c10 b-tagging algorithm, as measured in a simulated $t\bar{t}$ sample at $\sqrt{s} = 13\,\text{TeV}$.

Working Points / b-jet Efficiency	MV2c10 Cut	c-jets Rejection	light-jets Rejection
60%	0.94	34.5	1538.8
70%	0.82	12.2	381.3
77%	0.65	6.2	134.3
85%	0.18	3.1	33.5



(a) (b)

Figure 4.2: (a) MV2c10 BDT output for b- (solid blue), c- (dashed green) and light-(dotted red) jets. (b) The light-jet (dashed line) and c-jet rejection factors (solid line) as a function of the b-jet tagging efficiency of the MV2c10 b-tagging algorithm.

Due to the imperfect physics and detector modelling in simulation, a scaling factor is needed to applied to the MC samples, to correct the efficiencies measured in MC to the efficiencies measured in data. The efficiency in data for each b-tagging working point as shown in Table 4.6, is evaluated for each b, c and light jet flavor. The b-jet efficiency in data is extracted by two methods (tag-and-probe method and a combinatorial likelihood approach) using the high-purity sample of dileptonic $t\bar{t}$ events. The c-jet efficiency is conducted using semi-leptonic $t\bar{t}$ events where the hadronically decaying W boson has approximately 34% probability to produces a c-jet. A negative tag method is used to drive the light-jet efficiencies in data. These calibrations are derived as a function of jet p_T (and $|\eta|$ for light-jet), along with associated uncertainties considered in the analysis presented in this thesis.

4.7 Missing Transverse Energy

The vectorial sum of the transverse momentum of the collision products should be zero due to the conservation of momentum. The imbalance for the momentum in the transverse plane is known as miss transverse momentum [87] ($\boldsymbol{E}_T^{miss}$), and can arise from the stable particles in the final state, like neutrinos, or some other

such particles in theories beyond the SM, therefore the $\boldsymbol{E}_T^{miss}$ is an important variable in searches for exotic signatures. Fake $\boldsymbol{E}_T^{miss}$ can also arise from the SM particles which were mis-measured or unreconstructed due to the effect from the detector acceptance, thus $\boldsymbol{E}_T^{miss}$ is also an important variable of the overall event reconstruction performance.

The reconstructed $\boldsymbol{E}_T^{miss}$ in ATLAS can be characterized by two contributions. The first one is the hard-event signals which comprise the fully reconstructed and calibrated particles and jets, and can be referred to as hard objects. The reconstructed particles included electrons, photons, τ-leptons, and muons. The second one is the soft-event signals which comprise the reconstructed charged particle tracks (soft signals) associated with the hard scatter vertex but not with the hard objects, and can be referred to as soft signals.

The missing transverse momentum components $E_{x(y)}^{miss}$ are constructed from the components $p_{x(y)}$ of the transverse momentum vectors $\boldsymbol{P}_T$, and can be expressed as:

$$E_{x(y)}^{miss} = -\sum_{i\in(hard\,objects)} p_{x(y),i} \quad - \sum_{j\in(soft\,signals)} p_{x(y),j}. \tag{4.7}$$

The vector $\boldsymbol{E}_T^{miss}$ then can be expressed as:

$$\boldsymbol{E}_T^{miss} = (E_x^{miss}, E_y^{miss}), \tag{4.8}$$

and its magnitude E_T^{miss} is calculated as:

$$E_T^{miss} = |\boldsymbol{E}_T^{miss}| = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2}, \tag{4.9}$$

and its direction in the transverse plane can be given by the azimuthal angle θ^{miss}:

$$\theta^{miss} = tan^{-1}(E_y^{miss}/E_x^{miss}). \tag{4.10}$$

The dedicated reconstruction procedure for each kind of particles and jets have been discussed in the previous sections in this chapter. These procedures are actually independent of each other and can result a consequence that the same calorimeter signal used to reconstruct one object is also likely used to reconstruct another object, therefore introducing potentially double counting of the same signal during the reconstruction. In order to address this issue, the signal ambiguity resolution is adopted by requiring the explicit order for the E_T^{miss} reconstruction sequence for the hard objects contribution. For the analysis presented in this thesis, the order starts with electrons, followed by photons, hadronically decaying

τ-leptons, and finally jets. Muons yields basically no signal overlap with other reconstructed particles in the calorimeter thanks to the fact that muons are mainly reconstructed from ID and MS tracks, with corrections already applied based on their energy loss in the calorimeter.

Apart from E_T^{miss}, a track-based missing transverse momentum vector $E_{T,trk}^{miss}$ is constructed from the negative vector sum of the transverse momenta of all the reconstructed tracks that associated with the primary vertex. This quantity is very useful to suppress the multijet and non-collision backgrounds.

Chapter 5

Search for the Standard Model $VH(b\bar{b})$

5.1 Overview

The Higgs boson was discovered in 2012 by the ATLAS and CMS Collaborations [5, 6] from the analysis of proton-proton (pp) collisions produced by the LHC. After that, with the full Run 1 data collected at centre-of-mass energies of 7 TeV and 8 TeV, the properties of the discovered particle have been measured and were found to be compatible with those predicted by the SM within uncertainties [26, 88–90]. The observation of many of the Higgs production modes and decay channels predicted by the SM have been established, the bosonic decay channels have entered an era of precision measurements [9–14], the Higgs boson mass was measured by ATLAS as $mH = 124.98 \pm 0.28\,\text{GeV}$ from the combination of $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ analyses with Run 2 2015-2016 data [91]. The τ-lepton pairs decay was first observed in the combination of the ATLAS and CMS analyses [15]. The ggF and VBF production modes were observed following the analysis of Run 1 data. Recently, the ttH production mode was also observed by both ATLAS and CMS Collaborations [16, 17], and provided the directly observation of the coupling of the Higgs boson to top quarks.

The largest decay mode of the SM Higgs boson is Higgs decays into pairs of b-quarks, with a predicted branching fraction of 58% for $mH = 125\,\text{GeV}$ [18]. Probing $H \to b\bar{b}$ decay is very important to constrain the overall Higgs boson decay width [20, 21]. Despite the ggF production mode has the largest cross section at LHC, the overwhelm multijet backgrounds make the search in this production mode very challenging. The most sensitive production modes for probing $H \to b\bar{b}$

decays are the associated production of a Higgs and a W or Z boson [19](denoted as V), the leptonic decay modes of the vector boson lead to clean signatures that can be efficiently triggered on, while rejecting most of the multi-jet background events.

In 2012, the CDF and D0 Collaborations at Tevatron reported an excess of events in VH associated production in the mass range of 120 GeV to 135 GeV, with a global significance of 3.1 standard deviations, and a local significance of 2.8 standard deviations at a mass of 125 GeV [23]. With Run 1 data, ATLAS and CMS reported an excess of events in VH associated production, with a local significance of 1.4 and 2.1 standard deviations at a mass of 125 GeV [24, 25], respectively. The combination of these two analyses resulted in observed and expected significances of 2.6 and 3.7 standard deviations[26]. $H \to b\bar{b}$ decay searches have been also performed in the VBF [27–29] and ttH [32–34] production modes, but with significantly lower sensitivities.

This chapter mainly reports on the search for the SM Higgs boson in the VH production mode with Higgs decaying into a $b\bar{b}$ pair with the ATLAS detector in Run 2 of the LHC, using an integrated luminosity of 79.8 fb^{-1} with 2015, 2016 and 2017 data. Three lepton channels are considered based on the number of charged leptons, l (electrons or muons), referred to as 0-, 1-, 2- lepton channels, to explore signatures of $ZH \to \nu\nu b\bar{b}$, $WH \to l\nu b\bar{b}$ and $ZH \to llb\bar{b}$, respectively. Feynman diagrams for quark induced and gluon induced $VH(b\bar{b})$ productions are presented in Figure 5.1 and Figure 5.2, respectively. Whilst the VH production mode can help to reduce a lot of multijet background, there are still a number of background processes remaining in this search channel, and have much larger yields than signal events. The main backgrounds are $t\bar{t}$ (for all three lepton channels), W+jets (for 0- and 1- lepton channels), Z+jets (for 0- and 2- lepton channels), and single top-quark (for 1-lepton channel). To maximize the sensitivity to the Higgs boson signal, a boosted decision tree (BDT) is trained to separate signal events from backgrounds. The BDT output discriminant is built from variables that describe the kinematics of the selected events, and used as the main fit observable in a binned maximum-likelihood fit, referred to as global likelihood fit. The likelihood fit is performed to data simultaneously across the three channels in multiple analysis regions, in order to extract the signal yield and the main background normalizations. Two other analyses are used to validate this signal extraction method : the dijet-mass analysis, where the signal yield is extracted from a fit to the mass of the dijet system, and the diboson analysis, where the nominal multivariate analysis is modified to extract the $VZ, Z \to b\bar{b}$ diboson pro-

cess. The result of main multivariate analysis is also combined with the Run 1 $VH(b\bar{b})$ result [24], and also with the other searches for $H \to b\bar{b}$ decay and with the other searches in the VH production mode.

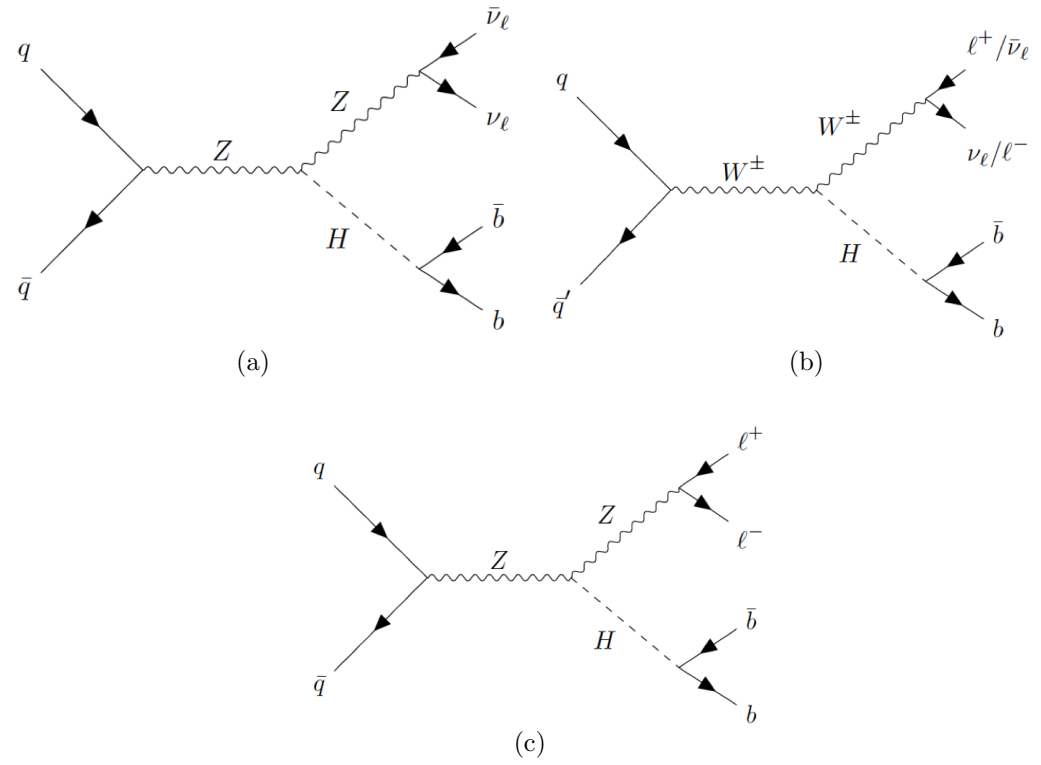


Figure 5.1: Feynman diagrams for the leading-order quark initiated SM $VH(b\bar{b})$ process in the 0-lepton (a), 1-lepton (b) and 2-lepton (c) channels.

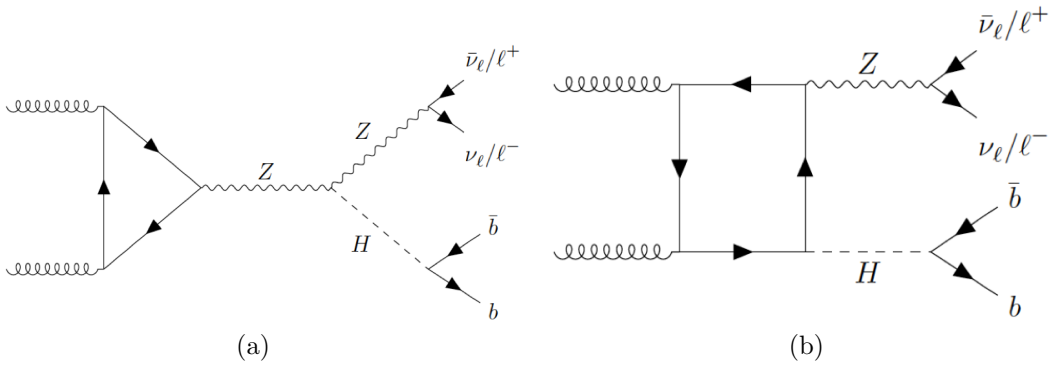


Figure 5.2: Feynman diagrams for the leading-order gluon initiated SM $VH(b\bar{b})$ process in the 0- and 2-lepton channels.

In this chapter, Section 5.2 presents the data and MC simulation samples used

for this analysis. Section 5.3 presents the object and event selections. The multivariate analysis is discussed in Section 5.4, including the details for the training and performance of the multivariate discriminant. Section 5.5 presents the estimation of multijet background. Some results from the efforts to further optimize the sensitivity of the 1-lepton analysis are presented in Section 5.6. Systematic uncertainties considered in this analysis are discussed in Section 5.7, an overview of the statistical analysis is summarized in Section 5.8. The results are presented in Section 5.9, including those from the cross-checks of the analysis and the combinations. Lastly, further improvements and prospects for the analysis beyond the current iteration are discussed in Section 5.10.

5.2 Dataset and Simulated Event Samples

The proton-proton collision data used in this analysis was collected by the ATLAS detector during the 2015, 2016 and 2017 running periods of the LHC. Events are selected only if they pass a filter requirement given by Good Run List (GRL), to ensure their quality and that all systems of the ATLAS detector were operating well when events were recorded. The selected events corresponds to a total integrated luminosity of 79.8 ± 1.6 fb^{-1} [92].

Monte Carlo samples are used to simulate the signal and most background processes, apart from the multijet contributions, which use a data-driven method as discussed in Section 5.5. All simulated processes are normalised using the most accurate theoretical cross-section predictions currently available and were generated at least to next-to-leading-order (NLO) accuracy. All samples of simulated events were passed through the ATLAS detector simulation [71] based on GEANT 4 [72] and were reconstructed with the standard ATLAS reconstruction software. The effects of multiple interactions in the same and nearby bunch crossings (pile-up) were modelled by overlaying minimum-bias events, simulated using the soft QCD processes of Pythia 8.186 [93] with the A2 [94] set of tuned parameters (tune) and MSTW2008LO [95] parton distribution functions (PDF). The EvtGen v1.2.0 program [96] was used to describe the decays of bottom and charm hadrons for all samples of simulated events, apart from those generated by Sherpa [97]. All the generators used for the simulation of the signal and background processes are summarized in Table 5.1, and shortly described as follows.

Simulated $VH \rightarrow Vb\bar{b}$ quark induced signal samples were generated using Powheg MiNLO + Pythia 8 applying the AZNLO tune with NNPDF3 par-

Table 5.1: The generators used for the simulation of the signal and background processes. If not specified, the order of the cross-section calculation refers to the expansion in the strong coupling constant (αs). The acronyms ME, PS and UE stand for matrix element, parton shower and underlying event, respectively. ($\star$) The events were generated using the first PDF in the NNPDF3.0NLO set and subsequently reweighted to the PDF4LHC15NLO set [98] using the internal algorithm in Powheg-Box v2. (†) The NNLO(QCD)+NLO(EW) cross-section calculation for the $pp \to ZH$ process already includes the $gg \to ZH$ contribution. The $qq \to ZH$ process is normalised using the cross-section for the $pp \to ZH$ process, after subtracting the $gg \to ZH$ contribution. An additional scale factor is applied to the $qq \to VH$ processes as a function of the transverse momentum of the vector boson, to account for electroweak (EW) corrections at NLO. This makes use of the VH differential cross-section computed with Hawk [48, 49].

Process	ME generator	ME PDF	PS and Hadronisation	UE model tune	Cross-section order
Signal, mass set to 125 GeV and $b\bar{b}$ branching fraction to 58%					
$qq \to WH$	Powheg-Box v2 [99] +	NNPDF3.0NLO$^{(\star)}$ [100]	Pythia 8.212 [93]	AZNLO [101]	NNLO(QCD)+
$\to \ell\nu b\bar{b}$	GoSam [102] + MiNLO [103, 104]				NLO(EW) [105–111]
$qq \to ZH$	Powheg-Box v2 +	NNPDF3.0NLO$^{(\star)}$	Pythia 8.212	AZNLO	NNLO(QCD)$^{(\dagger)}$+
$\to \nu\nu b\bar{b}/\ell\ell b\bar{b}$	GoSam + MiNLO				NLO(EW)
$gg \to ZH$	Powheg-Box v2	NNPDF3.0NLO$^{(\star)}$	Pythia 8.212	AZNLO	NLO+
$\to \nu\nu b\bar{b}/\ell\ell b\bar{b}$					NLL [47, 112–115]
Top quark, mass set to 172.5 GeV					
$t\bar{t}$	Powheg-Box v2 [116]	NNPDF3.0NLO	Pythia 8.230	A14 [117]	NNLO+NNLL [118]
s-channel	Powheg-Box v2 [119]	NNPDF3.0NLO	Pythia 8.230	A14	NLO [120]
t-channel	Powheg-Box v2 [119]	NNPDF3.0NLO	Pythia 8.230	A14	NLO [121]
Wt	Powheg-Box v2 [122]	NNPDF3.0NLO	Pythia 8.230	A14	Approximate NNLO [123]
Vector boson + jets					
$W \to \ell\nu$	Sherpa 2.2.1 [97, 124, 125]	NNPDF3.0NNLO	Sherpa 2.2.1 [126, 127]	Default	NNLO [128]
$Z/\gamma^* \to \ell\ell$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NNLO
$Z \to \nu\nu$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NNLO
Diboson					
$qq \to WW$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NLO
$qq \to WZ$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NLO
$qq \to ZZ$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NLO
$gg \to ZZ$	Sherpa 2.2.2	NNPDF3.0NNLO	Sherpa 2.2.2	Default	NLO

ton distribution functions. Gluon induced signal samples were simulated using POWHEG matrix element generator interfaced with PYTHIA 8 applying AZNLO tune with NNPDF3 PDFs. The SM Higgs boson mass is fixed to 125 GeV, the $b\bar{b}$ branching fraction is fixed to 58%. WH signal samples are normalised to the production cross section at next-to-next-to-leading order NNLO (QCD) and NLO (EW). The inclusive cross section of ZH production is calculated at NNLO (QCD) and NLO (EW), the cross section of gluon induced ZH production is then calculated at NLO (QCD), and quark induced production is taken as the difference of the two in order to avoid double counting.

Events containing W or Z bosons with jets (V+jets) are simulated with SHERPA 2.2.1 using the NNPDF3.0 NNLO PDFs with dedicated parton shower tuning developed by the SHERPA authors. Matrix elements were calculated for up to two partons at NLO and four partons at LO using the OPENLOOPS and COMIX matrix-element generators. The number of expected V + jets events is rescaled using the NNLO cross-sections. In order to generate sufficient high statistics, V+jets samples are sliced in the maximum of H_T and p_T^V at the parton level where the former is given by the scalar p_T sum of all parton-level jets with $p_T > 20$ GeV. Additionally, to obtain sufficient heavy-flavour final state statistics, the V+jets samples are generated by applying filters as summarised in Table 5.2. Samples are normalised using cross sections calculated at NNLO accuracy. The W + jets and Z + jets simulated background samples are decomposed according to the true flavour of the dijet pair used to reconstruct the Higgs candidate, leading to the following twelve sub-samples:

- Zbb and Wbb: the two jets are labelled as b-jet;
- Zcc and Wcc: the two jets are labelled as c-jet;
- Zl and Wl: the two jets are labelled as light-jet;
- Zbc and Wbc: one of the two jets is labelled as b-jet and the others as c-jet;
- Zbl and Wbl: one of the two jets is labelled as b-jet and the others as light-jet;
- Zcl and Wcl: one of the two jets is labelled as c-jet and the others as light-jet.

The scheme used to define the jet flavour, is based on a ΔR match between truth level hadrons and reconstructed jets. Final state hadrons with $p_T > 5$ GeV

Table 5.2: Heavy flavour filters used for V + jets, along with a simple description of their application.

Filter	Description
BFilter	at least 1 b-hadron with p_{T} >0 GeV and $\|\eta\|$ <4
CFilterBVeto	at least 1 c-hadron with p_{T} >4 GeV and $\|\eta\|$ <3 veto events which pass the BFilter
CVetoBVeto	veto events which pass the BFilter or the CFilterBVeto

and within $\Delta R < 0.3$ of the jet axis are assigned to each jet, each hadron is matched to only one jet, selecting the closest jet in ΔR space. If a truth b-hadron is matched to the jet, the jet is labelled as a b-jet, else if a truth c-hadron is matched to the jet, the jet is then labelled as a c-jet, otherwise the jet is labelled as a light-jet. A $V + HF$ category is defined as containing $V + bb$, $V + bc$, $V + bl$ and $V + cc$ events.

Top-quark pair production ($t\bar{t}$) is simulated using Powheg within the Powheg-BOX framework using NNPDF 3.0 PDFs and interfaced with Pythia 8 using NNPDF 2.3 PDFs for parton showering, with the A14 tune. The top quark mass was set to 172.5 GeV. The $t\bar{t}$ samples used in 0- and 1- lepton channels are generated with a filter to require that at least one of the W bosons decays leptonically (non-all-had), whilst the $t\bar{t}$ samples used for 2- and channels are generated with a filter to require that both of the W bosons decays leptonically (dilepton). Furthermore, for the $t\bar{t}$ samples used for 0-lepton channel, on top of the non-all-had filter, a number of E_T^{miss} filters are applied to increase the number of simulated events. All the samples are normalised using cross sections calculated at NNLO + next-to-next-to-leading-logarithm accuracy (NNLL).

Single top quark production (t, s and Wt channels) is simulated using Powheg with NNPDF 3.0 PDFs interfaced with Pythia 8 using NNPDF 2.3 PDFs for parton showering. Samples are normalised using cross sections calculated at NLO.

Semi-leptonic diboson samples are generated using Sherpa 2.2.1 interfaced with NNPDF 3.0 NNLO PDFs in a factorised approach where the boson pairs enter the matrix elements with zero-width and are produced on shell. The samples are produced up to NLO accuracy for $VV + 0j$ and $VV + 1j$ final states and are combined with multi-leg LO matrix elements for $VV + 2, 3j$ final states. In order to provide increased statistics for the diboson samples, in particular for the training of the Boosted Decision Tree in which the Standard Model $VZ \to b\bar{b}$ process is used as signal, additional samples are produced where one of the Z bosons is

forced to decay to $Z \to b\bar{b}$. The $Z \to b\bar{b}$ samples are combined with the inclusive $VZ \to q\bar{q}$ samples using appropriate event weights for the overlapping $Z \to b\bar{b}$ events.

In addition to the quark induced diboson samples, semi-leptonic loop-induced $gg \to VV$ samples are generated using SHERPA 2.2.2 interfaced with NNPDF 3.0 NNLO PDFs. The samples use LO accurate matrix elements for the $VV + 0j$ and $VV + 1j$ final states.

Two set of statistically independent MC samples are generated to reflect the different data running conditions and total integrated luminosity between 2015-2016 and 2017 data but with same generator settings. The MC events which are simulated and reconstructed using the 2015-2016 data running conditions are referred to as the mc16a MC samples, whilst the MC events for 2017 data are referred to as the mc16d MC samples. The number of events simulated in mc16d is approximately 1.2 times larger than the events simulated in mc16a to account for the larger integrated luminosity collected in 2017 compared with 2015 and 2016.

5.3 Object and Event Selection

5.3.1 Overlap removal procedure

The reconstruction and identification algorithms for the objects used for this analysis have already been discussed in Chapter 4, but such algorithms do not always result in unambiguous identifications, in order to remove the potential double counting of the objects used in this analysis, a procedure called as "overlap removal" is applied to the fully reconstructed and calibrated objects in the following steps:

- tau-electron: if $\Delta R(\tau_{had}, e) < 0.2$, the τ_{had} lepton is removed.
- tau-muon: if $\Delta R(\tau_{had}, \mu) < 0.2$, the τ_{had} lepton is removed, with the exception that if the τ_{had} lepton has $p_T > 50\,\text{GeV}$ and the muon is deemed to be of low quality, then the τ_{had} lepton is not removed.
- electron-muon: if a reconstructed muon shares an electron's ID track, the electron is removed.
- electron-jet: if $\Delta R(\text{jet}, e) < 0.2$, the jet is removed, since a jet is always expected from clustering an electron's energy deposits in the calorimeter. For

any surviving jets, if ΔR(jet, e) ¡ min(0.4,0.04 + 10 GeV/p_T^e), the electron is removed, such electrons are likely to originate from semileptonic b- or c-hadron decays.

- muon-jet: if ΔR(jet, μ) < 0.2 and the jet has fewer than three associated tracks or the muon energy constitutes most of the jet energy, the jet is removed. For any surviving jets, if ΔR(jet, μ) ¡ min(0.4,0.04 + 10 GeV/p_T^μ), the muon is removed.

- tau-jet: if ΔR(jet, τ_{had}) < 0.2, the jet is removed.

5.3.2 Analysis specific object definition

Considering the analysis specific requirements for the electrons, muons and jets, different categories are define for these objects.

For electrons, three categories, referred to as VH-loose, ZH-Signal and WH-Signal, are defined in the analysis. VH-loose electron criteria is defined to allow for the maximum electron selection efficiency for signal processes. Electron p_T is required to be greater than 7 GeV. The electron should be in the range of $|\eta| < 2.47$. Loose likelihood identification is applied in VH-loose criteria. LooseTrackOnly isolation is applied to reduce the non-prompt electrons. The isolation selection is chosen to keep 99% efficiency for real electrons. ZH-signal electron criteria requires a electron object with $p_T > 27\,\text{GeV}$ in addition to VH-loose electron criteria for the 2-lepton channel. In the 1-lepton analysis, tighter lepton selection is required to suppress multi-jet background, therefore tight likelihood identification and FixedCutHighPtCaloOnly isolation selection in addition to LooseTrackOnly requirement are required to define the WH-signal electron, this isolation requirement is optimized in dedicated $VH(b\bar{b})$ phase space with more details given in Section 5.5. The definitions of the requirements for each category are summarised in Tabel 5.3.

Table 5.3: Summary of electron selection requirements.

Electron Selection	p_T	Identification Quality	Isolation
Loose	$> 7\,\text{GeV}$	Loose	LooseTrackOnly
ZH-Signal	$> 27\,\text{GeV}$	Loose	LooseTrackOnly
WH-Signal	$> 27\,\text{GeV}$	Tight	LooseTrackOnly & FixedCutHighPtCaloOnly

Similar with electrons, three categories are defined for muons, and referred to as VH-Loose, ZH-Signal and WH-Signal. VH-loose muon criteria is defined

to keep muon from signal as much as possible. In VH-loose criteria muon is required with $p_T > 7\,\text{GeV}$, and pass Loose muon quality. LooseTrackOnly isolation is applied to reduce the non-prompt muons. The isolation selection is chosen to keep 99% efficiency for the signal muons. ZH-signal muon criteria requires muon object with $p_{\text{T}} > 27$ GeV and $|\eta| < 2.5$ in addition to the VH-loose muon criteria for the 2-lepton channel. In the 1-lepton analysis, tighter lepton selection is required to suppress multi-jet background. Therefore medium muon quality and FixedCutHighPtTrackOnly isolation selection in addition to LooseTrackOnly requirement are required to define the WH-signal muon, this isolation requirement is also optimized in dedicated $VH(b\bar{b})$ phase space with more details given in Section 5.5. The definitions of the requirements for each category are summarised inTabel 5.4.

Table 5.4: Summary of muon selection requirements.

Muon Selection	p_T	Identification Quality	Isolation
Loose	$> 7\,\text{GeV}$	Loose	LooseTrackOnly
ZH-Signal	$> 27\,\text{GeV}$	Loose	LooseTrackOnly
WH-Signal	$> 25\,\text{GeV}$	Medium	LooseTrackOnly & FixedCutHighPtTrackOnly

Jets used in this analysis are classified as either "signal jets" or "forward jets". Signal jets are eligible for b-tagging and used in reconstructing the Higgs boson. Signal jets are defined with the requirements of $|\eta| < 2.5$ and $p_T > 20\,\text{GeV}$, for jets with $|\eta| < 2.4$ and $p_T < 60\,\text{GeV}$, a requirement on JVT is also applied. Forward jets are defined with the requirements of $2.5 < |\eta| < 4.5$ and $p_T > 30\,\text{GeV}$. The full set of selection requirements are given in Table 5.5.

Table 5.5: Summary of jets selection requirements.

Jet Category	Selection Requirements
Forward jets	$p_T > 30\,\text{GeV}$ & $2.5 < \|\eta\| < 4.5$
Signal jets	$p_T > 20\,\text{GeV}$ & $\|\eta\| < 2.5$
	JVT ¿ 0.59 for jets with $p_T < 60\,\text{GeV}$ and $\|\eta\| < 2.4$

5.3.3 Event selections

As already discussed, data events used in this analysis are required to pass the GRL selections, to ensure their quality and that all systems of the ATLAS detector were operating well when events were recorded. Then, for both data

and MC simulation, events are categorized into three sub-channels, referred to as 0-, 1- and 2-lepton channel, by requiring exactly 0 VH-loose lepton, exactly 1 WH-signal lepton and exactly 2 VH-loose leptons with at least one ZH-signal lepton, respectively. In all three lepton channels, events are required to contain at least two signal jets. Exclusive categories of events, depending on the number of selected jets they contain, are defined in order to maximize the signal significance: events containing two jets comprise the 2-jet category, events with exactly three jets form the 3-jet category and events with three or more jets form the 3+-jet category. In the 0- and 1-lepton channels, the 2- and 3-jet categories are used, and events with four or more jets are rejected due to the high $t\bar{t}$ background contamination. A dedicated study for the potential sensitivity increase with using the 3+-jet category in 1 lepton channel by introducing a new specific cut was performed with more details shown in Section 5.6. In the 2-lepton channel, where the high jet multiplicity regions result in some additional sensitivity, the 2-jet and 3+-jet categories are used. In all three lepton channels, b-tagging is applied to all signal jets selected using the MV2c10 algorithm at the 70% efficiency working point. The b-tagging strategy, and efficiency working point have been optimized to maximize the expected signal significance. Events are categorized according to the number of b-tagged signal jets and only the 2-tag region is considered in this analysis, as this is the region that has the largest signal sensitivity. The leading b-tagged jet in the 2-tag category is required to have $p_T > 45\,\text{GeV}$.

5.3.3.1 0-lepton channel specific selection

Data events are recorded using lowest unprecaled E_T^{miss} triggers with online thresholds of 70 GeV for the data recorded in 2015, of 90 and 110 GeV for the data recorded in 2016 and of 110 GeV for the data recorded in 2017, depending on the data-taking period and the different trigger rates. Their efficiency was measured in W+jets, Z+jets and $t\bar{t}$ events in data using single-muon triggers, resulting in correction factors that are applied to the simulated events, ranging from 1.05 at the offline E_T^{miss} threshold of 150 GeV to a negligible deviation from unity at E_T^{miss} above 200 GeV. Tabel 5.6 shows the detailes for these E_T^{miss} triggers.

the reconstructed transverse momentum of the Z boson, p_T^Z, corresponds to E_T^{miss} in the 0-lepton channel, is requried to be greater than 150 GeV, due to the slow turn-on curve of the E_T^{miss} trigger. Further requirements are applied on the scalar sum of the p_T of the jets in the event (H_T), to remove a region which is mis-modelled in simulation due to a non-trivial dependence of the trigger

Table 5.6: MET triggers used during the 2015, 2016 and 2017 data collection period. The notation, (A, D3, D4,...) refer to the ATLAS collection periods in the year of 2016.

Trigger Name	Period	Threshold (GeV)	Description
HLT_xe70_mht_L1XE50	2015	70 GeV	Seeded using the level L1_XE50 LAr and Tile calorimeter triggers, calibrated at the EM scale, with a threshold of 50 GeV.
HLT_xe90_mht_L1XE50	2016 (A-D3)	90 GeV	
HLT_xe110_mht_L1XE50	2016 ($\geq$ D4)	110 GeV	
HLT_xe110_pufit_L1XE50	2017	110 GeV	

efficiency on the jet multiplicity. $H_T > 120\,\text{GeV}$ is applied to the 2-jets events, and $H_T > 150\,\text{GeV}$ is applied to the 3-jets events. The multijet background in 0-lepton channel is mainly due to the jet energy mis-measurements in the calorimeters, as a result, the fake missing transverse energy and momentum tend to be aligned with the mis-measured jet. In order to reduce the multijet background, four angular selection criteria (referred to as anti-QCD cuts) are required:

- $\Delta\phi(E_T^{miss}, E_{T,trk}^{miss}) < 90°$,
- $\Delta\phi(b_1, b_2) < 140°$,
- $\Delta\phi(E_T^{miss}, bb) > 120°$,
- $min[\Delta\phi(E_T^{miss}, jets)] > 20°$ for 2 jets, $> 30°$ for 3 jets.

Here ϕ is the azimuthal angle, $E_{T,trk}^{miss}$ is defined as negative vector the sum of the transverse momenta of the tracks reconstructed in the inner detector and associated to the primary vertex of the event. b_1 and b_2 are the two b-tagged jets forming the Higgs boson candidate's dijet system. The last selection is a requirement on the azimuthal angle between the E_T^{miss} vector and the closest jet. Thanks the anti-QCD cuts, the remaining multijet background in 0-lepton channel is found to be negligible with more details given in Section 5.5.

5.3.3.2 1-lepton channel specific selection

The transverse momentum of the W boson, p_T^W, is reconstructed as vectorial sum of E_T^{miss} and the charged lepton's transverse momentum and required to be greater than 150 GeV in 1-lepton channel, due to the much increased sensitivity and the reduced multijet background contamination in such high p_T^W region compare to the relative low p_T^W region. Despite not used in this iteration of the

analysis, an effort to include the $75\,\text{GeV} < p_T^W < 150\,\text{GeV}$ region (referred to as medium p_T^W region) in the 1-lepton channel has been studied. For this study, the details about the multi-jet reduction and estimation in medium p_T^W region can be found in Section 5.5, the details about the sensitivity increase in the global fit after adding the medium p_T^W region in the analysis is given in Section 5.10, after introducing the default results without the medium p_T^W region. For muon sub-channel, events are recorded using the same E_T^{miss} trigger as those used in the 0-lepton channel. The E_T^{miss} calculation at trigger level is relied on the calorimeter information, therefore muons are not included for this calculation. In events where a muon is present, the E_T^{miss} trigger is actually selecting events based on p_T^W, and is fully efficient for events with $p_T^W > 180\,\text{GeV}$. The overall signal efficiency for E_T^{miss} trigger in muon sub-channel is $\sim 98\%$, compared to $\sim 80\%$ efficiency for the combination of single-muon triggers, therefore E_T^{miss} trigger is used. A study about using the combination of E_T^{miss} trigger and signle-muon triggers has been performed. Only $\sim 2\%$ more signal events can be recovered by using the combination triggers, in that case, to simplify the analysis, only E_T^{miss} trigger is used in the muon sub-channel. For electron sub-channel, events are recorded using the lowest unprescaled single electron triggers in each data collection period and p_T thresholds started at 24 GeV in 2015 and increased to 26 GeV in 2016 and 2017. The lowest-threshold trigger in 2016 and 2017 includes isolation and identification requirements which are looser than any of the isolation and identification requirements applied in the offline analysis. These requirements are relaxed or removed for the higher-threshold triggers. Table 5.7 shows the details for these single electron triggers. In the electron sub-channel, an additional selection of $E_T^{miss} > 30\,\text{GeV}$ is applied to further reduce the multijet background. Events are categorised into the signal region (SR) or into a $W + HF$ events enriched control region ($W + HF$ CR), based on the selections on the invariant mass of the two b-tagged jets (m_{bb}), and on the reconstructed mass of a semi-leptonically decaying top-quark candidate (m_{top}). The $W + HF$ CR is obtained by applying two additional selection requirements: $m_{bb} < 75\,\text{GeV}$ and $m_{top} > 225\,\text{GeV}$, with more details given in Section 5.3.7.

5.3.3.3 2-lepton channel specific selection

The transverse momentum of the Z boson, p_T^Z, is reconstructed as vectorial sum of transverse momentum of two leptons, with a $p_T^Z > 75\,\text{GeV}$ cut applied due to low signal sensitivity in the lower p_T^Z regions. The 2-lepton channel is

Table 5.7: Single electron triggers used during the 2015, 2016 and 2017 data collection period.

Trigger Name	Period	Threshold (GeV)	Description
HLT_e24_lhmedium_L1EM20VH	2015	24 GeV	Seeded using L1EM20VH level 1 trigger calibrated at the EM scale with a threshold of 20 GeV, and require medium ID quality.
HLT_e60_lhmedium	2015	60 GeV	Medium ID likelihood required.
HLT_e120_lhloose	2015	120 GeV	Loose ID likelihood required.
HLT_e26_lhtight_nod0_ivarloose	2016 & 2017	26 GeV	Tight likelihood ID required, and variable loose isolation required
HLT_e60_lhmedium(_nod0)	2016 & 2017	60 GeV	Medium ID likelihood required
HLT_e140_lhloose(_nod0)	2016 & 2017	140 GeV	Loose ID likelihood required
HLT_e300_etcut	2017	300 GeV	No ID requirements.

then split into two regions, $75\,\text{GeV} < p_T^Z < 150\,\text{GeV}$ and $p_T^Z > 150\,\text{GeV}$. Events in electron sub-channel are recorded using the same lowest unprescaled single electron triggers as in the 1-lepton channel. For muon sub-channel, events are recorded using the lowest unprescaled single muon triggers in each data collection period and p_T thresholds started at 20 GeV in 2015 and increased to 26 GeV in 2017. Table 5.8 shows the details for these single muon triggers. The invariant mass of the di-lepton system must be consistent with the Z boson mass: $81\,\text{GeV} < m(ll) < 101GeV$, in order to suppresses backgrounds have a non-resonant lepton-pair, such as $t\bar{t}$ and multi-jet productions. For the selected di-muon events the two muons are further required to be of opposite charge; the requirement is not applied to di-electron events due to higher rate of charge misidentification. A top $e\mu$ control region is defined by applying the nominal selection but requiring an $e\mu$ lepton flavour combination instead of ee or $\mu\mu$, and requiring the two leptons to have opposite-sign charges, more details for this control region are given in Section 5.3.7.

Tabel 5.9 summarizes the signal events selection applied in each of the three channels.

5.3.4 Additional selections for dijet mass analysis

A dijet-mass analysis is performed as a cross-check to the main multivariate analysis, where the m_{bb} distribution is used as the main fit observable to extract the signal yields in the global fit. In order to increase the sensitivity for the dijet-mass analysis, a number of additional selection criteria are applied to the events

Table 5.8: Single muon triggers used during the 2015, 2016 and 2017 data collection period.

Trigger Name	Period	Threshold (GeV)	Description
HLT_mu20_iloose_L1MU15	2015	20 GeV	Seeded using L1MU15 level 1 trigger with a threshold of 15 GeV, and requiring loose isolation requirements.
HLT_mu40	2015 & 2016 (A)	40 GeV	No isolation requirements.
HLT_mu50	2015 & 2016 & 2017	50 GeV	No isolation requirements.
HLT_mu24_iloose(_L1MU15)	2016 (A, MC)	24 GeV	Loose isolation requirements
HLT_mu24_ivarmedium	2016 (A-D3)	24 GeV	Variable cone medium isolation requirements
HLT_mu26_ivarmedium	2016 ($\geq$ D4) & 2017	26 GeV	Variable cone medium isolation requirements

to further reduce the background contamination, and summarized in Table 5.10.

Considering the $H \to b\bar{b}$ decay, the relationship between the separation of the two b-quarks in η - ϕ space and the mass and p_T of the Higgs boson can be expressed as:

$$\Delta R(b, \bar{b}) \approx \frac{2m_H}{p_T^H}, \tag{5.1}$$

as also shown in Figure 5.3. With the increased Higgs boson p_T, the $\Delta R(b, \bar{b})$ is reduced. Assuming the Higgs has recoiled from the V boson, the Higgs boson p_T should be close to p_T^V. In that case, at higher p_T^V region, the signal events should have reduced ΔR separation, whilst the background events do not have the same feature. Therefore, to fully use the advantage from high p_T^V regime, the $p_T^V > 150\,\text{GeV}$ region is further separated into two regions : $150\,\text{GeV} < p_T^V < 200\,\text{GeV}$ and $p_T^V > 200\,\text{GeV}$, with different ΔR cut applied in different regions as shown in Table 5.10. The p_T^V separation and ΔR cuts are mainly inherited from previous iteration of the analysis, an effort to re-optimized the separation and cuts in 1-lepton channel has been performed with more details given in Section 5.6.

In 1-lepton channel, an additional cut on W boson's transverse mass $m_T^W < 120\,\text{GeV}$ is applied to further reduce the $t\bar{t}$ backgrounds that undergo dileptonic decays. The W boson's transverse mass, m_T^W, is defined as $m_T^W = \sqrt{2p_T^l E_T^{miss}(1 - cos(\Delta\phi(l, E_T^{miss})))}$, where the p_T^l is the lepton's transverse momentum.

In 2-lepton channel, in order to suppress the $t\bar{t}$ background, an additional cut is applied, with requiring $E_T^{miss}/\sqrt{S_T} < 3.5\sqrt{GeV}$, where S_T is defined as the scalar sum of the transverse momenta of all jets and leptons in the event.

Table 5.9: Summary of the signal event selection in the 0-, 1- and 2-lepton analyses.

Common Selections	
Jets	≥ 2 signal jets
b-jets	2 b-tagged signal jets
Leading jet p_T	$> 45\,\text{GeV}$
0 Lepton	
Trigger	E_T^{miss} as shown in Table 5.6
Jets	Exactly 2 or 3 jets
Leptons	Exactly 0 VH-loose lepton
E_T^{miss}	$> 150\,\text{GeV}$
H_T	$> 120\,\text{GeV}(2jets), > 150\,\text{GeV}(3jets)$
$\Delta\phi(E_T^{miss}, E_{T,trk}^{miss})$	$< 90°$
$\Delta\phi(b_1, b_2)$	$< 140°$
$\Delta\phi(E_T^{miss}, bb)$	$> 120°$
$min[\Delta\phi(E_T^{miss}, jets)$	$> 20°(2jet), > 30°(3jet)$
p_T^V regions	$> 150\,\text{GeV}$
1 Lepton	
Trigger	e channel: un-prescaled single electron as shown in Table 5.7
	μ channel: E_T^{miss} as shown in Table 5.6
Jets	Exactly 2 or 3 jets
Leptons	Exactly 1 WH-signal lepton, no additional VH-loose lepton
E_T^{miss}	$> 30\,\text{GeV}$ for e-sub channel
m_{top} & m_{bb}	$m_{top} < 225\,\text{GeV}$ or $m_{bb} > 75\,\text{GeV}$
p_T^V regions	$> 150\,\text{GeV}$
2 Lepton	
Trigger	e channel: un-prescaled single electron as shown in Table 5.7
	μ channel: un-prescaled single muon as shown in Table 5.8
Jets	Exactly 2 or 3+ jets
Leptons	Exactly 2 VH-loose lepton, at least one ZH-signal lepton
	Same flavor, opposite-charge for $\mu\mu$
m_{ll}	$81\,\text{GeV} < m_{ll} < 101\,\text{GeV}$
p_T^V regions	$75\,\text{GeV} - 150\,\text{GeV}, > 150\,\text{GeV}$

Table 5.10: Summary of the additional event selections in the 0-, 1- and 2-lepton channels of the dijet mass analysis.

Channel			
Selection	0-lepton	1-lepton	2-lepton
m_T^W	-	< 120 GeV	-
$E_T^{miss}/\sqrt{S_T}$	-	-	$< 3.5\sqrt{GeV}$
p_T^V regions			
p_T^V	(75, 150] GeV 2-lepton channel only	(150, 200] GeV	(200, ∞)
$\Delta R(b_1, b_2)$	< 3.0	< 1.8	< 1.2

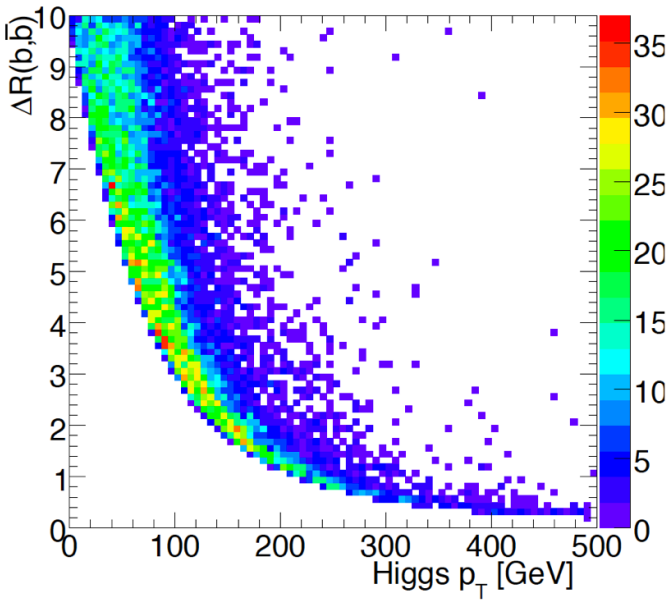


Figure 5.3: Distance in ΔR between the two b-quarks from the Higgs boson decay as a function of Higgs boson transverse momentum.

5.3.5 Truth b-jets tagging

The uncertainty in the expected number of events depends on the size of the simulated events. For the processes with large production cross-section but small selection efficiencies in the analysis, the production of the simulated events exceeding the integrated luminosity of the data is very challenging. For cases where such small efficiency is from the high rejection achieved by MV2c10 algorithm, a method call truth tagging is applied, to use the full simulated events of the samples, and keep the correct normalizations and shapes compared to selecting events directly based on the MV2c10 output (direct tagging).

For the 0 and 1 lepton channels, all V+jets samples with a c or light-jet filter and the WW MC sample are truth tagged in order to improve the statistical population of the V $+ll/cl/cc$ and WW templates provided to the likelihood fit. For the 2 lepton channel, where the WW contribution is much smaller than in the other two channels, truth tagging is used for the V $+ll/cl/cc$ templates only. In addition, for the 2 lepton channel the truth tagging is applied to any V $+ll/cl/cc$ events present in the V+b-jet filter samples. For all three channels all other samples are tagged using the direct tagging strategy, due to the relatively high production rate of b-jets within these remaining MC samples.

When using truth-tagging, all events pass the 2-tag requirement by construction. A combination of two jets in the event are randomly selected to be "tagged". The probability for a jet to be tagged is directly proportional to its b-tagging efficiency, which is a function of the jet's "real" flavour in MC, p_T and η. For a given tagging combination, a partial "truth-tagging" weight may be defined as the product of the b-tagging efficiencies of the two tagged jets times the product of one minus the efficiency of all untagged jets. The total truth-tagging event weight is taken to be the sum over all possible combinations, and the probability for selecting a given combination is directly proportional to its partial truth-tagging weight. For example, in an event with three jets, labeling the efficiency of the i^{th} jet as ε_i , the total truth tagging weight of the event is

$$\varepsilon_{tot} = \varepsilon_1\varepsilon_2(1-\varepsilon_3) + \varepsilon_1(1-\varepsilon_2)\varepsilon_3 + (1-\varepsilon_1)\varepsilon_2\varepsilon_3, \tag{5.2}$$

and the probability of selecting jets 1 and 2 as the tagged jets is

$$\frac{\varepsilon_1\varepsilon_2(1-\varepsilon_3)}{\varepsilon_{tot}}. \tag{5.3}$$

General good closure is found when comparing truth tagging to direct tagging.

In order to compensate the small remaining non-closure, large flavour composition priors are assigned to the ratios of each flavour to the Vbb process, with more deatils given in Section 5.7.

5.3.6 Jet energy corrections

In order to improve the b-jet energy measurement (scale and resolution), a few flavour-specific corrections are applied to b-tagged jets as shown in Figure 5.4, in addition to the standard JES correction as discussed in Section 4.5. The semileptonic decay of b- and c- hadron can produce muons which deposit very little fraction of their energy in the calorimeter. To correct for it, the muon-in-jet correction is used. When a muon with $p_T > 5\,\text{GeV}$ is found within $\Delta R = 0.4$ of a b-jet, the muon four momentum is added to the jet four momentum, while the energy deposited by the muon in the calorimeter is removed. If more than one muon is found within the jet cone, the muons are ordered according to the distance with respect to the jet axis and only the muon closest to the jet is used for the jet correction. Apart from muon-in-jet correction, a second correction, denoted as *PtReco*, is derived in bins of jet p_T. This correction takes into account the remaining expected difference from signal simulation between the reconstructed b-jets (with muon-in-jet correction applied) and the corresponding truth jets which formed by clustering final-state particles taken from the MC truth information, with muons and neutrinos included. This correction is also derived separately for jets with or without a lepton (muon or electron) found within $\Delta R = 0.4$ of the jet axis. The main feature of the PtReco correction is that for jets without matching a lepton, they increase the jet energy with around 12% at low p_T and decreases at high p_T to a plateau at around 1%, while for jets with matching a lepton, the corrections are about 10% larger across the jet p_T spectrum to account also for the missing neutrino energy.

In 2 lepton channel, the $ZH \to llb\bar{b}$ system can be fully reconstructed and the 2 leptons from Z boson have better momentum and energy resolution than those of b-jets. b-jet energy can then be adjusted by considering the balance of transverse momentum with a per-event kinematic likelihood fit, in place of the *PtReco* correction. The m_{bb} mass resolution is improved by 20-30% with respect to the muon-in-jet correction mass resolution, and the central value is moved closer to its nominal value as shown in Figure 5.4. The kinematic fit correction is applied to 2-jets and 3-jets events since the improvement is smeared out by the additional jets in the events with more jets.

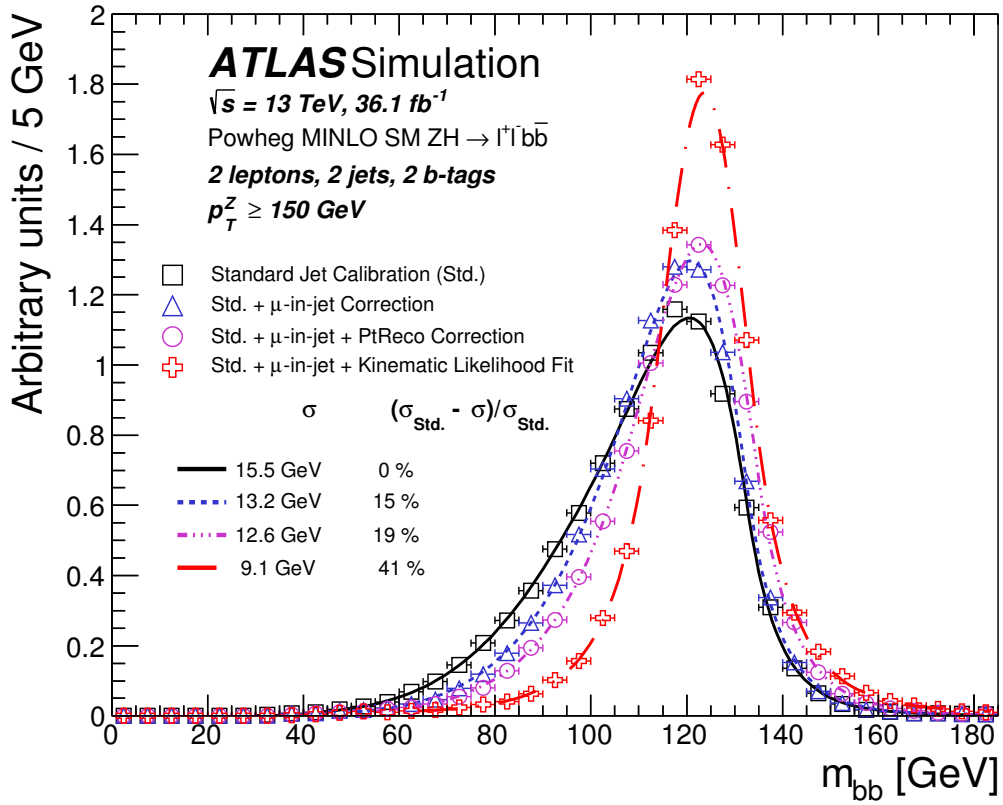


Figure 5.4: Comparison of the m_{bb} distributions as additional corrections are applied to the jet energy scale, shown for simulated events in the 2-lepton channel in the 2-jet and $p_T^Z > 150\,\text{GeV}$ region. A fit to a Bukin function is superimposed on each distribution, and the resolution values and improvements are reported in the legend.

5.3.7 Analysis regions

In the main version of the analysis, the BDT output discriminant is used in a binned maximum-likelihood fit, to extract the signal yield and the main background normalizations. A total of eight signal regions (SR) and six control regions (CR) are used in the fit and summarized in Table 5.11. The main purpose of controls regions is to help better constrain the modelling of background processes, with high purity for the dedicated background processes and negligible level of signal contamination.

Table 5.11: The distributions used in the global likelihood fit for the signal regions (SR) and control regions (CR) for all the categories in each channel, for the nominal multivariate analysis.

Channel	SR/CR	Categories			
		$75\,\text{GeV} < p_T^V < 150\,\text{GeV}$		$p_T^V > 150\,\text{GeV}$	
		2 jets	3 jets	2jets	3jets
0-lepton	SR	-	-	BDT	BDT
1-lepton	SR	-	-	BDT	BDT
2-lepton	SR	BDT	BDT	BDT	BDT
1-lepton	$W + HF$ CR	-	-	Yield	Yield
2-lepton	top $e\mu$ CR	m_{bb}	m_{bb}	Yield	m_{bb}

5.3.7.1 1-lepton $W + HF$ control region

In the 1-lepton channel, the normalization uncertainty on the $W + HF$ background is one of the largest systematic uncertainties from previous version of the analysis. Therefore a dedicated $W + HF$ CR is defined to better constrain the normalization of $W + HF$ background. To achieve a high $W + HF$ background purity in this CR, a cut on the reconstructed leptonically decaying top mass, m_{top} is introduced, with $m_{top} > 225\,\text{GeV}$ to reduced the dominated $t\bar{t}$ background in 1-lepton channel. In order to calculate m_{top}, the longitudinal momentum of the neutrino, p_z^ν, need to be determined first, by using W mass, m_W, as a constraint to solve the quadratic equation

$$p_z^\nu = \frac{1}{2(p_T^l)^2}\left[p_z^l X \pm E_l\sqrt{X^2 - 4(p_T^l)^2(E_T^{miss})^2}\right], \tag{5.4}$$

where

$$X = m_W^2 + 2p_x^l E_x^{miss} + 2p_y^l E_y^{miss}, \tag{5.5}$$

where $p_{x,y,z}^l$ are the x, y, and z components of the lepton's four momentum, and $E_{x,y}^{miss}$ are the x and y components of the missing transverse momentum. m_{top} is then reconstructed by selecting the jet from the two b-tagged jets and solution to p_z^ν which minimises the m_{top}. If p_z^ν has an imaginary solution, the E_T^{miss} is shifted such that the discriminant is equal to zero. Figure 5.5 shows the m_{top} distribution in the 1-lepton channel 2-jet and 3-jet SR and $W+HF$ CR. The $t\bar{t}$ background is peaked around the SM top mass in the SR, the $m_{top} > 225\,\text{GeV}$ cut is then selected to remove a large component of $t\bar{t}$ background and keep a significant number of $W+HF$ events in the meantime. To make sure the signal contribution in this CR is negligible, a cut on m_{bb} distributionm, $m_{bb} < 75\,\text{GeV}$, is requested, to remove $\sim$ 99.5% signal events. The $W+HF$ CR is cut from the SR, events passing these two cuts are placed into the $W+HF$ CR, otherwise they remain in the signal region, such that the two regions are fully orthogonal. As also shown in Figure 5.5, in the $W+HF$ CR, the $W+HF$ events purity is $\sim$ 80% ($\sim$ 75%) in the 2-jets (3-jets) region. The $W+HF$ CR is treated as one single bin in different jet categories and used only the yield information in the likelihood fit due to the limited statistics in this region.

5.3.7.2 2-lepton Top $e\mu$ control region

In the 2-lepton channel, the $t\bar{t}$ background is known as a flavor symmetric process. Therefore the high purity $t\bar{t}$ control region can be obtained by requiring different flavor of a pair of dilepton ($e\mu$ or μe), instead of requiring the same flavor as in SR. Lepton flavor does not expected to change the kinematics of $t\bar{t}$ background between SR and Top $e\mu$ CR, therefore the top background modeling in the SR can be constrained in Top $e\mu$ CR. An example of m_{bb} distribution in the Top $e\mu$ CR is shown in Figure 5.6. More than 99% events in this CR are from $t\bar{t}$ and Wt processes with almost 0 signal events contamination.

5.3.7.3 dijet mass analysis regions

In the dijet mass analysis, as already discussed in Section 5.3.4, an additional separation at $p_T^V = 200\,\text{GeV}$ is made in all there channels to exploit the larger sensitivity in the high p_T^V region. The signal and control regions used in the global likelihood fit for the dijet mass analysis are summarized in Table 5.12. In the

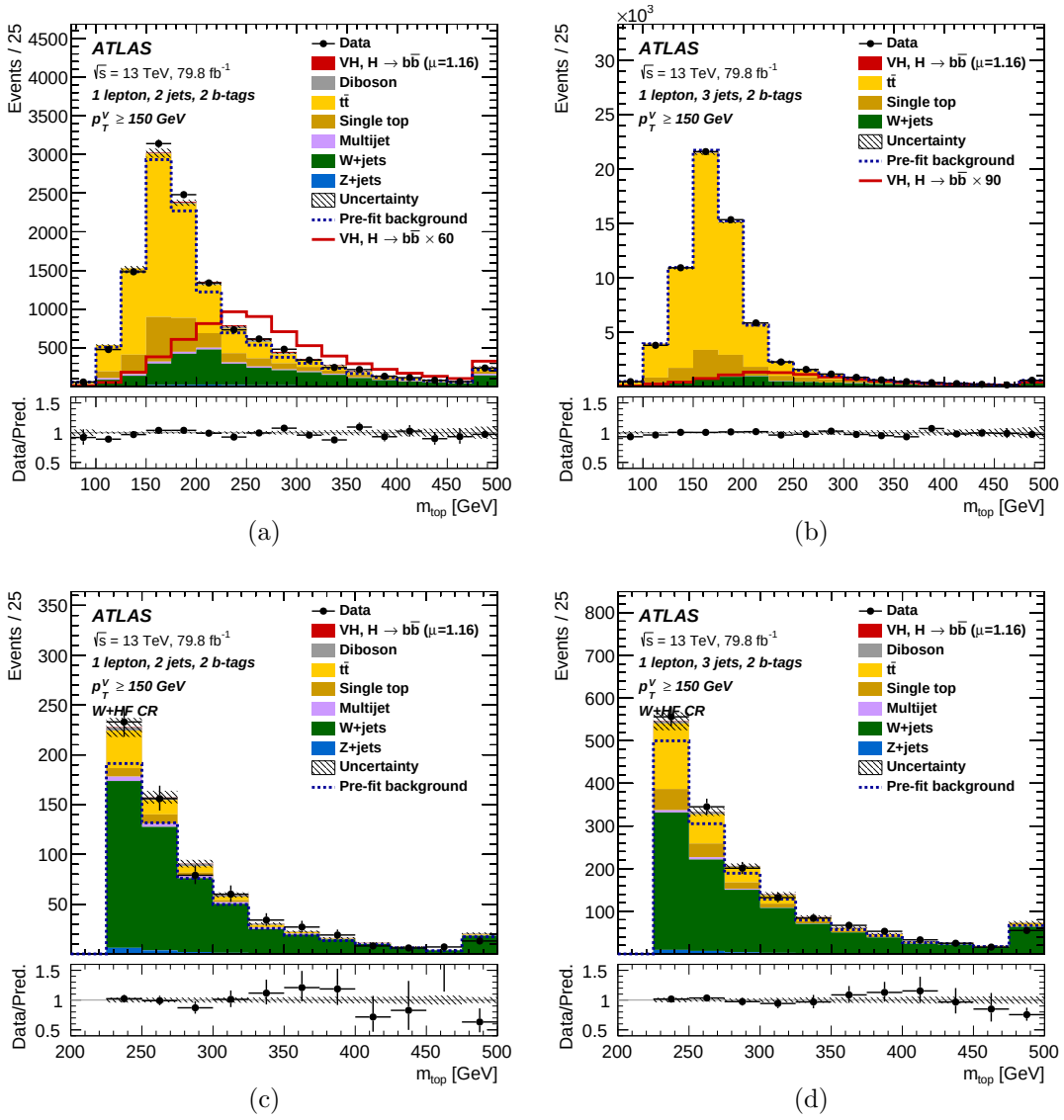


Figure 5.5: The m_{top} post-fit distributions from the global likelihood fit as described in Section 5.8 in the 1-lepton channel for 2-b-tag events, in the 2-jet SR (a), 3-jet SR (b), 2-jet $W + HF$ CR (c) and 3-jet $W + HF$ CR (d). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data, and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. In the $W + HF$ CRs, the unstacked unfilled histograms for the signal are not shown .The entries in overflow are included in the last bin. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

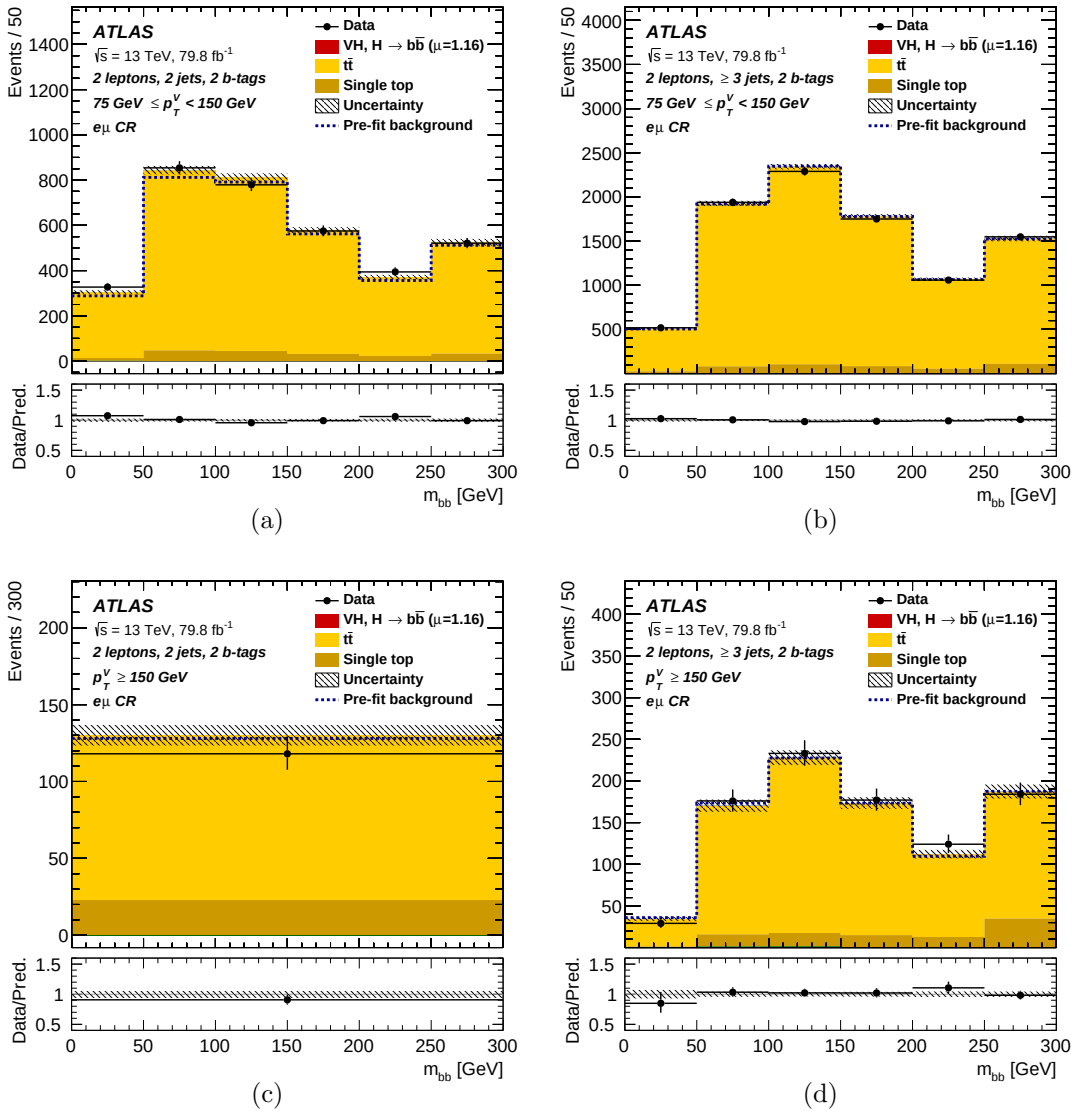


Figure 5.6: The m_{bb} post-fit distributions from the global likelihood fit as described in Section 5.8 in the 2-lepton channel for 2-b-tag events, in the 2-jet 75 GeV $< p_T^V <$ 150 GeV region (a), 3-jet 75 GeV $< p_T^V <$ 150 GeV (b), 2-jet $p_T^V >$ 150 GeV (c) and 3-jet $p_T^V >$ 150 GeV (d). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data. The entries in overflow are included in the last bin. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

1-lepton channel, the $W + HF$ CR is merged into signal region since the low m_{bb} region can already provide sufficient constraint for the $W + HF$ events. In the 2-lepton channel, the additional separation at $p_T^V = 200\,\text{GeV}$ is not considered for top $e\mu$ CR, in order to reduce the statistical uncertainties.

Table 5.12: The distributions used in the global likelihood fit for the dijet mass analysis, for the signal regions (SR) and control regions (CR) for all the categories in each channel. The two regions marked with * and $^\bullet$ are merged into a single region, to reduce statistical uncertainties.

Channel	SR/CR	Categories					
		$75\,\text{GeV} < p_T^V < 150\,\text{GeV}$		$150\,\text{GeV} < p_T^V < 200\,\text{GeV}$		$p_T^V > 200\,\text{GeV}$	
		2 jets	3 jets	2jets	3jets	2jets	3jets
0-lepton	SR	-	-	m_{bb}	m_{bb}	m_{bb}	m_{bb}
1-lepton	SR plus $W + HF$ CR	-	-	m_{bb}	m_{bb}	m_{bb}	m_{bb}
2-lepton	SR	m_{bb}	m_{bb}	m_{bb}	m_{bb}	m_{bb}	m_{bb}
2-lepton	top $e\mu$ CR	m_{bb}	m_{bb}	Yield*	$m_{bb}^\bullet$	Yield*	$m_{bb}^\bullet$

5.4 Multivariate Analysis

Multivariate analyses (MVAs) are used in a variety of high energy physics analyses to offer increased signal purity and background rejection. This is achieved through the combination of a well-chosen set of discriminating input variables which the multivariate algorithm is trained on, to construct a one dimensional discriminant. Such as the SM $VH(b\bar{b})$ analysis described in this thesis, the dijet mass, m_{bb}, is the variable which has largest discrimination between background and signal events, and provides largest sensitivity to the analysis, however, there are still a number of the other variables can separated signal from the background events and can be used to increase the sensitivity of the analysis, such as $\Delta R(b, \bar{b})$ and p_T^V. The algorithm is set up taking into account the available MC statistics, so that the final result does not depend on the random statistical fluctuations in the input distributions. On the other hand, multivariate algorithms must be trained and evaluated on separate MC samples to ensure an unbiased result: in this analysis a 2-fold cross-validation of the training is implemented. One training is performed using even (odd) event-numbered MC events, and then applied to odd

(even) events, thereby ensuring orthogonality between the samples the algorithm is trained on and evaluted on. The final discriminant is then build by summing the multivariate discriminant of the even and odd events since no difference in the physics is expected between them. For this analysis, a Boosted Decision Tree (BDT) provided by the TMVA package [129] of ROOT [130] is used, similarly to what was done in the Run 1 analysis. Due to varying kinematics and background compositions in each signal region of the analysis as shown in Table 5.11, a separate training is performed in each signal region in the aformentioned two-fold way to increase the sensitivity of the analysis.

5.4.1 Input variables

The input variables used in each channel are summarised in Table 5.13. They were chosen based on studies conducted during Run-1, where an iterative procedure was adopted to find the optimal set of variables and ranking to use. Initially, the BDT was constructed using simply the invariant mass of the two b-tagged jets and ΔR between them, which provide the most discriminating distributions. Each candidate variable was then added to the MVA in turn, with the variable offering the best improvement in significance being added to the MVA as the third variable. The final optimal MVA is then constructed when all variables have been studied and no further improvement is seen. This was done separately for each lepton channel. The similar procedure is repeated with Run 2 MC simulation events, and the default input variables inherited from Run 1 analysis have been proved still optimal for the analysis and thus kept in the training, only two more variables are added in 1-lepton channel training, m_{top} as already described in Section 5.3 and $\Delta Y(W,H)$ with more details given in below shortly. The input variables that are commonly used in all lepton channels are defined as follows:

- m_{bb}: invariant mass of the dijet system constructed from the two b-tagged jets
- $\Delta R(b_1, b_2)$: distance in η and ϕ between the two b-tagged jets
- $p_T^{b_1}$: transverse momentum of the b-tagged jet in the dijet system with the higher p_{T}
- $p_T^{b_2}$: transverse momentum of the b-tagged jet in the dijet system with the lower p_{T}

- p_T^V: transverse momentum of the vector bosos; given by E_T^{miss} in the 0 lepton channel, vectorial sum of E_T^{miss} and the transverse momentum of the lepton in the 1 lepton channel and vectorial sum of the transverse momenta of the two leptons in the 2 lepton channel

- $\Delta\phi(V, bb)$: distance in ϕ between the vector boson candidate, i.e. E_T^{miss} in the 0 lepton channel, E_T^{miss} and the lepton in the 1 lepton channel and the di-lepton system in the 2 lepton channel, and the Higgs boson candidate, i.e. the dijet system constructed from the two b-tagged jets

- $p_T^{jet_3}$: transverse momentum of the jet with the highest transverse momentum amongst the jets that are not b-tagged; only used for events with 3 or more jets

- m_{bbj}: invariant mass of the two b-tagged jets and the jet with the highest transverse momentum amongst the jets that are not b-tagged; only used for events with 3 or more jets

0 lepton channel uses two additional variables:

- $|\Delta\eta(b_1, b_2)|$: distance in η between the two b-tagged jets

- m_{eff} : scalar sum of E_T^{miss} and the p_T of all jets present in the event

1 lepton channel uses five additional variables:

- E_T^{miss}: missing transverse energy of the event

- $min[\Delta\phi(l, b)]$: distance in ϕ between the lepton and the closest b-tagged jet

- m_T^W : transverse mass of the W boson candidate, more details see 5.3.

- $\Delta Y(V, bb)$: difference in rapidity between the Higgs boson candidate and W boson candidate, the four-vector of the neutrino in the W boson decay is estimated as explained in Section 5.3 for m_{top}.

- m_{top}: reconstructed mass of the leptonically decaying top quark, more details see Section 5.3.

2 lepton channel uses three additional variables:

- E_T^{miss} significance: quasi-significance of the E_T^{miss} in the event, defined as $E_T^{miss}/\sqrt{S_T}$ with S_T the scalar sum of the p_T of the leptons and jets in the event.

- $|\Delta\eta(V, bb)|$: distance in η between the dilepton and dijet system of the b-tagged jets
- m_{ll}: invariant mass of the dilepton system

Table 5.13: Variables used for the multivariate discriminant in each of the categories.

Variable	0-lepton	1-lepton	2-lepton
p_T^V	$\equiv E_T^{miss}$	✓	✓
E_T^{miss}	✓	✓	
E_T^{miss} significance			✓
$p_T^{b_1}$	✓	✓	✓
$p_T^{b_2}$	✓	✓	✓
m_{bb}	✓	✓	✓
$\Delta R(b1, b2)$	✓	✓	✓
$\|\Delta\eta(b1, b2)$	✓		
$\Delta\phi(V, bb)$	✓	✓	✓
$\|\Delta\eta(V, bb)$			✓
m_{eff}	✓		
$min[\Delta\phi(l, b)]$		✓	
m_T^W		✓	
m_{ll}			✓
m_{top}		✓	
$\Delta Y(V, bb)$		✓	
	Only in 3-jet events		
$p_T^{jet_3}$	✓	✓	✓
m_{bbj}	✓	✓	✓

Since most of the kinematic variables have tails towards very high values, the range of the input variables is limited to a range that includes 99% of all signal events. All events above those limits will be artificially set to the maximum value. This procedure is introduced to avoid that the BDT wastes degrees of freedom to categorise the small number of events that accumulate in the tails of

these distributions. In addition, since statistics is crucial for the training of a multivariate algorithm to get a more stable, optimal performance, truth tagging as described in Section 5.3 is applied to all samples to increase the MC statistics in the training procedure.

5.4.2 Setup and training

The set of training parameters used for the BDT was optimized for the Run-1 analysis. A one-dimensional scan of each of the parameters was performed to obtain the optimal configuration shown in Table 5.14. It has been checked that this setup is still optimal for the Run-2 analysis as well. The BDT is trained using MC samples, combining the samples of the mc16a and mc16d production period. The VH samples are the signal template and the sum of all background samples is the background template for the training. For the diboson cross check analysis the diboson samples are used as the signal and the VH samples are added to the background template. No further changes are made for the diboson training.

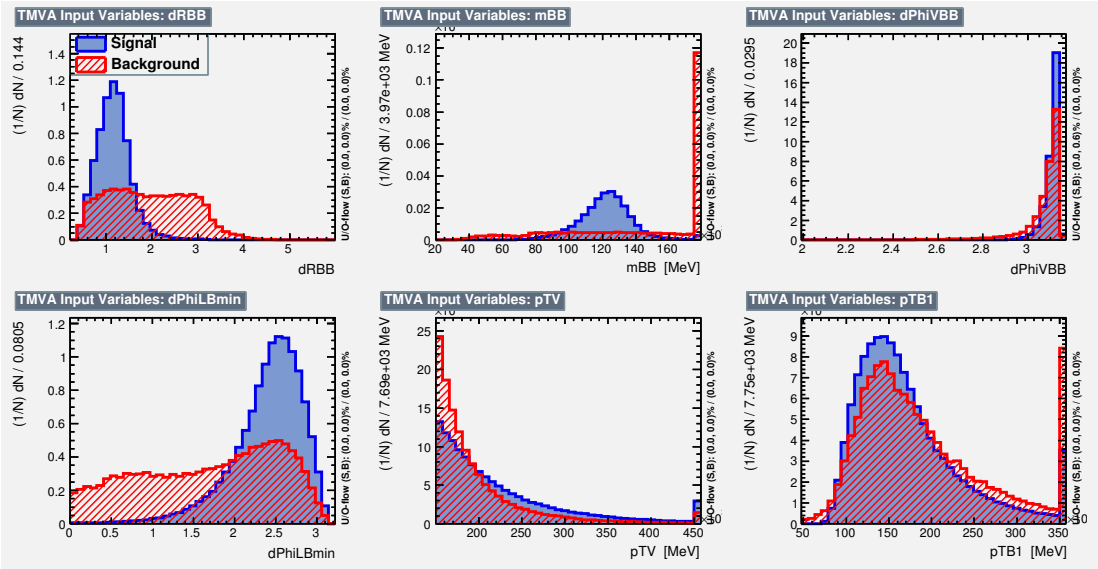
Table 5.14: BDT configuration parameters.

TMVA Setting	Value	Definition
BoostType	AdaBoost	Boost procedure
AdaBoostBeta	0.15	Learning rate
SeparationType	GiniIndex	Node separation gain
PruneMethod	NoPruning	Pruning method
NTrees	200	Number of trees
MaxDepth	4	Maximum tree depth
nCuts	100	Number of equally spaced cuts tested per variable per node
nEventsMin	5%	Minimum fraction of training events used in a node

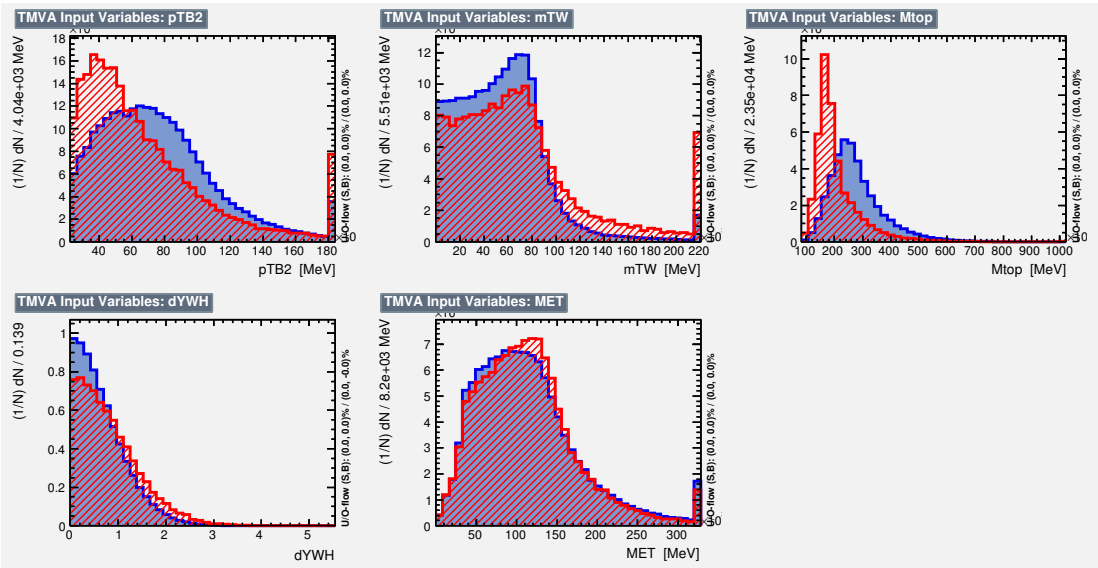
5.4.2.1 1-lepton training

The overall signal and background input distributions passed to the VH BDT traininig in the 1-lepton channel are shown in Figure 5.7 for 2-jets events and in Figure 5.8 for 3-jets events. Similar plots for the diboson traininig are shown in 5.9 and 5.10 for 2-jets and 3-jets events, respectively.

The VH BDT discriminants obtained for the signal (blue) and sum of all backgrounds (red) comparing the training events (dots) and testing events (histogram)

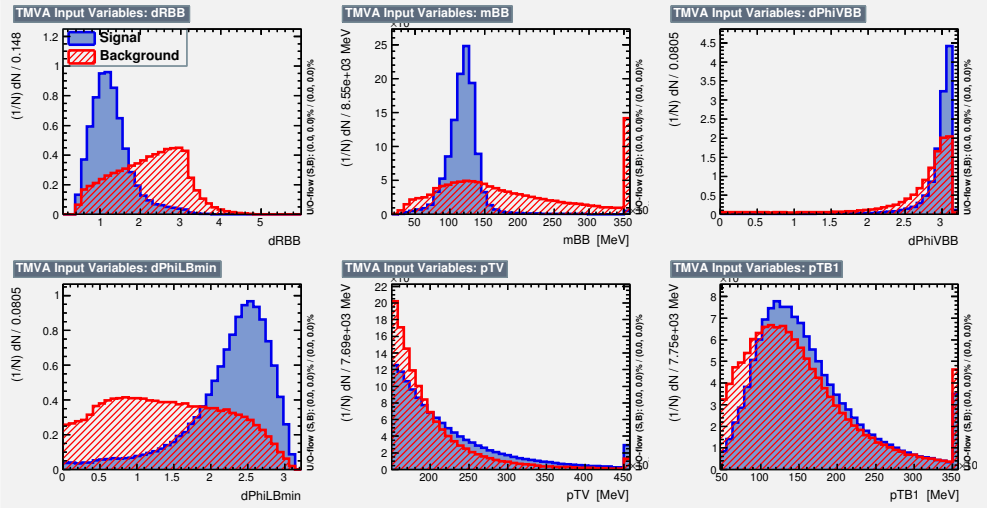


(a)

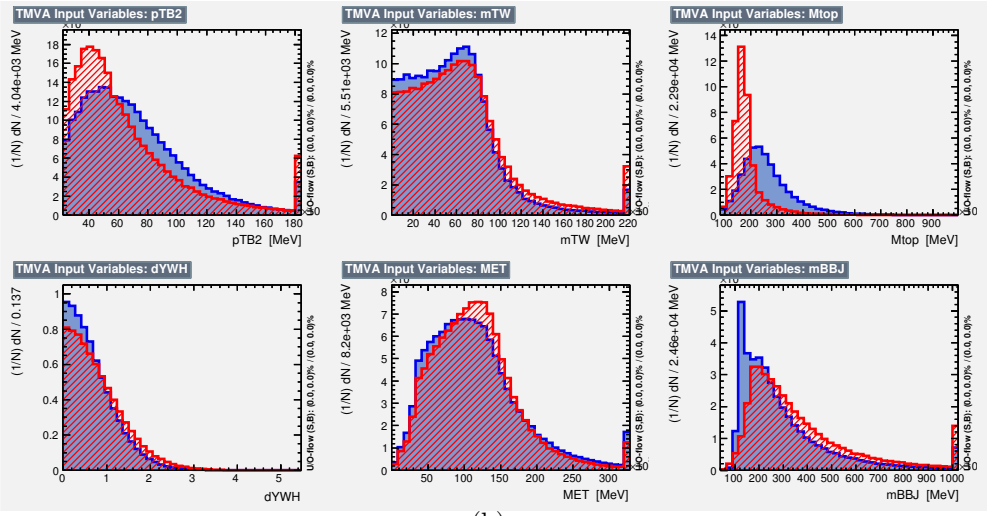


(b)

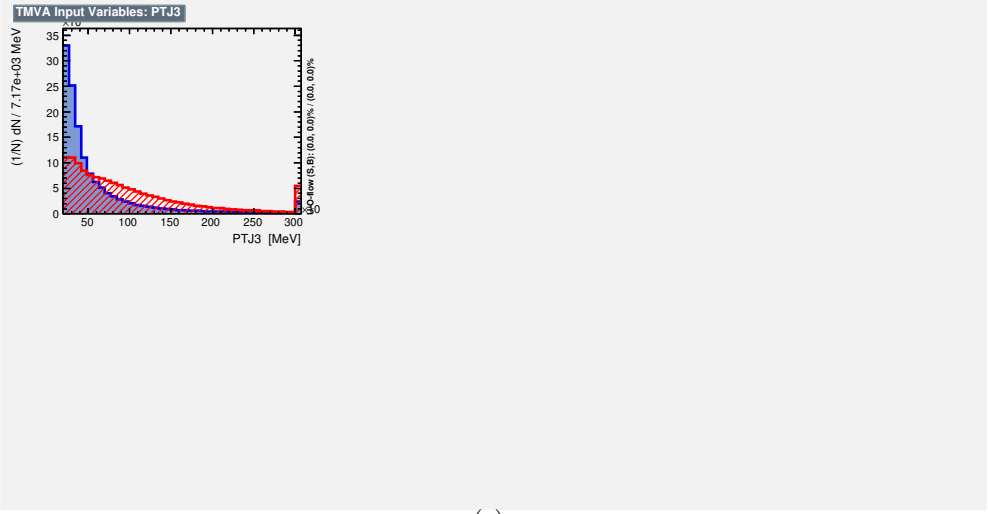
Figure 5.7: Distributions of input variables used in the VH BDT traininig for signal (blue) and background (red) samples in the 2-jet region of the 1-lepton channel.



(a)

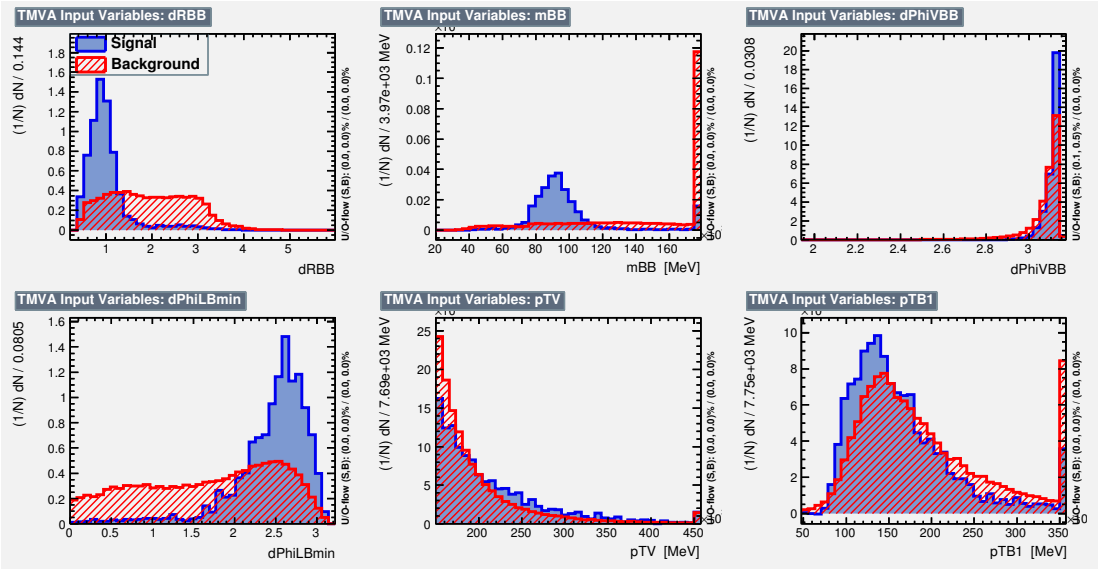


(b)

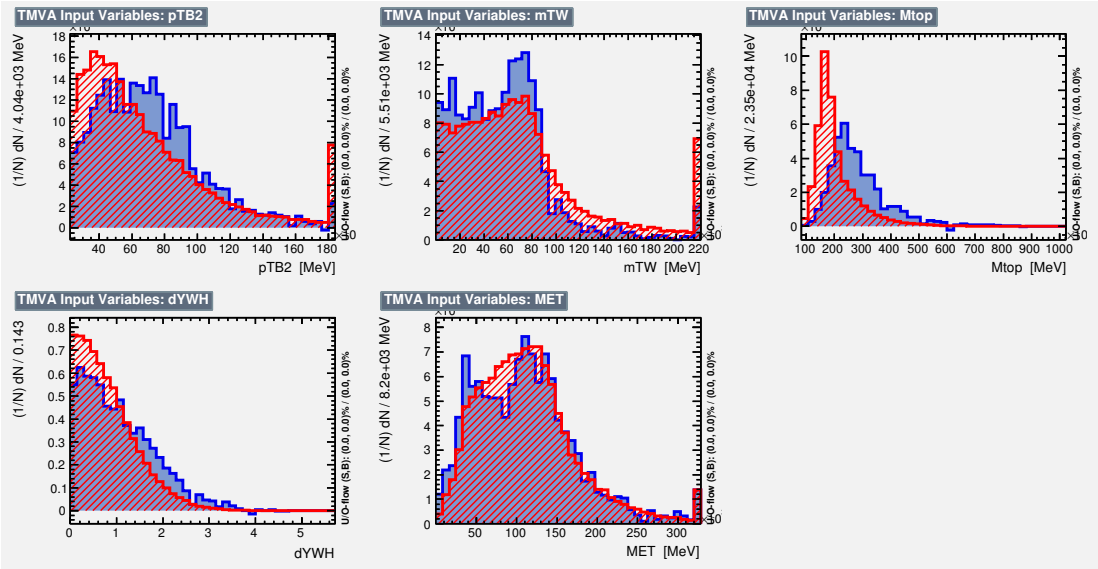


(c)

Figure 5.8: Distributions of input variables used in the VH BDT traininig for signal (blue) and background (red) samples in the 3-jet region of the 1-lepton channel.

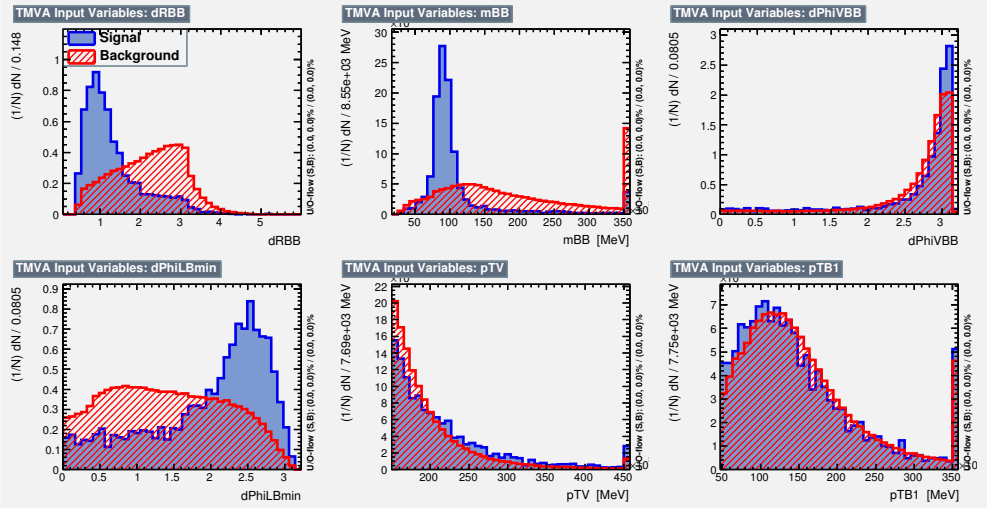


(a)

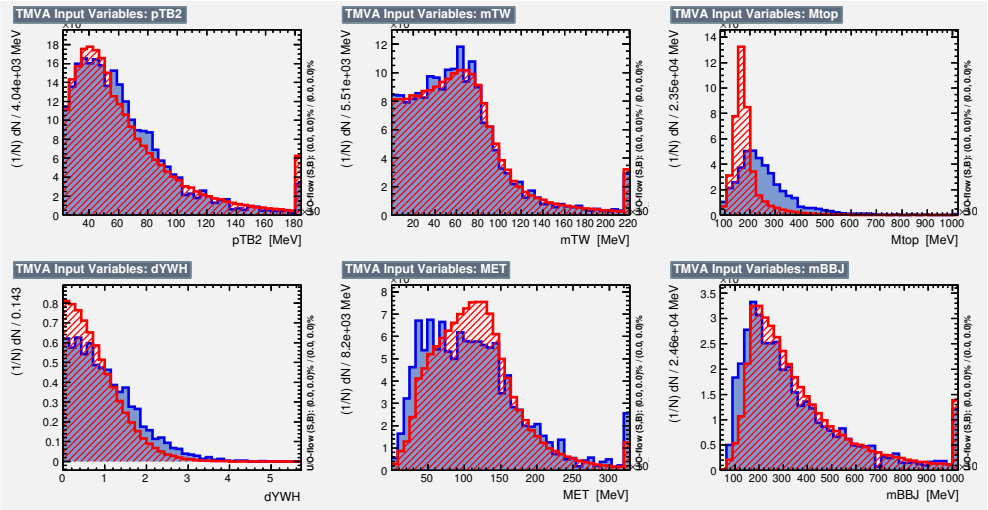


(b)

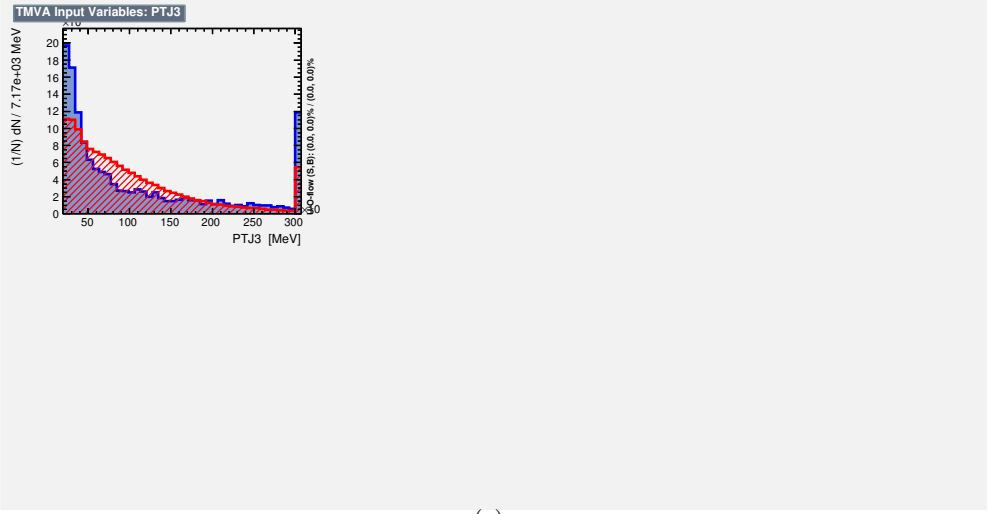
Figure 5.9: Distributions of input variables used in the diboson BDT training for signal (blue) and background (red) samples in the 2-jet region of the 1-lepton channel.



(a)



(b)



(c)

Figure 5.10: Distributions of input variables used in the diboson BDT traininig for signal (blue) and background (red) samples in the 3-jet region of the 1-lepton channel.

in 1-lepton channel are shown in Figure 5.11, both folds of the training, even (a) and odd (b) events training, are shown, i.e. the testing events are the events with odd (a) and even (b) event numbers. In all cases a reasonable agreement between the training and testing events is observed, indicating that the training is insensitive to statistical fluctuations of the training data set, i.e. the overtraining is under control. Similar plots for the diboson BDT are shown in Figure 5.12.

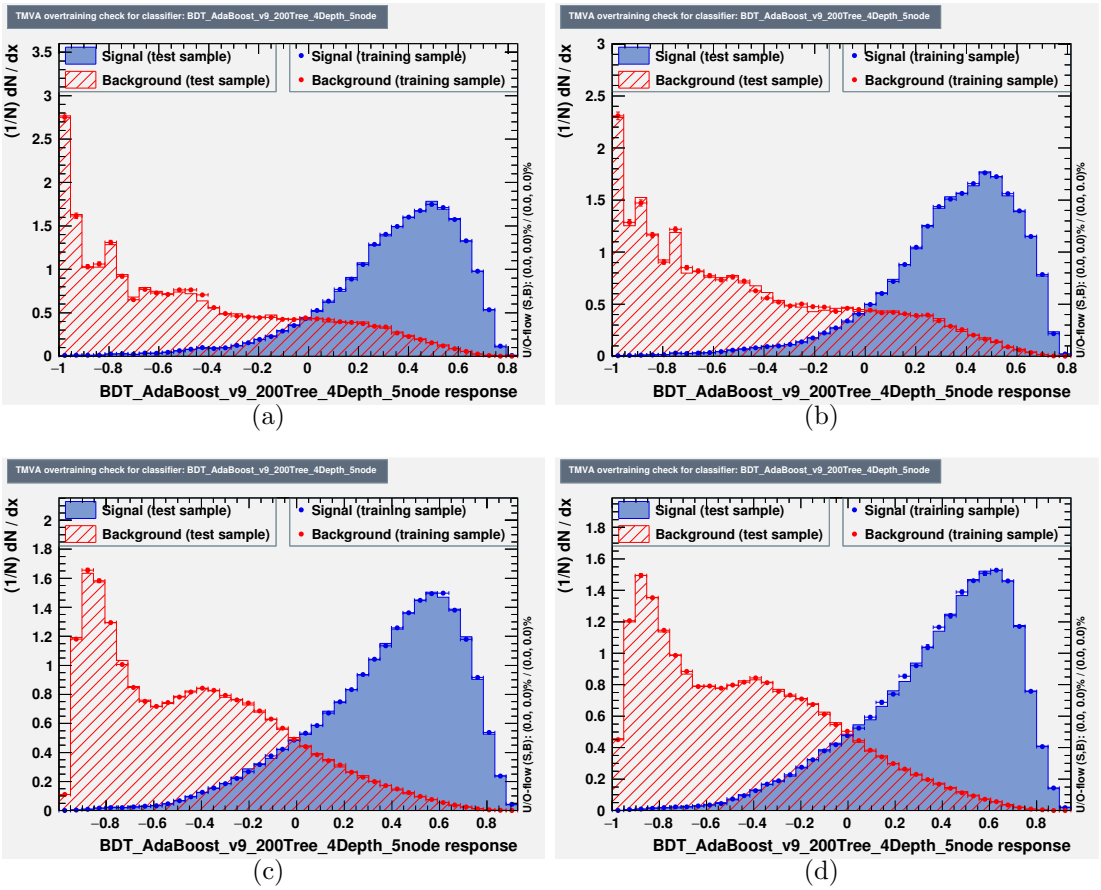


Figure 5.11: The VH BDT distributions of the signal (blue) and sum of all background (red) processes in the 1 lepton 2 jet region obtained while training (dots) and testing (histogram), with using even (a) and odd (b) event number for the training, and in the 3 jet region with using even (c) and odd (d) events for the training.

Correlations between the input variables are shown in Figures 5.13 for VH BDT trainning in both 2-jets and 3-jets regions, and for both signal and background processes. The similar plots for diboson training are shown in Figure 5.14.

The background rejection as a function of signal efficiency for the VH BDT training in 1-channel are shown in Figure 5.15. The similar plots for diboson BDT

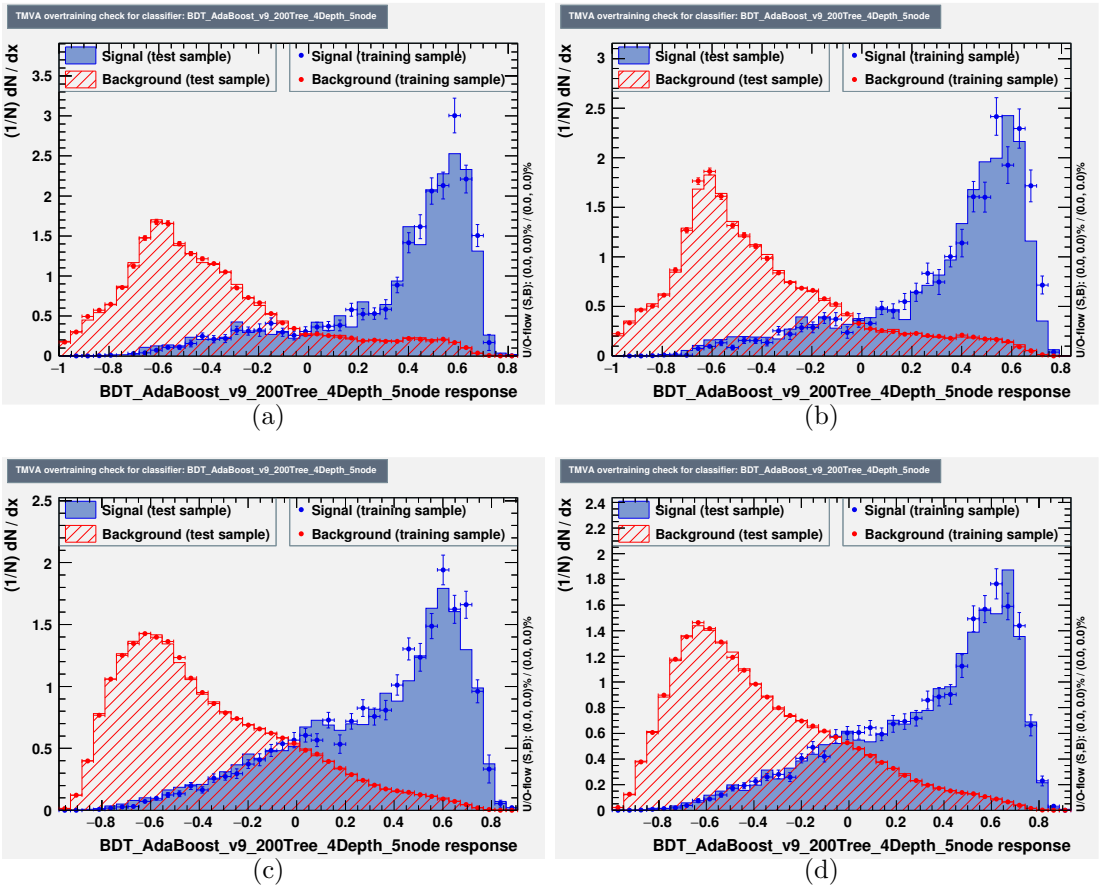


Figure 5.12: The diboson BDT distributions of the signal (blue) and sum of all background (red) processes in the 1 lepton 2 jet region obtained while training (dots) and testing (histogram), with using even (a) and odd (b) event number for the training, and in the 3 jet region with using even (c) and odd (d) events for the training.

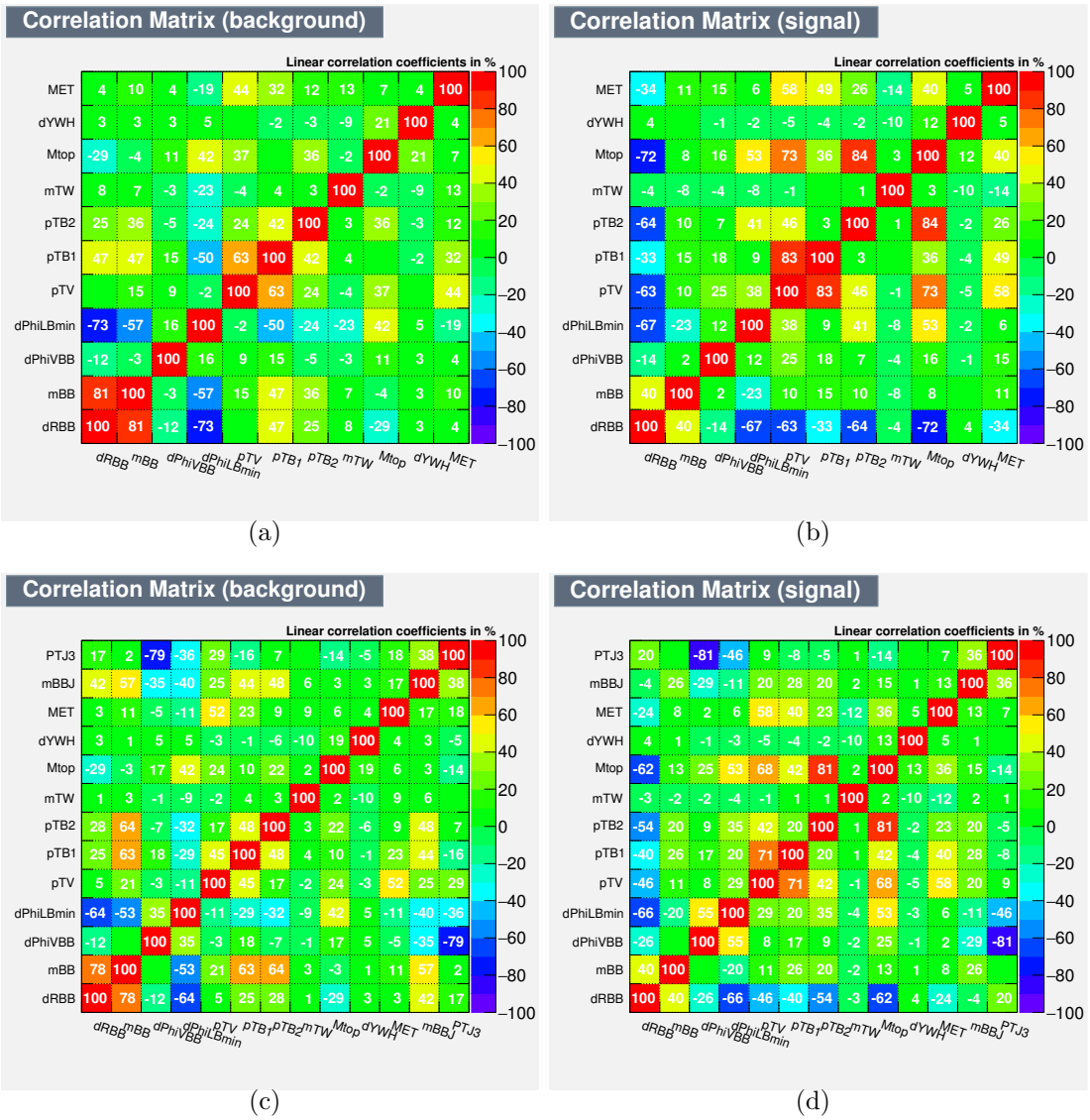


Figure 5.13: Correlation matrices of the VH BDT input variables in the 1-lepton 2-jets region for the sum of all background processes (a) and the signal process (b), and in the 3-jets for the sum of all background processes (c) and the signal process (d).

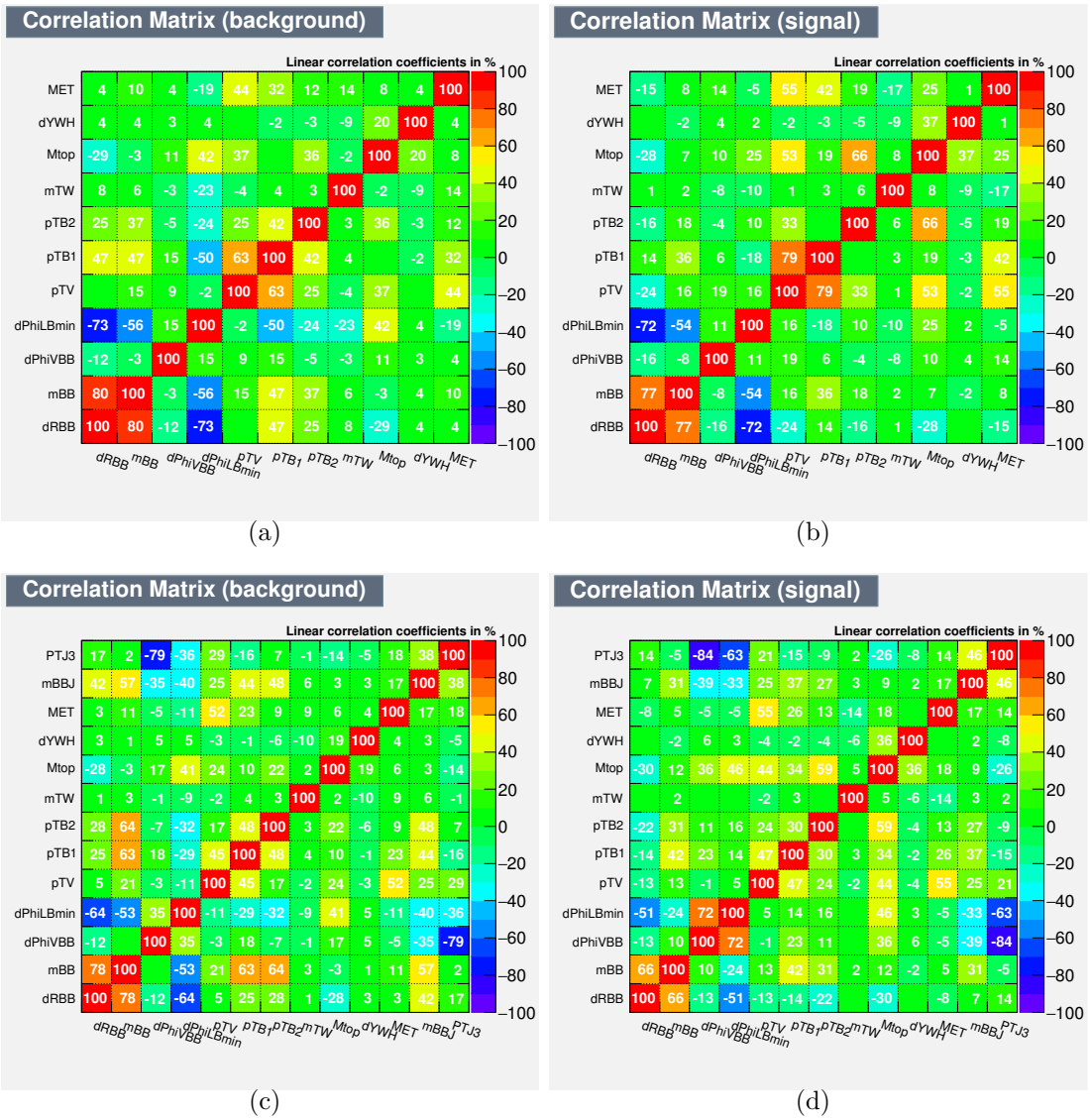
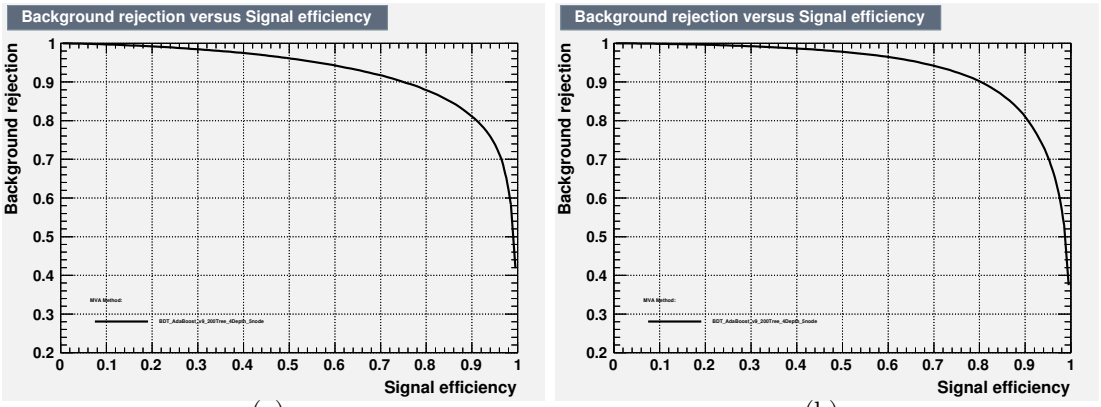


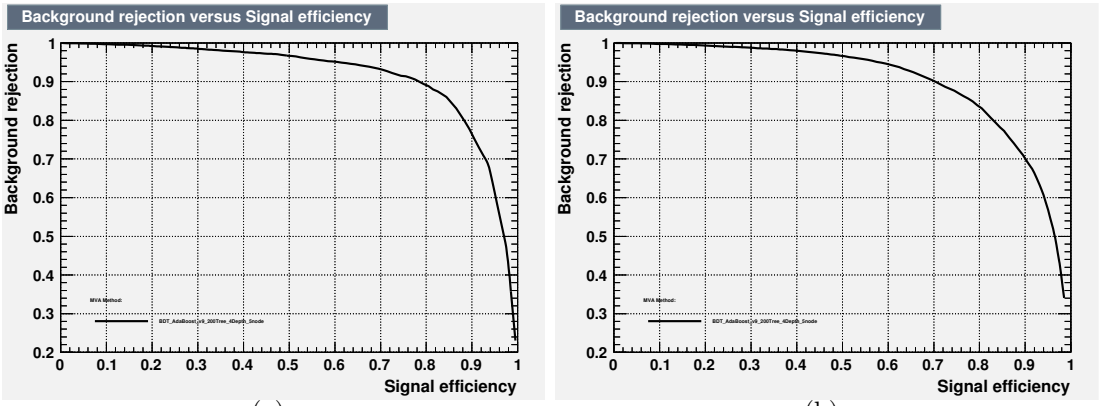
Figure 5.14: Correlation matrices of the diboson BDT input variables in the 1-lepton 2-jets region for the sum of all background processes (a) and the signal process (b), and in the 3-jets for the sum of all background processes (c) and the signal process (d).

training are shown in Figure 5.16.



(a) (b)

Figure 5.15: The background rejection as a function of signal efficiency for the VH BDT training in 1-lepton channel in 2-jets region (a) and 3-jets region (b).

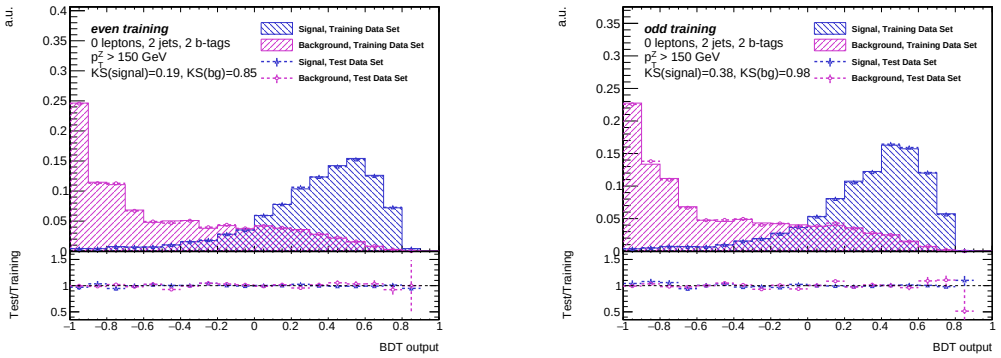


(a) (b)

Figure 5.16: The background rejection as a function of signal efficiency for the diboson BDT training in 1-lepton channel in 2-jets region (a) and 3-jets region (b).

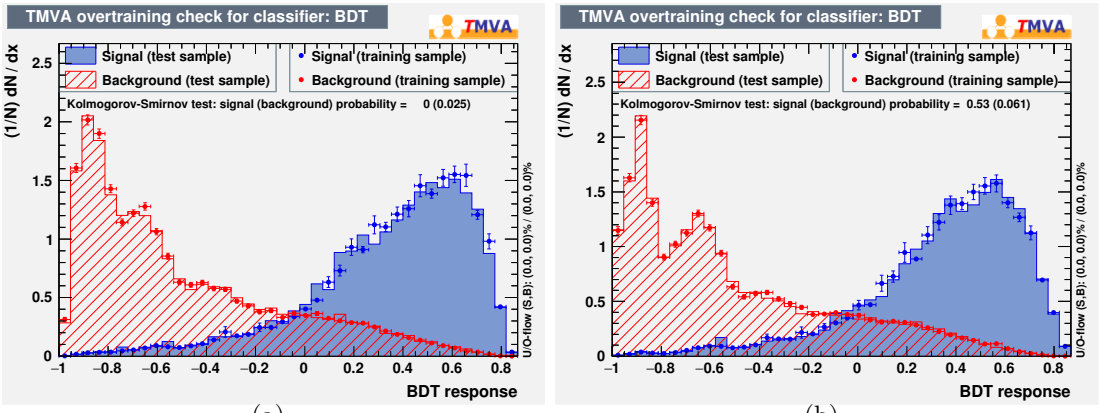
5.4.2.2 0- and 2- lepton training

The BDT traninig performance in 0- and 2- lepton channels are very similar as those in 1-lepton channel. Figure 5.17 shows the VH BDT discriminant obtained for the signal (blue) and sum of all backgrounds (red) comparing the training events (dots) and testing events (histogram) for 0-lepton channel in 2-jets region, similar plots for 2-lepton channel in $p_T^V > 150\,\text{GeV}$ 2-jets region are shown in Figure 5.18. In all cases training and testing are found to be in reasonable agreement.



(a) (b)

Figure 5.17: The VH BDT distributions of the signal (blue) and sum of all background (red) processes in the 0 lepton 2 jet region obtained while training (dots) and testing (histogram), with using even (a) and odd (b) events for the training.



(a) (b)

Figure 5.18: The VH BDT distributions of the signal (blue) and sum of all background (red) processes in the 2 lepton $p_T^V > 150\,\text{GeV}$ 2 jet region obtained while training (dots) and testing (histogram), with using even (a) and odd (b) events for the training.

5.4.3 BDT transformation

Since the output of the BDT is designed to maximise the separation of the signal and background populations, the optimal performance is not necessarily achieved with the default binning. For example, in the dijet mass tails, wider bins are required to reduce statistical uncertainty, but this is at a cost to the BDT sensitivity. Therefore, a transformation of the BDT output is studied and implemented in order to optimise the final analysis sensitivity.

As a general description, to remap the histograms entering the final fit, consider the function:

$$Z(I[k,l]) = Z(z_s, n_s(I[k,l]), N_s, z_b, n_b(I[k,l]), N_b), \tag{5.6}$$

where

- $I[k,l]$ is an interval of the histograms, containing the bins between bin k and bin l;
- N_s is the total number of signal events in the histogram;
- N_b is the total number of background events in the histogram;
- $n_s(I[k,l])$ is the total number of signal events in the interval $I[k,l]$;
- $n_b(I[k,l])$ is the total number of background events in the interval $I[k,l]$;
- z_s and z_b are parameters used to tune the algorithm.

There are several different possible Z functions exist to transform the BDT output, in the Run-1 analysis, the implementation of Transformation D was found to offer a significant decrease in the number of bins, whilst comparatively increasing the expected sensitivity:

$$Z = z_s n_s/N_s + z_b n_b/N_b. \tag{5.7}$$

The re-binning is then conducted using the following algorithm:

1. Starting from the last bin on the right of the original histogram, increase the range of the interval $I(k, last)$ by adding one after the other, the bins from the right to the left;
2. Calculate the value of Z at each step;

3. Once $Z(I[k_0, last]) > 1$ and the MC statistical uncertainty in the range is less than 20%, rebin all the bins in the interval $I(k_0, last)$ into a single bin;

4. Repeat steps 1-3, starting this time from the last bin on the right, not included in the previous remap (the new last is $k_0 - 1$), until k_0 in the first bin.

For the current analysis these sensitivity studies have been repeated and find $z_s = 10, z_b = 5$ as optimal parameters. Due to limited MC statistics of the diboson samples the parameters are changed to $z_s = 5, z_b = 5$ for the diboson cross check analysis.

5.5 Estimation of The Multi-jet Background

The background MC samples summarized in Table 5.1 are used to model the processes with W or Z boson decay into leptons, such processes (including W bosons from top-quark decays) are defined as electroweak backgrounds in this thesis. The multijet background provides no genuine leptonic signatures, but still has the potential to contribute a non-negligible background component due to the large cross-section. Using the Monte Carlo technology to achieve a good modelling of this background is also very difficult, therefore data driven approaches are used instead. In this section, the estimation of this background is discussed channel by channel.

5.5.1 0-lepton channel

In the 0-lepton channel the multijet background mainly enters due to jet energy mis-measurements. As a result, the fake missing transverse energy and momentum tend to be aligned with the mis-measured jet. As already discussed in Section 5.3.3, a set of anti-QCD cuts are applied to reduce the multijet background contamination. In order to estimate the remaining multijet contribution, the anti-QCD cuts are loosened by removing the $min[\Delta\phi(E_T^{miss}, jets)]$ cut. A fit to this distribution in the 3-jets region is then performed to extract the multijet yields. The multijet contribution is expected peaked at low $min[\Delta\phi(E_T^{miss}, jets)]$ region, and is parameterized with a falling exponential function $(A \cdot exp(-x/c))$ as predicated by a Pythia 8 MC sample generated with the A14 tune and NNPDF2.3LO PDFs. The parameter A and c are determined by the fit, while the template for the other electroweak background are taken directly from MC simulation. In

order to account for normalization differences between the electroweak MC background and data in this specific phase space region, a fit is perform to data in $min[\Delta\phi(E_T^{miss}, jets)] > 40°$ region while allowing the $W + jets$, $Z + jets$ and $t\bar{t}$ background normalization to float. Then the multijet yield can be extracted by fitting the exponentially falling multijet function and the scaled electroweak background templates to the data in the $min[\Delta\phi(E_T^{miss}, jets)] < 50°$ region. The post-fit $min[\Delta\phi(E_T^{miss}, jets)$ distribution in 3-jets region is shown in Figure 5.19, with the fitted parameters for the falling exponential function given in the caption.

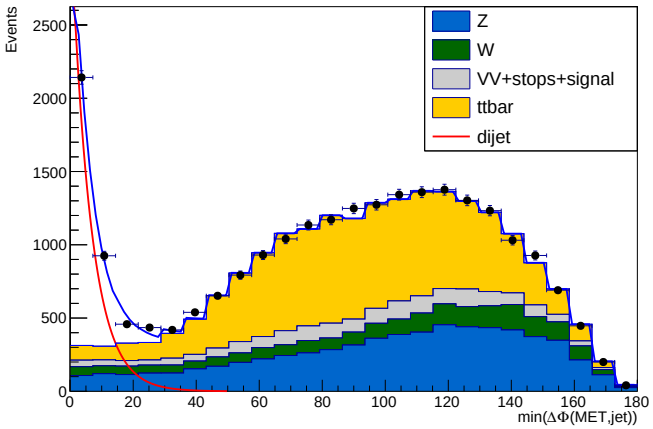


Figure 5.19: Post-fit $min[\Delta\phi(E_T^{miss}, jets)$ distribution in the 0-lepton channel 2-btag 3-jets region. The multijet is modeled using an exponential shape $A \cdot exp(-x/c)$, the fitted value of paramenter A is 3264.1 $\pm$ 130.4, while the fitted value of paramenter c is 6.27 $\pm$ 0.24.

After applying the nominal selection criteria with $min[\Delta\phi(E_T^{miss}, jets)] > 30°$ in 3-jets region, the residual muleijet contamination is found to be less than 10% of the expected signal contribution and negligible with respect to the total background. Furthermore, the BDT shape of the multijet background is studied by selecting the events within the $min[\Delta\phi(E_T^{miss}, jets)] < 20°$ region by subtracting the electroweak backgrounds from data and is found to have the similar shape to the one expected for the sum of all the electroweak backgrounds. The small multijet contribution is therefore can be absorbed in the floating normalization factors of the electroweak backgrounds in the global likelihood fit. In the 2-jets region, the similar fit can not be used since the events in the low value of $min[\Delta\phi(E_T^{miss}, jets)$ region have been already removed by the other anti-QCD cuts. However the multijet shape in 2-jets region is predicted by the MC to have the same exponential behavior as in the 3-jets region, the nominal anti-QCD selections are safe enough to reduce most of muleijet contribution in the 2-jets region. Therefore the multijet

background in the 0-lepton channel is found to be a small enough background and can be neglected in the global likelihood fit.

5.5.2 1-lepton channel

The multijet background contributes to both the electron and muon sub-channels. The dominant contribution to this background comes from the real electrons or muons from semileptonic decay of the heavy flavor hadrons. A second contribution in electron sub-channel stems from the $\gamma \to e^+e^-$ conversions where photons are produced in the decays of neutral pions or from π^0 Dalitz decay. These non-prompt leptons are not expected to be isolated, but still a non-negeligible fraction passes the isolation requirements. A robust procedure is necessary to estimate the contribution of this background both in the electron and muon sub-channels. Even though the medium p_T^V region is not included in the final global likelihood fit in this analysis, this section will discuss the multijet contribution in both high and medium p_T^V regions. The medium p_T^V region, with much enhanced multjet contribution compared to the high p_T^V region, can also provides a better appreciation of the quality of the modeling of the multijet background. This background is estimated separately not only in high and medium p_T^V regions, but also in the electron and muon sub-channels, and in the 2- and 3-jets categories, using the similar procedures.

5.5.2.1 Isolation requirements optimization

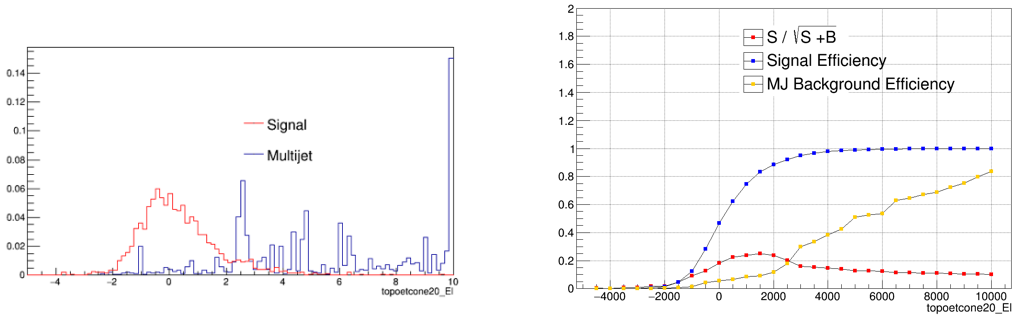
In an earlier version of the Run-2 analysis using 13.2 fb^{-1} of data, the LooseTrackOnly isolation working point was used in both electron and muon sub-channels, in addition, FixedCutTight isolation working point was used in electron sub-channel, which corresponds to the selections of $E_T^{cone0.2}/p_T < 0.06$ and $p_T^{varcone0.2}/p_T < 0.06$, FixedCutTightTrackOnly working point was used in the muon sub-channel, which corresponds to the selection of $p_T^{varcone0.3}/p_T < 0.06$. A study is performed to re-optimize the tighter isolation working points used in electron and muon sub-channels, while still keeping the LooseTrackOnly working point on top, with the purpose of reducing more multijet background while keeping similar signal acceptance in 1-lepton channel. Only high p_T^V region is considered in this study and the optimized isolation requirements are also used in the medium p_T^V region. This study is based on the WH signal and multijet background MC samples which are simulated and reconstructed using the 2015-2016 data running conditions, and normalized to 36.1 fb^{-1}. Two sets of multijet MC samples are

used, one is the same PYTHIA8 sample as those used in 0-lepton channel, another one is the SHERPA 2.2.1 multi b-jet sample generated with the NNPDF3.0NNLO PDFs. .

All the different E_T^{cone}, p_T^{cone} and $p_T^{varcone}$ variables with available cone sizes (0.2, 0.3, 0.4) are tested with the cut scan method. The optimal working point is selected based on the given signal events efficiency and multijet events rejection rate for a dedicated variable and cut value. Take the $E_T^{cone0.2}$ in electron sub-channel as an example as shown in Figure 5.20(a), the red histogram represents the $E_T^{cone0.2}$ distribution for the signal events, while the blue histogram represents the same distribution for the multijet events predicted by the PYTHIA8 MC sample. The cut value scan is performed to the distribution from leftmost (-4.5 GeV) to rightmost (10 GeV), with a step of 0.5 GeV, to achieve the corresponding signal efficiency, multijet background efficiency and the value of signal events divided by the square root of sum of signal and multijet background events, as shown in Figure 5.20(b), for example, from the points at $E_T^{cone0.2} = 2\,\text{GeV}$, we can read the message that the cut $E_T^{cone0.2} < 2\,\text{GeV}$ results $\sim 90\%$ signal efficiency, $\sim 10\%$ multijet efficiency, and $\sim 0.25\ S/\sqrt{(S+B)}$ (S represents the signal yields while B represents the multijet yields). Due to the very limited statistics for the multijet MC samples, only basic cuts were applied, with requiring exactly one WH-lepton in the event and $p_T^V > 150\,\text{GeV}$, while the nominal cuts are applied to the signal events. The 2-jets and 3-jets regions are combined together in this study.

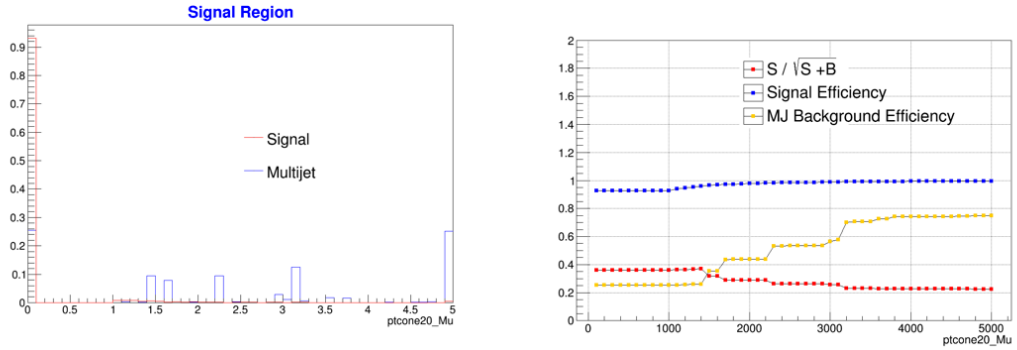
In the electron sub-channel, the $E_T^{cone0.2}$ provides the best discrimination between signal and multijet events. In order to keep at least 95% signal events, the cut at $E_T^{cone0.2} < 2.5\,\text{GeV}$ is selected which provides $\sim 15\%$ multijet events efficiency. In the muon sub-channel, the cut at $p_T^{cone0.2} < 1.25\,\text{GeV}$ is selected as shown in Figure 5.21, the multijet events are also predicted by the PYTHIA8 MC sample. $\sim 95\%$ signal events efficiency and $\sim 25\%$ multijet events efficiency are achieved with this cut.

Table 5.15 shows the detailed numbers of the signal and multijet events efficiencies for the default and new selected working points. The multijet events efficiencies are calculated for both the PYTHIA8 MC and SHERPA 2.2.1 multi b-jet MC samples. Loose-TrackOnly working point is applied on top for the efficiency calculation. For electron sub-channel, the new selected working point, $E_T^{cone0.2} < 3.5\,\text{GeV}$, provides $\sim 30\%$ decreased mutijet events efficiency with only $\sim 3\%$ signal loss, compared to the default FixedCutTight working point. In the muon sub-channel, the default FixedCutTrackOnly working point provides basically no signal loss at the cost of a quite bad multjet events rejection rate, while



(a) (b)

Figure 5.20: $E_T^{cone0.2}$ distributions for siganl (red) and multijet background (blue) events are shown in (a). The cut value scan results are shown in (b), the blue dots represent the signal efficnecy, the yellow dots represent the multijet background efficiency, while the red dots represent the vaule of $S/\sqrt{S+B}$, where S represents the signal yields and B represents the multijet yields.



(a) (b)

Figure 5.21: $p_T^{cone0.2}$ distributions for siganl (red) and multijet background (blue) events are shown in (a). The cut value scan results are shown in (b), the blue dots represent the signal efficnecy, the yellow dots represent the multijet background efficiency, while the red dots represent the vaule of $S/\sqrt{S+B}$, where S represents the signal yields and B represents the multijet yields.

the new selected working point, $p_T^{cone0.2} < 1.25\,\text{GeV}$, reduces the multijet events efficiency to $\sim 30\%$ with only $\sim 5\%$ signal loss.

Table 5.15: Signal events and multijet events efficiencies for the default and new selected isolation working points in both electron and muon sub-channels. LooseTrackOnly working point is applied on top for the efficiency calculation, the efficiency values for multijet events are given for different MC samples with statistical uncertainties.

Electron sub-channel			
Working Points	Signal events efficiency	multijet events efficiency (Pythia8 samples)	multijet events efficiency (Sherpa 2.2.1 multi b-jet samples)
FixCutTight	98%	38% ± 7%	58% ± 4%
$E_T^{cone0.2} < 3.5\,\text{GeV}$	95%	10% ± 4%	11% ± 2%
Muon sub-channel			
FixCutTrackOnly	99%	97% ± 2%	94% ± 2%
$p_T^{Cone0.2} < 1.25\,\text{GeV}$	95%	29% ± 8%	31% ± 5%

As already discussed, only very basic cuts are applied to the multijet MC samples in this study due to the very limited statistics and it could brings some biases for the results. In order to validate this approach, the achieved results are also tested with the data driven template fit method as descripted in Section 5.5.2.2. The result from template fit confirms the conclusion achieved by the MC samples, therefore these two new isolation working points are selected as the new default isolation requirements in 1-lepton channel.

5.5.2.2 Estimation of the multijet background

The real multijet contamination in the 1-lepton signal region cannot be extracted using MC simulations, both because the simulation is very statistically limited and because the simulation is not expected to reproduce fakes correctly. A template fit method is therefore employed to estimate the multijet contribution in the signal region, using data in a multijet enriched control region. The multijet enriched control region is defined using inverted lepton isolation cuts. Table 5.16 summarises both the isolation cuts applied in the signal region and the inverted selection used for the multijet enhanced control region. The transverse W-candidate mass (m_T^W) is chosen as the fit variable since this variable offers the best discrimination between multijet production and electroweak induced processes, while not being excessively sensitive to systematics. The multijet template for this variable is obtained in the inverted isolation region. The contribution from electroweak

background processes in the inverted isolation region is subtracted based on MC predictions. Systematic variations of the MC predictions are later applied as a source of systematic uncertainty. A fit to the m_T^W distribution is then applied in the signal region to extract the normalization factors for the multijet background. The template for the electroweak backgrounds in the signal region is obtained directly from MC predictions. Separate templates for the multijet contributions are obtained depending on lepton flavor (e/μ), jet multiplicity (2/3-jet regions) and p_T^V category (high and medium p_T^V regions). For each of these eight signal regions a corresponding multijet control region is thus defined. In the medium p_T^V region, dut to the much higher multijet contribution compared to the high p_T^V region, an additional $m_T^W > 20\,\text{GeV}$ is applied, in addition, single muon trigger is used in muon-sub channel since the E_T^{miss} trigger can not be used due to the low p_T^W threshold.

Table 5.16: Summary of differences in lepton isolation between the isolated and inverted isolation regions used for the template method. In each region, the events are requested to pass both of the two isolation criteria listed in the table.

	Isolated Region	Inverted Isolation Region
Electron	LooseTrackOnly	LooseTrackOnly
	$E_T^{cone0.2} < 3.5\,\text{GeV}$	$E_T^{cone0.2} > 3.5\,\text{GeV}$
Muon	LooseTrackOnly	LooseTrackOnly
	$p_T^{cone0.2} < 1.25\,\text{GeV}$	$p_T^{cone0.2} > 1.25\,\text{GeV}$

The statistics in the multijet enhanced control region is limited, so only 1 b-tag is required in the control region instead of requiring 2 b-tags as in the signal region, in order to reduce the impact of statistical fluctuations when deriving the template. The plots in Figure 5.22 and Figure 5.23 show the m_T^W distributions for the data and electroweak processes in the inverted isolation e/μ, 2/3-jet regions with requiring exactly 1 b-tag, in high and medium p_T^V regions, respectively. The approximate purity of the multijet events in each multijet enriched control region, calculated by using number of data events minus the number of electroweak background events and then divided by the number of data events, are summarized in Table 5.17. The purity in electron sub-channel floats from 50% to 70% in different categories, while in the muon sub-channel, the purity is a bit worse and floats from 30% to 65%. As a reference, Figure 5.24 and Figure 5.25 show also the m_T^W distributions for the data and electroweak processes in the inverted isolation 2

b-tags regions. As can be seen in the plots, the statistics is quite limited in such regions and the electroweak processes contamination is much larger than that in the 1 b-tag region.

Table 5.17: Summary of the approximate purity of the multijet events in each multijet enriched control region. The purity is calculated by using number of data events minus the number of electroweak background events, divided by the number of data events.

Electron sub-channel		
	High p_T^V	Medium p_T^V
1 b-tag 2-jet	70%	60%
1 b-tag 3-jet	55%	50%
Muon sub-channel		
1 b-tag 2-jet	40%	65%
1 b-tag 3-jet	30%	55%

The $t\bar{t}$ and W+jets processes are dominant in the signal region, and their normalization can have a significant impact on the mutlijet estimate. Their normalization is therefore extracted simultaneously to the multijet estimate itself. While the m_T^W variable provides discrimination mainly between processes without and with a W boson, the distributions of m_T^W for the $t\bar{t}$ and W+jets processes are not identical, since di-leptonic $t\bar{t}$ events induce a tail at high values of m_T^W. In order to avoid a bias onto the multijet estimate, separate normalization factors are extracted for the Top ($t\bar{t}$ +single top) and W+jet contributions. However, the m_T^W distribution alone only provides marginal separation between these two background components, so to determine their respective contribution a simultaneous fit is applied to the signal region and the W+HF enhanced region (the same used also in the main Higgs boson signal extraction fit). Since the relative W+jet / Top purity is very different in these two regions, a simultaneous fit to the two regions allows the extraction of the two separate normalizations with decent precision. The m_T^W distribution is then used in the fit basically only to disentangle the multijet contribution from both the Top and W+jets backgrounds. Due to the limited statistics, the m_T^W distribution is exploited in the signal region, while only the overall yield is used in the W+HF control region. To increase the statistical precision in the determination of the Top and W+jet normalization factors further, the fit is also applied simultaneously in the electron and muon channel, extracting

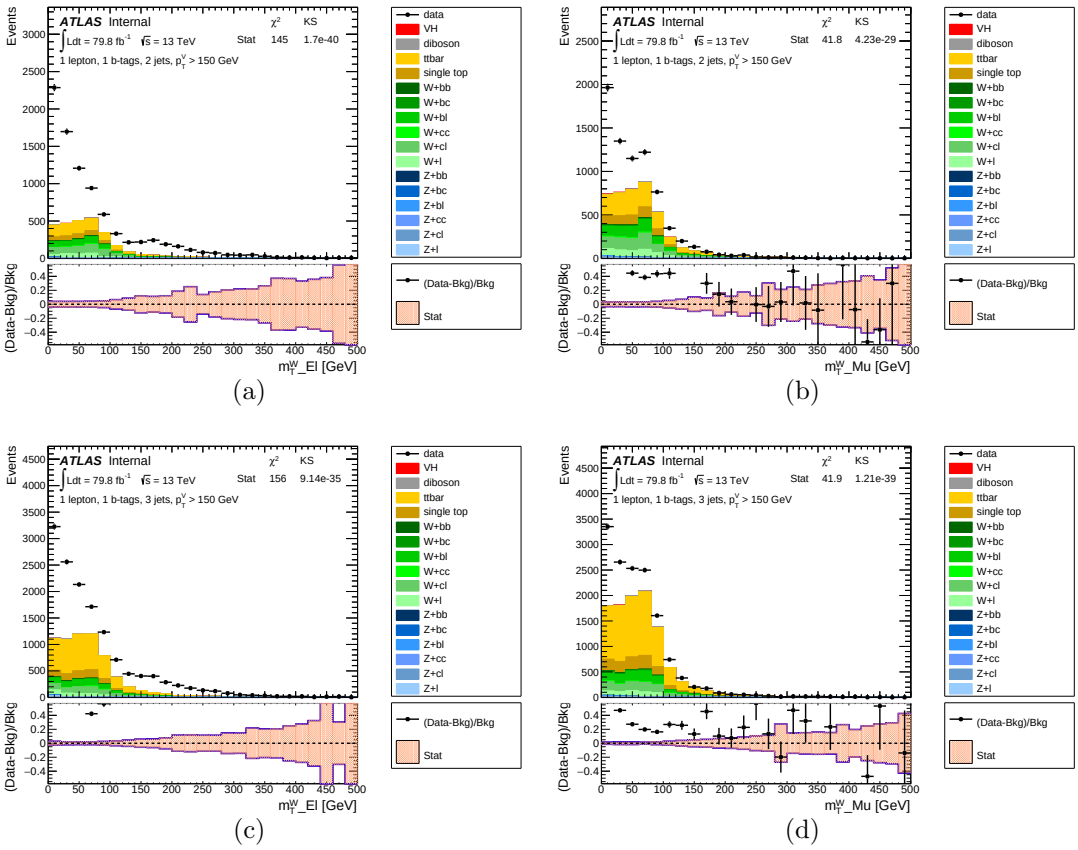


Figure 5.22: The m_T^W distribution in the inverted isolation 1-lepton $p_T^W > 150$ GeV region, requiring exactly 1 b-tag jet in electron sub-channel 2-jets region(a), muon sub-channel 2-jets region(b), electron sub-channel 3-jets region(c), and muon sub-channel channel 3-jets region(d).

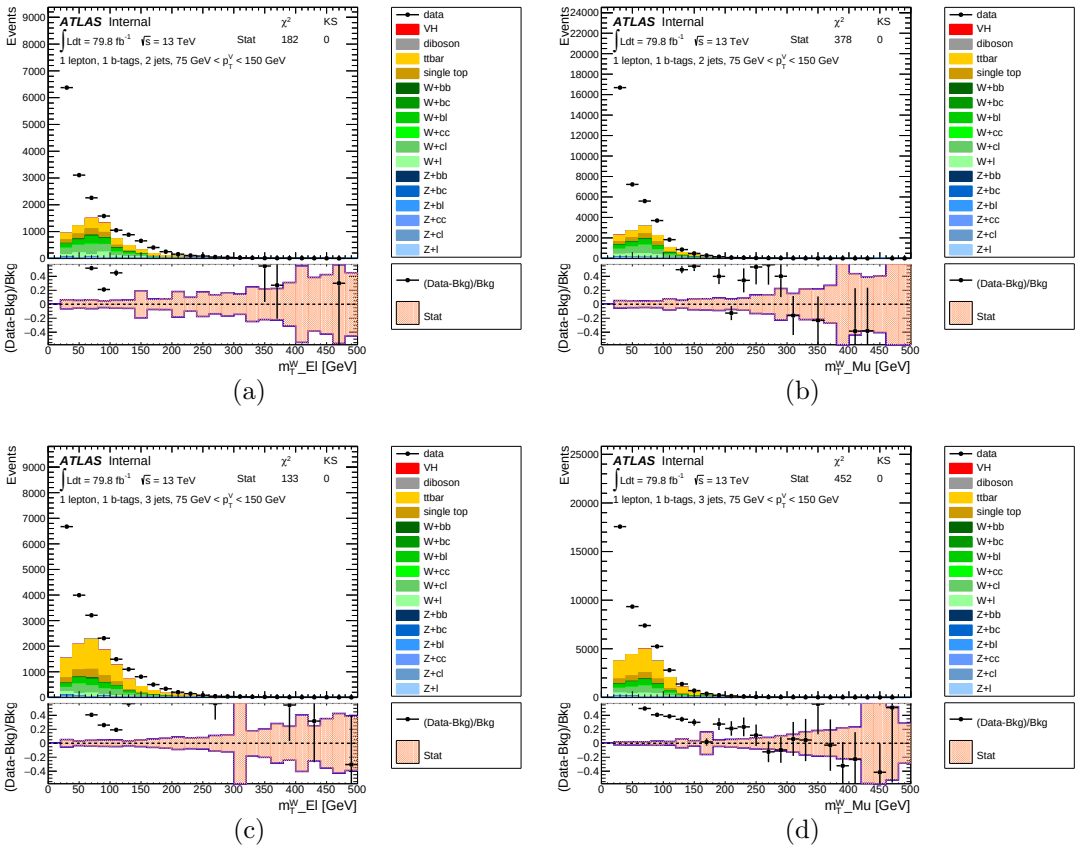


Figure 5.23: The m_T^W distribution in the inverted isolation 1-lepton 75 GeV < p_T^W < 150 GeV region, requiring exactly 1 b-tag jet in electron sub-channel 2-jets region(a), muon sub-channel 2-jets region(b), electron sub-channel 3-jets region(c), and muon sub-channel channel 3-jets region(d).

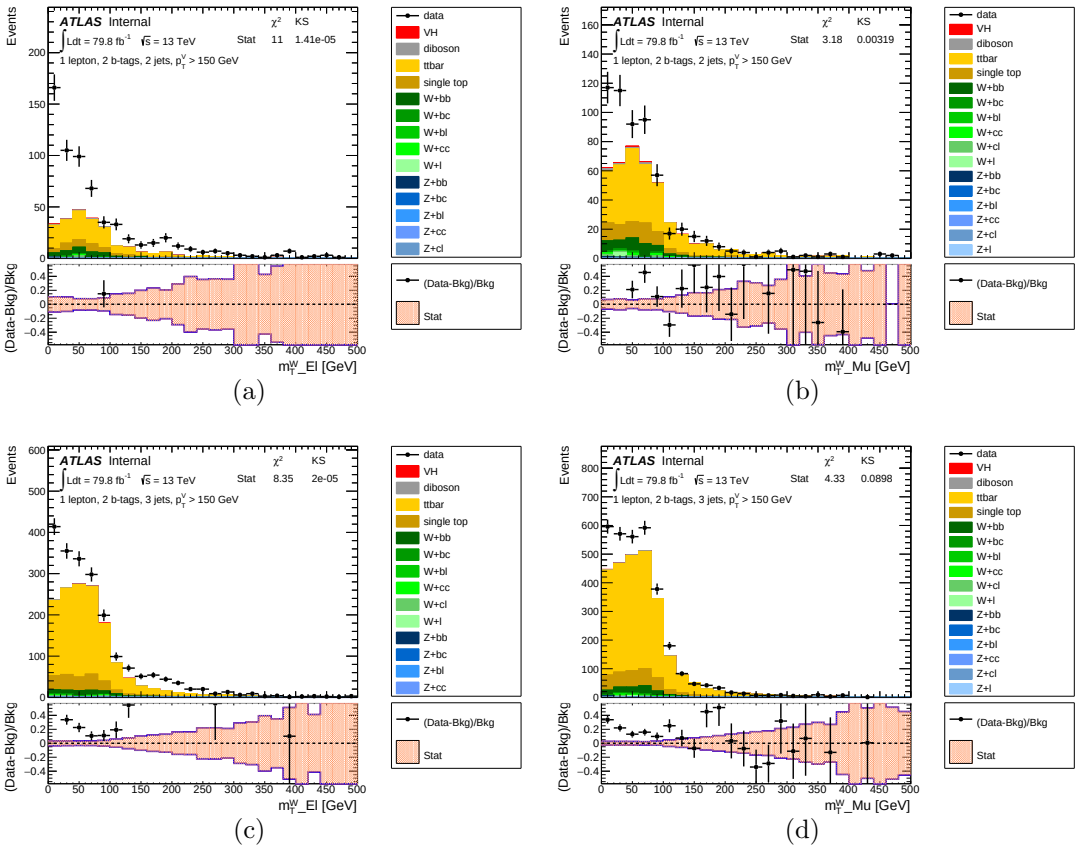


Figure 5.24: The m_T^W distribution in the inverted isolation 1-lepton $p_T^W > 150$ GeV region, requiring exactly 2 b-tag jets in electron sub-channel 2-jets region(a), muon sub-channel 2-jets region(b), electron sub-channel 3-jets region(c), and muon sub-channel channel 3-jets region(d).

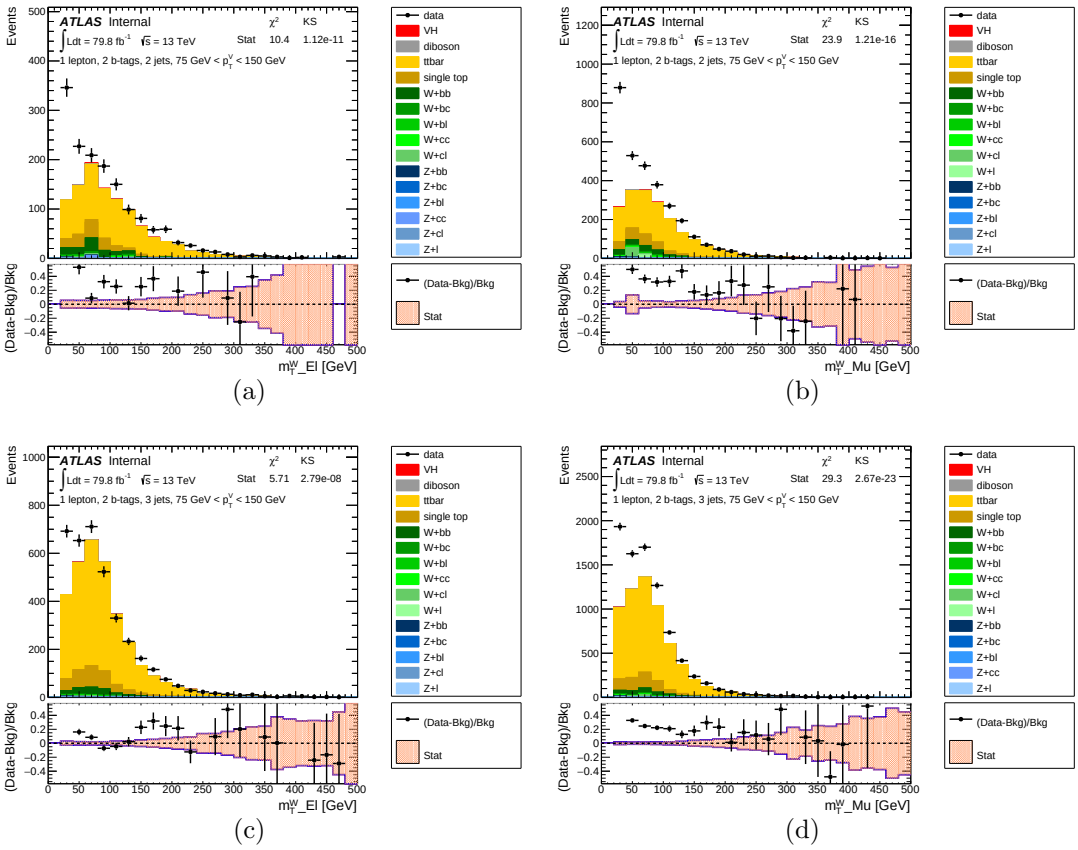


Figure 5.25: The m_T^W distribution in the inverted isolation 1-lepton 75 GeV < p_T^W < 150 GeV region, requiring exactly 2 b-tag jet in electron sub-channel 2-jets region(a), muon sub-channel 2-jets region(b), electron sub-channel 3-jets region(c), and muon sub-channel channel 3-jets region(d).

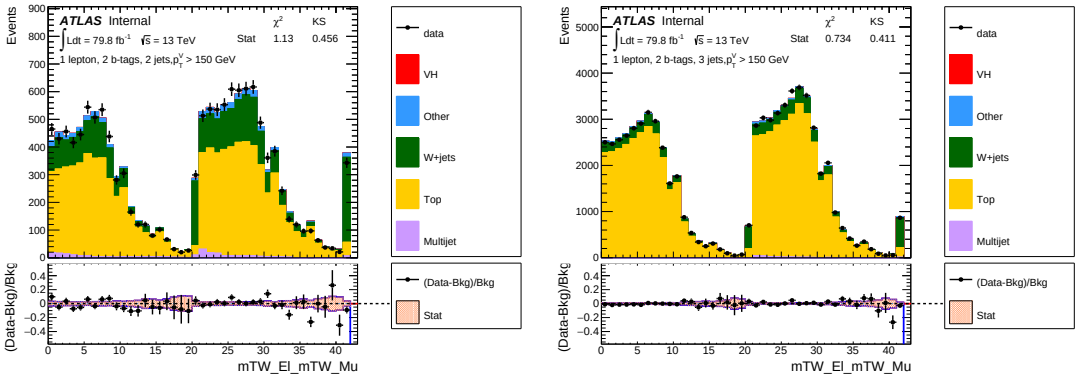
simultaneously the normalizations for the electron multijet, muon multijet, Top and W+jets components. The normalization factors extracted in the template fit for the Top and W+jets processes can be significantly different from unity: the difference from unity is later considered as a source of systematic uncertainty for the electroweak background subtraction procedure in the inverted isolation region.

Technically, the multijet fit is implemented as a template fit to a single region, with distributions/yields from different regions merged to adjacent intervals/bins of a single final distribution. The overall yield of the W+HF enhanced region is being represented by an additional bin at the extreme right of the m_T^W distribution. The electron channel is then put on the left in the final fit distribution, while the muon channel is put on the right. The binning of the m_T^W distribution is optimised in such a way to yield a roughly constant MC statistical uncertainty in each bin. Separate templates are used for the electron multijet, muon multijet, Top and W+jets components, and the normalization factor extracted for each contribution is presented in Table 5.18. Post-fit plots for the distribution exploited in the fit are shown in Figure 5.26 and Figure 5.27, for high and medium p_T^V region, respectively. Apart from the m_T^W distribution which is directly used in the template fit, Figure 5.28 to Figure 5.35 also show some the other post-fit plots for the distributions especially sensitive to the shape and normalization of the multijet background in both 2- and 3-jets regions, electron and muon sub-channels, and high and medium p_T^V regions. In these distributions, the normalization is fixed to the result derived from the template fit. In general, good agreement between data and sum of electroweak backgrounds from MC prediction and multijet background derived from template fit can be observed. In high p_T^V region, the multijet contribution in the 2-jets region is found to be 1.91% (2.76%) in electron (muon) sub-channel, while in the 3-jets region it is found to be 0.15% (0.43%). In the medium p_T^V region, the multijet contribution in the 2-jets region is found to be 3.57% (2.76%) in electron (muon) sub-channel, while in the 3-jets region it is found to be 0.85% (2.14%). The multijet fractions are summarized in Table 5.20 and Table 5.21 for high p_T^V and medium p_T^V regions, respectively.

To provide a better appreciation of the quality of the modeling of the multijet background, an "extended" medium p_T^V region is used, where the m_T^W cut has been removed, thus greatly enhancing the MJ contribution and the template fit has been re-performed without the m_T^W cut. Examples of distributions especially sensitive to the shape and normalization of the multijet background are shown in Fifure 5.36 to Fifure 5.39 for 2-jets and 3-jets regions and electron and muon subchannels. The great agreement between data and sum of electroweak backgrounds

Table 5.18: Summary of normalisation scale factors for Top ($t\bar{t}$ + single top) and W+jets derived from the template fit.

Region	Top ($t\bar{t}$ + single top)	W+jets
high p_T^V 2-tag, 2-jet	1.02 ± 0.02	1.27 ± 0.06
high p_T^V 2-tag, 3-jet	0.99 ± 0.006	1.13 ± 0.04
medium p_T^V 2-tag, 2-jet	1.05 ± 0.009	1.49 ± 0.05
medium p_T^V 2-tag, 3-jet	1.07 ± 0.004	1.10 ± 0.04



(a) (b)

Figure 5.26: The m_T^W distribution in the 1-lepton $p_T^W > 150\ GeV$ signal region, requiring exactly 2 b-tag jets in 2-jets region(a) and 3-jets region(b). Top ($t\bar{t}$ + single top) and W+jets normalisation factors derived from template fit are applied. Bins 1-21 correspond to the electron sub-channel, bins 22 to 42 correspond to the muon sub-channel, and bins 21 and 42 represent the $W + HF$ control regions in electron an and muon sub-channel, respectively.

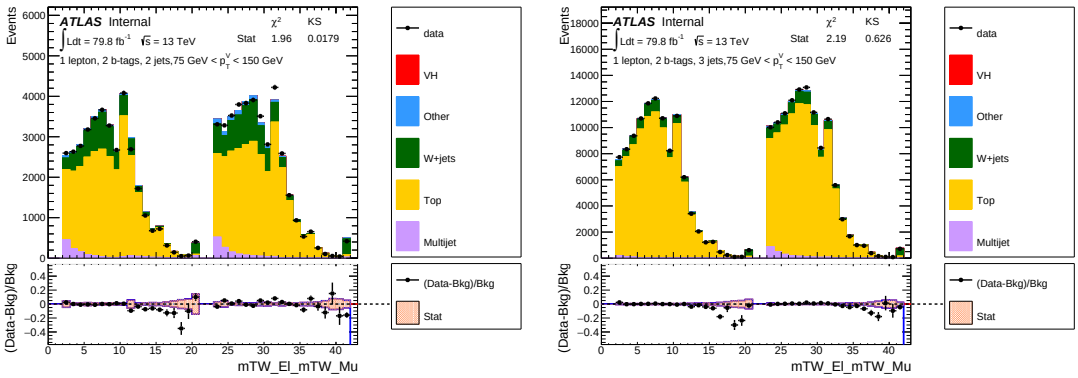
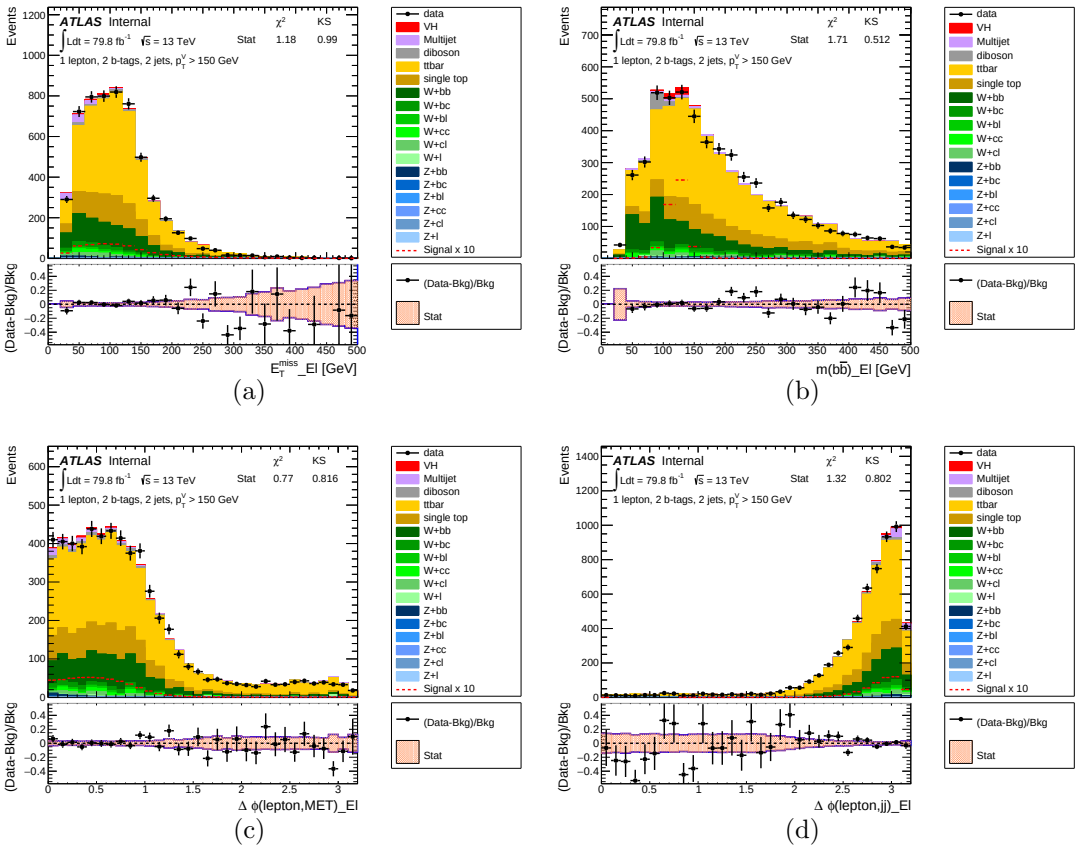


Figure 5.27: The m_T^W distribution in the 1-lepton 75 $GeV < p_T^W < 150$ GeV signal region, requiring exactly 2 b-tag jets in 2-jets region(a) and 3-jets region(b). Top ($t\bar{t}$ + single top) and W+jets normalisation factors derived from template fit are applied. Bins 1-21 correspond to the e only channel, bins 22 to 42 correspond to the μ only channel, and bins 21 and 42 represent the $W + HF$ control regions in electron an and muon sub-channel, respectively.

and multijet backgrounds in such multijet enhanced region indicates the strong robustness of this data-driven method. Also it can be seen that the $m_T^W > 20\,\text{GeV}$ can greatly reduce the multijet contribution in the medium p_T^V region.



(a) (b) (c) (d)

Figure 5.28: The distributions, for the 2-tag 2-jet $p_T^W > 150\ GeV$ category in electron sub-channel signal region, of (a) E_T^{miss} (b) $m_{b\bar{b}}$ (c) $\Delta\phi(l, E_T^{miss})$ (distance in ϕ between E_T^{miss} and lepton) and (d) $\Delta\phi(l, b\bar{b})$ (distance in ϕ between lepton and dijet system) are shown.

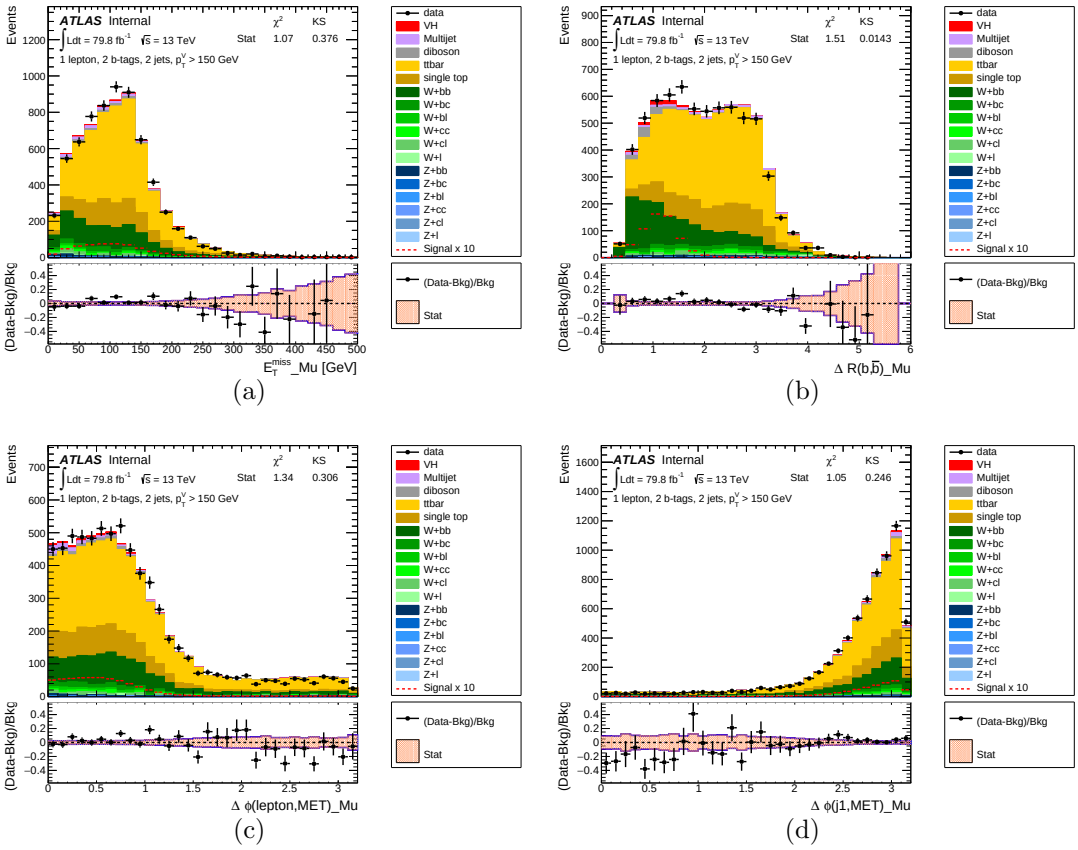


Figure 5.29: The distributions, for the 2-tag 2-jet $p_T^W > 150\ GeV$ category in muon sub-channel signal region, of (a) E_T^{miss} (b) $\Delta R b\bar{b}$ (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(b_1, E_T^{miss})$ (distance in ϕ between leading b-tagged jet and E_T^{miss}) are shown.

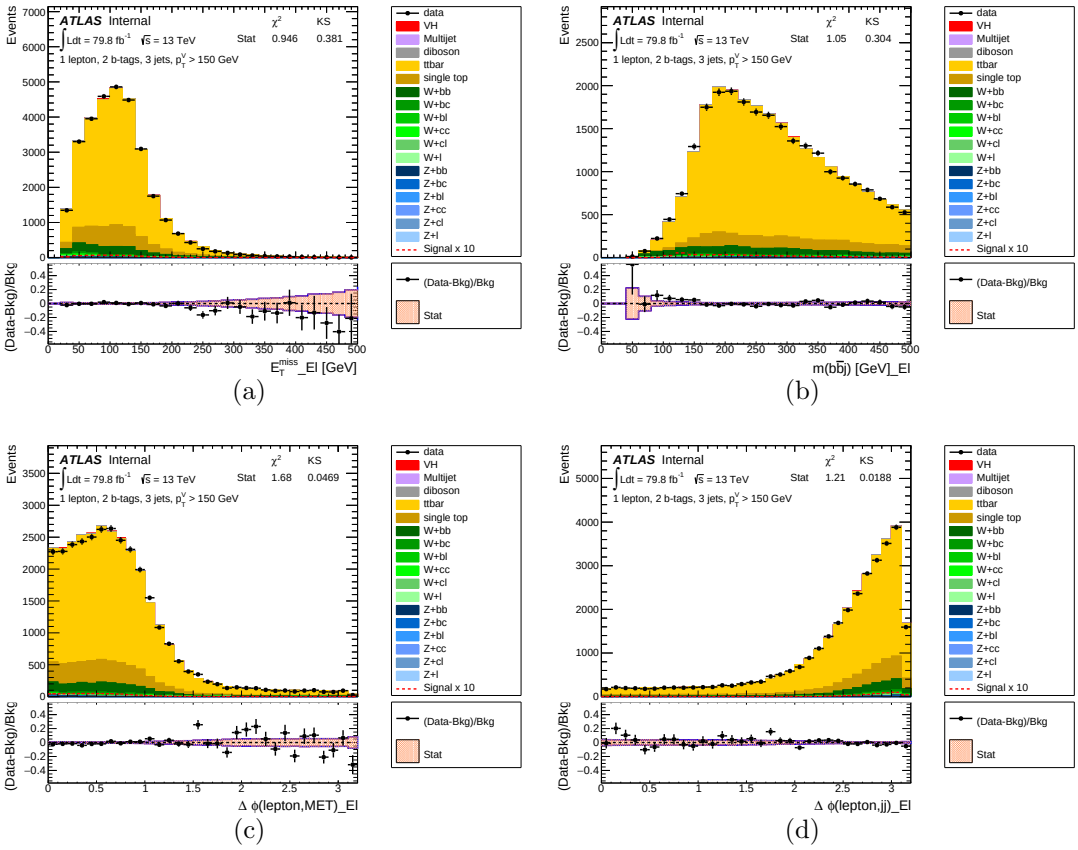


Figure 5.30: The distributions, for the 2-tag 3-jet $p_T^W > 150\ GeV$ category in electron sub-channel signal region, of (a) E_T^{miss} (b) m_{bbj} (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(l, b\bar{b})$ are shown.

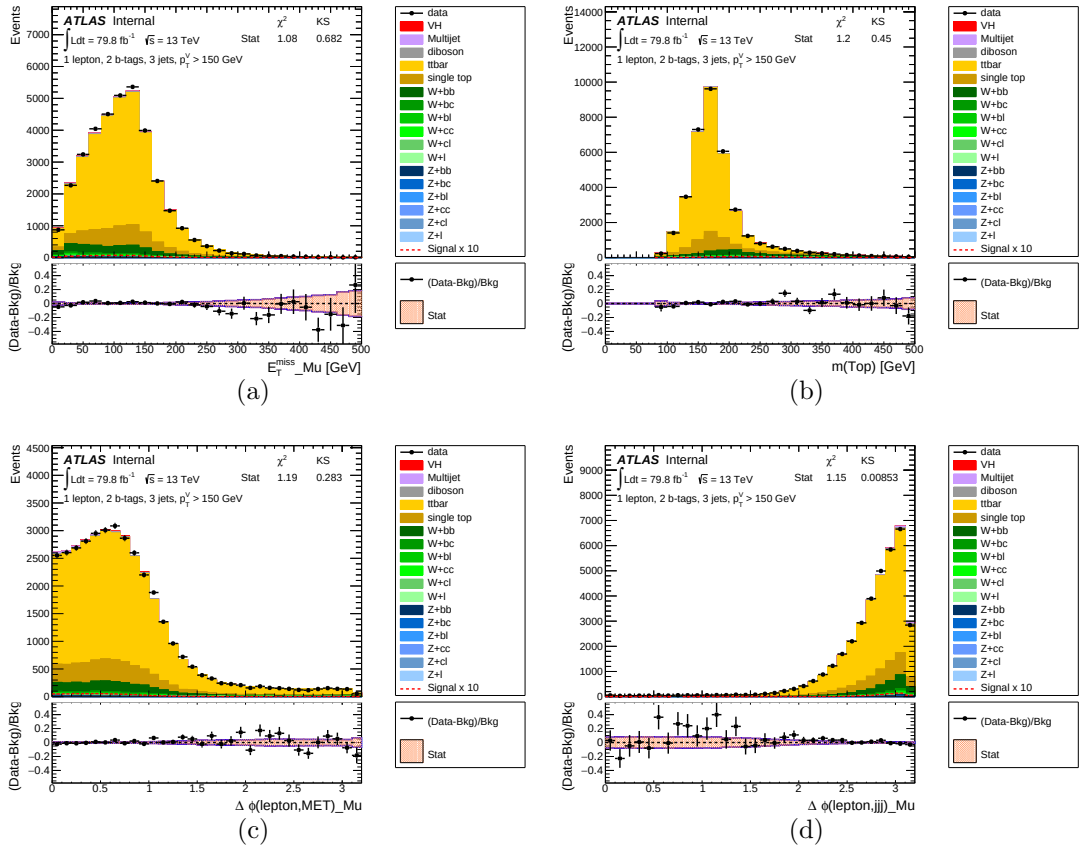


Figure 5.31: The distributions, for the 2-tag 3-jet $p_T^W > 150\ GeV$ category in muon sub-channel signal region, of (a) E_T^{miss} (b) m_{top} (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(l, bbj)$ (distance in ϕ between lepton and bbj system) are shown.

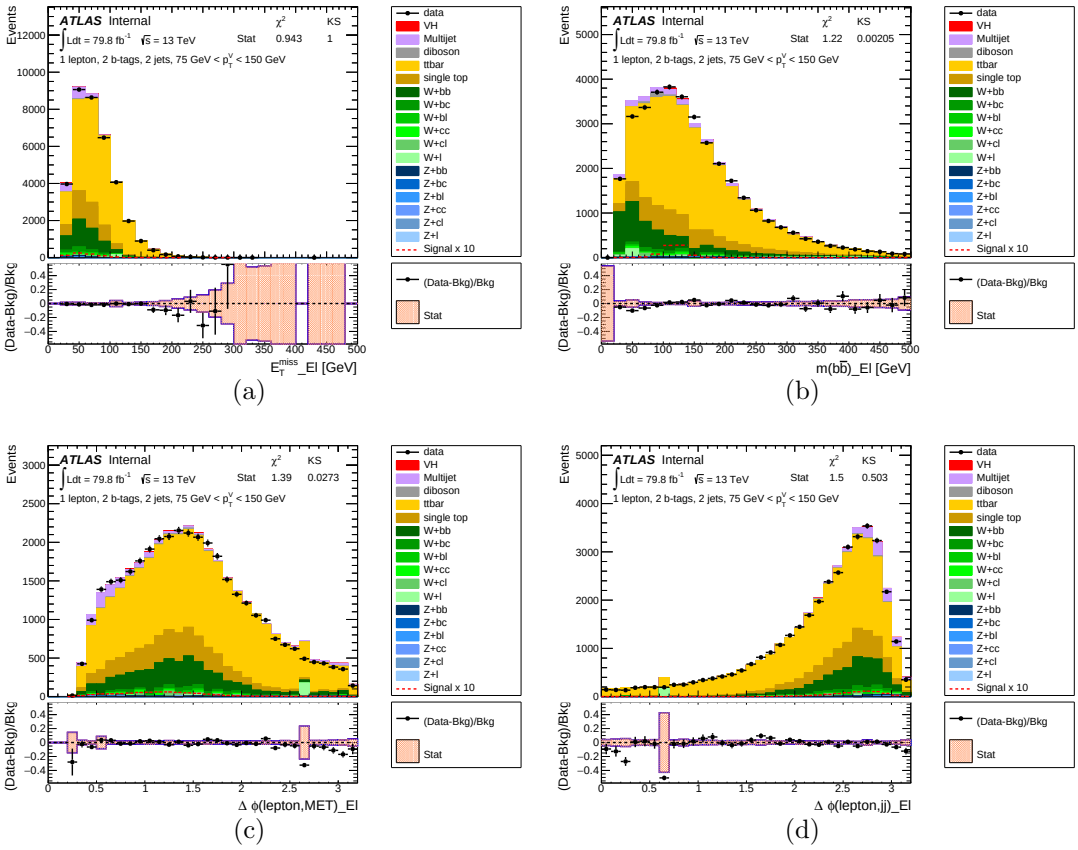


Figure 5.32: The distributions, for the 2-tag 2-jet 75 $GeV < p_T^W < 150\ GeV$ category in electron sub-channel signal region, of (a) E_T^{miss} (b) $m_{b\bar{b}}$ (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(l, b\bar{b})$ are shown.

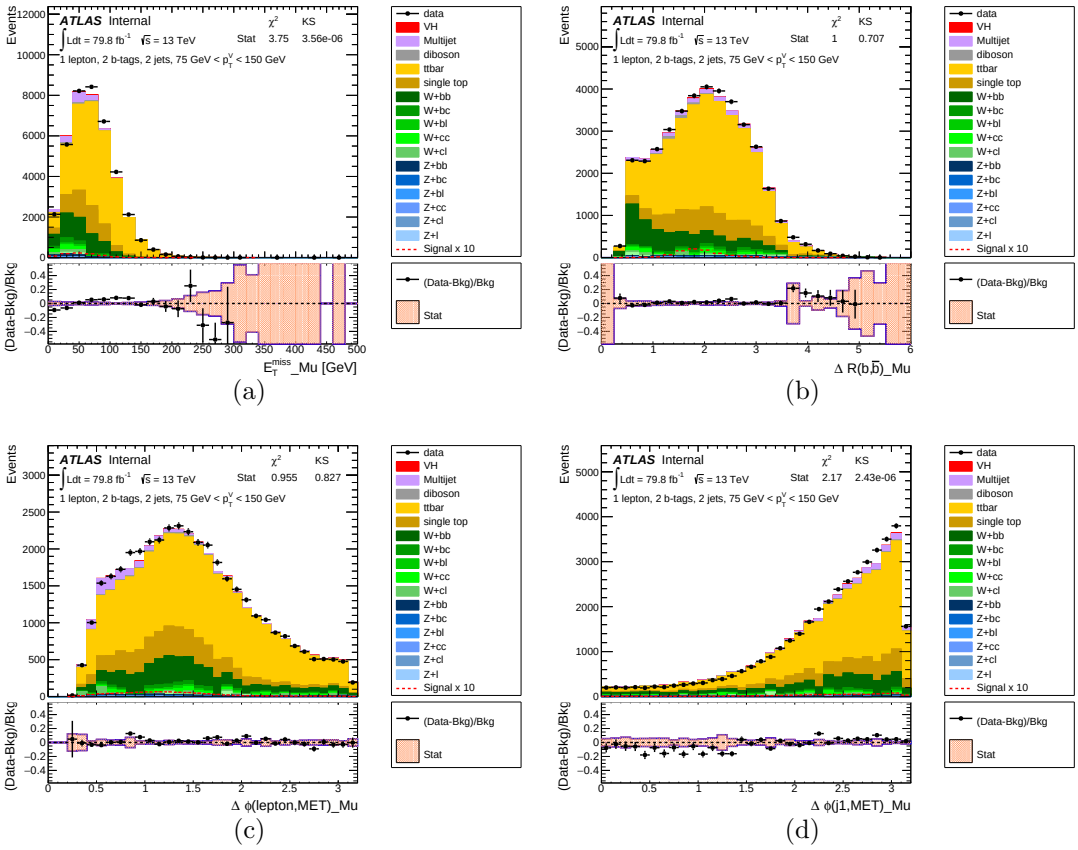


Figure 5.33: The distributions, for the 2-tag 2-jet 75 $GeV < p_T^W < 150$ GeV category in muon sub-channel signal region, of (a) E_T^{miss} (b) $\Delta R(b\bar{b})$ (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(b_1, E_T^{miss})$ are shown.

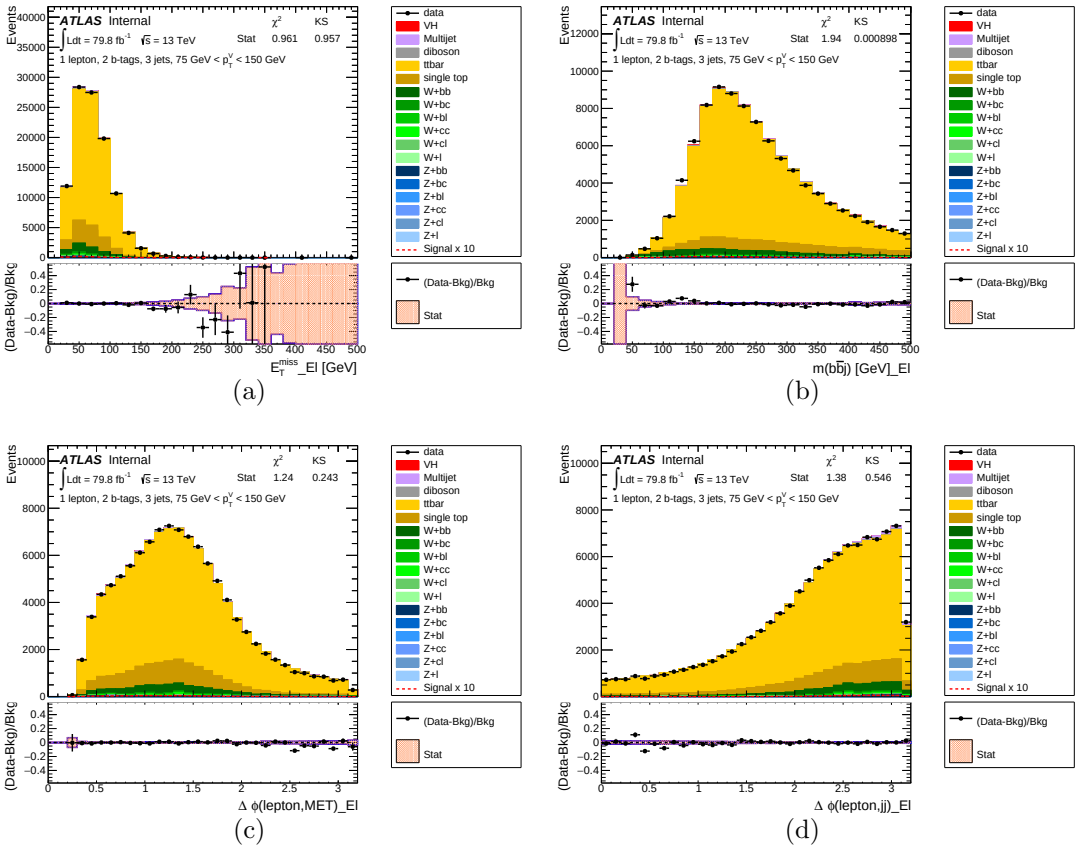


Figure 5.34: The distributions, for the 2-tag 3-jet 75 $GeV < p_T^W < 150$ GeV category in electron sub-channel signal region, of (a) E_T^{miss} (b) m_{bbj} (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(l, b\bar{b})$ are shown.

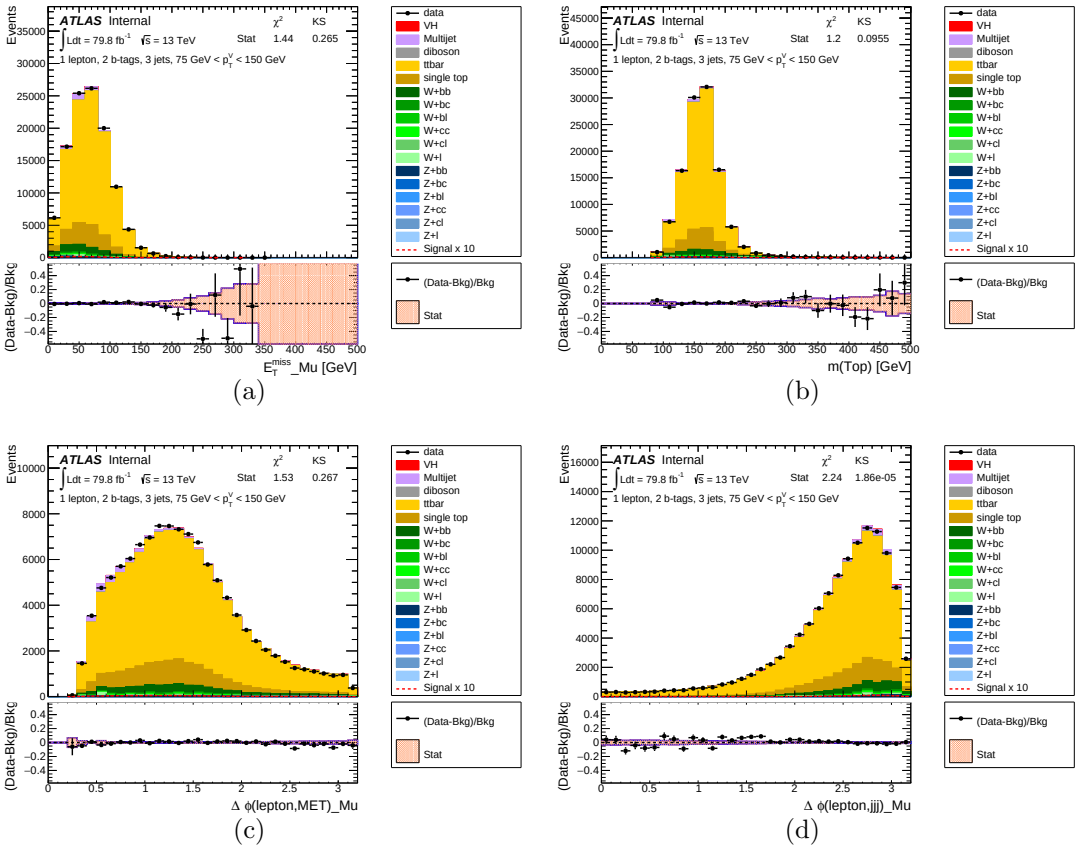


Figure 5.35: The distributions, for the 2-tag 3-jet 75 $GeV < p_T^W < 150$ GeV category in muon sub-channel signal region, of (a) E_T^{miss} (b) m_{top} (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(l, bbj)$ are shown.

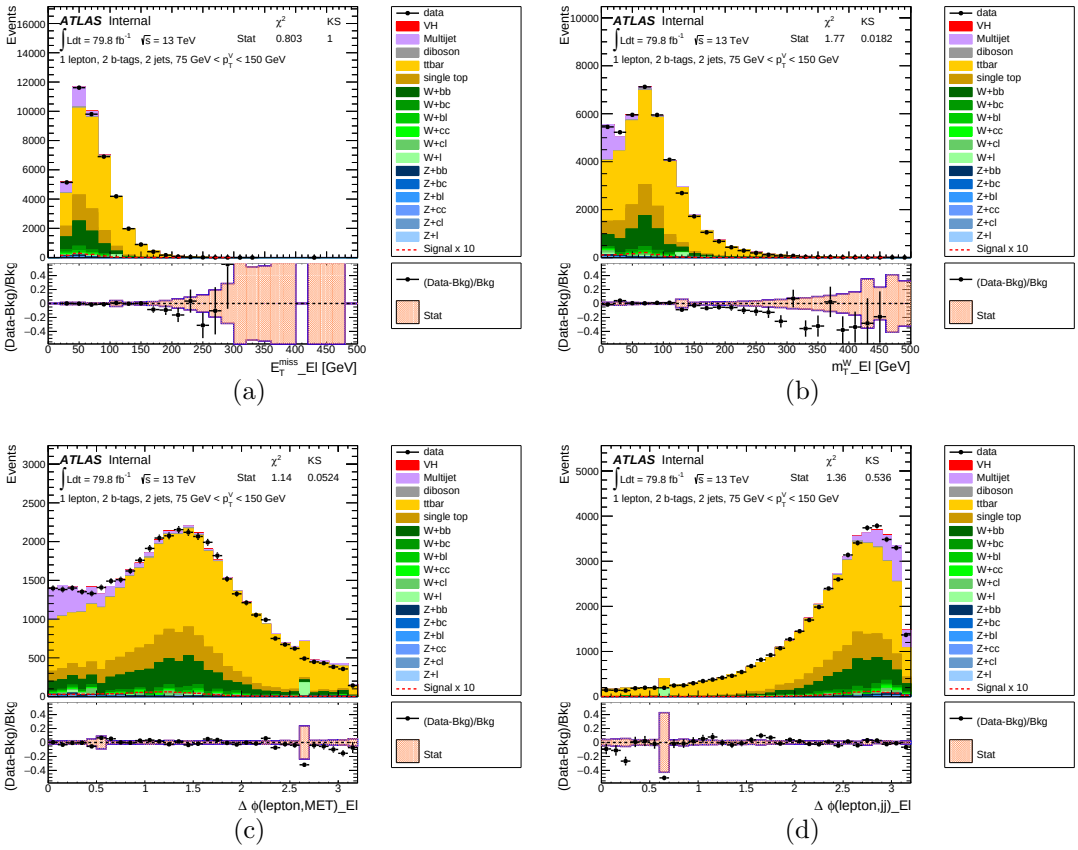


Figure 5.36: The distributions, for the 2-tag 2-jet 75 $GeV < p_T^W < 150$ GeV category in electron sub-channel signal region without $m_T^W > 20\,\text{GeV}$ cut applied, of (a) E_T^{miss} (b) m_T^W (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(l, b\bar{b})$ are shown.

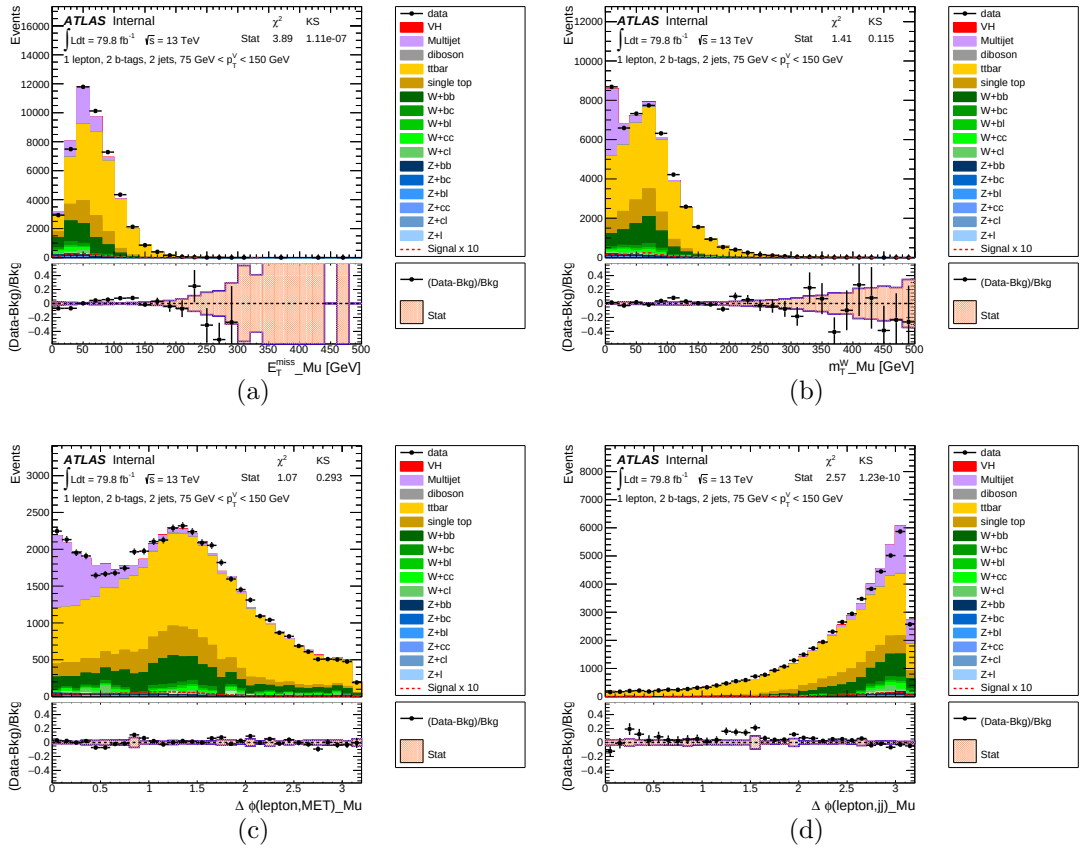


Figure 5.37: The distributions, for the 2-tag 2-jet 75 $GeV < p_T^W < 150$ GeV category in muon sub-channel signal region without $m_T^W > 20\,\text{GeV}$ cut applied, of (a) E_T^{miss} (b) m_T^W (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(l, b\bar{b})$ are shown.

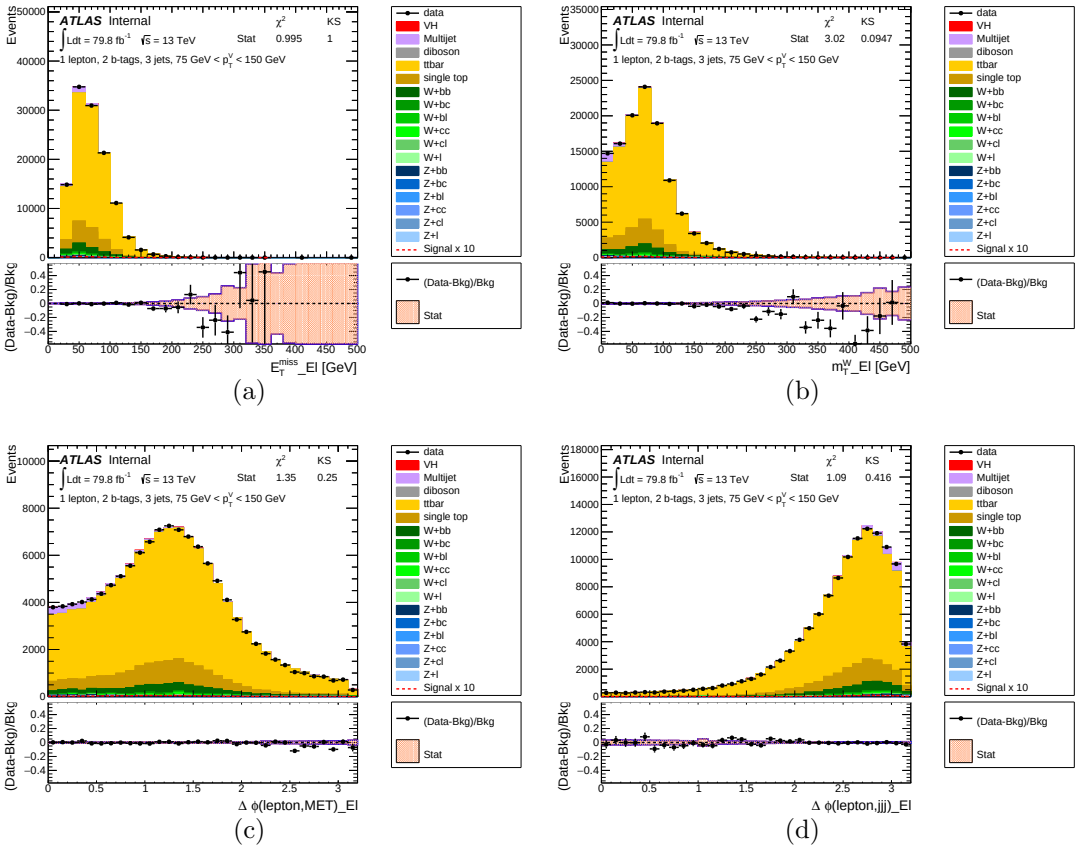


Figure 5.38: The distributions, for the 2-tag 3-jet 75 $GeV < p_T^W < 150$ GeV category in electron sub-channel signal region without $m_T^W > 20\,\text{GeV}$ cut applied, of (a) E_T^{miss} (b) m_T^W (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(l, bbj)$ are shown.

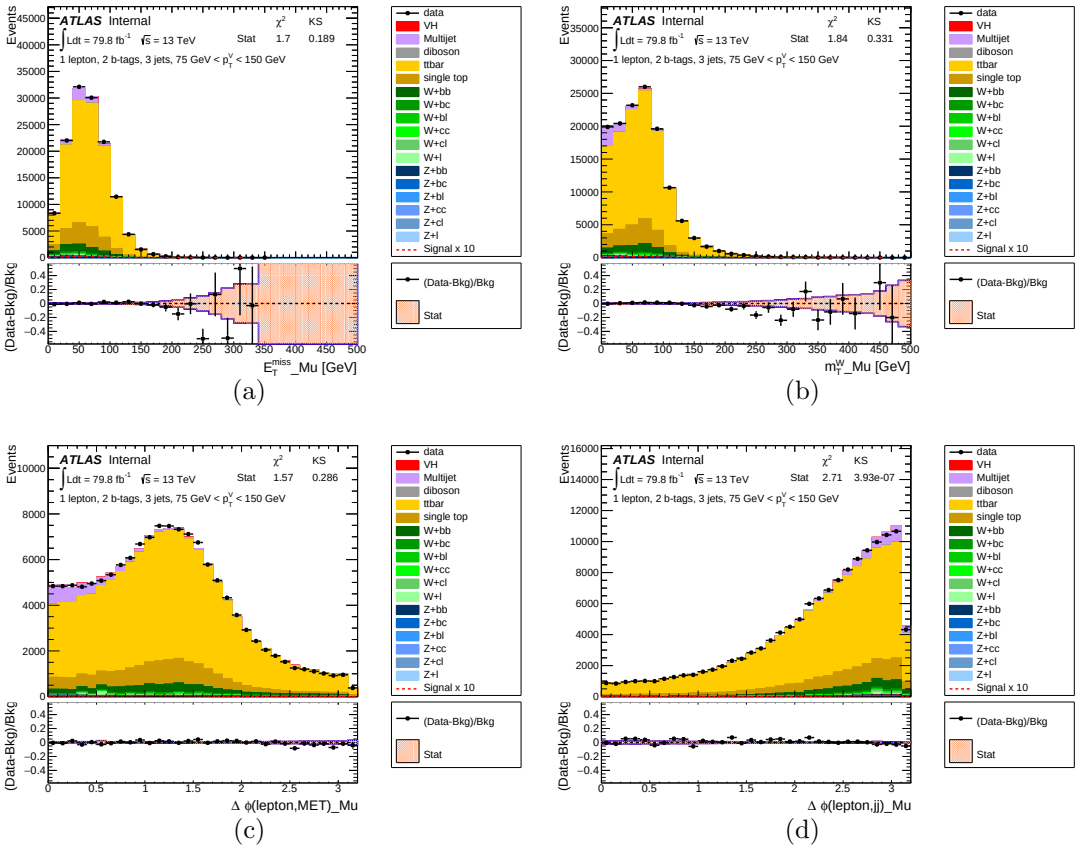


Figure 5.39: The distributions, for the 2-tag 3-jet 75 $GeV < p_T^W < 150$ GeV category in muon sub-channel signal region without $m_T^W > 20$ GeV cut applied, of (a) E_T^{miss} (b) m_T^W (c) $\Delta\phi(l, E_T^{miss})$ and (d) $\Delta\phi(l, bbj)$ are shown.

5.5.2.3 Systematics uncertainties

Systematic uncertainties can have impacts on the multijet estimation in two ways : either changing the fit distribution used in the template fit, therefore impacting the extracted multijet yields, or changing the multijet BDT distributions used in the global likelihood fit. A number of sources of systematic uncertainty are considered, and uncorrelated between electron and muon sub-channels, between 2- and 3- jets regions, and between high p_T^V and medium p_T^V categories. The respective variations are added in quadrature for the normalization uncerainties, or considered as separate shape uncertainties. In this section, the systematic uncertainties that impact the shape will be discussed first, since most of these are also considered for the normalization uncertainties.

Shape Uncertainties

In order to evaluate the shape uncertainty of the MJ background estimate, a number of shape systematic uncertainties are considered:

- The impact of the choice of lepton trigger on the MJ estimate is evaluated, as this may introduce a bias in the inverted isolation region. This systematic effect only on the electron sub-channel channel and medium p_T^V muon sub-channel, since in the high p_T^V muon sub-channel the E_T^{miss} trigger is used rather than the single muon trigger. Instead of using the combination of triggers, listed in Sec 5.3.3, simply the lowest p_T trigger is used. This corresponds to the trigger selections for each data period listed in Table 5.19.

Table 5.19: Reduced triggers used to evaluate possible trigger bias in inverted isolation region.

Dataset	Single electron trigger	Single muon trigger
2015	`e24_lhmedium_L1EM20VH`	`mu20_iloose_L1MU15`
2016-2017	`e26_lhtight_nod0_ivarloose`	`mu26_ivarmedium`

- An evaluation of the uncertainty introduced by the extrapolation from the full inverted isolation region to the signal region is considered. A reduced inverted-isolation region is defined, with additional isolation cuts applied to the inverted isolation region defined in Table 5.16. In the electron sub-channel, this is defined with additionally requiring $E_T^{cone0.2} < 12\,\text{GeV}$ in high p_T^V region and $< 6\,\text{GeV}$ in medium p_T^V region, and in the muon sub-channel,

$p_T^{cone0.2} < 2.9\,\text{GeV}$ in high p_T^V region and $< 2.1\,\text{GeV}$ in medium p_T^V region. The additional cuts are optimized to keep about half of data events in the full inverted regions for both electron and muon sub-channels.

- The impact of using the normalization factors extracted in the template fit for the Top and W+jets processes in the electroweak background subtraction procedure in the inverted isolation region is evaluated. The nominal MJ template shape is evaluated without applying the normalization factors, for this systematic, the template shape is evaluated with applying the normalization factors and and the difference in shape taken as the systematic uncertainty.

These systematic uncertainties are implemented as shape only systematic uncertainties by normalizing the variation to the nominal MJ yield. Plots in Figure 5.40 to Figure 5.41 show the shape comparison for the nominal BDT and the main shape systematics variations in the high and medium p_T^V region for both electron and muon sub-channels.

Normalisation Uncertainty

The sources of systematic uncertainty that have an impact on the BDT shape are also considered to derive an uncertainty on the estimated multijet normalization. For each individual contribution, the positive and negative differences from the fitted nominal multijet yield are separately added in quadrature, and the results are added in quadrature to the statistical uncertainty of the nominal fit to give the overall normalization uncertainty, separately in the high and medium p_T^V regions, in the 2- and 3- jets regions, and in the the electron and muon sub-channels. The negative uncertainties are restricted to be at most equal to the nominal values. In cases where the fitted nominal multijet yield is equal to zero, half of the positive error is used in the global fit as nominal value as well as symmetric error. In addition to the sources considered for the shape uncertainties, a few more are considered exclusively for the normalization uncertainty:

- In the high p_T^V region, including the $E_T^{miss} < 30\,\text{GeV}$ region in the template fit (electron sub-channel only), which induces a significant change to the m_T^W distribution both for the multijet component derived from the inverted isolation region in data and for the electroweak background components estimated using MC simulations. In the medium p_T^V region, including the $m_T^W < 20\,\text{GeV}$ region in he template fit, to probe the potential mismodelling due to the additional m_T^W cut.

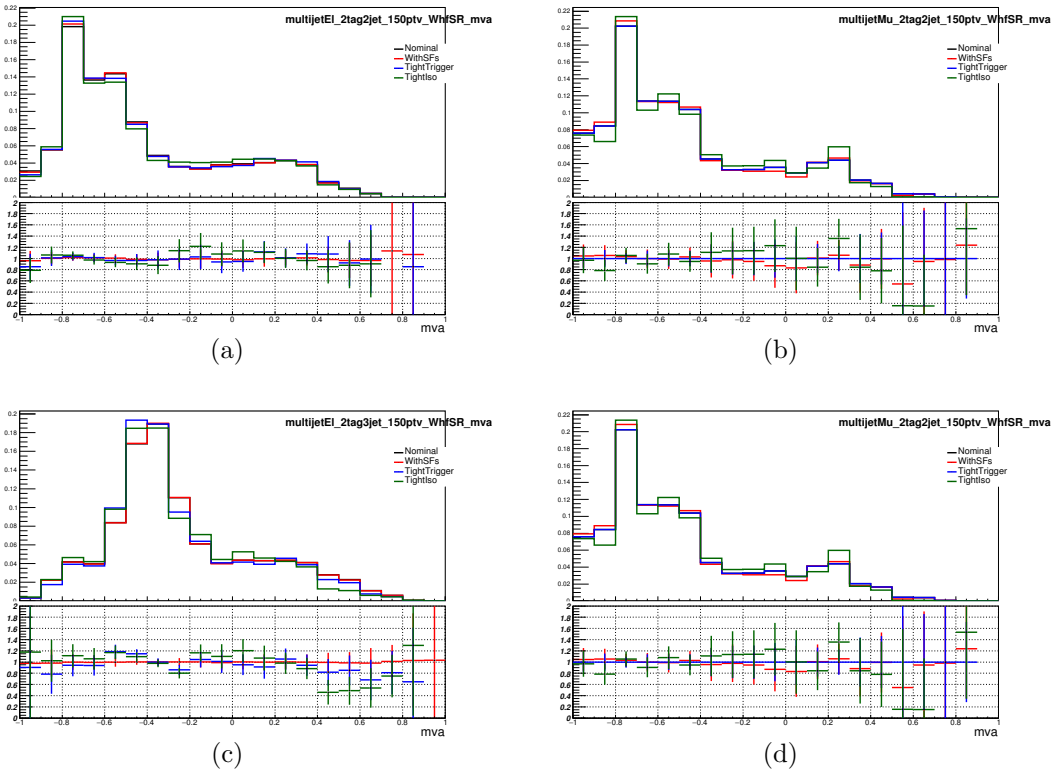


Figure 5.40: The multijet BDT shape comparison for the nominal (in black) and shape variations in the high p_T^V region, electron sub-channel 2-jets region (a), muon sub-channel 2-jets region (b), electron sub-channel 3-jets region (c), and muon sub-channel 3-jets region (d). The green histograms indicate the impact of using the reduced inverted isolation region, the red histograms indicate the impact of using the Top and W+jets normalization factors in the inverted isolation region, and the histograms in blue indicate the impact of using the lowest lepton p_T trigger.

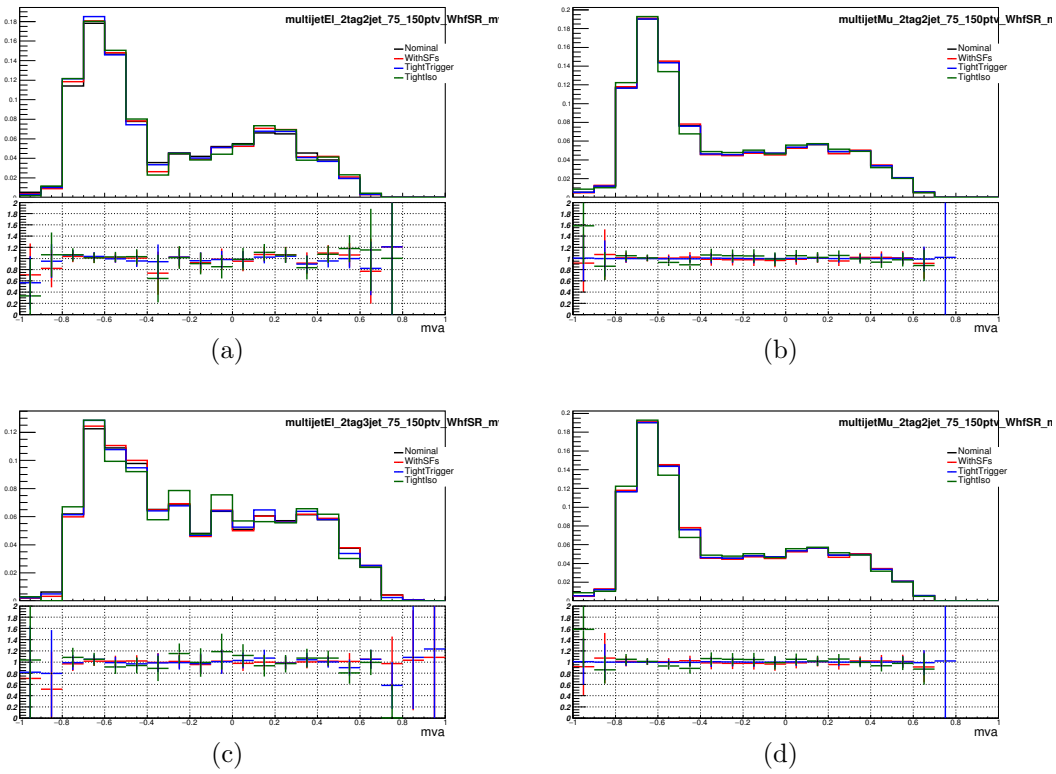


Figure 5.41: The multijet BDT shape comparison for the nominal (in black) and shape variations in the medium p_T^V region, electron sub-channel 2-jets region (a), muon sub-channel 2-jets region (b), electron sub-channel 3-jets region (c), and muon sub-channel 3-jets region (d). The green histograms indicate the impact of using the reduced inverted isolation region, the red histograms indicate the impact of using the Top and W+jets normalization factors in the inverted isolation region, and the histograms in blue indicate the impact of using the lowest lepton p_T trigger.

- Using an alternative distribution instead of m_T^W in the template fit. In the 2-jets category, $\Delta\phi(lepton, b\bar{b})$ is selected and in 3 jets category, $\Delta\phi(lepton, bbj)$ is selected thanks to the good discrimination between multijet and electronweak backgrounds provided by these variables.

The combination of these uncertainties gives rise to the fractions of the multijet contribution compared to the total background and their uncertainties are presented in Table 5.20 and Table 5.21 for high and medium p_T^V region, respectively. In high p_T^V region, the main contribution to the systematic uncertainties is from using the reduced inverted isolation region in both electron and muon subchannels, while in the medium p_T^V region, the main contributors are: the change of variable used in the template fit and the removal of the $m_T^W > 20\,\text{GeV}$ cut.

Table 5.20: Summary of MJ fractions, along with their associated uncertainty in the 2-jets and 3-jets high p_T^V regions ($W + HF$ CR and SR are combined) separately.

Region	MJ Fractions (%)	MJ norm. uncertainty
2-tag, 2-jet, e	$1.91^{+1.96}_{-1.91}$	-100% / +105%
2-tag, 2-jet, μ	$2.76^{+2.06}_{-1.65}$	-60% / +75%
2-tag, 3-jet, e	$0.15^{+0.24}_{-0.15}$	-100% / +160%
2-tag, 3-jet, μ	$0.43^{+1.10}_{-0.43}$	-100% / +260%

Table 5.21: Summary of MJ fractions, along with their associated uncertainty in the 2-jets and 3-jets medium p_T^V regions ($W + HF$ CR and SR are combined) separately.

Region	MJ Fractions (%)	MJ norm. uncertainty
2-tag, 2-jet, e	$3.57^{+0.44}_{-0.79}$	-12% / +22%
2-tag, 2-jet, μ	$2.76^{+1.19}_{-0.64}$	-25% / +40%
2-tag, 3-jet, e	$0.85^{+0.37}_{-0.31}$	-40% / +45%
2-tag, 3-jet, μ	$2.14^{+0.26}_{-1.03}$	-50% / +12%

5.5.2.4 Multijet estimation in dijet mass analyis

In the di-jet analysis, due to the additional cuts and the diffetent analysis categories compared to multivariate analysis, the independent multijet estimation is needed. The general strategy is very similar to what was used for the multivariate analysis, the relevant differences are presented in this section.

Briefly, the multijet background is estimated with the same template fit ethod as in the MVA. However, the template fits to the m_T^W distributions do not have as good a performance in terms of discrimination between the Top and W+jets backgrounds, this is because the latter is obtained in the MVA thanks to the distinction between signal and $W + HF$ control regions, which is not applied in the dijet-mass analysis.

Therefore, a preliminary fit is performed in each analysis region to a variable showing good discrimination between these two backgrounds. The variable showing the best performance in this respect was found to be $\Delta R(b, \bar{b})$. The fit is performed over the combined electron-muon $\Delta R(b, \bar{b})$ distribution with two free normalization factors. The MJ background, known to be small, is neglected at this stage, but the fitted normalization factors are used to provide only the relative fractions of Top and W+jets backgrounds, from which the global shape of the electroweak background that is used in the subsequent template fit involving the MJ background is obtained.

A template fit is next performed in each analysis region to a variable showing

good discrimination between multijet and electroweak backgrounds. This variable is traditionally m_T^W in the multivariate analysis, but it was found that other variables could provide a similar or even better discrimination in the dijet mass analysis (based on the statistical errors of the fits). Here, the azimuthal angle between the lepton and the missing transverse energy, $\Delta\phi(l, E_T^{miss})$, was found to provide the best overall performance, considering the various analysis regions. Fits to the m_T^W distributions are nevertheless used in the assessment of systematic errors. Each template fit is performed simultaneously over the separate electron and muon distributions with three free scale factors, one for the electroweak background, one for the multijet background in the electron sub-channel, and similarly one in the muon sub-channel, with all scale factors constrained to remain non-negative.

Such multijet scale factors should be determined in each of the analysis regions. However, the statistics are quite limited for $p_T^V > 200\,\text{GeV}$, leading to results overly sensitive to statistical fluctuations. Therefore, multijet scale factors are determined for $p_T^V > 150\,\text{GeV}$ and applied in all analysis regions in this p_T^V range.

The resulting MJ fractions are given in Table 5.22, separately for electrons and muons as well as for their combination. The MJ fractions are small, less than 1% except in the medium p_T^V region in the 2-jets category where they are at the 3% level. The Top, W+jets and multijet scale factors obtained in the template fits are used in Figure 5.42 to Figure 5.43 to show the agreement of the simulation with the data.

Table 5.22: Fractions of multijet background in percent, separately for electrons and muons as well as combined, for 2- and 3-jet events. The errors represent the combined statistical and systematic uncertainties.

Region	$75 - 150\,\text{GeV}$	$> 150GeV$
Electrons 2 jets	2.6 (+0.6 -0.4)	0.0 (+2.1 -0.0)
Muons 2 jets	3.0 (+1.6 -0.7)	0.6 (+1.1 -0.6)
Combined 2 jets	2.8 (+0.9 -0.4)	0.4 (+1.1 -0.4)
Electrons 3 jets	0.0 (+1.1 -0.0)	0.0 (+0.9 -0.0)
Muons 3 jets	1.5 (+1.0 -0.1)	0.0 (+0.7 -0.0)
Combined 3 jets	0.8 (+0.7 -0.0)	0.0 (+0.6 -0.0)

The following systematic uncertainties in the normalization of the MJ background were considered:

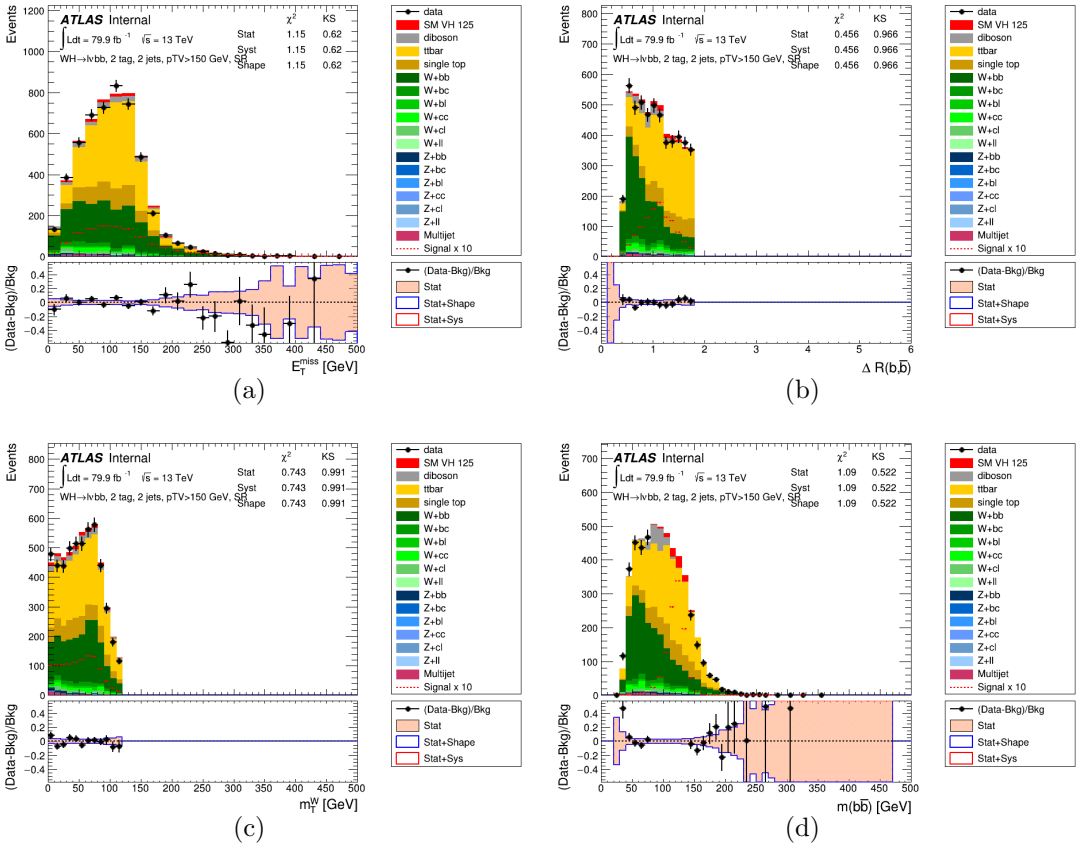


Figure 5.42: Distributions in the high p_T^V region and in the 2-jets category with the Top, W+jets and multijet scale factors applied. The electron and muon sub-channels are combined.

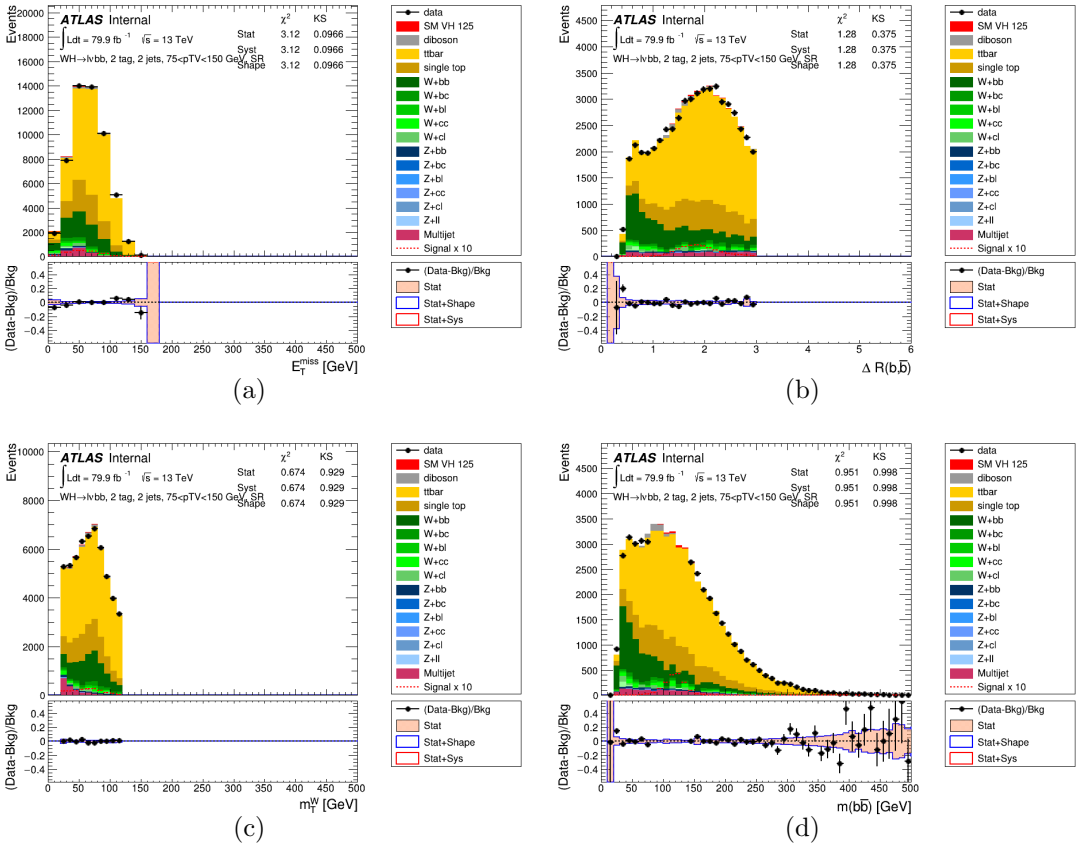
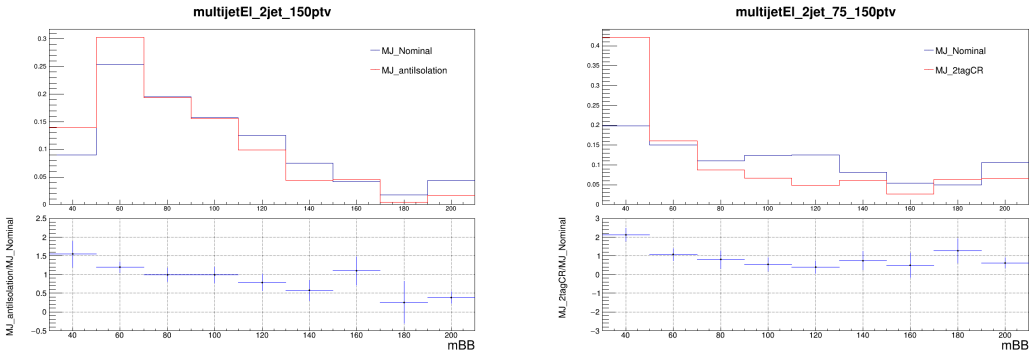


Figure 5.43: Distributions in the medium p_T^V region and in the 2-jets category with the Top, W+jets and multijet scale factors applied. The electron and muon sub-channels are combined.

- The m_T^W distribution is used in the template fits instead of $\Delta\phi(l, E_T^{miss})$.
- The multijet templates are obtained from data in the 1-tag control regions after subtraction of the electroweak background. To normalize the electroweak background in a given 1-tag CR, an "ad hoc" scale factor is applied, simply taken to be the ratio of data to simulation in the corresponding 2-tag signal region. This is replaced by a similar ratio calculated in the 1-tag signal region.
- The shape of the electroweak background in a template fit is affected by the relative contributions of the Top and W+jets backgrounds. These fractions are obtained from a fit to the $\Delta R(b\bar{b})$ distribution. The fitted Top and W+jets fractions are modified by the corresponding fitted errors, taking into account their anti-correlation.
- Instead of using the full CRs, only the halves of MJ events closest to the signal regions in terms of value of the isolation variable are used
- The 2-tag CRs are directly used instead of the 1-tag CRs (at the expense of reduced statistics).
- In the medium pTV region, the $m_T^W > 20GeV$ cut is removed
- For $p_T^V > 150\,\text{GeV}$ and in the electron sub-channel, the $E_T^{miss} > 30GeV$ cut is removed.
- Only the lowest unprescaled single-lepton triggers, which involve isolation criteria, are used. (The muon sub-channel is unaffected for $p_T^V > 150GeV$, where E_T^{miss} triggers are used instead.)

The shape of the $m_{b\bar{b}}$ distribution of the MJ background is also affected by some of the aforementioned systematic uncertainties, namely those related to: the choice of "ad-hoc" scale factors; the shape of the electroweak background; the size of the CRs; the choice of 2-tag rather than 1-tag CRs; the single-lepton triggers. For each of these systematic uncertainty sources, the ratio of the varied to nominal $m_{b\bar{b}}$ distributions is computed and is found to be significantly different from being uniform in only a few cases: the choice of 2-tag CRs in the medium p_T^V region in the electron sub-channel and the reduction of the size of the CRs for 2-jets events in the high p_T^V region. They cover all the other variations and are implemented in the global fit as shape-only systematics. These two variations are shown in Figure 5.44 for 2-jets events in the electron sub-channel.



(a) (b)

Figure 5.44: Nominal and systematically varied distributions, with their ratio in the bottom panels. The systematic variations are, for the electron sub channel: the reduction of the size of the CRs in the high p_T^V region (left); the choice of 2-tag rather than 1-tag CRs in the medium p_T^V region (right).

5.5.3 2-lepton channel

In the 2-lepton channel the multijet background is highly suppressed by requiring an event with two isolated leptons, and a dilepton invariant mass close to that of a Z boson. Any residual QCD background is estimated using the template method, which fits the expected EW background contributions estimated from MC simulations, and an exponential model for the multijet background, to same-sign charged data events over the m_{ll} distribution. An estimate is then made of the fraction of the background in a mass window around the Z boson peak in the signal region that could be attributed to multijet events based on the assumption that the opposite sign and same sign events are symmetric for multijet events. Inside a dilepton mass window $71\,\text{GeV} < m_{ll} < 121\,\text{GeV}$ the upper limit of the expected MJ contamination as a fraction of the total electroweak background is estimated to be 0.34% and 0.08% for the electron and muon sub-channels, respectively. In the $100\,\text{GeV} < m_{b\bar{b}} < 140\,\text{GeV}$ mass window, the residual multijet contamination is found to be less than 10% of the signal contribution, and found to have a BDT shape similar to the one expected for the sum of the remaining backgrounds. This is thus small enough to have a negligible impact on the signal extraction and so is not included in the global likelihood fit.

5.6 1-lepton Channel Optimization

The author is mainly working on the 1-lepton channel analysis. In this section, some studies with the purpose of improving the sensitivity and robustness of the

1-lepton channel analysis are presented.

5.6.1 $W \to \tau_{had}\nu$ channel study

In the default 1-lepton channel analysis, only electron and muon sub-channels are considered and a dedicated $W \to \tau\nu$ channel is not included. About 35% τ lepton undergo leptonic decay and present an electron or muon in the final state, the default 1-lepton channel selections can already cover such events efficiently. In the other hand, about 65% τ lepton undergo hadronic decay, and present a τ_{had} jet in the final state. The default 0-lepton channel actually has some sensitivities for such events since no τ_{had} veto (vetoing events with τ_{had} jet presented) selection applied in this channel. In this study, we want to test if a channel explicitly selecting hadronic τ decays could bring additional sensitivity for this analysis.

This study is based on the WH signal MC samples which are simulated and reconstructed using the 2015-2016 data running conditions, and normalized to 36.1 fb^{-1}. The first step is applying the default 0-lepton selections on these WH signal events, then for the events do not pass the selections, a set of dedicated requirements to select the signal events with W decays to τ_{had} and neutrino (referred as τ_{had} selection) are considered to check how many events can be recovered. The possible triggers can be used in the τ_{had} selection are summarized in Table 5.23, including the signal τ_{had} triggers, $\tau_{had} + E_T^{miss}$ triggers and E_T^{miss} triggers as used in the 0-lepton channel.

Apart from the trigger requirements, exactly one medium τ_{had} jet with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.5$ (excluding $1.37 < |\eta| < 1.52$) is also required in the τ_{had} selection.

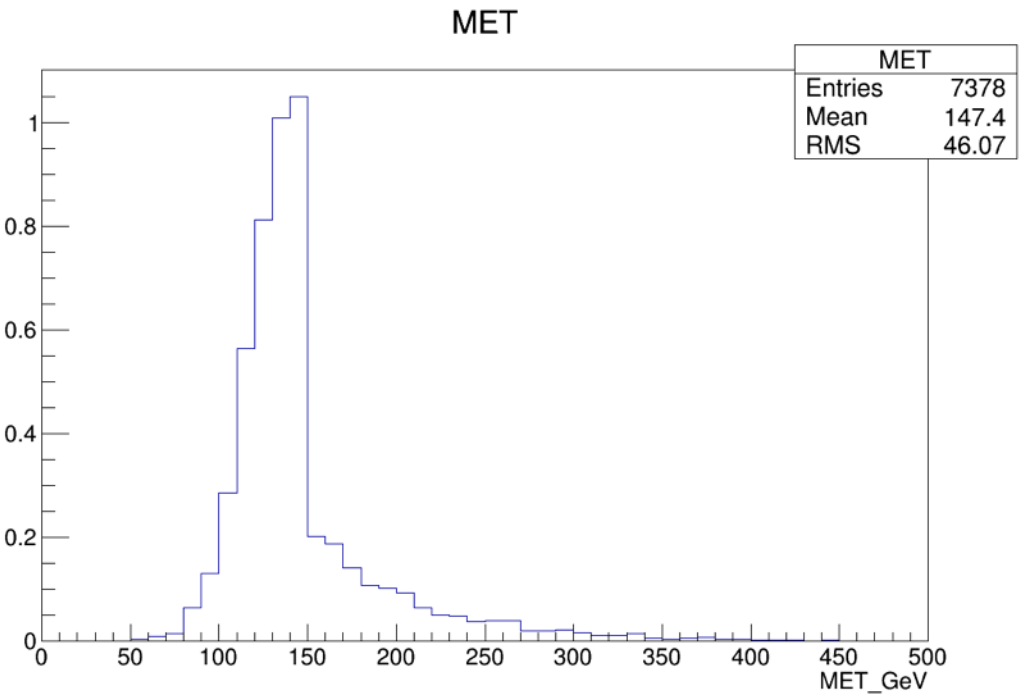
The total yield of signal events that fail the default 0-lepton selections but pass the τ_{had} selection is 10.25 ± 0.15 (statistical uncertainty). For such events, there are 3.86 events pass only the E_T^{miss} trigger. Figure 5.45 shows the offline E_T^{miss} distribution for these events. Most of these events have offline E_T^{miss} less than 150 GeV, and would be difficult to use as they are in the turn-on of the trigger. In that case, the E_T^{miss} trigger is discarded in the τ_{had} selection.

For the signal events pass the τ_{had} selection: there are total 4.78 events pass the $\tau_{hadd} + E_T^{miss}$ trigger requirement, when adding the $p_T^V > 150\,\text{GeV}$ cut for harmonizing with the other default channels, 3.83 events left; there are total 1.62 events pass the single τ_{had} trigger but fail the $\tau_{had} + E_T^{miss}$, when adding the $p_T^V > 150\,\text{GeV}$ cut, 1.07 events left.

In total, 4.9 ± 0.07 WH signal events can be recovered by the τ_{had} selection in

Table 5.23: Summary of the possible triggers can be used for the τ_{had} selection.

Single τ_{had} trigger	
Data period	Trigger name
2015 - 2016 (A)	HLT_tau80_medium1_tracktwo_L1TAU60
2016 (B-D3)	HLT_tau125_medium1_tracktwo
2016 ($\geq$ D4)	HLT_tau160_medium1_tracktwo
τ_{had} + E_T^{miss} trigger	
All	HLT_tau35_medium1_tracktwo_xe70_L1XE45
E_T^{miss} trigger	
2015	HLT_xe70_mht_L1XE50
2016 (A-D3)	HLT_xe90_mht_L1XE50
2016 ($\geq$ D4)	HLT_xe110_mht_L1XE50



(a)

Figure 5.45: Offline E_T^{miss} distribution for evnets passing the τ_{had} selection and passing only the E_T^{miss} trigger requirements.

$p_T^V > 150\,\text{GeV}$ region, which brings only 4% increase compared to the WH signal events selected by the default channels. Furthermore, more criteria need to be added in the τ_{had} selection, such as increased offline τ_{had} jet p_T cut to be able to really use such τ_{had} triggers and the selections to reduce the multijet contribution. The conclusion is then made that the dedicated $W \to \tau_{had}\nu$ channel is helpless for increasing the analysis sensitivity and therefore is not considered in the analysis.

5.6.2 τ_{had} removal

Even though the dedicated $W \to \tau_{had}\nu$ channel is useless in the analysis, the τ_{had} veto (vetoing events with τ_{had} jet presented) shall be capable to remove quite a lot $t\bar{t}$ events and thus bring some additional sensitivity in 1-lepton channel. Consider a typical $t\bar{t}$ event, in where both W bosons undergo leptonic decays, and one of them decays to an electron or muon and neutrino, while another one decays to a τ and neutrino, and the τ lepton then undergos hadronic decay and present a τ_{had} jet in the event. After reconstruction, such events have typically one electron or muon, 2 b-tagged jets, one τ_{had} jet and sufficiently E_T^{miss} presented in the final state. The τ_{had} veto helps to remove the such events with no expected signal loss. Table 5.24 shows the WH signal and different background processes efficiencies in high p_T^V region when adding the τ_{had} veto requirement.

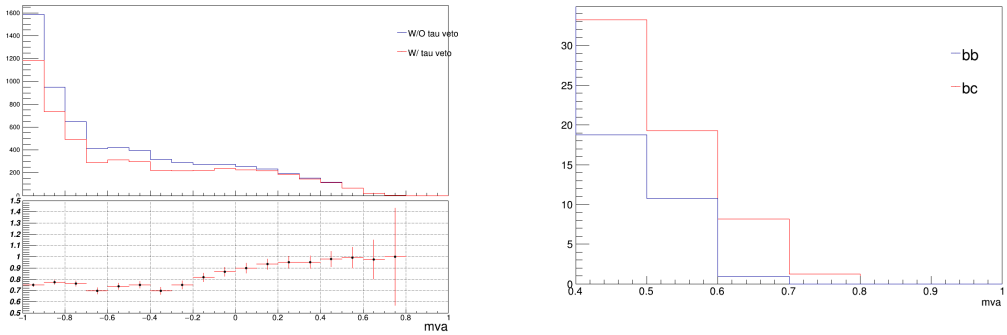
Table 5.24: WH signal and background events efficiencies in high p_T^V region when adding the τ_{had} veto requirement.

Region	WH signal	$t\bar{t}$	single top	W + HF
2tag2jet	99.7%	79.4%	93.8%	99.6%
2tag3jet	99.7%	93.2%	97.4%	99.3%

As can be seen in the table, the τ_{had} veto has basically no effect on WH signal events and W + HF events. For $t\bar{t}$ events, the effect is mainly on the 2-jet region, in where $\sim$ 20% events have been removed, while in 3-jet region, the effect is smaller and $\sim$ 7% events have been removed. τ_{had} veto removes also $\sim$ 6% (3%) single top events in 2- (3-) jet region.

Figure 5.46 (a) shows the BDT distributions in 2-tag 2-jet region for both the $t\bar{t}$ events with and without τ_{had} veto applied. As can be seen, the effect is more visible in relative low BDT region (BDT < 0.2), while very limited $t\bar{t}$ events have been removed in the particular high BDT region (BDT > 0.4) and the improvement of analysis sensitivity is therefore modest. The reason can be explained by the $t\bar{t}$

truth flavor component achieved by the same truth matching scheme as discussed in the Section 5.2. Figure 5.46 (b) shows the BDT distribution in particular high BDT region (BDT > 0.4) region without τ_{had} veto applied for both bb and bc $t\bar{t}$ events. bb means both the b-tagged jets in the $t\bar{t}$ event match the truth b-hadrons, while bc means one b-tagged jets matches a truth b-hadron while another one matches a truth c-hadron. As can be seen, the bc events are clearly dominated in such high BDT region, however, only $\sim 2\%$ of the $t\bar{t}$ events removed by the τ_{had} veto in the 2-jet region are bc events, therefore the τ_{had} veto has modest effect on the high BDT region. Apart from that, the τ_{had} related experimental uncertainties need to be considered when introducing the τ_{had} veto in the analysis. Synthesize all the considerations, τ_{had} veto is then not adopted in order to keep the analysis simple. However, the removal of 20% $t\bar{t}$ events in the 2-jet signal region is crucial for such a systematics uncertainty dominated analysis. And also the BDT distributions shown in Figure 5.46 are from the default training, the BDT re-training with τ_{had} veto applied on top may brings additional sensitivity in the analysis, so the τ_{had} veto is definitely worth more detailed studies in the next round of the analysis.



(a) (b)

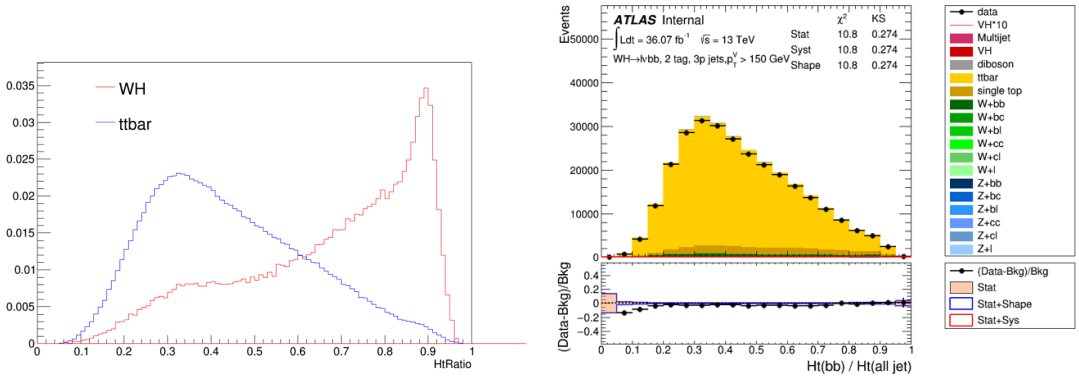
Figure 5.46: (a): BDT distributions in 2-tag 2-jet region for both the $t\bar{t}$ events with (red) and without (blue) τ_{had} veto applied. (b): BDT distributions in 2-tag 2-jet region, for events with particular high BDT value (BDT > 0.4) without τ_{had} veto applied for both bb (blue) and bc (red) $t\bar{t}$ events.

5.6.3 $t\bar{t}$ reduction cut study

As discussed in Section 5.3, the default 1-lepton channel selections reject the events with more than 3 jets due to the high $t\bar{t}$ background contamination in that region. In this study, a new $t\bar{t}$ reduction cut is investigated, to see if additional signal sensitivity can be achieved by using this cut in 3+-jet region instead of simply

removing the events with more than 3 jets. The MC samples used in this study are simulated and reconstructed using the 2015-2016 data running conditions, and normalized to 36.1 fb^{-1}.

Consider a signal event, the initial state radiation (ISR) jets are likely to have low p_T values, while the b-tagged jets from $H \to b\bar{b}$ decay are likely to have relative high p_T values as they carry the kinematical energy of the Higgs boson. For a $t\bar{t}$ event, the b-tagged jets used for reconstructing the Higgs candidate have in average as much p_T as the jets from the W bosons. Under this consideration, a new variable, called as HtRatio, is build with the ratio of scalar sum of p_T of two b-tagged jets and scalar sum of p_T of all the jets in the event. Figure 5.47(a) shows the HtRatio distributions for both WH signal and $t\bar{t}$ events in 2-tag 3+-jet $p_T^V >$ 150 GeV region, and clear discrimination can be seen as expected. Figure 5.47(b) shows also the data simulation comparison in such region and in general quite good modelling of this variable observed which indicates the cut on this variable can be safely used to separate the signal from the $t\bar{t}$ events and this distribution can be also used directly in the BDT training.



(a) (b)

Figure 5.47: (a): HtRatio distrubitons for signal (red) and $t\bar{t}$ events (blue) in 2-tag 3+-jet $p_T^V >$ 150 GeV region, default MVA selections are applied, the yields are normalized to unity. (b): HtRatio distrubitons in 2-tag 3+-jet region for data and all the MC simulations. Default MVA selections are applied, no normalization factors are applied to the MC events.

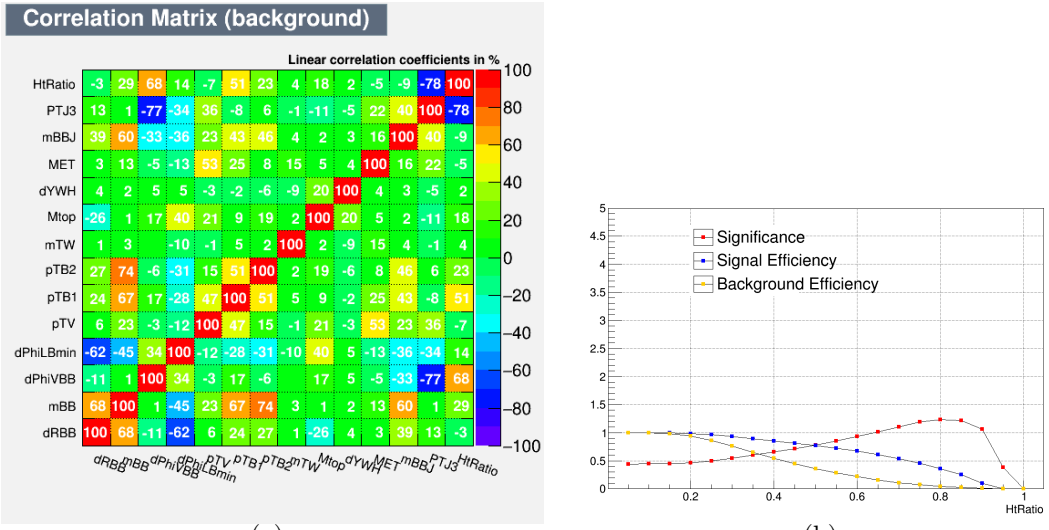
In order to quantify the improvement from the new HtRatio variable, the log-likelihood ratio statistical only sensitivity, S, is used and calculated on a bin-by-bin basis for a given distribution,

$$S = \sqrt{\sum_{i=1}^{n}(2 \times ((s_i + b_i) \times ln(1 + s_i/b_i) - s_i))}, \tag{5.8}$$

where n is the total number of bins in the distribution, s_i is the signal yield in bin i, and b_i is the background yield in bin i.

In the default 1-lepton 2-tag 3-jet region, the total signal yield is 58.35 while total $t\bar{t}$ background yield is 22095.3. The S calculated with BDT distribution is 1.65, these numbers are referred to as the baseline numbers. Different configurations are tested with HtRatio variable then. First attempt is to re-train the BDT in 3-jet region with adding the HtRatio distribution into the input variables list. The S calculated with the retrained BDT output in 3-jet region is the same as the baseline. The correlations between input variables for the background in the BDT training are shown in Figure 5.48 (a). The HtRatio and p_T^{j3} are highly correlated, which indicates HtRatio is useless in the 3-jet region training due to the present of p_T^{j3} in the input variables list. The second attempt is to re-train the BDT in 3+-jet region with adding the HtRatio distribution into the input variables list, the S calculated in this category is still the same as the baseline number even thought a better signal and background separation observed compared to the default case in 3-jet region. The non-improved sensitivity is mainly due to the much higher $t\bar{t}$ contamination in the 3+-jet region. So the remaining possible approach is re-training the BDT still in 3+-jet region but with a proper HtRatio cut applied on top. To achieve the proper cut, a cut scan approach is implemented to the HtRatio distribution in 3+-jet region and the corresponding S is calculated with the m_{bb} distribution instead of the BDT to avoid the huge amount of works to retrain the BDT for every individual HtRatio cut value. As shown in Figure 5.48 (b), the cut value scan is performed to the HtRatio distribution from leftmost to rightmost with a step of 0.05. For every cut value, a new m_{bb} distribution is built to calculate the S in the range of 30 GeV to 200 GeV with 10GeV/bin. Table 5.25 shows also the detailed numbers of WH signal events, $t\bar{t}$ background events and statistical sensitivities in the 3-jet region, in the 3+-jet region without any HtRatio cut and in the 3+-jet region with HtRatio > 0.75 cut. Compared with the baseline numbers, the HtRatio > 0.75 cut keeps the similar signal yield, reduces $\sim 25\%$ $t\bar{t}$ background and increase the S by $\sim 60\%$, and therefore is selected as the cut applied on top before the BDT training in the 3+-jet region.

The BDT is then re-trained in 3+-jet region with adding HtRatio distribution in the input variables list and applying HtRatio > 0.75 cut on top. Figure 5.49 shows the BDT distributions of signal and sum of all background processes while training and testing in both the default case in 3-jet region and in 3+-jet region with HtRatio cut. Clearly better signal and background separation can be seen in the latter case. The sensitivity calculated with new BDT out distributions in



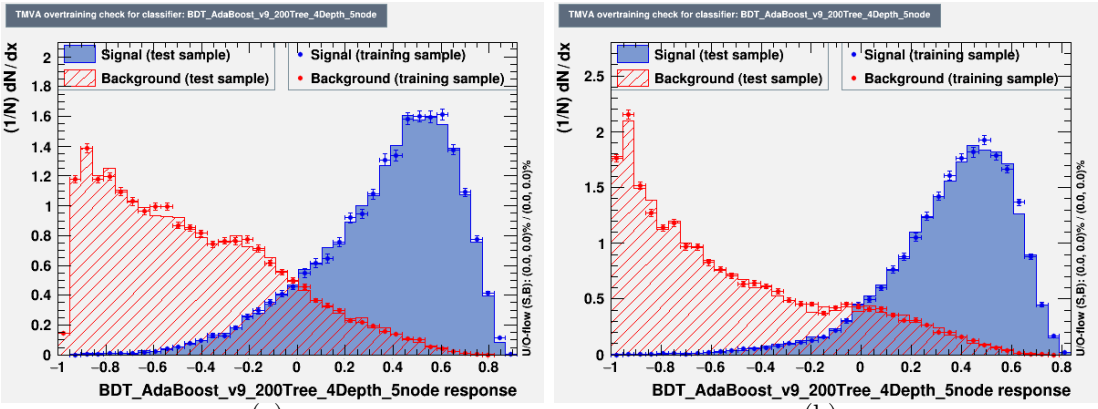
(a) (b)

Figure 5.48: (a): Correlation matrices of the VH BDT input variables in the 1-lepton 3-jet region for the sum of all background processes, the HtRatio and P_T^{j3} are highly correlated. (b): HtRatio cut scan results in 3+-jet region.

Table 5.25: Summary of WH signal events, $t\bar{t}$ background events and statistical sensitivity in different regions.

Region	WH signal events	ttbar events	Sensitivity (S)
2-tag 3-jet (Baseline)	58.35	22095.3	0.71
2-tag 3+-jet (No HtRatio cut)	129.44	259073	0.43
2-tag 3+-jet (HtRatio $>$ 0.75)	58.25	16838.75	1.08

3+-jet region is 1.79, $\sim$ 9% improvement achieved compared with the baseline number (1.65). However, when combined with 2-jet region which provides the most sensitivity in the analysis, the overall improvement is quite limited. The S calculated in default 2-jet region is 2.85, combined with the S from the new 3+-jet region by adding the individual sensitivity in quadrature, the overall S is 3.36, compared with the overall S in the default 2 and 3-jet region, 3.29, the improvement is only $\sim$ 2%.



(a) (b)

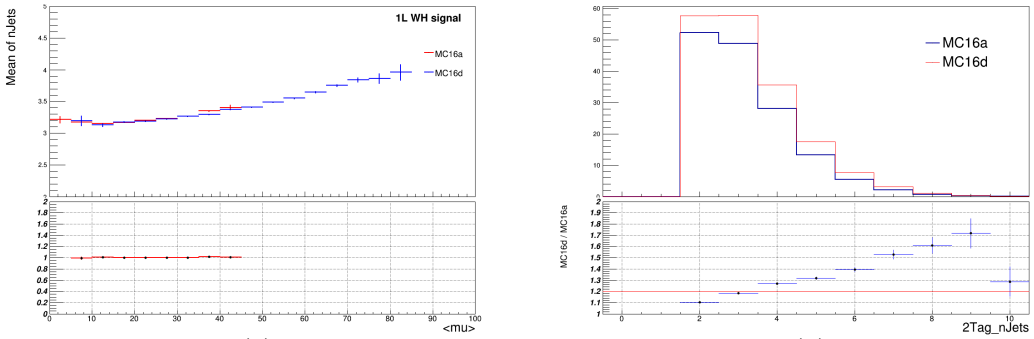
Figure 5.49: BDT distributions of signal (blue) and sum of all background processes (red) while training (dots) and testing (histogram) in default 3-jet region (a), and in 3+-jet region with adding HtRatio distribution in the input variables list and applying HtRatio $>$ 0.75 cut on top.

In conclusion, HtRatio cut can be only used for events with 3 or more than 3 jets. The study shows that using 3+-jet region with HtRatio cut is clearly better than using only 3-jet region. But since 2-jet region provides the most sensitivity in the analysis, the overall statistical only sensitivity improvement is only $\sim$ 2%. In that case, there is no need to complicate the analysis and the therefore the HtRatio cut and 3+-jet region are not adopted in the 1-lepton analysis.

5.6.4 Pile-up jet suppression

As discuss in Section 3.4, the mean number of interactions per bunch crossing ($< \mu >$) in data17 is much larger than the $< \mu >$ in data15 and 16. It indicates that the effect from pile-up events may more visible on data17. Figure 5.50 (a) shows the 1-lepton signal events $< \mu >$ versus jet multiplicity for both the MC16a and MC16d samples. MC16a events are simulated and reconstructed using the 2015-2016 data running conditions and normalized to 36.1 fb^{-1}, while MC16d

events are simulated and reconstructed using the 2017 data running conditions and normalized to 43.8 fb^{-1}. As can be seen, jet multiplicity is increased with the increase of $< \mu >$ and yield the result shown in Figure 5.50 (b). This plot shows the 1-lepton WH signal number of jets distributions in 2-tag region for both the MC16a and MC16d samples, the bottom pad of the plot shows the ratio of MC16d events yield and MC16a events yield, the red line at 1.2 represents the luminosity ratio of data17 and data15-16. As can be seen, in the 2- and 3-jets signal region, the increased WH signal events are less than the expected events from the luminosity ratio, due to the jet multiplicity migration. The increase of the analysis sensitivity then will be smaller than the one expected from the increase of the luminosity. A study is therefore performed in 1-lepton channel to suppress the pile-up jets further and increase the analysis sensitivity.



(a) (b)

Figure 5.50: (a): 1-lepton signal events $< \mu >$ versus jet multiplicity for both the MC16a and MC16d samples. (b): 1-lepton WH signal number of jets distributions for both the MC16a and MC16d samples, the bottom pad of the plot shows the ratio of MC16d events yield and MC16a events yield, the red line at 1.2 represents the luminosity ratio of data17 and data15-16.

The default pile-up jet suppression requirement applied to only the jets with $|\eta| < 2.4$ and $p_T < 60\,\text{GeV}$, with requiring JVT > 0.59. There is no cut applied to the forward jets and also the signal jets with $2.4 < |\eta| < 2.5$. Different ways are considered apart from the the default JVT cut and the results are shown one by one in the following.

Apply the ForwardJVT (FJVT) cut to the forward jet FJVT is a new technique developed for the suppression of pile-up jets in the forward region with $p_T < 50\,\text{GeV}$ as discussed in Section 4.5. In this study, the tight working point is adopted. Table 5.26 shows the comparison of WH signal event yield, total background event yield and the statistical only sensitivity (S) calculated with

Equation 5.8 with the m_{bb} distribution before (referred to as Default) and after (referred to as Option1) applying the FJVT cut. The sensitivities are calculated separately in 2- and 3-jet signal regions with di-jet mass analysis selections applied, and then combined in quadrature as the final sensitivity. As can be seen, after applying the FJVT cut, the increase of signal (background) yield is about 5% (10%) in 2-jet signal region, while the sensitivity remains the same.

Table 5.26: Comparison of WH signal event yield, total background event yield and the statistical only sensitivity (S) before (Default) and after (Option1) applying the FJVT cut.

	Default	Option1
2-jet signal region		
Signal yield	57.4	60.3
Background yield	5975.3	6565.5
3-jet signal region		
Signal yield	56.8	57.5
Background yield	35134.3	37341.5
Sensitivity (S)	2.21 ± 0.07	2.21 ± 0.07

Apply the tight JVT cut to the non b-tagged signal jets. The default JVT working point used in this analysis is the medium working point. In this study, tight JVT working point is tested and applied to the non b-tagged jet. The b-tagging algorithm itself provides already the strong suppression of the pile-up jet, so there is no need to tighten the JVT cut for the jets that already being b-tagged. Table 5.27 shows the comparison of WH signal event yield, total background event yield and the statistical only sensitivity (S) calculated with Equation 5.8 with the m_{bb} distribution before (referred to as Default) and after (referred to as Option2) applying the tight JVT cut to the non b-tagged jets. As can be seen, after applying this cut, the increase of signal (background) yield is about 10% (22%) in 2-jet signal region, and the sensitivity reduced about 1.4% due to the much higher increase of background yield than the signal yield.

Raise the p_T cut for the non b-tagged signal jets. The default p_T cut for the signal jet is $p_T > 20\,\text{GeV}$. In this study, raising the p_T cut to 30 GeV is tested for the non b-tagged jets. Table 5.28 shows the comparison of WH signal event yield, total background event yield and the statistical only sensitivity

Table 5.27: Comparison of WH signal event yield, total background event yield and the statistical only sensitivity (S) before (Default) and after (Option2) applying the tight JVT cut to the non b-tagged jets.

	Default	Option2
2-jet signal region		
Signal yield	57.4	63.5
Background yield	5975.3	6565.5
3-jet signal region		
Signal yield	56.8	57.3
Background yield	35134.3	37341.5
Sensitivity (S)	2.21 ± 0.07	2.18 ± 0.07

(S) calculated with Equation 5.8 with the m_{bb} distribution before (referred to as Default) and after (referred to as Option3) raising the jet p_T cut to 30 GeV for the non b-tagged jets. As can be seen, after raising this cut, the increase of signal (background) yield is about 36% (87%) in 2-jet signal region, and the sensitivity reduced about 5.0% due to the much higher increase of background yield than the signal yield, in particular for the $t\bar{t}$ background that the yield increased more than 100% in 2-jet signal region.

Table 5.28: Comparison of WH signal event yield, total background event yield and the statistical only sensitivity (S) before (Default) and after (Option3) raising the jet p_T cut to 30 GeV for the non b-tagged jets.

	Default	Option3
2-jet signal region		
Signal yield	57.4	78.4
Background yield	5975.3	11166.9
3-jet signal region		
Signal yield	56.8	56.6
Background yield	35134.3	58831.2
Sensitivity (S)	2.21 ± 0.07	2.10 ± 0.07

Raise the p_T cut for the non b-tagged signal jets in the region of 2.4

$< |\eta| <$ **2.5**. The default JVT requirement works only for the jet with $|\eta| < 2.4$. In this study, raising the p_T cut to 30 GeV is tested for the non b-tagged signal jets with $2.4 < |\eta| < 2.5$. Table 5.29 shows the comparison of WH signal event yield, total background event yield and the statistical only sensitivity (S) calculated with Equation 5.8 with the m_{bb} distribution before (referred to as Default) and after (referred to as Option4) raising the jet p_T cut to 30 GeV for the non b-tagged jets with $2.4 < |\eta| < 2.5$. As can be seen, after raising this cut, the increase of signal (background) yield is about 2% (3%) in 2-jet signal region, and the sensitivity remains the same.

Table 5.29: Comparison of WH signal event yield, total background event yield and the statistical only sensitivity (S) before (Default) and after (Option4) raising the jet p_T cut to 30 GeV for the non b-tagged jets with $2.4 < |\eta| < 2.5$.

	Default	Option4
2-jet signal region		
Signal yield	57.4	58.6
Background yield	5975.3	6136.4
3-jet signal region		
Signal yield	56.8	56.9
Background yield	35134.3	35901.6
Sensitivity (S)	2.21 $\pm$ 0.07	2.21 $\pm$ 0.07

In conclusion, in order to further suppress the pile-up jets and increase the analysis sensitivity, different ways have been tested in 1-lepton channel, including using the FJVT requirement, using the tighter JVT requirement and using tighter jet p_T requirement. None of them bring real increase of the sensitivity for this analysis, some of them even harm the sensitivity a lot due to the much higher increase of background yield than signal yield. The same studies have been also performed in 0- and 2- lepton channels, and the similar conclusions achieved. In that case, the default pile-up jets suppression cuts have been kept and no other actions adopted in this analysis.

5.6.5 Dijet-mass analysis selection and categorization optimization

The default dijet-mass analysis additional selections and categories as discussed in Section 5.3.4 are mainly inherited from the Run 1 analysis without careful re-optimization. The dijet-mass analysis is performed as an important cross-check to the main multivariate analysis, and has typically 10% to 15% lower sensitivity compared to the multivariate analysis. The dijet-mass analysis shall be also play an more important role when the analysis moving towards the precision measurement with more data. In that case, an effort to optimize the dijet-mass analysis additional selections and categories in 1-lepton channel has been made, and the results are shown in this section. This study is based on the MC samples which are simulated and reconstructed using the 2015-2016 data running conditions, and normalized to 36.1 fb^{-1}. The sensitivity values quoted in this study are statistical only sensitivity calculated with Equation 5.8.

By default, two set of additional cuts on m_T^W and ΔR_{bb} are used 1-lepton channel dijet-mass selection as shown in Table 5.10, in order to improve the analysis sensitivity. Figure 5.51 shows the m_T^W distributions comparison for signal and total background with only MVA selection applied in different dijet-mass signal regions, the m_T^W distributions have been scaled to the same (unit) area in order to highlight the shape differences. Figure 5.52 shows the data and MC simulations comparison in the same regions with the same selection applied. As can be seen, the $m_T^W < 120\,\text{GeV}$ cut is designed to reduce the $t\bar{t}$ backgrounds that underdo the dileptonic decays. To test if the cut value is optimal in the analysis, the cut value scan method that used in $t\bar{t}$ reduction cut study (5.6.3) is also adopted in this study. As shown in Figure 5.53, the cut value scan is performed to the m_T^W distribution from leftmost to rightmost with a step of 10 GeV. For every cut value, new m_{bb} distributions are built in different dijet-mass signal regions to calculate the sensitivity (S) in the m_{bb} range of 30 GeV to 200 GeV with 10GeV/bin. The scan result shows the default $m_T^W < 120\,\text{GeV}$ cut has already given the best significance with a reasonable signal efficiency. Without any m_T^W, the calculated significance is 2.10, while with $m_T^W < 120\,\text{GeV}$ cut, the significance is 2.14 with only less than 5% signal loss. The default m_T^W cut is therefore kept in 1-lepton channel.

The same cut value scan method is then performed to the ΔR_{bb} distributions. The p_T^V categories are first kept as default.The default ΔR_{bb} cuts are $\Delta R_{bb} < 1.8$ in $150\ \text{GeV} < p_T^V < 200\ \text{GeV}$ region and $\Delta R_{bb} < 1.2$ in $p_T^V > 200\ \text{GeV}$ region. Figure 5.54 shows the ΔR_{bb} distributions comparison for signal and total

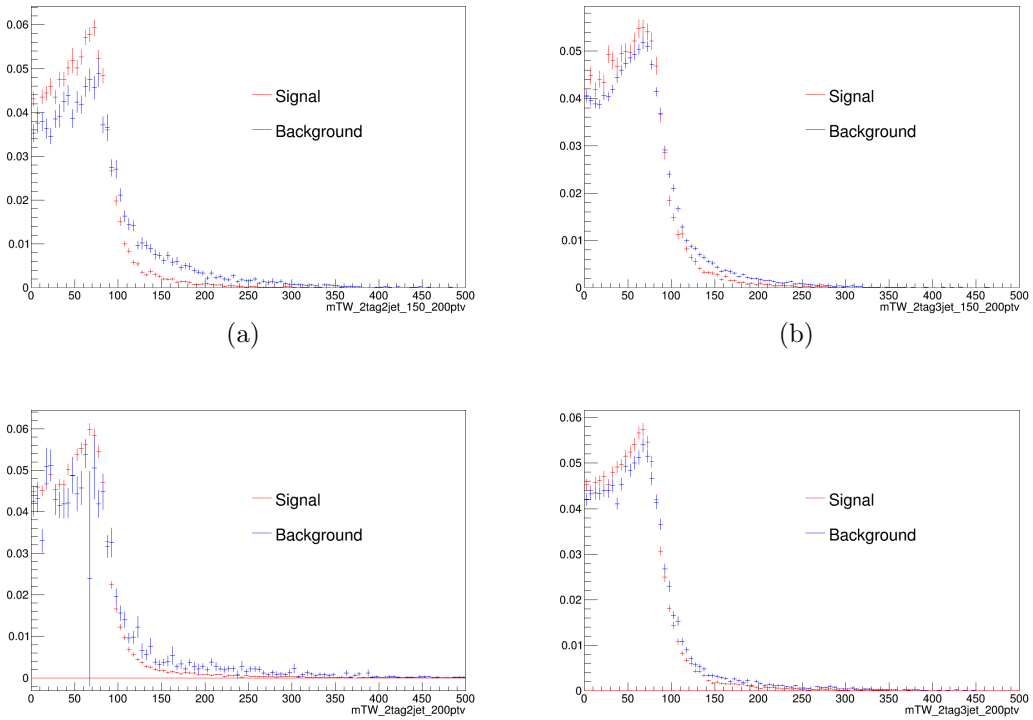


Figure 5.51: m_T^W distributions comparison for signal and total background with only MVA selection applied in different dijet-mass signal regions : 2-jet 150 GeV $< p_T^V <$ 200 GeV (a), 3-jet 150 GeV $< p_T^V <$ 200 GeV (b), 2-jet $p_T^V >$ 200 GeV (c), 3-jet $p_T^V >$ 200 GeV (d). The m_T^W distributions have been scaled to the same (unit) area in order to highlight the shape differences.

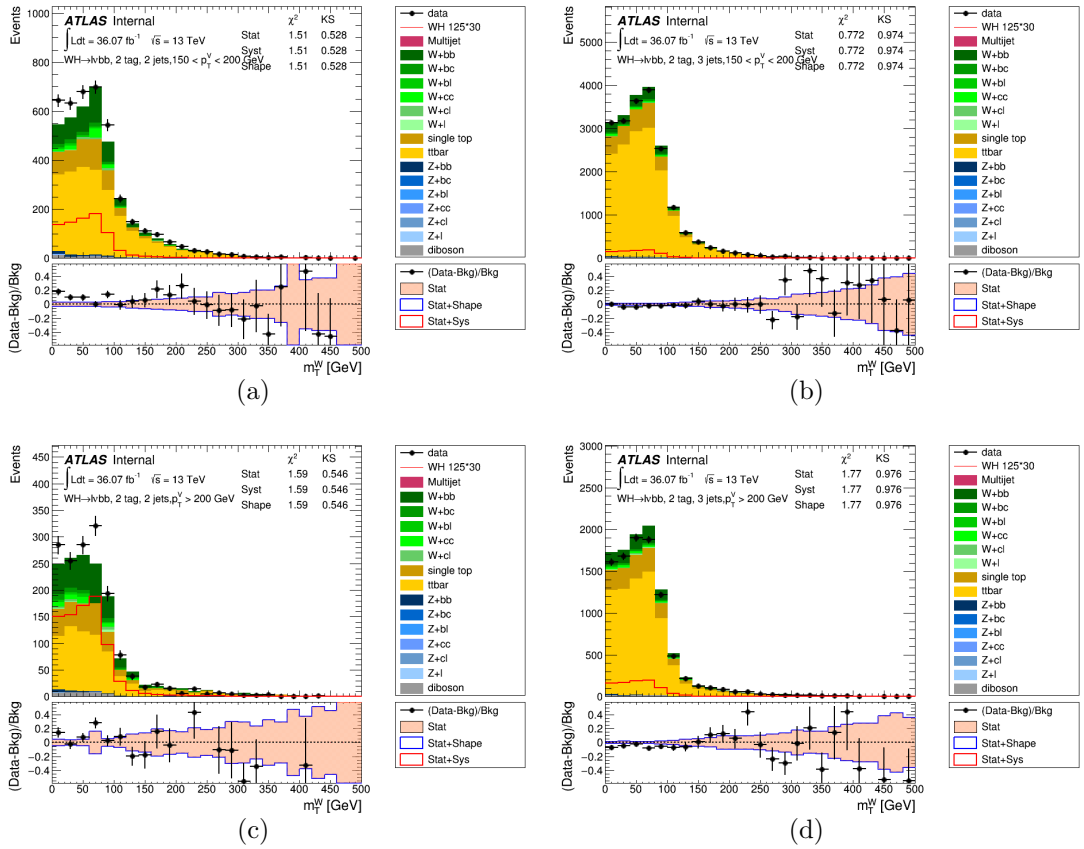


Figure 5.52: m_T^W distributions for data and all the MC simulations with only MVA selection applied in different dijet-mass signal regions : 2-jet 150 GeV $< p_T^V <$ 200 GeV (a), 3-jet 150 GeV $< p_T^V <$ 200 GeV (b), 2-jet $p_T^V >$ 200 GeV (c), 3-jet $p_T^V >$ 200 GeV (d). no normalization factors are applied to the MC events.

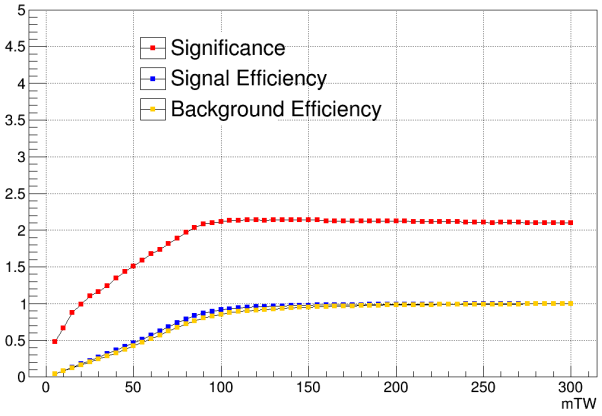


Figure 5.53: The m_T^W cut value scan results, the blue dots represent the signal efficiency, the yellow dots represent the background efficiency, while the red dots represent the value of the significance.

background with MVA selections and $m_T^W < 120\,\text{GeV}$ cut applied in different dijet-mass signal regions, the ΔR_{bb} distributions have been scaled to the same (unit) area in order to highlight the shape differences. Figure 5.55 shows the data and MC simulations comparison in the same regions with the same selection applied. As shown in Figure 5.56 and Figure 5.57, the cut value scan is performed to the ΔR_{bb} distribution from leftmost to rightmost with a step of 0.1. For every cut value, new m_{bb} distributions are built in different dijet-mass signal regions to calculate the sensitivity (S) in the m_{bb} range of 30 GeV to 200 GeV with 10GeV/bin. Figure 5.56 shows the scan result in 150 GeV $< p_T^V <$ 200 GeV region, while Figure 5.57 shows the scan result in $p_T^V >$ 200 GeV region. Figure 5.58 shows the 2-dimension (2D) scan results that combining these two p_T^V regions, the x axis represents the ΔR_{bb} cut in the 150 GeV $< p_T^V <$ 200 GeV region, and the y axis represents the ΔR_{bb} cut in $p_T^V >$ 200 GeV region. The numbers in the plot represents the combined significance by adding the significance calculated in two different p_T^V regions in quadrature. As shown in the plot, the upper ΔR_{bb} cut between 1.3 and 1.8 in 150 GeV $< p_T^V <$ 200 GeV region, and upper ΔR_{bb} cut between1.2 and 1.4 in $p_T^V >$ 200 GeV region constitute the highest significance region (in red). The combination of 1.5 and 1.2 yields the best significance 2.49 compared to significance 2.45 with the default ΔR_{bb} cuts.

Apart from the ΔR_{bb} cuts optimization in the default p_T^V categorization. The p_T^V categorization itself is also optimized with additional split at $p_T^V = 250\,\text{GeV}$. The choice of 250 GeV is also in order to fit the bins of Simplified Template Cross Sections (STXS) framework [131]. Two additional p_T^V categorizations are tested:

- Set1 [150 GeV - 250 GeV], [250 GeV - ∞]
- Set2 [150 GeV - 200 GeV], [200 GeV - 250 GeV], [250 GeV - ∞]

The same ΔR_{bb} cut value scan method is also performed in the different p_T^V categories. Table 5.30 summaries the results from the different p_T^V categories. As can be see, the ΔR_{bb} cuts which yield the best significance in Set1 are $\Delta R_{bb} < 1.4$ and $\Delta R_{bb} < 1.2$ in the corresponding p_T^V categories, compared with the default p_T^V categories with optimized ΔR_{bb} cuts, the increase of the significance is about 2.8%. For Set2, the ΔR_{bb} cuts which yield the best significance in are $\Delta R_{bb} < 1.5$, $\Delta R_{bb} < 1.4$ and $\Delta R_{bb} < 1.2$ in the corresponding p_T^V categories, compared with the default p_T^V categories with optimized ΔR_{bb} cuts, the increase of the significance is about 3.6%.

In conclusion, considering only the significance, in the default p_T^V categories, the upper ΔR_{bb} cuts 1.5 and 1.2 yield the best significance, 2.49, compared with

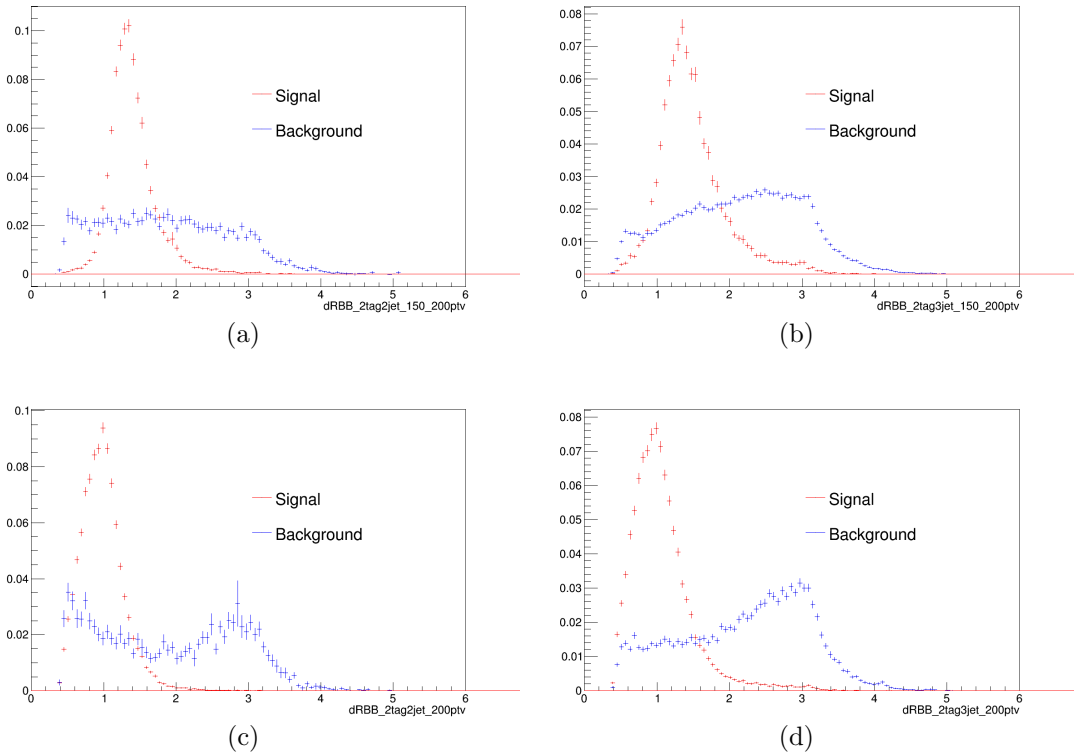


Figure 5.54: ΔR_{bb} distributions comparison for signal and total background with only MVA selection and $m_T^W < 120GeV$ cut applied in different dijet-mass signal regions : 2-jet 150 GeV $< p_T^V <$ 200 GeV (a), 3-jet 150 GeV $< p_T^V <$ 200 GeV (b), 2-jet $p_T^V >$ 200 GeV (c), 3-jet $p_T^V >$ 200 GeV (d). The ΔR_{bb} distributions have been scaled to the same (unit) area in order to highlight the shape differences.

Table 5.30: Summary of the upper ΔR_{bb} cut results in different p_T^V categories.

	p_T^V categories	Optimized upper ΔR_{bb}	Significance
Default	[150 GeV - 200 GeV], [200 GeV - ∞]	1.5; 1.2	2.49
Set1	[150 GeV - 250 GeV], [250 GeV - ∞]	1.4; 1.2	2.56
Set2	[150 GeV - 200 GeV], [200 GeV - 250 GeV], [250 GeV - ∞]	1.5; 1.4; 1.2	2.58

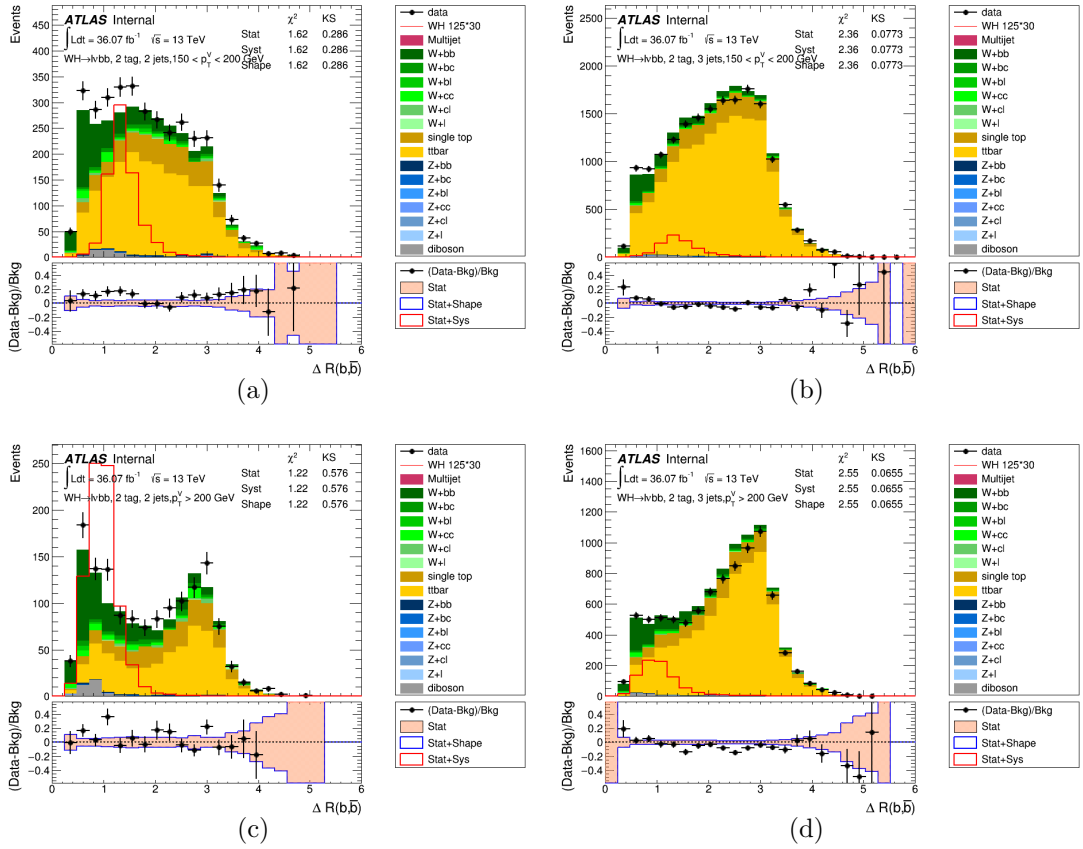


Figure 5.55: ΔR_{bb} distributions for data and all the MC simulations with only MVA selection and $m_T^W < 120GeV$ cut applied in different dijet-mass signal regions : 2-jet 150 GeV $< p_T^V <$ 200 GeV (a), 3-jet 150 GeV $< p_T^V <$ 200 GeV (b), 2-jet $p_T^V >$ 200 GeV (c), 3-jet $p_T^V >$ 200 GeV (d). no normalization factors are applied to the MC events.

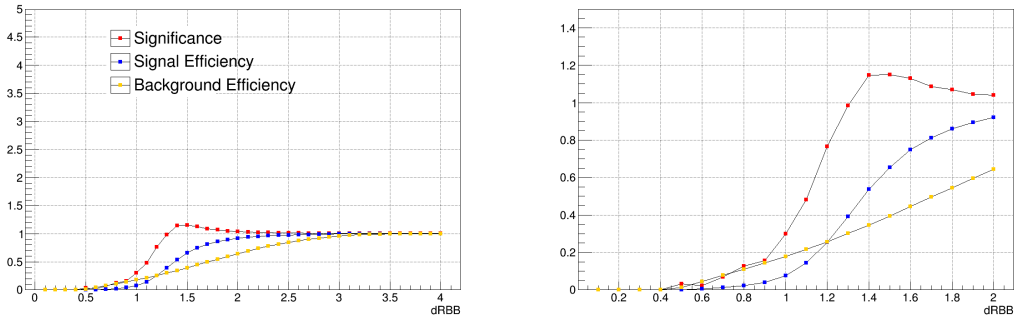


Figure 5.56: The ΔR_{bb} cut value scan results in 150 GeV $< p_T^V <$ 200 GeV region, the blue dots represent the signal efficiency, the yellow dots represent the background efficiency, while the red dots represent the value of the significance. (b) is the zoomed plot from (a) to show only the results in the range of $0 < \Delta R_{bb} < 2.0$.

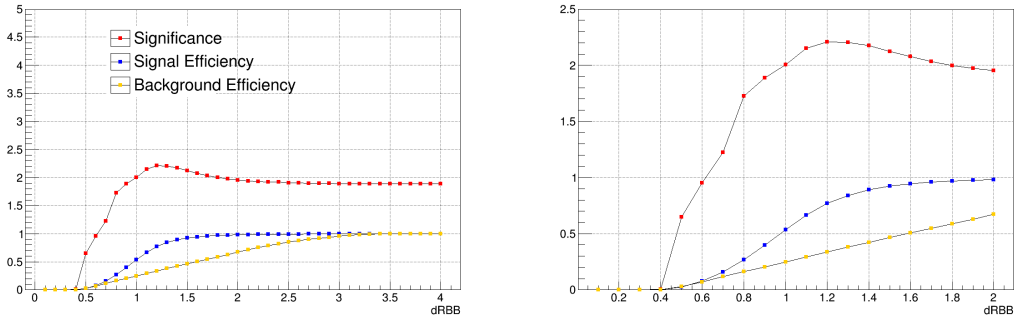


Figure 5.57: The ΔR_{bb} cut value scan results in $p_T^V > 200$ GeV region, the blue dots represent the signal efficiency, the yellow dots represent the background efficiency, while the red dots represent the value of the significance. (b) is the zoomed plot from (a) to show only the results in the range of $0 < \Delta R_{bb} < 2.0$.

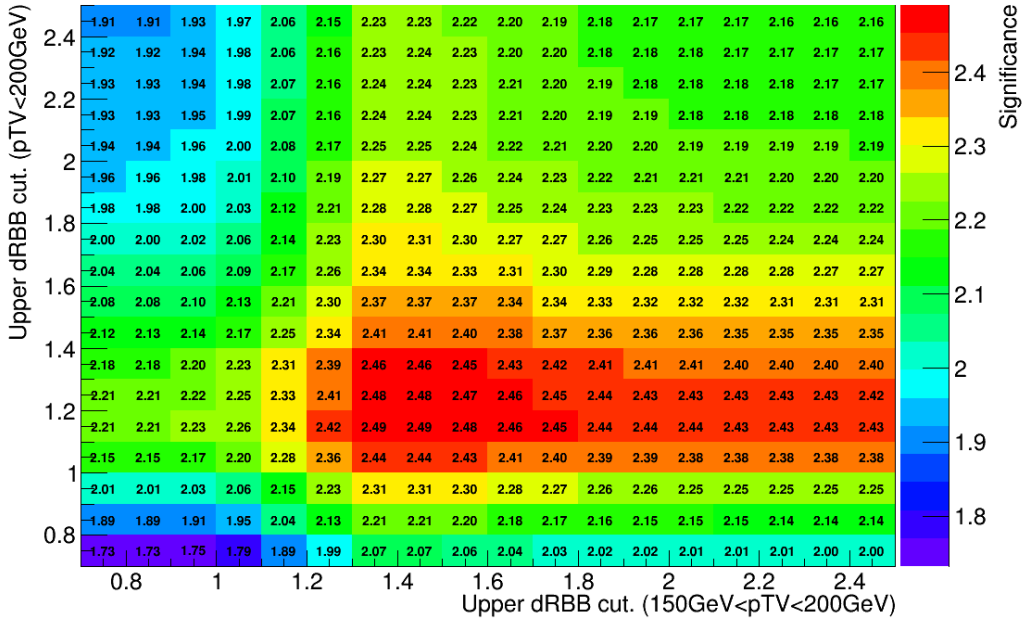


Figure 5.58: 2D ΔR_{bb} scan results, the x axis represents the ΔR_{bb} cut in the 150 GeV $< p_T^V <$ 200 GeV region, and the y axis represents the ΔR_{bb} cut in $p_T^V >$ 200 GeV region. The numbers in the plot represents the combined significance by adding the significance calculated in two different p_T^V regions in quadrature.

significance (2.45) from the the default ΔR_{bb} cuts (1.8 and 1.2), about 1.6% increase of the significance is achieved. For the different p_T^V categories optimization, the categories of [150 GeV - 200 GeV], [200 GeV - 250 GeV], [250 GeV - ∞] yields the best significance of 2.58, with the optimized upper ΔR_{bb} cuts 1.5, 1.4 and 1.2. This results about 3.6% increase of significance compared with the default p_T^V categories with optimized ΔR_{bb} cuts, and about 5.3% increase of significance compared with the default p_T^V categories and default ΔR_{bb} cuts.

By default, there are no lower ΔR_{bb} cuts used in the dijet-mass analysis, in this study, the optimization of the lower ΔR_{bb} cut is also performed, and the result confirms that there is no need to introduce the lower ΔR_{bb} cuts in the analysis.

This study is based on only the significance, for the safety of the signal efficiency, such optimized ΔR_{bb} cuts may still need to loosen a bit. In the other hand, apart from the multijet background, all the other signal and background modelling uncertainties use in the dijet-mass analysis are inherited from the multivariate analysis. Due to the different phase space, such uncertainties may not accurate in the dijet-mass analysis, especially for the p_T^V and m_{bb} shape uncertainties. To make the final decision, the dedicated stduies of the dijet-mass analysis modelling uncertainties are necessarily to be performed, and the extensive fit studies are also needed. This preliminary cut optimization study shows clearly a direction to consider to improve the sensitivity and robustness of the dijet-mass analysis in the next round of the analysis.

5.7 Systematic Uncertainties

In this section, the systematics uncertainties that are considered in this analysis are summarized. The sources of these systematic uncertainty can be roughly divided into four groups: those of experimental nature, those related to the modelling of the simulated backgrounds, those associated with the Higgs boson signal simulation, and those related to the estimation of multi-jet background. The last one has been discussed in Section 5.5.2. In the following, Section 5.7.1 presents the summary of the experimental systematic uncertainties, while Section 5.7.2 and Section 5.7.3 present the summary of simulated background and signal modelling systematic uncertainties, respectively.

5.7.1 Experimental systematic uncertainties

Several sources of experimental systematic uncertainties are considered in this analysis, as outlined below and summarized in Table 5.31.

Luminosity: The luminosity uncertainty [92] is 2.1%, 2.6%, 2.4 % for 2015 data (3.2 fb^{-1}), 2016 data (32.9 fb^{-1}) and 2017 data (43.6 fb^{-1}), respectively. The total uncertainty for the combined 2015-2017 dataset is 2.0%.

Pile-up reweighting: The pile-up weight is applied to MC events to correct the pile-up difference between MC and data. This weight is calculated with the distribution of average number of interactions per bunch crossing (μ). Before calculating the pile-up weight, the agreement of μ between data and MC can be improved by scaling the μ of MC by a measured factor of 1.03 [132]. Due to the discrete nature of the values of μ used in MC, it is more practical to scale the μ in data (which is a continuous variable) by the inverse scale factor, 1/1.03. The factor is measured with an uncertainty. The pile-up reweighting uncertainty is then derived by recalculating the pile-up weight by using the $\pm 1\sigma$ values of the nominal factor (1 and 1/1.08).

Lepton: Uncertainties on the efficiencies of lepton trigger, reconstruction, identification and isolation are considered, along with the uncertainties on the lepton energy scale and resolution, for both electrons [77] and muons [78]. These uncertainties are found to have only very small effect on the final result.

Jets: The most prominent sources of jet-related uncertainty are the uncertainties from the jet energy scale (JES) and jet energy resolution (JER) [82, 133]. These uncertainties are also the one of the dominant experimental uncertainties in the analysis. The many sources of JES uncertainties are decomposed into 23 uncorrelated components which are treated as independent. An additional specific uncertainty on the energy calibration of b- and c-jets, along with the uncertainty on JVT, are considered.

E_T^{miss}: The uncertainties in the resolution and energy scale of the leptons and jets are propagated to the calculation of E_T^{miss}, The additional uncertainties in the E_T^{miss} come from also the resolution, scale and reconstruction efficiency of the tracks used to compute the soft term [87], as well as the modelling of the underlying event. E_T^{miss} trigger scale factors are derived by using the $W(\mu\nu)$ + jets events. Uncertainties on these scale factors are derived by taking account for the statistical fluctuations in their determination, differences in their measurement with alternative physics processes (for example $t\bar{t}$), and the kinematic dependence of E_T^{miss} trigger efficiency on the offline scalar sum of all final state jets within the

event.

Flavor-tagging: The uncertainties come from the b-tagging MC simulation to data efficiency correction factors are also one of the dominant experimental uncertainties in the analysis. These correction factors are derived for b-jets, c-jets and light-flavour jets separately. All three correction factors depend on jet p_T (or p_T and $|\eta|$) and have uncertainties estimated from many sources. These are decomposed into uncorrelated components which are then treated independently, resulting in three uncertainties for b-jets and for c-jets, and five for light- flavour jets [134–136]. The approximate size of the uncertainty in the tagging efficiency is 2% for b-jets, 10% for c-jets and 40% for light jets. Additional uncertainties are considered in the extrapolation of the b-jet efficiency calibration to jets with p_T above 300 GeV and in the misidentification of hadronically decaying τ lepton as b-jets.

5.7.2 Simulated background uncertainties

Three areas are broadly covered by the simulated backgrounds modelling uncertainties: normalization, acceptance differences that affect the relative normalization between analysis regions with a common background normalization, and the differential distributions of the most important kinematic variables. These uncertainties are derived either from particle-level comparisons between nominal and alternative samples, or from comparisons to data in control regions. Detector-level simulation comparisons whenever these are available are used as cross check and good agreement is found compared with the particle-level comparisons. Acceptance uncertainties are estimated by normalizing all the nominal and alternative samples to the same production cross-section. The size of these uncertainties are derived by adding the differences between the nominal and alternative samples in quadrature. Shape uncertainties are estimated by scaling all the nominal and alternative samples to the same normalization. These uncertainties are considered in each of the analysis regions separately. The shape differences between each alternative generator and nominal sample are compared and the largest one is taken as the shape uncertainty. Shape uncertainties are derived only for the m_{bb} and p_T^V distributions, as these two variables are the highest ranked variables in the VH BDT training and have only very weak correlation. It was also found that the overall shape variation of BDT discriminant can be covered by considering only the changes induced in these two variables by an alternative generator. The simulated backgrounds modelling systematic uncertainties considered in the

Table 5.31: Summary of the experimental sytematic uncertainties considered in the analysis.

Systematic uncertainty	Short description
Event	
Luminosity	uncertainty on total integrated luminosity
Pileup Reweighting	uncertainty on pileup reweighting
Electrons	
EL_EFF_Trigger_Total_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty
EL_EFF_Reco_Total_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty
EL_EFF_ID_Total_1NPCOR_PLUS_UNCOR	ID efficiency uncertainty
EL_EFF_Iso_Total_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty
EG_SCALE_ALL	energy scale uncertainty
EG_RESOLUTION_ALL	energy resolution uncertainty
Muons	
MUON_EFF_TrigStatUncertainty	trigger efficiency uncertainty
MUON_EFF_TrigSystUncertainty	
MUON_EFF_RECO_STAT	reconstruction and ID efficiency uncertainty for muons with $p_{\mathrm{T}} > 15$ GeV
MUON_EFF_RECO_SYS	
MUON_EFF_RECO_STAT_LOWPT	reconstruction and ID efficiency uncertainty for muons with $p_{\mathrm{T}} < 15$ GeV
MUON_EFF_RECO_SYST_LOWPT	
MUON_ISO_STAT	isolation efficiency uncertainty
MUON_ISO_SYS	
MUON_TTVA_STAT	track-to-vertex association efficiency uncertainty
MUON_TTVA_SYS	
MUON_ID	momentum resolution uncertainty from inner detector
MUON_MS	momentum resolution uncertainty from muon system
MUON_SCALE	momentum scale uncertainty
MUON_SAGITTA_RHO	charge dependent momentum scale uncertainty
MUON_SAGITTA_RESBIAS	
Jets	
JET_23NP_JET_EffectiveNP_1	energy scale uncertainty from the in situ analyses splits into 8 components
JET_23NP_JET_EffectiveNP_2	energy scale uncertainty from the in situ analyses splits into 8 components
JET_23NP_JET_EffectiveNP_3	energy scale uncertainty from the in situ analyses splits into 8 components
JET_23NP_JET_EffectiveNP_4	energy scale uncertainty from the in situ analyses splits into 8 components
JET_23NP_JET_EffectiveNP_5	energy scale uncertainty from the in situ analyses splits into 8 components
JET_23NP_JET_EffectiveNP_6	energy scale uncertainty from the in situ analyses splits into 8 components
JET_23NP_JET_EffectiveNP_7	energy scale uncertainty from the in situ analyses splits into 8 components
JET_23NP_JET_EffectiveNP_8restTerm	energy scale uncertainty from the in situ analyses splits into 8 components
JET_23NP_JET_EtaIntercalibration_Modeling	energy scale uncertainty on eta-intercalibration (modeling)
JET_23NP_JET_EtaIntercalibration_TotalStat	energy scale uncertainty on eta-intercalibrations (statistics/method)
JET_23NP_JET_EtaIntercalibration_NonClosure_highE	energy scale uncertainty on eta-intercalibrations (non-closure)
JET_23NP_JET_EtaIntercalibration_NonClosure_negEta	
JET_23NP_JET_EtaIntercalibration_NonClosure_posEta	
JET_23NP_JET_Flavor_Composition	energy scale uncertainty on VV and VH sample's flavour composition
JET_23NP_JET_Flavor_Response	energy scale uncertainty on samples' flavour response
JET_23NP_JET_Pileup_OffsetNPV	energy scale uncertainty on pile-up (NPV dependent)
JET_23NP_JET_Pileup_PtTerm	energy scale uncertainty on pile-up (pt term)
JET_23NP_JET_Pileup_RhoTopology	energy scale uncertainty on pile-up (density ρ)
JET_23NP_JET_PunchThrough_MC16	energy scale uncertainty for punch-through jets
JET_23NP_JET_SingleParticle_HighPt	energy scale uncertainty from the behaviour of high-pT jets
JET_JER_SINGLE_NP	energy resolution uncertainty
JET_SR1_JET_EtaIntercalibration_NonClosure	
JET_SR1_JET_GroupedNP_1	
JET_SR1_JET_GroupedNP_2	
JET_SR1_JET_GroupedNP_3	
JET_JvtEfficiency	JVT efficiency uncertainty
FT_EFF_Eigen_B0	
FT_EFF_Eigen_B1	
FT_EFF_Eigen_B2	
FT_EFF_Eigen_C0	
FT_EFF_Eigen_C1	b-tagging efficiency uncertainties: 3 components for b jets, 3
FT_EFF_Eigen_C2	for c jets and 5 for light jets
FT_EFF_Eigen_L0	
FT_EFF_Eigen_L1	
FT_EFF_Eigen_L2	
FT_EFF_Eigen_L3	
FT_EFF_Eigen_L4	
FT_EFF_Eigen_extrapolation	b-tagging efficiency uncertainty on the extrapolation to high-p_{T} jets
FT_EFF_Eigen_extrapolation_from_charm	b-tagging efficiency uncertainty on tau jets
MET	
METTrigStat	trigger efficiency uncertainty
METTrigTop/Z	
MET_SoftTrk_ResoPara	track-based soft term related longitudinal resolution uncertainty
MET_SoftTrk_ResoPerp	track-based soft term related transverse resolution uncertainty
MET_SoftTrk_Scale	track-based soft term related longitudinal scale uncertainty
MET_JetTrk_Scale	track MET scale uncertainty due to tracks in jets

analysis are summarized in Table 5.32 and Table 5.33, and the key details of how the uncertainties are estimated are reported below for each simulated background sample.

$t\bar{t}$ production. $t\bar{t}$ is a dominant background in all three channels. The acceptance and shape uncertainties are derived from comparing the nominal sample (Powheg+Pythia8) to the alternative samples with different matrix-element generation (MadGraph5_aMC@NLO+Pythia8), parton-shower generation (Powheg+Herwig7) and settings of the nominal generator designed to increase or decrease the amount of radiation. Due to the clearly different regions of phase space probed, the characteristics of $t\bar{t}$ in 0- and 1-lepton channel ((jointly referred to as 0+1 lepton in the following) are quite different to that in 2-lepton channel. For $t\bar{t}$ events in 0+1 lepton, some of the objects from $t\bar{t}$ decay have been missed and not reconstructed. While most of $t\bar{t}$ events in 2-lepton channel undergo the di-leptonic decay and can be fully reconstructed. Therefore different overall floating normalization factors are considered for 0+1 lepton and 2-lepton channels, and acceptance uncertainties are derived separately and taken as uncorrelated between the 0+1 and 2-lepton channels. For the 0+1 lepton channels, the 1-lepton channel 3-jet region provides the main constraint of $t\bar{t}$ normalization due to the quite high $t\bar{t}$ purity ($> 75\%$) in that region. Two extrapolation uncertainties are then applied in the 2-jet region (2-to-3-jet ratio) and 0-lepton region (0-to-1-lepton ratio) separately, by considering the normalization ratios of these regions. An additional acceptance uncertainty is considered in the normalization ratio of W + HF control region and signal region. These uncertainties are estimated as double ratio

$$\frac{Acceptance[Region_A(nominalMC)]}{Acceptance[Region_B(nominalMC)]} \Big/ \frac{Acceptance[Region_A(alternativeMC)]}{Acceptance[Region_B(alternativeMC)]}. \quad (5.9)$$

The differences between the nominal and each of the alternative samples are summed in quadrature to provide an overall uncertainty. For the 2-lepton channel, due to the powerful constraint of $t\bar{t}$ normalization provided by the Top $e\mu$ control region in 2 and 3+-jet regions, two floating normalization factors are used separately in these two regions. Shape uncertainties are also estimated separately in 0+1 and 2-lepton channels. The difference between the nominal sample and MadGraph5_aMC@NLO+Pythia8 sample results the largest variation, and is therefore considered as the shape systematic uncertainty.

V + jets production. The V + jets backgrounds are divided into three different components based on the jet flavour labels of the two b-tagged jets in

Table 5.32: Summary of the systematic uncertainties in the background modelling for Z + jets, W + jets, $t\bar{t}$, single top quark and multi-jet production. An "S" symbol is used when only a shape uncertainty is assessed. The regions for which the normalisations float independently are listed in brackets.

Z + jets	
$Z + ll$ normalisation	18%
$Z + cl$ normalisation	23%
$Z + bb$ normalisation	Floating (2-jet, 3-jet)
$Z + bc$-to-$Z + bb$ ratio	30 – 40%
$Z + cc$-to-$Z + bb$ ratio	13 – 15%
$Z + bl$-to-$Z + bb$ ratio	20 – 25%
0-to-2 lepton ratio	7%
m_{bb}, p_T^V	S
W + jets	
$W + ll$ normalisation	32%
$W + cl$ normalisation	37%
$W + bb$ normalisation	Floating (2-jet, 3-jet)
$W + bl$-to-$W + bb$ ratio	26% (0-lepton) and 23% (1-lepton)
$W + bc$-to-$W + bb$ ratio	15% (0-lepton) and 30% (1-lepton)
$W + cc$-to-$W + bb$ ratio	10% (0-lepton) and 30% (1-lepton)
0-to-1 lepton ratio	5%
W + HF CR to SR ratio	10% (1-lepton)
m_{bb}, p_T^V	S
$t\bar{t}$ (all are uncorrelated between the 0+1 and 2-lepton channels)	
$t\bar{t}$ normalisation	Floating (0+1 lepton, 2-lepton 2-jet, 2-lepton 3-jet)
0-to-1 lepton ratio	8%
2-to-3-jet ratio	9% (0+1 lepton only)
W + HF CR to SR ratio	25%
m_{bb}, p_T^V	S
Single top quark	
Cross-section	4.6% (s-channel), 4.4% (t-channel), 6.2% (Wt)
Acceptance 2-jet	17% (t-channel), 55% ($Wt \to bb$), 24% ($Wt \to oth$)
Acceptance 3-jet	20% (t-channel), 51% ($Wt \to bb$), 21% ($Wt \to oth$)
m_{bb}, p_T^V	S (t-channel, $Wt \to bb$, $Wt \to oth$)
Multi-jet (1-lepton)	
Normalisation	60 – 100% (2-jet), 90 – 140% (3-jet)
BDT template	S

Table 5.33: Summary of the systematic uncertainties in the background modelling for diboson production. "PS/UE" indicates parton shower / underlying event. An "S" symbol is used when only a shape uncertainty is assessed. When determining the $(W/Z)Z$ diboson production signal strength, the normalisation uncertainties in ZZ and WZ production are removed.

ZZ	
Normalisation	20%
0-to-2 lepton ratio	6%
Acceptance from scale variations (var.)	10 – 18% (Stewart–Tackmann jet binning method)
Acceptance from PS/UE var. for 2 or more jets	5.6% (0-lepton), 5.8% (2-lepton)
Acceptance from PS/UE var. for 3 jets	7.3% (0-lepton), 3.1% (2-lepton)
m_{bb}, p_T^V, from scale var.	S (correlated with WZ uncertainties)
m_{bb}, p_T^V, from PS/UE var.	S (correlated with WZ uncertainties)
m_{bb}, from matrix-element var.	S (correlated with WZ uncertainties)
WZ	
Normalisation	26%
0-to-1 lepton ratio	11%
Acceptance from scale var.	13 – 21% (Stewart–Tackmann jet binning method)
Acceptance from PS/UE var. for 2 or more jets	3.9%
Acceptance from PS/UE var. for 3 jets	11%
m_{bb}, p_T^V, from scale var.	S (correlated with ZZ uncertainties)
m_{bb}, p_T^V, from PS/UE var.	S (correlated with ZZ uncertainties)
m_{bb}, from matrix-element var.	S (correlated with ZZ uncertainties)
WW	
Normalisation	25%

the event. The main background contributions (V + bb, V + bc, V + bl and V + cc) are jointly considered as the V + HF background. W + HF is a dominate background in the 0- and 1- lepton channels, while the Z + HF is a dominate background in the 0- and 2-lepton channels. Their overall normalization, separately in the 2- and 3-jet regions, is free to float in the global likelihood fit. The remaining flavour components, V + cl and V + ll, constitute less than 1% of the total background in each region, and therefore only uncertainties in the normalization of these backgrounds are considered. Acceptance uncertainties are estimated for the relative normalizations of the different regions that share a common floating normalization factor. For W + HF background, the 1-lepton W + HF control region provides the best constraint of the normalization. Two extrapolation uncertainties are then applied in the 1-lepton signal region (W + HF CR to SR ratio) and 0-lepton region (0-to-1-lepton ratio) separately, by considering the normalization ratios of these regions. For Z + HF background, the 2-lepton channel provides the best constraint of the normalization, extrapolation uncertainty is then applied in the 0-lepton channel (0-to-2-lepton ratio).

Uncertainties are also considered in the relative normalization of the four heavy-flavour components that constitute the V + HF background. These uncertainties are estimated by comparing the bc, cc and bl yields to the dominant bb yields, and are estimated separately for the 0- and 1-lepton channels for the W + HF backgrounds and separately for the 0-lepton, 2-lepton 2-jet and 2-lepton 3-jet regions for the Z + HF background.

The acceptance and normalization uncertainties are all calculated by adding the differences between the nominal Sherpa 2.2.1 sample and its associated systematic variations in quadrature, including a variation of:

- the renormalisation scale by factors of 0.5 and 2.
- the factorisation scale by factors of 0.5 and 2.
- the CKKW merging scale from 30 GeV to 15 GeV.
- the parton-shower/resummation scale by factors of 0.5 and 2.
- alternative sample produced with MadGraph+Pythia8.

For Z + HF background, shape uncertainties are estimated by comparing the Z + jets background to data in control regions with high Z + jets purity and depleted signal. These control regions are defined in the 2-lepton channel 1- and 2-tag regions, with the m_{bb} region around the Higgs boson mass excluded in the

2-tag case. The E_T^{miss} significance cut ($E_T^{miss}/\sqrt{S_T} < 3.5\sqrt{\text{GeV}}$) used in the dijet mass analysis is also required to remove the residual $t\bar{t}$ contamination. For the W + HF background, shape uncertainties are estimated with the same systematic uncertainty sources as for the acceptance and normalization and uncertainties. The W + HF control region is not used due to the limited number of data events.

Single top production. Uncertainties are derived in the normalization, acceptance and shape in the Wt and t-channels. In the s-channel, only normalization uncertainties are considered since the overall negligible contribution. The normalization uncertainties are taken from the variations of the renormalization and factorization scales, αs and PDFs. For Wt and t-channels, the nominal samples (Powheg+Pythia8) are compared to alternative samples, which are similar to those used in the $t\bar{t}$ case, to derive the acceptance and shape uncertainties.

- Alternative matrix element generation (MadGraph5_aMC@NLO+Herwig++).
- Alternative parton shower generation (Powheg+Herwig++).
- Nominal samples with increased and reduced radiation tunes

For Wt-channel, an additional uncertainty is considered to assess the interference between the Wt and $t\bar{t}$ production processes, by using a diagram subtraction scheme instead of the nominal diagram removal scheme. In additional, the modelling uncertainties for Wt channel are also based on the flavour of the two b-tagged jets, due to the different regions of phase space being probed when there are two b-jets (bb) present compared with events where there are fewer b-jets present (other).

Diboson production. For the WW production, only normalization uncertainty is assigned due to the overall negligible contribution. For the WZ and quark induced ZZ productions, uncertainties are derived in the normalization, acceptance and shapes of the m_{bb} and p_T^V distributions by comparing the nominal sample (*Sherpa*2.2.1) to the alternative samples with varied factorization, renormalization and resummation scales, and using the Stewart-Tackmann method to calculate scale variation uncertainties for the acceptance in the jet multiplicity categories. Additional uncertainties are estimated in the parton-shower and underlying-event model by considering the difference between Powheg+Pythia8 and Powheg+Herwig++, as well as changes in the Pythia8 parton-shower tune. A systematic uncertainty in the m_{bb} shape distribution results from the comparison of Sherpa 2.2.1 and Powheg+Pythia8. For the WZ production, Acceptance uncertainties are estimated for the ratio of

0-to-1 lepton channels and for the ratio of the 2-to-3 jet regions. For the ZZ production, acceptance uncertainties are estimated for the ratio of the 0-to-2 lepton channels and of the 2-to-3 jet regions. The semi-leptonic loop-induced ZZ productions use the same systematic uncertainties as those used for the quark induced ZZ productions in a correlated manner.

5.7.3 Signal uncertainties

The signal samples are normalized using their inclusive cross-sections. To correct the sizeable impact of the NLO (EW) corrections to the p_T^V distributions, an additional scale factor calculated using the HAWK generator is applied as a function of p_T^V. Table 5.34 summarize the systematic uncertainties considered for the modelling of the signal.

Table 5.34: Summary of the systematic uncertainties in the signal modelling. "PS/UE" indicates parton shower / underlying event. An "S" symbol is used when only a shape uncertainty is assessed.

Signal	
Cross-section (scale)	0.7% (qq), 27% (gg)
Cross-section (PDF)	1.9% ($qq \to WH$), 1.6% ($qq \to ZH$), 5% (gg)
Branching ratio	1.7 %
Acceptance from scale variations (var.)	2.5 – 8.8% (Stewart–Tackmann jet binning method)
Acceptance from PS/UE var. for 2 or more jets	2.9 – 6.2% (depending on lepton channel)
Acceptance from PS/UE var. for 3 jets	1.8 – 11%
Acceptance from PDF+αs var.	0.5 – 1.3%
m_{bb}, p_T^V, from scale var.	S
m_{bb}, p_T^V, from PS/UE var.	S
m_{bb}, p_T^V, from PDF+αs var.	S
p_T^V from NLO EW correction	S

Uncertainties in the calculations of the VH production cross-sections and the $H \to b\bar{b}$ branching ratio are assigned following the recommendations of the LHC Higgs Cross Section working group. The uncertainties in the overall VH production cross-section from missing higher-order terms in the QCD perturbative expansion are obtained by varying the renormalization scale μ_R and factorisation scale μ_F independently, from 1/3 to 3 times their original value. The PDF+αs uncertainty in the overall VH production cross-section is calculated from the 68% CL interval using the PDF4LHC15_nnlo_mc PDF set. The latest LHC Higgs working

group recommendations do not distinguish between the uncertainties in $qq \to ZH$ production and $gg \to ZH$ production. To get the scale uncertainties for these two processes separately, the uncertainty in $qq \to ZH$ production is assumed to be identical to the uncertainty in WH production. The $gg \to ZH$ production uncertainty is then derived such that the sum in quadrature of the $qq \to ZH$ and $gg \to ZH$ production uncertainties equal to the overall ZH production scale uncertainty. The PDF+αs uncertainty is known larger for WH production than the ZH production, therefore the method used for the scale uncertainty cannot be used for this uncertainty.

Systematic uncertainty in the overall VH cross-section that stems from missing higher-order EW corrections is estimated as the maximum variation among three quantities: the maximum size expected for the missing NNLO EW effects, the size of the NLO EW correction and the uncertainty in the photon induced cross-section relative to the total $(W/Z)H$ cross-section. The $H \to b\bar{b}$ branching ratio uncertainty is calculated by considering the missing higher-order QCD and EW corrections and the uncertainties in the b-quark mass and the value of αs.

Acceptance and shape uncertainties are estimated by comparing the nominal samples to those generated with weights corresponding to varied factorization and renormalization scales applied. The Stewart-Tackmann method is used to calculate the scale variation uncertainties in the acceptance in the jet multiplicity categories. Uncertainties that stem from the parton-shower and underlying-event models are estimated by considering the difference between Powheg+MiNLO+Pythia8 and Powheg+MiNLO+Herwig7, as well as changes in the Pythia8 parton-shower tune. The PDF+αs uncertainty in the acceptance between regions and in the m_{bb} and p_T^V shapes is estimated by applying the PDF4LHC15_30 PDF set and its uncertainties, according to the PDF4LHC recommendations.

5.8 Statistical Analysis

A global likelihood fit procedure [137] is performed to data in order to the extract the signal significance and strength. In this section, an overview of the global likelihood fit procedure is presented, along with descriptions of the key items related to this analysis.

5.8.1 Likelihood function

A likelihood function is obtained from the probability of data for a given certain hypothesis. In this analysis, the hypothesis is represented by the signal strength parameter, that is defined as the ratio of Higgs signal rate (SM Higgs boson production cross-section times branching ratio into $b\bar{b}$) to its SM prediction and can be expressed as:

$$\mu = \frac{\sigma \cdot BR}{\sigma_{SM} \cdot BR_{SM}}. \tag{5.10}$$

The binned likelihood function is defined as a product of Poisson probability terms, and can be expressed as:

$$\mathcal{L}(\mu) = \prod_{i=1}^{nbins} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)}, \tag{5.11}$$

when considering only the statistical uncertainties. Where nbins is the total number of bins, n_i is the observed number of data events in bin i, s_i (b_i) is the expected number of signal (background) events in bin i. As discussed in Section 5.7, a number of sources of systematic uncertainty are considered in this analysis and could have effect on the signal strength measurement, therefore a vector of nuisance parameters (NP), $\boldsymbol{\theta}$, is introduced to the likelihood function and allow for additional degrees of freedom in the likelihood. The likelihood function can be then modified as:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{nbins} \frac{(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{n_i}}{n_i!} e^{-(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))} \times \mathcal{L}_{AUX}(\boldsymbol{\theta}). \tag{5.12}$$

Each systematic uncertainty θ_i corresponds to an element of $\boldsymbol{\theta}$, and $\mathcal{L}_{AUX}(\boldsymbol{\theta})$ is the Gaussian or log-normal probability density functions of the prior uncertainty on each NP θ, the latter one being used for normalisation uncertainties to prevent normalisation factors from becoming negative in the fit. For example, as shown in Table 5.33, the $t\bar{t}$ 0-to-1 lepton ratio is one of the NPs, with a 8% prior uncertainty derived from physics studies shown in Section 5.7. The priors and the auxiliary function play a critical role to constrain the NPs within their uncertainties by penalizing large deviations in the likelihood. The floating NPs, such as $t\bar{t}$ normalization uncertainty, have no prior uncertainty and therefore no such auxiliary likelihood function assigned. The statistical uncertainties of simulated MC events are introduced through one nuisance parameter per bin, using the Beeston-Barlow technique [138].

5.8.2 Test statistics

The likelihood function ratio can be defined as:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}, \tag{5.13}$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the parameters that maximise the likelihood, and $\hat{\hat{\boldsymbol{\theta}}}$ is the value of $\boldsymbol{\theta}$ that maximise the likelihood for a given μ value. From the definition, it is clear that $0 \leq \lambda(\mu) \leq 1$ and $\lambda(\mu) \ll 1$ corresponds to a bad agreement between data and the given value of μ, the test statistic used in this analysis is then defined as:

$$t_\mu = -2ln\lambda(\mu), \tag{5.14}$$

the higher values of t_μ indicate the larger incompatibility between the data and the given value of μ. This test statistics is introduced to test the background-only hypothesis with $\mu = 0$ against the alternative hypotheses that μ is assumed to be positive. Rejecting such background-only hypothesis then leads to the discovery of a signal.

$$t_0 = \begin{cases} -2ln\frac{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{5.15}$$

The requirement of $\hat{\mu} \geq 0$ indicates that data are considered in disagreement with the background only hypothesis only if an non-negative signal strength fluctuation is observed, the negative $\hat{\mu}$ may also indicates some evidence against the background-only model but does not show that the measured data contain signal events. The higher values of t_0 indicates the larger incompatibility between the data and the background-only hypothesis. This incompatibility can be expressed with a p-value and can be defined as:

$$p_0 = \int_{t_{0,obs}}^{\infty} f(t_0)|0)dt_0, \tag{5.16}$$

where $t_{0,obs}$ is the measured value of the test statistic from the data, and $f(t_0)|0)$ is the probability distribution function of the test statistic itself, under background-only hypothesis. A small p_0 value therefore corresponds to a low false positive probability. p_0 can be also converted into standard deviations of the Gaussian distribution using the normal inverse cumulative distribution function

$$Z = \Phi^{-1}(1 - p_0). \tag{5.17}$$

A p_0 value of 1.35×10^{-4} corresponds to a 3 σ deviation from the background-only hypothesis, a p_0 value of 2.87×10^{-7} corresponds to a 5 σ deviation from the background-only hypothesis. In the context of high energy physics, 3 σ deviation is requested to claim the evidence of a new signal and 5 σ derivation is used as the benchmark deviation required for the discovery for a new signal. The expected significance quoted in this analysis are obtained in the same way as the observed results by replacing the data in each input bin by the prediction from simulation with all NPs set to their best-fit values, as obtained from the fit to the data, except for the signal strength parameter, which is kept at its nominal value.

5.8.3 Uncertainty on signal strength

The fitted $\hat{\mu}$ value is obtained by maximizing the likelihood function with respect to all parameters. The uncertainty on $\hat{\mu}$ is determined through a scan of the likelihood function values as a function of μ. The $\pm 1\sigma$ uncertainty on $\hat{\mu}$ is defined by determining the points in which the logarithm of the likelihood increases (decreases) by 1/2 with respect to the maximum value. In this analysis, there are two methods used to determine the effect of each NP, on the $\hat{\mu}$ measurement. First is the breakdown method that redo the likelihood fit and evaluate the uncertainty on μ without a systematic (or group of systematics) uncertainty, and subtract the resulting uncertainty quadratically from the full uncertainty. Second is the ranking method that fix the corresponding individual NP to its fitted value modified upwards or downwards by its fitted uncertainty, and perform the fit again, with all the other parameters allowed to vary to extract the new fitted μ value, the different between the $\hat{\mu}$ and new fitted μ value is taken as the effect of the individual NP on the $\hat{\mu}$ measurement.

5.8.4 Asimov dataset

Before performing the fit to the real data events, it is always very useful to construct a representative dataset, the Asimov dataset, from the MC simulation to check the expected performance of the fit. The Asimov dataset corresponds to the nominal simulated dataset, therefore when performing the fit to the Asimov dataset, all NPs should remain at their nominal value, but it is possible to check the constraints and correlations of the NPs. When observably differences are

found between fit to Asimov dataset and to real data for the pulls, constraints and correlations of the NPs, their sources are investigated and the fit model may changed by, for example, introducing additional NPs, in order to providing the fit with the degrees of freedom required to adjust the MC expectation to the observed data. The Asimov dataset is also very useful to tune and optimize the analysis based on the MC simulation.

5.8.5 Treatment of the nuisance parameters in the likelihood fit

5.8.5.1 Correlation of the Systematic Uncertainties

It is possible to decide to correlate or uncorrelate NPs in the fit model. It is clear that the NPs related to the different systematics effects need to be treated as uncorrelated, such as the NPs for the $t\bar{t}$ extrapolation uncertainties and b-tagging related uncertainties. To correlate two NPs is equivalent to the assumption that information on one of them can affect the other, it is important to study such correlations case by case, since one of the NP may be strongly constrained by the likelihood fit and propagate this strong constraint to the second NP, and causing potential bias in the final result. In the other hand, keeping the NPs uncorrelated may represent a more conservative approach, since it increases the number of degrees of freedom in the fit. It is also important to study the correlation behavior between NPs in the likelihood fit, to understand if a pull in one NP is related to the other NPs and make sure the fit model is reasonable. To evaluate the correlation between NPs, the Hessian matrix is constructed first, the correlation matrix is then extracted from the covariance matrix which is obtained from the inversion of the Hessian matrix.

5.8.5.2 Smoothing of the Systematic Uncertainties

Shape uncertainties are implemented in the likelihood fit as alternative templates for the discriminating variable relative to the nominal prediction, therefore can be suffered from statistical fluctuation in the simulation. The shape uncertainties are propagated in the analysis in two different ways: by shifting weights or by re-selecting events. An example of the former case is the b-tagging efficiency uncertainty , where a scale factor is used to correct the simulation efficiency to data, this weight is shifted up (down) and the change in the final distribution is noted as the +1 (-1) σ shift. The jet energy scale (JES) uncertainty is an example

of the latter case. The jet energies are shifted and therefore events can migrate in and out of the acceptance. Again the difference in the final variable is noted as the 1 σ error but if the variations are small and/or the sample statistics are small, the MC statistical uncertainty can make up a substantial part of this supposed systematic difference. If there are multiple JES uncertainties, as in this analysis, then this MC uncertainty should not be included in each one.

To mitigate these effects, two so-called smoothing algorithms are used to merge consecutive bins in the MC templates. First, bins from one extremum to the next are merged until no local extrema remain in the BDT distribution (or up to on in the m_{bb} distribution for the di-jet mass analysis). Merging is performed at each step of this iterative process where the difference between merged and unmerged templates is the smallest. Second, the bins resulting from this first algorithm are sequentially merged, starting from the upper end of the distribution, until the statistical uncertainty in each of the merged bins, calculated in the nominal template, is smaller than 5%. In each of these sets of bins, the integrals of the nominal and systematically shifted distributions are compared to give the $\pm 1\sigma$ variation. This value is then used as the associated uncertainty for all the nominal bins in the set.

5.8.5.3 Pruning of the Systematic Uncertainties

Several of the uncertainties described in Section 5.7 have a negligible effect on the distributions entering in the fit. In addition, limited statistics in the MC nominal distributions can produce systematic templates with large fluctuations, introducing artificial variations in the fit. Therefore, uncertainties are removed following a pruning procedure, which is carried out for each category/sub-channel in each region and performed as follows:

- Neglect the normalisation uncertainty for a given sample in a region if either of the following is true: the variation is less than 0.5%; both up and down variations have the same sign.

- Neglect the shape uncertainty for a given sample in a given region if the following is true: not one single bin has a deviation over 0.5% after the overall normalisation is removed; if only the up or the down variation is non-zero and passed the previous pruning steps.

- Neglect the shape and normalisation uncertainties for a given sample in a given region if the sample is less than 2% of the total background: if

the signal $< 2\%$ of the total background in all bins and the shape and normalisation error are each $< 0.5\%$ of the total background; if at least one bin has a signal contribution $> 2\%$ of the total background, only in those bins where the shape and normalisation error are each $< 2\%$ of the signal yield.

5.9 Results

In this section, the results from the likelihood fit are presented. Section 5.9.1 shows the results from the main multivariate analysis BDT_{VH} fit with Run 2 data, Section 5.9.2 and Section 5.9.3 present the results from the BDT_{VZ} fit of the diboson analysis and m_{bb} fit of the dijet-mass analysis, respectively. The results from the combination of the main VH multivariate analysis and the previously published analysis of Run 1 data, the other searches for $b\bar{b}$ decays of the Higgs boson and the other searches in the VH production mode are presented in Section 5.9.4.

5.9.1 Results of the SM Higgs boson search at $\sqrt{s} = 13$ TeV

5.9.1.1 Post-fit distributions and yields

Figure 5.59 to 5.61 show the post-fit distributions for the variables used as input to the global likelihood fit in the three channels in both signal and control regions. The post-fit distributions are obtained by applying the the best fit μ and $\boldsymbol{\theta}$ to the simulated MC events.

Figure 5.62 to 5.63 show some other post-fit distributions for the variables not used directly as input to the global likelihood fit in the three channels in both signal and control regions. When applying the post-fit μ and $\boldsymbol{\theta}$ from the fit with the BDT_{VH} to the other variables, the nuisance parameters arising from MC statistical uncertainties are not included due to the complexity in translating the MC statistical uncertainties from one distribution to another.

The post-fit signal and background yields are shown in Table 5.35 and Table 5.36 for all signal regions and control regions, respectively. The post-fit normalisation factors of the floating backgrounds in the global likelihood fit are shown in Table 5.37.

Figure 5.64 shows the data, background and signal yields, where final-discriminant bins in all regions are combined into bins of log(S/B). S and B

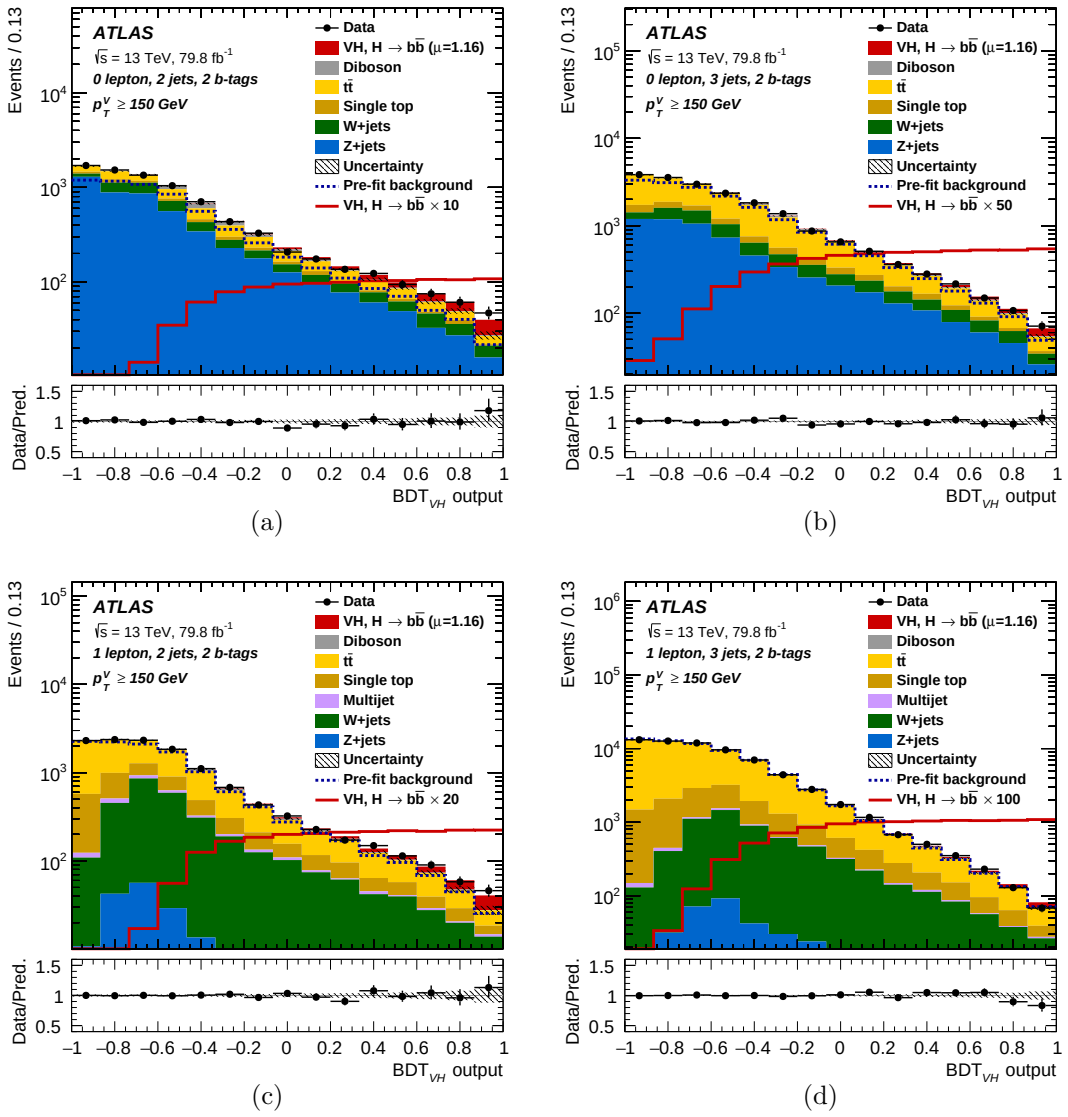


Figure 5.59: The BDT_{VH} post-fit distributions from the global likelihood fit in the 0-lepton 2-jet SR (a), 0-lepton 3-jet SR (b), 1-lepton 2-jet SR (c), 1-lepton 3-jet SR (d). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data, and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

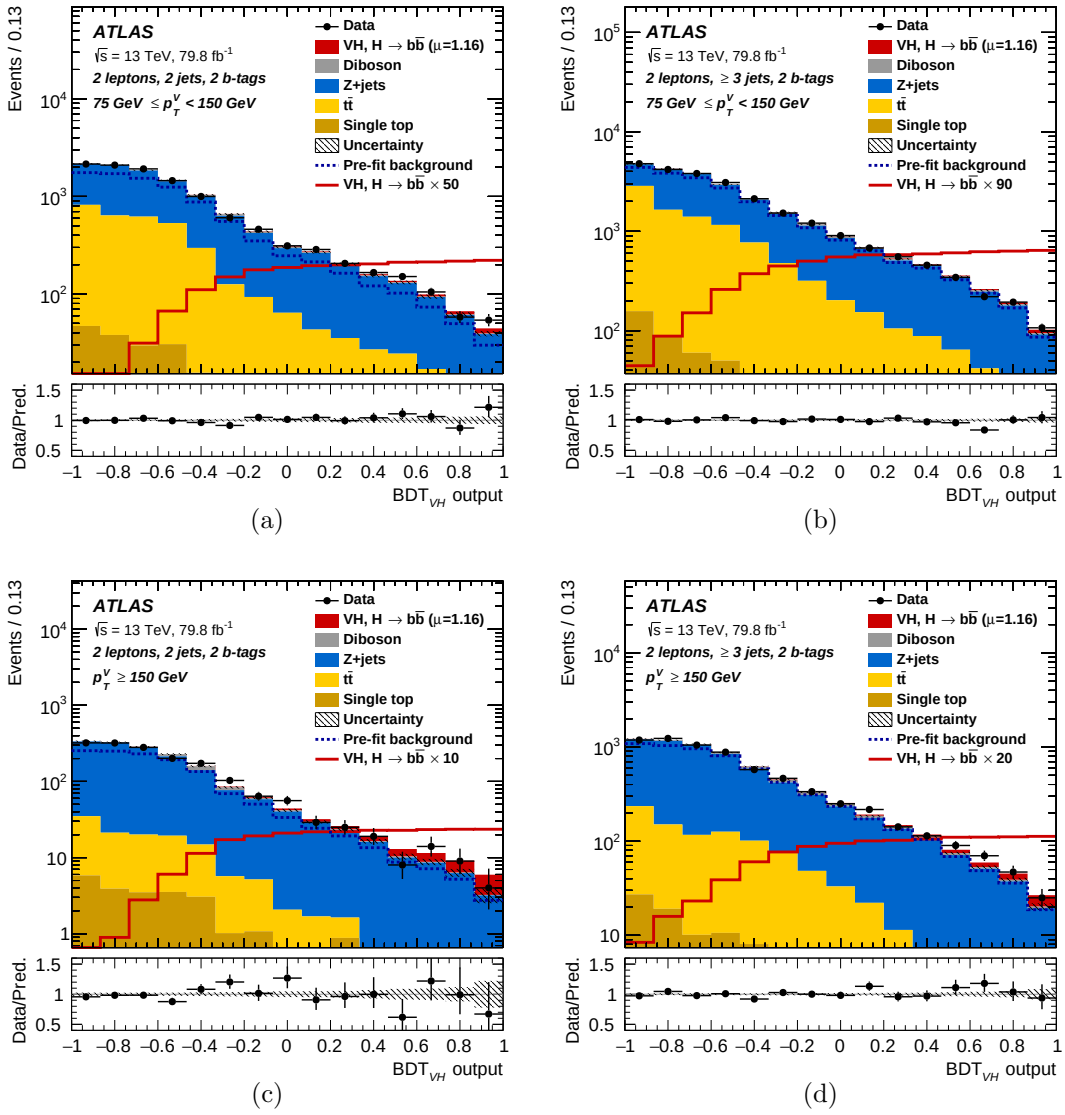


Figure 5.60: The BDT_{VH} post-fit distributions from the global likelihood fit in the 2-lepton channel, in the 2-jet $75\,\mathrm{GeV} < p_T^V < 150\,\mathrm{GeV}$ region (a), 3-jet $75\,\mathrm{GeV} < p_T^V < 150\,\mathrm{GeV}$ (b), 2-jet $p_T^V > 150\,\mathrm{GeV}$ (c) and 3-jet $p_T^V > 150\,\mathrm{GeV}$ (d). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data, and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

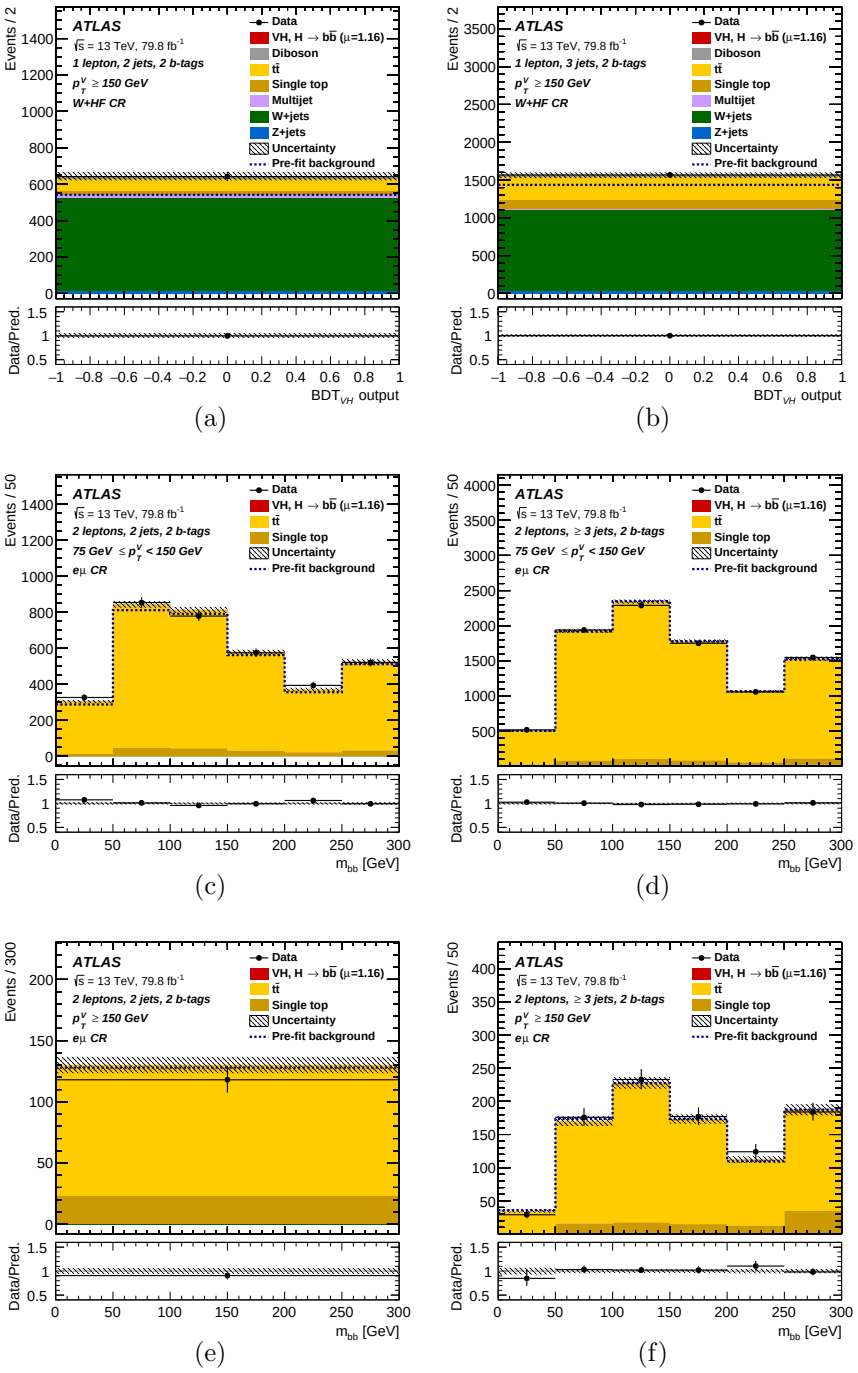


Figure 5.61: The BDT_{VH} post-fit distributions from the global likelihood fit in the 1-lepton channel W+HF CR, in the 2-jet region (a), 3-jet region (b). The m_{bb} post-fit distributions from the global likelihood fit in the 2-lepton Top $e\mu$ CR, in the 2-jet $75\,\mathrm{GeV} < p_T^V < 150\,\mathrm{GeV}$ region (c), 3-jet $75\,\mathrm{GeV} < p_T^V < 150\,\mathrm{GeV}$ (d), 2-jet $p_T^V > 150\,\mathrm{GeV}$ (e) and 3-jet $p_T^V > 150\,\mathrm{GeV}$ (f). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data. The entries in overflow are included in the last bin. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

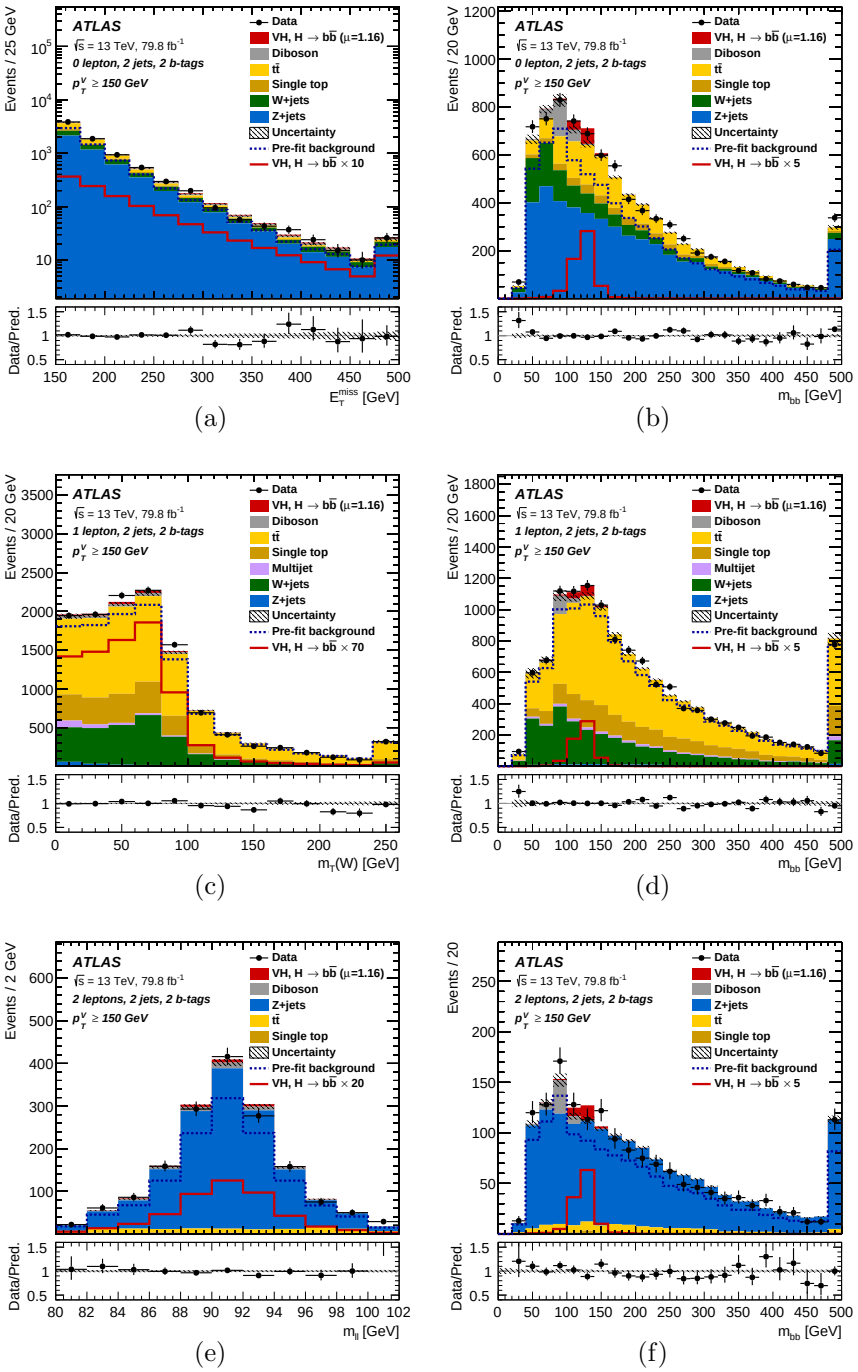


Figure 5.62: The post-fit distributions for E_T^{miss} (a), m_T^W (c), m_{ll} (e) and m_{bb} (right) in the 0-lepton (top), 1-lepton (middle) and 2-lepton (bottom) channels for 2-jet, 2-b-tag events in the high p_T^V region. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds normalized to the signal yield extracted from data ($\mu = 1.16$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The entries in overflow are included in the last bin. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

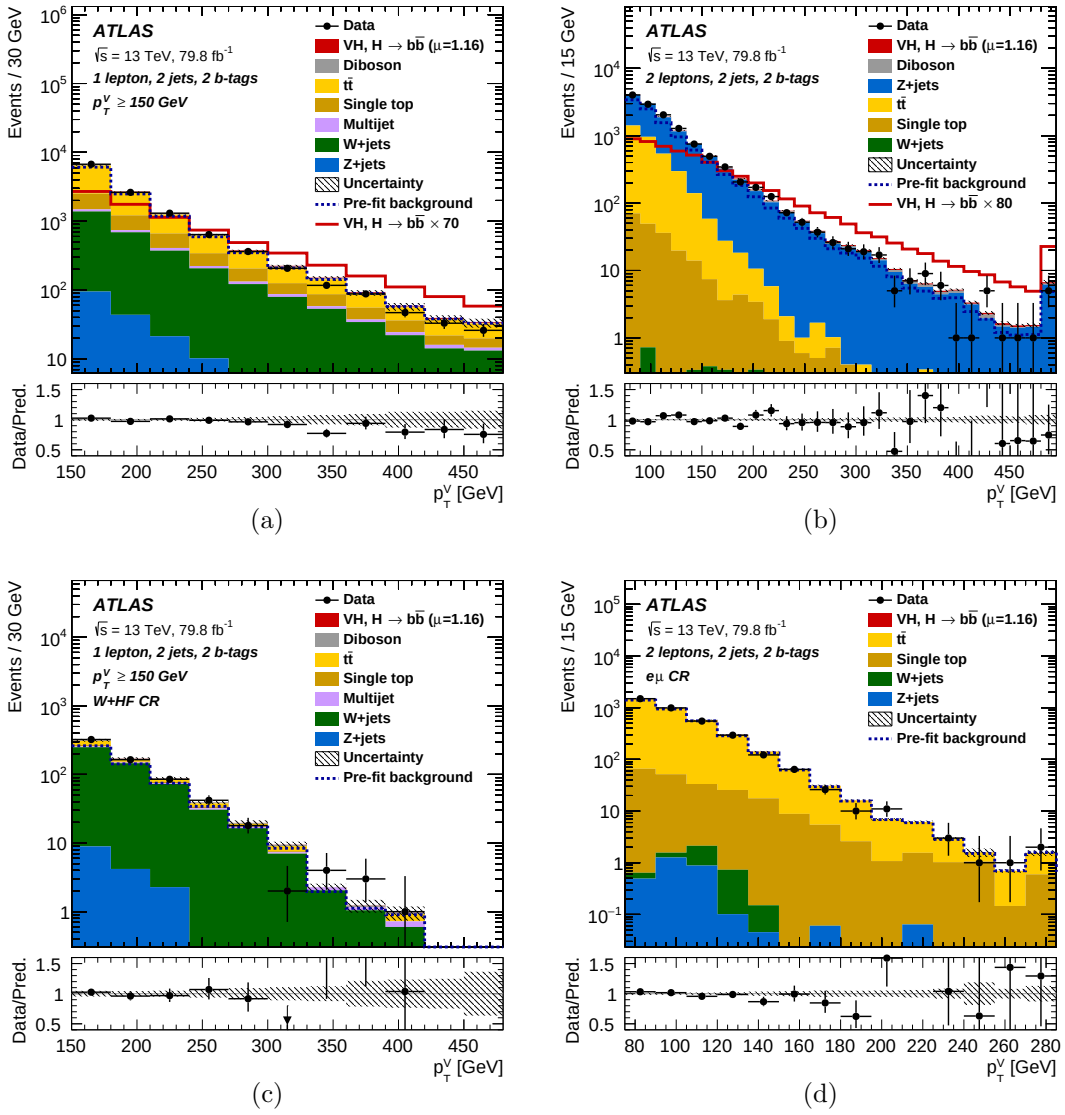


Figure 5.63: Distributions of the p_T^V for all 2-jet signal and control regions in 1 and 2-lepton channels. Shown are the data (points with error bars) and expectation (histograms). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds normalized to the signal yield extracted from data ($\mu = 1.16$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend for the signal regions. In the W + HF and $e\mu$ CRs, the unstacked unfilled histograms for the signal are not shown. The dashed histogram shows the total pre-fit background. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty for the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

Table 5.35: The Higgs boson signal, background and data yields for each signal region category in each channel after the full selection of the multivariate analysis. The signal and background yields are normalised to the results of the global likelihood fit. All systematic uncertainties are included in the indicated uncertainties. An entry of "–" indicates that a specific background component is negligible in a certain region, or that no simulated events are left after the analysis selection.

	0-lepton $p_T^V > 150$ GeV, 2-b-tag		1-lepton $p_T^V > 150$ GeV, 2-b-tag		2-lepton 75 GeV $< p_T^V < 150$ GeV, 2-b-tag		2-lepton $p_T^V > 150$ GeV, 2-b-tag	
Process	2-jet	3-jet	2-jet	3-jet	2-jet	≥3-jet	2-jet	≥3-jet
$Z + ll$	17 ± 11	27 ± 18	2 ± 1	3 ± 2	14 ± 9	49 ± 32	4 ± 3	30 ± 19
$Z + cl$	45 ± 18	76 ± 30	3 ± 1	7 ± 3	43 ± 17	170 ± 67	12 ± 5	88 ± 35
Z + HF	4770 ± 140	5940 ± 300	180 ± 9	348 ± 21	7400 ± 120	14160 ± 220	1421 ± 34	5370 ± 100
$W + ll$	20 ± 13	32 ± 22	31 ± 23	65 ± 48	< 1	< 1	< 1	< 1
$W + cl$	43 ± 20	83 ± 38	139 ± 67	250 ± 120	< 1	< 1	< 1	< 1
W + HF	1000 ± 87	1990 ± 200	2660 ± 270	5400 ± 670	2 ± 0	13 ± 2	1 ± 0	4 ± 1
Single top quark	368 ± 53	1410 ± 210	2080 ± 290	9400 ± 1400	188 ± 89	440 ± 200	23 ± 7	93 ± 26
$t\bar{t}$	1333 ± 82	9150 ± 400	6600 ± 320	50200 ± 1400	3170 ± 100	8880 ± 220	104 ± 6	839 ± 40
Diboson	254 ± 49	318 ± 90	178 ± 47	330 ± 110	152 ± 32	355 ± 68	52 ± 11	196 ± 35
Multi-jet e sub-ch.	–	–	100 ± 100	41 ± 35	–	–	–	–
Multi-jet μ sub-ch.	–	–	138 ± 92	260 ± 270	–	–	–	–
Total bkg.	7850 ± 90	19020 ± 140	12110 ± 120	66230 ± 270	10960 ± 100	24070 ± 150	1620 ± 30	6620 ± 80
Signal (post-fit)	128 ± 28	128 ± 29	131 ± 30	125 ± 30	51 ± 11	86 ± 22	28 ± 6	67 ± 17
Data	8003	19143	12242	66348	11014	24197	1626	6686

represent fitted signal and background yields in each analysis bin, respectively.

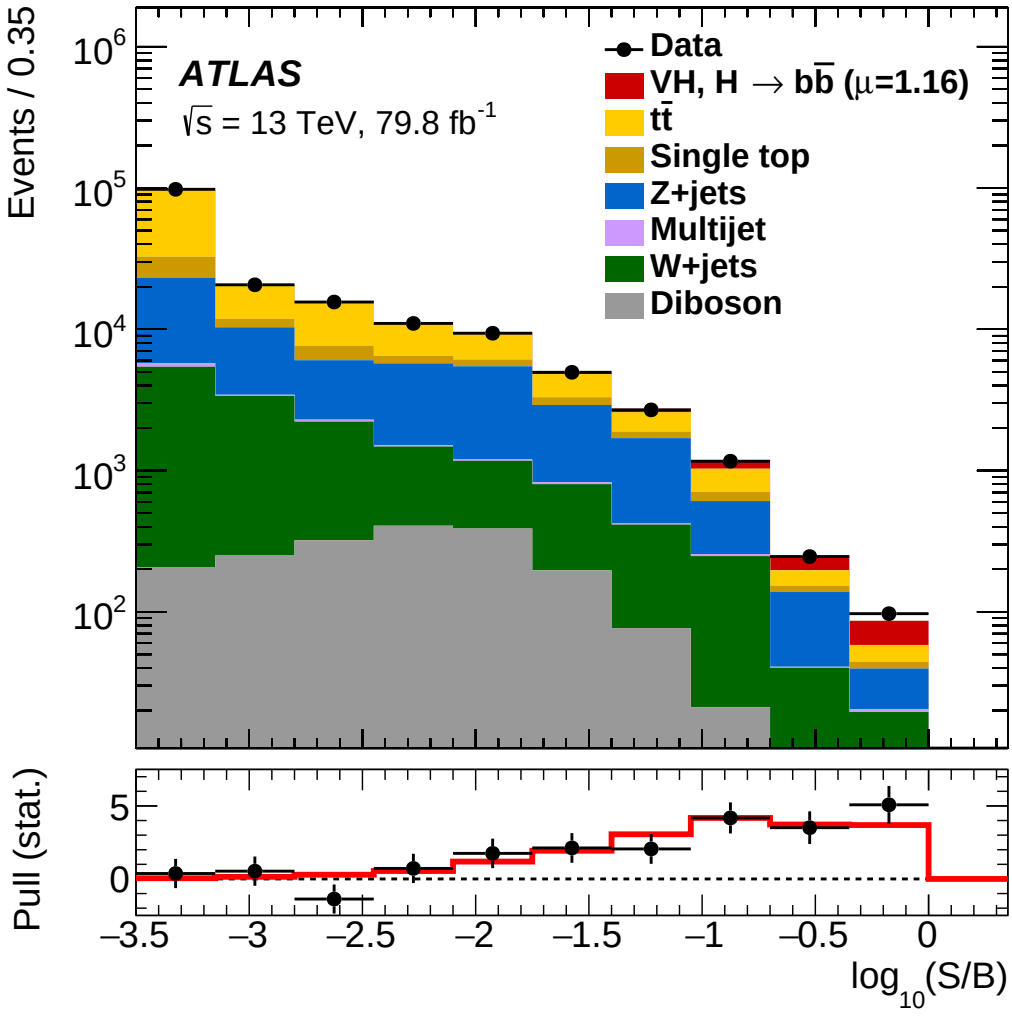


Figure 5.64: Event yields as a function of log (S/B) for data, background and a Higgs boson signal with $m_H = 125$ GeV. Final-discriminant bins in all regions are combined into bins of log (S/B), with S being the fitted signal and B the fitted background yields. The Higgs boson signal contribution is shown after rescaling the SM cross-section according to the value of the signal strength extracted from data ($\mu = 1.16$). In the lower panel, the pull of the data relative to the background (the statistical significance of the difference between data and fitted background) is shown with statistical uncertainties only. The full line indicates the pull expected from the sum of fitted signal and background relative to the fitted background.

5.9.1.2 Signal strength and significance

For a Higgs boson mass of 125 GeV, when all lepton channels are combined, the probability p_0 of obtaining a signal at least as strong as the observation from

background alone is $5.3 \cdot 10^{-7}$, whilst the expected value is $7.3 \cdot 10^{-6}$. The observation corresponds to an excess with a significance of 4.9 standard deviations, to be compared with an expectation of 4.3 standard deviations. The fitted value of the signal strength is:

$$\mu_{VH}^{bb} = 1.16^{+0.27}_{-0.25} = 1.16 \pm 0.16(\text{stat.})^{+0.21}_{-0.19}(\text{syst.}).$$

Combined fits are also performed with floating signal strength parameters separately for the three lepton channels, or the WH and ZH production processes, but leaving all other NPs with the same correlation scheme as for the nominal result. The results of these fits are shown in Table 5.38 and Figure 5.65. The compatibility of the signal strength parameters measured in the three lepton channels is 80%. The compatibility of the signal strength across different analysis regions in the fit is evaluated by repeating the fit with different signal strength parameters assigned to each of such N regions, while keeping the rest of the likelihood definition unchanged. Under the hypothesis that the true underlying signal strength parameter values are the same, the difference in the values of profiled -2ln$\mathcal{L}$ between the likelihood fit in the nominal and in the new configuration is expected to be distributed according to a χ^2 distribution with number of degrees of freedom equal to N - 1. The corresponding p-value is thus quoted as a measure of the compatibility. The WH and ZH production modes have observed (expected) significances of 2.5 (2.3) and 4.0 (3.5) standard deviations, respectively, with a linear correlation between the two signal strengths of -1%.

5.9.1.3 Systematic uncertainties breakdown and ranking

The effects of systematic uncertainties on the measurement of the signal strength are presented in Table 5.39. The total statistical uncertainty is defined as the uncertainty in μ when all the NPs are fixed to their best-fit values. The total systematic uncertainty is defined as the difference in quadrature between the total uncertainty in μ and the total statistical uncertainty. As presented in the table, the analysis is now systematically limited, the systematic uncertainties due to the modelling of the signal play a dominant role, followed by the uncertainty due to the limited size of the simulated samples, the modelling of the backgrounds and the b-tagging uncertainty.

Impact of systematic uncertainties for the fitted Higgs boson signal strength μ are presented in Figure 5.66 with the ranking method, the systematic uncertainties

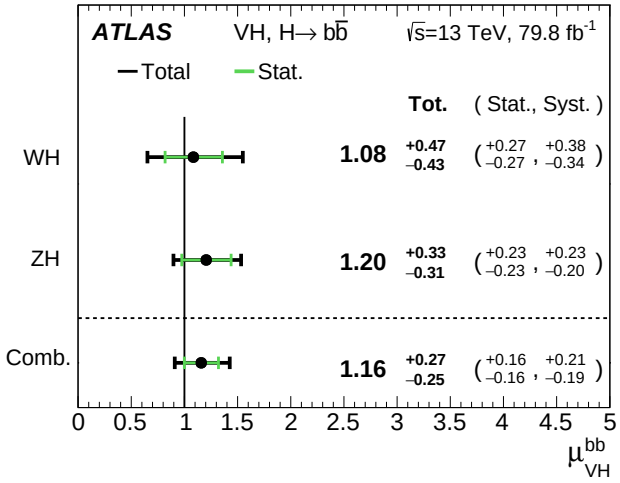


Figure 5.65: The fitted values of the Higgs boson signal strength μ_{VH}^{bb} for $m_H = 125$ GeV for the WH and ZH processes and their combination. The individual μ_{VH}^{bb} values for the WH and ZH processes are obtained from a simultaneous fit with the signal strength for each of the WH and ZH processes floating independently. The probability of compatibility of the individual signal strengths is 84%.

are listed in decreasing order of their impact on μ. As shown in the figure, the three leading contributions are from the systematic uncertainties of W + jets p_T^V, Z + jets m_{bb} shape and Diboson m_{bb} shape. The large data sample in the 0- and 2-lepton mass sidebands allows the fit to pull and constrain the nuisance parameter on the m_{bb} shape of the Z + HF background. The pull corrects a mismodelling, observed in Z + HF enriched sideband regions, of the m_{bb} distribution by the simulation.

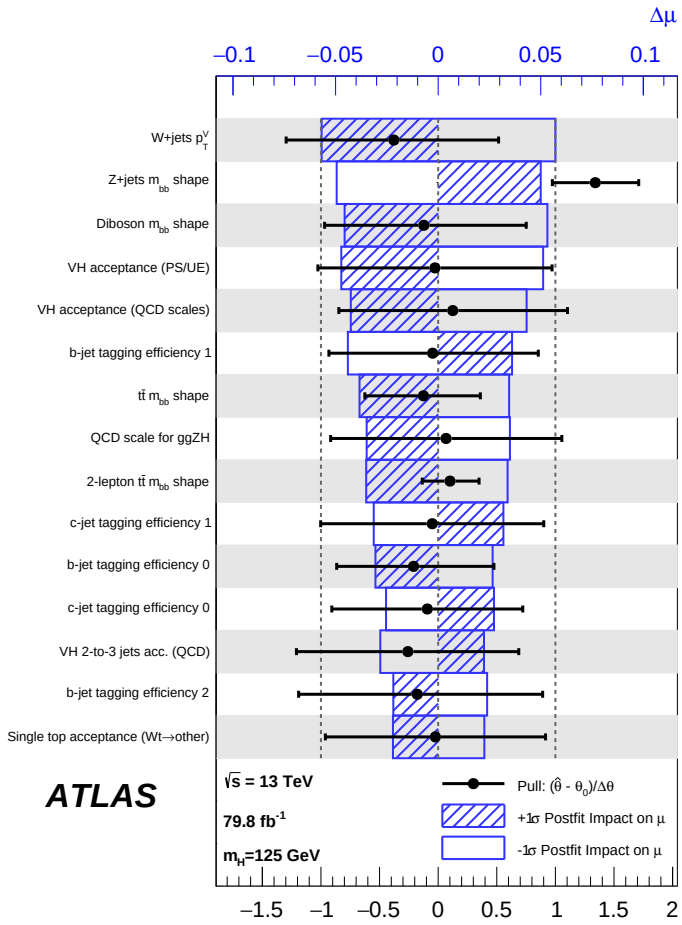


Figure 5.66: Impact of systematic uncertainties for the fitted Higgs boson signal strength μ for the nominal MVA analysis applied to the 13 TeV data. The systematic uncertainties are listed in decreasing order of their impact on μ. The boxes show the variations of μ, referring to the top x-axis, when fixing the corresponding individual nuisance parameter θ to its fitted value modified upwards or downwards by its fitted uncertainty, and performing the fit again, with all the other parameters allowed to vary, so as to take correlations between systematic uncertainties properly into account. The hatched and open areas correspond to the upwards and downwards variations, respectively. The filled circles, referring to the bottom x-axis, show the deviations of the fitted nuisance parameters from their nominal values , expressed in terms of standard deviations with respect to their nominal uncertainties. The associated error bars show the fitted uncertainties of the nuisance parameters, relative to their nominal uncertainties.

Table 5.36: The fitted signal and background yields for each control region category in each channel (W + HF in the 1-lepton channel, $e\mu$ events in the 2-lepton channel), corresponding to the selection applied to the control regions for the multivariate analysis. The yields are normalised by the results of the global likelihood fit. All systematic uncertainties are included in the indicated uncertainties. An entry of "–" indicates that a specific background component is negligible in a certain region, or that no simulated events are left after the analysis selection.

	1-lepton		2-lepton			
	$p_T^V > 150$ GeV, 2-b-tag		75 GeV $< p_T^V <$ 150 GeV, 2-b-tag		$p_T^V > 150$ GeV, 2-b-tag	
Process	2-jet	3-jet	2-jet	≥3-jet	2-jet	≥3-jet
Z + HF	15 ± 1	33 ± 3	3 ± 0	2 ± 0	< 1	< 1
$W + ll$	2 ± 2	4 ± 3	–	–	–	–
$W + cl$	8 ± 4	14 ± 7	–	< 1	–	–
W + HF	498 ± 34	1044 ± 92	3 ± 0	8 ± 1	< 1	3 ± 0
Single top quark	24 ± 5	122 ± 23	189 ± 90	450 ± 210	22 ± 7	93 ± 27
$t\bar{t}$	68 ± 18	307 ± 77	3243 ± 98	8690 ± 210	107 ± 7	807 ± 37
Diboson	13 ± 4	23 ± 8	–	< 1	–	< 1
Multi-jet e sub-ch.	8 ± 9	4 ± 3	–	–	–	–
Multi-jet μ sub-ch.	7 ± 5	13 ± 13	–	–	–	–
Total bkg.	644 ± 23	1563 ± 39	3437 ± 58	9153 ± 95	130 ± 7	905 ± 27
Signal (post-fit)	< 1	2 ± 1	< 1	< 1	< 1	< 1
Data	642	1567	3450	9102	118	923

Table 5.37: Factors applied to the nominal normalisations of the $t\bar{t}$, $W + HF$, and $Z + HF$ backgrounds, as obtained from the global likelihood fit. The errors represent the combined statistical and systematic uncertainties.

Process	Normalisation factor
$t\bar{t}$ 0- and 1-lepton	0.98 ± 0.08
$t\bar{t}$ 2-lepton 2-jet	1.06 ± 0.09
$t\bar{t}$ 2-lepton 3-jet	0.95 ± 0.06
$W + HF$ 2-jet	1.19 ± 0.12
$W + HF$ 3-jet	1.05 ± 0.12
$Z + HF$ 2-jet	1.37 ± 0.11
$Z + HF$ 3-jet	1.09 ± 0.09

Table 5.38: Measured signal strengths with their combined statistical and systematic uncertainties, expected and observed p_0 and significance values (in standard deviations) from the combined fit with a single signal strength, and from a combined fit where each of the lepton channels has its own signal strength, using 13 TeV data.

Signal strength	Signal strength	p_0		Significance	
		Exp.	Obs.	Exp.	Obs.
0-lepton	$1.04^{+0.34}_{-0.32}$	$9.5 \cdot 10^{-4}$	$5.1 \cdot 10^{-4}$	3.1	3.3
1-lepton	$1.09^{+0.46}_{-0.42}$	$8.7 \cdot 10^{-3}$	$4.9 \cdot 10^{-3}$	2.4	2.6
2-lepton	$1.38^{+0.46}_{-0.42}$	$4.0 \cdot 10^{-3}$	$3.3 \cdot 10^{-4}$	2.6	3.4
VH, $H \to b\bar{b}$ combination	$1.16^{+0.27}_{-0.25}$	$7.3 \cdot 10^{-6}$	$5.3 \cdot 10^{-7}$	4.3	4.9

Table 5.39: Breakdown of the contributions to the uncertainty in μ.

Source of uncertainty		σ_μ
Total		0.259
Statistical		0.161
Systematic		0.203
Experimental uncertainties		
Jets		0.035
$E_{\mathrm{T}}^{\mathrm{miss}}$		0.014
Leptons		0.009
b-tagging	b-jets	0.061
	c-jets	0.042
	light-flavour jets	0.009
	extrapolation	0.008
Pile-up		0.007
Luminosity		0.023
Theoretical and modelling uncertainties		
Signal		0.094
Floating normalisations		0.035
Z + jets		0.055
W + jets		0.060
$t\bar{t}$		0.050
Single top quark		0.028
Diboson		0.054
Multi-jet		0.005
MC statistical		0.070

5.9.2 Results of the diboson analysis

The diboson analysis targets diboson production with a Z boson decaying into a pair of b-quarks and produced in association with either a W or Z boson. This process has a signature that is similar to the one considered in this analysis, and therefore provides an important validation of the VH result. The cross-section is about nine times larger than for the SM Higgs boson with a mass of 125 GeV, the m_{bb} distribution peaks at lower values, and the p_T^{bb} spectrum is softer. The BDT$_{VZ}$ is used to extract the diboson signal. In the diboson-analysis fits, the normalization of the diboson contributions is allowed to vary with a multiplicative scale factor μ_{VZ} with respect to the SM prediction, except for the small contribution from WW production, which is treated as a background and constrained within its uncertainty. The overall normalization uncertainties for the WZ and ZZ processes are removed, while all other systematic uncertainties are kept identical to those in the nominal fit used to extract the Higgs boson signal. A SM Higgs boson with $mH = 125$ GeV is included as a background, with a production cross-section at the SM value with an uncertainty of 50%. The diboson and Higgs boson BDTs provide sufficient separation between the VZ and VH processes that they only have a weak direct correlation (<1%) in their results.

5.9.2.1 Post-fit distributions

Figure 5.67 and 5.68 show the post-fit BDT$_{VZ}$ distributions in the three channels in the signal regions.

Figure 5.69 shows the data, background and VZ diboson signal yields, where final-discriminant bins in all regions are combined into bins of log(S/B). S and B represent fitted signal and background yields in each analysis bin, respectively.

5.9.2.2 Signal strength

The fitted value of the signal strength of the diboson analysis is:

$$\mu_{VZ}^{bb} = 1.20^{+0.20}_{-0.18} = 1.20 \pm 0.08(\text{stat.})^{+0.19}_{-0.16}(\text{syst.}),$$

which is in good a agreement with the Standard Model prediction. Combined fits are also performed with floating signal strength parameters separately for the three lepton channels, or the WZ and ZZ production processes. The results of these fits are shown in Figure 5.70 and 5.71.

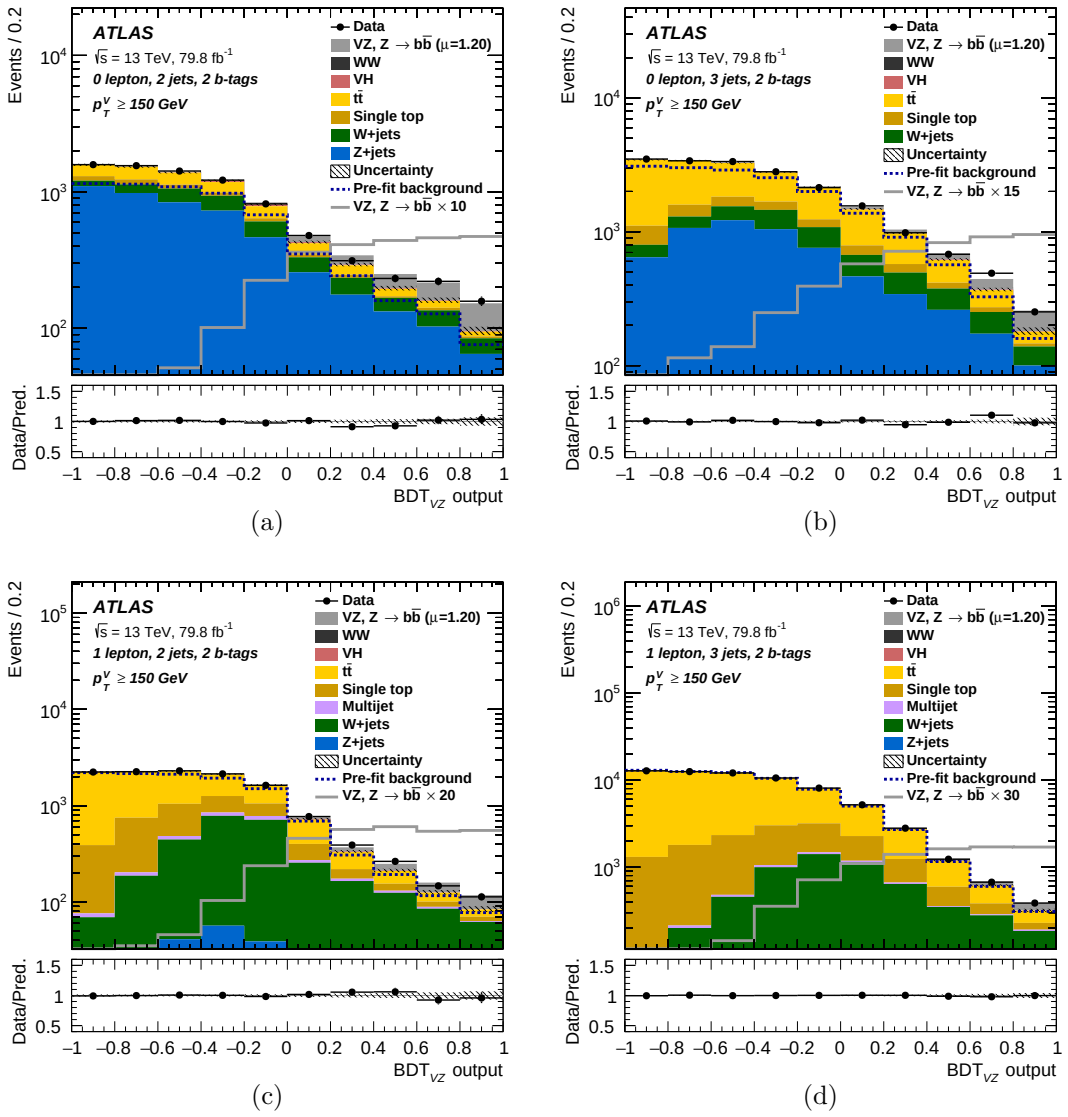


Figure 5.67: The BDT_{VZ} post-fit distributions from the global likelihood fit in the 0-lepton 2-jet SR (a), 0-lepton 3-jet SR (b), 1-lepton 2-jet SR (c), 1-lepton 3-jet SR (d). The background contributions after the global likelihood fit are shown as filled histograms. The VZ diboson signal is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data, and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

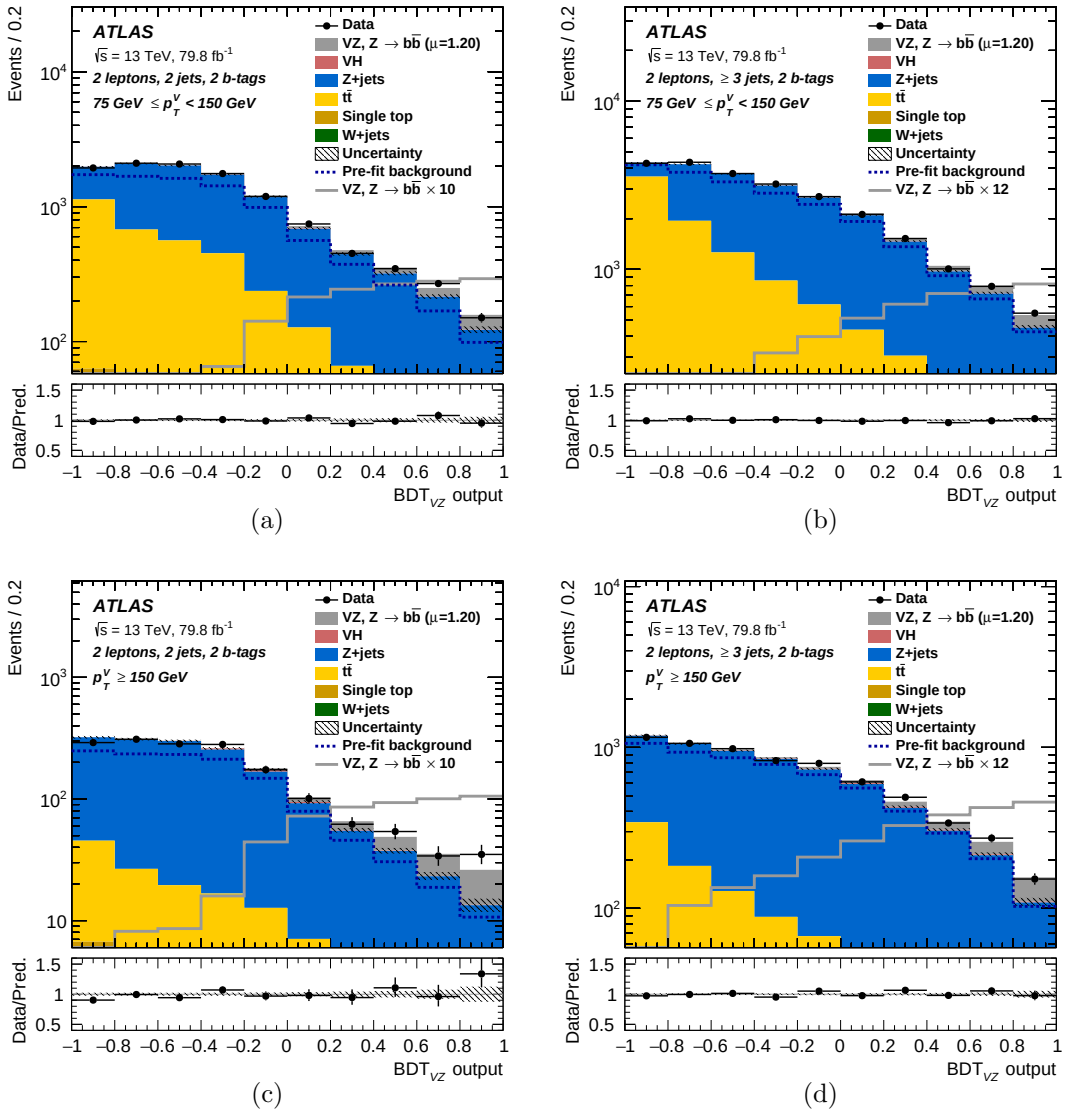


Figure 5.68: The BDT_{VZ} post-fit distributions from the global likelihood fit in the 2-lepton channel, in the 2-jet $75\,\mathrm{GeV} < p_T^V < 150\,\mathrm{GeV}$ region (a), 3-jet $75\,\mathrm{GeV} < p_T^V < 150\,\mathrm{GeV}$ (b), 2-jet $p_T^V > 150\,\mathrm{GeV}$ (c) and 3-jet $p_T^V > 150\,\mathrm{GeV}$ (d). The background contributions after the global likelihood fit are shown as filled histograms. The VZ diboson signal is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data, and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

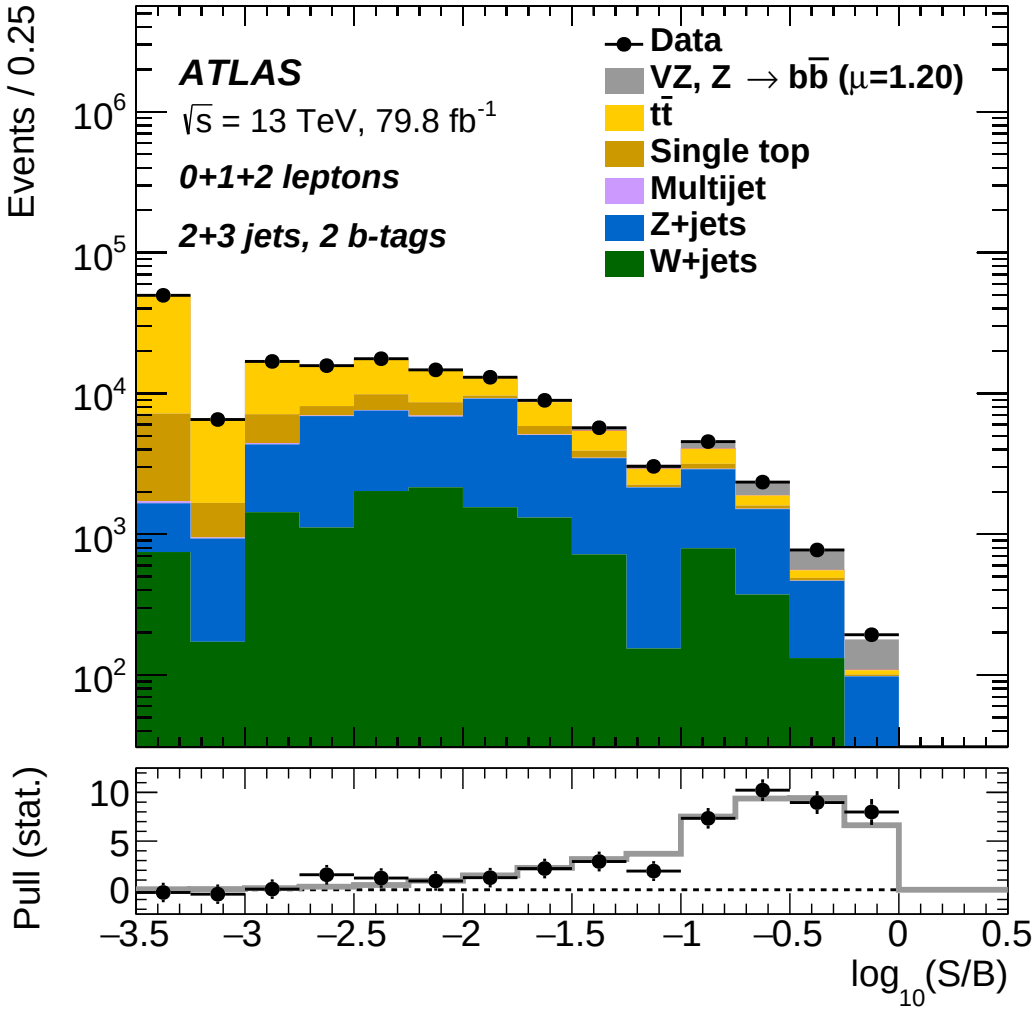


Figure 5.69: Event yields as a function of log (S/B) for data, background and VZ diboson signal. Final-discriminant bins in all regions are combined into bins of log (S/B), with S being the fitted VZ diboson signal and B the fitted background yields. The VZ diboson signal contribution is shown after rescaling the SM cross-section according to the value of the signal strength extracted from data ($\mu = 1.20$). In the lower panel, the pull of the data relative to the background (the statistical significance of the difference between data and fitted background) is shown with statistical uncertainties only. The full line indicates the pull expected from the sum of fitted signal and background relative to the fitted background.

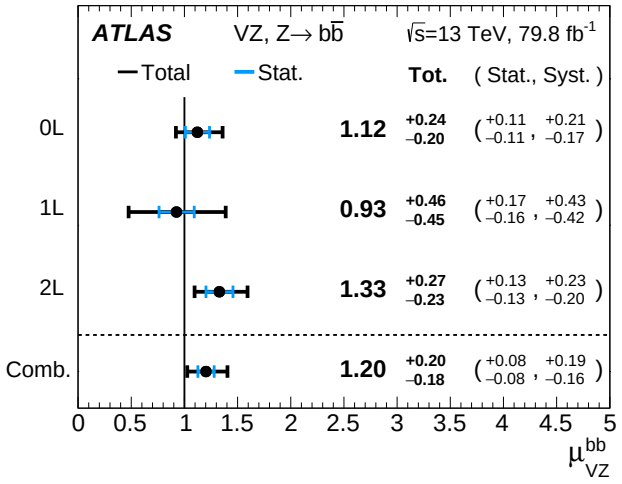


Figure 5.70: The fitted values of the diboson signal strength μ_{VZ}^{bb} for the 0-, 1-, 2-lepton channels and their combination. The individual μ_{VZ}^{bb} values for lepton channels are obtained from a simultaneous fit with the signal strength for each of the lepton channels floating independently. The probability of compatibility of the individual signal strengths is 64%.

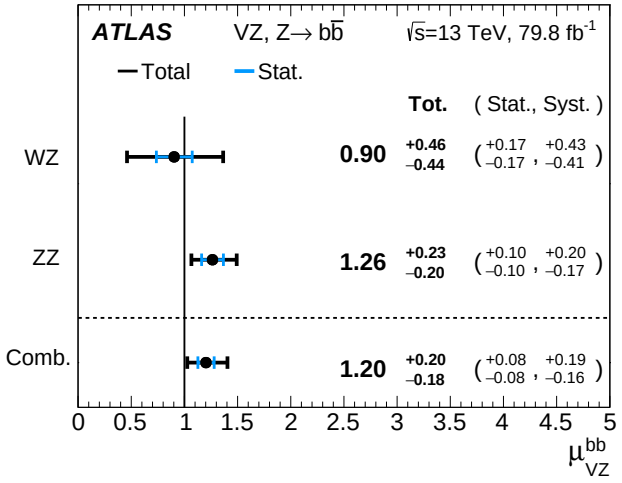


Figure 5.71: The fitted values of the diboson signal strength μ_{VZ}^{bb} for the WZ and ZZ processes and their combination. The individual μ_{VZ}^{bb} values for the WZ and ZZ processes are obtained from a simultaneous fit with the signal strength for each of the WZ and ZZ processes floating independently. The probability of compatibility of the individual signal strengths is 47%.

5.9.3 Results of the dijet-mass analysis

In the dijet-mass analysis, the BDT_{VH} discriminant is replaced by the m_{bb} variable as the main input used in the global fit, and the number of signal regions is increased as a consequence of splitting the event categories with $p_T^V > 150\,\mathrm{GeV}$ in two in each of the three lepton channels.

5.9.3.1 Post-fit distributions

Figure 5.72 to 5.74 show the post-fit m_{bb} distributions in the three channels in the signal regions.

Figure 5.75 shows the m_{bb} distribution summed over all channels and regions, weighted by their respective values of the ratio of fitted Higgs boson signal and background yields and after subtraction of all backgrounds except for the WZ and ZZ diboson processes.

5.9.3.2 Signal strength and significance

For a Higgs boson mass of 125 GeV, when all lepton channels are combined, the observed excess has a significance of 3.6 standard deviations, to be compared to an expectation of 3.5 standard deviations. The fitted value of the signal strength is:

$$\mu_{VH}^{bb} = 1.06^{+0.36}_{-0.33} = 1.06 \pm 020(\mathrm{stat.})^{+0.30}_{-0.26}(\mathrm{syst.}).$$

Combined fits are also performed with floating signal strength parameters separately for the three lepton channels. Good agreement is also found when comparing the values of signal strengths in the individual channels from the dijet-mass analysis with those from the multivariate analysis, as shown in Figure 5.76.

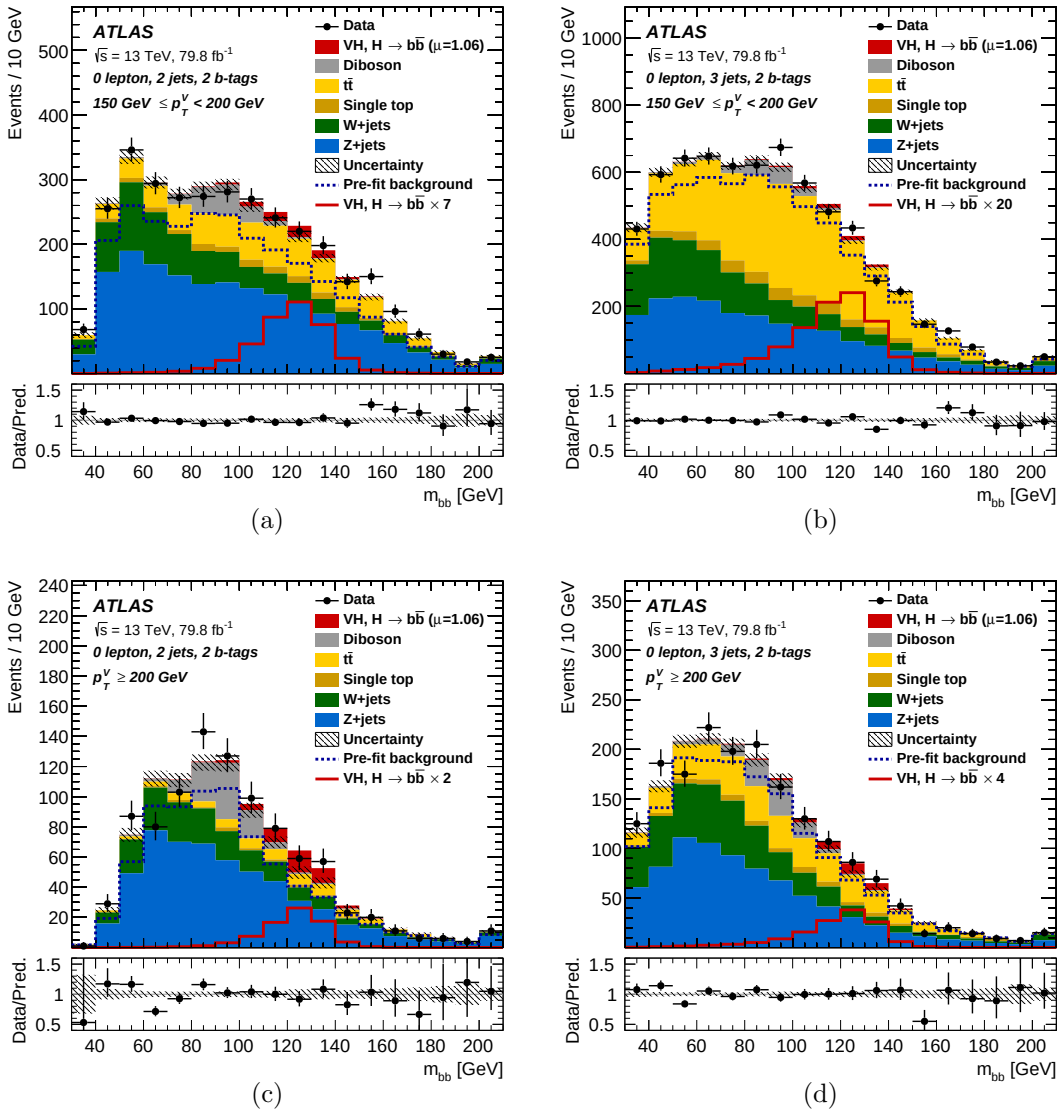


Figure 5.72: The m_{bb} post-fit distributions from the global likelihood fit in the 0-lepton channel,as obtained with the dijet-mass analysis. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data ($\mu = 1.06$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The entries in overflow are included in the last bin. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

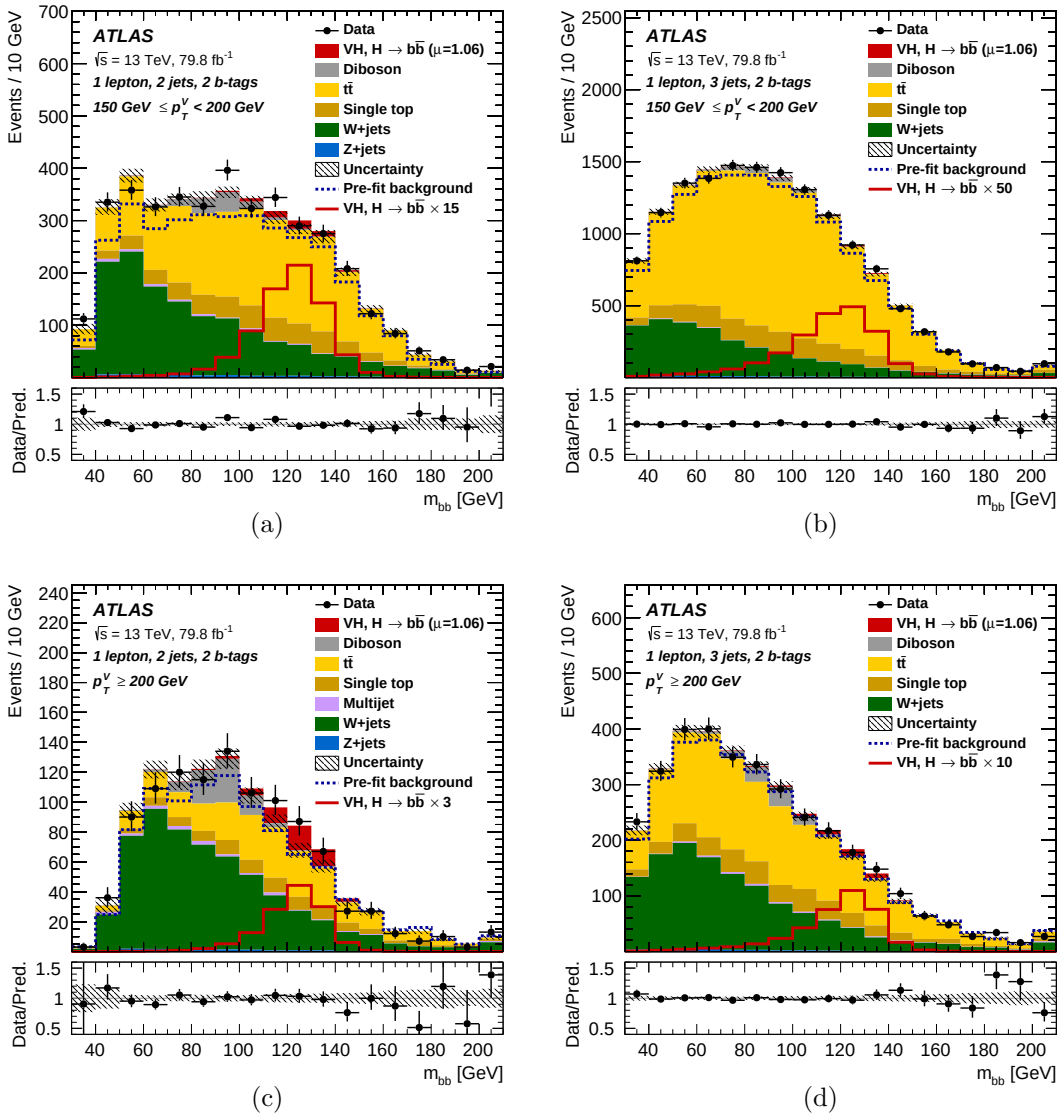


Figure 5.73: The m_{bb} post-fit distributions from the global likelihood fit in the 1-lepton channel, as obtained with the dijet-mass analysis. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data ($\mu = 1.06$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The entries in overflow are included in the last bin. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

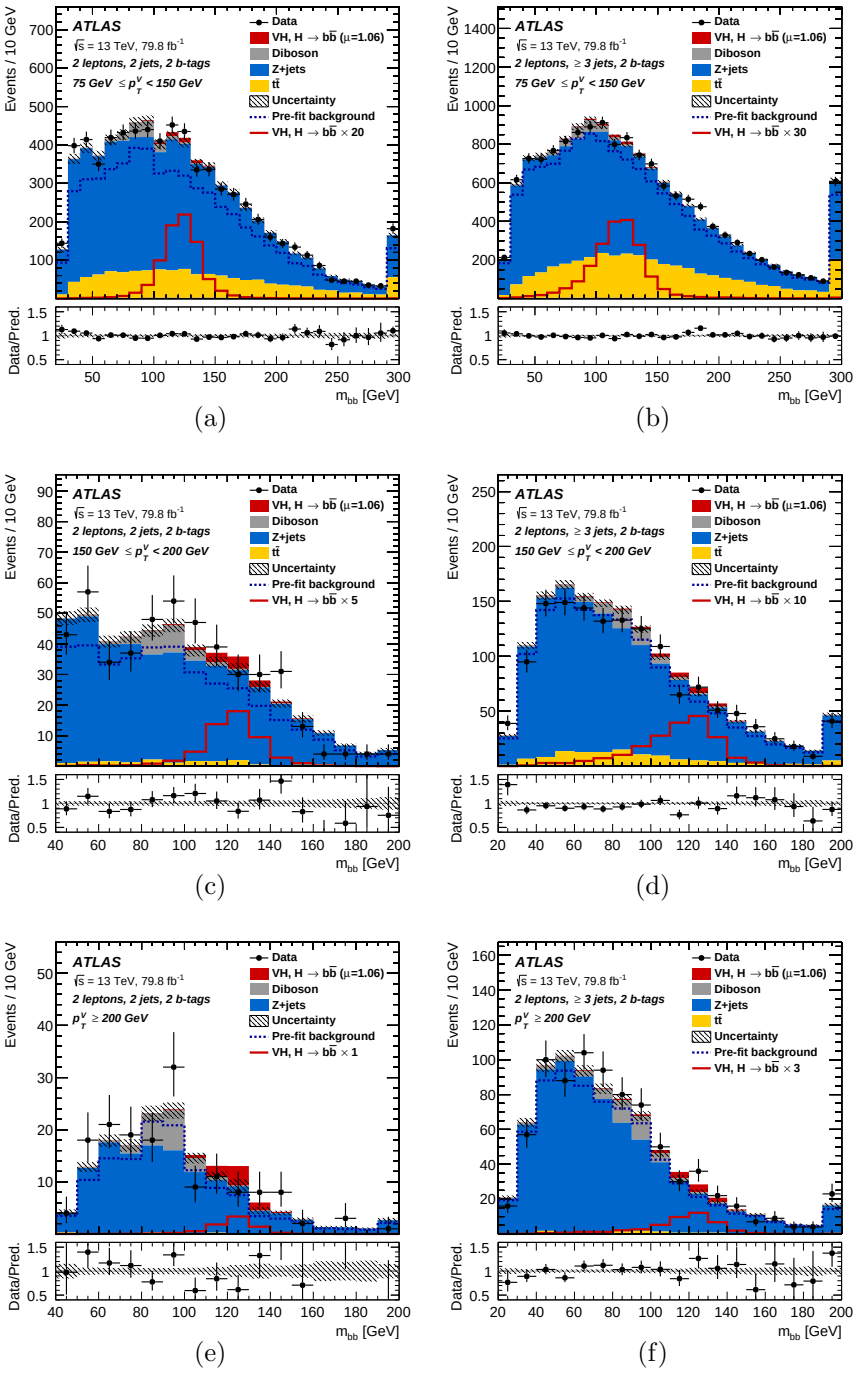


Figure 5.74: The m_{bb} post-fit distributions from the global likelihood fit in the 2-lepton channel, as obtained with the dijet-mass analysis. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data ($\mu = 1.06$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The entries in overflow are included in the last bin. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.

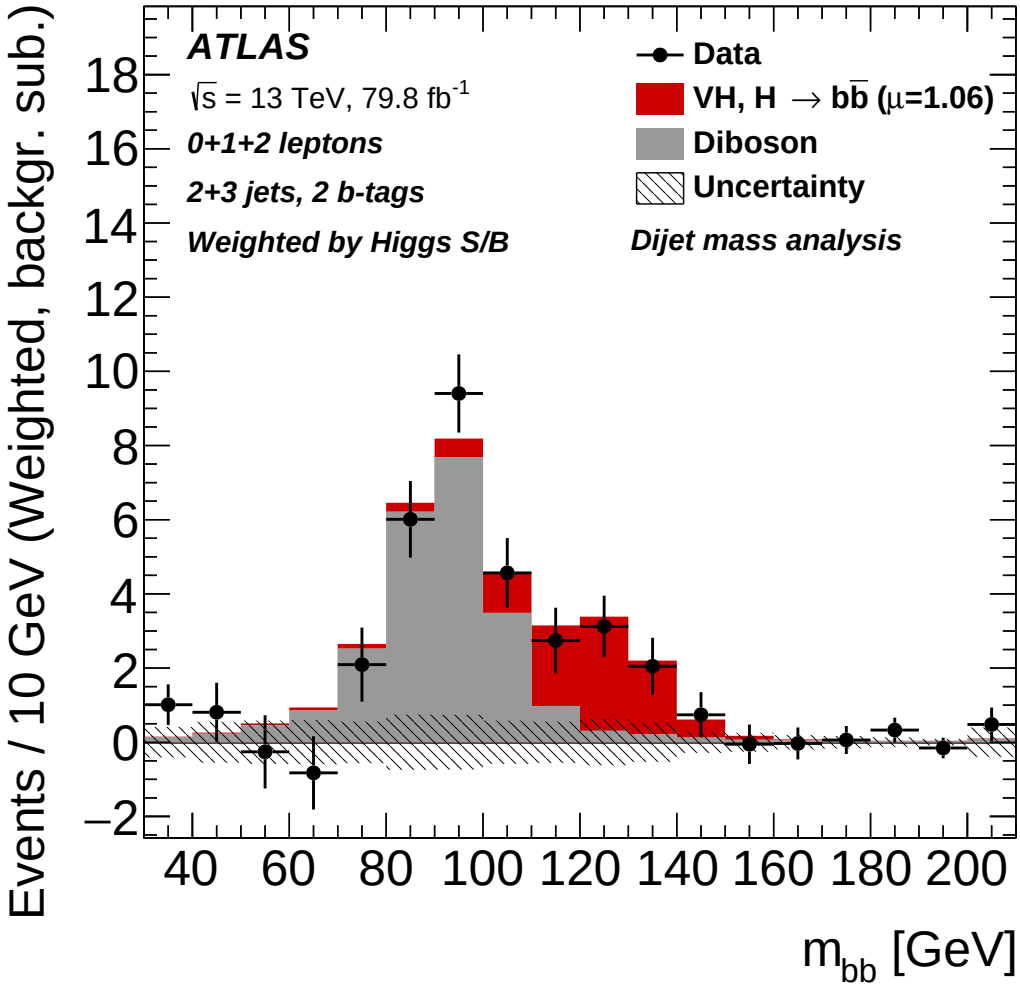


Figure 5.75: The distribution of m_{bb} in data after subtraction of all backgrounds except for the WZ and ZZ diboson processes, as obtained with the dijet-mass analysis. The contributions from all lepton channels, p_T^V regions and number-of-jets categories are summed and weighted by their respective S/B, with S being the total fitted signal and B the total fitted background in each region. The expected contribution of the associated WH and ZH production of a SM Higgs boson with m_H=125 GeV is shown scaled by the measured signal strength ($\mu = 1.06$). The size of the combined statistical and systematic uncertainty for the fitted background is indicated by the hatched band.

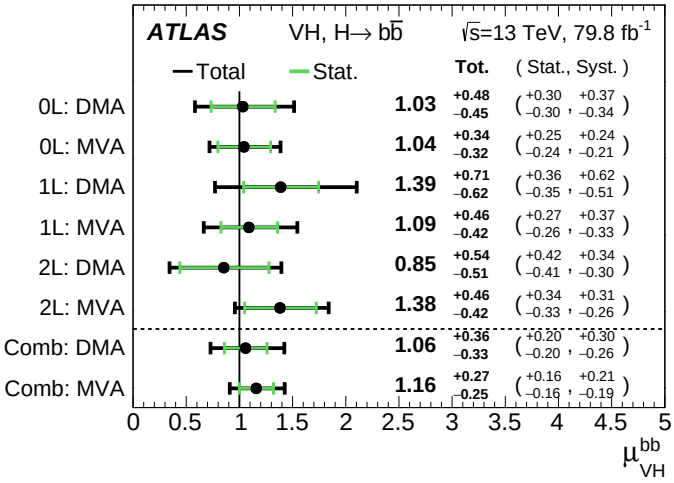


Figure 5.76: The fitted values of the Higgs boson signal strength μ_{VH}^{bb} for m_H=125 GeV for the 0-, 1- and 2-lepton channels and their combination, using the 13 TeV data. The results are shown both for the nominal multivariate analysis (MVA) and for the dijet-mass analysis (DMA). The individual μ_{VH}^{bb} values for the lepton channels are obtained from a simultaneous fit with the signal strength for each of the lepton channels floating independently.

5.9.4 Results of combination

5.9.4.1 Run 1 and Run 2 combination for VH, $H \to b\bar{b}$

The results of the main multivariate analysis of the 13 TeV data are combined with those from the data recorded at 7 TeV and 8 TeV to improve the precision of the measurement. Several studies were carried out on the correlation and compatibility of the 13 TeV results and the 7 TeV and 8 TeV results. Studies on the correlation of the experimental systematic uncertainties between the 7 TeV, 8 TeV and 13 TeV analyses were performed for the dominant uncertainties. In most cases, the impact of correlations was found to be negligible. Only a b-jet-specific jet energy scale, and theory uncertainties in the Higgs boson signal (overall cross-section, branching fraction and p_T^V dependent NLO EW corrections) are correlated across the different centre-of-mass energies.

The Run 1 and Run2 VH, $H \to b\bar{b}$ combination yields an observed significance of 4.9 standard deviations, to be compared with an expectation of 5.1 standard deviations. The measured signal strength is:

$$\mu_{VH}^{bb} = 0.98^{+0.22}_{-0.21} = 0.98 \pm 0.14(\text{stat.})^{+0.17}_{-0.16}(\text{syst.}).$$

Combined fits are also performed with floating signal strength parameters separately for the WH and ZH production processes, with the results shown in Figure 5.77. The compatibility of the signal strength parameters measured in WH and ZH production processes is 72%.

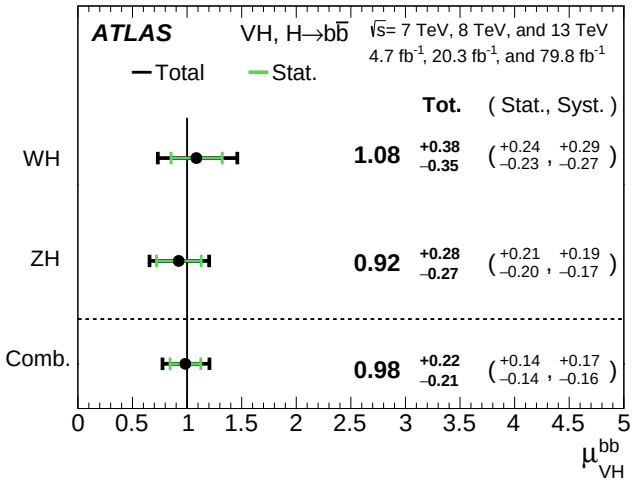


Figure 5.77: The fitted values of the Higgs boson signal strength μ_{VH}^{bb} for $m_H = 125$ GeV for the WH and ZH processes and their combination, using 7 TeV, 8 TeV and 13 TeV data. The individual μ_{VH}^{bb} values for the WH and ZH processes are obtained from a simultaneous fit with the signal strength for each of the WH and ZH processes floating independently. The probability of compatibility of the individual signal strengths is 72%.

5.9.4.2 Observation of $H \to b\bar{b}$

The VH, $H \to b\bar{b}$ result is further combined with results of the searches for the Standard Model Higgs boson decaying into a $b\bar{b}$ pair produced in vector-boson fusion (VBF) and in association with a $t\bar{t}$ pair ($t\bar{t}H$) for both Run 1 and Run 2, to improve the precision of the measurement of the $H \to b\bar{b}$ decay. As the analysis targeting the VBF production mode has a significant contribution from gluon-gluon fusion (ggF) events, it is therefore referred to as the VBF+ggF analysis in the following. The only NP correlated across the six analyses is the $H \to b\bar{b}$ branching fraction that affects the SM prediction. A few other NPs are correlated across some of the analyses, based on the dedicated studies for the combinations of Run 1 results [15], of analyses of the $t\bar{t}H$ production mode [16], and of Run 2 results. Assuming the relative production cross-sections are those predicted by the SM for a Higgs boson mass of 125 GeV, the observed significance for the $H \to b\bar{b}$ decay is 5.4 standard deviations, to be compared with an expectation

of 5.5 standard deviations. With an additional assumption that the production cross-sections are those predicted by the SM, combing all channels, the fitted value of the signal strength of the branching fraction into $b\bar{b}$ is:

$$\mu_{H\to b\bar{b}} = 1.01 \pm 0.20 = 1.01 \pm 0.12(\text{stat.})^{+0.16}_{-0.15}(\text{syst.}).$$

The significance values for the combined global likelihood fit and for the independently VBF+ggF, $t\bar{t}H$ and VH channels are presented in Table 5.40. The main contribution is from the VH channel, the VBF+ggF and $t\bar{t}H$ channels yield an observed significance of 1.5 standard derivation and 1.9 standard derivation, respectively. The combined fits are also performed with the signal strengths floated independently for each of the production processes in both Run 1 and Run 2 or combined. The results are shown in Figure 5.78 and 5.79.

Table 5.40: Expected and observed significance values (in standard deviations) for the $H \to b\bar{b}$ channels fitted independently and their combination using the 7 TeV, 8 TeV and 13 TeV data.

Channel	Significance	
	Exp.	Obs.
VBF+ggF	0.9	1.5
$t\bar{t}H$	1.9	1.9
VH	5.1	4.9
$H \to b\bar{b}$ combination	5.5	5.4

5.9.4.3 Observation of VH production

The Run 2 $VH, H \to b\bar{b}$ result is also combined with other results in the VH production mode, for the case of the Higgs boson decaying into two photons ($H \to \gamma\gamma$) or into four leptons via ZZ* ($H \to ZZ^* \to 4l$) with Run 2 79.8 fb^{-1} data. For a Higgs boson mass of 125 GeV, assuming the relative branching fractions of the three decay modes considered to be as predicted by the SM, the observed significance for VH production is 5.3 standard deviations, to be compared with an expectation of 4.8 standard deviations. The significance values for the combined global likelihood fit, and for a fit where these three decay modes have their own signal strength are shown in Table 5.41. The main contribution

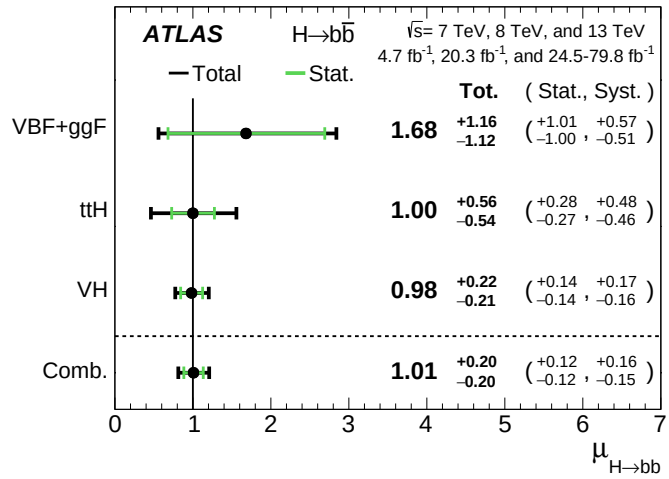


Figure 5.78: The fitted values of the Higgs boson signal strength $\mu_{H\to b\bar{b}}$ for m_H=125 GeV separately for the VH, $t\bar{t}H$ and VBF+ggF analyses and their combination, using the 7 TeV, 8 TeV and 13 TeV data. The individual $\mu_{H\to b\bar{b}}$ values for the different production modes are obtained from a simultaneous fit with the signal strengths for each of the processes floating independently. The probability of compatibility of the individual signal strengths is 83%.

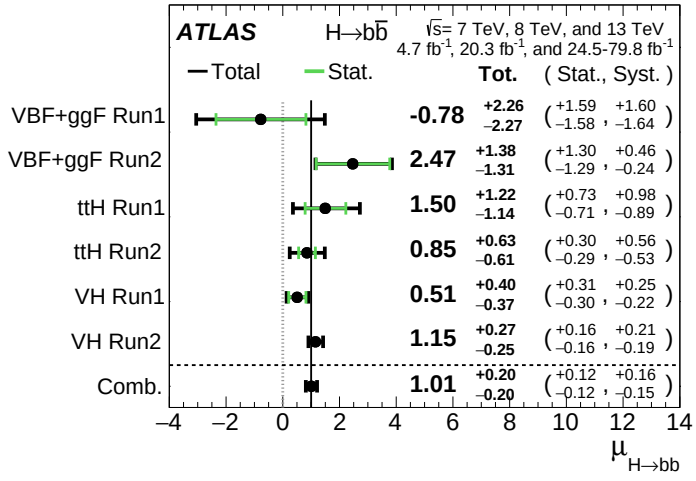


Figure 5.79: The fitted values of the Higgs boson signal strength $\mu_{H\to b\bar{b}}$ for m_H=125 GeV separately for the VH, $t\bar{t}H$ and VBF+ggF analyses in both Run 1 and Run 2, and their combination, using the 7 TeV, 8 TeV and 13 TeV data. The individual $\mu_{H\to b\bar{b}}$ values for the different production modes are obtained from a simultaneous fit with the signal strengths for each of the processes floating independently. The probability of compatibility of the individual signal strengths is 54%.

is from the $b\bar{b}$ channel, the $\gamma\gamma$ and $4l$ channels yield an observed significance of 1.9 standard derivation and 1.1 standard derivation, respectively. Assuming the branching fractions are as predicted by the SM, the fitted value of the VH signal strength for all channels combined is:

$$\mu_{VH} = 1.13^{+0.24}_{-0.23} = 1.13 \pm 0.0.15(\text{stat.})^{+0.18}_{-0.17}(\text{syst.}).$$

Figure 5.80 shows the signal strengths obtained from the fit where individual signal strengths are fitted for the three decay Modes and their combination. The probability of compatibility of the individual signal strengths is 96%.

Table 5.41: Expected and observed significance values (in standard deviations) for the VH production channels from the combined fit and from a combined fit where each of the lepton channels has its own signal strength, using 13 TeV data.

Channel	Significance	
	Exp.	Obs.
$H \to ZZ^* \to 4\ell$	1.1	1.1
$H \to \gamma\gamma$	1.9	1.9
$H \to b\bar{b}$	4.3	4.9
VH combined	4.8	5.3

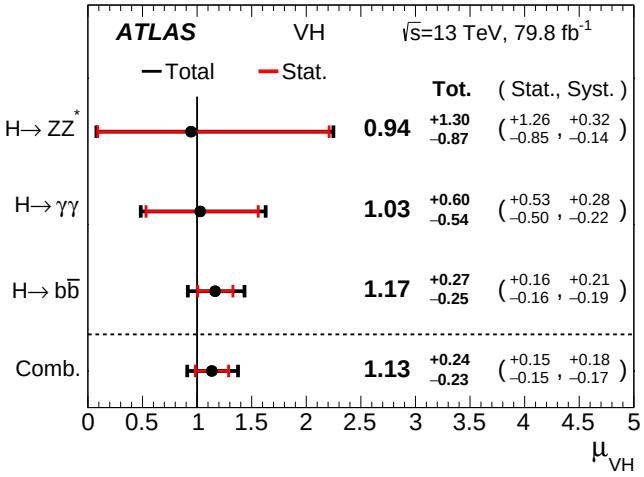


Figure 5.80: The fitted values of the Higgs boson signal strength μ_{VH} for m_H=125 GeV separately for the $H \to b\bar{b}$, $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ decay modes, along with their combination. The individual μ_{VH} values for the different decay modes are obtained from a simultaneous fit with the signal strengths for each of the processes floating independently. The probability of compatibility of the individual signal strengths is 96%.

5.10 Two Further Improvements in 1-lepton Channel

With Run 2 79.8 fb^{-1} data, for a Higgs boson mass of 125 GeV, the observed excess has a significance of 4.9 standard deviations, to be compared to an expectation of 4.3 standard deviations. Due to time constraints, some of the efforts to improve the analysis sensitivity are not included in the current official results. In this section, two further improvements in 1-lepton channel are tested in the likelihood fit and the results are presented. Section 5.10.1 shows the fit results by adding the 1-lepton medium p_T^V region in the fit. Section 5.10.2 presents the fit results by using extended $t\bar{t}$ samples in 1-lepton channel. Only the conditional fit to data with $\mu = 1$ are performed, and only the expected significance are calculated and compared when quantifying the improvements with respect to the default results.

5.10.1 Adding 1-lepton medium p_T^V region in the fit

The multijet estimation and the corresponding uncertainties in the 1-lepton medium p_T^V region has been discussed in Section 5.5.2. The multijet fraction in 1-lepton medium p_T^V 2- (3-) jet signal region are 3.57% (0.85%) and 2.76% (2.14%)

for electron and muon sub-channels, respectively. These results are used when including the medium p_T^V region in the likelihood fit. For the other backgrounds and signal processes, the same MC samples are used as those used in the high p_T^V region. The signal and background modelling uncertainties in the medium p_T^V region are not re-derived by the dedicated studies, but using the same uncertainties as derived in the high p_T^V region.

The fit is tested first by treating all the uncertainties between high and medium p_T^V regions as correlated. When adding the medium p_T^V region in the fit, the $t\bar{t}$ m_{bb}, p_T^V shape uncertainties and the 2-to-3-jet ratio uncertainties are high constrained compared with the high p_T^V only fit due to the the very high $t\bar{t}$ statistics in the medium p_T^V region. In order to prevent such constrains to be propagated to the high p_T^V region, such uncertainties are treated as uncorrelated in the fit. When decorrelating these uncertainties, the p_T^V shape uncertainty in medium p_T^V region is also highly pulled with a strong correlation with $t\bar{t}$ floating normalization. The fit is then tested with decorrelating also the $t\bar{t}$ floating normalizations between high and medium p_T^V region, and the p_T^V shape uncertainty is then no longer pulled with inconsistent $t\bar{t}$ normalization observed between high and medium p_T^V regions. Considering all the observations above, The $t\bar{t}$ related uncertainties as listed below are treated as uncorrelated between the high and medium p_T^V regions in the fit:

- Floating normalization.
- 2-to-3-jet ratio.
- m_{bb} shape uncertainties.
- p_T^V shape uncertainties.

The other uncertainties are also tested and no strong constraints and pulls observed, so correlation scheme is kept for these uncertainties.

The 1-lepton only conditional likelihood fit to data with $\mu = 1$ is performed first with adding the medium p_T^V region in. The post-fit BDT$_{VH}$ distributions in the medium p_T^V are shown in Figure 5.81, the blinding produced is performed from right to left of the BDT$_{VH}$ distributions in signal regions and 60% signal is blinded.

The breakdown of the effects of systematic uncertainties on the signal strength are presented in Table5.42, to be compared with the breakdown table (5.43) from default 1-lepton channel high p_T^V region only fit (the fit is also performed with conditional $\mu = 1$ for consistency.) As can be seen, when adding the medium p_T^V

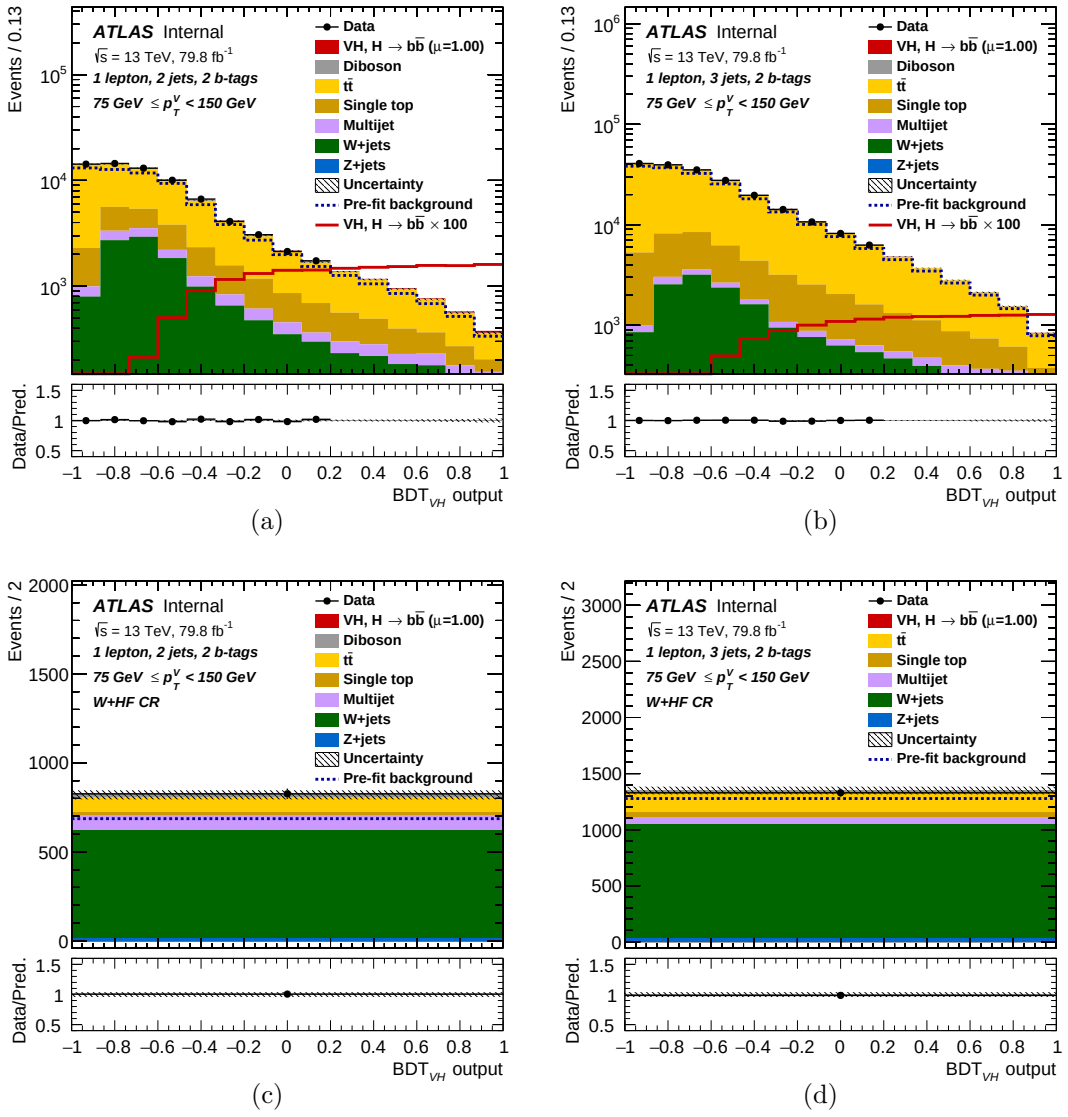


Figure 5.81: The BDT_{VH} post-fit distributions from the 1-lepton channel conditional likelihood fit ($\mu = 1$) in the 1-lepton medium p_T^V region, 2-jet SR (a), 3-jet SR (b), 2-jet W+HF CR (c), 3-jet W+HF CR (d).

region in the fit, the uncertainties from both data statistics and systematics are reduced, the total uncertainties of μ reduced from 1 ± 0.44 to 1 ± 0.41. Table 5.44 shows the comparison of the expected significances in 1-lepton channel fit with and without medium p_T^V regions included. 8.2% significance increase can be achieved by adding the medium p_T^V region in the fit. The combined global fit with 3 lepton channels are also performed with and without the 1-lepton medium p_T^V region, the expected significance are also shown in Table 5.44, as can be seen, the 1-lepton channel medium p_T^V region brings 5.5% additional sensitivity in combined global fit.

Table 5.42: Breakdown of the contributions to the uncertainty in μ for the 1-lepton channel conditional fit ($\mu = 1$) with high and medium p_T^V regions included in the fit.

POI SigXsecOverSM	Central Value 1	
Set of nuisance parameters	Impact on error	
Total	+0.430 / -0.391	±0.410
DataStat	+0.244 / -0.238	±0.241
FullSyst	+0.354 / -0.310	±0.332
Floating normalizations	+0.049 / -0.044	±0.046
Multi Jet	+0.035 / -0.033	±0.034
Modelling: single top	+0.110 / -0.100	±0.105
Modelling: ttbar	+0.094 / -0.086	±0.090
Modelling: W+jets	+0.068 / -0.065	±0.066
Modelling: Z+jets	+0.003 / -0.002	±0.002
Modelling: Diboson	+0.065 / -0.062	±0.064
Modelling: VH	+0.173 / -0.080	±0.126
Detector: lepton	+0.007 / -0.003	±0.005
Detector: MET	+0.041 / -0.037	±0.039
Detector: JET	+0.073 / -0.064	±0.068
Detector: FTAG (b-jet)	+0.040 / -0.031	±0.036
Detector: FTAG (c-jet)	+0.110 / -0.093	±0.102
Detector: FTAG (l-jet)	+0.031 / -0.026	±0.028
Detector: FTAG (extrap)	+0.019 / -0.018	±0.019
Detector: PU	+0.005 / -0.004	±0.005
Lumi	+0.024 / -0.010	±0.017
MC stat	+0.140 / -0.141	±0.140

Table 5.43: Breakdown of the contributions to the uncertainty in μ for the 1-lepton channel conditional fit ($\mu = 1$) with only high p_T^V regions included in the fit.

POI	Central Value	
SigXsecOverSM	1	
Set of nuisance parameters	Impact on error	
Total	+0.462 / -0.424	±0.443
DataStat	+0.270 / -0.262	±0.266
FullSyst	+0.375 / -0.333	±0.354
Floating normalizations	+0.058 / -0.066	±0.062
Multi Jet	+0.023 / -0.026	±0.024
Modelling: single top	+0.094 / -0.087	±0.091
Modelling: ttbar	+0.079 / -0.066	±0.072
Modelling: W+jets	+0.141 / -0.143	±0.142
Modelling: Z+jets	+0.007 / -0.007	±0.007
Modelling: Diboson	+0.056 / -0.055	±0.055
Modelling: VH	+0.177 / -0.074	±0.125
Detector: lepton	+0.010 / -0.005	±0.007
Detector: MET	+0.010 / -0.008	±0.009
Detector: JET	+0.062 / -0.034	±0.048
Detector: FTAG (b-jet)	+0.065 / -0.047	±0.056
Detector: FTAG (c-jet)	+0.106 / -0.086	±0.096
Detector: FTAG (l-jet)	+0.037 / -0.032	±0.035
Detector: FTAG (extrap)	+0.021 / -0.019	±0.020
Detector: PU	+0.003 / -0.001	±0.002
Lumi	+0.026 / -0.010	±0.018
MC stat	+0.143 / -0.148	±0.146

Table 5.44: Expected significance from the 1-lepton fit and combined global fit with and without 1-lepton channel medium p_T^V region included in the fit.

Fit	Expected significance
1-lepton channel fit without medium p_T^V region	2.32
1-lepton channel fit with medium p_T^V region	2.51
Combined global fit without 1-lepton medium p_T^V region	4.33
Combined global fit with 1-lepton medium p_T^V region	4.57

5.10.2 Using extended $t\bar{t}$ MC samples

The default $t\bar{t}$ MC sample used in 1-lepton channel is simulated with POWHEG and interfaced with PYTHIA8, and are generated with a filter at generator level using truth information to require that at least one of the W bosons decays leptonically (non-all-had). Apart from such default samples, three set of $t\bar{t}$ samples are also generated with same generator but different generator level filters:

- both of the W bosons decay leptonically (dilepton)
- non-all-had 100 GeV $< p_T^W <$ 200 GeV
- non-all-had $p_T^W >$ 200 GeV

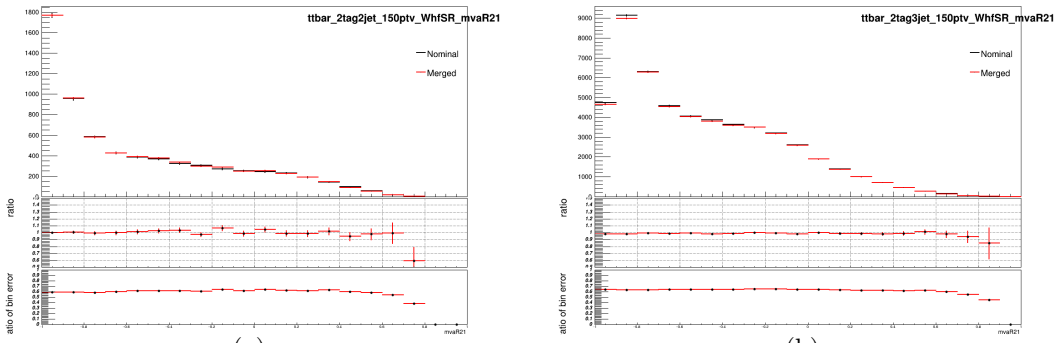
The extended new filter samples are combined with the default $t\bar{t}$ sample by following the procedures listed below:

- Part of the dilepton filter events are duplicated with the dilepton events produced with the default non-all-had filter, in that case, in order to avoid using the same events twice, the dilepton events in default $t\bar{t}$ sample are removed before combined with the high statistics dilepton filter events.
- the p_T^W filter $t\bar{t}$ events are produced independently with respect to the default events, in that case, the events are combined based on the produced event numbers as shown in Table 5.45. As can be seen, the number of events with p_T^W between 100 GeV and 200 GeV in the default sample and 100 GeV $< p_T^W <$ 200 GeV filter sample are basically the same, while the ratio for the numbers of events with $p_T^W >$ 200 GeV in default sample and $p_T^W >$ 200 GeV filter sample is about 1/3. In order to maximize the statistics increase, the event weights used for combining the default sample and p_T^W filter sample are: 0.5 for events with 100 GeV $< p_T^W <$ 200 GeV in the default sample; 0.5 for events produced with 100 GeV $< p_T^W <$ 200 GeV filter; 0.25 for events with $p_T^W >$ 200 GeV in the default sample; 0.75 for events produced with $p_T^W >$ 200 GeV filter.

Figure 5.82 shows the comparison of $t\bar{t}$ BDT$_{VH}$ distributions by using only the default $t\bar{t}$ samples and by using the comination of the default sample and all the filter samples discussed above in 2-jet and 3-jet signal regions. The middle pad shows the ratio of these two distributions and very good agreement can be achieved. The bottom pad shows the ratio of the bin errors, as can be seen, the error reduced about 40% by using the new filter $t\bar{t}$ samples.

Table 5.45: Event numbers for events with 100 GeV $< p_T^W <$ 200 GeV and $p_T^W >$ 200 GeV for the default and p_T^W filter $t\bar{t}$ events.

Event numbers	non-all-had default	non-all-had 100 GeV $< p_T^W <$ 200 GeV	non-all-had $p_T^W >$ 200 GeV
100 GeV $< p_T^W <$ 200 GeV	6273765	5818519	-
$p_T^W >$ 200 GeV	1461029	-	4029814



(a) (b)

Figure 5.82: Comparison of $t\bar{t}$ BDT$_{VH}$ distributions by using only the default $t\bar{t}$ samples and the by using all the filter samples in 1-lepton 2-jet (a) and 3-jet (b) signal regions.

The effect by using the new filter $t\bar{t}$ samples is also tested in the 1-lepton conditional ($\mu = 1$) likelihood fit, and compared with the default 1-lepton fit results. The breakdown of the effects of systematic uncertainties on the signal strength are presented in Table 5.46, to be compared with the breakdown table (5.43) from default 1-lepton channel fit (the fit is also performed with conditional $\mu = 1$ for consistency.). As can be seen, when using the new filter $t\bar{t}$ samples, the effect from MC stat is reduced from ± 0.146 to ± 0.114, and the total uncertainties of μ is reduced from 1 ± 0.44 to 1 ± 0.43. The expected significance from the fit is 2.43, compared with the expected signnificance from default 1-lepton channel fit (2.32), 4.7% significance increase achieved by using the new filter $t\bar{t}$ samples.

Table 5.46: Breakdown of the contributions to the uncertainty in μ for the 1-lepton channel conditional fit ($\mu = 1$) with high and medium p_T^V regions included in the fit.

POI SigXsecOverSM	Central Value 1	
Set of nuisance parameters	Impact on error	
Total	+0.450 / -0.412	±0.431
DataStat	+0.268 / -0.261	±0.265
FullSyst	+0.362 / -0.319	±0.340
Floating normalizations	+0.061 / -0.067	±0.064
Multi Jet	+0.022 / -0.025	±0.023
Modelling: single top	+0.093 / -0.086	±0.089
Modelling: ttbar	+0.077 / -0.064	±0.070
Modelling: W+jets	+0.143 / -0.143	±0.143
Modelling: Z+jets	+0.007 / -0.008	±0.007
Modelling: Diboson	+0.056 / -0.055	±0.055
Modelling: VH	+0.174 / -0.075	±0.124
Detector: lepton	+0.009 / -0.005	±0.007
Detector: MET	+0.022 / -0.021	±0.021
Detector: JET	+0.063 / -0.041	±0.052
Detector: FTAG (b-jet)	+0.065 / -0.044	±0.054
Detector: FTAG (c-jet)	+0.093 / -0.075	±0.084
Detector: FTAG (l-jet)	+0.035 / -0.031	±0.033
Detector: FTAG (extrap)	+0.024 / -0.021	±0.022
Detector: PU	+0.016 / -0.010	±0.013
Lumi	+0.027 / -0.011	±0.019
MC stat	+0.113 / -0.115	±0.114

Chapter 6

Conclusion and Outlook

The search for Standard Model VH, $H \to b\bar{b}$ has been carried out using a dataset corresponding to an integrated luminosity of 79.8 fb^{-1} collected by the ATLAS experiment in proton-proton collisions from Run 2 of the LHC. Extensive work has been carried out to improve the analysis sensitivity and the understanding of the main uncertainties. A data-driven method has been developed to estimate the multijet background in the 1-lepton channel, along with the detailed studies to assign proper uncertainties on the estimation. Improvements to the sensitivity and robustness of the analysis have been carried by extensive studies, such as a study of the 1-lepton channel medium p_T^V region and a study of pile up jets suppression cuts. These have been combined with the works on validating and updating the MVA training, extensive fit studies to ensure the robustness of the final results. The combined BDT$_{VH}$ fit for the main multivariate analysis yields an excess over the expected background with a significance of 4.9 standard deviations compared with an expectation of 4.3. The measured signal strength relative to the SM prediction for $m_H = 125$ GeV is found to be $\mu_{VH}^{bb} = 1.16^{+0.27}_{-0.25} = 1.16 \pm 0.16(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$, in good agreement with the SM prediction. The result is validated with a BDT$_{VZ}$ fit for the diboson analysis, and the measured signal strength is $\mu_{VZ}^{bb} = 1.20^{+0.20}_{-0.18} = 1.20 \pm 0.08(\text{stat.})^{+0.19}_{-0.16}(\text{syst.})$. The result is also cross-checked with an m_{bb} fit for the dijet-mass analysis, the VH, $H \to b\bar{b}$ signal was observed with a significance of 3.6 standard deviations compared with an expectation of 3.5, the measured signal strength is $\mu_{VH}^{bb} = 1.06^{+0.36}_{-0.33} = 1.06 \pm 020(\text{stat.})^{+0.30}_{-0.26}(\text{syst.})$, in good agreement with the result of the main multivariate analysis.

This main multivariate analysis result is first combined with previous result based on the Run 1 data collected at centre-of-mass energies of 7 TeV and 8 TeV. An excess over the expected background is observed with a significance of 4.9

standard deviations compared with an expectation of 5.1. The measured signal strength relative to the SM prediction for $m_H = 125$ GeV is found to be $\mu_{VH}^{bb} = 0.98^{+0.22}_{-0.21} = 0.98 \pm 0.14(\text{stat.})^{+0.17}_{-0.16}(\text{syst.})$.

Results for the SM Higgs boson decaying into a $b\bar{b}$ pair in the VH, $t\bar{t}H$ and VBF+ggF production modes at centre-of-mass energies of 7 TeV, 8 TeV and 13 TeV are also combined, assuming the relative production cross-sections of these processes to be as predicted by the SM. An excess over the expected background is observed with a significance of 5.4 standard deviations compared with an expectation of 5.5. The result provides an observation of the $H \to b\bar{b}$ decay mode. Assuming the SM production strengths, the measured signal strength is $\mu_{H\to b\bar{b}} = 1.01 \pm 0.20 = 1.01 \pm 0.12(\text{stat.})^{+0.16}_{-0.15}(\text{syst.})$, consistent with the value in the SM of the Yukawa coupling to bottom quarks.

The Run 2 $VH, H \to b\bar{b}$ result is also combined with the results of other Run 2 searches for the Higgs boson decaying into either four leptons (via ZZ^*) or diphotons in the VH production mode, assuming the relative branching fractions of the three decay modes to be as predicted by the SM. An excess over the expected background is observed with a significance of 5.3 standard deviations compared with an expectation of 4.8. This provides a direct observation of the Higgs boson being produced in association with a vector boson. Assuming the SM branching fractions, the measured signal strength is $\mu_{VH} = 1.13^{+0.24}_{-0.23} = 1.13 \pm 0.15(\text{stat.})^{+0.18}_{-0.17}(\text{syst.})$, consistent with the SM prediction.

The observation of $H \to b\bar{b}$ decays and VH production has been established with the results presented in this thesis. All the measurements are consistent with SM predictions so far. Nevertheless, the uncertainties for the measurement of $H \to b\bar{b}$ decays are still at the level of 20%, which does not rule out new physics beyond SM in term of the large $H \to b\bar{b}$ branching ratio predicted by the SM. With more data delivered by the LHC in the future, more precision measurements are absolutely needed to probe in more details of the Higgs boson properties and to detect any sign of the new physics. As can be seen in Figure 6.1, about 60 fb^{-1} data were recorded by ATLAS during 2018 data-taking. In total, ATLAS recorded about 140 fb^{-1} data during Run 2 data-taking. Another 150 fb^{-1} data are expected for Run 3 data-taking, and the goal of the total integrated luminosity for HL-LHC is 3000 fb^{-1}. One of the straightforward way for the precision measurements is using the simplified template cross section (STXS) framework [131] to measure the cross section of the $H \to b\bar{b}$ decays as a function of the Higgs boson p_T with reduced theoretical uncertainties, as we know the Higgs boson p_T spectrum is highly sensitive to new physics [139], especially in the high Higgs p_T regime. Apart

from the current jet reconstruction techniques used in research work presented in this thesis, the boosted analysis techniques provide another excellent opportunity to improve the analysis sensitivity at high p_T phase space. Any deviation from the SM provided by the precision measurements may indicates the new physics and open a new window for the better understanding of the Higgs boson, particle physics, and the world we are living in.

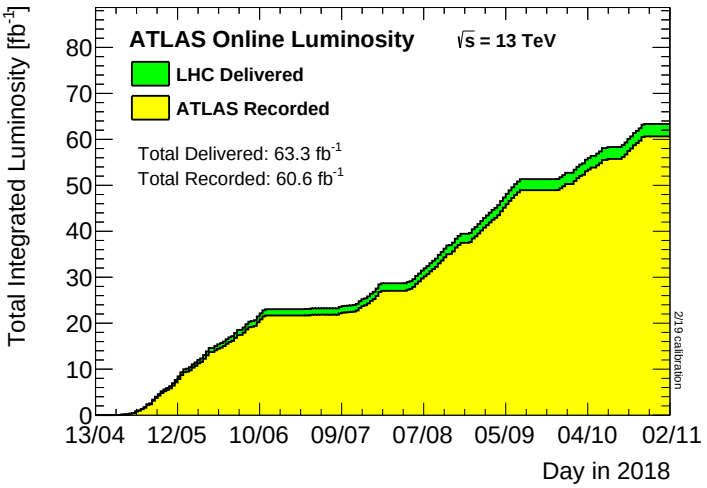


Figure 6.1: Cumulative luminosity versus time delivered by LHC (green) and recorded by ATLAS (yellow) during stable beams for pp collisions at $\sqrt{s} = 13\,\text{TeV}$ for the year of 2018.

References

[1] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", in: *Phys. Rev. Lett.* 13 (1964), pp. 321–323, DOI: 10.1103/PhysRevLett.13.321.

[2] Peter W. Higgs, "Broken symmetries, massless particles and gauge fields", in: *Phys. Lett.* 12 (1964), pp. 132–133, DOI: 10.1016/0031-9163(64)91136-9.

[3] Peter W. Higgs, "Broken symmetries and the masses of gauge bosons", in: *Phys. Rev. Lett.* 13 (1964), pp. 508–509, DOI: 10.1103/PhysRevLett.13.508.

[4] G.S. Guralnik, C.R. Hagen, and T.W.B. Kibble, "Global conservation laws and massless particles", in: *Phys. Rev. Lett.* 13 (1964), pp. 585–587, DOI: 10.1103/PhysRevLett.13.585.

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

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

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

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

[9] ATLAS Collaboration, "Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector", in: (2018), arXiv: 1802.04146 [hep-ex].

[10] ATLAS Collaboration, "Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector", in: *JHEP* 03 (2018), p. 095, DOI: 10.1007/JHEP03(2018)095, arXiv: 1712.02304 [hep-ex].

[11] ATLAS Collaboration, "Measurement of inclusive and differential cross sections in the $H \to ZZ^* \to 4\ell$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", in: *JHEP* 10 (2017), p. 132, DOI: 10.1007/JHEP10(2017)132, arXiv: 1708.02810 [hep-ex].

[12] CMS Collaboration, "Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 8$ TeV", in: *Eur. Phys. J. C* 76 (2016), p. 13, DOI: 10.1140/epjc/s10052-015-3853-3, arXiv: 1508.07819 [hep-ex].

[13] CMS Collaboration, "Measurement of the transverse momentum spectrum of the Higgs boson produced in pp collisions at $\sqrt{s} = 8$ TeV using $H \to WW$ decays", in: *JHEP* 03 (2017), p. 032, DOI: 10.1007/JHEP03(2017)032, arXiv: 1606.01522 [hep-ex].

[14] CMS Collaboration, "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV", in: *JHEP* 11 (2017), p. 047, DOI: 10.1007/JHEP11(2017)047, arXiv: 1706.09936 [hep-ex].

[15] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", in: *JHEP* 08 (2016), p. 045, DOI: 10.1007/JHEP08(2016)045, arXiv: 1606.02266 [hep-ex].

[16] ATLAS Collaboration, "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector", in: *Submitted to Phys. Lett. B* (2018), arXiv: 1806.00425 [hep-ex].

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

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

[19] S. L. Glashow, Dimitri V. Nanopoulos, and A. Yildiz, "Associated production of Higgs bosons and Z particles", in: *Phys. Rev. D* 18 (1978), pp. 1724–1727, DOI: 10.1103/PhysRevD.18.1724.

[20] R. Lafaye et al., "Measuring the Higgs Sector", in: *JHEP* 08 (2009), p. 009, DOI: 10.1088/1126-6708/2009/08/009, arXiv: 0904.3866 [hep-ph].

[21] LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", in: *CERN-2013-004* (2013), arXiv: 1307.1347 [hep-ph].

[22] John Ellis, Veronica Sanz, and Tevong You, "Complete Higgs Sector Constraints on Dimension-6 Operators", in: *JHEP* 07 (2014), p. 036, DOI: 10.1007/JHEP07(2014)036, arXiv: 1404.3667 [hep-ph].

[23] CDF Collaboration and DZero Collaboration, 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].

[24] ATLAS Collaboration, "Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector", in: *JHEP* 01 (2015), p. 069, DOI: 10.1007/JHEP01(2015)069, arXiv: 1409.6212 [hep-ex].

[25] CMS Collaboration, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks", in: *Phys. Rev. D* 89 (2014), p. 012003, DOI: 10.1103/PhysRevD.89.012003, arXiv: 1310.3687 [hep-ex].

[26] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", in: *JHEP* 08 (2016), p. 045, DOI: 10.1007/JHEP08(2016)045, arXiv: 1606.02266 [hep-ex].

[27] ATLAS Collaboration, "Search for the Standard Model Higgs boson produced by vector-boson fusion and decaying to bottom quarks in $\sqrt{s}$ = 8 TeV pp collisions with the ATLAS detector", in: *JHEP* 11 (2016), p. 112, DOI: 10.1007/JHEP11(2016)112, arXiv: 1606.02181 [hep-ex].

[28] CMS Collaboration, "Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$", in: *Phys. Rev. D* 92 (2015), p. 032008, DOI: 10.1103/PhysRevD.92.032008, arXiv: 1506.01010 [hep-ex].

[29] ATLAS Collaboration, *Search for Higgs boson production via weak boson fusion and decaying to $b\bar{b}$ in association with a high-energy photon using the ATLAS detector*, ATLAS-CONF-2016-063, 2016, URL: https://cds.cern.ch/record/2206201.

[30] ATLAS Collaboration, "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", in: *Eur. Phys. J. C* 75 (2015), p. 349, DOI: 10.1140/epjc/s10052-015-3543-1, arXiv: 1503.05066 [hep-ex].

[31] ATLAS Collaboration, "Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", in: *JHEP* 05 (2016), p. 160, DOI: 10.1007/JHEP05(2016)160, arXiv: 1604.03812 [hep-ex].

[32] ATLAS Collaboration, "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", in: *Phys. Rev. D* (2017), arXiv: 1712.08895 [hep-ex].

[33] CMS Collaboration, "Search for a standard model Higgs boson produced in association with a top-quark pair and decaying to bottom quarks using a matrix element method", in: *Eur. Phys. J. C* 75 (2015), p. 251, DOI: 10.1140/epjc/s10052-015-3454-1, arXiv: 1502.02485 [hep-ex].

[34] CMS Collaboration, "Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV", in: (2018), arXiv: 1804.03682 [hep-ex].

[35] CMS Collaboration, "Inclusive Search for a Highly Boosted Higgs Boson Decaying to a Bottom Quark–Antiquark Pair", in: *Phys. Rev. Lett.* 120 (2018), p. 071802, DOI: 10.1103/PhysRevLett.120.071802, arXiv: 1709.05543 [hep-ex].

[36] W. N. Cottingham and D. A. Greenwood, *An introduction to the Standard Model of particle physics*, University of Bristol, UK: Cambridge University Press, 2007.

[37] C. Patrignani and Particle Data Group, "Review of Particle Physics", in: *Chinese Physics C* 40 (2016), p. 100001, URL: http://stacks.iop.org/1674-1137/40/i=10/a=100001.

[38] Steven Weinberg, "A model of leptons", in: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266, DOI: 10.1103/PhysRevLett.19.1264.

[39] A. Salam, "Electromagnetic and weak interactions", in: *Phys. Rev. Lett.* 13 (1964), p. 168.

[40] D. M. Webber et al., "Measurement of the Positive Muon Lifetime and Determination of the Fermi Constant to Part-per-Million Precision", in: *Phys. Rev. Lett.* 106 (4 Jan. 2011), p. 041803, DOI: 10.1103/PhysRevLett.106.041803, URL: https://link.aps.org/doi/10.1103/PhysRevLett.106.041803.

[41] Timo Antero Aaltonen et al., "Combination of CDF and D0 W-Boson Mass Measurements", in: *Phys. Rev.* D88.5 (2013), p. 052018, DOI: 10.1103/PhysRevD.88.052018, arXiv: 1307.7627 [hep-ex].

[42] Morad Aaboud et al., "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].

[43] Ringaile Placakyte, "Parton Distribution Functions", in: *Proceedings, 31st International Conference on Physics in collisions (PIC 2011): Vancouver, Canada, August 28-September 1, 2011*, 2011, arXiv: 1111.5452 [hep-ph].

[44] Guido Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", in: *Nucl. Phys.* B126 (1977), pp. 298–318, DOI: 10.1016/0550-3213(77)90384-4.

[45] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory", in: *Sov. J. Nucl. Phys.* 15 (1972), [Yad. Fiz.15,781(1972)], pp. 438–450.

[46] A. M. Cooper-Sarkar, "PDF Fits at HERA", in: *PoS* EPS-HEP2011 (2011), p. 320, DOI: 10.22323/1.134.0320, arXiv: 1112.2107 [hep-ph].

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

[48] 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].

[49] 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].

[50] Morad Aaboud et al., "Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector", in: *Phys. Lett.* B784 (2018), pp. 345–366, DOI: 10.1016/j.physletb.2018.07.050, arXiv: 1806.00242 [hep-ex].

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

[52] A. Augusto Alves Jr. et al., "The LHCb Detector at the LHC", in: *JINST* 3 (2008), S08005, DOI: 10.1088/1748-0221/3/08/S08005.

[53] The ALICE Collaboration et al., "The alice experiment at the CERN LHC", in: *Journal of Instrumentation* 3 (Aug. 2008), S08002, DOI: 10.1088/1748-0221/3/08/S08002.

[54] The LHCf Collaboration et al., "The LHCf detector at the CERN Large Hadron Collider", in: *Journal of Instrumentation* 3.08 (2008), S08006, URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08006.

[55] J. L. Pinfold, "The MoEDAL Experiment at the LHC - a New Light on the Terascale Frontier", in: *Journal of Physics Conference Series*, vol. 631, Journal of Physics Conference Series, July 2015, p. 012014, DOI: 10.1088/1742-6596/631/1/012014.

[56] The TOTEM Collaboration, "The TOTEM Experiment at the CERN Large Hadron Collider", in: *Journal of Instrumentation* 3.08 (2008), S08007, URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08007.

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

[58] "ATLAS central solenoid: Technical design report", in: (1997).

[59] "ATLAS barrel toroid: Technical design report", in: (1997).

[60] "ATLAS endcap toroids: Technical design report", in: (1997).

[61] ATLAS Collaboration, "The ATLAS Inner Detector commissioning and calibration", in: *Eur. Phys. J. C* 70 (2010), p. 787, DOI: 10.1140/epjc/s10052-010-1366-7, arXiv: 1004.5293 [hep-ex].

[62] M. Capeans et al., "ATLAS Insertable B-Layer Technical Design Report", in: (2010).

[63] G. Aad et al., "Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions", in: *Eur. Phys. J.* C70 (2010), pp. 723–753, DOI: 10.1140/epjc/s10052-010-1354-y, arXiv: 0912.2642 [physics.ins-det].

[64] G. Aad et al., "Readiness of the ATLAS Tile Calorimeter for LHC collisions", in: *Eur. Phys. J.* C70 (2010), pp. 1193–1236, DOI: 10.1140/epjc/s10052-010-1508-y, arXiv: 1007.5423 [physics.ins-det].

[65] "ATLAS muon spectrometer: Technical design report", in: (1997).

[66] Peter Jenni et al., "ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity", in: (2008).

[67] Sebastian White, "The ATLAS zero degree calorimeter", in: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 617.1 (2010), 11th Pisa Meeting on Advanced Detectors, pp. 126–128, ISSN: 0168-9002, DOI: https://doi.org/10.1016/j.nima.2009.09.120, URL: http://www.sciencedirect.com/science/article/pii/S0168900209019044.

[68] A. Mapelli, "ALFA: Absolute Luminosity For ATLAS - Development of a scintillating fibre tracker to determine the absolute LHC luminosity at ATLAS", in: *Nuclear Physics B - Proceedings Supplements* 197.1 (2009), 11th Topical Seminar on Innovative Particle and Radiation Detectors (IPRD08), pp. 387–390, ISSN: 0920-5632, DOI: https://doi.org/10.1016/j.nuclphysbps.2009.10.110, URL: http://www.sciencedirect.com/science/article/pii/S0920563209008317.

[69] "ATLAS first level trigger: Technical design report", in: (1998).

[70] "ATLAS high-level trigger, data acquisition and controls: Technical design report", in: (2003).

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

[72] 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.

[73] Wolfgang Lukas, "Fast Simulation for ATLAS: Atlfast-II and ISF", in: *Journal of Physics: Conference Series* 396.2 (2012), p. 022031, URL: http://stacks.iop.org/1742-6596/396/i=2/a=022031.

[74] ATLAS Collaboration, "Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2", in: *The European Physical Journal C* 77.10 (Oct. 2017), p. 673, DOI: 10.1140/epjc/s10052-017-5225-7, URL: https://doi.org/10.1140/epjc/s10052-017-5225-7.

[75] ATLAS Collaboration, "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC", in: *Eur. Phys. J. C* 77.5 (2017), p. 332, DOI: 10.1140/epjc/s10052-017-4887-5, arXiv: 1611.10235 [physics.ins-det].

[76] Georges Aad et al., "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].

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

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

[79] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s} = 13$ *TeV*, ATLAS-CONF-2017-029, 2017, URL: https://cds.cern.ch/record/2261772.

[80] Walter Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, 2008, URL: https://cds.cern.ch/record/1099735.

[81] 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].

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

[83] ATLAS Collaboration, *Selection of jets produced in* 13 *TeV proton–proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015, URL: https://cds.cern.ch/record/2037702.

[84] ATLAS Collaboration, "Performance of pile-up mitigation techniques for jets in *pp* collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector", in: *Eur. Phys. J. C* 76.11 (2016), p. 581, DOI: 10.1140/epjc/s10052-016-4395-z, arXiv: 1510.03823 [hep-ex].

[85] ATLAS Collaboration, *Forward Jet Vertex Tagging: A new technique for the identification and rejection of forward pileup jets*, ATL-PHYS-PUB-2015-034, 2015, URL: http://cds.cern.ch/record/2042098.

[86] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, URL: https://cds.cern.ch/record/2273281.

[87] ATLAS Collaboration, "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV", in: (2018), arXiv: 1802.08168 [hep-ex].

[88] ATLAS Collaboration, "Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC", in: *Phys. Lett. B* 726 (2013), p. 88, DOI: 10.1016/j.physletb.2014.05.011, arXiv: 1307.1427 [hep-ex], Erratum: in: *Phys. Lett. B* 734 (2014), p. 406, DOI: 10.1016/j.physletb.2014.05.011.

[89] ATLAS Collaboration, "Evidence for the spin-0 nature of the Higgs boson using ATLAS data", in: *Phys. Lett. B* 726 (2013), p. 120, DOI: 10.1016/j.physletb.2013.08.026, arXiv: 1307.1432 [hep-ex].

[90] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", in: *JHEP* 06 (2013), p. 081, DOI: 10.1007/JHEP06(2013)081, arXiv: 1303.4571 [hep-ex].

[91] ATLAS Collaboration, *Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, ATLAS-CONF-2017-046, 2017, URL: https://cds.cern.ch/record/2273853.

[92] ATLAS Collaboration, "Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC", in: *Eur. Phys. J. C* 76.12 (2016), p. 653, DOI: 10.1140/epjc/s10052-016-4466-1, arXiv: 1608.03953 [hep-ex].

[93] 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].

[94] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012, URL: https://cds.cern.ch/record/1474107.

[95] A. D. Martin et al., "Parton distributions for the LHC", in: *Eur. Phys. J. C* 63 (2009), p. 189, DOI: 10.1140/epjc/s10052-009-1072-5, arXiv: 0901.0002 [hep-ph].

[96] D. J. Lange, "The EvtGen particle decay simulation package", in: *Nucl. Instrum. Meth. A* 462 (2001), pp. 152–155, DOI: 10.1016/S0168-9002(01)00089-4.

[97] 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].

[98] 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].

[99] 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].

[100] 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].

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

[102] 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].

[103] 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].

[104] 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].

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

[106] Oliver Brein, Abdelhak Djouadi, and Robert Harlander, "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders", in: *Phys. Lett.* B579 (2004), pp. 149–156, DOI: 10.1016/j.physletb.2003.10.112, arXiv: hep-ph/0307206 [hep-ph].

[107] Giancarlo Ferrera, Massimiliano Grazzini, and Francesco Tramontano, "Associated WH production at hadron colliders: a fully exclusive QCD calculation at NNLO", in: *Phys. Rev. Lett.* 107 (2011), p. 152003, DOI: 10.1103/PhysRevLett.107.152003, arXiv: 1107.1164 [hep-ph].

[108] Oliver Brein et al., "Top-Quark Mediated Effects in Hadronic Higgs-Strahlung", in: *Eur. Phys. J.* C72 (2012), p. 1868, DOI: 10.1140/epjc/s10052-012-1868-6, arXiv: 1111.0761 [hep-ph].

[109] Giancarlo Ferrera, Massimiliano Grazzini, and Francesco Tramontano, "Higher-order QCD effects for associated WH production and decay at the LHC", in: *JHEP* 04 (2014), p. 039, DOI: 10.1007/JHEP04(2014)039, arXiv: 1312.1669 [hep-ph].

[110] Giancarlo Ferrera, Massimiliano Grazzini, and Francesco Tramontano, "Associated ZH production at hadron colliders: the fully differential NNLO QCD calculation", in: *Phys. Lett.* B740 (2015), pp. 51–55, DOI: 10.1016/j.physletb.2014.11.040, arXiv: 1407.4747 [hep-ph].

[111] John M. Campbell, R. Keith Ellis, and Ciaran Williams, "Associated production of a Higgs boson at NNLO", in: *JHEP* 06 (2016), p. 179, DOI: 10.1007/JHEP06(2016)179, arXiv: 1601.00658 [hep-ph].

[112] 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].

[113] B. Hespel, F. Maltoni, and E. Vryonidou, "Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM", in: *JHEP* 06 (2015), p. 065, DOI: 10.1007/JHEP06(2015)065, arXiv: 1503.01656 [hep-ph].

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

[115] Robert V. Harlander, Stefan Liebler, and Tom Zirke, "Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-Doublet Model", in: *JHEP* 02 (2014), p. 023, DOI: 10.1007/JHEP02(2014)023, arXiv: 1307.8122 [hep-ph].

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

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

[118] 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].

[119] Simone Alioli et al., "NLO single-top production matched with shower in Powheg: s- and t-channel contributions", in: *JHEP* 09 (2009), Erratum: JHEP **02** (2010) 011, p. 111, DOI: 10.1007/JHEP02(2010)011,10.1088/1126-6708/2009/09/111, arXiv: 0907.4076 [hep-ph].

[120] 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].

[121] Nikolaos Kidonakis, "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production", in: *Phys. Rev.* D83 (2011), p. 091503, DOI: 10.1103/PhysRevD.83.091503, arXiv: 1103.2792 [hep-ph].

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

[123] 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].

[124] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini, "Scattering amplitudes with open loops", in: *Phys. Rev. Lett.* 108 (2012), p. 111601, DOI: 10.1103/PhysRevLett.108.111601, arXiv: 1111.5206 [hep-ph].

[125] T. Gleisberg and S. Höche, "Comix, a new matrix element generator", in: *JHEP* 12 (2008), p. 039, DOI: 10.1088/1126-6708/2008/12/039, arXiv: 0808.3674 [hep-ph].

[126] Steffen Schumann and Frank Krauss, "A Parton shower algorithm based on Catani-Seymour dipole factorisation", in: *JHEP* 03 (2008), p. 038, DOI: 10.1088/1126-6708/2008/03/038, arXiv: 0709.1027 [hep-ph].

[127] Stefan Höche et al., "QCD matrix elements + parton showers: The NLO case", in: *JHEP* 04 (2013), p. 027, DOI: 10.1007/JHEP04(2013)027, arXiv: 1207.5030 [hep-ph].

[128] S. Catani et al., "Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO", in: 103 (2009), p. 082001, DOI: 10.1103/PhysRevLett.103.082001, arXiv: 0903.2120 [hep-ph].

[129] Andreas Hoecker et al., *TMVA: Toolkit for multivariate data analysis*, 2007, arXiv: physics/0703039.

[130] I. Antcheva et al., "ROOT - A C++ framework for petabyte data storage, statistical analysis and visualization", in: *Computer Physics Communications* 180.12 (2009), 40 YEARS OF CPC: A celebratory issue focused on quality software for high performance, grid and novel computing architectures, pp. 2499–2512, ISSN: 0010-4655, DOI: https://doi.org/10.1016/j.cpc.2009.08.005, URL: http://www.sciencedirect.com/science/article/pii/S0010465509002550.

[131] D. de Florian et al., *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*, 2016, arXiv: 1610.07922 [hep-ph].

[132] ATLAS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector at the LHC", in: *Phys. Rev. Lett.* 117 (2016), p. 182002, DOI: 10.1103/PhysRevLett.117.182002, arXiv: 1606.02625 [hep-ex].

[133] 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 (2013), p. 2306, DOI: 10.1140/epjc/s10052-013-2306-0, arXiv: 1210.6210 [hep-ex].

[134] ATLAS Collaboration, "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV", in: (2018), arXiv: 1805.01845 [hep-ex].

[135] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach wth the ATLAS detector*, ATLAS-CONF-2018-001, 2018, URL: https://cds.cern.ch/record/2306649.

[136] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2018-006, 2018, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-006/.

[137] Glen Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", in: *Eur. Phys. J.* C71 (2011), [Erratum: Eur. Phys. J. C73 (2013) 2501], p. 1554, DOI: 10.1140/epjc/s10052-011-1554-0, arXiv: 1007.1727 [physics.data-an].

[138] Roger J. Barlow and Christine Beeston, "Fitting using finite Monte Carlo samples", in: *Comput. Phys. Commun.* 77 (1993), pp. 219–228, DOI: 10.1016/0010-4655(93)90005-W.

[139] Anke Biektter et al., "Vices and virtues of Higgs effective field theories at large energy", in: *Phys. Rev.* D91 (2015), p. 055029, DOI: 10.1103/PhysRevD.91.055029, arXiv: 1406.7320 [hep-ph].

Titre : Observation du mode de désintégration H→bb et de la production associée de VH avec le détecteur ATLAS

Mots clés : LHC, expérience ATLAS, boson de Higgs, production associée de VH, quark b

Résumé : Une recherche du boson de Higgs du Modèle Standard produit en association avec un boson W ou Z et se désintégrant en une paire quark-antiquark b a été effectuée avec le détecteur ATLAS. Les données de collisions proton-proton utilisées ont été accumulées durant le Run 2 du Grand Collisionneur de Hadrons du CERN à une énergie dans le centre de masse de 13 TeV, et correspondent à une luminosité intégrée de 79.8 fb^{-1}. Trois canaux contenant zéro, un ou deux leptons chargés (électrons ou muons) sont considérés, correspondant à chacune des désintégrations leptoniques d'un boson W ou Z: Z→*νν*, W→l*ν* et Z→ll. Pour un boson de Higgs de masse 125 GeV, un excès d'événements par rapport aux bruits de fonds des autres processus du Modèle Standard est observé avec un niveau de signification statistique de 4.9 déviations standard, à comparer à 4.3 attendues. Le rapport du nombre d'événements observé au nombre attendu est mesuré être 1.16 +0.27/-0.25 = 1.16 +/-0.16(stat) +0.21/-0.19(syst). Ce résultat est combiné avec d'autres d'ATLAS sur la recherche du boson de Higgs se désintégrant dans le mode bbbar, utilisant des données du Run 1 et du Run 2. Le niveau de signification mesuré (attendu) pour ce mode de désintégraion est de 5.4 (5.5) déviations standard, ce qui en constitue la première observation directe. De plus, une combinaison des résultats du Run 2 sur la recherche de la production associée du boson de Higgs et d'un boson W ou Z conduit à un niveau de signification observé (attendu) de 5.3 (4.8) déviations standard, et donc à la première observation de ce mode de production.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

université PARIS-SACLAY | ÉCOLE DOCTORALE PHENIICS

Title : Observation of H→bb decays and VH production with the ATLAS detector

Keywords : LHC, ATLAS experiment, Higgs boson, VH associated production, b-quark

Abstract : A search for the Standard Model Higgs boson produced in association with a W or Z boson, and decaying to a bb pair has been performed with ATLAS detector. The data were collected in proton-proton collisions during Run 2 of the Large Hadron Collider at a centre-of-mass energy of 13 TeV, and correspond to an integrated luminosity of 79.8 fb^{-1}. Three channels containing zero, one and two charged leptons (electrons or muons) have been considered to target each of the leptonic decays of the W or Z boson, Z→$\nu\nu$, W→lν et Z→ll, referred to as the 0-lepton, 1-lepton and 2-lepton channels, respectively. For a Higgs boson mass of 125 GeV, an excess of events over the expected background from other Standard Model processes is found with an observed significance of 4.9 standard deviations, compared to an expectation of 4.3 standard deviations. The ratio of the measured signal events to the Standard Model expectation equal 1.16 +0.27/-0.25 = 1.16 +/-0.16(stat) +0.21/-0.19(syst). The result is also combined with the other results from the searches for the Higgs boson in the bb decay mode in Run 1 and Run 2, the combination yields an observed (expected) significance of 5.4 (5.5) standard deviations, and therefore provides a direct observation of the Higgs boson decay into a bb pair. In addition, a combination of Run 2 results searching for the Higgs boson produced in association with a W or Z boson yields an observed (expected) significance of 5.3 (4.8) standard deviations, and therefore provides a direct observation of Higgs boson being produced in association with a W or Z boson.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France