CPPM-T-2015-07

AIX-MARSEILLE UNIVERSITÉ

Faculté des Sciences – Département de Physique

Ecole Doctorale 352: Physique et Sciences de la Matière (ED352)

Centre de Physique des Particules de Marseille

Thèse présentée pour obtenir le grade universitaire de docteur

Discipline : Physique et Sciences de la Matière
Spécialité: Physique des Particules et Astrophysique

Kazuya MOCHIZUKI

Search for the Higgs boson in the $WH \to \ell\nu b\bar{b}$ channel with the ATLAS detector
Development of high performance b-jet identification algorithms

Soutenue le 27/10/2015 devant le jury:

Meenakshi NARAIN	Brown University, Etats-Unis d'Amérique	Rapporteur
Yuji ENARI	University of Tokyo, ICEPP, Japon	Rapporteur
Eric KAJFASZ	Aix-Marseille Université, CPPM, France	Examinateur, président du jury
Christian WEISER	Universitaet Freiburg, Allemagne	Examinateur
Laurent VACAVANT	Aix-Marseille Université, CPPM, France	Directeur de thèse
Yann COADOU	Aix-Marseille Université, CPPM, France	Co-directeur de thèse

Acknowledgments

To finish my PhD, I had many helps from various aspects in my research and private life. The reason I can be here as a PhD cannot be explained without their warm supports. Here, I would like to tell my gratitude to them from my heart.

First of all, I would like to express my deep gratitude to my supervisors for their patience, motivation, and immense knowledge. They both gave continuous support on my PhD and related research. Dr. L. Vacavant always led me with his very bright scopes while being a great team-leader. Dr. Y. Coadou always helped me with his careful observation and thoughtful explanation, for not only public but also private aspects. They made my PhD experience beyond what I imagined, and I can never thank enough for this experience under their supervision. Still, there are a lot I could not absorb from them. I will continue to chase them throughout my research life so that I can be like them someday.

Besides my supervisors, I would like to thank my thesis committee: Prof. M. Narain, Prof. Y. Enari, and Dr. C. Weiser, for their insightful comments including hard questions. My sincere thanks go to Dr. E. Kajfasz for organizing the thesis committee and CPPM throughout my PhD period.

The ATLAS CPPM group was always active in many physics analysis and hardware developments in the collaboration, and I would like to thank all of the staff members. Among them, I am particularly grateful to Dr. G. Aad, who helped my b-tagging research with his sharp and experienced suggestions, and Dr. Y. Nagai, who gave me unceasing advice on the analysis presented in this thesis.

Mr. E. Le Guirrec, Ms. E. Knoops and Mr. T. Mouthy helped me from the computational and technical aspects. I could not process the analysis jobs presented in this thesis without their continuous management of the grid and the local computing farm. For a lot of support to solve many issues in my administrative aspects, I am grateful to Ms. E. Garnier.

I was lucky to have my friends in the lab, who helped my PhD life to be bright through the stimulating discussions and all the fun we had together. Thank you, T. Calvet, V. Ellajosyula, A.Hadef, S. Kahn, W. Kanso, D. Madaffari, A. Morda, Y. Rodina, R. E. Ticse Torres, M. Ughetto, and R. Wolff.

Last but not the least, I would like to express my thanks to my family and relatives for supporting me since my childhood until now. I could finally place a period on my long life as a student. As a partner, Ms. N. Yamamoto has supported me since the beginning of my research life in high energy physics. Aligatou with love, until now and from now on.

Contents

List of Figures

List of Tables

Introduction

The Standard Model (SM) of particle physics describes precisely the elementary particles and the interactions between them. It comprises a mechanism which gives masses to the elementary particles, especially to the weak force mediators, W and Z bosons. This is the so-called Brout-Englert-Higgs (BEH) mechanism, introduced about half a century ago. The BEH mechanism predicts a field, which interacts with the elementary particles and thereby gives masses to all particles, except the massless ones. At the same time, a particle, the so-called Higgs boson, was predicted by this mechanism, however its mass was not predictable by the theory.

After the discovery of top-quark and tau-neutrino at Fermilab in 1995 and 2000, respectively, the Higgs boson became the last missing piece of the SM. In spite of many searches for the Higgs boson, performed over decades by many experiments, the signal was not observed. The most recent direct searches were performed by the LEP collider at CERN and the Tevatron collider at Fermilab using electron-positron collisions and proton-antiproton collisions, respectively, before the LHC started its stable operation since 2010. In July 2012, the ATLAS and CMS collaboration announced the discovery of a new particle at mass of about 125 GeV, which was eventually confirmed as compatible with the Higgs boson predicted by the SM. Although all results available currently show a full compatibility of the Higgs boson properties with the SM prediction, the $H \to b\bar{b}$ decay mode has not been seen clearly yet contrary to the other decay modes ($H \to \gamma\gamma, WW^*$, ZZ^*, and $\tau\tau$).

Despite the largest branching fraction of $H \to b\bar{b}$ at $m_H = 125$ GeV, this decay mode has yet to be discovered, due to the overwhelming multijet background production in pp collisions. In order to suppress the background and trigger the signal events efficiently, the $H \to b\bar{b}$ searches exploit the associated production with vector bosons decaying leptonically ($Z \to \nu\nu, \ell\ell$ and $W \to \ell\nu$, with $\ell = e, \mu$).

This thesis presents a search for the Higgs boson in the $WH \to \ell\nu b\bar{b}$ channel, using proton-proton collision data at $\sqrt{s} = 8$ TeV taken with the ATLAS detector in the year 2012, corresponding to an integrated luminosity of 20.3 fb^{-1}. A particular focus is placed on the author's significant contributions made in the $WH \to \ell\nu b\bar{b}$ search: additional analysis region with muon events triggered by the missing transverse momentum. The additional analysis region was successfully implemented in the ATLAS published results and improved the sensitivity of the $WH \to \ell\nu b\bar{b}$ search by about 5%. Although the improvement was not very large, this gain was very important since the data taking had been stopped during the long shutdown while no clear $H \to b\bar{b}$ signal has been observed. An excess over the background-only hypothesis has been found with a significance of 1.8 σ while 1.5 σ was expected. The observed (expected) upper limit on the cross-section times branching ratio for $WH \to \ell\nu b\bar{b}$ at 95% confidence level was found to be 2.35 (1.37) times the SM prediction at $m_H = 125$ GeV.

The $WH \to \ell\nu b\bar{b}$ search highly depends on the performance of the so-called b-tagging: the identification of jets stemming from b-quarks' fragmentation. It is very crucial to improve the b-tagging performance for not only $H \to b\bar{b}$ searches but also any other physics analyses such as top quark property measurements since the top quark decays to weakly to b-quark ($t \to Wb$). This thesis presents development of high performance b-tagging algorithms based on boosted decision trees. At the same b-tagging efficiency, the new algorithm improves the background rejection, particularly for the c-quark originating jets and the one with the best c-jet rejection is now used as the default b-tagging algorithm in the ATLAS experiment in Run-2 analyses.

This thesis is organized as follows. The theoretical and experimental background of the studies presented in this thesis is given in chapter 1, followed by the description of the experimental setup of the ATLAS experiment at the LHC in chapter 2. The particle identification strategies in the ATLAS experiment are discussed in chapter 3. Chapters 4 and 5 present development of the high performance

b-tagging algorithms and a search for the Higgs boson in the $WH \to \ell\nu b\bar{b}$ channel, respectively. Finally, chapter 6 concludes the studies in this thesis.

1. Theoretical and experimental introduction

Humans have thought about this question since recorded history: what is matter made of? As of 2015, our knowledge has come to reach to the idea and verify that it is made by elementary particles, called quarks and leptons, which are subatomic particles. However, we do not know yet whether these particles we call "elementary" are really elementary or not.

Our best knowledge about these particles and interactions between them is summarized in a theory called the *Standard Model* (SM) of particle physics. It comprises also the mechanism that gives masses to the particles. The Higgs boson, the corresponding particle of the Higgs field has been searched for decades.

In July 2012, a new boson has been observed at a mass of about $125\ \text{GeV}$ [1, 2], by the ATLAS [3] and CMS [4] collaborations. The compatibility of the newly discovered boson with the SM prediction has been tested. All results available as of summer 2015 show a compatibility with the predicted SM Higgs boson. However, the $b\bar{b}$ decays of the Higgs boson ($H \to b\bar{b}$) have not been seen yet as clearly as the bosonic decays ($H \to \gamma\gamma, WW^*, ZZ^*$). In order to fully conclude whether the Higgs boson is the one predicted by the SM or not, direct observation using $b\bar{b}$ decay modes is a crucial task, and it should be followed by measurements using this decay modes.

This chapter briefly describes the Standard Model and the BEH mechanism in section 1.1. Predictions of the Higgs boson are explained in section 1.2 followed by a brief historical review of the searches for the Higgs boson, performed until the discovery at Large Hadron Collider (LHC) experiments in section 1.3.

1.1. The Standard Model

The Standard model of particle physics describes very precisely three interactions between all known particles: strong, electromagnetic, and weak interactions. Almost all of experimental results are consistent with the SM predictions within uncertainty [5].

An underlying principle of the SM is a gauge symmetry given by:

$$SU(3)_C \times SU(2)_L \times U(1)_Y, \tag{1.1}$$

where C, L, and Y refer to color, left-handed, and hyper-charge, respectively. The strong interaction, quantum chromo-dynamics (QCD) [6, 7], is described by the $SU(3)_C$ symmetry, and the electromagnetic and weak interactions are unified in an $SU(2)_L \times U(1)_Y$ symmetry, introduced by Glashow, Salam, and Weinberg [8–10].

1.1.1. Fundamental particles

In the SM, all particles can be categorized into two types: fermions and bosons. Fermions have spin-half, forming matter, while bosons have integer-spin.

Fermions

Fermions are divided into quarks and leptons, with three generations for both of them. Quarks are further categorized in up and down types. In the up type, there are up(u), charm(c), and top(t) quarks, and in the down type, there are down(d), strange(s), and bottom(b) quarks in the first, second and third generation, respectively. There are neutral and charged leptons. Charged leptons are electron(e), muon(μ), and tau(τ) and the neutral leptons are neutrinos, ν_e, ν_μ, and ν_τ in the first, second, and third generation, respectively. All these leptons and quarks have their antiparticle. The left-handed quarks and leptons are in weak isospin doublets:

$$\begin{aligned} L_q^{(1)} &= \begin{pmatrix} u \\ d' \end{pmatrix}_L, & L_e &= \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \\ L_q^{(2)} &= \begin{pmatrix} c \\ s' \end{pmatrix}_L, & L_\mu &= \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \\ L_q^{(3)} &= \begin{pmatrix} t \\ b' \end{pmatrix}_L, & L_\tau &= \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \end{aligned} \tag{1.2}$$

for the three generations. However, the right-handed (denoted as R) up and down type quarks and charged leptons are in weak isospin singlets:

$$\begin{aligned} R_u^{(1)} &= u_R, & R_d^{(1)} &= d_R, & R_e &= e_R, \\ R_u^{(2)} &= c_R, & R_d^{(2)} &= s_R, & R_\mu &= \mu_R, \\ R_u^{(3)} &= t_R, & R_d^{(3)} &= b_R, & R_\tau &= \tau_R. \end{aligned} \tag{1.3}$$

In the SM, neutrinos are allowed to be only left-handed, thereby, treated as massless particles. The neutrino oscillation result, first presented by the Super-Kamiokande collaboration [11] in 1998, is an evidence that the neutrinos have non-zero masses, however, very close to zero. Neutrino oscillations can be accommodated in the SM with a mixing matrix similar to the one described below for the quarks.

The mass eigenstate of the left-handed down type quarks (d', s', b') are connected with the flavor eigenstates (d, s, b) as below:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.4}$$

where V_{CKM} is the so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix [12, 13]:

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.5}$$

From measurements performed by many experiments [5], the magnitude of each coefficient looks as follows:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}. \tag{1.6}$$

For quarks, there are quantum number with three states, thereby, called color charges: r, g, and b, referring to red, green, and blue. In QCD, a system becomes stable when it is colorless (color confinement). Therefore, hadrons composed of quarks are always colorless, hence the visible matter. Baryons are colorless made of three colors: mixture of r, g, and b, e.g. a proton (u_r, u_g, d_b). Mesons are also colorless but made of a pair of a color and its anti-color, e.g. a charged pion ($u_r, \bar{d}_r$).

Table 1.1 summarizes all the fermions with their properties.

Table 1.1: Quarks and leptons in the SM for the three generations. There is an antiparticle for each particle. All the particles in the table have already been observed experimentally [5][a]. The uncertainty on the electron and muon masses are omitted.

	Charge [e]	1^{st}		2^{nd}		3^{rd}	
Quarks	$2/3$	u	$(2.3^{+0.7}_{-0.5}$ MeV$)$	c	$(1.275 \pm 0.025$ GeV$)$	t	$(173.2 \pm 0.9$ GeV$)$
	$-1/3$	d	$(4.8^{+0.5}_{-0.3}$ MeV$)$	s	$(95 \pm 5$ MeV$)$	b	$(4.18 \pm 0.03$ GeV$)$
Leptons	0	ν_e	$(<2$ eV$)$	ν_μ	$(<0.19$ MeV$)$	ν_τ	$(<18.2$ MeV$)$
	-1	e	$(0.511$ MeV$)$	μ	$(105.6$ MeV$)$	τ	$(1776.8 \pm 0.2$ MeV$)$

Bosons

The interactions between fermions are mediated by gauge bosons, which have spin-1. The electromagnetic and strong interactions are mediated by photons (γ) and gluons (g), respectively, and they are massless. Contrarily, the weak interaction is mediated by weak bosons ($W^{\pm}, Z$) which are massive: this is a consequence of electroweak symmetry breaking explained in section 1.1.2. These bosons are listed in table 1.2 with their masses and electric charges.

Table 1.2: Gauge bosons in the SM and their corresponding interactions, electric charges and masses [5]. All listed particles have spin-1 and have been observed experimentally.

Interaction	mediator	electric charge [e]	Mass [GeV]
Strong	8 gluons (g)	0	0
Electromagnetic	photon (γ)	0	0
Weak	$W^{\pm}$	± 1	80.38 ± 0.01
	Z	0	91.188 ± 0.002

1.1.2. The electroweak theory and BEH mechanism

The electroweak theory is based on the $SU(2)_L \times U(1)_Y$ Lagrangian:

$$\mathcal{L}_{SU(2)_L \times U(1)_Y} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermions}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \tag{1.7}$$

where

$$\begin{aligned} \mathcal{L}_{\text{gauge}} &= -\frac{1}{4} F^i_{\mu\nu} F^{\mu\nu i} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \\ \mathcal{L}_{\text{fermions}} &= i\bar{\psi}_j \gamma^\mu D_{\mu,j} \psi_j, \quad j = L, R. \end{aligned}$$

The $\mathcal{L}_{\text{gauge}}$ explains kinetic energies of the gauge fields, and $\mathcal{L}_{\text{fermions}}$ explains fermion kinetic energies. $F^i_\mu, i = 1, 2, 3$ and B_μ are the $SU(2)_C$ and $U(1)_Y$ gauge fields, respectively and they are

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon_{ijk} W^j_\mu W^k_\nu, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu. \end{aligned}$$

The operator, D_μ, is called covariant derivative and given by the equations below with a generator T_i([b]),

$$\begin{aligned} D_{\mu L} &= \partial_\mu - igT_i W^i_\mu + i\frac{g'}{2} Y B_\mu, \\ D_{\mu R} &= \partial_\mu + i\frac{g'}{2} Y B_\mu, \end{aligned}$$

where g and g' are the gauge coupling constants for $SU(2)_L$ and $U(1)_Y$ fields [14].

The particles in the SM are originally massless. The gauge invariance of the theory could not be maintained due to the massive weak bosons observed. In order to solve this problem, masses of the gauge bosons and fermions are introduced via the so-called Higgs field in the Brout-Englert-Higgs (BEH) mechanism [15–18].

The complex scalar Higgs doublet,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{1.8}$$

is introduced in addition to the other gauge fields, with the potential:

$$V(\phi) = \frac{\mu^2}{2}\phi^\dagger\phi + \frac{\lambda}{4}(\phi^\dagger\phi)^2. \tag{1.9}$$

Using them, the Lagrangian of this field is:

$$\mathcal{L}_{\text{Higgs}} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi), \tag{1.10}$$

and is invariant under $SU(2)_L \times U(1)_Y$ transformation.

When $\mu^2 < 0$, a corresponding single complex scalar field is illustrated in fig 1.1. One of the solutions, which minimizes the potential $V(\phi)$, can be chosen as:

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}, \qquad v = \sqrt{-\frac{\mu^2}{\lambda}}. \tag{1.11}$$

The field ϕ develops a non-zero vacuum expectation value: $v/\sqrt{2} = \mu/\lambda$. In the perturbative calculations, expansions around the ground state is involved. Without loss of generality, the field now can be rewritten as:

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{1.12}$$

---

[b] $T_i = \frac{\sigma_i}{2}$, where $\sigma_i, i = 1, 2, 3$ are the 2×2 Pauli matrices. The commutation relations for the group are given by: $[T_i, T_j] = i\epsilon_{ijk}T_k, \quad [T_i, Y] = 0.$

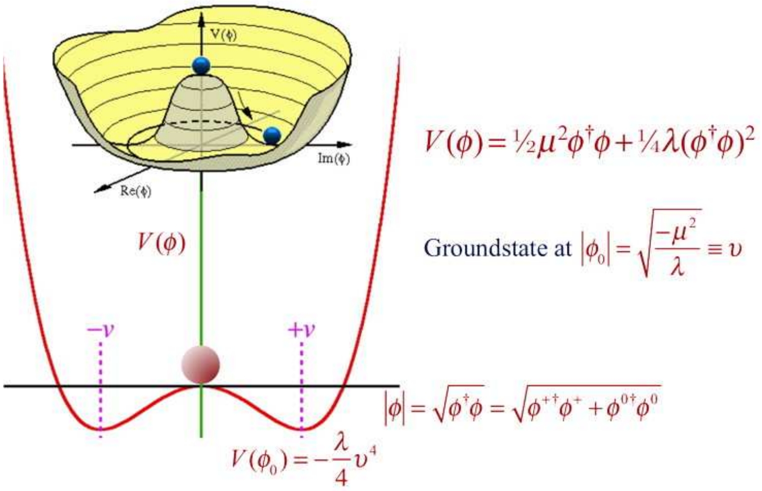


Fig. 1.1: The Higgs potential $V(\phi)$ for a single complex scalar field.

where $h(x)$ is the so-called Higgs field, and the Lagrangian of eq 1.10 now becomes:

$$\mathcal{L}_{\text{Higgs}} = (\partial_\mu h)^2 + \frac{1}{4}g^2 W_\mu W^\mu (v+h)^2 + \frac{1}{8}\left(\sqrt{g^2+g'^2}\right)^2 Z_\mu Z^\mu (v+h)^2 - V(\frac{1}{2}(v+h)^2), \tag{1.13}$$

where second and third terms correspond to W and Z boson, and the W and Z boson masses can be written as:

$$m_W = \frac{1}{2}gv, \qquad m_Z = \frac{1}{2}v\sqrt{g^2+g'^2}, \tag{1.14}$$

and they are 80.38 ± 0.01 GeV and 91.188 ± 0.002 GeV from the measurements, respectively. The Lagrangian above further contains the term of a scalar boson with a mass of:

$$m_h = \sqrt{-2\mu^2} = \sqrt{2\lambda}v, \tag{1.15}$$

which interacts with the massive electroweak bosons and with itself. Since μ is a free parameter the Higgs boson mass cannot be predicted theoretically. This boson is referred to as Higgs boson, and a compatible boson has been observed in 2012 at the LHC.

There is also a relation between m_W and m_Z:

$$\frac{m_W}{m_Z} = \frac{g}{\sqrt{g^2+g'^2}} = \cos\theta_W \tag{1.16}$$

where θ_W is the Weinberg angle which describes the mixing of W^3 and B_μ fields to form the neutral bosons: the massive Z and the massless photon.

The last term in eq 1.7 can be written as:

$$-\mathcal{L}_{\text{Yukawa}} = \mathbf{y_u}\bar{L}_{q_i}\phi^\dagger u_R + \mathbf{y_d}\bar{L}_{q_i}\phi^\dagger d_R + \mathbf{y_e}\bar{L}_i\phi^\dagger e_R, \tag{1.17}$$

where $\mathbf{y_u}, \mathbf{y_d}$, and $\mathbf{y_e}$ are the so-called Yukawa-coupling constants and given as 3×3 matrices due to the three generations of up and down types of quarks and charged leptons, respectively.

1.2. Predictions on the Standard Model Higgs boson

This section describes Higgs boson production mode and their production cross-section in section 1.2.1, and the decay modes of the Higgs boson in section 1.2.2.

1.2.1. Production modes of the Higgs boson

There are mainly four types of production processes of the Higgs boson.

- Gluon fusion process
  Gluon fusion ($gg \to H$ or simply noted ggH) process through a heavy quark loop is the leading production mechanism of the SM Higgs boson at both the LHC and Tevatron. The lowest order (LO) Feynman diagram of gluon fusion process is shown in fig 1.2.

- Vector boson fusion process
  The electroweak gauge bosons are emitted from quarks and fusion into the Higgs boson: this production of the Higgs boson associated with two hard jets in the forward and backward regions of the detector is called vector boson fusion (VBF) process, and its LO diagram is shown in fig 1.2. Bounds on non-standard couplings between Higgs and electroweak gauge bosons can be imposed from precision studies of this channel.

- Higgs-strahlung process
  Production of the Higgs boson associated with a vector boson (W/Z) production is called associated production or Higgs-strahlung process and its LO diagram is shown in fig 1.3. For the $H \to b\bar{b}$ search, in which the final state of the Higgs boson is composed of hadronic jets, this process provides the most sensitive final state signatures since the leptonic decays of associated W or Z bosons can be exploited for triggering and background rejection.

- ttH production
  The Higgs production mode in association with $t\bar{t}$ production is called ttH. In this production mode, the Higgs boson is produced by fermion pair without any loop process. Hence, it is very important to understand Yukawa coupling.

  There is also a similar process called bbH, replacing top quarks with b-quarks. However, the production cross-section of the b-quark coupling to the Higgs boson is weaker than that of top quark.

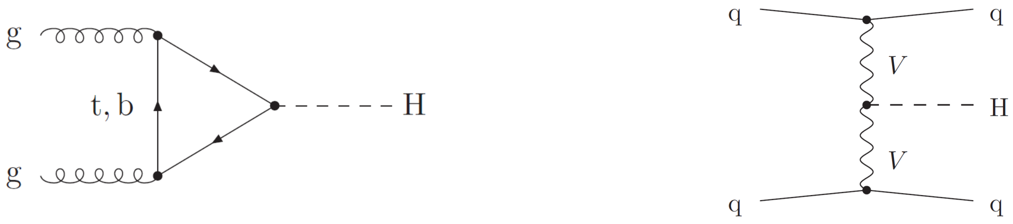


Fig. 1.2: Feynman diagram contributing to $gg \to H$ at lowest order (left). Feynman diagram for $qq \to qqH$ at LO (right), where V denotes the vector bosons (W or Z).

Once the Higgs boson mass is known, the production cross-section of these modes described above can be predicted by the SM at a given center-of-mass energy and given colliding particles. Figure 1.5

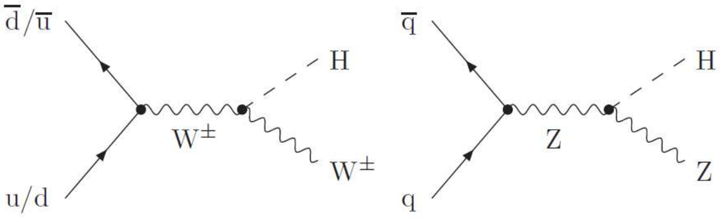


Fig. 1.3: The LO diagrams for $qq \to WH$ (left) and $qq \to ZH$ (right). For the W-boslon associated production, the W-boson charge is determined by the two quarks involved in the interaction.

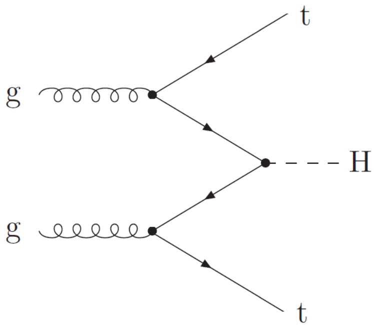


Fig. 1.4: The LO diagrams for $gg \to t\bar{t}H$.

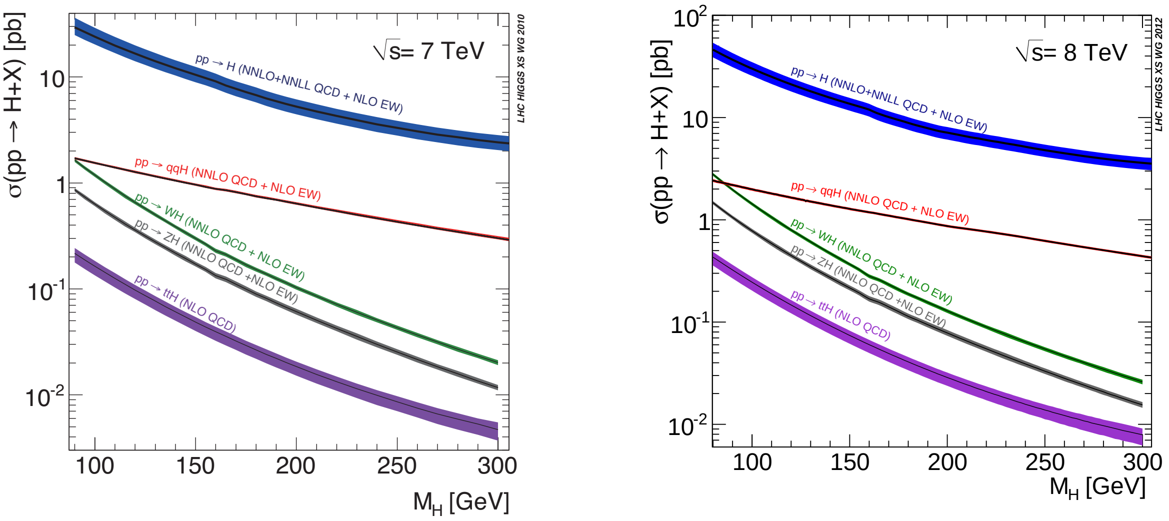


Fig. 1.5: The production cross-section of the SM Higgs boson as a function of its mass for pp collisions at the center-of-mass energy of 7 TeV (left) and 8 TeV (right) [19].

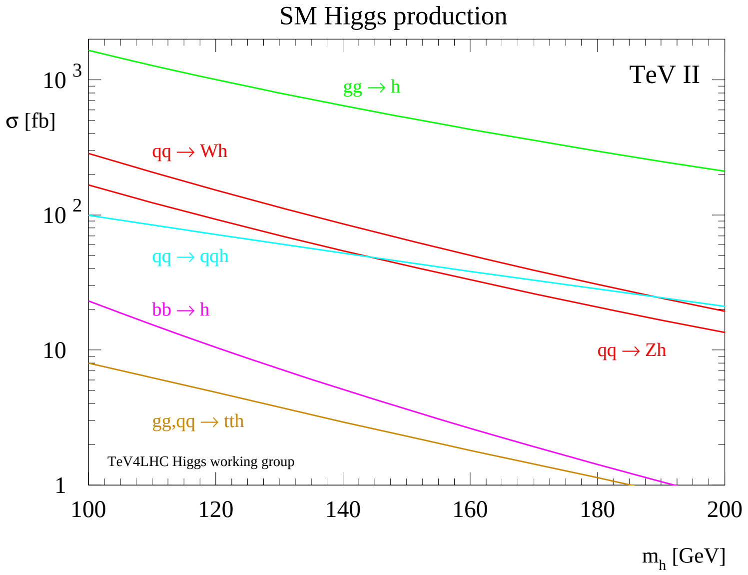


Fig. 1.6: The production cross-section of the SM Higgs boson as a function of its mass for $p\bar{p}$ collisions at the center-of-mass energy of 1.96 TeV [20].

shows the production cross-section of the Higgs boson as a function of its mass under the LHC Run 1 conditions. The same for Tevatron Run II condition is shown in fig 1.6.

Under both Tevatron and LHC conditions, the dominant production process is the ggH process. The second leading process is VBF at the LHC, while the (W/Z)-associated production has larger cross-section than VBF in the low mass region ($m_H \lesssim 145$ GeV) at Tevatron. The reason is that the antiquark is needed to generate the W/Z bosons. In the pp collisions at the LHC, antiquarks arise from sea-quarks, however, the proton antiquark PDFs are much smaller than those of antiprotons due to the valence antiquarks in the antiprotons.

1.2.2. Higgs boson decays

Figure 1.7 shows branching fraction of the Higgs boson decays as a function of m_H. In the mass region above 130 GeV, the $H \to WW^*$ channel dominates the Higgs boson decay, while in the mass region below 130 GeV, $H \to b\bar{b}$ dominates the Higgs boson decay.

Figure 1.8 shows the Higgs boson production cross-section times branching fraction as a function of m_H. In spite of the largest branching fraction of $H \to b\bar{b}$ ($\sim 58\%$) at m_H 125 GeV, $H \to b\bar{b}$ is not indicated in the plot since an inclusive search for the $H \to b\bar{b}$ decay is not feasible due to the overwhelming multijet background produced with a huge rate ($> 10^5$ pb) at the LHC. Hence, the associated production with a vector boson decaying leptonically ($Z \to \nu\nu, \ell\ell$ and $W \to \ell\nu$, with $\ell = e, \mu$) is exploited to search for $H \to b\bar{b}$. In the search presented in chapter 5, these leptonic decays of vector bosons are used for triggering and background reduction, thanks to the low fake rate for the high p_{T} electrons or muons, or large missing transverse momentum.

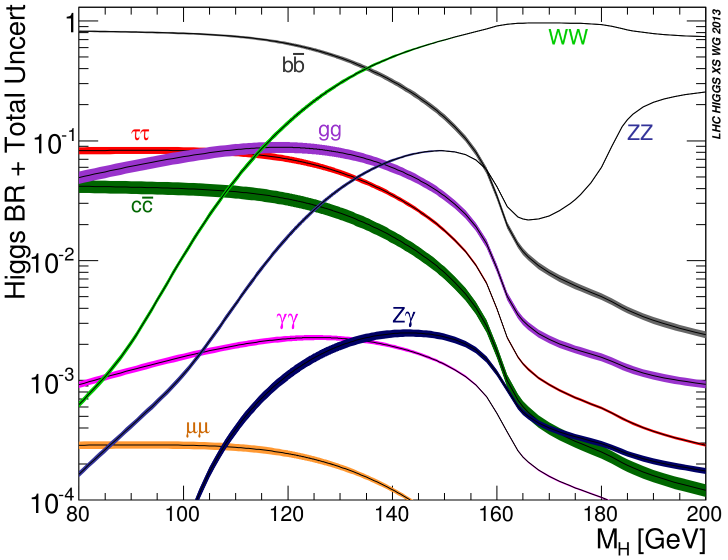


Fig. 1.7: The SM Higgs boson branching fraction as a function of m_H [19].

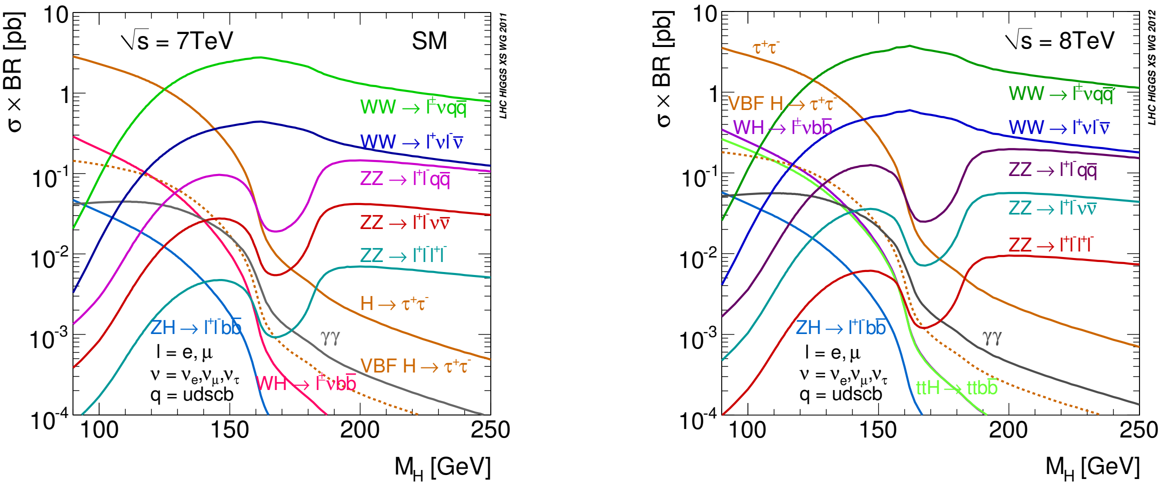


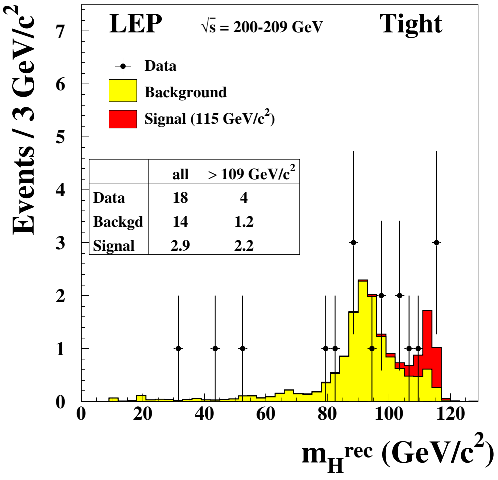
Fig. 1.8: The production cross-section of the SM Higgs boson times branching fraction as a function of m_H at the center-of-mass energy of 7 TeV (left) and 8 TeV (right). For $H \to \tau^+\tau^-$ both inclusive (solid line) and VBF (dashed line) mode are plotted [19].

1.3. Search for the Standard Model Higgs boson

To search for and study properties of the Higgs boson is very important, since it remained undiscovered for more than 10 years as the last missing building block of the SM after the discovery of the top quark and tau-neutrino. Since its prediction proposed about half a century ago [15–18], many searches have been performed over decades in many experiments. This section reviews the history of the Higgs boson searches performed in LEP, Tevatron, and LHC experiments.

1.3.1. Search at LEP and Tevatron

Searches for the Higgs boson were performed in the LEP experiments using e^+e^- collision data at center-of-mass energies of 200–209 GeV [21]. In the e^+e^- collisions, the production mode of the Higgs boson is dominated by the $e^+e^- \to ZH$ process. Due to the small cross-section for the high mass region at these center-of-mass energies, the search targeted relatively low mass region ($m_H <$ 120 GeV). In this low mass range, $H \to b\bar{b}$ firstly dominates the decay mode (75% at $m_H = 115$ GeV) and the Higgs decays to $\tau\tau$, WW^*, gg, and $c\bar{c}$ follow next. The searches were performed using mainly the decay modes (final state) of $ZH \to q\bar{q}b\bar{b}$ (four-jet), $ZH \to \nu\nu b\bar{b}$ (two-jet and missing transverse energy), $ZH \to \ell\ell b\bar{b}$ (two-jet and two-lepton with same flavor), and $ZH \to \tau^+\tau^- b\bar{b}$ or $qq\tau^+\tau^-$ (two-jet and two-tau), where ℓ denotes an electron or muon. Figure 1.9 shows the reconstructed dijet mass of the Higgs boson candidates in the combination of all-four LEP experiments (ALEPH, Delphi, L3, and OPAL). Since the statistical significance of the signal excess over the background-only hypothesis was not strong (less than 2.5σ), a limit on the mass of the SM Higgs boson was set, shown in fig 1.10. As of March 2003, the Higgs boson lower mass limit was set at around 114 GeV at 95% CL.



	all	> 109 GeV/c²
Data	18	4
Backgd	14	1.2
Signal	2.9	2.2

Fig. 1.9: The reconstructed dijet mass distribution of Higgs boson candidates at the LEP experiments. The yellow (red) distribution shows the expected background (Higgs boson signal) contribution along the data points. The distribution shows the combined results from all analysis channels and the four experiments [21].

Using the $p\bar{p}$ collision data at a center-of-mass energy of 1.96 TeV, the searches in a mass range which was broader than LEP experiments, were performed at Tevatron experiments (CDF and D0). The searches were performed using mainly the VH associated production for $H \to b\bar{b}$ and WW decay

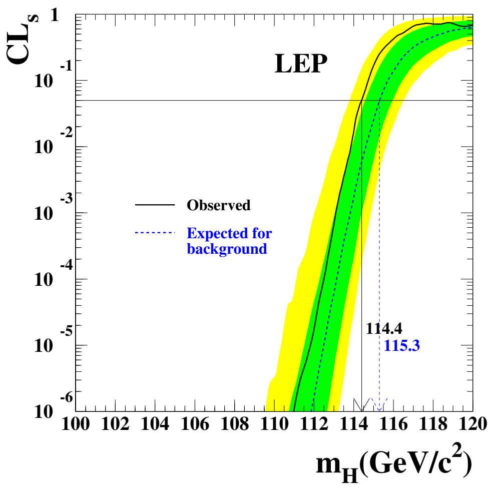


Fig. 1.10: The ratio, $CL_S = CL_{s+b}/CL_b$, for the Higgs boson signal plus background hypothesis as a function of m_H from the LEP results. The observed (solid line) and expected background-only hypothesis (dashed line) are obtained by combination of all analysis channels in the four experiments [21].

channels, and the gluon fusion production for $H \to \gamma\gamma$, WW^*, and $\tau\tau$ decay channels. Figure 1.11 shows the combined results from the two experiments and all the analysis channels for 95% CL upper limit on the SM Higgs boson signal strength as a function of m_H. As of July 2010, the SM Higgs boson was excluded also in the mass range 158–175 GeV, which was the situation before the LHC operation at a center-of-mass energy of 7 TeV got on the rails.

1.3.2. Discovery at LHC

The LHC commissioning for high luminosity operation at a center-of-mass energy of 7 TeV started to be achieved in Autumn, 2010. The searches have been performed in the main decay channels ($H \to b\bar{b}, \tau^+\tau^-, WW^*, ZZ^*$, and $\gamma\gamma$). In mainly the bosonic decay channels ($H \to \gamma\gamma$, $H \to WW^*$, and $H \to ZZ^*$) where the background is less compared to the fermionic decay channels ($H \to b\bar{b}, \tau^+\tau^-$), an excess was observed at a mass of about 125 GeV. In July 2012, ATLAS and CMS collaborations announced the discovery of a new particle which could be the SM Higgs boson. Figure 1.12 shows the reconstructed invariant mass distributions of the Higgs boson candidates in the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels. Figure 1.13 shows the ATLAS combined result of searches in the bosonic decay channels using data taken by early 2012 at center-of-mass energies of 7 and 8 TeV, corresponding to integrated luminosities of about 5 and 6 fb^{-1}, respectively. In the region, $m_H < 560$ GeV, the SM Higgs boson is excluded except the mass range around 125 GeV where the excess was observed. The local p_0 value as a function of m_H along with the corresponding significance is shown in fig 1.14, and the statistical significance of the excess over the background-only hypothesis was more than 5σ at $m_H \sim 125$ GeV, which led to the announcement of the discovery.

After the discovery, the excess became more significant as data increased. Using full LHC Run 1 data, the property measurement of the Higgs boson was performed in the bosonic decay channels. Figure 1.15 shows the results of the hypothesis test obtained in the $H \to ZZ^*$ channel, using $J^P = 0^+, 0^-$ spin-parity hypotheses, where J and P stand for spin and parity, respectively. By combining

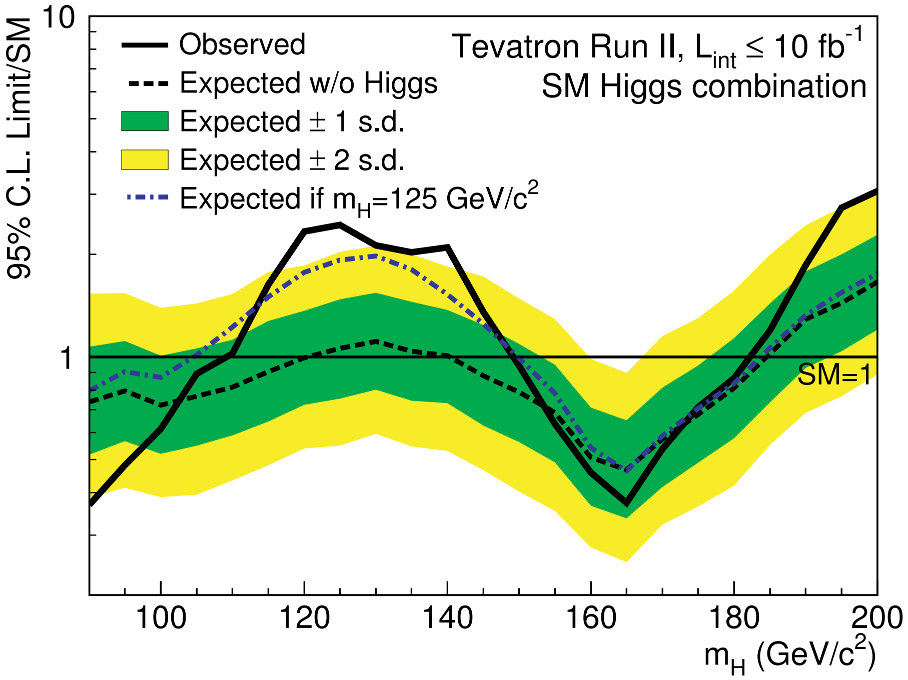


Fig. 1.11: The observed (solid line) and expected (dashed line) 95% CL upper limit on the SM Higgs boson signal strength as a function of m_H, obtained by combination of searches performed in the CDF and D0 expriments using $p\bar{p}$ collisions at the center-of-mass energy of 1.96 TeV [22].

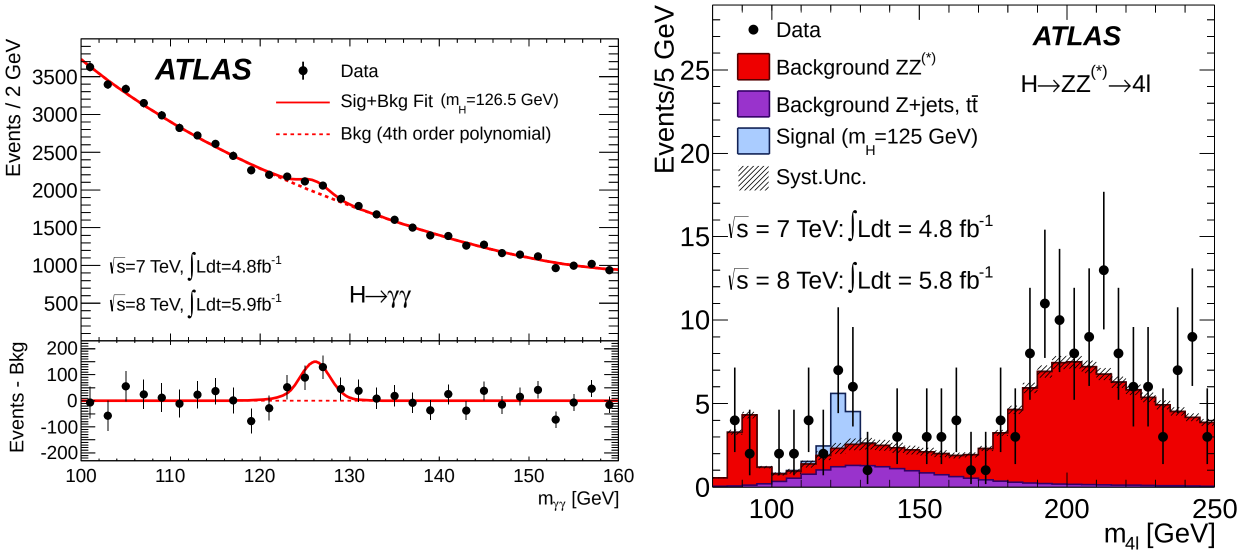


Fig. 1.12: Reconstructed mass distributions of the Higgs boson candidates in the $H \to \gamma\gamma$ (left) and $H \to ZZ^* \to 4\ell$ (right) channels. The background-subtracted distribution is shown in the lower panel of the left plot. The Higgs boson signal excess for $m_H \sim 125$ GeV can be clearly seen for both channels [1].

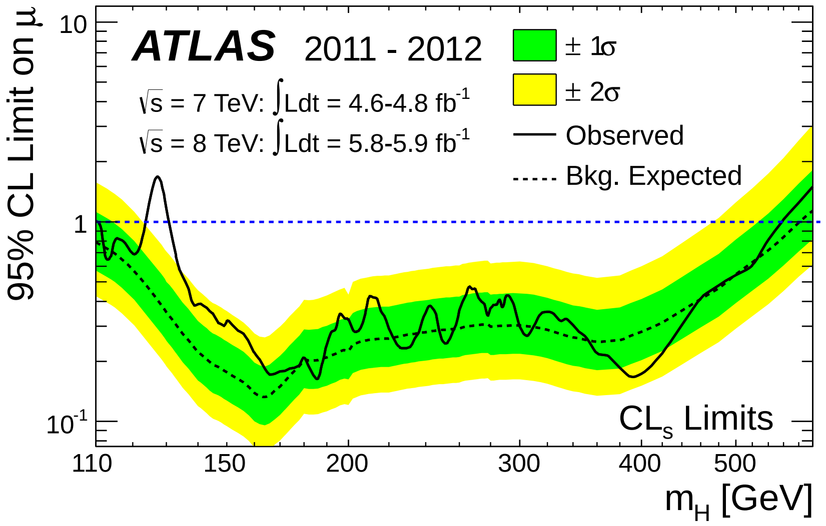


Fig. 1.13: The observed (solid line) and expected (dashed line) 95% CL upper limit on the SM Higgs boson signal strength as a function of m_H, obtained by combination of searches in bosonic decay channels ($H \to \gamma\gamma$, WW^*, and ZZ^*) [1].

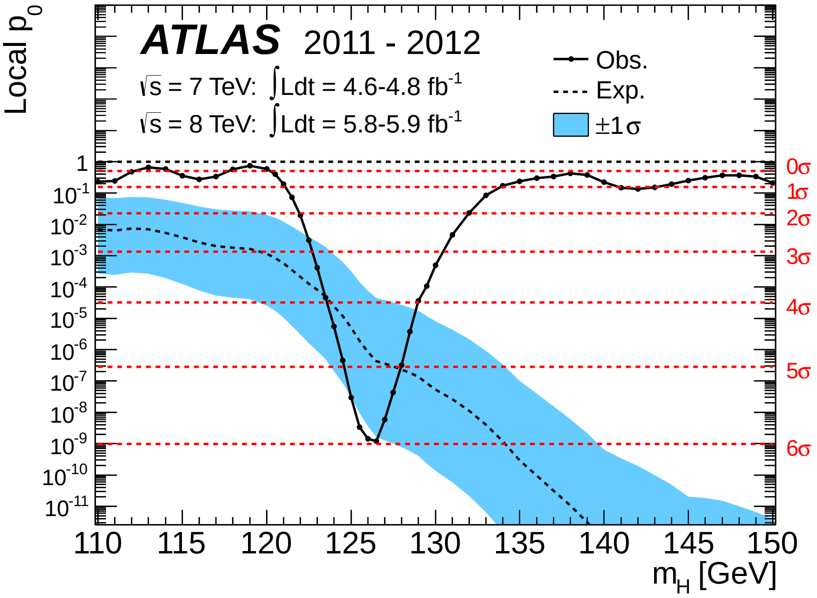


Fig. 1.14: The observed local p_0 value (solid line) as a function of m_H in the low mass range. The expected values (dashed line) with the SM Higgs boson hypothesis at the corresponding mass is also shown with the blue band for $\pm1\sigma$ [1].

the reults of the $H \to \gamma\gamma$, WW^*, and ZZ^* channels, the compatibility of the Higgs boson spin with the SM hypothesis is found to be preferred by data, while excluding the other spin hypotheses [23].

1.3.3. Search for $VH \to Vbb$ at Tevatron

Soon after the discovery of the Higgs boson candidate at around $m_H \sim 125$ GeV, the Tevatron experiments also published their full data search for the $VH \to Vb\bar{b}$ channel, related to the fact that the branching fraction of the Higgs boson is dominated by $H \to b\bar{b}$ at the corresponding mass. Figure 1.16 shows the dijet mass distribution after background-subtraction using data collected at CDF and D0 experiments. A broad excess over the background-only hypothesis was observed, with a 2.8σ significance at $m_H = 125$ GeV. The exclusion limit set for the signal strength as a function of m_H is shown in fig 1.17.

1.4. Summary

The Standard Model is a very successful theory which explains the known particles and the interactions in between them. The last building block of the SM particles, the Higgs boson, was directly observed by ATLAS and CMS experiments at the LHC in July 2012, followed by the confirmation of the SM Higgs boson properties, such as spin and CP. However, the discovery and the property measurements were performed using the bosonic decay channels and the fermionic decay modes were not yet confirmed as fully as the bosonic decay modes. To observe the fermionic decay modes of the Higgs boson and perform property measurements are very important pieces for the SM confirmation and also the exclusion or observation of the new physics contamination to the Higgs boson properties. The $H \to b\bar{b}$ decay mode has the largest branching fraction at $m_H = 125$ GeV (∼58%) and has yet to be observed.

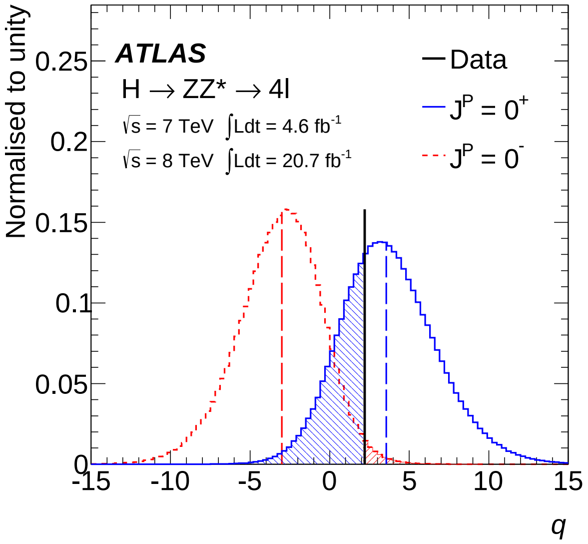


Fig. 1.15: The expected distributions of log-likelihood ratio between the two hypotheses, $q = \log\Big(\mathcal{L}(J^P = 0^+)/\mathcal{L}(J^P = 0^-)\Big)$. The SM hypothesis 0^+ (blue) distribution is compared with that of the alternative 0^- (red) hypothesis. The observed value is indicated by the vertical solid line and the expected medians by dashed lines. The colored areas correspond to the integrals of the expected distributions up to the observed value and are used to computed the p_0 values for the rejection of each hypothesis [23].

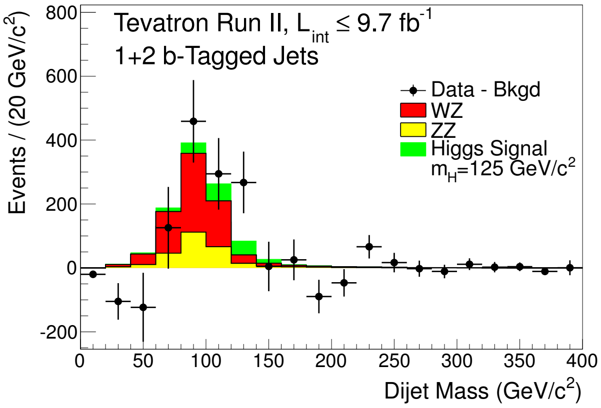


Fig. 1.16: Background-subtracted distribution of the reconstructed m_{jj}, summed over all input channels. The signal and background contributions are fit to the data, and the fitted background is subtracted except the diboson (WZ and ZZ) background components. The fitted VZ background and VH ($m_H = 125$ GeV) signal contributions are shown with filled histograms [24].

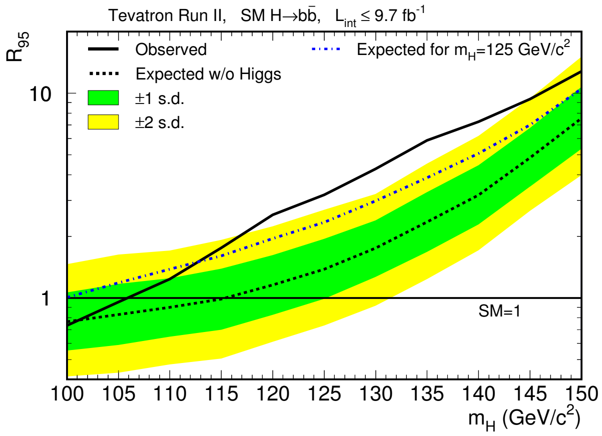


Fig. 1.17: The observed (solid line) and expected (dashed line) 95% CL upper limit on the SM Higgs boson signal strength as a function of m_H, obtained by combination of searches in $VH \to Vbb, V = W, Z$. The dot-dashed line shows the median expected limit if the SM Higgs boson is present at $m_H = 125$ GeV [24].

2. The ATLAS experiment at the Large Hadron Collider

This chapter describes a series of accelerators including the Large Hadron Collider and the ATLAS detector built at the European Organization for Nuclear Research (CERN). Additionally, the software tools and computing facilities used for the analyses and simulation of collision data are discussed.

2.1. Large Hadron Collider

The Large Hadron Collider (LHC) [25] is the world's largest synchrotron, built underground at about 100 m depth near Geneva, Switzerland. Its diameter is about 9 km, lying across the border between France and Switzerland.

2.1.1. Towards the LHC

The LHC accelerator is built in a circular tunnel originally constructed for the Large Electron Positron (LEP) collider. Although very clean e^+e^- collision events were provided by LEP, the synchrotron radiation loss was significant in case of the energy increase considered, since the electrons and positrons are very light.

One possible way to avoid this synchrotron radiation loss for the e^+e^- collider experiment is to use a linear collider. However, this linear structure does not allow to reuse the bunches after collision so that the project becomes technically difficult. For these linear collider projects, the International Linear Collider (ILC) [26] and Compact Linear Collider (CLIC) [27] projects are being planned, but none of them has been approved for construction as of 2015.

In order to avoid the dis-advantage of a linear collider approach, the proton-proton circular collider approach was chosen for the higher energy collider at CERN. Due to the heavy mass of the proton, which is about 1836 times heavier than an electron, the synchrotron radiation loss is negligible in the proton collider, since the magnitude of the synchrotron radiation loss is proportional to m^{-4}, where m is the mass of the accelerated particle. At the energy frontier, the Tevatron accelerator was running proton-antiproton collisions at Fermilab, USA. Although the proton-antiproton collisions provided very encouraging results like top-quark discovery, the difficulty was to acquire a very high luminosity for the antiproton beam. However, the proton-proton collisions can also provide similar physics results. Taking advantage of the higher luminosity obtained by proton-proton beams, it was decided to select proton-proton collisions to be performed at CERN.

Figure 2.1 shows the comparisons of the production cross-sections of the Standard Model (SM) processes for the different center-of-mass energies and colliding particles ($pp/p\bar{p}$ collisions).

The main purpose of the LHC is to study physics phenomena at very high energy (nominally $\sqrt{s} = 14$ TeV), starting from the SM physics processes to the Higgs boson production, and new particles or signatures from so-called beyond-the-standard-model physics. At the four collision points of the LHC, the ATLAS [3], ALICE [29], CMS [4], and LHCb [30] experiments are placed to study the head-on collisions of the protons or heavy ions. The ATLAS and CMS detectors are multi-purpose detectors while the ALICE and LHCb detectors are made for specific purposes which are heavy ion collisions and B-physics, respectively.

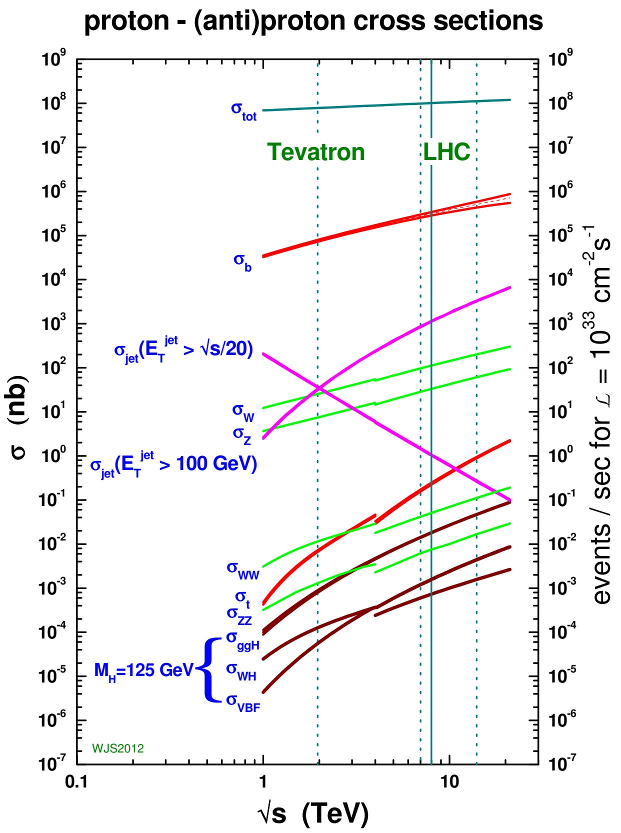


Fig. 2.1: Cross-sections for a few typical processes at Tevatron (proton-antiproton collisions) and LHC colliders (proton-proton collisions) shown as a function of a center-of-mass energy [28].

2.1.2. The accelerator chain

Figure 2.2 shows the accelerator complex at CERN. First, protons are extracted by ionizing hydrogen gas, then injected into a linear accelerator (LINAC2). The Booster raises up the beam energy to 1.8 GeV, and the Proton Synchrotron (PS) brings up the beam energy to 26 GeV. Proton beams are then injected in the LHC after being accelerated to 450 GeV by the Super Proton Synchrotron (SPS). Finally in the LHC main ring, the beam is accelerated to the colliding beam energy, nominally 7 TeV.

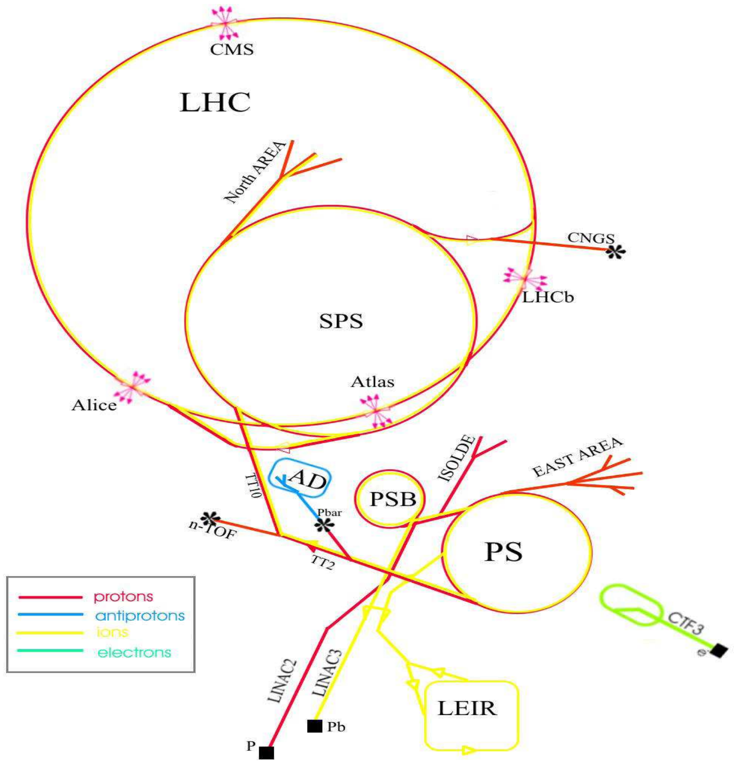


Fig. 2.2: The CERN accelerator complex.

The synchrotron is made of roughly three components: the accelerating cavities, the bending magnets, and the focusing magnets. The LHC is composed of 1232 dipole magnets, which change the beam direction. These dipole magnets are also equipped with sextupole, octupole and decapole magnets in order to correct for small imperfections in the magnetic field at the extremities. The quadrupole magnets are used to squeeze the beams for the collisions. By design, the magnetic field of the dipole magnets is 8.3 T corresponding to the beam energy of 7 TeV per proton. These magnets are superconducting and kept at 1.9 K with super-fluid helium. In order to accelerate two proton beams circulating in opposite directions, the LHC has two beam pipes in one magnet, generating two opposite magnetic fields around each beam pipe. Figure 2.3 shows the structure of the LHC dipole magnet and two beam pipes lying in the cryostat, and the main parameters of the LHC are shown in table 2.1.

2.1.3. Luminosity

The number of events per second, generated at a collider for a specific physics process, is given by $N_{\text{events}} = \mathcal{L} \cdot \sigma$, where $\mathcal{L}$ is the instantaneous luminosity of the accelerator, and σ is the production cross-section of any physics process of interest. The instantaneous luminosity is determined by accel-

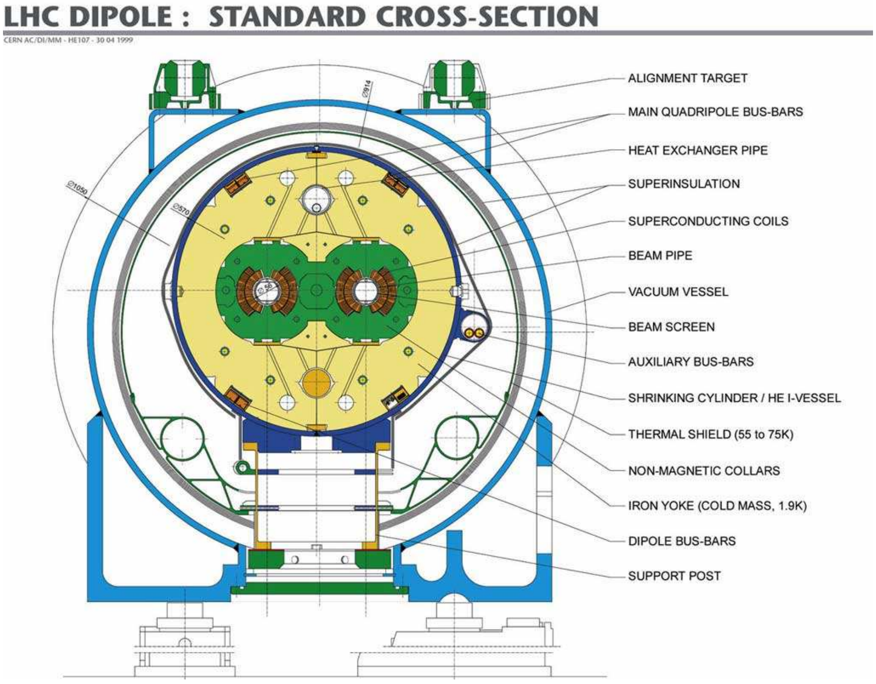


Fig. 2.3: A cross-sectional view of the LHC accelerator.

Table 2.1: The main parameters of the LHC. Design values are compared to the actual values in each operation year.

	Design	2010	2011	2012
Beam energy [TeV]	7.0	3.5	3.5	4.0
Number of bunches	2808	368	1380	1380
Bunch spacing [ns]	25	150	50	50
Protons per bunch [10^{11}]	1.15	1.2	1.5	1.6–1.7
Integrated luminosity [fb^{-1}]	100	0.045	5.1	21.3
Peak luminosity [$10^{32}\text{cm}^{-2}\text{s}^{-1}$]	100	2.1	37	77
Collisions per bunch crossing (Mean μ)	19	3	9	21

erator characteristics:

$$\mathcal{L} = \frac{N_p^2 k_b f_{\text{rev}}}{4\pi\sigma_x\sigma_y} F \tag{2.1}$$

where N_p is the number of protons per bunch, k_b is the number of bunches per beam, f_{rev} is the revolution frequency of the bunches, and σ_x, σ_y are the widths of the Gaussian distribution of the beam in the transverse plane[a]. The factor F accounts for the impact of crossing angle and other characteristics of the beam configuration.

The integrated luminosity over a certain period of time, L_{int}, is then defined as $L_{\text{int}} = \int \mathcal{L} dt$. The total number of events in a dataset of integrated luminosity L_{int} is given by $N_{\text{events}}^{\text{Tot.}} = L_{\text{int}} \cdot \sigma$.

Pileup interactions

Due to the high luminosity provided by the LHC, there are multiple number of interactions per one bunch crossing: these interactions are the so-called in-time-pileup interactions[b]. Such number of interactions per bunch crossing follows a Poison distribution with mean value, μ, with fluctuation from bunches to bunches. The mean number of interactions averaged over all bunch crossings and averaged over the data analyzed is referred to as $<\mu>$. In data, μ is calculated by:

$$\mu = \frac{\mathcal{L} \cdot \sigma_{\text{inel}}}{n_{\text{bunch}} \cdot f_{\text{rev}}}, \tag{2.2}$$

where σ_{inel} is the total inelastic cross-section and n_{bunch} is number of colliding bunches. The uncertainty on μ depends on the luminosity uncertainty and the total inelastic cross-section.

The μ values in each operation year is also indicated in the last row in table 2.1. Although it is a consequence of high luminosity, the large number of pileup interactions in one event can affect many measurements of many physics objects, especially the ones measured by calorimetry: photons, electrons, tau-jets, jets, and missing transverse momentum. The tracking is also affected by the pileups due to the high occupancy in the detector, and primary vertex reconstruction is affected by many interactions occurring close to each other.

2.1.4. The LHC operation

The circulation of the proton beams in the LHC started in September 2008. However soon after the first circulation, an accident happened: a connection between superconducting magnets melted. More than one year was spent to study and repair the problem and put in place a new protection system, and the operation restarted in November 2009. The beam energy was raised gradually from the injection energy 450 GeV, while overtaking the highest energy record established by the Tevatron accelerator. Finally the beam energy at which the LHC was operated in the year 2011 (2012) became 3.5 (4.0) TeV. Figure 2.4 shows the integrated luminosity that was recorded by the ATLAS experiment in 2011 and 2012. Until the end of 2011 (2012), about 5 (21) fb^{-1} of data were recorded by the ATLAS experiment at $\sqrt{s} = 7$ (8) TeV.

The data taking has been performed in several levels of units. Starting from smaller units, they are:

- Event: it is corresponding to one hard scattering and the particles coming out from the collision. However, the events contain several pp interaction points due to the pileup interactions.

---

[a] A plane perpendicular to the beam axis.

[b] There is also another one, called out-of-time pileup. They are left-over effects from the previous bunch crossing in the calorimeter, since its readout integration time is longer than bunch crossing rate (nominally 40 MHz).

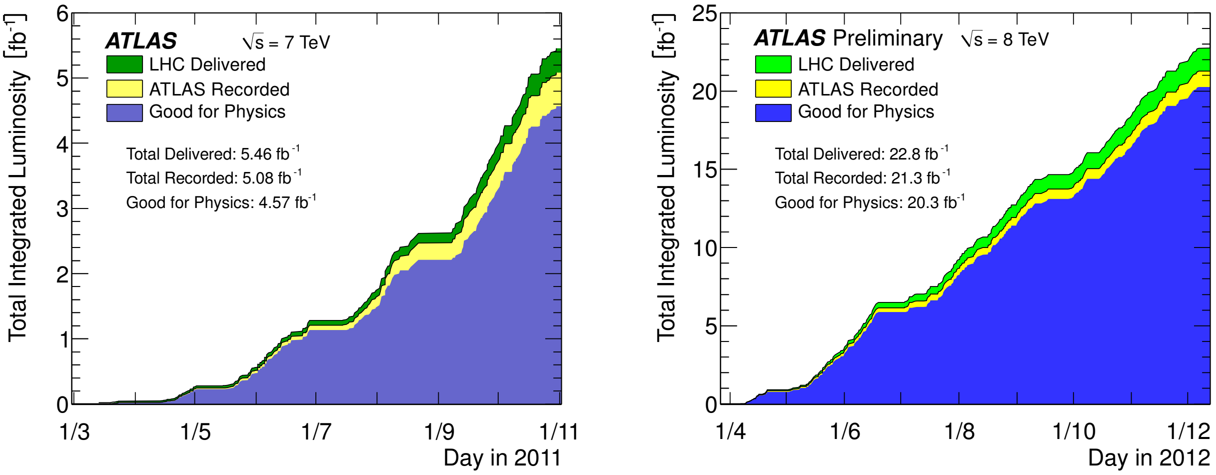


Fig. 2.4: The integrated luminosity recorded by the ATLAS experiment as a function of time in the year 2011 (left) and 2012 (right) [31]. The difference between LHC Delivered and ATLAS Recorded reflects the time difference from the start of stable beams until the LHC requests ATLAS to put the detector in a safe standby mode to allow beam dump or beam studies. Another difference between the ATLAS Recorded and Good for Physics reflects the DAQ inefficiency and so-called “warm start” inefficiency. The warm start occurs due to the undergoing preparation of silicon detector high-voltage while the stable beam flag is raised.

- Luminosity-block: a luminosity-block contains events and its time span is typically 1–2 minutes. In order to use events taken in the ideal condition in which all the sub-detectors are fully working, luminosity-blocks with some sub-detectors in bad conditions are removed using a list, called good run list (GRL), from physics analyses. The GRL stores all the sub-detectors’ conditions.
- Run and fill: a run, containing many luminosity-blocks inside, is typically about 10 hour long. A fill, containing from one to several runs, starts with the new beam injections into the LHC and ends with the beam dumps due to their lifetime.
- Period: a Period is composed of up to a few tens of runs and starting from the end of short shutdowns which were made for the maintenance of the accelerator and detectors, and detector calibration. It is typically spanning several weeks.
- Year: it is the calendar year, and a year has several data taking periods.

2.2. The ATLAS detector

This section describes the ATLAS detector located at CERN Point-1, which is indicated as Atlas in fig 2.2. The collaboration consists of more than 3000 scientists from over 35 countries. The detector is placed in an underground cavern, and the main control room is located on the surface, almost right above the detector. The overview of the ATLAS detector is shown in fig 2.5. Similarly to the other collider experiments, the ATLAS detector is composed of four main systems which are tracking, calorimeter, muon and the magnet systems. The tracking and muon systems in ATLAS are called Inner Detector (ID) and Muon Spectrometer (MS), respectively.

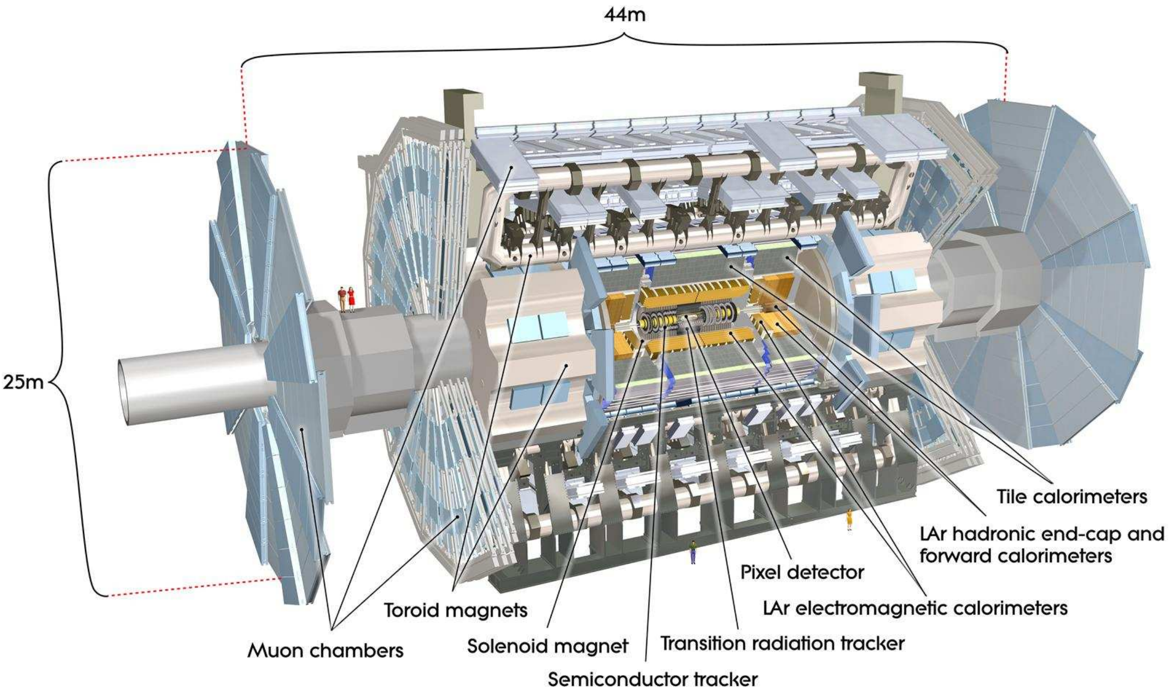


Fig. 2.5: Overview of the ATLAS detector. It is composed of three main sub-detectors which are the Inner Detector, the calorimeter, and the Muon Spectrometer.

The experimental coordinates and setup are summarized in section 2.2.1, followed by the description of the sub-detector components. The solenoid and toroidal magnet systems work together with the Inner Detector and the Muon Spectrometer, respectively. Therefore, these magnet systems are also described in the corresponding sections.

2.2.1. Experimental coordinates and quantities

The right handed-system is used for the coordinates of the ATLAS experiment. The $+x$ direction points at the center of the LHC ring, the $+y$ direction points at the sky vertically. The beams from the LHC are along the z-axis, and the $+z$ direction corresponds to the counterclockwise rotation when the LHC is seen from the sky.

The cylindrical coordinate system (r, ϕ, θ) is also used at the same time, r for the radius from the beam axis (z-axis) on the x-y plane, ϕ for the azimuthal angle from x-axis ($-\pi < \phi \leq +\pi$) in the transverse plane, and θ for the polar angle from z-axis.

Rapidity y, in which the distance Δy is Lorentz-invariant, is defined as the equation below:

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

where E is the energy of the particle and p_z is the z component of the particle momentum. In case the rest-mass of the particle is negligible, y in eq 2.3 becomes the pseudo-rapidity:

$$\eta = -\frac{1}{2} \ln \left(\tan \frac{\theta}{2} \right). \tag{2.4}$$

and is frequently used instead of θ. The distance between two particles, ΔR, in the pseudo-rapidity-azimuth space is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \tag{2.5}$$

where $\Delta\eta$ and $\Delta\phi$ are the angular separation between the two particles in the corresponding angular variable.

The particle transverse momentum, p_T, is defined as:

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}. \tag{2.6}$$

Since ATLAS is a hadron collider experiment, the longitudinal energy of the initial colliding partons cannot be measured. However, the vectorial sum of particle momenta on the transverse plane should be conserved. This provides the missing transverse momentum, which can be a hint of neutrinos' momenta. The missing transverse momentum magnitude, $E_{\mathrm{T}}^{\mathrm{miss}}$, is defined as:

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2}. \tag{2.7}$$

Details about $E_{\mathrm{T}}^{\mathrm{miss}}$ are discussed in section 3.6.

2.2.2. Inner detector

This system is located at the innermost part of the detector, which is inside the superconducting solenoid. The role of the Inner Detector is to reconstruct the tracks of the charged particles using the hit information in the sub-detectors. The magnetic field generated by the solenoid bends the charged particle tracks, and their curvatures are used to reconstruct their momenta. Using these tracks fitted to the Inner Detector hits, the interaction points and the decay position of the short-lived particles are reconstructed. This excellent tracking and vertexing resolution provided by the Inner Detector is a crucial ingredient for b-tagging presented in chapter 4 and also for analyses based on b-tagging such as the $H \to b\bar{b}$ search presented in chapter 5. The track p_{T} reconstruction threshold is nominally 0.5 GeV but different thresholds are used depending on the physics objects. The detailed method of tracking or vertexing is discussed in section 3.1. The Inner Detector covers $-\pi < \phi \leq +\pi$ on the $r - \phi$ plane and the pseudo-rapidity range $|\eta| < 2.5$ on the $y - z$ plane. Number of channels and $|\eta|$ coverages of the three components of the Inner Detector are summarized in Table 2.2. Figure 2.6 summarizes the layer positions with respect to pseudo-rapidity, $|\eta|$. The requirements on the alignment precision and the intrinsic accuracy are summarized in table 2.3. The resolution in the combined use of the whole Inner Detector system is discussed in section 3.1.

Table 2.2: The ATLAS Inner Detector parameters, number of channels and maximum η coverage, are shown.

Detector	channels [10^6]	$\lvert\eta\rvert$ coverage
Pixel (Run 1)	80	2.5
SCT	6.6	2.5
TRT	0.35	2.0

Table 2.3: Intrinsic measurement accuracy and mechanical alignment tolerances for the Inner Detector sub-systems. The numbers in the table correspond to the single-module accuracy for pixels, to the effective single-module accuracy for the SCT, and to the drift-time accuracy of a single straw for the TRT [3].

Item	intrinsic accuracy (μm)	alignment tolerances (μm) radial (r)	axial (z)	azimuth (r-ϕ)
Pixel				
Layer-0	10 (r-ϕ) 115(z)	10	20	7
Layer-1 and -2	10 (r-ϕ) 115(z)	20	20	7
Disks	10 (r-ϕ) 115(r)	20	100	7
SCT				
Barrel	17 (r-ϕ) 580(z)	100	50	12
Disks	17 (r-ϕ) 580(r)	50	200	12
TRT	130			30

The Inner Detector is required to be made with minimal substantial quantity of material in order to reduce the multiple scattering and generation of secondary particles before any particles reach the calorimeter. Figure 2.7 shows the material quantity as a function of $|\eta|$, expressed in radiation length[c]. The superconducting solenoid is made of Al-stabilized NbTi conductor, which is optimally structured for the material reduction in front of the calorimeter, contributing a total of ~0.66 radiation length at nominal incidence. The solenoid is designed to provide a 2.0 T axial magnetic field with a nominal operation current of 7.73 kA.

---

[c] The radiation length (X_0) is a mean distance over which a high energy electron (> 10 MeV) loses all but $1/e$ of its energy by bremsstrahlung, and 7/9 of the mean free path for pair production by a high-energy photon.

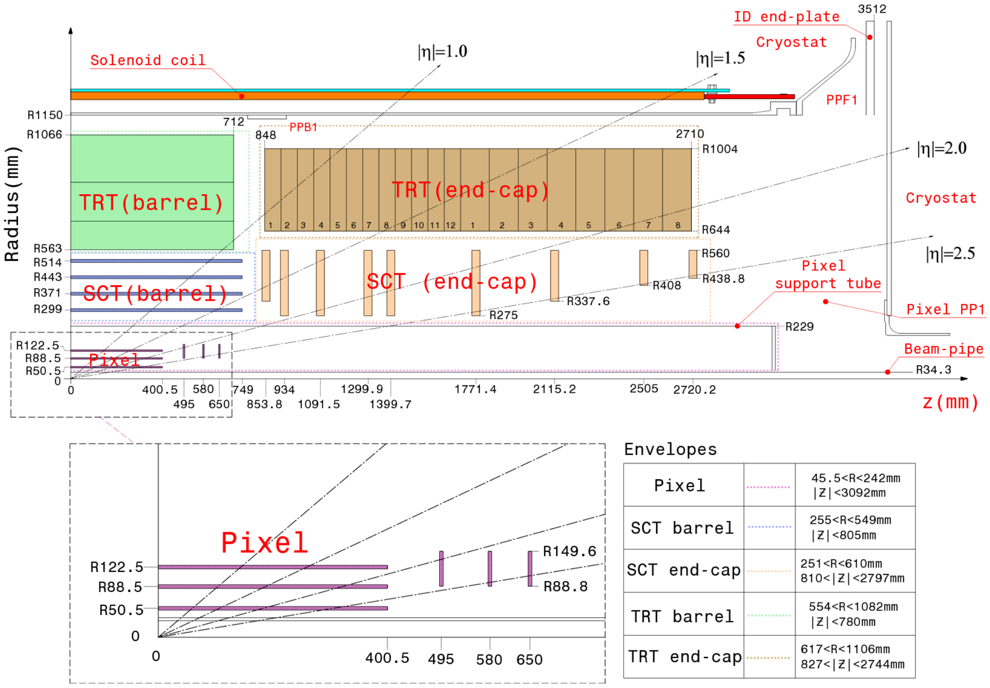


Fig. 2.6: The quarter section of the Inner Detector showing each of the major detector elements with its active envelopes. The labels PP1, PPB1 and PPF1 indicate the patch-panels for the services [3].

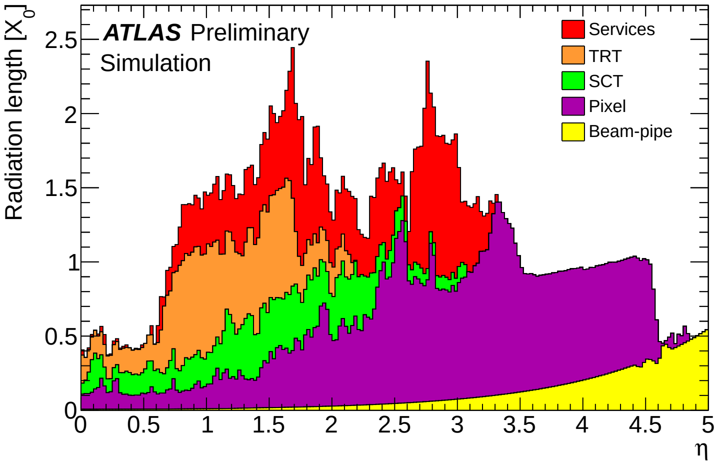


Fig. 2.7: The material distribution as a function of η, in radiation length (X_0). The contributions of the different detector elements, including the services and thermal enclosures are shown separately by filled color areas [32].

The Inner Detector consists of three independent but complementary sub-detectors which are the pixel detector, semi-conductor tracker, and the transition radiation tracker, described in the following.

To maintain an adequate noise performance after radiation damage, the pixel and SCT sensors must be kept at low temperature (approximately -5 to -10 °C) implying coolant temperatures of about -25°C.

Pixel Detector

The pixel detector is located at the closest position to the collision point. Therefore, the pixel detector is required to have the highest radiation tolerance together with the most precise measurement among all the sub-detector components. Since it has the greatest impact on the resolution of the track parameters and the vertices, it is the most important detector component for the identification of short-lived particles, such as τ-leptons and so-called b-hadrons. Thus, installation of the Insertable B-Layer improves the b-tagging performance significantly and it is discussed in section 4.6.2. The barrel part consists of three layers surrounding the beam axis at r =50.5, 88.5, and 122.5 mm. The endcap-disks also consist of three layers but in both the $+z$ and $-z$ sides.

The pixel sensors are made of silicon semi-conductor with p-n junction[d], and the reverse-bias voltage enlarges the depletion region. The sensors are 250 μm thick detectors, using oxygenated n-type wafers with readout pixels on the n^+-implanted side of the detector. When a charged particle traverses the depletion region, a number of electron-hole pairs (typically $\sim$20 k), which are proportional to the energy loss of the particle, are created. These pairs are amplified and then detected as a signal. When the signal lies above a threshold, a hit is recorded with a position and time.

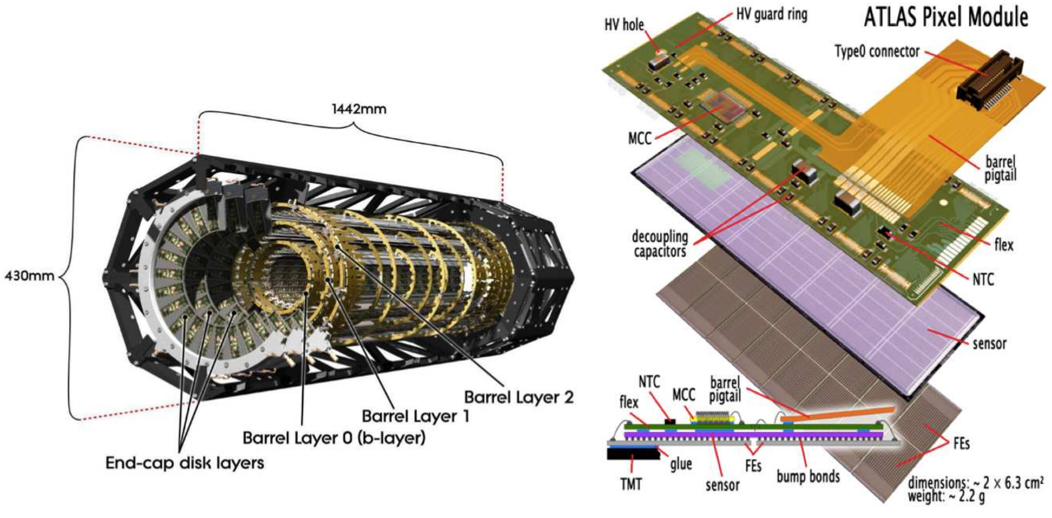
All modules are identical with the size of 19×63 mm^2. A pixel is sized at 50×400 $\mu\mathrm{m}^2$ in $r\phi \times z$ providing excellent spacial resolution. Each module is composed of 46080 pixels directly bump-bonded with the 16 front-end chips for readout. The size of the pixels at the front-end inter-chip regions is 50×600 $\mu\mathrm{m}^2$. Figure 2.8 shows a schematic of assembled layers of pixel detector (a) and a barrel pixel module (b). Every front-end chip has 2880 channels to amplify the signals from the sensor pixels. The number of pixel modules is 1744, and the total number of channels is about 80 million. The layout of front-end chips creates an insensitive region caused by the unread sensors. This is solved by the “ganged pixel” method: the pixels uncovered by the front-end chips are connected with the neighbor pixels (see fig 2.9).

In the barrel region, 13 modules are attached to a stave. The staves are mounted on a cylindrical support with 20° tilt angle to form a layer. Three layers, forming the barrel part, are placed at radii shown in fig 2.6. The innermost pixel layer was called B-layer in Run 1. The endcap disks are made of 8 sectors, and each sector has six pixel modules directly mounted. Figure 2.10 shows photographed examples of the barrel stave and disk sector.

The Pixel detector is capable to provide information about energy loss, so-called dE/dx, using Time-over-Threshold (ToT) [35]. By measuring the time during which the signal lies above the threshold in 25 ns pitches, the total charge can be estimated. This dE/dx measurement will give a hint to identify traversing particles together with the particles momenta measured by the track curvatures. However, except a few particular analyses, this functionality is not in use due to the small impact on main physics analyses.

---

[d] The boundary between p-type (in which the careers are the holes) and n-type (in which the careers are the electrons) semiconductors is called p-n junction. When the p-n junction is made, the conduction electrons annihilate with the electron holes and cause a region where no career exists. This region is called depletion region and the voltage caused in the depletion region is called built-in voltage. The voltage with the opposite direction to built-in voltage is called “reverse-biasing”.



(a) (b)

Fig. 2.8: (a) Schematic of the ATLAS Pixel Detector. The detector consists of three barrel layers with three discs in each endcap sides. The detector modules are mounted on carbon fiber support structures with integrated cooling circuits. (b) Assembly and cut views of an ATLAS Pixel Detector module. The bump-bonded front-end chips are indicated as FEs [33].

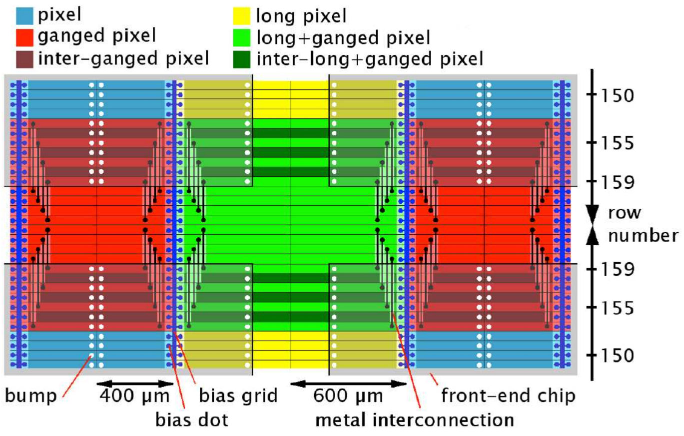


Fig. 2.9: The inter-chip region of the pixel sensor. The ganged pixels are introduced around the inter-chip regions [34].

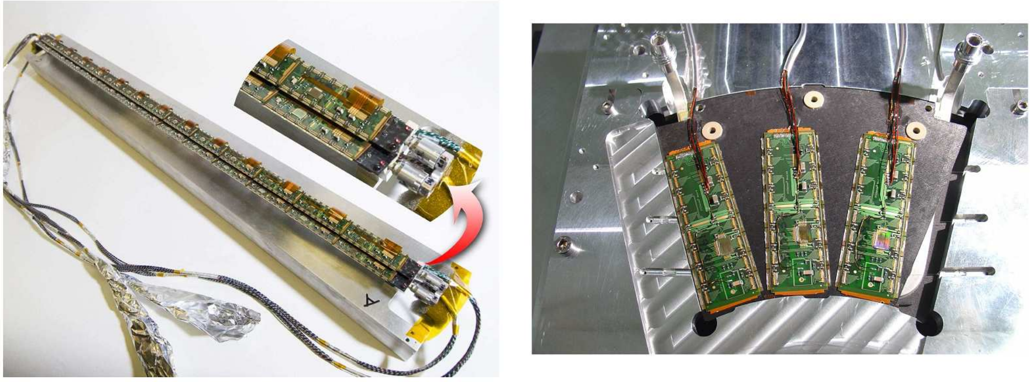


Fig. 2.10: Close up of bi-stave loaded with modules (left). Pixel disk sector during the attachment of modules (right). There are also three modules on the back of the sector [3].

Insertable B-layer

The Insertable B-Layer (IBL) [36] is a new innermost layer of the pixel detector which was installed in 2014 for the LHC Run 2. The beam pipe is replaced with one with a smaller radius in order to save the space for the IBL. Its purposes are to increase the tracking precision and cope with the increased number of pileup events and ration damages in the other pixel layers. Since the IBL is directly attached to the beam pipe, there is no endcap part. The IBL is installed at a mean radius of 33.25 mm covering width in $|z| \lesssim 330$ mm: $|\eta| < 2.58$. Since it is located closer to the interaction region, the pixels in IBL are sized shorter than for other pixel layers: 50×250 μm^2. There are 224 modules on 14 staves, forming about 6 million channels. Figure 2.11 (top) shows a 3D view of IBL modules in the ATLAS geometry. After the insertion of the IBL, Run 2 data taking started and is currently under operation. Figure 2.11 (bottom) shows a collision event display with IBL recorded in June 2015. The material distribution is also changed accordingly to the insertion of the IBL in Run 2 (see fig 2.12). A significant improvement of the impact parameter resolution is expected thanks to the IBL (cf. section 3.1.1).

Semi-Conductor Tracker

The second innermost sub-detector is the Semi-Conductor Tracker (SCT) which is made of silicon micro strips. The SCT barrel and endcap parts are composed of four layers and nine disks (for each side), respectively. The total number of modules is 2112 for the four barrel layers and 988 for the 9 endcap disks for each side.

Figure 2.13 shows a barrel module of the SCT which is made of the sensor, the readout chip, and the base-board. One SCT sensor consists of 768 strips in 80 μm pitch, and each strip is 128 mm long. The pair of sensors are glued on both sides of the base-board with 40 mrad stereo angle, which allows the 2D measurement of the position. High voltage is applied to the sensors via the conducting base-board. Contrarily to the pixel sensor, the SCT sensor is made of p-in-n type semiconductor. The SCT wafer is 285 μm thick with silicon strips implanted on one side. The strips are highly doped p-type semiconductor and the wafer is n-type silicon due the small amount of doping impurities.

The endcap module is trapezoidal in order to shape the disk. The silicon strips are also radially aligned. There are three types of modules for the disks, and they are called inner, middle, and outer types. The pitch between the strips are 54–69, 70–95, and 71–90 μm for inner, middle, and outer types, respectively.

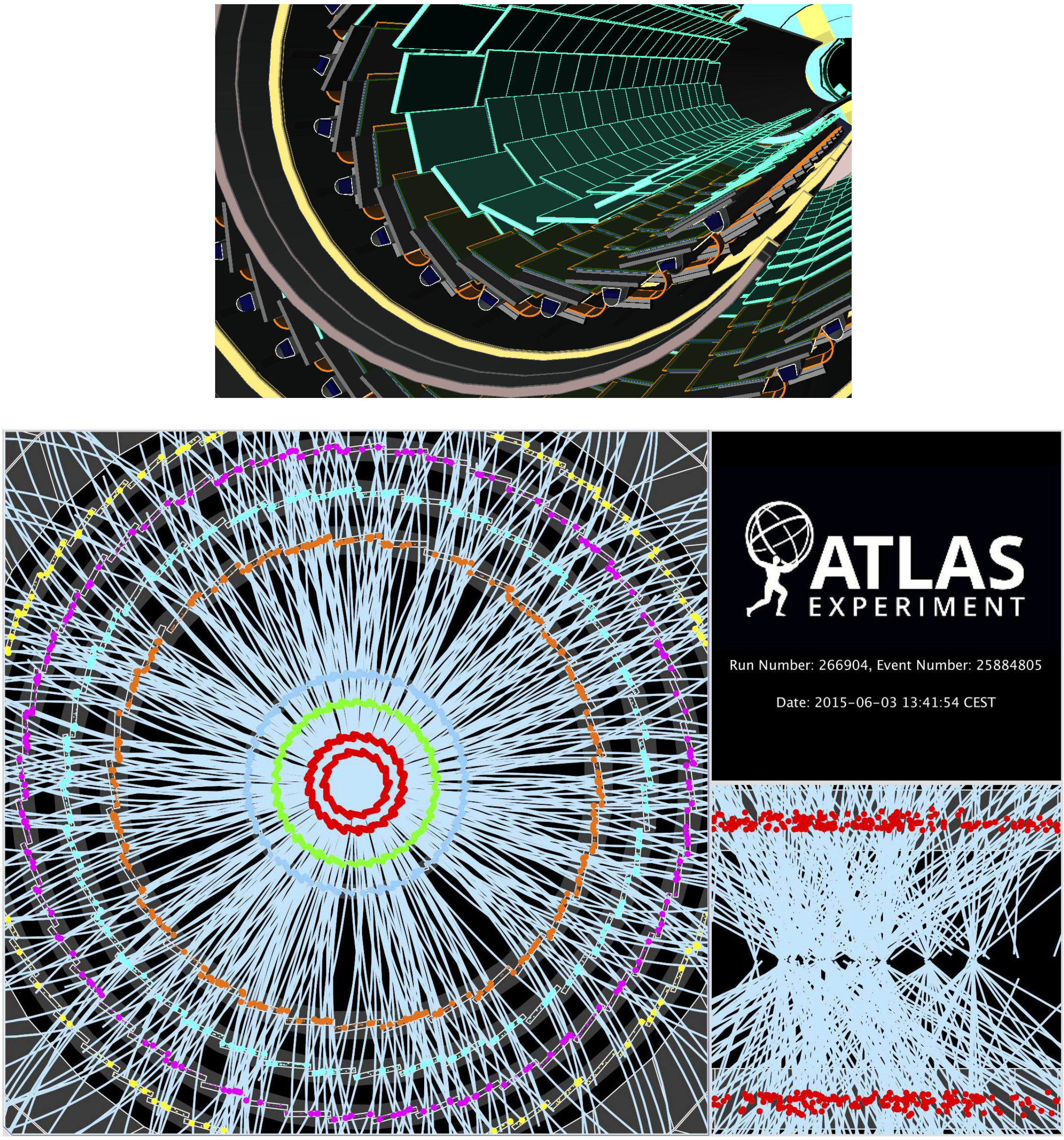


Fig. 2.11: The IBL modules highlighted in cyan in a 3D view, together with the other pixel layers (top) [36]. A collision event display after the IBL insertion (bottom), the IBL is highlighted in red together with the old innermost layer [37].

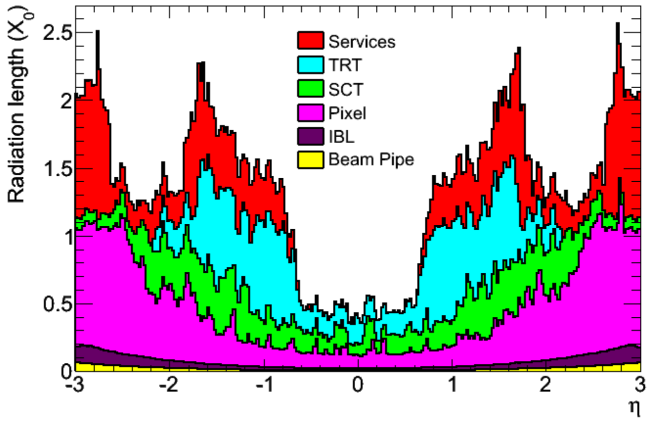


Fig. 2.12: Simulated material distribution expressed in the radiation length as a function of η. The IBL material amount is compared to the other Inner Detector components [36].

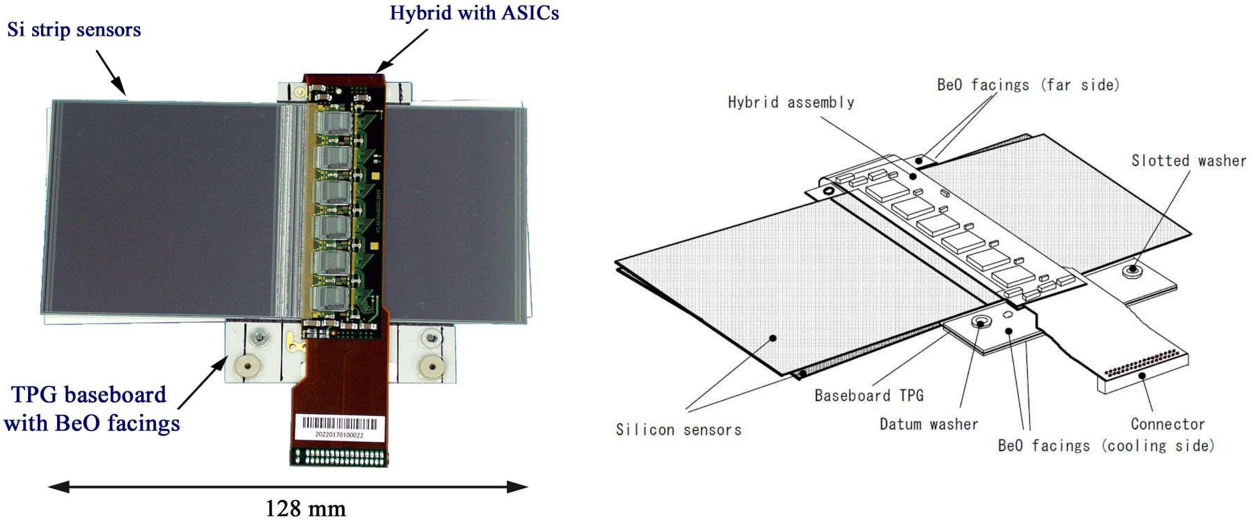


Fig. 2.13: Photograph (left) and drawing (right) of the SCT barrel module, showing its components. The high thermal conductivity between the coolant and sensors are kept by the base-board which is made of thermal pyrolytic graphite (TPG). The TPG base-board has also excellent electric conductivity so that it is used for applying the bias voltage to the micro strip sensors [38, 39].

Fig. 2.14: Photograph of TRT endcap wheel during assembly. The inner and outer C-fiber rings can be seen, as well as the first layer of straws and the first stack of polypropylene radiator foils beneath it [3].

Transition Radiation Tracker

The outermost part of the Inner Detector is the Transition Radiation Tracker (TRT), which is designed to work at room temperature contrarily to the pixel detector and the SCT. Taking advantage of a light and cheap gas detector, it extends the tracking volume radially ($r_{\mathrm{outer}} = 1082$ mm), allowing to have better momentum resolution. The TRT is composed of many drift tube gas chambers. The gas is mainly xenon (Xe) mixed with CO_2 and O_2. The Xe gas works as the absorber for the X-rays while CO_2 works as the stabilizer for the gas amplification. At the center of the straw tube which has 4 mm diameter, a gold coated tungsten wire (30 μm diameter) is strung for read out.

For the purpose of reducing the barrel hit occupancy, the barrel part of the TRT is divided at $z = 0$, and the readout circuits are placed at the outer edge for both $+z$ and $-z$ sides. Each barrel part consists of 52,544 tubes which are parallel to the beam axis, and the transition radiation material (polypropylene fiber) is interleaved in between the tubes. The barrel (endcap) tubes cover $|\eta|$ <0.7 (2.0) where a minimum of about 36 (22) straws are traversed for a track with $p_{\mathrm{T}} > 0.5$ GeV. The endcap part is composed of 122,880 tubes aligned radially in the $+z$ and $-z$ sides, respectively. Figure 2.14 shows the endcap wheel of TRT with the readout circuits along the outer ring.

The TRT provides continuous tracking to enhance pattern recognition and improve the momentum resolution within $|\eta| < 2.0$ and electron identification complementary to that of the calorimeter over a wide range of energies. The measurement of the drift time of the ions generated by incident charged particles in each tube allows to obtain 130 μm resolution per tube. When a high energy charged parti-

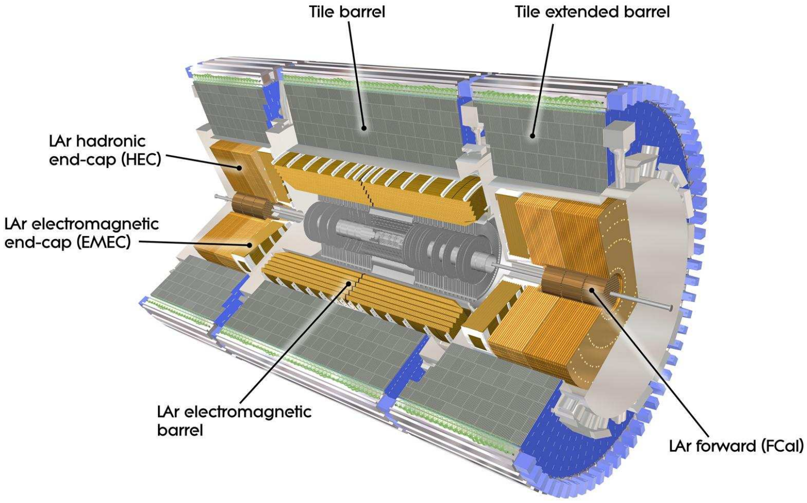


Fig. 2.15: The overview of the ATLAS calorimeter system.

cle goes through the layer of polypropylene fiber which has 15 μm diameter, a transition radiation[e] is emitted. This transition radiation helps the identification of particles, especially electrons from pions. The particle identification of electrons or photons using TRT is detailed in section 3.2.

2.2.3. Calorimetry

The purpose of the calorimeter system is to measure the energy of the incident particles and identify electrons, photons, and jets. It is also very important to limit punch-through of the hadrons into the muon system. The overview of the calorimeter is illustrated in fig 2.15. In the ATLAS detector, liquid argon (LAr) is used for the electromagnetic calorimeter and forward hadronic calorimeter. Scintillation tiles are used for the barrel part of the hadronic calorimeter. The calorimeter covers a large η-range ($|\eta| < 4.9$) in order to have excellent hermeticity which allows the accurate measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$. Together with the large η-coverage, the thickness will also ensure a good measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$, which is important for many physics signatures such as $VH \to Vbb$ where $V = W/Z$ with $W \to \ell\nu$ and $Z \to \nu\nu$. Using different techniques, the calorimeters are suited to the widely varying requirements of the physics processes of interest and of the radiation environment over this η-range. Figure 2.16 illustrates the electromagnetic and hadronic calorimeter positions and the forward calorimeters.

The calorimetry is performed in three main sub-systems, which are the electromagnetic calorimeter, the hadronic calorimeter, and the forward calorimeter explained in the following sub-sections.

Electromagnetic calorimeter

The electromagnetic (EM) calorimeter estimates the energy detecting the shower from photons and electrons which are generated through photon conversions and bremsstrahlung. The EM calorimeter consists of lead plates and LAr layers for the photon conversion and the detection, respectively. The

---

[e] Transition radiation is the emission of light in the range of visible to X-ray when a charged particle traverses a boundary between materials of different densities. Since the intensity of the emitted radiation is proportional to the γ factor, $\gamma(= E/m)$, this effect is significant for electrons which have a light mass.

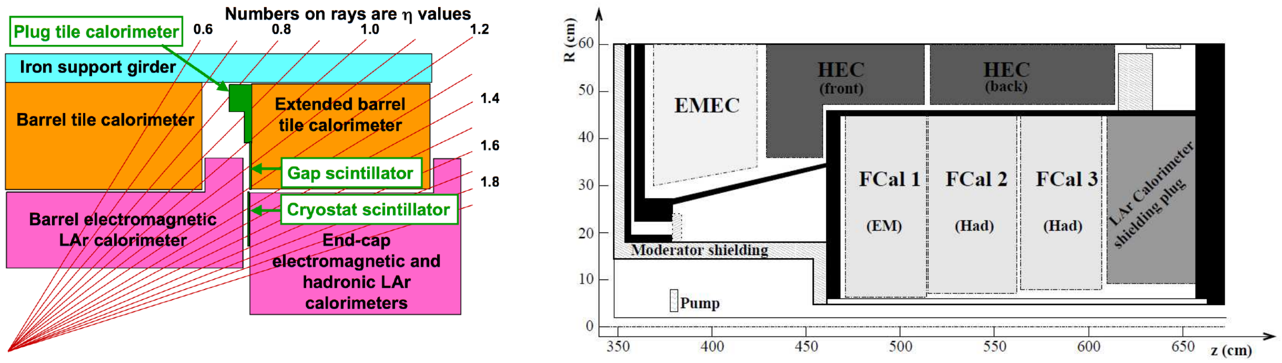


Fig. 2.16: Schematic of the transition region between the barrel and endcap cryostats, where additional scintillator elements are installed to provide corrections for energy lost in inactive material (not shown), such as the LAr cryostat and the inner-detector services (left). Schematic diagram showing the three FCal modules located in the endcap cryostat (right). The material in front of the FCal and the shielding behind it are also shown. The diagram has a larger vertical scale for clarity [3].

barrel part covers $|\eta|$ <1.475 and the two endcaps cover 1.375< $|\eta|$ <3.2, each housed in their own cryostat. The central solenoid and the LAr calorimeter share a common vacuum vessel, thereby eliminating the two vacuum walls inside the calorimeter. Similarly to the TRT barrel, the EM calorimeter barrel is divided by a small gap at $z = 0$. The Electromagnetic Endcaps (EMEC) of LAr calorimeters are divided into coaxial wheels. The outer (inner) wheel covers $1.375 < |\eta| < 2.5$ $(2.5 < |\eta| < 3.2)$. The accordion structure of the EM calorimeter provides hermetic coverage in ϕ, without azimuthal cracks, while enabling fine segmentation in $\eta - \phi$ and readout in depth (see fig 2.17). The lead thickness in the absorber plates has been optimized as a function of η in terms of EM calorimeter performance in energy resolution. Over the region devoted to precision physics ($|\eta| < 2.5$) the EM calorimeter is segmented in three sections in depth. In the range of $|\eta| < 1.8$, a pre-sampler detector corrects for the energy lost by electrons and photons upstream of the calorimeter.

Hadronic calorimeters

The ATLAS hadronic calorimeter is composed of two sub-systems.

Tile calorimeter: the scintillation tile hadronic calorimeter is placed directly outside the LAr EM calorimeter. The coverage is $|\eta| < 1.0$ $(0.8 < |\eta| < 1.7)$ for the (extended) barrel(s) as shown in fig 2.16. It is also a sampling calorimeter using steel as the absorber and scintillating tiles as the active material. There are 64 modules deployed azimuthally in the barrel and extended barrels. The inner and outer radii of the barrel and extended barrels are 2.28 m and 4.25 m, respectively. Two sides of the scintillating tiles are read out by wavelength shifting fibers into two separate photo-multiplier tubes.

Hadronic endcap calorimeter: the Hadronic Endcap Calorimeter (HEC) consists of two independent wheels per endcap. The closer to the interaction point is the front wheel and the other is the back wheel. These two sets of front and back wheels cover $1.5 < |\eta| < 3.2$. Each wheel is composed of 32 identical wedge-shaped modules which are divided into two segments in depth, for a total of four layers per endcap. The front two endcap wheels are built from 25 mm parallel copper plates, while

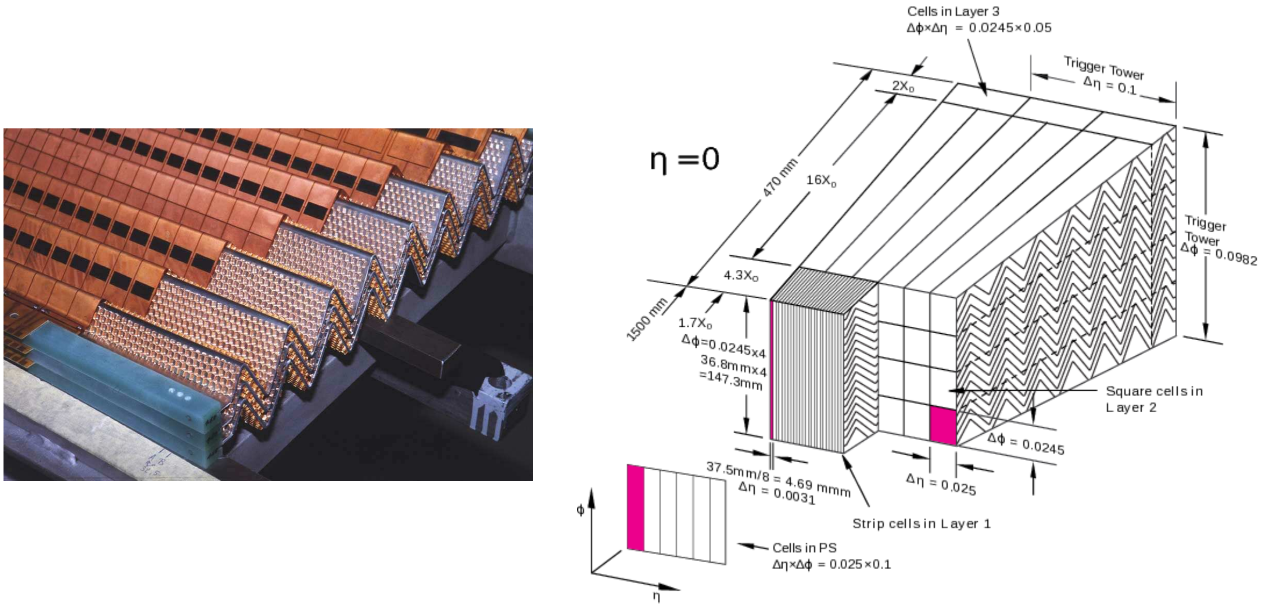


Fig. 2.17: The photographed accordion shape of LAr calorimeter (left), the honey-comb structured spacers are also visible. Sketch of a barrel module where the different layers are clearly visible with the ganging of electrodes in ϕ (right). The granularity in η and ϕ of the cells of each of the three layers and of the trigger towers is also shown.

back wheels use 50 mm copper plates. The inner and outer radii of the copper plates are 0.475 and 2.03 m, respectively. For only the front wheels, the inner copper plates' radii become 0.372 m in the overlapping region with forward calorimeter. In between the copper plates, there are LAr gaps which provide the active medium for the sampling calorimeter.

Forward calorimeter

The Forward Calorimeter (FCal) is integrated into the endcap cryostats, as this provides clear benefits, which are hermetic coverage of calorimeter and the reduced radiation background in the muon spectrometer. FCal covers the very forward region of $3.1 < |\eta| < 4.9$, and is composed of three modules in each endcap: the first is a LAr EM calorimeter, using copper for the absorber while the last two, made of tungsten, measure predominantly the energy of hadronic interactions. Each module consists of a metal matrix, with regularly spaced longitudinal channels filled with the electrode structure consisting of concentric rods and tubes parallel to the beam axis.

2.2.4. Muon spectrometer

The purpose of the Muon Spectrometer (MS) is to measure the momenta of the charged particles escaping through the calorimeter, which are mainly muons. The Muon Spectrometer works together with the toroidal magnet, which generates the different direction of the magnetic field from that of the solenoid. Figure 2.18 shows the structure of the MS system together with the toroidal magnets. Three MS layers cover the barrel part and each endcap part, respectively. There are eight toroidal magnets for the barrel part, and two endcap magnets. The toroidal magnetic field changes with the pseudo-rapidity, η, as illustrated in fig 2.19. The barrel part of the toroidal magnets covers $|\eta| < 1.4$ with a bending power of 1.5–5.5 Tm, while the endcap part covers $1.6 < |\eta| < 2.7$ with a bending power of 1.0–7.5 Tm. Over the transition region ($1.4 < |\eta| < 1.6$), magnetic deflection is provided by a combination of barrel and endcap fields (see fig 2.19). The MS measurement is greatly affected by multiple scattering for the low p_{T} particles as it is located in the outermost region. Muon momenta down to about 3 GeV can be measured by the MS alone. As well as in the case of the Inner Detector, the MS momentum resolution becomes worse for the very high momentum particles because the curvature will be very small. The MS provides excellent charge identification at very high p_{T} up to 3 TeV. The momentum resolution provided by the MS is a few percent level for muons with $10 \lesssim p_{\mathrm{T}} \lesssim 100$ GeV, and it increases in $p_{\mathrm{T}} > 100$ GeV due to the multiple scattering. For a muon with $p_{\mathrm{T}} \sim 1$ TeV, the momentum resolution is about 10% [3].

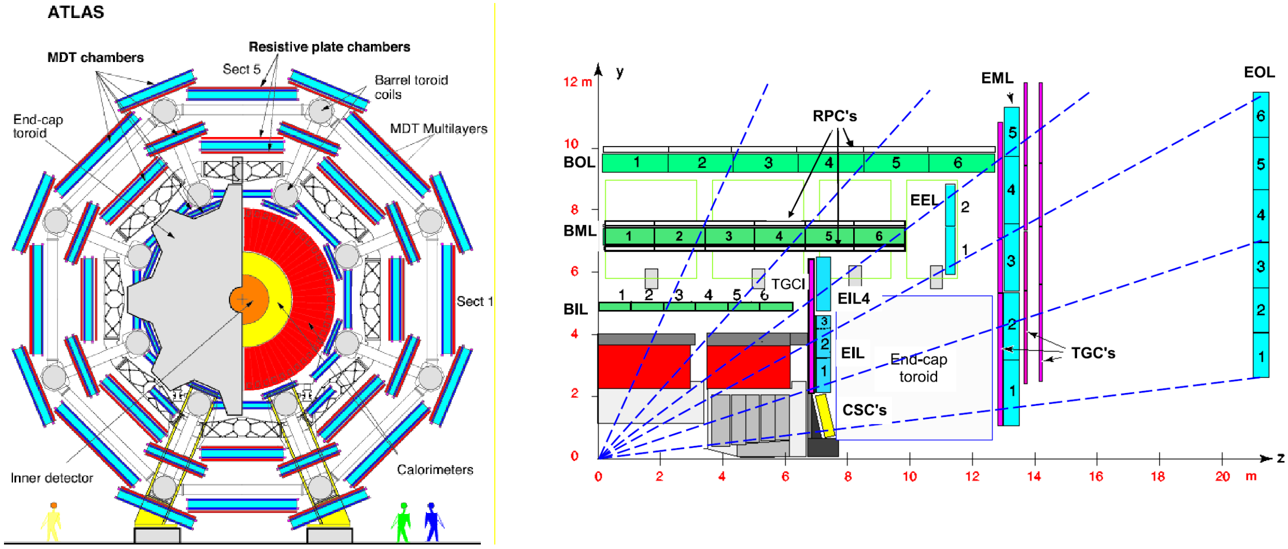


Fig. 2.18: The x-y projection of the muon spectrometer (left) [40]. Cross-section of the muon system in a plane containing the beam axis (bending plane). Infinite-momentum muons would propagate along straight trajectories which are illustrated by the dashed lines and typically traverse three muon stations (right) [3].

The MS system is composed of four sub-detectors: they are Monitored Drift Tube chambers (MDT), Cathode Strip Chambers (CSC), Resistive Plate Chambers (RPC), and Thin Gap Chambers (TGC). The MDT (CSC) detector measures the tracks precisely in the central (forward) η-range of $|\eta| < 2.0$ ($2.0 < |\eta| < 2.7$), while the RPC (TGC) detector provides the trigger signals in the central (forward) η-range of $|\eta| < 1.05$ ($1.05 < |\eta| < 2.4$).

The MDT is composed of two drift tubes using mixed gas made of Ar (97%) and CO_2 (3%). It has an excellent resolution of 80 μm, but the MDT cannot be used with a counting rate higher than 150 Hz/cm^2. Therefore, the CSC is used in the forward region where the counting rate is higher.

The CSC can be used with counting rate of up to 1000 Hz/cm^2 with very high spacial and time resolution. The CSC wheel is composed of eight modules of Multi Wire Proportional Chamber (MWPC). The wires are strung in the pitch of 2.54 mm. Although it cannot reconstruct the tracks, it provides pointing resolution of 60 μm.

The RPCs are gas chambers with spacial and time resolution of 1 mm and 1 ns, respectively. The 4.5 kV voltage is applied to the chambers which contain a mixed gas made of $C_2H_2F_2$ (97%) and C_4H_{10} (3%). RPC-1 and -2 are used for low p_{T} (6–9 GeV) trigger while RPC-3 is used for high p_{T} (9–35 GeV) trigger.

The TGC is based on MWPC, with the copper wires strung radially in the R direction and the carbon strips strung in the ϕ direction. This structure provides a 2D position readout with very fast response

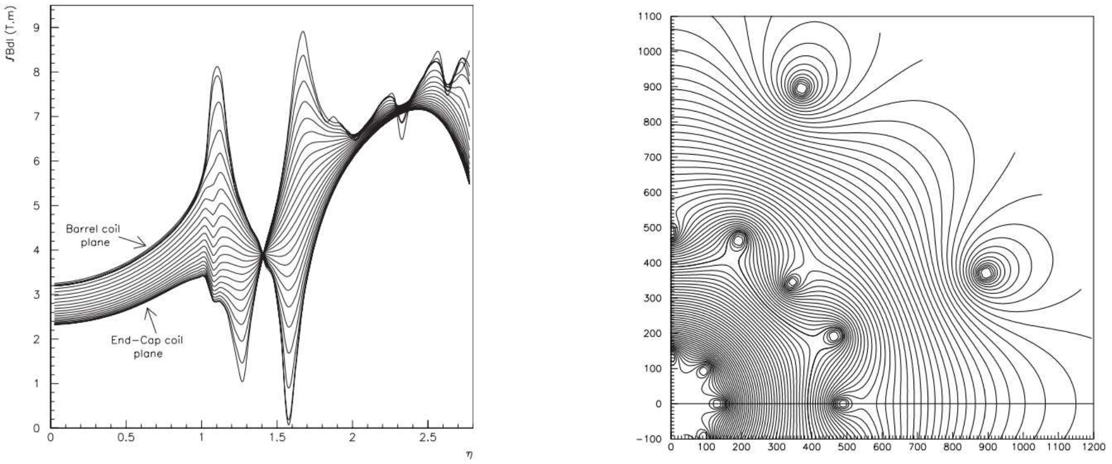


Fig. 2.19: The bending power dependence on pseudo rapidity η (left). The structure of the toroidal magnetic field in the transition region on the x-y plane (right) [41].

for triggering.

2.3. Trigger

The very high collision rate provided by the LHC requires efficient signal selection and rejection of uninteresting events at the same time. The space on disk needed for an event is typically 1 MB so that the necessary space per second becomes $\mathcal{O}(1)$ PB, which is not feasible to store all the data. The trigger system is responsible for selecting interesting events and reducing the event recording rate down to about 500–1000 Hz from the bunch crossing rate (nominally 40 MHz) which is much higher than the experimental acceptance.

The ATLAS trigger system is organized in three levels. The first level is based on custom-made hardware and uses coarse-granularity calorimeter and muon information. The second and third levels are often referred to as high level trigger (HLT), and implemented as software algorithms and can use the full detector granularity. Figure 2.20 shows the overview of the ATLAS trigger and data-acquisition (DAQ) flow.

Level-1 trigger

The level-1 trigger is based on the calorimeter and the muon system. It is in charge of the event rate reduction down to about 100 kHz from a collision rate of 40 MHz. The trigger decision has to be performed within a waiting (decision) time of 2.5 μs. During the decision time, data from all the sub-detectors will be sent to the pipelines for further processing in the level-2 computing farms. The decision on the level-1 trigger is based on the E_{T} or p_{T} threshold and the multiplicity of the physics objects caught in the sub-detectors. They are electrons, muons, photons, jets, hadronic tau-decays, and $E_{\mathrm{T}}^{\mathrm{miss}}$. The algorithms to compute E_{T} of these objects use calorimeter towers of 0.1×0.1 granularity in the $\Delta\eta \times \Delta\phi$ plane (coarser granularity for $|\eta| > 2.5$) [42]. A region of interest (ROI) is associated to the objects which pass the level-1 trigger threshold and this information is sent to the level-2 trigger for further processing but not used for the level-1 trigger decision.

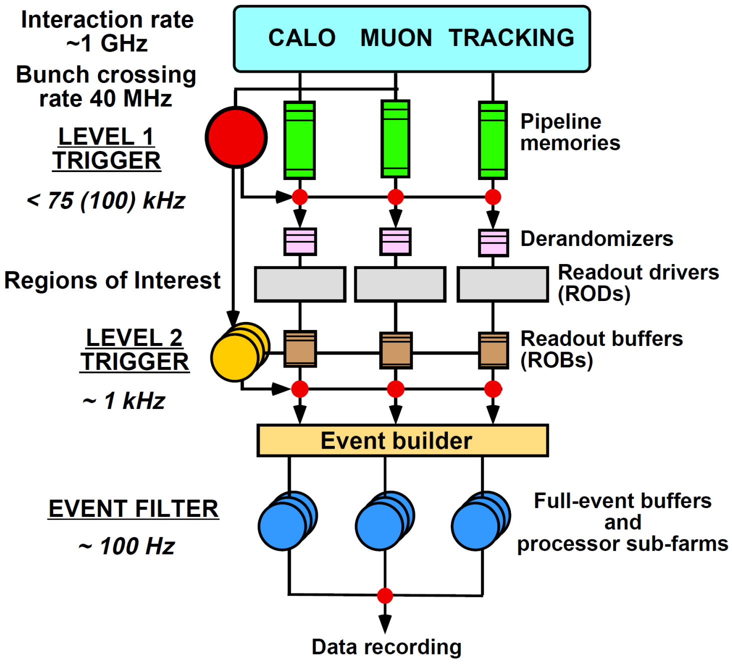


Fig. 2.20: Block diagram of the Trigger/DAQ system [42]. Along with the levels of trigger system, the original design rate acceptances are indicated.

The level-1 trigger decision is based on 256 bits, which store the object multiplicities and E_{T} thresholds. Using this 256 bits information, level-1 trigger menus are formed by a combination of certain level-1 object thresholds/multiplicity. Each level-1 trigger menu can be also associated with prescale factors[f].

Level-2 trigger

The level-2 trigger algorithms are run on a computing farm which consists of $\mathcal{O}(1000)$ PCs. It is designed to reduces the trigger rate down to a few kHz. The level-2 latency is designed to be about 10 μs with average processing time of about 40 ms, including the data transfer.

For further rejection of the events, the ROI information associated to level-1 trigger-object is used in the level-2 together with the other information which was not available in the level-1, such as tracks from the Inner Detector.

Event filter

The Event Filter (EF) uses the information sent from level-2 as the seeds for a more precise analysis. The algorithms are close to the offline reconstruction which provides further rejection at the EF. The EF reduces the rate down to less than 1 kHz, with average processing time of about 4 s.

Once the event is selected by EF, it is stored permanently at CERN Computing Center (called Tier-0) and processed by the full offline reconstruction software, and the latest version of object calibration, etc. Finally the processed data are spread on the grid computing centers located all over the world.

---

[f] In order to reduce the trigger rate, data amount to be further processed is controlled by a factor called prescale, which stores the event among N events. The prescale factors are defined for a luminosity-block.

2.4. Simulation

The simulation of the collision events are based on the random number generation, so called Monte Carlo (MC) method. The term MC is always used to refer to MC simulation. Figure 2.21 shows the comparison of the full chain flow and fast chain flow. The actual data taken from the detector ends into the same data format. In this section, computing tools (software or resources) for the simulation are discussed.

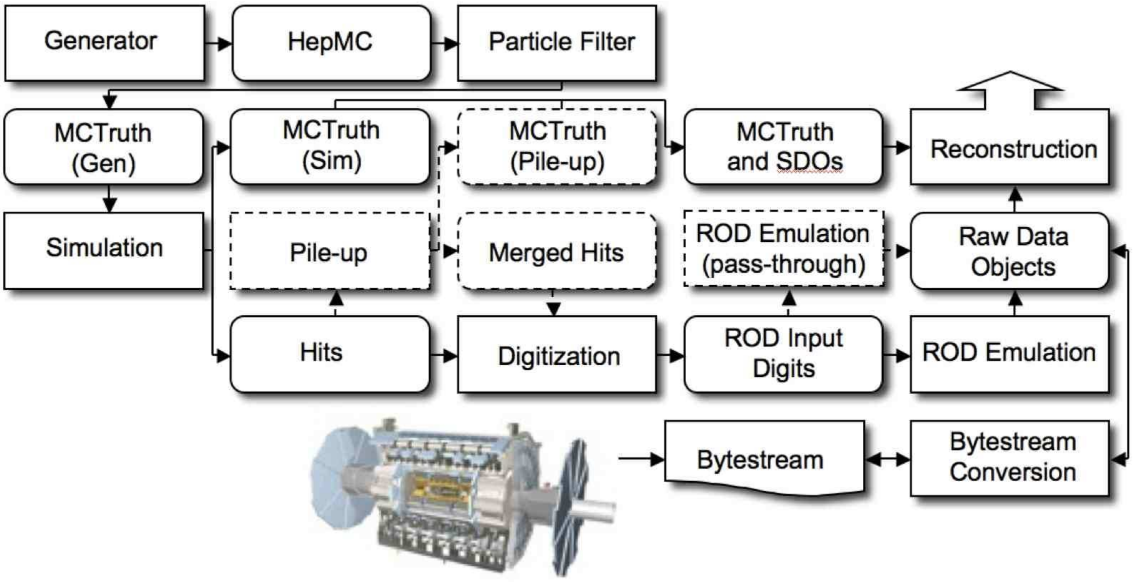


Fig. 2.21: The flow of the ATLAS simulation software, from event generators (top left) to reconstruction (top right). Algorithms are placed in square-cornered boxes and persistent data objects are placed in rounded boxes. The optional pileup portion of the chain, used only when events are overlaid, is dashed. Generators are used to produce data in HepMC format. Monte Carlo truth is saved in addition to energy depositions in the detector (hits). This truth is merged into Simulated Data Objects (SDOs) during the digitization. Also, during the digitization stage, Read Out Driver (ROD) electronics are simulated [43].

2.4.1. Event generation

Physical processes in inelastic collisions are provided with random number within certain conditions of final states and certain kinematic constraints. In order to simulate the physics processes with the LHC collisions, *event generators* are first used to create a finite number of events as they are predicted to occur in the experiment. For each event, the four momenta of the final state particles are provided.

The event generation is done step by step through the *hard process*, *parton shower*, and *hadronization*. First in the hard process, the main interaction with partons is considered. Second in the parton shower process, photon and gluon emissions are generated using QED and QCD, respectively. Finally in hadronization, bunches of particles are generated according to the estimation provided experimentally, as the hadronization process with non-perturbative theory is not achieved to simulate. The event generators used in this thesis are PYTHIA6 [44], PYTHIA8 [45], HERWIG [46, 47], ALPGEN [48], ACERMC [49], MC@NLO [50], SHERPA [51], and POWHEG [52–54]. The first five use exact leading-order (LO) matrix elements while the last three use next-to-leading-order (NLO) matrix elements. For the tau decay, an add-on program called TAUOLA [55] is used, in which the polarization of taus is taken into account. Contrarily to the tau-decays, there was no uniform treatment for the b-hadron decays in Run 1. This was identified as a source of different b-jet efficiency in many samples. The EVTGEN [56] package, therefore, is exploited for simulated samples used in Run 2 analyses, however, not in the sample used in this thesis.

2.4.2. Detector simulation

After the generation of the collision events, the final state particles pass through the process called detector simulation. In the ATLAS experiment, two approaches of detector simulation are developed, which are called *full simulation* and *fast simulation*.

The full simulation based on GEANT4 [57] computes the interaction between final state particles and detector components or materials, which produces the *hit* information. The *digitization* process is performed using hits and the detector response is emulated. Finally, the reconstruction process runs, using the digitized detector response, to create the physical object segments such as tracks, and calorimeter energy cluster.

Although full simulation describes the real data better, it requires huge computing resources, e.g. 15 minutes per event. To overcome this computing resource issue, the fast simulation called ATLFAST was developed [43]. It aims to simplify the showering simulation in the calorimeters which takes 80% of the full simulation time. The approach taken was to remove the low energy electromagnetic particles and replace them with pre-simulated showers. Different fast simulation softwares for three main sub-detector systems (the Inner Detector, the calorimeter, and the Muon Spectrometer) were developed and they can be switched with the full simulation, system by system: e.g. full simulation for the tracking and muon reconstruction with fast simulation for the calorimetry. The improvement with this example configuration in time is very large: 10 times faster than the full simulation.

2.5. Data processing framework

As mentioned in the previous section, huge computing power and flexibility are necessary to simulate the large scale samples used for the data analysis. In order to meet these needs, various computing resources including software and hardware are developed.

2.5.1. ATHENA framework

All the steps shown in fig 2.21 are handled by the ATLAS data reconstruction framework, called ATHENA [43]. It is based on the GAUDI framework originally developed in the LHCb experiment. The ATHENA framework is mainly written in two computing languages: C++ and PYTHON. The C++ parts, called modules, consist of classes used for the repetitive calculation, since they are fast after compilation. All these compiled modules are configurable via the PYTHON part, called “job option”. The combined use of these two parts enables the handy development by reducing the number of compilation, and the framework becomes more flexible for a lot of use cases.

2.5.2. Distributed computing

For the requirement of large computing power for the sample production, the Worldwide LHC Computing Grid (WLCG) project has been launched. It consists of more than 140 computing centers in about 35 countries, with the four LHC experiments. Each institute provides their computing resources of both data storage and the CPUs.

On WLCG, a software system which handles all the production and analysis was developed, the so-called *production and distributed analysis* (PANDA). All the jobs based on ROOT and ATHENA can be run with PANDA on WLCG.

In the ATLAS experiment, data taken by the detector is along the hierarchy of **Tier-0**, **Tier-1**, and **Tier-2**. Together with the less than 20% of whole grid CPU to perform first data processing, Tier-0 has a mass disk and tape storage system, which is responsible for the prompt data collection directly from the detector. Then Tier-0 distributes these data to Tier-1 centers around the world for further processing and analysis. Tier-2 receives processed data from Tier-1, in the data formats of Event Summary Data (ESD) and Analysis Object Data (AOD). In ATLAS Run 1, the ESD is processed into AOD, and then into the format called D3PD which can be analyzed without the ATHENA framework.

The ATHENA software is designed to have backward compatibility and the updates are being validated every night (nightly versions). The software coordinators provide recommendations about stable releases to perform analysis. The versions used in this thesis are Release 17 and 20 for chapter 4. Release 17 was used for the all Run 1 analyses including the one presented in chapter 5.

3.Particle and object identification

In this chapter, the reconstruction and identification of all the physics objects using the detector outputs are discussed. The physics objects are tracks, vertices, photons, electrons, muons, tau-jets, jets, and missing transverse momentum.

Figure 3.1 shows a partial transverse cut view of the ATLAS detector with many particles traversing.

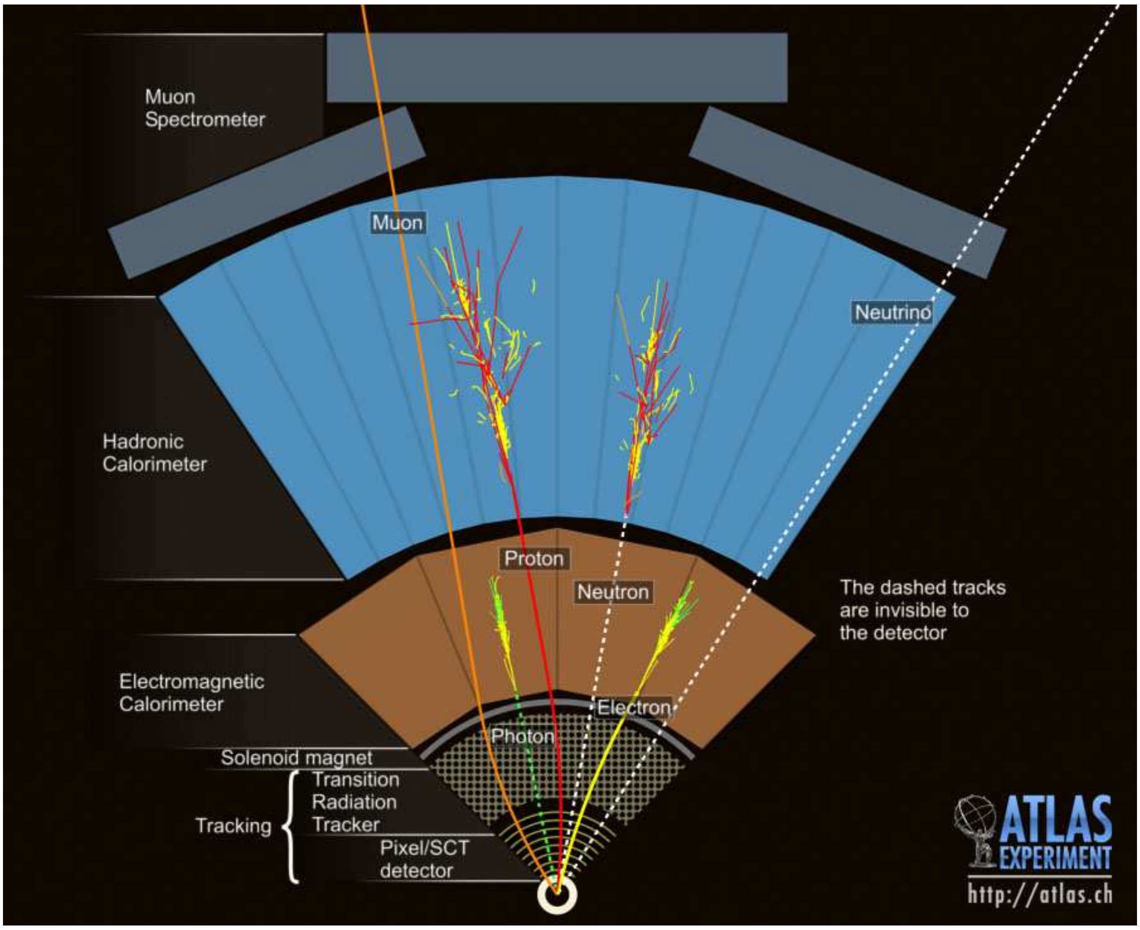


Fig. 3.1: An overview of particle identification in the ATLAS detector. The solid and dashed curves show the tracks of charged and neutral particles. Arising from the interaction region (beam axis), the muon goes through the whole detector while being tracked by the Inner Detector and the Muon Spectrometer, the electron and the photon are caught by mainly the EM calorimeter with and without being tracked, respectively, the proton and the neutron are trapped by mainly the hadronic calorimeter with and without leaving a track, respectively, and the neutrino passes through the entire detector without leaving its signature in any of the sub-detectors.

After the discussion on the tracking and vertexing in section 3.1, electron and photon identification is described in section 3.2, followed by the muon identification in section 3.3. A brief review of hadronic tau decay identification in section 3.4 is followed by the jet reconstruction technique explained in section 3.5. Finally using all these objects, the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is discussed in section 3.6.

3.1. Tracks and primary vertex

As mentioned in section 2.2.2, the Inner Detector is the key detector for *tracking* and *vertexing*, which are both crucial for b-tagging as described in the next chapter. In this section, the methods of tracking and vertexing are briefly discussed.

3.1.1. Tracks

From the hit information recorded in the Inner Detector, reconstruction of the track of the charged particles can be performed by connecting spatial points with a helix considering the magnetic field generated by the superconducting solenoid. First, the three dimensional space points of the hits are obtained from the pixel modules and double sided SCT modules. The raw timing information from the TRT is calibrated into drift circular curves. Second, various track finding algorithms find tracks arising from the interaction region.

In the default track finding, called *inside-out* tracking, space points from the three pixel layers and the first SCT layer are combined to form track seeds, and these seeds are extended, while removing outliers and fake tracks by applying quality cuts [58]. Then these track candidates are extended outward to the TRT, and the final refitting is performed using the information from the pixels, SCT, and TRT detector.

A converse strategy, called *outside-in* tracking, is also employed for finding the tracks stemming from long-lived particle decays, such as so-called V_0 particle decays[a], photon conversions, and the interaction between a particle from collision and detector material. Outside-in tracking uses TRT seeded track-segments extrapolated inward to the SCT and pixel hits.

Track parameters

Once the track is reconstructed, the track parameters are defined for each track. To describe the track of a charged particle in a magnetic field, five parameters are needed: transverse impact parameter, longitudinal impact parameter, azimuthal angle, polar angle, and inverse transverse momentum.

- *Azimuthal angle*: ϕ is the angle between $\overrightarrow{p_{\mathrm{T}}}$ and the x-axis in the transverse plane, where $\overrightarrow{p_{\mathrm{T}}}$ is defined at the track's perigee which is given below. The range of ϕ is defined to be within $-\pi < \phi \leq +\pi$. A similar angle, ϕ_0, is defined as the angle between the x-axis and a vector pointing the perigee from the coordinate system origin $\mathcal{O}(0,0,0)$, as shown in fig 3.2.
- *Transverse impact parameter*: d_0 is the track's distance of closest approach to the interaction point in the transverse plane, where the point defining d_0 on the track is called the perigee. However, the d_0 is firstly defined as a closest approach to $\mathcal{O}$ as shown in fig 3.2, since the position of the interaction point is unknown when the tracking algorithm runs. Its sign is positive when $\phi_0 - \phi = \frac{\pi}{2}$ is true and negative otherwise.
- *Longitudinal impact parameter*: z_0 is the z-coordinate of the track's perigee.

---

[a] Electrically neutral strange hadrons with relatively long lifetime are called V_0 particles, due to the signature of V-shaped track pair left by their decays. They often decay far from the interaction region but before the calorimeter, typically $K_S^0 \to \pi^+\pi^-$ and $\Lambda \to p\pi^-$.

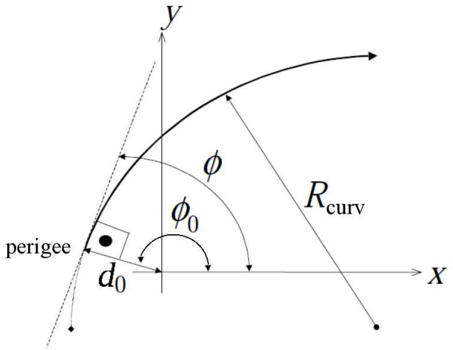


Fig. 3.2: The definition of track parameters, d_0, ϕ_0, ϕ, and $R_{\rm curv}$.

- *Polar angle*: θ is the angle between $\vec{p}$ and the z-axis in the r-z plane defined in $0 \leq \theta < \pi$. The relation between θ and η was shown in eq 2.4.
- *Inverse transverse momentum*: $Q/p_{\rm T}$ involves the electric charge of a particle, since the charge is determined by the curvature. It is calculated with the curvature radius $R_{\rm curv}$ of the track. This means the quantity $Q/p_{\rm T}$ is measured with Gaussian error but not $p_{\rm T}$. The $p_{\rm T}$ resolution is almost proportional to the spacial resolution of the detector, since $R_{\rm curv}$ is measured by fitting to the detectors' hit points. In a magnetic field B, the relation between $Q/p_{\rm T}$ and $R_{\rm curv}$ is given by the equation below:

  $$Q/p_{\rm T} = (0.3BR_{\rm curv})^{-1}, \tag{3.1}$$

  where the dimensions of Q, $p_{\rm T}$, and $R_{\rm curv}$ are given in [e], [GeV], and [m], respectively.

Later on, and notably for b-tagging, these track parameters are recomputed with respect to the reconstructed hard-scatter vertex position.

The resolution of a track parameter X can be expressed as a function of track $p_{\rm T}$:

$$\sigma_X(p_{\rm T}) = \sigma_X(\inf)(1 \oplus \frac{p_X}{p_{\rm T}}) \tag{3.2}$$

where $\oplus$ denotes addition in quadrature, $\sigma_X(\inf)$ is the asymptotic resolution when the track $p_{\rm T}$ is an infinite, and p_X is the critical transverse momentum for which the intrinsic, and multiple scattering terms contribute equally.

Figure 3.3 shows the transverse and longitudinal impact parameter resolution as a function of track $p_{\rm T}$. Comparisons has been made to quantify the improvement mainly brought by the IBL. The impact parameter resolution is improved significantly with the IBL by a factor of close to two. Most notably for the longitudinal impact parameter resolution, improvement on the $\sigma_X(\inf)$ in the equation above can be seen in the high $p_{\rm T}$ regions. The improvement factor is about 37%, thanks to the shorter pixel size in z direction as described in section 2.2.2. Due to the same size of the IBL pixels as the other layers, there is almost no difference in the transverse impact parameter resolution in the high $p_{\rm T}$ regions. However, significant improvement is seen in the low $p_{\rm T}$ regions. This is due to the smaller radius for the IBL (factor 1.3 to the previous B-layer), and the gain in material (IBL is about twice lighter in radiation length).

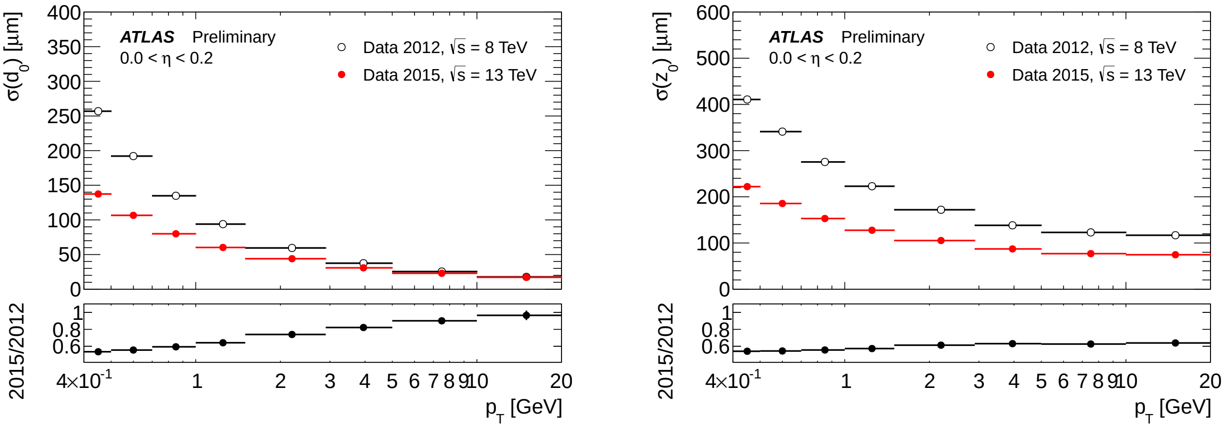


Fig. 3.3: The transverse (left) and longitudinal (right) impact parameter resolutions as a function of p_{T}, comparing them in Run 1 without IBL (black open circles) and in Run 2 with IBL (red points). The ratio of Run 1 data to Run 2 data is shown in the lower panel.

3.1.2. Vertexing

The extrapolation of the tracks allows us to reconstruct the vertices which are corresponding to the interaction point of pp collisions.

Primary vertex

To determine the position of the primary vertex of the hard collision allows to measure the initial parameters of charged particles precisely.

The reconstruction of the primary vertices is performed in two steps. First, the vertex finding algorithm associates track to vertex candidates. Second, the vertex fitting determines the vertex position and their uncertainties. The vertex fitting also involves the refitting of the associated tracks which improves the accuracy of the track parameters [59].

In the vertex finding algorithm, an iterative approach is used where the vertex seeds are found by looking at the local maximum in the distribution of the z coordinates. Vertex fitting is performed by the "adaptive vertex fitting" algorithm which is an extension of the χ^2 based Kalman-filter.

The outlying tracks are down-weighted rather than rejected. Tracks which are incompatible with the vertex by more than 7σ are used to seed a new vertex. Finally, the list of vertices are ordered in $\sum(p_{\mathrm{T}}^{\mathrm{track}})^2$ where $p_{\mathrm{T}}^{\mathrm{track}}$ is the p_{T} of tracks associated to the corresponding vertex. The vertex with the highest $\sum(p_{\mathrm{T}}^{\mathrm{track}})^2$ is chosen as the signal vertex while the others are recognized as pileup vertices. The primary vertex identification efficiency depends on the physics processes: it decreases down to about 99% and 92% for $t\bar{t}$ and $Z \to ee$ processes, respectively at around $\mu \sim 40$.

Unbiased transverse impact parameter

After reconstruction of the vertices, all track parameters are recomputed with respect to the reconstructed primary vertex position, e.g. d_0 with respect to the vertex, d_0', is calculated to obtain the track's distance of closest approach to the primary vertex. Using an approximation of $R_{curv} \gg d_0$ so as to estimate the track to be a line, d_0' is calculated as:

$$d_0' = d_0 - (x_V \sin\phi_0 + y_V \cos\phi_0), \tag{3.3}$$

where x_V and y_V are the x-, and y-position of the reconstructed primary vertex. In stead of this approximation above, full-fledged propagation of the helix is used for the b-tagging.

However, the d_0' value directly calculated above is biased since the vertex position is fitted using the corresponding track. To avoid this bias, the unbiased vertex position, which is fitted without the corresponding track, is estimated by using the track's weight based on the error. Finally, the track's distance of closest approach to the unbiased position of the vertex, d_0^{ub}, is obtained and used in the following.

Pileup vertices

As described in section 2.1.3, at the LHC, multiple collisions happen for each bunch crossing, called pileup interactions. The mean number of collisions per crossing depends on the luminosity, and is shown in fig 3.4. Also, multiple number of interaction vertices are reconstructed in each event and shown in fig3.5.

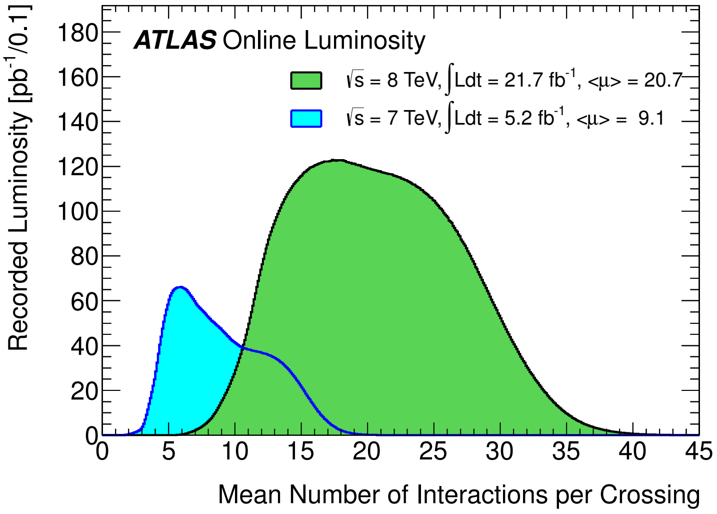


Fig. 3.4: Luminosity-weighted distribution of the mean number of interactions per bunch crossing, in the 2011 and 2012 full data [31].

Figure 3.6 shows a typical high pileup event observed in the 2012 data.

3.2. Electrons and photons

Electrons and photons leave very similar showers in the EM calorimeter. Therefore, the identification strategies for them are deeply connected. This section describes the identification methods of electrons and photons.

3.2.1. Electrons

Each electron leaves a track in the Inner Detector. This allows to have the electron reconstruction seeded by an EM cluster associated to the track of a charged particle in the Inner Detector.

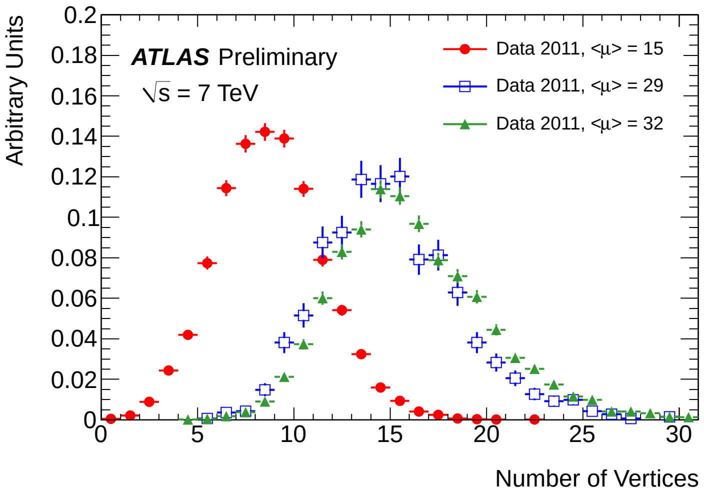


Fig. 3.5: Number of reconstructed vertices in 2011 data for different $<\mu>$ values [58].

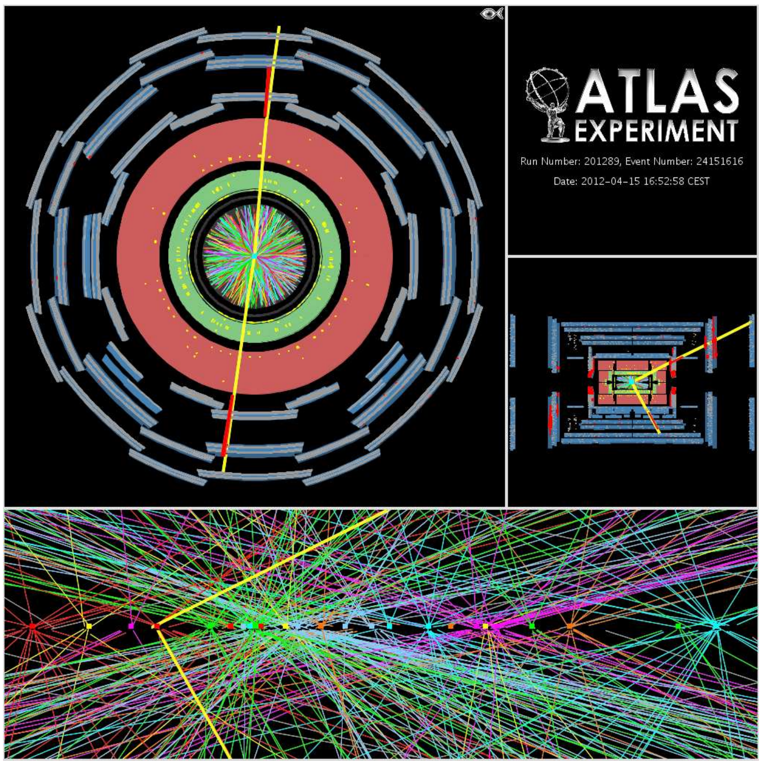


Fig. 3.6: An event candidate of Z boson decaying to dimuon ($Z \to \mu\mu$) with 25 reconstructed vertices. The two muon tracks are indicated with yellow lines [60].

The clustering is performed using the *sliding-window* algorithm [61] in the EM calorimeter. The sliding-window algorithm works in three steps: tower building, pre-cluster (seed) finding, and clustering.

1. Tower building
   First, the entire EM calorimeter is divided into 200×256 squares of 0.025×0.025 in the η-ϕ plane ($|\eta| < 2.5$). Then, contributions from all layers are summed in each square, forming a 'tower'.

2. Sliding-window pre-cluster (seed) finding
   A window of 5×5 (in units of the tower size 0.025×0.025) scans over each element of the tower grid defined above, in steps of $\Delta\eta$ and $\Delta\phi$. When this window scan finds sum of E_{T} inside above a threshold (3 GeV), a pre-cluster is formed. The position of the pre-cluster is determined as the energy-weighted barycenter of all cells in the window, but a smaller window size (3×3) is used to obtain a precise position.

3. EM cluster formation
   At the end, an EM cluster is formed by taking all the cells within the window, whose center is seeded by the pre-cluster's position, but the central position is different in each layer of the calorimeter. The cluster size is optimized for different particle types and shown in table 3.1. In the barrel region the ϕ size of the cluster is enlarged considering the fact that the electron trajectory is curved by the magnetic field, while the same sizes are used in the endcap region where the effect of the magnetic field is smaller.

Table 3.1: Cluster size given in $N_\eta^{\mathrm{cluster}} \times N_\phi^{\mathrm{cluster}}$ for different particle types in the barrel and endcap regions of the EM calorimeter [61].

Particle type	Barrel	endcap
Electron	3×7	5×5
Converted photon	3×7	5×5
Unconverted photon	3×5	5×5

Among tracks matched within $\Delta\eta < 0.5$ and $\Delta\phi < 0.1$ to an EM cluster, the track with the minimum ΔR to the cluster is associated, and this object formed by the cluster and the matched track is registered in the electron container which stores all electron candidates.

However, this matching is not yet perfect to reject the misidentified electrons coming from photon conversion. Such contamination is significantly reduced by the B-Layer hit requirement. If the electron has $p_{\mathrm{T}} > 15$ GeV, the cluster reconstruction is expected to be 100% efficient. Since electrons lose energy with bremsstrahlung, tracking accuracy is lower compared to other heavier particles such as muons or pions. Typically, seven to ten high-threshold hits[b] in TRT are expected for electrons with an energy above 2 GeV.

The electron reconstruction method described here is the standard one and there are also algorithms for soft electrons and forward electrons, for which the track information is not available. However, these are not used in this thesis.

These reconstructed electrons are further selected using shower shape variables in the calorimeter and TRT information in order to reject background (mainly jets). The electrons in this thesis are tagged using both cut-based and likelihood algorithms using 13–18 input variables [32].

---

[b] Hits registering higher signal than 6 keV sampled in Time-over-Threshold in 25 ns time bins.

3.2.2. Photons

Since photons are electrically neutral, they don't leave any hit in the Inner Detector, and thereby are invisible in the tracking volume, unless photon conversion happens. Depending on whether the photon object has a conversion vertex in the Inner Detector or not, they are categorized into *converted* and *unconverted* photons [62]. The former (latter) can be reconstructed as a cluster object with (without) a conversion vertex in the Inner Detector. In addition, it is very important to distinguish photons from neutral hadron decays, such as $\pi^0 \to \gamma\gamma$. In the WH analysis presented in this thesis, photons are not directly used, however, they are used for the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation and can also enter in the game as a background to electrons.

3.3. Muons

Muons are reconstructed with very clear signature and provide excellent momentum resolution unless the momentum is above a few TeV scale. The identification efficiency for muons is very high.

There are four muon categories, which are *Standalone*, *Segment-tagged*, *Calo*, and *Combined* muons:

- Standalone muons
  Muons tracked by only the Muon Spectrometer are called standalone muons. The extrapolation of the tracks to the collision point allows to calculate the initial momenta of the muons. This extrapolation takes into account multiple scattering and energy loss in the calorimeter. The typical muon energy loss in the calorimeter is about $4\ \mathrm{GeV}$.

- Calo-muons
  Muons detected without the Muon Spectrometer are called calo-muons. When the Inner Detector finds a track with a matched calorimeter energy deposit corresponding to a minimum ionizing particle, it is tagged as a calo-muon. It is only used when $|\eta^{\mu}| < 0.1$, due to a hole in Muon Spectrometer coverage.

- Segment-tagged muons
  When an Inner Detector track matches to a track-segment in the Muon Spectrometer, it is tagged as a segment-tagged muon.

- Combined muons
  When a Muon Spectrometer track matches to an Inner Detector track, it is tagged as a combined muon, which is the best quality muon type.

There are two major combined muon reconstruction algorithms in ATLAS: *Staco* and *Muid* algorithms. The Staco algorithm performs an association between the Inner Detector track and standalone track with a χ^2 test, defined by the difference between the respective track parameters weighted by the combined covariance matrices, while the Muid algorithm refits the full track using hits in the Inner Detector and the Muon Spectrometer. The analysis performed in this thesis uses the Muid algorithm for combined muons.

Muons are selected using the following requirement criteria to the Inner Detector tracks [63]:

- At least 1 pixel hit

- At least 5 SCT hits

- At most 2 expected hits: active Pixel or SCT sensors traversed by the track but without hit.

- At least 9 TRT hits when the muon lies in $0.1 < |\eta| < 1.9$.

Since the muons are the dominant constituent for the secondary cosmic rays, cosmic muons can contaminate the collision events. At the ground level, about 70% of the cosmic rays are muons. There are two large shafts over the ATLAS detector, which were used to install the detector components, allowing the cosmic muons to reach the ATLAS detector easily. To suppress such muons originating in cosmic rays, a longitudinal impact parameter($|z_0^\mu|$) selection is normally applied for muons (see section 5.3) in physics analyses.

3.4. Taus

As shown in section 1.1.1, tau is the heaviest lepton, decaying via the weak interaction. Since it has a large enough mass, it can decay not only leptonically but also hadronically. The branching fraction of leptonic decay ($\tau \to \ell\nu\nu, \ell = e, \mu$) is roughly 17.5% for each of electron and muon channels. For hadronic decays, a decay mode with one charged pion, so-called 1-prong, has branching fraction of about 46%, while one with three charged pions, so-called 3-prong, has about 12.5%.

It has a mass of 1.777 GeV and life time of 2.9×10^{-13} s, corresponding to $ct = 87.03$ μm, where t is the lifetime of the tau particle. With $p_{\mathrm{T}}^{\tau} = 50$ GeV, a tau flies over a distance of about 2 mm in the transverse plane. Because of non-negligible flight distance with relatively heavy mass, it can also be background to b-jets.

Identification of leptonic decays of taus is very difficult since the visible component of the final state is exactly the same as the prompt electrons or muons, such as $W \to e\nu$ or $W \to \mu\nu$. In the WH analysis presented in section 5, the leptonic decays of taus in $W \to \tau\nu$ can enter in the analysis regions.

The final state of hadronically decaying tau is very similar to jets, thus is called a tau-jet. A tau-jet is observed as a narrower jet compared to quark- or gluon-jets, since taus decay weakly. For the tau-jet identification, this narrow jet topology is mainly used. A narrow shape calorimetric cluster associated with a small number of tracks seeds the tau finding algorithm using $\Delta R < 0.2$ matching between the cluster and the track(s). The tau reconstruction runs over the calorimeter clusters within $|\eta| < 2.5$ because of the Inner Detector coverage. Although the primary background of tau identification is jets originated in quarks or gluons, electrons can also be a major background, since the finding algorithm is similar to the one used for electron finding [64].

In ATLAS, boosted decision trees (BDT) are trained against several background components (jets and electrons).

When it decays to multi-prong (mainly 3-prong) final state, it leaves similar signature to b- or c-jets due to the secondary vertex presence.

Cluster information for photons is also useful for both 1- and 3-prong tau identification, because the final state of hadronic tau decay often contains π^0, which decays to two photons immediately ($\pi^0 \to \gamma\gamma$).

3.5. Jets

Jets are frequently observed objects in high energy hadron collisions, due to the color confinement. They are collimated sprays of particles originating from fragmentation and hadronization of energetic quarks or gluons.

There are two constrains for jet reconstruction:

- Collinear safety: jets should not be changed by collinear splitting such as gluon emission from a quark.
- Infrared safety: jets should not be changed by soft emission such as very low particles produced in hadronization.

In addition, jet reconstruction algorithms should not depend on the geometry of the detector and should be highly efficient. Based on these requirements, roughly two types of jet reconstruction algorithms are developed: the *cone* algorithms which were mainly used in Tevatron experiments, and *sequential recombination* algorithms mainly used in LHC experiments.

In the ATLAS experiment, the anti-k_t algorithm, which is one of the sequential recombination algorithms, is mainly used. A precise explanation of the anti-k_t algorithm can be found in [65]. Figure 3.7 shows jets clustered by the anti-k_t algorithm, shown in the η-ϕ plane.

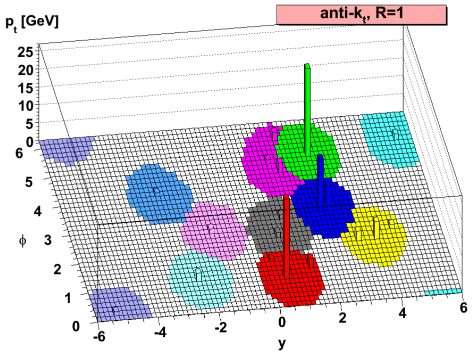


Fig. 3.7: A sample parton-level event, together with many random soft momentum particles (called ghost), clustered with the anti-k_t algorithm [65].

The noise-suppressed topological clusters[c] of energy in the calorimeter are used as inputs for the anti-k_t algorithm [61], and these jets, with a cone size parameter of 0.4, are called AntiKt4EMTopo jets and used as the baseline in ATLAS.

Tracks can also be used as inputs for the jet clustering algorithm, and currently the application of these jets, called track-jets, is being studied in ATLAS. An advantage of track jets is robustness against pileup interactions since tracks can be specifically chosen for an interesting primary vertex.

Pileup jets

Jets are always composed of several particles and therefore can be easily contaminated by particles originated in pileup interactions. Furthermore, jet reconstruction is entirely dependent on the calorimetry, where discrimination of originated vertices is difficult.

In order to identify the origin of jets, tracks associated to jets are exploited to form a quantity called jet vertex fraction (JVF) [66]. For a given jet in an event, JVF is defined as follows:

$$\mathrm{JVF} = \frac{\displaystyle\sum_{\mathrm{tracks}\in\mathrm{jet}\cap\mathrm{PV}} p_{\mathrm{T}}^{\mathrm{track}}}{\displaystyle\sum_{\mathrm{tracks}\in\mathrm{jet}} p_{\mathrm{T}}^{\mathrm{track}}}, \tag{3.4}$$

---

[c] Group of calorimeter cells that are designed to follow the shower development.

where the denominator is the sum of p_T of tracks associated to the jet, and the numerator is the sum of p_T of tracks associated to both the jet and the primary vertex in the event.

The left sketch in fig 3.8 shows an example case of JVF for two vertices and two jets, and the right plot in fig 3.8 shows simulated JVF distributions for jets from hard scatter vertices and pileup vertices.

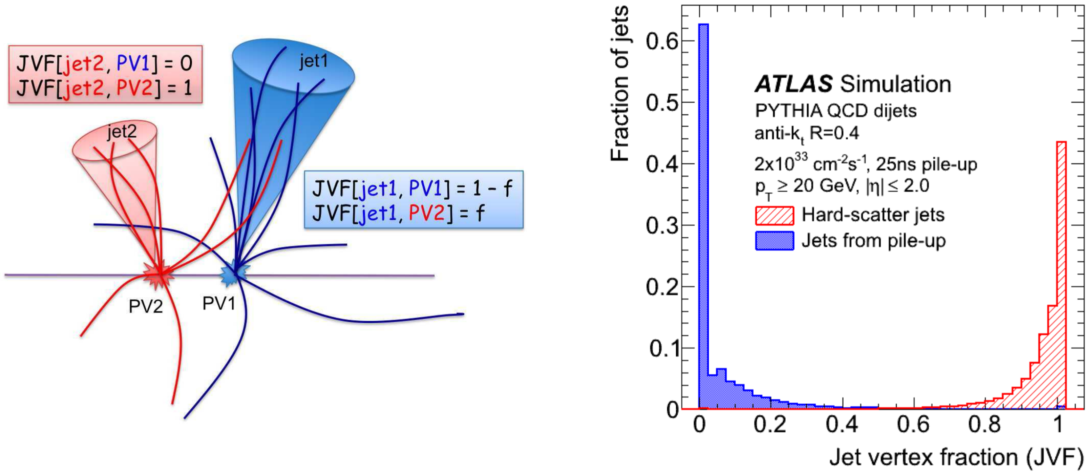


Fig. 3.8: A sketch of the JVF quantities for a case with two jets and two vertices (left). The JVF distributions (right) for the blue histogram for jets originated in pileup vertex and the red histogram for jets originated in hard scattering vertex [66].

In Run 1, JVF worked well for the reduction of pileup-originated jets. However, its robustness against the number of primary vertices was not perfect. In order to recover this weak point of JVF, the jet vertex tagger (JVT) has been developed for Run 2, which is based on the 2D likelihood using two newly developed variables: JVF corrected for the number of reconstructed vertex in the event, and ratio of summed p_T of the tracks originated from the hard-scatter in the jet to $p_\mathrm{T}^\mathrm{jet}$ [67].

b-tagging

Jets originating from b-quark (b-hadron) can be separated from the other jets, and this is particularly important to study the Higgs boson decays to b-quark pairs ($H \to b\bar{b}$) or top quarks ($t \to bW$). The signatures based on the b-hadron lifetime are used for b-tagging. All the details can be found in chapter 4.

3.6. Missing transverse momentum

Protons are composite particles, so that the initial partons' center-of-mass energy along the beam axis cannot be measured on an event-by-event basis. However, the momentum conservation in the transverse plane is useful since the initial partons have negligible transverse momenta in comparison to those along the beam axis. This momentum conservation in the transverse plane provides a quantity called missing transverse momentum, by reversing a vectorial sum of the transverse momenta of all objects in the event. Its magnitude is $E_\mathrm{T}^\mathrm{miss}$.

High energy neutrinos can result in a large $E_\mathrm{T}^\mathrm{miss}$, since they usually do not interact with any of the material in the ATLAS detector. Hence, $E_\mathrm{T}^\mathrm{miss}$ measurement is a very important key for a lot of physics analyses, such as analyses involving $W \to \ell\nu$ and $Z \to \nu\nu$. If a certain event has one neutrino, ideally $E_{\mathrm{T}}^{\mathrm{miss}}$ corresponds to the neutrino's p_{T}. However, if the number of neutrinos is more than one, it is not possible to obtain the transverse momenta carried by individual neutrinos without further assumptions.

$E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed using energy deposits in the EM and hadronic calorimeters, with corrections for momenta of individual objects, such as photons, electrons, muons, tau-jets, and jets.

In addition to the energy deposited by particles arising from the hard scattering, there are many other sources of $E_{\mathrm{T}}^{\mathrm{miss}}$, such as noise in the detector, the underlying event, the pileup interactions, and the jet energy resolution. If high p_{T} objects pass through dead modules or service parts in the calorimeter, they can also affect the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement.

In order to achieve precise $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement, it is very important to calculate the energy deposits as precisely as possible in the calorimeters with their coverage as large as possible. Hence, the ATLAS calorimeter has a good hermeticity ($|\eta| < 4.9$). The calibration and noise suppression improves the energy measurement using calorimeters. All physics analyses involving $E_{\mathrm{T}}^{\mathrm{miss}}$, therefore, have to use data taken under the best condition of all the sub-detectors, using GRL described in section 2.1.

Another similar quantity is a track-based $E_{\mathrm{T}}^{\mathrm{miss}}$, with its magnitude expressed as $p_{\mathrm{T}}^{\mathrm{miss}}$, which is calculated using the tracks reconstructed in the Inner Detector [68].

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

In ATLAS, there are two main $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction algorithms: cell-based and object-based algorithms. In the analysis presented in chapter 5, the latter is used.

The concept of the object-based algorithm is to use reconstructed and calibrated object for $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The calorimeter cells are associated with reconstructed physics objects, such as electrons, photons, tau-jets, jets, and muons. In this order, the overlap removal is performed. Cells not associated with any physics object are also taken into account and their $x(y)$ component of $E_{\mathrm{T}}^{\mathrm{miss}}$ is expressed as $E_{x(y)}^{\mathrm{miss(CellOut)}}$. Energies of all these physics objects listed above are corrected and calibrated: all dedicated correction and calibration techniques used for the analysis are applied for these physics objects and those are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation.

The cell calibration is replaced by each dedicated object calibration mentioned above, which generally has higher accuracy. This calculation is called *refined calibration* [69]. This calibration also includes pileup suppression in the $E_{x(y)}^{\mathrm{miss(CellOut)}}$ term using tracks or the jet area method and in the $E_{x(y)}^{\mathrm{miss(jet)}}$ term using tracks.

After the corrections and calibrations mentioned above, the $x(y)$ component of $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated by:

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss}(e)} + E_{x(y)}^{\mathrm{miss}(\gamma)} + E_{x(y)}^{\mathrm{miss}(\tau)} + E_{x(y)}^{\mathrm{miss(jet)}} + E_{x(y)}^{\mathrm{miss(Softjet)}} + E_{x(y)}^{\mathrm{miss(CellOut)}} + E_{x(y)}^{\mathrm{miss}(\mu)},$$

where $E_{x(y)}^{\mathrm{miss}(e)}$, $E_{x(y)}^{\mathrm{miss}(\gamma)}$, and $E_{x(y)}^{\mathrm{miss}(\tau)}$ are the terms for electrons, photons, and tau-jets, respectively. The $E_{x(y)}^{\mathrm{miss(jet)}}$ $\left(E_{x(y)}^{\mathrm{miss(SoftJet)}}\right)$ is the term for jets with $p_{\mathrm{T}} > 20$ GeV ($7 < p_{\mathrm{T}} < 20$ GeV). In order to avoid double counting, the negative sum of reconstructed muon objects, $E_{x(y)}^{\mathrm{miss}(\mu)}$, is made with a subtraction of the muon energy loss in the calorimeter [70].

4.Study of b-jet identification

The identification technique of jets originating in b-quarks, so-called b-tagging, is a very crucial ingredient for many physics programs at the LHC, such as the $H \rightarrow b\bar{b}$ search, the top quark studies, or the searches for new physics. This chapter explains the methodology of the b-tagging in section 4.1, describes briefly its main ingredients in section 4.2, and reviews the existing basic b-tagging algorithms in the ATLAS experiment [71] in section 4.3. Explanations of multivariate-analysis methods (MVA) are presented before a review of existing MVA taggers in section 4.4. New algorithm developments are presented in section 4.5 which combine basic algorithms' outputs using Boosted Decision Trees (BDT) and provide the best b-tagging performance in ATLAS. The performance of various algorithms in Run 1 and Run 2 is compared in section 4.6. Section 4.7 briefly discuss data-to-MC calibration of b-tagging in ATLAS, and summary is given in section 4.8.

Notations

The term "b-jet" is used for a jet stemming from a b-quark fragmentation. In the hadronization process, the b-quark forms a hadron which is called b-hadron such as $B^{\pm}$, B^0, etc. Likewise, jets originated from c-quarks (c-hadrons) are called c-jets. The jets originated from gluons, and u-, d-, or s-quarks are called light jets. The jets produced by pileup interactions are also categorized as light jets.

When tagging, a selection cut is applied on the tagging output distribution and jets passing the cut are defined as tagged jets. The b-jet efficiency, ϵ_b, is therefore defined as the fraction of tagged jets out of the true b-jets ($\epsilon_b = N_b^{\text{tagged}}/N_b^{\text{true}}$), and c- and light jet efficiencies are also defined in the same way. The selection cut defines the so-called *"working point"*, and these working points are referred to in ATLAS using the b-jet efficiency measured in the $t\bar{t}$ simulation sample. The $\epsilon_b = 70\%$ working point is typically used in many physics analyses in ATLAS. The cut defining each working point is also called as a *global cut* since the target efficiency there is determined inclusively in $p_{\text{T}}^{\text{jet}}$ or η^{jet}. Therefore, the efficiency obtained with the global cut is called *global efficiency*.

There are another way to define a b-tagging selection using a set of cuts defined in each $p_{\text{T}}^{\text{jet}}$ bin: these $p_{\text{T}}^{\text{jet}}$-dependent cuts are called *flat efficiency* cuts and they are defined in such a way that the signal efficiency in each $p_{\text{T}}^{\text{jet}}$ bin is the same as the target efficiency. The flat efficiency cuts are not only used for $p_{\text{T}}^{\text{jet}}$ but also other properties of events.

The light jet rejection factor, R_{light}, is defined as the inverse of the light jet efficiency, ϵ_{light}: $R_{\text{light}} = 1/\epsilon_{\text{light}}$. The discrimination power of the classifier is expressed by the receiver operating characteristic (ROC) curve, which is made by sliding the cut value on the discriminant variable, each cut corresponding to a set of ϵ_b and $R_{\text{bkg}} = R_l, R_c$.

4.1. Characteristics of b-jets

The b-jets have very distinguishable characteristics explained below. The b-tagging is performed using the features of the b-, c-, and light jets.

The b-jet characteristics are caused by the b-hadron's mass, lifetime, and if should be the case its semi-leptonic decay. The b-jet has relatively high track multiplicity originating in the high mass of the b-hadron ($m_B \sim 5$ GeV). The b-hadron's lifetime, $c\tau \sim 450$ μm, is long enough to be observed as a displaced vertex in the detector ($\beta\gamma c\tau \sim 1$ mm when the b-hadron momentum is around 10 GeV). Once the secondary vertex is reconstructed, useful information will be available, such as its mass, track multiplicity, flight direction, and decay length of the secondary vertex. Even when the secondary vertex reconstruction is not successful, the b-hadron lifetime information is inherited by the tracks which are displaced from the primary vertex due to the secondary or tertiary decays. The b-hadrons tend to decay sequentially to c-hadrons and then to hadrons with strangeness. A schematic view of the cascade decay of the b-hadron is shown in fig 4.1. This cascade decay provides distinct characteristics which is fully exploited by the *Jet Fitter* algorithm described in section 4.3.3. The semi-leptonic decay of the b-, or c-hadron also supplies very useful information to identify b-jets. However, this is statistically limited because of the branching ratio which is about 20 % for each lepton flavor (e, μ), including $b \to c$ cascade decays.

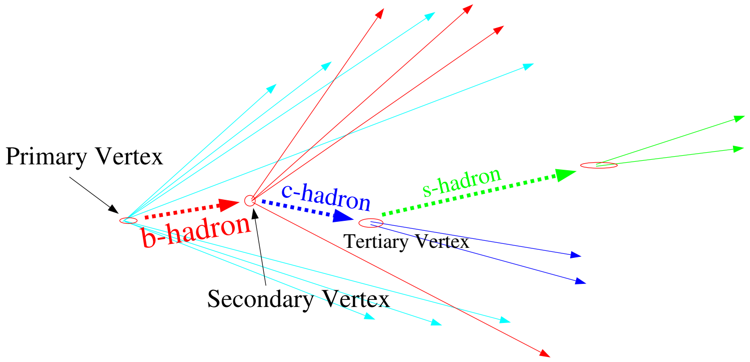


Fig. 4.1: A schematic view of the cascade decay of the b-hadron [72].

Although the c-jets are very similar to the b-jets, the features are weaker because of the relatively short lifetime (e.g. $c\tau_{D^{\pm}} \sim 320\mu$m) and low mass (e.g. $m_D \sim 1.9$ GeV) of c-hadrons, in comparison with b-hadrons.

The light jets have opposite features, in comparison with the b-jets, such as lower track multiplicity, less displaced tracks from the primary vertex, and a lower mass at the secondary vertex, if any is found.

4.2. Ingredients for b-tagging

Currently in the ATLAS experiment, b-tagging is performed using the lifetime features of the b-hadrons, but a method based on the semi-leptonic decays (so-called soft lepton tagging) is also used for specific purposes (calibration) and is studied for a more general use. Before reviewing the basic algorithms in the next section, this section describes the ingredients for b-tagging, such as primary vertex, track-to-jet association, and the track selections.

4.2.1. Jets and primary vertex

Jets are reconstructed by anti-k_{t} algorithm using inputs of topological clusters of energy in the calorimeter [61], with the cone size parameter $R = 0.4$ as described in section 3.5. Jets are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$, due to the jet calibration procedure threshold and the tracker's coverage, respectively. The primary vertex used for b-tagging is reconstructed and selected as described in the previous chapter in section 3.1.2.

4.2.2. Truth labeling

Jets are labeled using truth particles registered in the simulated samples. To match the reconstructed jet and the truth particle, $\Delta R = 0.3$ is used. The truth particles used in this chapter are partons with $p_{\mathrm{T}} > 5$ GeV. When the jet is matched to a b-quark, it is labeled as b-flavor. If not matched to b-quark but c-quark, it is labeled as a c-flavor. If not matched to both b- and c- quarks but tau, it is labeled as tau. If not matched to any of b-, c-quarks, or tau, it is labeled as light flavor. Labeling using the b- and c-hadrons instead of b- and c-quarks was recently tested and found to be slightly better, however it is not used in this chapter because of availability.

4.2.3. Track-to-jet association

Jets in ATLAS are first reconstructed using calorimeter information as explained in section 3.5. Then, tracks associated to the jets are going to be inputs for the b-tagging algorithms.

The track-to-jet association used to be done using $\Delta R(\mathrm{track}, \mathrm{jet}) < 0.4$. However, the fixed ΔR scheme is not an optimal association because the high p_{T} jets have very collimated tracks compared to the low p_{T} jets. This effect changes the track's distance from the jet axis depending on the jet p_{T}, resulting in that the low p_{T} jets lose the signal tracks and the high p_{T} jets are contaminated by background tracks such as track from pileup interactions.

Track-to-jet association was optimized by introducing variable ΔR cone size depending on the jet p_{T} with the dependence expressed by the following equation:

$$\Delta R(p_{\mathrm{T}}) = a_0 e^{a_1 + a_2 p_{\mathrm{T}}} \tag{4.1}$$

where a_0, a_1, and a_2 are optimized coefficients, $a_0 = 0.239$, $a_1 = -1.22$, and $a_2 = -1.64 \cdot 10^{-5}$, so as to collect on average 95% of the b-hadron decay products [73]. Figure 4.2 shows the ΔR cone size as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$. The cone size is 0.45 (0.26) at $p_{\mathrm{T}}^{\mathrm{jet}} = 20(150)$ GeV.

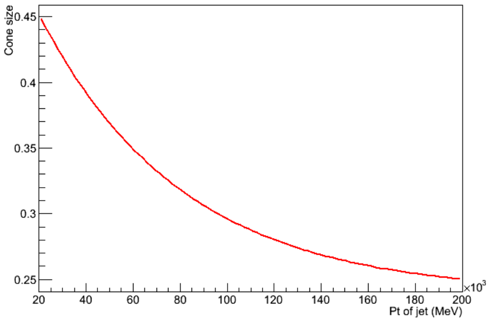


Fig. 4.2: The track-to-jet association in b-tagging with the variable ΔR cone size as a function of the jet p_{T} [74].

4.2.4. Track selection

After the tracks are associated to the jets, quality cuts are applied for the tracks in order to reject irrelevant tracks stemming from pileup interactions or poorly reconstructed tracks and to have better separation between b- and light jets.

The optimal quality cuts are different for each algorithm. Table 4.1 shows typical examples of quality cuts for three main algorithms. These cuts are also useful to remove the tracks stemming from long-lived particles such as K_S ($c\tau = 2.68$ cm) and Λ ($c\tau = 7.89$ cm), from photon conversion, and from the hadronic material interactions.

Table 4.1: The typical track quality selection for three main algorithms of Impact Parameter (IP), secondary vertex (SV), and Jet Fitter algorithm. The variables, d_0 and z_0, are the ones defined in section 3.1.1.

Criteria	IP	SV	Jet Fitter
p_{T} [GeV]	>1.0	>0.7	>0.5
$\lvert d_0\rvert$ [mm]	<1.0	<5.0	<7.0
$\lvert z_0 - z_{\mathrm{PV}}\rvert\sin\theta$ [mm]	<1.5	<25	<10
number of B-Layer hits	≥ 1	≥ 0	≥ 0
number of Pixel hits	≥ 2	≥ 1	≥ 1
number of SCT hits	≥ 0	≥ 4	≥ 4
number of silicon hits	≥ 7	≥ 7	≥ 7
number of shared hits	–	≤ 1	–

4.3. Basic b-tagging algorithms in the ATLAS experiment

In this section, the b-tagging algorithms based on the tracks, secondary vertex, and the decay chain fit, are reviewed in sections 4.3.1–4.3.3. They provide the input information for the downstream MVA taggers.

In addition to the existing algorithms, the multi-secondary vertexing approach, which aims at the reconstruction of the decay vertices of multiple number of b-hadrons in a jet (e.g. $g \to b\bar{b}$), is also under development.

4.3.1. Track-based algorithm: IP tagger

Several b-tagging algorithms based on the track impact parameter (IP) were developed in the ATLAS experiment. It is particularly important because of the limited secondary vertex reconstruction efficiency. Depending on the type of the track impact parameter used to build the tagger, two variations of taggers are constructed.

- **IP2D tagger**: a likelihood-based tagger using the transverse impact parameter, d_0.
- **IP3D tagger**: a likelihood-based tagger based on the 2D correlation between d_0 and z_0.

Figure 4.3 sketches the definition of the impact parameter (d_0) and the decay length (L_{xy}). The discrimination is based on the fact that the tracks in the b-jet tend to have larger d_0 due to the secondary or tertiary decays. Although the absolute value of d_0 is the unbiased one defined in the section 3.1, the sign of d_0 is differently defined in b-tagging, given by the inner product of two vectors shown in the equation below,

$$\mathrm{sign}(d_0) = \mathrm{sign}(\vec{d_0} \cdot \vec{j}_{xy}) \tag{4.2}$$

where $\vec{d_0}$ is a vector from the primary vertex to the point which defines d_0 on the track, and $\vec{j}_{xy}$ is a vector of the jet axis on the transverse plane. This is to take advantage of the lifetime information

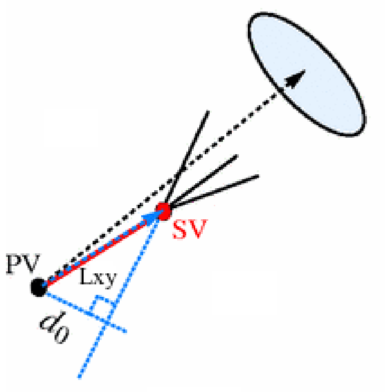


Fig. 4.3: A sketch of the definition of the transverse impact parameter, d_0, and the transverse decay length, L_{xy} [75]. There are primary vertex (PV), secondary vertex (SV), and jet axis shown with the black filled circle, red filled circle, and the black dotted arrow, respectively. As explained in section 3.1, d_0 is the track's closest approach to the primary vertex, but its sign is differently assigned for b-tagging (for the details, see text). The distance between the primary vertex and the secondary vertex, L_{xy}, is also signed.

reflected in the sign of d_0 since the tracks stemming from the secondary decay of the b-hadron tend to distribute more on positive side with respect to the jet-axis while the tracks stemming from the light jet tend to equally populate positive and negative sides.

In order to take advantage of the correlation between the track impact parameter and its error and to give high weight to well-measured tracks, the significance of the impact parameter is used to construct the likelihood probability for three hypotheses of b-, c-, and light jets. The transverse and longitudinal impact parameter significance are defined as ratios of $d_0/\sigma(d_0)$ and $z_0/\sigma(z_0)$, respectively, where $\sigma(x)$ is the error of the x measurement with $x = d_0, z_0$. The track-level likelihood probability of being a track in the b-jet, p_b^{track}, is given by the equation below,

$$p_b^{\text{track}}(d_0/\sigma(d_0)) = \frac{\mathcal{P}_b(d_0/\sigma(d_0))}{\mathcal{P}_b(d_0/\sigma(d_0)) + \mathcal{P}_c(d_0/\sigma(d_0)) + \mathcal{P}_u(d_0/\sigma(d_0))} \tag{4.3}$$

where $\mathcal{P}_b$, $\mathcal{P}_c$, and $\mathcal{P}_u$ are the probability density functions (PDF) of the tracks in b-, c-, and light jets. The track level probabilities of being c- and light jets are also calculated as in the b-jet case. The reference PDFs are constructed using $t\bar{t}$ samples for all flavor jets, which can be seen in fig 4.4 and 4.5 for IP2D and IP3D taggers, respectively.

Using the probabilities defined above, the log-likelihood-ratio is taken for all the tracks, and combined into the jet level log-likelihood-ratio (w_{jet}) using the equation below:

$$w_{\text{jet}} = \sum_{\text{track}} \log(\frac{p_b^{\text{track}}}{p_u^{\text{track}}}) \tag{4.4}$$

To maximize the separation, the reference PDFs are independently constructed in the several track qualities, which are called grades. The grades are defined using the quality information of the tracks in the detector layers: the missing hit while expected, the shared hit where it is shared by at least two tracks, and the split hit where it is split due to the overlapping tracks. In Run 1, six exclusive grades were considered while the number of grades were re-optimized for 14 in Run 2 configuration [76],

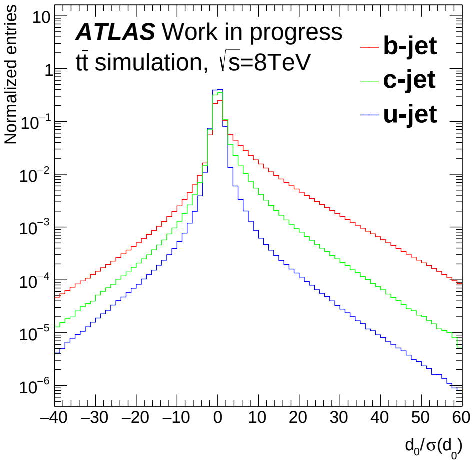


Fig. 4.4: The reference PDFs, constructed using $t\bar{t}$ sample, for IP2D tagger for the good grade tracks. The b-, c-, and light jet distributions are shown with red, green, and blue, respectively.

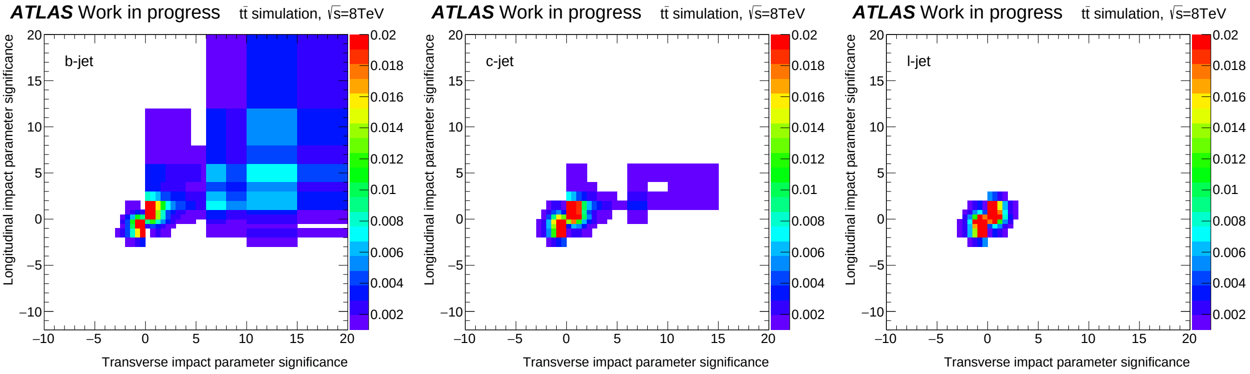


Fig. 4.5: The reference PDFs for IP3D tagger, , for the good grade tracks. The b-, c-, and light jet tracks are in the left, middle, and right plots, respectively.

in order to take advantage of the IBL insertion. These two grade strategies in Run 1 and Run 2 are shown in table 4.2. Most of the b-tagging tracks (∼85%) belongs to the best quality grade, good.

Table 4.2: List of exclusive track grades exploited for IP2D and IP3D tracks in Run 1 and Run 2 conditions. In Run 2, the L0 and L1 refer to the first pixel layer (IBL) and the second pixel layer (formerly called B-layer in Run 1). As it goes down, the track quality improves.

#	Description
	Run 1
0	No hit in B-layer; while expected
1	More than one shared hits
2	A shared hit in B-layer
3	A shared hit in one of two outer Pixel layers
4	A shared hit in one of SCT layers
5	Good. A track not falling in any of categories above
	Run 2
0	No hit in the fist two layers; expected hit in L0 and L1
1	No hit in the fist two layers; hit in L0 and no expected hit in L1
2	No hit in the fist two layers; no expected hit in L0 and expected hit in L1
3	No hit in the fist two layers; no expected hit in L0 and L1
4	No hit in L0; expected hit in L0
5	No hit in L0; no expected hit in L0
6	No hit in L1; expected hit in L1
7	No hit in L1; no expected hit in L1
8	Shared hits in both L0 and L1
9	A shared hit in one of two outer Pixel Layers
10	Two or more shared SCT hits
11	Split hits in both L0 and L1
12	A split hit in one of outer two Pixel layers
13	Good. a track not falling in any of categories above

Figure 4.6 shows the output weight distributions for IP2D and IP3D taggers.

4.3.2. Secondary vertex based algorithm: SV tagger

Two algorithms based on the secondary vertex (SV) properties were developed, which are a simple one and sophisticated one, called SV0 and SV1 taggers, respectively. Both taggers rely on an *inclusive secondary vertex finder* which tries to reconstruct a single vertex from secondary b-hadron and tertiary c-hadron decays.

The SV0 tagger is a very simple tagging algorithm which takes the 3D decay length significance ($L_{xyz}/\sigma(L_{xyz})$) as its discriminant. Similarly to the d_0, the sign of L_{xyz} is defined by the inner product of two vectors shown in the equation below:

$$\mathrm{sign}(L_{xyz}) = \mathrm{sign}(\overrightarrow{SV}_{xyz} \cdot \vec{j}_{xyz}) \tag{4.5}$$

where $\overrightarrow{SV}_{xyz}$ is a vector pointing the secondary vertex from the primary vertex, and $\vec{j}_{xyz}$ is the 3D momentum vector of the jet axis. The xy-projection of $\overrightarrow{SV}_{xyz}$, $\overrightarrow{SV}_{xy}$, is shown with a red dotted

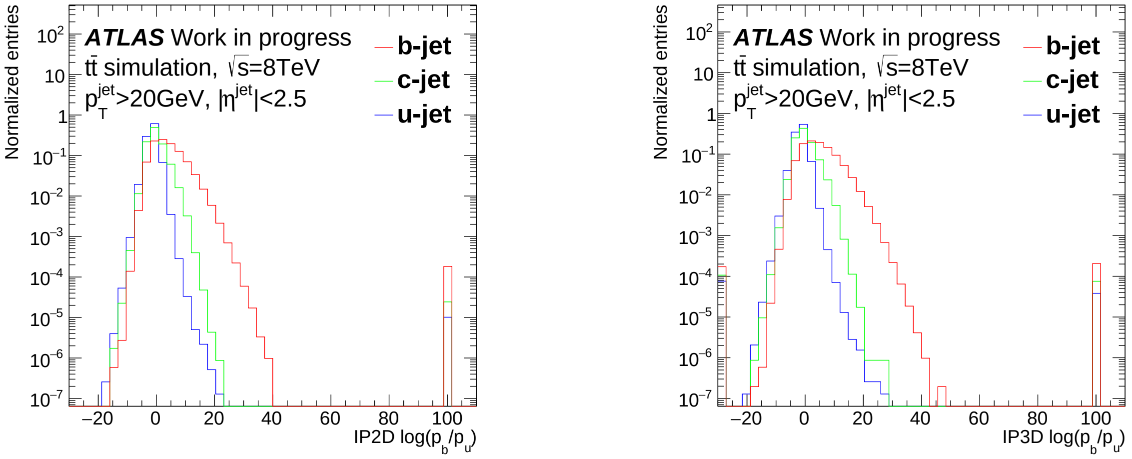


Fig. 4.6: The output jet weight distributions of IP2D (left) and IP3D (right) taggers. The b-, c-, and light jet tracks are shown with red, green, and blue lines, respectively.

arrow in fig 4.3. As well as the L_{xyz}, the sign of L_{xy} is also defined by the 2D jet-axis vector, $\overrightarrow{j}_{xy}$, and the 2D vector pointing the secondary vertex from the primary vertex, $\overrightarrow{SV}_{xy}$.

The SV1 tagger is a likelihood tagger based on the secondary vertex properties. The likelihood probability is constructed using the four variables below, with three PDFs:

- N_{2tracks}: the number of two-track pairs that can form a vertex.
- $\Delta R(\overrightarrow{SV}_{xyz}, \overrightarrow{j}_{xyz})$: the angular distance between the secondary vertex and the jet-axis direction.
- m_{SV}: the invariant mass of the tracks associated to the secondary vertex.
- f_{SV}: the fraction of the track momentum sum at the secondary vertex to the track momentum sum of the jet.

The first two variables are shown in fig 4.7, and the last two variables are combined together in the 2D-PDF shown in fig 4.8. Contrarily to the IP taggers which are constructed by single PDF, the SV1 likelihood is constructed from several PDFs (variables). The b-jet likelihood probability is given by the equation below,

$$p_b = \frac{\mathcal{L}_b}{\mathcal{L}_b + \mathcal{L}_c + \mathcal{L}_u}, \tag{4.6}$$

where,

$$\mathcal{L}_b = \prod_{k=1}^{n_{\text{var}}} \mathcal{P}_b^k(x_k),\ \mathcal{L}_c = \prod_{k=1}^{n_{\text{var}}} \mathcal{P}_c^k(x_k),\ \mathcal{L}_u = \prod_{k=1}^{n_{\text{var}}} \mathcal{P}_u^k(x_k) \tag{4.7}$$

and $\mathcal{P}_b^k(x_k)$, $\mathcal{P}_c^k(x_k)$, and $\mathcal{P}_u^k(x_k)$ are the PDFs for the kth input for b-, c-, and light jets. The c- and light jet probabilities are also calculated in the same manner as the b-jet case. In the current use, $n_{\text{var}} = 3$ in the SV1 tagger because of the 2D PDF, which is made of secondary vertex mass and the energy fraction, and both are transformed as $m_{\text{SV}}/(1 + m_{\text{SV}})$ and $(f_{\text{SV}})^{0.7}$, where m_{SV} is given in GeV unit. The output jet weight by log-likelihood-ratio, w, is given by, $w = \log(p_b/p_u)$, and the distributions for all three flavors are shown in fig 4.9.

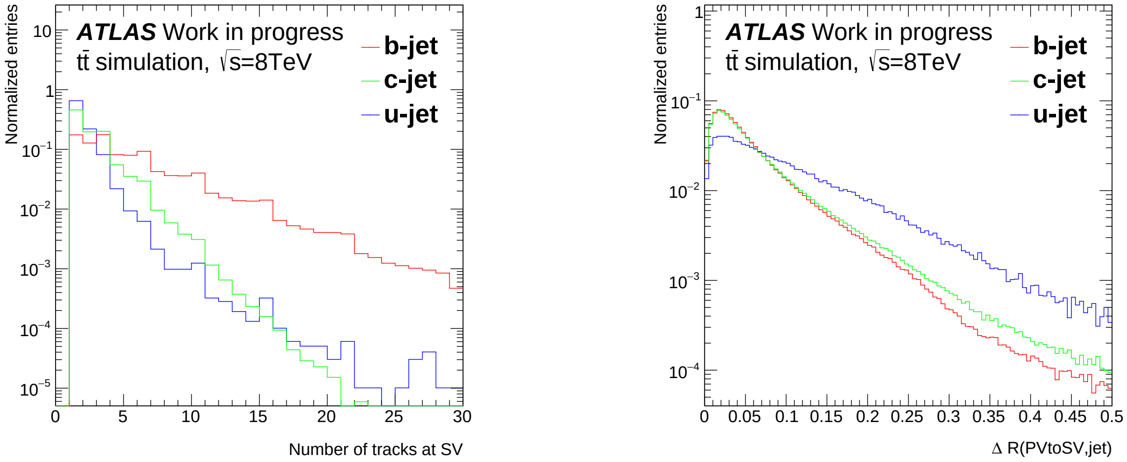


Fig. 4.7: The reference PDFs, constructed using $t\bar{t}$ sample, for SV1 tagger for the number of two track vertices (left), and $\Delta R(\vec{SV}_{xyz}, \vec{j}_{xyz})$ (right). The b-, c-, and light jet tracks are shown with red, green, and blue, respectively.

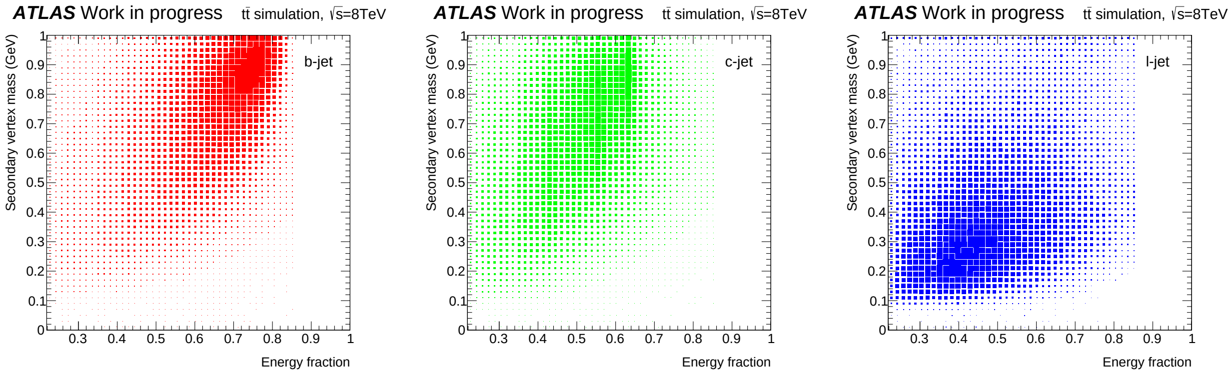


Fig. 4.8: The reference PDFs, constructed using $t\bar{t}$ sample, for SV1 tagger for the 2D correlation of the secondary vertex energy fraction and the secondary vertex mass, both transformed (see text for details). The b-, c-, and light jet tracks are shown with red, green, and blue, in the left, middle, and right plot, respectively. The size of the boxes represents the population density of the jets.

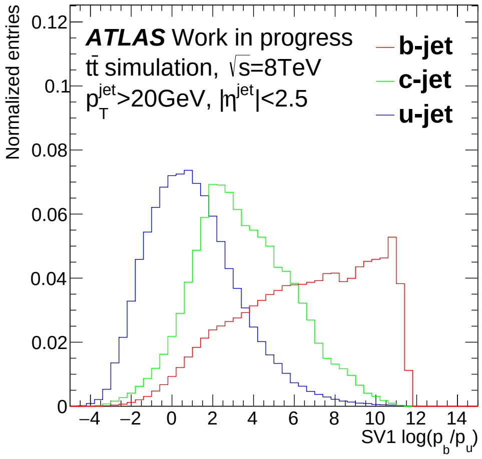


Fig. 4.9: The SV1 output weight distribution. The b-, c-, and light jet tracks are shown with red, green, and blue lines, respectively.

4.3.3. Decay chain fitting algorithm: Jet Fitter

The most sophisticated reconstruction of the b-hadron decay is performed by the Jet Fitter algorithm [77]. Compared to the inclusive secondary vertexing algorithm, the Jet Fitter exploits the topological structure of the weak cascade decay of b- and c-hadrons inside a jet. An assumption has been taken such that the b-, and c-hadron decay vertices lie on the same line as the b-hadron's flight path. A Kalman filter is used to find such a line, and the JetFitter algorithms attempts to fit the cascade decay chain of the b-hadron to c-hadron to etc. Figure 4.10 shows the conceptual difference between the inclusive secondary vertex decay fit and cascade decay chain fit. This algorithm is particularly powerful in the following points:

Fig. 4.10: The inclusive secondary vertexing fits all the displaced tracks to an inclusive vertex (left). The Jet Fitter algorithm performs a multiple number of vertices fit using the b-hadron flight direction as constrained as the jet-direction (right) [77].

- It can reconstruct the incomplete topology under the condition of limited tracking and secondary vertexing efficiency, since the decay producing only one charged track, so-called 1-track vertex, can be reconstructed.
- Evaluating the compatibility of the given set of tracks with the b-c-hadron like cascade topology enables to achieve a powerful separation against light jets.

- The constraint on the b-hadron flight direction reduces the degrees of freedom of the fit, thus, it increases the separation between b- and c-hadron decays: ideally, a b-jet has two displaced vertices inside while a c-jet has one vertex inside.

The hypothesis made here matches the particle kinematics in the b-jets due to the hard b-quark fragmentation function. The lateral displacement of the c-hadron decay vertex from the b-hadron flight path is very close to be discriminated under the limited tracking resolution.

This algorithm's outputs were merged by neural network originally (see section 4.4.3), now combined by boosted decision trees by MV2 algorithm (see section 4.5).

4.4. Multivariate classifiers and taggers

Before describing the MVA taggers, the two major MVA approaches are briefly explained in this section, which are the neural networks and the boosted decision trees.

4.4.1. Neural networks

Artificial neural networks (NN or ANN) is one of the multi-variate analysis techniques based on *machine learning*. Most NNs have a structure called multi-layer perceptron (MLP), which is made of input, hidden, and output layers, as shown in fig 4.11. The NN is composed of neurons, which are called also nodes, and connections in between them. All the neurons are interpreted as a function which gives back one output from a given set of input variables. The weights of connections are optimized by the so-called back propagation algorithm. The output is given by applying several weights to the input variables, so that normally the application speed of the NN is faster than BDT. The neuron function, ρ, can be expressed by two functions of κ and α: $\rho = \alpha \circ \kappa$, where κ and α are called *synapse* and *activation* functions, and usually sigmoid function is chosen for α ($\alpha(x) = 1/(1+e^{-x})$), and κ for the jth node in lth layer is selected as, $\kappa(\mathbf{y}^l_n, \mathbf{w}^l_{nj}) = w^l_{0j} + \sum_{i=1}^{n} y^l_i w^l_{ij}$ where $\mathbf{y}^l_n$ and $\mathbf{w}^l_{nj}$ are the input tuples and weights composed of n components, and w^l_{0j} comes from the bias node. The bias node

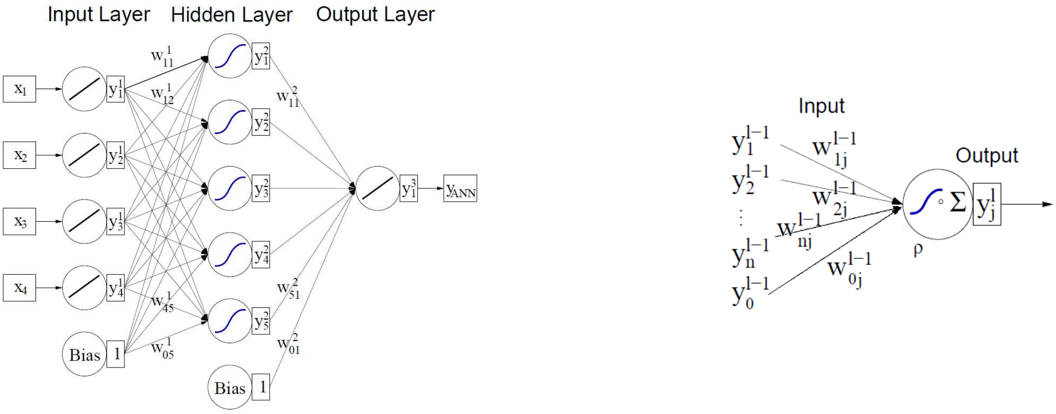


Fig. 4.11: An example of MLP with one hidden layer (left). Single neuron j in layer l with n input connections. The incoming connections carry a weight of $w_{ij}^{(l-1)}$ [78].

always emits 1 so that it is directly added to the argument of the activation function. This allows to control the behavior of the layer by shifting the response of the activation function output.

4.4.2. Boosted decision trees

Boosted decision trees (BDT) are a set of binary structured decision trees, which is called forest, using the *boosting* technique.

Decision trees

The decision trees are machine learning technique developed first in data mining and pattern recognition in various fields. The first formalization was done by Breiman *et al.* [79], proposing the “Classification And Regression Trees” algorithm. The basic principle is to extend a simple cut-based analysis into an MVA by continuing to split events until they satisfy the terminating condition (e.g. number of events in a leaf node). Conceptually, a decision tree does not immediately give up the events without easy feature to recognize, but instead it tries to find some other features which may help the discrimination.

The decision tree structure is shown in fig 4.12. It is composed of the internal splitting nodes and the end-nodes, so-called leaves. Starting from the root node, a sequence of binary splits using the discriminating variable x_i is applied to the data, repetitively splitting the events by the cuts with variables, x_j and x_k, until the stop condition is satisfied.

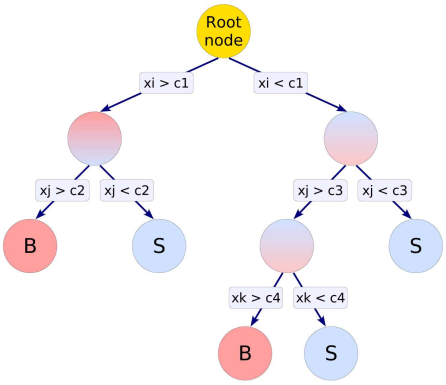


Fig. 4.12: Schematic view of a decision tree with a depth 3 [78].

Discriminating the signals from backgrounds, each split uses the best separating variable and its value at the node. During the training, the cut optimization is done by selecting the best set of cut variable and its cut value after scanning all the possible sets of variables and their values, for the events at each node: the best split, S^*, between signal and background is obtained by choosing from all the splits, S, which maximizes the decrease of impurity, Δi,

$$\Delta i(S^*) = \max_{S \in \text{splits}} \Delta i(S), \tag{4.8}$$

where $\Delta i(S)$ is defined as,

$$\Delta i(S) = i - \min[p_{\mathrm{P}} \cdot i_{\mathrm{P}}, p_{\mathrm{F}} \cdot i_{\mathrm{F}}], \tag{4.9}$$

and p_{P} (p_{F}) is the fraction of events passing (failing) the split S. The impurity is typically defined as,

$i = p \cdot (1 - p)$, called *Gini-index*, where p is the signal purity defined as $s/(s+b)$, in which s (b) is the weighted number of signal (background) events.

When optimizing a cut value on a variable over its full range, the number of intervals to evaluate the cuts is called *nCuts* and it is 200 in actual use for the MV2 development. The decision trees have the levels (so-called depth) of splits depending on how many cuts are applied to reach the leaf nodes at the deepest level. The minimum population required for each leaf node is called minimum node size and expressed by the fraction of the training sample.

The same variable may be used at several nodes, which allows the decision tree to find intervals of interest. It is also possible that some particular variable(s) is(are) not used at all: this feature makes the decision trees robust for the addition of the less discriminating variables.

A notable feature of the decision trees is human-readability: any event can be tracked down from the root node to the leaf node by looking at each cut made at each splitting node. Therefore, a tree can be interpreted in terms of physics, defining selection rules, rather than only of a mathematical object [80].

The leaf nodes of the tree are labeled as S for signal and B for background depending on the majority of events that end up in the respective nodes. The output of the ith decision tree, $p_i(\mathbf{x})$, is defined as the end-node's purity given by

$$p_i(\mathbf{x}) = \frac{s_i(\mathbf{x})}{s_i(\mathbf{x}) + b_i(\mathbf{x})}, \tag{4.10}$$

where $\mathbf{x}$ is the tuple of the input variables, and s_i (b_i) is the weighted number of signal (background) events in the ith node found in the training process. The events with signal-like signature will give values close to 1, while the events with background-like signature will give values close to 0.

Averaging is a helpful method for combining the several decision trees. It was firstly introduced as the so-called V-fold cross-validation [79]. The training sample, $\mathcal{L}$, is divided into V subsets of equal size ($\mathcal{L} = \{\mathcal{L}_1, ..., \mathcal{L}_V\}$), then a decision tree, T_v, is trained on a set of sub-sample, $\mathcal{L} - \mathcal{L}_v$, and tested on the sub-sample $\mathcal{L}_v$. These V decision trees' outputs are combined by

$$\frac{1}{V}\sum_{v=1}^{V} T_v. \tag{4.11}$$

It provides the enormous advantage, such as smoothing out the output distribution since one decision tree output distribution is spiky. Based on this idea, many other averaging techniques have been developed. Combining this averaging technique with boosting technique, the boosted decision trees became a very powerful classification algorithm.

Boosting

Boosting is a way of enhancing the classification performance of typically weak MVA methods, which can be applied to any classification algorithms like neural network. It works very well especially with weak classifiers which provides a simple but less discriminating outputs.

Boosted classifier is a collection of child classifiers (in this case, decision trees) trained on the same sample, but with different weights. The goal of boosting is to provide a new and more stable classifier with a smaller error rate and better performance by combining weak classifiers.

Considering a training sample $\mathcal{T}_k$ containing N_k events with ith event associated with a weight w_i^k and input variables $\mathbf{x}_i$, and a class label $y_i = +1(-1)$ for signal (background), the training of the BDT is done by sequentially performing the following procedure for N_{tree} times after initializing the first sample $\mathcal{T}_1$. In the kth iteration,

1. Train a decision tree T_k on $\mathcal{T}_k$.
2. Assign a weight α_k to T_k.
3. Update $\mathcal{T}_k$ into $\mathcal{T}_{k+1}$.

The weight α_k is called boost weight and given by

$$\alpha_k = \beta \cdot \ln \frac{1-\epsilon_k}{\epsilon_k} \tag{4.12}$$

where β is a free boosting parameter to adjust boosting strength (1 in the original algorithm), and ϵ_k is the mis-classification (error) rate defined by

$$\epsilon_k = \frac{\sum_{i=1}^{N_k} w_i^k \cdot \mathrm{isMissclassified}_k(i)}{\sum_{i=1}^{N_k} w_i^k} \tag{4.13}$$

where $\mathrm{isMissclassified}_k(i)$ returns 1 (0) when $y_i \cdot (T_k(i) - 0.5) \leq 0$ is true (false). Updating the $\mathcal{T}_k$ into $\mathcal{T}_{k+1}$ is done by applying a so-called boost factor for the ith event's weight in the $\mathcal{T}_k$ sample,

$$w_i^{k+1} = w_i^k \cdot e^{\alpha_k \cdot \mathrm{isMissclassified}_k(i)}, \tag{4.14}$$

in the *AdaBoost* (adaptive boost) algorithm [81], and $\alpha_k = 1$ in the *GradientBoost* algorithm.

The output of the algorithm is given by taking a weighted majority vote of the sequence of MVA algorithms. The AdaBoost output, $y_{\mathrm{Boost}}^{\mathrm{Ada}}(\mathbf{x})$, is given by

$$y_{\mathrm{Boost}}^{\mathrm{Ada}}(\mathbf{x}) = \frac{1}{\sum_i^{N_{\mathrm{tree}}} \alpha_i} \cdot \sum_i^{N_{\mathrm{tree}}} \alpha_i \cdot p_i(\mathbf{x}), \tag{4.15}$$

where the sum is over all the trees. For the GradientBoost, the output, $y_{\mathrm{Boost}}^{\mathrm{Grad}}(\mathbf{x})$, is given by the following equation,

$$y_{\mathrm{Boost}}^{\mathrm{Grad}}(\mathbf{x}) = \frac{2}{1 + \exp\left(-2 \cdot \sum_i^{N_{\mathrm{tree}}} p_i(\mathbf{x})\right)} - 1. \tag{4.16}$$

The GradientBoost output distribution, therefore, lies in between the interval of $(-1, +1)$.

AdaBoost has been introduced to classification techniques in early 90's, and in many cases this simple strategy results in dramatic performance gain.

4.4.3. MVA-taggers before MV2

At the end of the b-tagging flow, upstream taggers' information is collected and combined by MVA taggers. The MVA taggers provide the best separation between b- and other flavor jets while projecting the various information into single output distribution.

The first attempt to combine the information using MVA was made by the **JetFitterCombNN** tagger, which was based on neural network provided by the JETNET package [82]. It combines IP3D taggers output weight shown in section 4.3.1, together with the five Jet Fitter output variables, which are N_{Vertex}, $N_{\mathrm{SingleTracks}}$, N_{TrkAtVx}, m_{SV}, f_{SV}, and 3D decay length significance, listed in table 4.4. Figure 4.13 shows the JetFitterCombNN output distributions for all flavor jets.

Then, on top of JetFitterCombNN, the **MV1** tagger was developed using the neural network technique implemented in the TMVA package [78]. The MV1 tagger combines three algorithms outputs, which are IP3D, SV1, and JetFitterCombNN taggers, thus it was not a simple MVA combination but

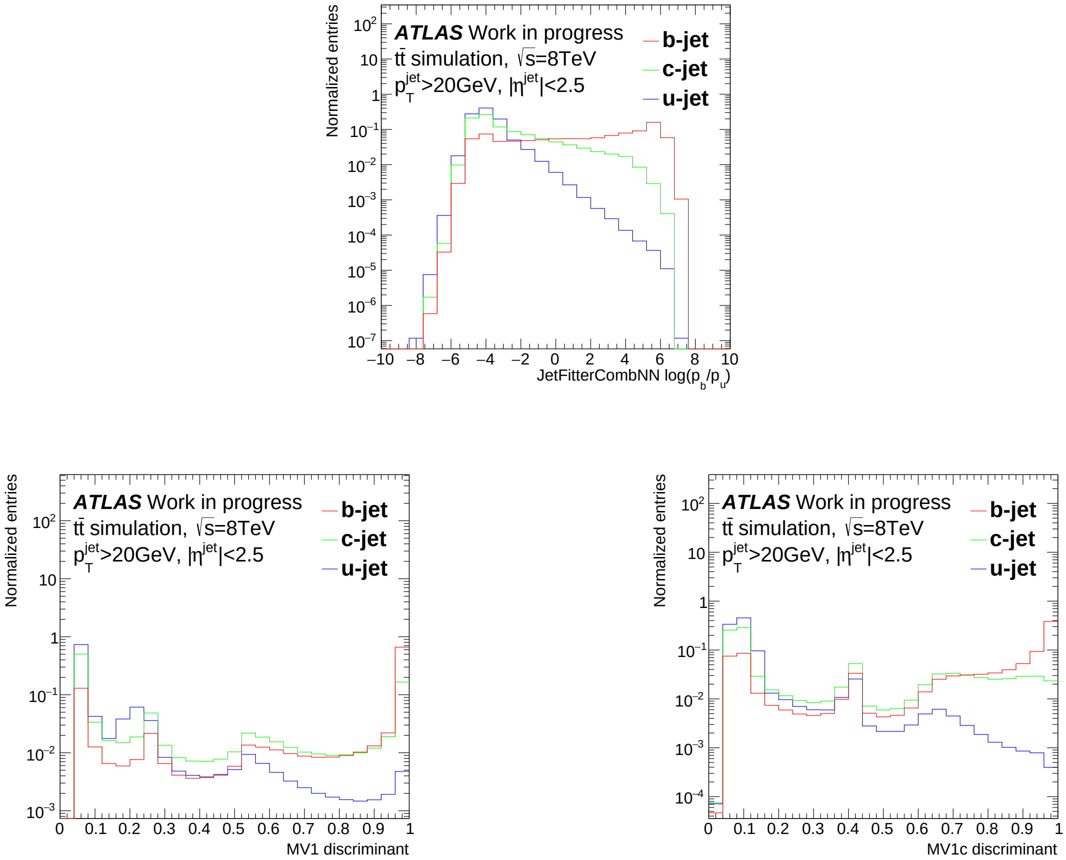


Fig. 4.13: The output distributions of JetFitterCombNN, MV1, and MV1c taggers in top, left, and right plot, respectively. The *b*-, *c*-, and light jets are shown with red, green, and blue lines, respectively.

a cascade and this will be simplified by the MV2 tagger. The MV1 output distribution is shown in fig 4.13 for all flavor jets. It was used for almost all the analyses with b-tagging in Run 1, and achieved very high classification performance, with a light-jet rejection factor of about 150 when the b-tagging efficiency is about 70%. As a variant of MV1, the **MV1c** tagger was also developed [83], and used for the $H \to b\bar{b}$ search as described in section 5. The MV1c tagger was trained against also the c-jet background, while the MV1 tagger was trained only against the light jet background. The MV1c output distribution is also shown in fig 4.13. It has better c-jet rejection than MV1 with a modest cost in light jet rejection. In Run 1 condition at the 70% b-jet efficiency working point, the MV1c tagger has 4% more c-jet rejection and 6% less light jet rejection, in comparison with for MV1 tagger.

4.5. Development of MV2 tagger

Although MV1 achieved unprecedented performance in Run 1, as mentioned in the last section it was a cascade of MVA algorithms so that re-tuning or optimization of the entire b-tagging flow was complicated (see fig 4.14). Also, as shown in fig 4.13, the output shape had several peaks and very sharp rise in the signal side, which makes the cut value determination difficult especially for the high purity (low efficiency) working point. In addition, more variables for the MVA input were expected to increase the discrimination power. For these reasons, development of the successor of the MV1 algorithm was attempted. In section 4.5.1 and 4.5.2, studies were performed in late Run 1 period using the prototype of MV2 algorithm[a]. The list of input variables are re-optimized in section 4.5.3 and the performance of the final version of MV2 algorithms are shown in section 4.6.

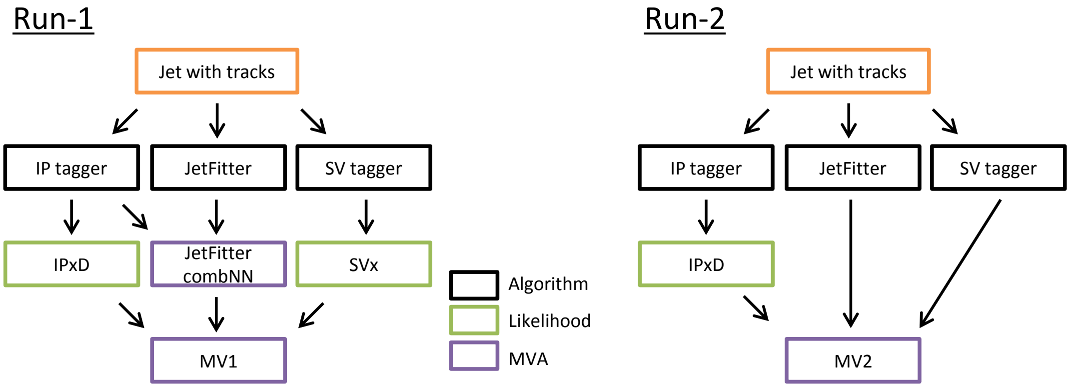


Fig. 4.14: The ATLAS b-tagging flowchart for Run 1 and Run 2. In Run 1 (left), intermediate taggers based on likelihood and MVA methods were used. In Run 2 (right), intermediate MVA tagger, JetFitterCombNN, is removed as well as the SVx likelihood tagger. However, the IPxD likelihood tagger is kept also in Run 2 in order to combine track-level information to jet-level ones.

4.5.1. The BDT output dependence on $p_{\mathrm{T}}^{\mathrm{jet}}$

In this research [83], a first attempt to develop the successor of MV1(c) is investigated. The first approach was to use the Boosted Decision Trees (BDT), due to the huge increase in the NN training

---

[a] These studies were performed using ATLAS software Release 17. Others were done with Release 20.

time when the number of input variables increased. However, there was one negative feature of BDT output distribution, which was the unwanted correlation especially to the $p_{\mathrm{T}}^{\mathrm{jet}}$. In the NN taggers, different strategy to reduce this unwanted correlation was used, in which the bin number defined on the $p_{\mathrm{T}}^{\mathrm{jet}}$-$|\eta^{\mathrm{jet}}|$ histogram is used as the input variable to exploit the correlation between the other input variables and the two, but it was not applicable for BDT since the BDT directly cuts on these numbers, causing very peaky output distribution. One of the possible solutions to avoid this correlation was to build individual BDT classifiers in the several $p_{\mathrm{T}}^{\mathrm{jet}}$ intervals. Ten BDT classifiers were trained and applied in 10 $p_{\mathrm{T}}^{\mathrm{jet}}$ bins from 20 to 500 GeV, with the boundaries at 30, 40, 50, 60, 75, 90, 110, 140, and 200 GeV. Although this approach improved the classification performance, the consumption of computing resources in both RAM and CPU significantly increased because of the 10 BDTs. Also, the usage of this new tagger's output was not as straight-forward as the typical b-tagging cut shown in fig 4.15 since the tagger's output was given by bonding the 10 individual outputs, and the response in the individual $p_{\mathrm{T}}^{\mathrm{jet}}$ bins can be very different from adjacent bins. This required alternative $p_{\mathrm{T}}^{\mathrm{jet}}$ dependent cuts on the taggers output distribution. Due to these reasons, the multiple BDT approach was discarded, and the strategy was shifted to single BDT per tagger.

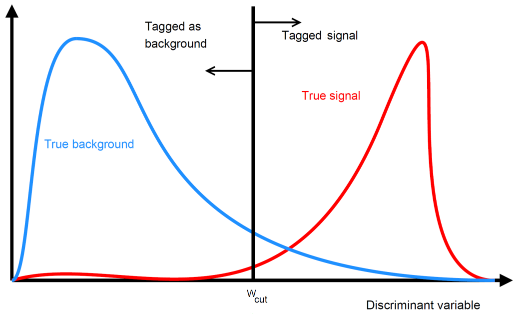


Fig. 4.15: The standard cut approach for general b-tagging. The cut value is obtained for the target efficiency of the true signal events, without consideration of $p_{\mathrm{T}}^{\mathrm{jet}}$ correlation of the discriminant.

However, the single BDT approach produces unwanted $p_{\mathrm{T}}^{\mathrm{jet}}$ correlation. The rejection factor as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ sharply goes down below MV1 as $p_{\mathrm{T}}^{\mathrm{jet}}$ increases, since the BDT learns that the b-jets populate in relatively high p_{T} compared to light jets, due to the $p_{\mathrm{T}}^{\mathrm{jet}}$ profiles of the $t\bar{t}$ sample in the training, as shown in fig 4.16. To minimize the $p_{\mathrm{T}}^{\mathrm{jet}}$ correlation of the BDT output, several weighting schemes were applied and tested.

- *No weight* : No weighting is applied concerning $p_{\mathrm{T}}^{\mathrm{jet}}$ nor η^{jet}.
- *Ratio weight*: Ratio of the 2D spectra of the light jets to the b-jets is applied in such a way that the b-jet 2D spectrum of (p_{T}, $|\eta|$) will be the same as that of light jet. The c-jet 2D spectrum is also treated as that of the b-jet.
- *Flattening weight*: The 2D ($p_{\mathrm{T}}^{\mathrm{jet}}$,$|\eta^{\mathrm{jet}}|$) spectra of all flavors are flattened individually in all the flavors.
- *Inverted CDF weight*: The cumulative distribution function (CDF) on the 2D plane of (p_{T},$|\eta|$) is reverted ($1/\mathrm{CDF}(p_{\mathrm{T}}, |\eta|)$) and then applied as the weight.

Figure 4.17 shows the comparison of the several weighting schemes evaluated in the $t\bar{t}$ sample. The 'no weight' scheme seems to be the best performing in the ROC curve comparison (in fig 4.17 left), but its

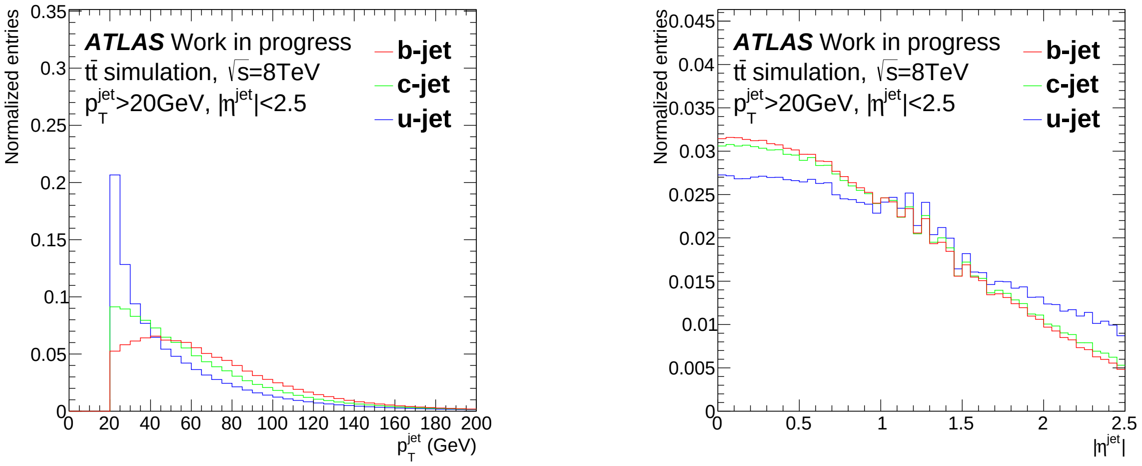


Fig. 4.16: The $p_{\mathrm{T}}^{\mathrm{jet}}$ (left) and $|\eta^{\mathrm{jet}}|$ (right) spectra of the $t\bar{t}$ sample used for the training for $\sqrt{s} = 8$ TeV. The b-, c-, and light jet are shown with red, green, and blue lines, respectively.

light jet rejection factor sharply declines as $p_{\mathrm{T}}^{\mathrm{jet}}$ increases with the global cut for $\epsilon_b = 70\%$ (in fig 4.17 right). Among tested variations, the inverted CDF weighting scheme performs better compared to the other schemes in both the ROC curve and the global cut performance comparisons. For the prototype version, the inverted CDF weight was selected due to the high performance. However, the inverted CDF weighting scheme does not fix the b- and light jet $(p_{\mathrm{T}}^{\mathrm{jet}}, |\eta^{\mathrm{jet}}|)$ spectra the same way. This can potentially cause the biased selection for the selection of b- and light jet rejection, hence, the ratio weighting scheme was selected for the final version.

4.5.2. The c-jet background fraction in the training

The MV1c tagger improved the c-jet background rejection with a modest cost in the light jet rejection. However, the fraction of the c-jet background fraction was arbitrarily taken from the training $t\bar{t}$ sample so that the balance of discrimination powers for the light and c-jet backgrounds was not optimal. Since it is desirable to have only one tagger or algorithm to separate b-, c-, and light jets, this study [83] also attempted to separate three flavors as much as possible. The approach was to introduce the three individual BDTs which are responsible for the discrimination of b/c, b/light, and c/light. However, the usage of this combination again required alternative 2D cut on the two output values, and it required really fine tuning to obtain the valuable cut values for them. Therefore, the strategy was shifted back to the MV1c-like single BDT approach with controlled c-jet fraction in the training sample, while such MVA approach with three output probabilities using single classifier is also studied and preliminary study is summarized in appendix A.2.

Several variations with different c-jet fraction were trained and tested, which were `MV2c00`, `MV2c10`, `MV2c20`, `MV2c30`, and `MV2c100`, where the numbers following 'MV2c' are the relative fraction of the c-jet events added in the training compared to that of b-jet which is the same as that of light jet. Hence, for the `MV2c`X tagger, the c-jet background fraction is $X/(100 + X)$. Figure 4.18 shows the ROC curve comparison of all the variations of c-jet background fraction in both light and c-jet rejection, and fig 4.19 shows comparison in the light and c-jet rejection as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ using a working point for $\epsilon_b = 70\%$.

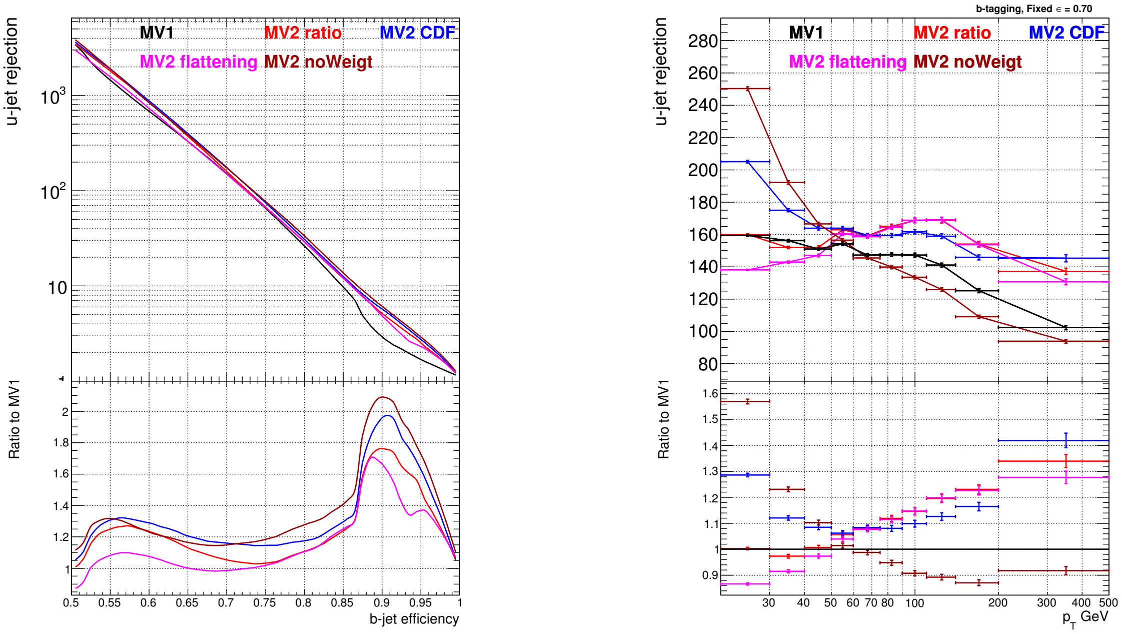


Fig. 4.17: The ROC curves (left) and the p_T dependence of the light jet rejection using the global cut for $\epsilon_b = 70\%$ (right). The brown, red, blue, and magenta curves show the performance of the training with no weight, ratio weight, inverted CDF weight, and the flattening weight schemes, respectively. The MV1 is also shown as the reference with the black curve and the ratios of the variations to MV1 are shown in the bottom panels.

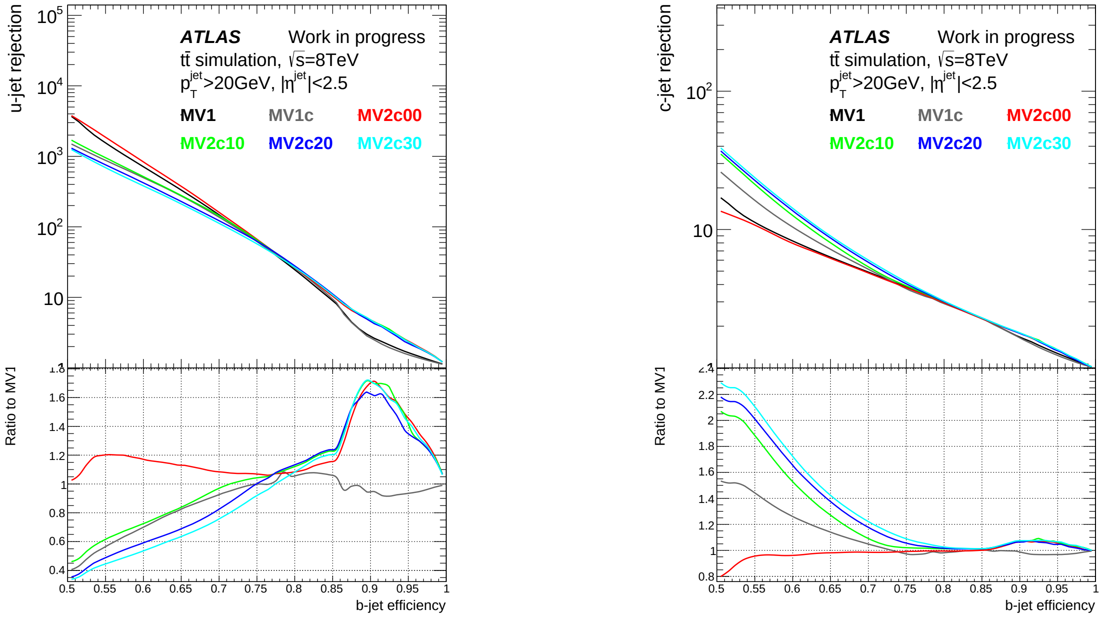


Fig. 4.18: The ROC curve comparison for variations of c-jet background fraction of MV2 for the light jet background (right) and the c-jet background (right). The MV2c00, MV2c10, MV2c20, MV2c30, and MV2c100 taggers are shown with red, green, orange, yellow, cyan, and brown curves, respectively. The MV1 and MV1c taggers are also shown as reference with black and gray curves, respectively, and the ratios to MV1 are shown in the bottom panels.

As it is clearly visible in fig 4.18 and 4.19 there is a trade-off between the c- and light jet background rejection depending on the c-jet background fraction difference. In between MV2c30 and MV2c100 there is no significant improvement in c-jet rejection while the light jet rejection decreases sharply. In order to adapt to many needs from the physics goals, it was decided to keep three variations of MV2 taggers, which are **MV2c00**, **MV2c10**, and **MV2c20**.

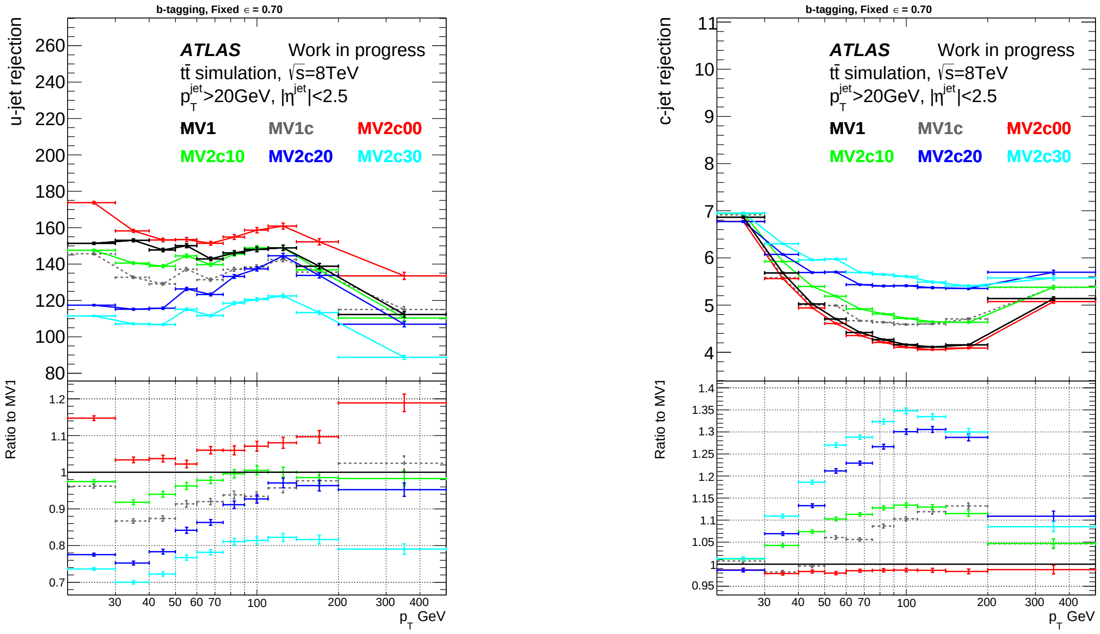


Fig. 4.19: The performance comparison in the rejection factor as a function of p_T^{jet} with the global cut for $\epsilon_b = 70\%$ for the light jet background (right) and the c-jet background (right). The MV2c00, MV2c10, MV2c20, MV2c30, and MV2c100 taggers are shown with red, green, orange, yellow, cyan, and brown curves, respectively. The MV1 and MV1c taggers are also shown as reference with black and gray curves, respectively, and the ratios to MV1 are shown in the lower panel.

4.5.3. Final configuration

The final configuration is optimized at the end of long shut down period.

List of input variables

The list of input variables was also optimized in this thesis, in particular:

- The intermediate MVA outputs, JetFitterCombNN probabilities, were removed to achieve a smooth flow for optimization.
- Although it was not necessarily, another intermediate MVA outputs, SV1 likelihood probabilities, were removed to simplify the b-tagging algorithm flow.
- Contrarily to the other MVA algorithms, the IP2D and IP3D likelihood probabilities are kept, in order to combine track-level probabilities into jet-level ones. The jet-level probabilities were directly fed to the BDTs in Run 1, while the log-likelihood-ratios were chosen for Run 2 due to the improvement observed.

- SV0 tagger's inputs were replaced with those from the SV1 tagger due to the improvement observed.

- The list of JetFitter inputs were refined due to the availability of some variable (χ^2/N_{DOF} to drop) and an improvement observed (e.g. use of ΔR while dropping $\Delta\phi$ and $\Delta\eta$).

- In order to make use of the correlation between the kinematic properties ($p_{\mathrm{T}}^{\mathrm{jet}}$ and η^{jet}) and other input variables, the $p_{\mathrm{T}}^{\mathrm{jet}}$ and η^{jet} were added for Run 2.

Table 4.3 summarizes the transition of the list of input variables from Run 1 to Run 2 and the input variables are described in table 4.4.

4.6. The *b*-tagging performance in Run 1 and Run 2

The improved b-tagging performance by MV2 taggers are shown and compared to the existing taggers in fig 4.20–4.22 and 4.23–4.25 in Run 1 and Run 2 configurations, respectively. In the ROC curve comparisons in fig 4.20 and 4.23, there is a peaking improvement in the very high b-jet efficiency ($0.86 \lesssim \epsilon_b \lesssim 0.98$) region. This is because there are two regimes for any taggers' weights: in first regime, linked with lifetime, the b- and light jets are discriminated clearly, while in the second regime, the b- and light jets are less separated because of the limited detector resolution. The ROC curve drop in the high efficiency corresponds to the transition between the two regimes and depends on the taggers' performance.

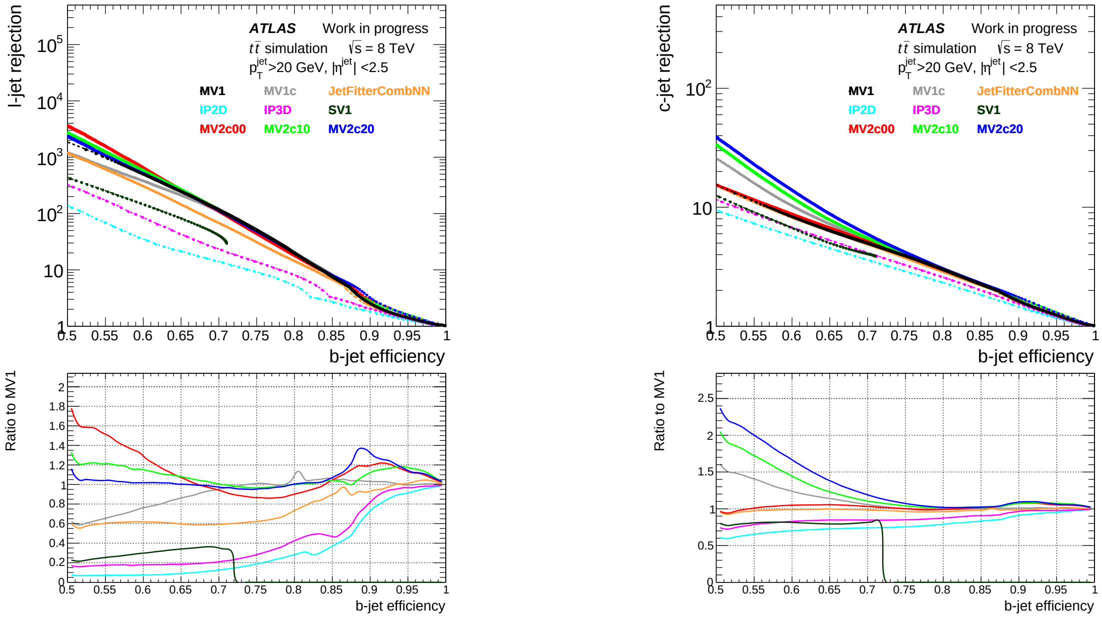


Fig. 4.20: The ROC curve comparison of the various taggers in the $t\bar{t}$ sample with Run 1 condition for the light jet background (left) and the c-jet background (right).

Table 4.5 shows the rejection factors at several b-tagging efficiency working points (50%, 60%, 70%, and 80% b-jet efficiency), comparing the numbers in Run 1 and Run 2 configurations.

Table 4.3: The list of MV2 input variables optimized at the end of Run 1 (prototype version) and long shutdown period (final version). The final set of MV2 input variables shown in the appendix A.

Algorithm	Input variable	Prototype	Final set
jet kinematics	$p_{\mathrm{T}}^{\mathrm{jet}}$	–	✓
	η^{jet}	–	✓
IP2D	p_b	✓	–
	p_u	✓	–
	p_c	✓	–
	$\log(p_b/p_u)$	–	✓
	$\log(p_b/p_c)$	–	✓
	$\log(p_c/p_u)$	–	✓
IP3D	p_b	✓	–
	p_u	✓	–
	p_c	✓	–
	$\log(p_b/p_u)$	–	✓
	$\log(p_b/p_c)$	–	✓
	$\log(p_c/p_u)$	–	✓
SV0	3D decay length significance	✓	–
	N_{TrkAtVx}	✓	–
	$N_{\mathrm{TwoTrackVx}}$	✓	–
	m_{SV}	✓	–
	f_{SV}	✓	–
SV1	p_b	✓	–
	p_u	✓	–
	p_c	✓	–
	N_{TrkAtVx}	–	✓
	$N_{\mathrm{TwoTrackVx}}$	–	✓
	m_{SV}	–	✓
	f_{SV}	–	✓
	L_{xy}	–	✓
	L_{xyz}	–	✓
	$L_{xyz}/\sigma(L_{xyz})$	–	✓
	$\Delta R(\overrightarrow{SV_{xyz}}, \overrightarrow{j_{xyz}})$	–	✓
Jet Fitter	3D decay length significance	✓	✓
	m_{SV}	✓	✓
	f_{SV}	✓	✓
	N_{TrkAtVx}	✓	✓
	N_{Vertex}	✓	✓
	$N_{\mathrm{SingleTracks}}$	✓	✓
	$N_{\mathrm{TwoTrackVx}}$	✓	✓
	$\Delta\phi$	✓	–
	$\Delta\eta$	✓	–
	ΔR	–	✓
	χ^2/N_{DOF}	✓	–
JetFitterCombNN	p_b	✓	–
	p_u	✓	–
	p_c	✓	–
Number of input variables		27	24

Table 4.4: The list of MV2 input variables and their description.

Algorithm	Input variable	Description
IP2D, IP3D	p_b	output likelihood probability for b-jet hypothesis
	p_u	output likelihood probability for light jet hypothesis
	p_c	output likelihood probability for c-jet hypothesis
	$\log(p_b/p_u)$	log-likelihood ratio between b- and light jet hypotheses
	$\log(p_b/p_c)$	log-likelihood ratio between b- and c-jet hypotheses
	$\log(p_c/p_u)$	log-likelihood ratio between c- and light jet hypotheses
SV0, SV1	p_b	output likelihood probability for b-jet hypothesis
	p_u	output likelihood probability for light jet hypothesis
	p_c	output likelihood probability for c-jet hypothesis
	N_{TrkAtVx}	number of associated tracks at the SV
	$N_{\text{TwoTrackVx}}$	number of two-track pairs forming SV candidates
	m_{SV}	invariant mass of the SV
	f_{SV}	energy fraction of charged tracks associated to the SV
	L_{xy}	transverse decay length
	L_{xyz}	3D decay length
	$L_{xyz}/\sigma(L_{xyz})$	3D decay length significance
	$\Delta R(\overrightarrow{SV_{xyz}}, \overrightarrow{j_{xyz}})$	angular distance between jet and SV's flight direction
Jet Fitter	3D decay length significance	Significance of the average distance
		between the PV displaced vertices
	m_{SV}	invariant mass of charged particles associated
		to all displaced vertices
	f_{SV}	energy fraction of charged particles associated
		with all displaced vertices to the jet
	N_{TrkAtVx}	number of tracks associated to the displaced vertex
		with at least two tracks
	N_{Vertex}	number of vertices with more than 1 tracks
	$N_{\text{SingleTracks}}$	number of vertices with single track
	$N_{\text{TwoTrackVx}}$	number of two-track pairs forming a vertex candidate
		(prior to decay chain fit)
	$\Delta\phi$	$\Delta\phi$ between jet and vectorial sum of all SVs' momenta
	$\Delta\eta$	$\Delta\eta$ between jet and vectorial sum of all SVs' momenta
	ΔR	$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, where $\Delta\phi$ and $\Delta\eta$ are given above
	χ^2/N_{DOF}	decay chain fit quality
JetFitterCombNN	p_b	output likelihood probability for b-jet hypothesis
	p_u	output likelihood probability for light jet hypothesis
	p_c	output likelihood probability for c-jet hypothesis

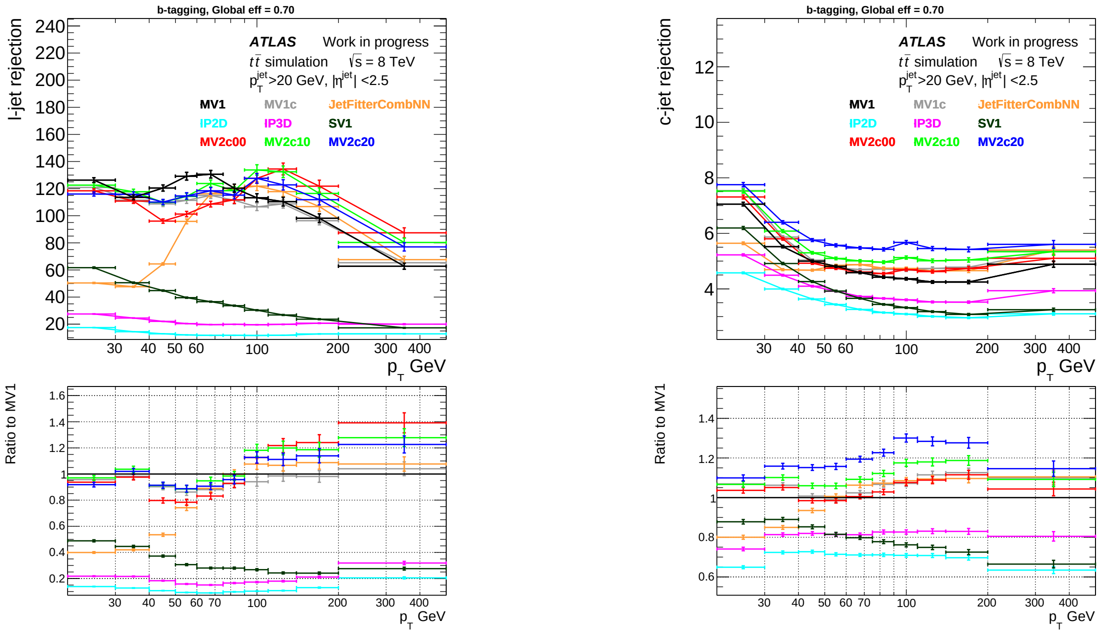


Fig. 4.21: The performance comparison in the rejection factor as a function of p_T^{jet} with the global cut for $\epsilon_b = 70\%$ in the $t\bar{t}$ sample with Run 1 condition for the light jet background (left) and the c-jet background (right).

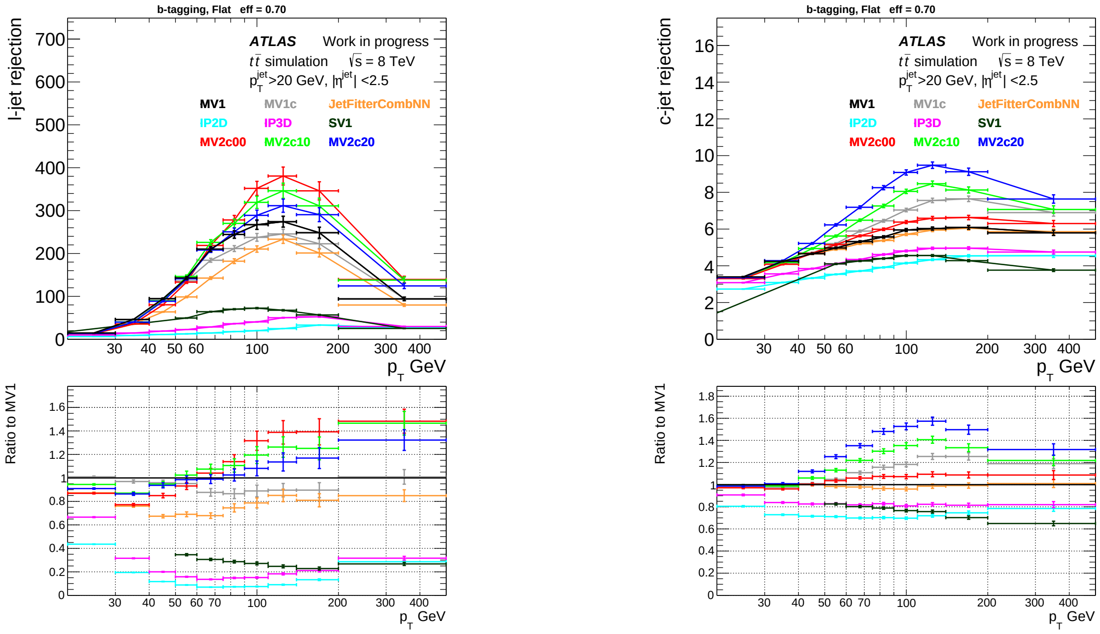


Fig. 4.22: The performance comparison in the rejection factor as a function of p_T^{jet} with the flat efficiency cuts for $\epsilon_b = 70\%$ in the $t\bar{t}$ sample with Run 1 condition for the light jet background (left) and the c-jet background (right).

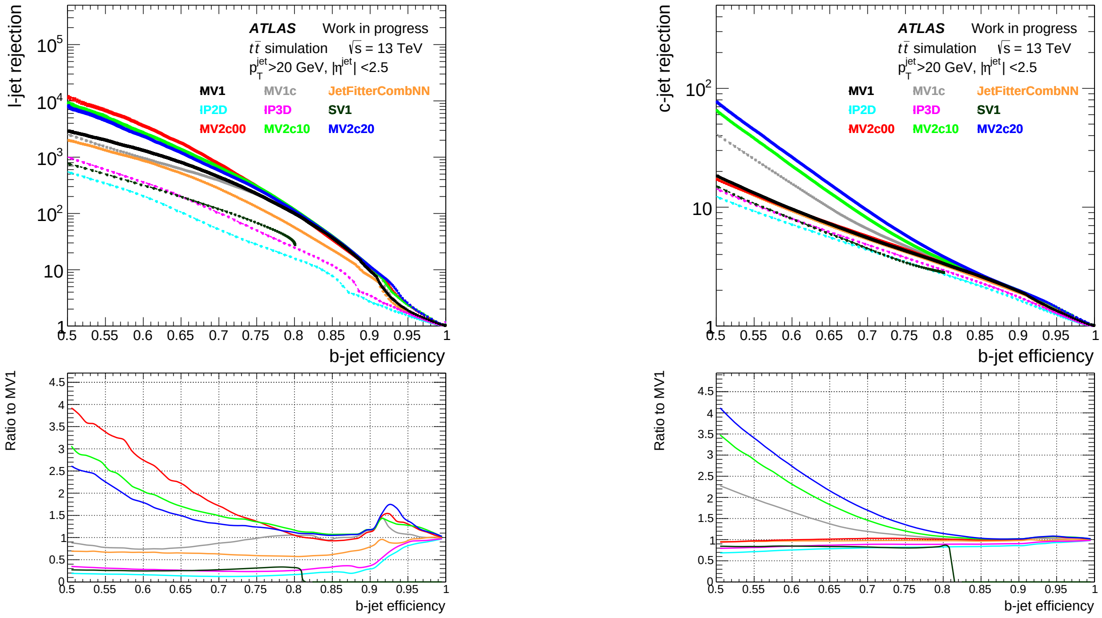


Fig. 4.23: The ROC curve comparison of the various taggers in the Run 2 condition in $t\bar{t}$ sample for the light jet background (left) and the c-jet background (right).

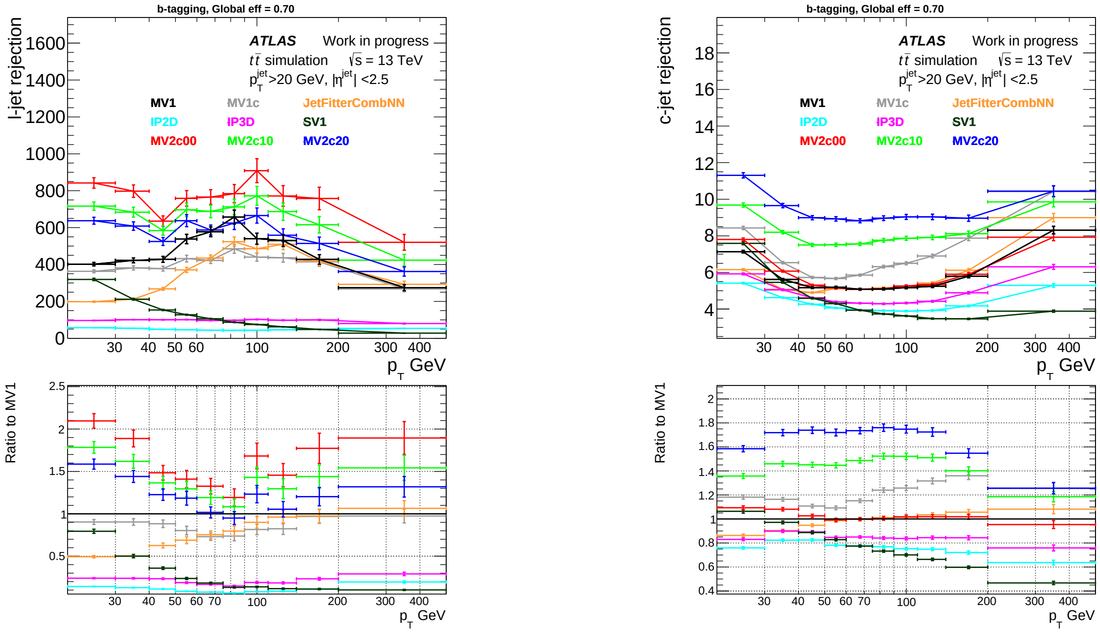


Fig. 4.24: The performance comparison in the rejection factor as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ with the global cut for $\epsilon_b = 70\%$ in the $t\bar{t}$ sample with Run 2 condition for the light jet background (left) and the c-jet background (right).

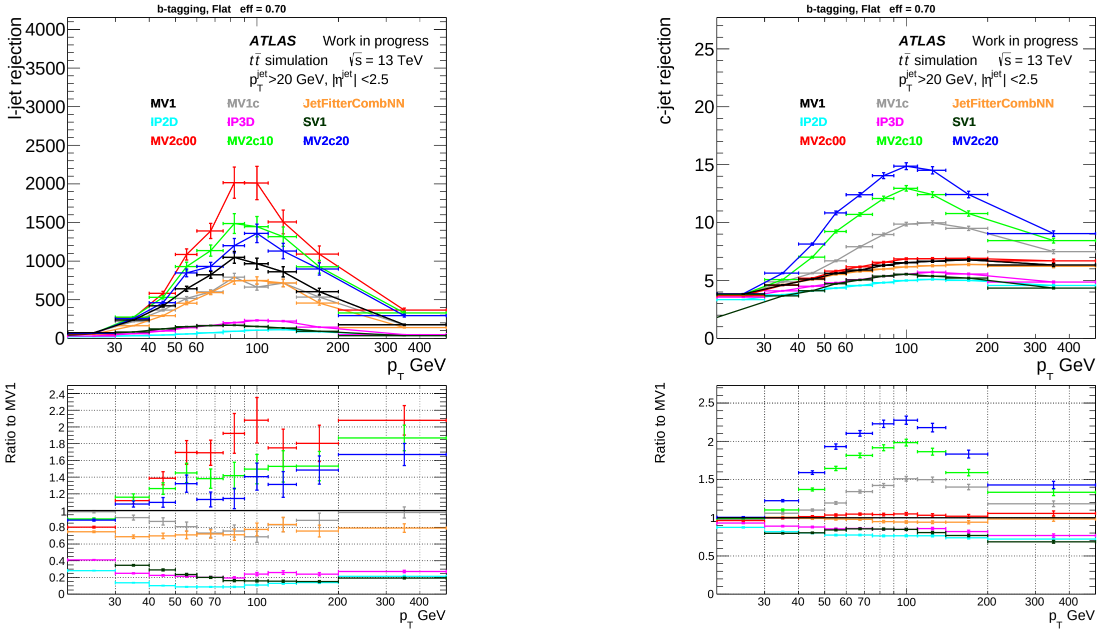


Fig. 4.25: The performance comparison in the rejection factor as a function of p_T^{jet} with the flat efficiency cuts for $\epsilon_b = 70\%$ in the $t\bar{t}$ sample with Run 2 condition for the light jet background (left) and the c-jet background (right).

4.6.1. Pileup dependence of the b-tagging performance

In this subsection, pileup dependence of the b-tagging performance has been checked, using the light jet rejection factor as a function of mean number of interactions per bunch crossing ($< \mu >$) and number of reconstructed vertices (n_{PV}). Figure 4.26 shows the pileup dependence of the b-tagging performance using 70% b-jet efficiency working point. Using the 70% b-jet efficiency working point, the light jet rejection factor slightly decreases when both $< \mu >$ and n_{PV} increase. Similarly but using flat efficiency cuts, the pileup dependence has been checked and shown in fig 4.27. A significant decrease of the light jet rejection is seen when n_{PV} increases: the decrease goes down to about 30% below the low $n_{\text{PV}}(<10)$ region. This significant decrease seen only in the flat efficiency cuts, when n_{PV} increases, is not yet fully understood and studies in details are underway.

4.6.2. Impact of the IBL insertion

The b-tagging performance in Run 2 is significantly lifted up from Run 1. Although this is due to various changes, the main ones affecting the b-tagging performance are:

- Increased center-of-mass energy from 8 to 13 TeV, for which the detector was originally designed.
- Insertion of the IBL (see section 2.2.2). Additional pixel layer improves the tracking performance significantly, greatly affecting the b-tagging performance.
- Improved b-tagging algorithms.
- Difference in the pileup conditions.

Differences in the jet kinematics (p_T^{jet} and η^{jet}) and mean number of interactions per bunch crossing from Run 1 to Run 2 are shown in fig 4.28. The distributions of mean number of interactions per

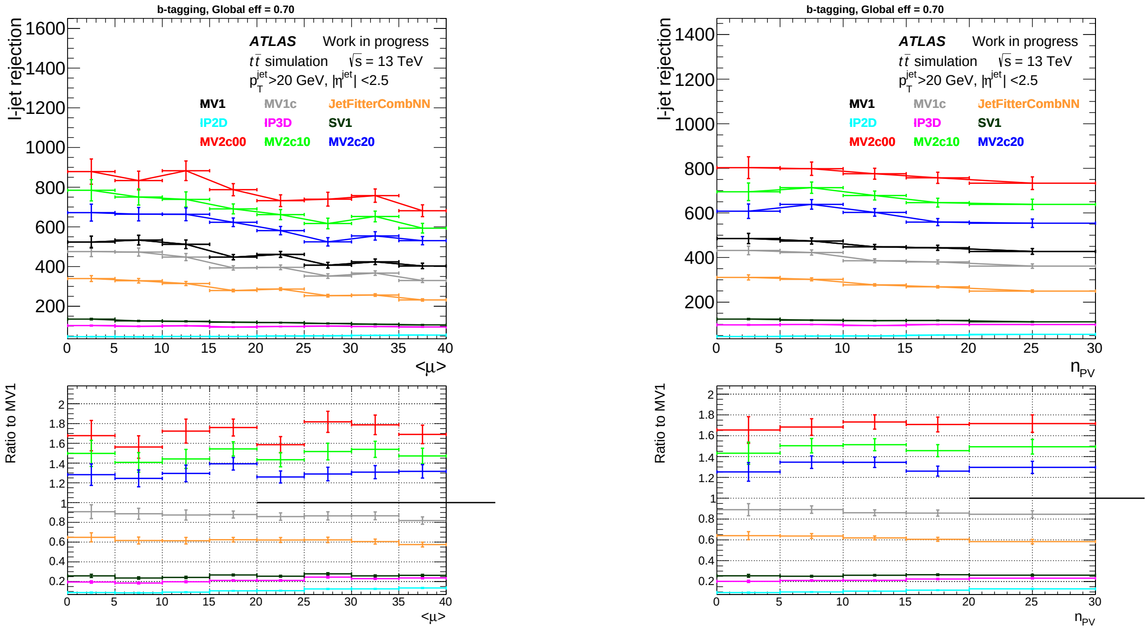


Fig. 4.26: The b-tagging performance with a global cut for 70% b-jet efficiency working point. Shown are light jet rejection factor as a function of mean number of interactions per bunch crossing (left) and number of reconstructed vertices (right).

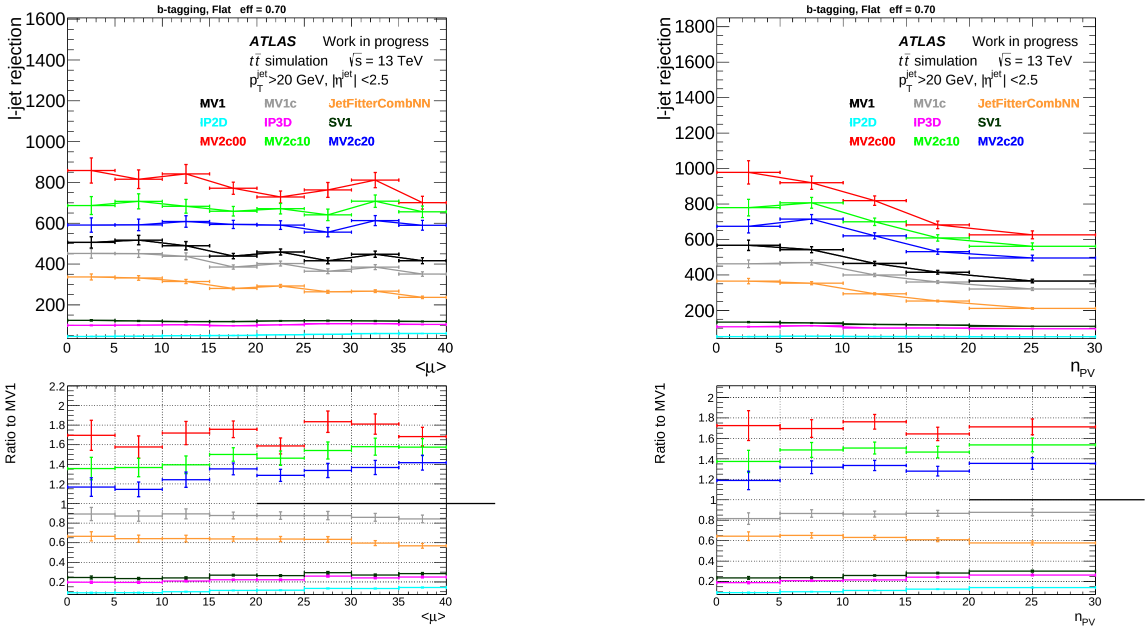


Fig. 4.27: The b-tagging performance a set of cuts for flat 70% b-jet efficiency. Shown are light jet rejection factor as a function of mean number of interactions per bunch crossing (left) and number of reconstructed vertices (right). In both plots, the b-jet efficiency in every bin is kept at 70%.

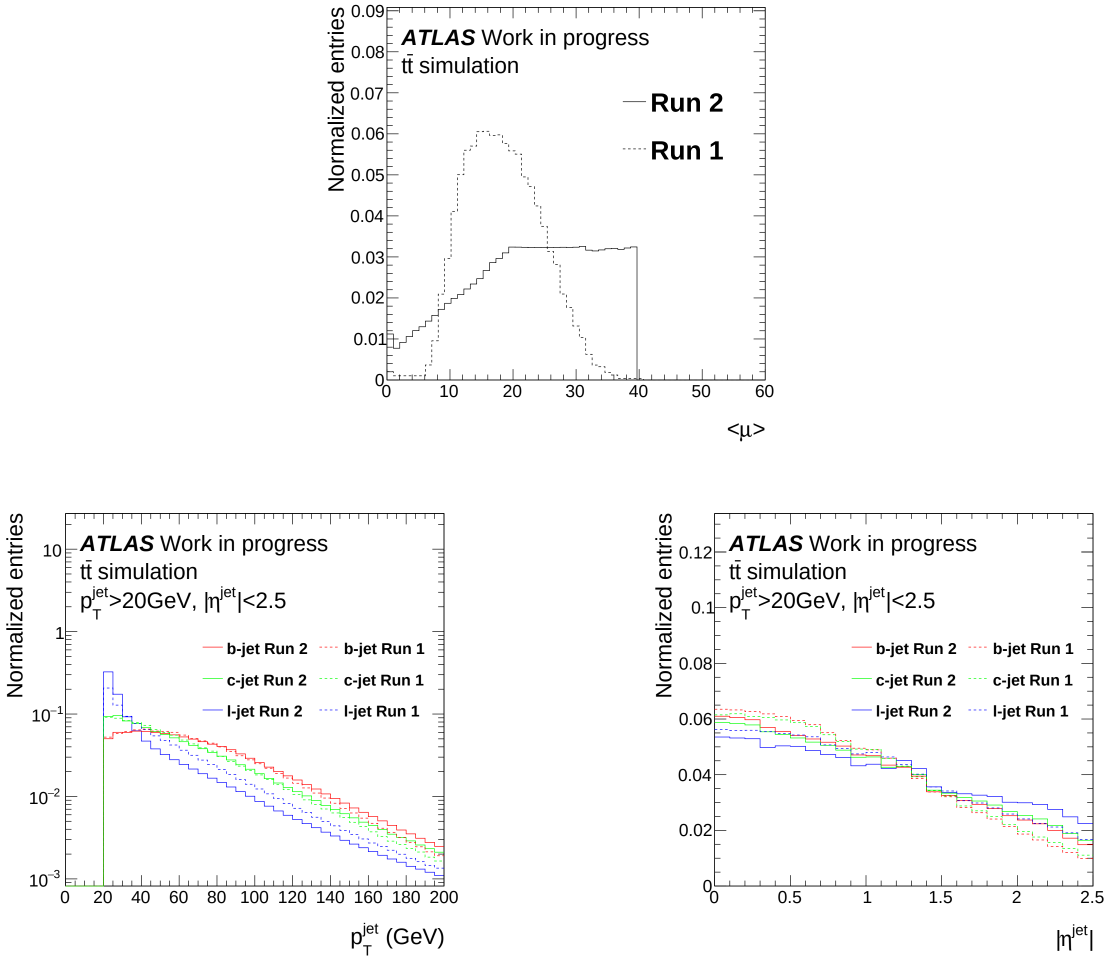


Fig. 4.28: Difference in the distributions of mean number of interactions per bunch crossing (top). Difference in $p_{\mathrm{T}}^{\mathrm{jet}}$ (bottom left) and η^{jet} (bottom right) in Run 1 and Run 2 configurations, separately for each flavors are shown.

bunch crossing are artificially chosen in the simulated samples, and a correction is applied in order to match with the one observed in data in actual analyses but not here. The light jet p_{T} spectrum is softer than that in Run 1 configuration, due to the increased radiation and the pileup interactions. The b- and c-jet p_{T} spectra are harder in Run 2 configuration. In η^{jet} distributions, all the flavors are populated slightly more in the forward region in Run 2 configuration.

In order to quantify the impact of the IBL insertion, a reweighting is applied to the Run 1 sample in such a way that the jet kinematics in Run 1 sample are the same as the Run 2 sample: the scale factors applied to the Run 1 sample are obtained by 2D maps in $|\eta^{\mathrm{jet}}|$–$p_{\mathrm{T}}^{\mathrm{jet}}$ plane for each jet (b-, c-, and light) flavor. Additional correction to the Run 1 sample is applied for the mean number of interactions per bunch crossing in the similar way as the kinematic corrections but in 1D map inclusively in all flavors. Using the weighted Run 1 sample, the impact of the IBL insertion is checked using the MV2c20 tagger and shown in fig 4.29. The IBL improves the light jet rejection by a factor up to about five at the same b-jet efficiency, and c-jet rejection factor up to about twice higher.

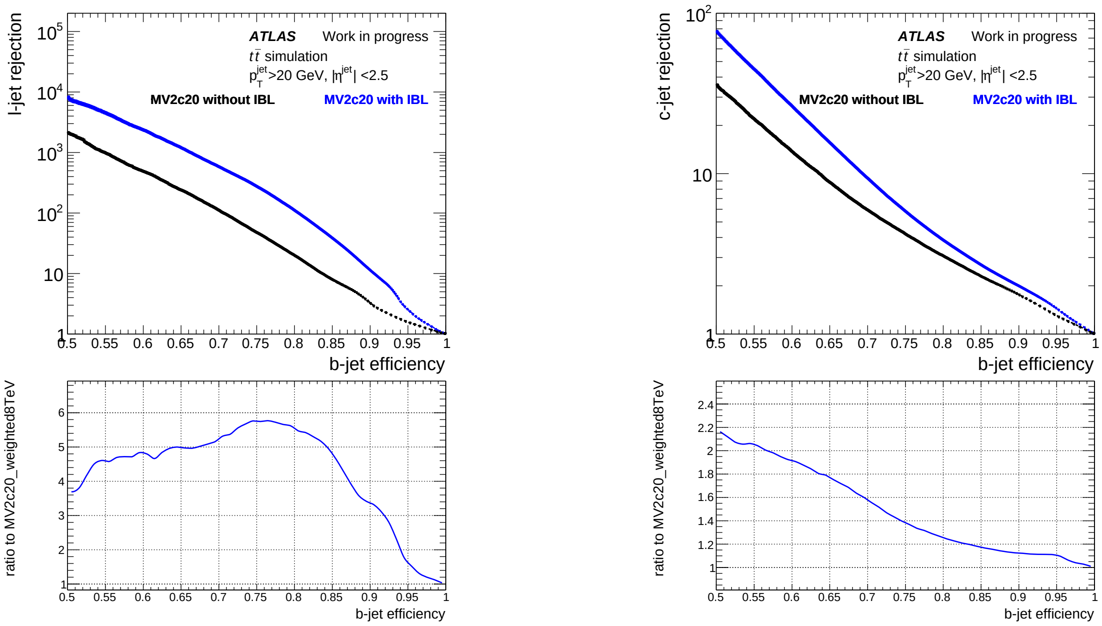


Fig. 4.29: The ROC curve comparison of the MV2c20 tagger with (blue curve) and without (black curve) IBL, in the $t\bar{t}$ sample for the light jet background (left) and the c-jet background (right).

Comparison with the IBL-TDR

Using IP3D tagger, a comparison with IBL-TDR [36] has been checked. Figure 4.30 compares the light jet rejection as a function of b-jet efficiency, with and without IBL, comparing also the difference from the studies made in 2010. The rejection factors are also compared in table 4.6. There is significant improvement from the study made in 2010. The possible reasons are:

- Optimized tracking including the pixel clustering using neural network method.
- Optimized algorithm of IP3D for Run 2 with IBL, such as track selection and track grading scheme.

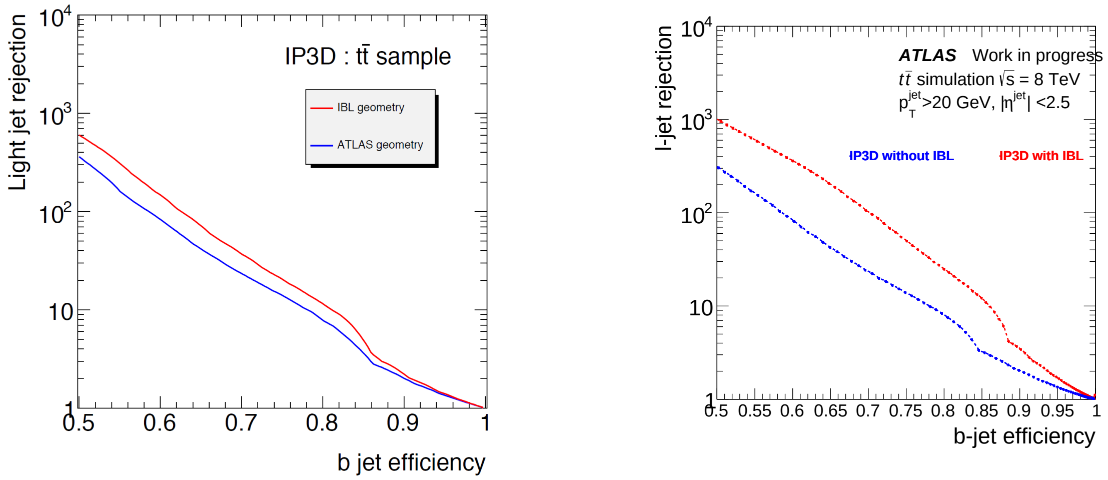


Fig. 4.30: The ROC curve comparison of the IP3D tagger with (red curve) and without (blue curve) IBL, in the $t\bar{t}$ sample for the light jet background. The left plot is from a study made in the IBL-TDR and right plot is from studies made here.

4.7. Data-to-simulation calibration of b-tagging

Without a very specific selection on samples, performance of the b-tagging algorithms can be measured in only the simulated sample due to the jet truth labeling based on the so-called "truth" information. Moreover, the simulation may not fully reproduce the data sufficiently well, and the remaining differences have to be measured. Exploiting dedicated methods, to select a very particular set of data samples, enables to measure the b-tagging performance in data. The simulated performance which is different from the data one should be corrected using the so-called data-to-simulation scale factors (SF) derived in several dedicated calibration methods.

4.7.1. b-tagging calibration in Run 1

The b-, c-, and light jet tagging efficiencies measured in Run 1 data are shown below [71].

b-jet efficiency

Although there are several ways to measure the b-jet tagging efficiency in data such as using muon-in-jet samples, the most common and accurate ones are the methods using $t\bar{t}$ sample, taking advantage of large $t\bar{t}$ cross-section at the LHC. There were so-called tag-and-probe and kinematic selection method [84] for the b-jet efficiency measurement using $t\bar{t}$ samples. A new method called combinatorial likelihood approach, which is an extension of kinematic selection one, is developed [85].

The sample used for this analysis was dileptonic $t\bar{t}$ events collected in 2012. In order to extract the efficiency corresponding to a certain b-tagging requirement, the following equations are used. For the case with exactly two jets in the event:

$$f_{2\text{tags}} = f_{bb}\epsilon_b^2 + f_{bj}\epsilon_j\epsilon_b + (1 - f_{bb} - f_{bj})\epsilon_j^2 \tag{4.17}$$

$$\begin{aligned} f_{1\text{tag}} =& 2f_{bb}\epsilon_b(1-\epsilon_b) + f_{bj}[\epsilon_j(1-\epsilon_b) + (1-\epsilon_j)\epsilon_b] \\ &+ (1 - f_{bb} - f_{bj})2\epsilon_j(1-\epsilon_j) \end{aligned} \tag{4.18}$$

where ϵ_b (ϵ_j) is the b-jet (non b-jet) efficiency, $f_{1\text{tag}}$ ($f_{2\text{tags}}$) is the fraction of events with 1 (2) tagged jets, and f_{bb} (f_{bj}) is the fraction of events with a true bb (bj) jet pair. The b-tagging efficiency can be obtained by maximizing a likelihood function, taking both $f_{1\text{tag}}$ and $f_{2\text{tags}}$ from data, with f_{bb}, f_{bj}, and ϵ_j from simulation. Figure 4.31 shows the result of efficiency measurement in both data and simulation and data-to-simulation SFs in ten $p_{\text{T}}^{\text{jet}}$ bins. The SF varies in the range 0.965–1.008 with total uncertainties of 0.018–0.084.

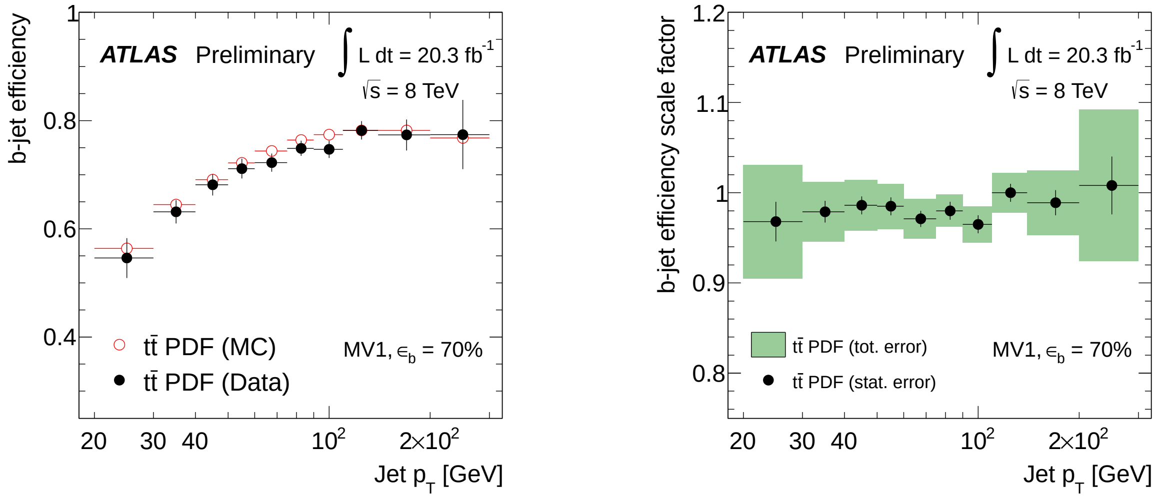


Fig. 4.31: The p_{T} dependence of b-jet efficiencies (left) and b-jet efficiency SFs (right) for the MV1 algorithm at 70% b-jet efficiency working point. The error bars on data points in the left plot represents the total statistical and systematic uncertainties. In the right plot, the statistical only (black lines) and total errors (green shaded region) are shown [85].

c-jet efficiency

There are also several ways to measure the c-jet tagging efficiency: so-called $W + c$ method [86] and D^* method [87]. Although the $W + c$ method is expected to improve the c-jet efficiency measurement, more statistics is needed. Hence, the results provided for Run 1 physics analyses were provided using the D^* calibration method.

This method uses jets containing D^* ($D^{*\pm}$) mesons for c-jet efficiency measurement by comparing yield of D^* before and after the b-tagging requirement. The $D^{*\pm} \to D^0\pi^\pm$ ($D^0 \to K^\mp\pi^\pm$) decay mode offers distinctive kinematic features, resulting in a modest combinatorial background. The contamination of D^* mesons originating from b-hadron decays is measured with a fit to the D^0 pseudo-proper time distribution [87]. Figure 4.32 shows the measured tagging efficiency of c-jets for data and simulation for MV1 at the 70% b-tagging efficiency working point. The data-to-simulation SF varies in the range 0.86–0.97 depending on $p_{\text{T}}^{\text{jet}}$, with a 8–15% uncertainty.

Light jet efficiency

The mistag rate, defined as the fraction of light-jets which are tagged, is measured in an inclusive jet sample using the *negative tag* method [87]. Light-flavor jets are tagged as b-jets mainly because of the finite resolution of the inner detector and the presence of tracks from displaced vertices of long-lived particles or material interactions. Fake secondary vertices and impact parameters of tracks in light jets are expected to distribute approximately symmetrically around the primary vertex while those from

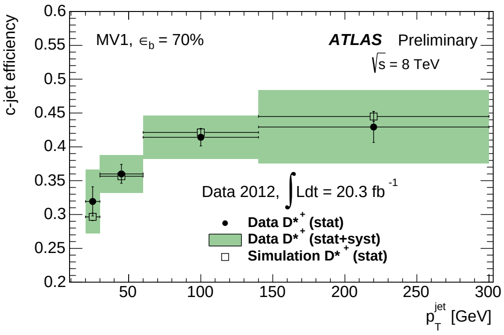


Fig. 4.32: The c-jet tagging efficiency in data and simulation for jets containing D^*, for the 70% working point of the MV1 tagger [87].

true SVs are not. The mistag rate is estimated by using the tracks' impact parameter and/or secondary vertex decay lengths which are negatively signed. A correction has been applied for jet-direction mismeasurement for heavy flavor jets. For light-jets, another correction has been applied for tracks or SVs originating from long-lived particles and materiel interactions. The measured mistag rates in data and simulation for the MV1 tagging algorithm at 70% efficiency working point for central jets ($|\eta^{\text{jet}}| < 1.2$) are shown in Fig 4.33. For the chosen working point, the mistag rate ranges from 0.5% to 2.5%. The

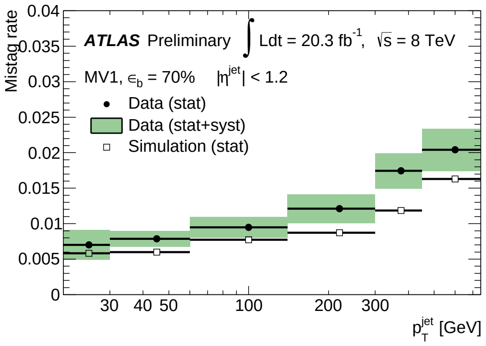


Fig. 4.33: The mistag rate in data and simulation for the MV1 algorithm at 70% efficiency working point, obtained with the negative tag method, for jets with $|\eta| < 1.2$ [87].

data/simulation SFs are slightly larger than unity, with relative total uncertainties ranging 15%–43%.

New calibration approach: calibration of entire tag weight

The calibration analyses for b-tagging discussed above were performed for the cumulative operating points: the SFs are provided for jets whether to be tagged or not for a certain operating point. However, a new calibration method, so-called continuous calibration [88], is developed. This method calibrates the entire tagging output distribution to data and done for MV1c tagger: the tagging output distribution is divided into five bins with boundaries of the working points (100%, 80%, 70%, 60%, 50% and 0%). To provide this, an extension of the combinatorial likelihood approach was used for b-jet tagging efficiency, and D^* and mis-tag rate calibration methods are used for the c-jet and light jet efficiency measurements. The biggest advantage coming from this new approach is a capability to directly feed the tagging output to some MVAs in physics analyses: the analysis presented in the next chapter also profit from this.

4.7.2. Towards b-tagging calibration in Run 2 data

In order to use the newly developed algorithm, MV2c20 presented in this thesis, in ATLAS physics analyses in Run 2, a calibration has to be performed using Run 2 data. Although the statistics from early 2015 data has not yet allowed to perform a full calibration, as a first step, a compatibility between data and simulation is checked using dileptonic $t\bar{t}$ enhanced sample in the $e\mu$-channel [89].

The MV2c20 discriminant distributions in data and simulated samples are compared in fig 4.34. Given the uncertainty mainly coming from the limited statistics from early Run 2 data, corresponding to an integrated luminosity of 85 pb^{-1}, a good agreement is found between the data and simulated events.

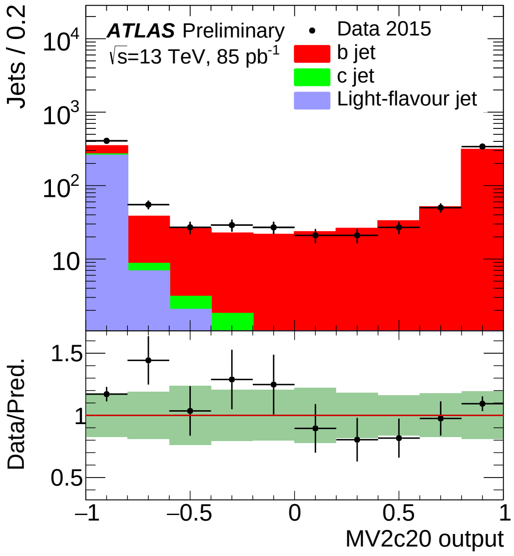


Fig. 4.34: MV2c20 discriminant after kinematic selection to enhance dileptonic ($e\mu$-channel) $t\bar{t}$ events in Run 2 data. Along with the data points, the simulated distribution is shown and divided according to the flavor of the jets: b-, c-, and light flavors are shown with red, green and azure colors. Lower panel shows the data-to-simulation ratio together with the total statistical and systematic uncertainty on the simulation with the light green shade [89].

4.8. Summary

The MV2 taggers were developed and improve b-tagging performance in both Run 1 and Run 2. The $p_{\mathrm{T}}^{\mathrm{jet}}$ correlation of the BDT output is improved. The c-jet background fraction is optimized, and three variations were introduced to supply the balance of c- and light jet background rejections optimal for various analyses: MV2c00, MV2c10, and MV2c20 taggers. The MV2c20 will be used as the baseline b-tagging algorithm in the ATLAS experiment in Run 2. The MV2m tagger is developed and the preliminary study shows suitable working points for c-tagging by 1D cut. The impact of the increased center-of-mass energy to the b-tagging performance was found to be small (within a few percent level). The b-tagging performance dramatically improves in Run 2 condition compared to that of Run 1, due to the IBL insertion, and the optimized b-tagging algorithms.

Table 4.5: The rejection factors for c-jet and light jet background for the b-jet efficiency 50, 60, 70, and 80% working points measured in $t\bar{t}$ sample in Run 1 and Run 2 conditions. The reweighting made on Run 1 sample did not impact b-tagging performance (within a few percent level), hence the number here are only for non-reweighted Run 1 sample.

tagger	rejection for	b-jet efficiency	Run 1	Run 2
MV1	light jet	50%	1862	2948
		60%	507	1335
		70%	118	451
		80%	20	101
	c-jet	50%	15	18.5
		60%	8.3	9.7
		70%	5.0	5.5
		80%	3.0	3.4
MV1c	light jet	50%	1120	2540
		60%	385	978
		70%	111	392
		80%	20	105
	c-jet	50%	26	41
		60%	10	16
		70%	5.3	6.6
		80%	2.96	3.4
MV2c00	light jet	50%	3598	11753
		60%	655	3639
		70%	111	772
		80%	18	106
	c-jet	50%	15.5	17.2
		60%	8.8	9.6
		70%	5.2	5.7
		80%	2.98	3.4
MV2c10	light jet	50%	2723	9105
		60%	588	2754
		70%	118	672
		80%	20	118
	c-jet	50%	33	65
		60%	12	22
		70%	5.5	8.0
		80%	3.05	3.6
MV2c20	light jet	50%	2353	8163
		60%	519	2387
		70%	115	590
		80%	20	112
	c-jet	50%	39	77
		60%	14	26
		70%	6.0	9.4
		80%	3.1	3.8

Table 4.6: The light jet rejection factors in without-IBL and with-IBL conditions, comparing studies made in 2010 in IBL-TDR and 2015. The numbers are for IP3D tagger's 60% working point.

Study	Without IBL	With IBL	Ratio
IBL-TDR	83±1.5	147±3.4	1.8
2015	84±0.5	363±4.2	4.3

5. Search for $WH \to \ell\nu b\bar{b}$

This section describes a search for the Standard Model (SM) Higgs boson produced in association with a leptonically decaying W boson ($WH \to \ell\nu b\bar{b}$, $\ell = e, \mu$). The result was published [90] and presented [91] together with other searches for Higgs boson production in association with a Z boson ($ZH \to \nu\nu b\bar{b}$ and $ZH \to \ell\ell b\bar{b}$, $\ell = e, \mu$). The term lepton refers here to electron and/or muon.

As described in chapter 1, an inclusive search for $H \to b\bar{b}$ is not feasible due to the overwhelming multijet background coming from strong interactions in pp collisions. The leptonic decays of the vector bosons (W/Z) are therefore used for triggering and background rejection purposes.

The author worked for the $VH(H \to b\bar{b})$ search mentioned above, particularly being in charge of the $WH \to \ell\nu b\bar{b}$ channel, with two significant contributions described below. The data and simulated samples used for this analysis are explained in section 5.1, followed by section 5.2 which describes the triggers in this analysis. The object selection and the event selection are defined in section 5.3 and 5.4, respectively. Section 5.5 illustrates the multi-variate analysis method. The background estimation strategies are described in section 5.6, including the author's significant contribution to the multijet background estimation. Another and unique author's contribution to this analysis, the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel in the WH analysis, is presented in section 5.7. The systematic uncertainties and their impact on the final fit results are discussed in sections 5.8–5.10. At the end in section 5.11, a conclusion drawn from the results is given.

5.1. Data and simulation samples

Data used for this analysis were collected in the year 2012 at a center-of-mass energy of 8 TeV, corresponding to an integrated luminosity of 20.3 fb^{-1}. The published result [90] is presented in combination with the result from data collected in the year 2011 at a center-of-mass energy of 7 TeV, corresponding to an integrated luminosity of 4.6 fb^{-1}.

5.1.1. Data taking and cleaning

Using a good-run-list as described in section 2.1.4, luminosity-blocks with some sub-detectors in bad conditions are removed from this analysis.

During the data taking and also during the various analyses, troubles on sub-detector parts are reported. Removal of these problematic events are called cleaning process, as described in the following items:

- Cleaning for poorly reconstructed jets.
  Occasionally, jets associated to non-collision-originated energy deposits in the calorimeter are reconstructed. This is due to many sources such as detector noise or problems, LHC beam conditions and cosmic-ray showers. Events containing such poorly reconstructed jets are removed.

- Corrupted tile veto.
  The corrupted data rejection algorithm was not perfect and some events were found to contain fake 400 GeV energy deposit in a particular channel of the tile calorimeter. The events with such a corrupted data are removed.

- Hot tile cell cleaning.
  In the data taking period B, there was a hot tile calorimeter cell that had not been masked in the reconstruction process. Events containing jets pointing to such hot tile cell are removed.

5.1.2. Simulated samples

The ATLFAST-II simulation [43] is used to produce samples for signal and background processes. This fast simulation is based on parametrized simulation for the calorimeter responses studied using the the full Monte Carlo (MC) detector simulation with the GEANT4 program [57], while the full simulation is used for the tracking and muon systems. The simulated signal and background samples and event generators used are listed in table 5.1.

Table 5.1: Event generators used for the simulation of the signal and background processes.

Process	Event generator
Signal	
$q\bar{q} \to WH \to \ell\nu b\bar{b}$	PYTHIA8
$q\bar{q} \to ZH \to \ell\ell b\bar{b}$	PYTHIA8
$gg \to ZH \to \ell\ell b\bar{b}$	POWHEG+PYTHIA8
$V+$ jets	
$W \to \ell\nu$	SHERPA 1.4.1
$Z/\gamma* \to \ell\ell$	SHERPA 1.4.1
$Z \to \nu\nu$	SHERPA 1.4.1
Top-quark	
$t\bar{t}$	POWHEG+PYTHIA
t-channel single top	ACERMC+PYTHIA
s-channel single top	POWHEG+PYTHIA
Wt-channel single top	POWHEG+PYTHIA
Diboson	
WW	POWHEG+PYTHIA8
WZ	POWHEG+PYTHIA8
ZZ	POWHEG+PYTHIA8

WH signal samples

The signal WH sample is generated using PYTHIA8 [45] with the CTEQ6L1 [92] parton distribution functions (PDFs). The parton shower, hadronization, and multiple parton interactions are simulated using the AU2 tune [93, 94]. The PHOTOS program [95] is used for QED final-state radiation. For $q\bar{q}$ initiated WH production, the total production cross-section and associated uncertainties are computed at next-to-next-to-leading order (NNLO) in QCD [96], and with electroweak corrections at next-to-leading order (NLO) [97]. Additional normalization-preserving differential electroweak NLO corrections are applied as a function of the transverse momentum of the vector boson [98]. The branching ratio calculation of $H \to b\bar{b}$ has been done using HDECAY [99]. Within the range from 100 GeV to 150 GeV with intervals of 5 GeV, the different masses of the Higgs signal process are generated. All the leptonic decays of W bosons are considered, as leptonic decays of τ leptons can also be selected in this analysis.

Background samples

The main background processes are W+jets and $t\bar{t}$ production. SHERPA 1.4.1 [51] is used with the CT10 [100] PDFs to simulate (W/Z)+jets at leading order in QCD, with massive b- and c-quarks. To improve the statistical precision in the most sensitive (high p_{T}^{W}) regions, more events with heavy flavor content or large vector boson p_{T} are generated. For $t\bar{t}$ samples, the POWHEG [52–54] generator is used with the CT10 PDFs, interfaced with PYTHIA6 [44]. In this analysis, the final normalization of these dominant backgrounds are constrained by the data, but theoretical cross-sections are used to optimize the selection. The cross-sections are calculated at NNLO for (W/Z)+jets [101]. The $t\bar{t}$ production cross-section is also calculated at NNLO including re-summations of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [102].

Other backgrounds are single top quark production and diboson production (WW, WZ, and ZZ). For single top quark production, the s-channel exchange process and Wt production are simulated with POWHEG, while the t-channel exchange process is simulated with the ACERMC [49] generator interfaced with PYTHIA6, using the CTEQ6L1 PDFs and the Perugia2011C tune [93, 94]. The cross-sections used are from refs [103–105]. The POWHEG generator interfaced with PYTHIA8 with CT10 PDFs and the AU2 tune is used for diboson processes [106].

The PYTHIA8 generator is used to simulate events from minimum-bias interactions with MSTW 2008NLO PDFs [107] and the A2 tune [93, 94]. They are overlaid on the simulated signal and background events according to the luminosity profile of the recorded data. These are so-called "pileup" interactions, and simulated both within the same bunch crossing as the hard-scattering process and in neighboring bunch crossings. These resulting events are then processed through the same reconstruction programs as the data.

V+jet event categorization

Simulated jets are labeled using ΔR association with hadrons with $p_{\mathrm{T}} > 5$ GeV, with a cone radius of 0.4. For this association, the priority is given to b-hadron, c-hadron, and taus in this order. If none of these particles are found within the jet cone, the jet is labeled as light jet, thus corresponding to a jet originating from u, d, s, gluon, or pileup interaction. Simulated V+jet ($V = W$ or Z) events are categorized according to the labels of the two jets that are used to reconstruct the Higgs boson candidate. If one of those jets is a b-jet, the event is categorized as Vb. If not and one of the jets is a c-jet, the event is categorized as Vc. Otherwise, the event is categorized as Vl. Further sub-divisions are defined according to the flavor of the other jet from the pair, using the same precedence order: Vbb, Vbc, Vbl, Vcc, Vcl.

Corrections to the simulation

The MC simulation does not always describe data precisely. For some quantities, which have been found to differ from data, corrections were applied as scale factor (SF): the basic idea is to apply the $x(\mathrm{data})/x(\mathrm{MC})$ ratio to the MC simulation, where x is the observed number of events or efficiency in a particular bin of a variable.

In order to correct the difference between the mean number of interactions per bunch crossing observed in data and the predefined values in simulation, a correction, called pileup reweighting, is applied. Another weight is also applied to correct for the simulated beam-spot size along the z axis, so as to match those observed in data.

The electron energies are calibrated in data by making use of reference processes such as $Z \to ee$, so as to reproduce those observed in the simulation. Additional corrections are applied to the simulated electrons and muons to fix small differences in the trigger efficiencies, identification efficiencies including isolation requirement, and the energy and momentum resolutions, so as to match the ones observed in data.

A correction has been applied to the W+jet background sample due to a mis-modeling. This correction is based on the $\Delta\phi(\text{jet}_1, \text{jet}_2)$ and described in section 5.6.2.

As described in the previous section, the b-tagging efficiency difference in data and simulation is corrected in the simulation samples.

5.2. Triggers

Triggers used for this analysis are mainly single lepton triggers (SLT), with additional E_T^miss triggers (XET) to recover the events missed by the single muon trigger. Table 5.2 summarizes the triggers used for this analysis and details are as follows.

Table 5.2: Triggers used for the analysis. Single electron and muon triggers are used, in combination of low and high p_T threshold triggers together with and without online track isolation requirement, respectively. The E_T^miss trigger used is a combination of several triggers, depending on the data taking period.

Event filter menu	online object	EF $p_\text{T}(E_\text{T})$ threshold [GeV]	online track isolation
`e24i_vh_medium1`	electron	24	yes
`e60_medium1`	electron	60	no
`mu24i_tight`	muon	24	yes
`mu36_tight`	muon	36	no
`xe80_tclcw`	E_T^miss	80	–
`xe80T_tclcw_loose`	E_T^miss	80	–
`xe80_tclcw_loose`	E_T^miss	80	–

SLT consists of single electron triggers (SET) and single muon triggers (SMT). SET has two event filter (EF) menus of `e24i_vh_medium1` and `e60_medium1` with minimum p_T threshold of 24 and 60 GeV, respectively. As well as SET, SMT has two EF menus of `mu24i_tight` and `mu36_tight` with minimum p_T threshold of 24 and 36 GeV, respectively. In order to reduce the EF trigger rate, relative track isolation is required for `e24i_vh_medium1` (`mu24i_tight`) in addition to the p_T threshold: the isolation requirement is defined as $\sum_{\Delta R<0.2} p_\text{T}^\text{trk}/p_\text{T} < 0.1(0.12)$. The SET and SMT efficiencies are measured using a *tag-and-probe* method applied to $Z \to ee$ and $Z \to \mu\mu$ events, respectively [108, 109]. Small differences in the data and simulation efficiencies are corrected by applying data/simulation scale factors to the simulated samples.

XET consists of three EF menus: `xe80_tclcw`, `xe80T_tclcw_loose`, and `xe80_tclcw_loose`. These three EF menus are used depending on the reconstructed E_T^miss range and data taking period. The usage and the chains of XET are outlined in tables 5.3 and 5.4. This complex usage comes from

Table 5.3: The XET usage for this analysis.

Event filter menu	data taking period in 2012	E_T^miss range	luminosity [fb^{-1}]
`xe80_tclcw`	A–B5	$E_\text{T}^\text{miss} > 160$ GeV	1.91
`xe80T_tclcw_loose`	A–B5	$E_\text{T}^\text{miss} < 160$ GeV	2.12
`xe80_tclcw_loose`	B6–L	any	18.1

Table 5.4: The XET chains from level-1 to EF. The right column BGRP7 is to skip the first three bunches in the bunch train. The numbers for the thresholds are given in GeV.

Event filter menu	level-1 threshold	level-2 threshold	EF threshold	BGRP7
`xe80_tclcw`	50	55	80	no
`xe80T_tclcw_loose`	40	45	80	yes
`xe80_tclcw_loose`	40	45	80	no

the efforts of maximization of acceptance in data taken in the year 2012. In the early periods in 2012, a mixture of two $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers was used depending on the $E_{\mathrm{T}}^{\mathrm{miss}}$ range: `xe80_tclcw` and `xe80T_tclcw_loose`. The trigger `xe80T_tclcw_loose` is designed to skip the first three bunches in the bunch train, in order to reduce the level-1 trigger rate. Thanks to the successful level-1 calorimeter noise suppression[a] in the forward region ($|\eta| > 2.5$) of the calorimeter, `xe80_tclcw_loose` is activated, starting from period B6.

Since the main target of XET is the search for $ZH \to \nu\nu b\bar{b}$, the efficiency dependence of XET on offline reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ is measured with a loosened $ZH \to \nu\nu b\bar{b}$ selection [74], using $W \to \mu\nu$+jets and $Z \to \mu\mu$+jets events collected with SMT, parametrized with $E_{\mathrm{T}}^{\mathrm{miss}}$ without muon momentum contribution, as detailed in section 5.7. Under this condition, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is fully efficient for $E_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV, 97% efficient for $E_{\mathrm{T}}^{\mathrm{miss}} > 120$ GeV. Important selections in this XET efficiency measurement condition are included in the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered $WH \to \mu\nu b\bar{b}$ channel in order to apply the same $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger SF.

5.3. Object selection

In physics analyses with the ATLAS experiment, the reconstructed particles in the events are stored as *candidates*, which are selected with the loosest level of identification requirements in order to maximize the acceptance. From these candidates, the reconstructed *objects* are chosen with tight requirements.

In the VH analysis, the lepton objects are classified in three levels, *loose, medium*, and *tight* and jets are classified in two categories of *forward* and *central*. For the multijet (MJ) background estimation using a data driven method, loosely isolated leptons are used and defined as MJ-leptons. Such leptons form the MJ control region (CR). Although the definitions of those MJ-leptons are also introduced here, the details of the MJ estimation method are discussed in section 5.6.1.

Primary vertex

Charged particles are reconstructed with a p_{T} threshold of 400 MeV. Due to the pileup interactions, an event has multiple primary vertices (PVs). Among them, the vertex with leading $\sum_{\mathrm{track}} (p_{\mathrm{T}}^{\mathrm{track}})^2$ is selected as the actual PV and used in this analysis, where $p_{\mathrm{T}}^{\mathrm{track}}$ is p_{T} of a track associated to the vertex.

As described in section 3.1, when the vertexing algorithm finds no vertex candidate, the beam-spot is assigned to the PV position. In order to avoid to use such default beam-spot or poorly reconstructed PVs, selected PV in the event is required to be associated with more than 2 tracks.

---

[a] The tower threshold to be considered for $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is raised from 1.2 GeV to 10 GeV, due to the strong dependence on the pileup interactions in the forward region of the calorimeter.

The lepton categories are prepared for two purposes: to select the high quality lepton object and to reject additional lepton objects. The tight electrons and muons are defined for the former purpose, and the loose electrons and muons are for the latter. The selection criteria for these objects are detailed below.

Electrons

Loose electrons are selected with $E_{\mathrm{T}} > 7$ GeV, $|\eta| < 2.47$ and requirements listed below.

- Reconstruction is seeded by cluster-based algorithm or track-based algorithm described in section 3.2.1.
- No bad cell exists in the calorimeter energy cluster.
- The 'very-loose likelihood' requirement [32] is fulfilled, which combines shower-shape information, track-quality criteria, the matching quality between track and associated energy-cluster in the calorimeter, TRT information, and a criterion to help identify electrons originating from photon conversions.
- Track isolation is required to be less than 0.1, where track isolation is defined as the ratio of the scalar p_{T} sum of additional tracks within a cone of radius $\Delta R = 0.2$ centered on the electron track to the electron E_{T}.

Among loose electrons, the ones with $E_{\mathrm{T}} > 25$ GeV are selected as medium electrons. Furthermore, tight electrons are selected from medium ones with the following requirements.

- 'very-tight likelihood' requirement [32].
- Track isolation defined above is required to be less than 0.04.
- Calorimeter isolation is required to be less than 0.04, where calorimeter isolation is defined as the ratio of the scalar E_{T} sum of additional calorimeter clusters within a cone of radius $\Delta R = 0.3$ around the electron to the electron E_{T}.

MJ control electrons are selected by following the loose selection but changing the isolation and identification selections as follows:

- 'medium++' requirement [110] in addition to very loose likelihood requirement used for loose selection.
- The track isolation requirement defined above is loosened to be in between 0.05 and 0.5.
- The calorimeter isolation requirement defined above is loosened to be below 0.07.

Electron MJ events in CR must be triggered by `e24i_vh_medium1` since the isolation range for MJ electron spans across the isolation requirement made at the trigger level.

Muons

In order to maximize the acceptance, four types of loose muons are employed:

(1) Combined muons: muons reconstructed in both the Muon Spectrometer and the Inner Detector.

(2) Segment-tagged muons: muons reconstructed in the Inner Detector whose tracks are matched to segments in the Muon Spectrometer.

(3) Standalone muons: muons identified in the MS with no associated track in the Inner Detector

(4) Calo-muons: Calo-muons are identified with ID tracks with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 0.1$ where there is a lacking MS coverage, if their calorimeter energy deposits are consistent with a minimum ionizing particle.

Combined, segment-tagged, and standalone muons are required to have $p_{\mathrm{T}} > 7$ GeV, $|\eta| < 2.7$ and to fulfill the 'tight track fit quality' criteria [111]. Since combined, segment-tagged, and calo-muons are reconstructed with ID track, the following selection on the tracks are applied.

- Inner Detector track hit requirements defined in section 3.3.
- $|d_0| < 0.1$ mm and $|z_0| < 10$ mm, where d_0 and z_0 are the transverse and longitudinal impact parameters, respectively, as defined in section 3.1.1. This cut is used to reject muons from pileup interactions and cosmic muons passing through the ATLAS detector coincidentally.
- Track isolation is required to be less than 0.1, where track isolation is defined as the ratio of the scalar p_{T} sum of additional tracks within a cone of radius $\Delta R = 0.2$ centered on the muon track to the muon p_{T}.

Among combined and loose muons, medium muons are selected with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. Furthermore, tight muons are chosen out of medium ones with the following isolation requirements.

- Track isolation defined above is required to be less than 0.04.
- Calorimeter isolation is required to be less than 0.04, where calorimeter isolation is defined as the ratio of the scalar E_{T} sum of additional calorimeter clusters within a cone of radius $\Delta R = 0.3$ around the muon to the muon p_{T}.

MJ control muons are selected by loosening the selection, changing isolation selections as below:

- The track isolation requirement defined above is loosened to be in between 0.05 and 0.5.
- The calorimeter isolation requirement defined above is loosened to be below 0.07.

Figure 5.1 summarizes the isolation criteria for loose, tight, and MJ-leptons.

Loosely selected leptons

The author also worked for a study to gain sensitivity by including additional signal regions using loosely selected leptons. Such leptons were classified into three categories, according to their p_{T} and track isolation criteria:

- $15 < p_{\mathrm{T}} < 25$ GeV and track isolation smaller than 0.04.
- $p_{\mathrm{T}} > 25$ GeV and track isolation greater than 0.04 but smaller than 0.1.
- $15 < p_{\mathrm{T}} < 25$ GeV and track isolation greater than 0.04 but smaller than 0.1.

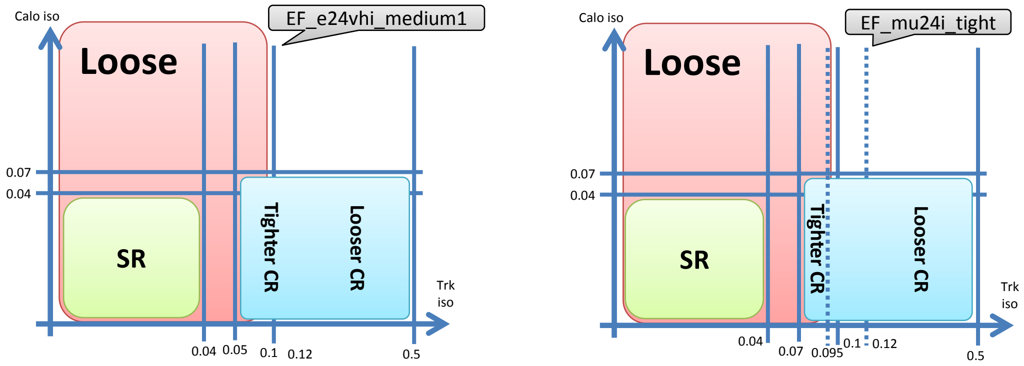


Fig. 5.1: The lepton isolation requirement is shown for electrons (left) and muons (right) in the iso_{calo} versus iso_{track} plane. The red box shows the region used for loose leptons and the green box for the tight lepton region. The blue box represents the region used for the MJ-lepton selection. One can clearly see that only some multijet leptons satisfy the loose lepton criteria. For electrons, the Tight CR is used for the nominal template and the loose for the systematic variation. For muons, the inclusive CR is used for the nominal while Tight and Loose are used individually for the Up and Down systematic variation of the iso_{track} cut.

The sensitivity gain using these regions were studied and improvements over tight-lepton-only analysis was seen when these categories were added. However, the gains from these three categories were very small compared to the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel. Thus, a decision was taken that the time and efforts should be focused on the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel, and to revisit these loosely selected leptons in Run 2. This study is summarized in appendix C.

Jets

Jets used in this analysis are reconstructed by collection of noise-suppressed topological clusters of energy in the calorimeters [61], using the anti-k_t algorithm with a cone size parameter of 0.4. In order to suppress contamination of jets from pile-up interactions, the jet vertex fraction (JVF) is required to be at least 50%. As described in section 3.5, for a given jet, JVF is defined as the ratio of scalar p_{T} sum of tracks associated to both the jet and PV to the scalar p_{T} sum of tracks associated to the jet. This JVF requirement is applied only to jets with $p_{\mathrm{T}} < 50$ GeV and $|\eta| < 2.4$. Jet energies are corrected for the contribution of the pileup interactions using a jet-area based technique [112]. These corrections are calibrated using p_{T}- and η-dependent correction factors determined from simulation, with residual corrections from in-situ measurements applied to data [113]. Further adjustments are performed using jet internal properties, which improve the energy resolution without changing the average calibration (called global sequential calibration [113]). Due to the b-tagging application, in which the tracks reconstructed by the Inner Detector are necessary, jets are selected with $|\eta| < 2.5$ (central jets). For the central jet, the minimum p_{T} threshold is 20 GeV. Forward jets, selected with $p_{\mathrm{T}} > 30$ GeV and $2.5 < |\eta| < 4.5$, are used for reduction of $t\bar{t}$ background events. The precise usage of forward jets is explained in section 5.4.

b-jets

For the central jets, a b-tagging requirement is applied using the MV1c tagger described in section 4.4.3. Thanks to the newly developed calibration method, so-called "continuous" calibration [88], the b-tagging category for a given jet is classified into three levels depending on the output weight of the MV1c tagger. In this analysis, $loose$, $medium$, and $tight$ working points are used, which correspond to the 80%, 70%, and 50% efficiency working points measured in $t\bar{t}$ simulated sample, respectively. For the tight (loose) operating point, the rejection factors are 26 (3) and 1400 (30) against c- and light-jets, respectively. Jets satisfying loose b-tagging requirement are called $tagged$ jets, while those not satisfying it are called *un-tagged* ones. The greatest benefit from the "continuous" tagging is that the entire $MV1c$ distribution is calibrated so that the tagging weight itself can be used as an input to the MVA. The calibration of the tagging weight is performed in intervals of working points, thus the $MV1c$ value falls in one of the five bins (100%–80%, 80%–70%, 70%–60%, 60%–50%, and 50%–0%). In the MVA, BDT takes this binned $MV1c$ value as input variable for the two central jets.

Overlap removal

To avoid double counting, an overlap removal procedure is applied among loose leptons and jets in the following order.

(1) Forward and central jets are discarded if a loose electron exists around jets within $\Delta R = 0.4$ around the jet.

(2) Forward and central jets are removed if the jet has less than 4 associated tracks and a loose muon is found within $\Delta R = 0.4$ around the jet. In this case, a muon is likely to have showered in the calorimeter. If the jet has more tracks, the loose muon is removed.

(3) Remaining loose electrons with $\Delta R < 0.2$ to a loose muon are removed, except if the muon is identified only in the calorimeter, in which case the muon is removed.

After this overlap removal procedure, the remaining objects in the event are used for the analysis.

Missing transverse momentum

The magnitude of missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$, is reconstructed as the negative vector sum of the transverse momenta of calorimeter energy clusters within $|\eta| = 4.9$. As described in section 3.6, the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation takes into account the corrections for the reconstructed objects like jets, electrons, tau-jets, and photons, using the calibrated energies of these objects. These objects are contributing to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation after a proper overlap removal defined above.

Additionally, a missing transverse momentum reconstructed without muon momenta, $E_{\mathrm{T}}^{\mathrm{miss}}(\not{\mu})$, is employed for the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel.

5.4. Event selection

This analysis is optimized for $WH \to \ell\nu b\bar{b}$ with $m_H = 125$ GeV. The search is carried out with two strategies, multi-variate analysis (MVA) and cut-based analysis. Unless specifically described, the selection detailed here is common to the MVA and cut-based analyses. All the cleaning selection described in section 5.1.1 are properly applied to all data.

The events must have exactly one tight lepton with no additional loose lepton. The selected tight lepton must be matched with the online object[b] which fired the SLT. In the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel, this 'trigger matching requirement' is eliminated.

Events with forward jets are discarded for rejection of $t\bar{t}$ background which is likely to have forward jets due to its high jet multiplicity. The analysis is first categorized into 2- and 3-jet bins, according to the number of central jets in the event. In the 3-jet bin, the additional jet with lowest $p_{\mathrm{T}}^{\mathrm{jet}}$ is allowed only when it is un-tagged. The jet with highest p_{T} (leading jet) in the event must have $p_{\mathrm{T}}^{\mathrm{jet}} > 45$ GeV.

This analysis has the following b-tagging categories depending on the two central jets' MV1c weights: these are LL, MM and TT 2-tag regions, 1-tag region, and the 0-tag region shown in Fig 5.2. TT events have two tagged jets, where both jets satisfy the tight tagging requirement. MM (LL) events have two tagged jets, where both jets pass the medium (loose) tagging requirement, but at least one of two jets fails the tight (medium) tagging requirement. Events with one jet passing the loose tagging requirement but the other not passing the loose requirement form the 1-tag category. Events with both central jets not passing the loose tagging requirement form the 0-tag category. This b-tagging categorization improves the sensitivity with respect to what would be obtained using a single category, such as TT+MM, with the LL category providing constraints on the backgrounds not containing two real b-jets.

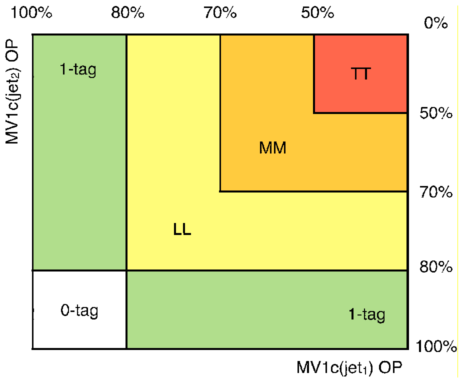


Fig. 5.2: A sketch of b-tagging categories in this analysis.

The analysis is further categorized according to the transverse momentum of the W boson, p_{T}^{W}, to take advantage of better signal-to-background ratio at high p_{T}^{W}. p_{T}^{W} is reconstructed as the vector sum of the tight lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$. Different categorization on p_{T}^{W} is defined for MVA and cut-based as detailed in this section below.

Additional topological and kinematic selections are applied for background rejection and sensitivity enhancement. An event must have $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV only when $p_{\mathrm{T}}^{W} > 120$ GeV. Events with $p_{\mathrm{T}}^{W} < 120$ GeV are required to have $H_{\mathrm{T}} > 180$ GeV, where H_{T} is defined as the scalar sum of tight lepton p_{T}, p_{T} of the two central leading jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$. This reduces mainly the MJ background. The definition of H_{T} is the same also for the 3-jet bin. These selections are outlined in table 5.5. An angular separation between the two jets of the dijet system, ΔR_{jj}, is required to be above 0.7 for reduction of W+jet background contamination. To increase the signal acceptance, this requirement is removed in the highest p_{T}^{W} bin, where the background amount is the smallest.

The MVA has looser selection than cut-based analysis in order to maximize the information available to the final discriminant.

---

[b] Online objects are the objects used for triggering described in section 5.2. These objects are not using the full granularity and/or full accuracy of the detector.

MVA selection

In the MVA selection, events are categorized in only two p_{T}^{W} intervals with a boundary at 120 GeV, but the detailed p_{T}^{W} information is used in the final discriminant. These regions are called low and high p_{T}^{W} regions. Although an upper limit for ΔR_{jj} is not set for the MVA selection, this information is used in the final discriminant.

Cut-based selection

In the cut-based selection, events are categorized into five p_{T}^{W} intervals, with boundaries at 90, 120, 160, and 200 GeV. The W transverse mass, m_{T}^{W}, is reconstructed using the p_{T} of the tight lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$,

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

and required to be smaller than 120 GeV when $p_{\mathrm{T}}^{W} < 120$ GeV. This requirement reduces contamination from the $t\bar{t}$ background. When an event has $p_{\mathrm{T}}^{W} > 200(160 < p_{\mathrm{T}}^{W} < 200)$ GeV, the event must have $E_{\mathrm{T}}^{\mathrm{miss}} > 50(20)$ GeV. The ΔR_{jj} range limits are set depending on p_{T}^{W} to take advantage of the increasing collimation of the dijet system, especially for $t\bar{t}$ background reduction. These range limits are summarized in table 5.5.

Other than these selections, additional selections are applied for the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel, in order to apply the same trigger efficiency scale factor as in the $ZH \to \nu\nu b\bar{b}$ analysis. The details are explained in section 5.7.

5.5. Multivariate analysis

Although the dijet mass (m_{bb}) provides the best discrimination between signal and background, the sensitivity is improved by making use of additional kinematic, topological and b-tagging properties of the selected events in MVA. In this analysis, Boosted Decision Trees (BDT) implemented in the TMVA [78] package are employed for the MVA method, properly taking into account the correlations between the input variables.

The BDTs are trained to separate the $WH,\ H \to b\bar{b}$ signal from the sum of the expected background processes. The input variables are chosen in order to maximize the separation, while avoiding the use of variables not improving the performance significantly. The variable selection starts from the dijet mass. The additional variables are tried one at a time and the one yielding the best separation gain is kept. The final set of variables is obtained by repeating this procedure until finding a variable not resulting in a significant performance gain.

Table 5.5: Summary of kinematic requirements in the cut-based analysis and MVA.

	cut-based					MVA	
p_{T}^{W} bin	0	1	2	3	4	0	1
p_{T}^{W} [GeV]	0–90	90–120	120–160	160–200	> 200	0–120	> 120
ΔR_{jj}	0.7–3.4	0.7–3.0	0.7–2.3	0.7–1.8	< 1.4	< 0.7 (when $p_{\mathrm{T}}^{W} < 200$ GeV)	
m_{T}^{W} [GeV]	< 120			–	–	–	–
H_{T} [GeV]	> 180			–	–	> 180	–
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	–			> 20	> 50	–	> 20

For $WH \to \ell\nu b\bar{b}$ with $m_H = 125$ GeV, four BDTs are built in low and high p_{T}^{W} bins in each of the 2- and 3-jet bins. This set of four BDTs is also built for the mass scan of the Higgs boson and diboson cross-check analysis, by just changing the signal sample in the training from $m_H = 125$ GeV to other masses or diboson sample. Trainings are performed in combined 2-tag signal region (LL+MM+TT) and also in combination of electron and muon channels since none of the input variables are lepton-flavor specific. The final selection of 11 or 13 input variables for the BDTs is:

- m_{T}^{W}
- p_{T}^{W}
- $E_{\mathrm{T}}^{\mathrm{miss}}$
- m_{bb}
- $\Delta R(b_1, b_2)$
- $\Delta\phi(W, bb)$
- $\min[\Delta\phi(\ell, b)]$
- $p_{\mathrm{T}}^{b_1}$
- $p_{\mathrm{T}}^{b_2}$
- $MV1c^{b_1}$
- $MV1c^{b_2}$
- $p_{\mathrm{T}}^{j_3}$ (only for 3-jet bin)
- m_{bbj} (only for 3-jet bin)

$p_{\mathrm{T}}^{b_1}$ and $p_{\mathrm{T}}^{b_2}$ are the p_{T} of the b-tagged leading and sub-leading jets, and m_{bb} is the invariant mass of the dijet system of the two b-tagged jets. The angular separation $\Delta R(b_1, b_2)$ is the distance between the two b-tagged jets, $\Delta\phi(W, bb)$ is the angular separation in the transverse plane between the W and dijet system, $\min[\Delta\phi(\ell, b)]$ is the angular separation in the transverse plane between the tight lepton and closest b-tagged jet. $MV1c^{b_1}$ and $MV1c^{b_2}$ are the MV1c b-tagging weight for the leading and sub-leading b-tagged jets. For the 3-jet bin, additional variables are used. They are $p_{\mathrm{T}}^{j_3}$ and m_{bbj}, which are the p_{T} of the non b-tagged jet and the invariant mass of the tri-jet system.

The input variables of the BDTs are compared between data and simulation and good agreement is found within the assigned uncertainties. Selected input-variable distributions are shown in appendix B, and BDT output distributions in the 1-tag control regions are shown in fig 5.3. In this figure, as for all figures in the following, the MJ background is estimated as described in section 5.6.1, the other simulated samples are estimated with the corrections explained in section 5.6.

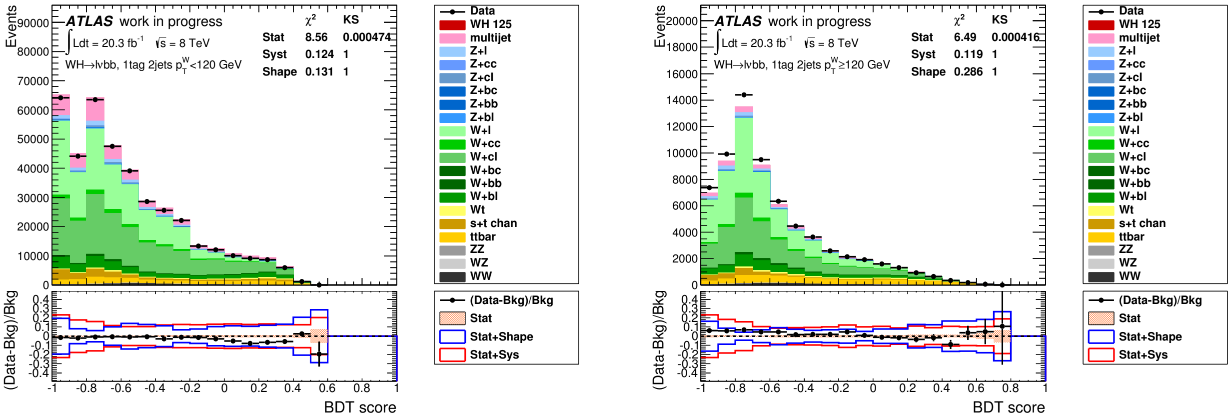


Fig. 5.3: Pre-fit BDT output distributions in the 1-tag control region in the low p_{T}^{W} (left) and high p_{T}^{W} regions. Events selected are triggered by SLT.

5.6. Background estimation

This section describes the modeling of the three main background components in this analysis, which are the multijet, $t\bar{t}$, and W+jet backgrounds. For the multijet background, a data-driven method is used, while the other dominant $t\bar{t}$ and W+jet backgrounds are estimated with improved MC simulations using dedicated methods which are explained below.

5.6.1. Multijet background

The multijet (MJ) events arise due to QCD interaction in pp collisions which has a huge production cross-section. These events produce fakes of the signal object(s) for the analysis such as electrons, muons and $E_{\mathrm{T}}^{\mathrm{miss}}$. Although these events have small fake rates, production rates are huge, resulting in a non-negligible background contribution in the analysis. The estimation of MJ events using MC is impractical, since the fake processes of the MJ events are very difficult to model precisely and it is not feasible to produce enough simulated events due to the very low fake rate. Therefore, the MJ background contribution is estimated using data directly.

5.6.1.1. Origin of the multijet background

As described in section 3, the identification principles of electrons and muons are very different, thereby the origins of the electron- and muon-fakes are also different. Below are the possible processes which can contribute for the electron-fake signature in the detector:

- An electron stemming from photon-conversion, where the other electron has very soft momentum and is lost from the reconstruction, leaves the electron signature in the TRT and the EM calorimeter.
- A real electron stemming from semi-leptonic decay of the b- or c-hadrons, while the other decay products are soft, thereby the electron leaves an isolated signature.
- Non-electron backgrounds such as photons, hadrons or jets which leave electron-like signatures in the detector.

Since the last case has its origin in hadron or jet originated processes, the electron-fake yields more than the muon-fake case in a hadron collider experiments.

The possible muon-fake sources are listed below:

- So-called sail-through of hadrons (mainly $\pi^{\pm}$): charged hadrons go trough the calorimeter and cryo-system and leave muon-like signatures in the MS.
- So-called punch-through of hadrons (mainly $\pi^{\pm}$): charged hadrons go deeper in the calorimeter or cryo-system and interact with the materials, creating the charged particles which can leave muon-like signatures in the MS.
- Pion or kaon decays in flight: a real muon stemming from a decay in the tracking volume when the neutrino carries out a very soft momentum fraction.
- A real muon stemming from semi-leptonic decays of the b- or c-hadrons, as for the electron-fakes.

5.6.1.2. Estimation method

Among several ways to estimate the MJ contribution in the signal region, a variant of the *jet electron* method [114], so-called *multijet fit* method, is used in this analysis. The assumption is made that shapes of any kinematic property of the multijet background in the signal region can be described by the ones in the control region, which is dominated by the MJ background, requiring a loosely isolated lepton in the event. The selection criteria for these loosely isolated leptons are described in section 5.3. Below shows the procedure for the template shape and normalization estimation by the fit method:

1. Obtain the shape of the multijet background by subtracting the electroweak (EW) background contribution from the data distribution in the control region. The fraction of the EW background contribution is supposed to be very small ($f_{\mathrm{EM}} \ll 1$), since the CR is dominated by the MJ background.

2. The shape obtained from 1, together with the EW background contribution, is used to fit the data $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the signal region, and the normalization scale factor (SF) for MJ background (MJ-SF) is obtained. Here the EW backgrounds are treated as one component, and both the EW and MJ components are floated, but the normalization SF for the EW component (EW-SF) is close to unity (typically, 0.90<EW-SF<1.1), thereby the EW-SF is not applied before the global fit.

For both the subtraction in the control region and the fitting in the signal region, the EW subcomponents are improved by scale factors obtained from a preliminary global fit. Considering the very different sources for the MJ background, the MJ fits are performed separately for electron and muon channels, as well as the 2- and 3-jet bins due to the very different event kinematic property (see fig 5.4). This fit is performed inclusively in the p_{T}^{W} regions but separately in each b-tagging category. However, the combined 2-tag category does not have enough statistics for the control region, thereby the 1-tag category template is added to the combined 2-tag category template due to the similar kinematic properties for the 1-tag and 2-tag categories.

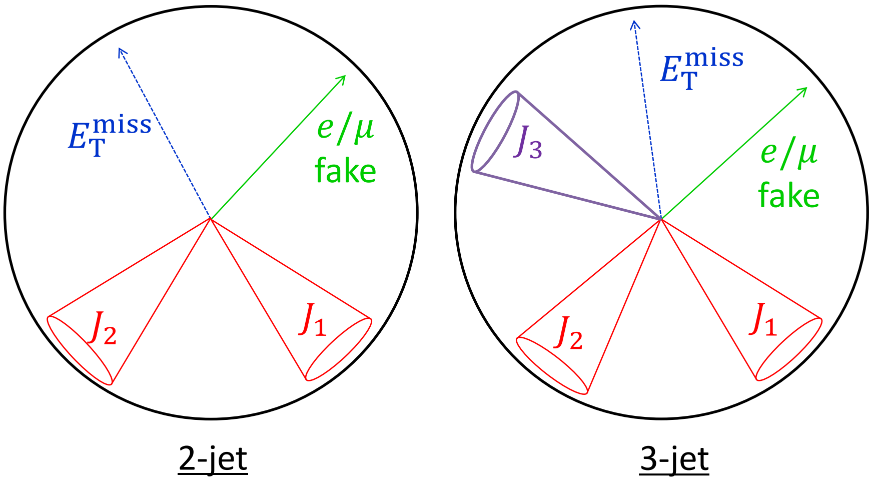


Fig. 5.4: The topology for 2(3)-jet event in the transverse plane in the left (right) sketch.

5.6.1.3. Pseudo 2-tag template using 1-tag events

In order to estimate the BDT distribution shape for the MJ events in the 2-tag regions, events in the 1-tag category are promoted to the 2-tag category by assigning an emulated $MV1c$ value to the un-tagged jet, drawn from the appropriate $MV1c$ distribution observed in the corresponding 2-tag MJ template. This $MV1c$ distribution depends on the order in $p_{\mathrm{T}}^{\mathrm{jet}}$ (leading or sub-leading) of the un-tagged jet and on the $MV1c$ value of the tagged jet. To cope with residual differences observed in some distributions between these pseudo 2-tag and the actual 2-tag templates, a reweighting is applied according to the $MV1c$ value of the tagged jet, and for each electron channel, according to $\Delta R(\text{jet}_1, \text{jet}_2)$ and p_{T}^{W}. This procedure is applied in each of 2- and 3-jet, LL, MM, and TT categories.

5.6.1.4. Multijet fit result

Figures 5.5 and 5.6 show the MJ template subtraction in the control region and the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution after the MJ fit in the signal region in 2- and 3-jet bins, respectively. As can be seen, kinematics in 1-tag and 2-tag regions are similar.

Table 5.6 shows the obtained MJ-SF from the fit in the signal regions. After applying these SFs to the templates obtained from the subtraction in the control regions, the shapes and normalizations of the multijet background are obtained and used for the global fit. The MJ background in this analysis is concentrated at low p_{T}^{W}, and ranges in the 2-jet 2-tag sample with $p_{\mathrm{T}}^{W} < 120$ GeV from 11% of the total background in the LL category to 6% in the TT category. The main purpose to include the $p_{\mathrm{T}}^{W} < 120$ GeV intervals is to provide constraints on the largest backgrounds (W+jets and $t\bar{t}$) in the global fit. In the low p_{T}^{W} interval, the MJ background in the electron channel is twice as large as the one in the muon channel. Also, a mis-modeling was found in the low p_{T}^{W} electron channel, therefore, only the muon channel is kept for this p_{T}^{W} interval so that the most reliable constraints on the non-MJ backgrounds can be provided. The resulting loss on the upper limit on the Higgs boson signal strength at $m_H = 125$ GeV is 0.6%.

Table 5.6: The results of the multijet fit in the b-tagging categories in 2- and 3-jet bins, electron and muon channels, separately.

		electron		muon	
		EW	MJ	EW	MJ
2jet	pretag	0.98	0.62	0.97	1.60
	0-tag	0.98	0.62	0.98	1.55
	1-tag	0.95	0.58	0.93	1.67
	2-tag	0.99	0.052	0.93	0.21
3jet	pretag	0.91	0.83	0.97	2.29
	0-tag	0.91	0.84	0.89	2.46
	1-tag	0.90	0.79	0.89	2.06
	2-tag	0.98	0.06	0.94	0.18

5.6.2. Other backgrounds

The background processes of $t\bar{t}$ and W+jets events, which give the dominant contribution in this analysis, were floated freely in the global fit. Although the basic properties of these dominant backgrounds are estimated from the MC simulation, some dedicated corrections for these simulations are applied prior to the global fit in order to achieve a better description of the data.

$t\bar{t}$ background

From an unfolded measurement, it appeared that the POWHEG generator interfaced with PYTHIA predicts the p_{T} of the $t\bar{t}$ system to be too hard [115]. A correction accounting for this discrepancy is therefore applied at the level of generated top quarks in the $t\bar{t}$ production process.

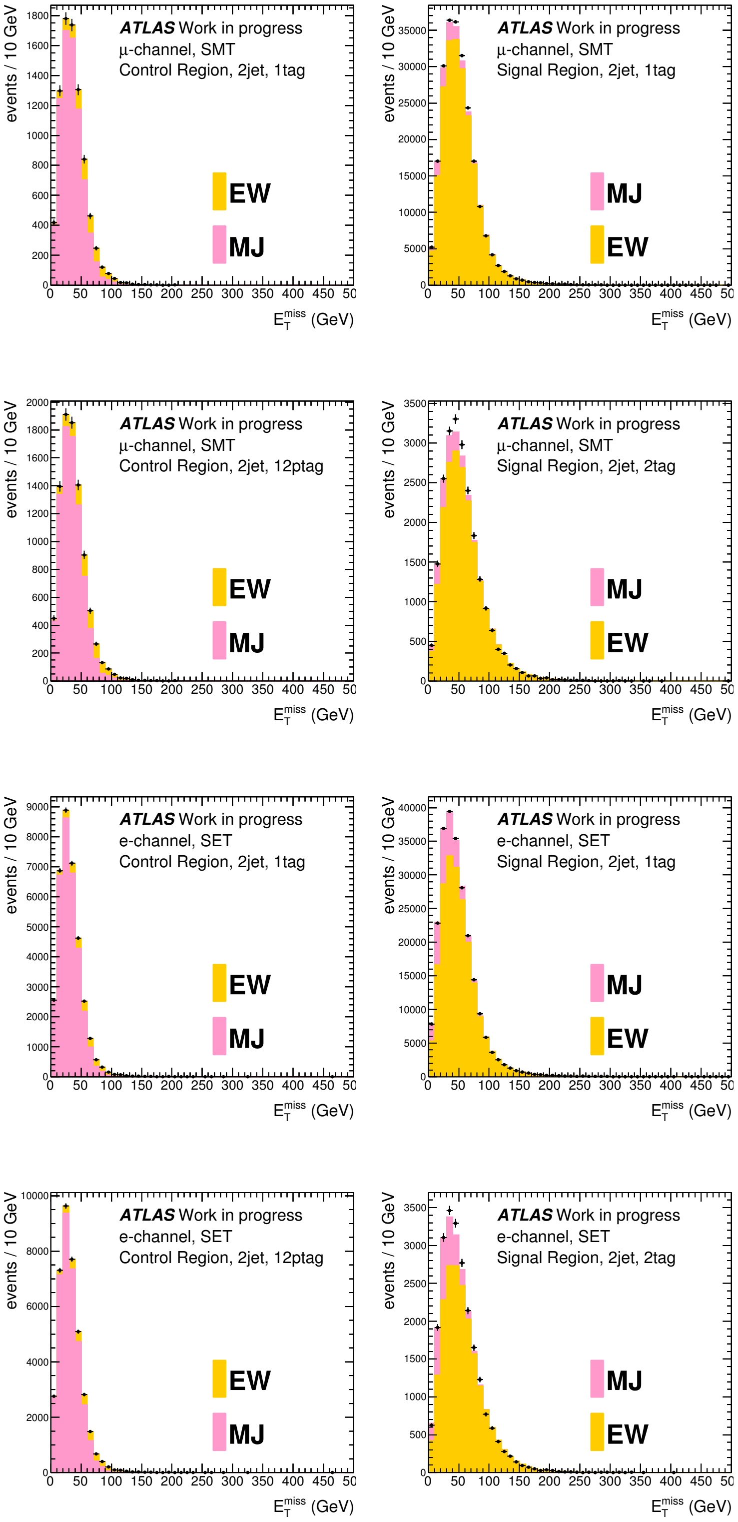


Fig. 5.5: The EW component subtraction to obtain MJ template in the control region (left) and the MJ fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the signal region (right). The $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions are shown in the 2-jet 1-tag and 2-jet 2-tag region muon (electron) channel in the first (third) and second (forth) row, respectively. The yellow distribution shows the merged EW background while the magenta distribution shows the MJ background. The black points show the data distribution. The EW background contribution is subtracted from the data distribution on the left plots, then the MJ distribution shape is obtained. To recover the lack of statistics in the 2-tag template, the 1-tag template promoted to 2-tag is merged with the 2-tag templates and used for the MJ fit.

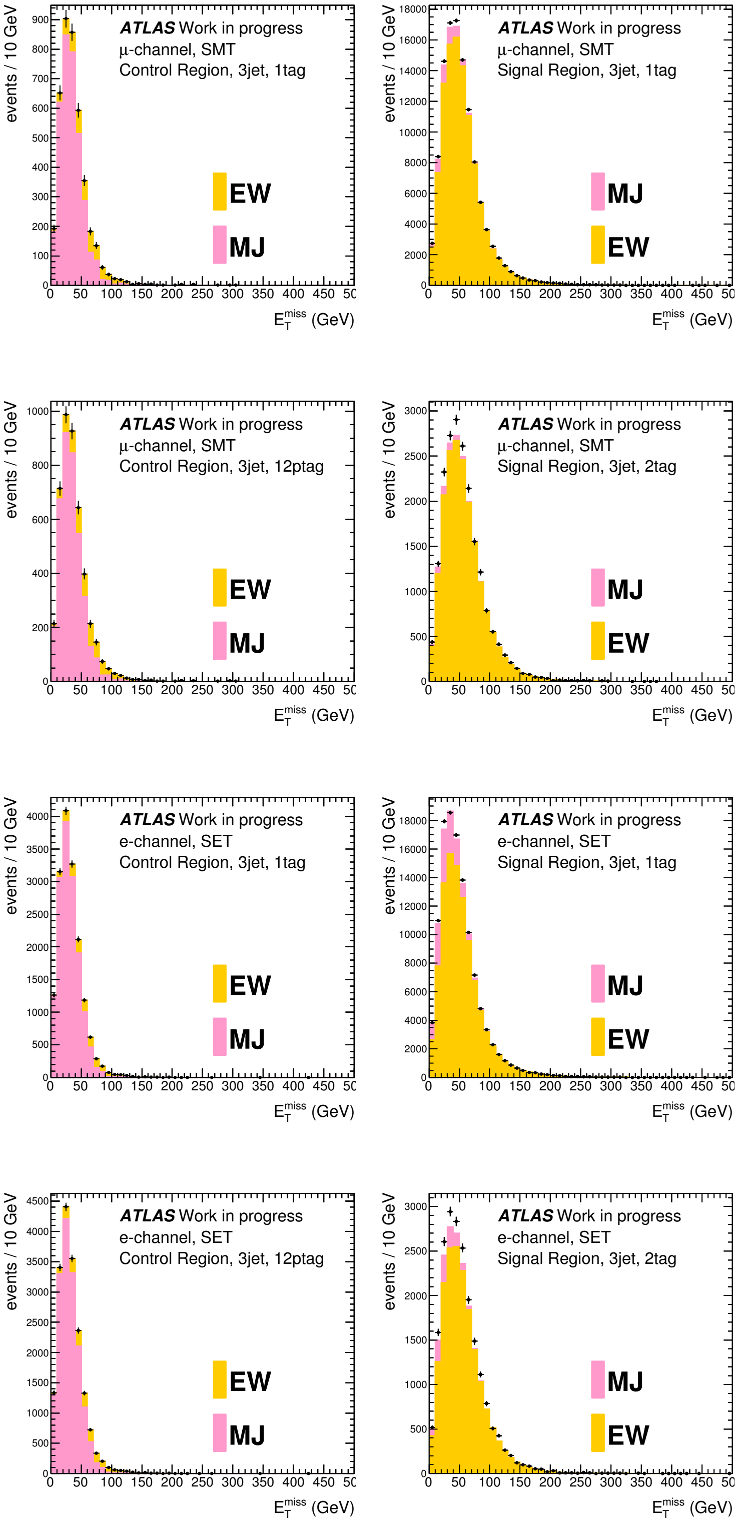


Fig. 5.6: The EW component subtraction to obtain MJ template in the control region (left) and the MJ fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the signal region (right). The $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions are shown in the 3-jet 1-tag and 3-jet 2-tag region muon (electron) channel in the first (third) and second (forth) row, respectively. The yellow distribution shows the merged EW background while the magenta distribution shows the MJ background. The black points show the data distribution. The EW background contribution is subtracted from the data distribution on the left plots, then the MJ distribution shape is obtained. To recover the lack of statistics in the 2-tag template, the 1-tag template promoted to 2-tag is merged with the 2-tag templates and used for the MJ fit.

W+jets background

The 0-tag sample has enough statistics to allow detailed investigations of the modeling of the W+jet background generated with SHERPA. Given that the search is performed in intervals of p_{T}^{W}, with the higher intervals having the higher sensitivity, an accurate modeling of the p_{T}^{W} distribution is important.

Figure 5.7 shows the p_{T}^{W} spectrum for W+jets production, in which data has a softer spectrum than MC simulation. The source of the mis-modeling was found to be a strong correlation between p_{T}^{W} and $\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2)$. The $\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2)$ distribution is therefore used for reweighting in order to address this mis-modeling (fig 5.7). The Wl and Wcl samples are re-weighted based on parametrized fits to the ratio of data to simulation in the $\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2)$ variable in the 0-tag region where these backgrounds dominate. After this reweighting, the modeling of the whole p_{T}^{W} distribution is greatly improved, as can be seen in fig 5.7. This reweighting also improves the other kinematic properties modeling, most notably the dijet mass. It is applied for the Wl and Wcl samples in all analysis regions, due to the modeling improvement observed also in the 1-tag control region. The Wb and Wcc background contributions in the 0-tag and 1-tag control regions are too small to allow conclusive studies of their modeling, so that no reweighting is applied to these background. However, an associated systematic uncertainty is assessed instead, as explained in section 5.8.

5.7. $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel

For the muon channel ($WH \to \mu\nu b\bar{b}$), events triggered by $E_{\mathrm{T}}^{\mathrm{miss}}$ are used in addition to SLT events in order to recover the events missed by SMT. In the former analysis, $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon events were considered only when $p_{\mathrm{T}}^{W} > 160$ GeV [116]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon events became available down to $p_{\mathrm{T}}^{W} > 120$ GeV, thanks to the well described $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger properties derived for the $ZH \to \nu\nu b\bar{b}$ analysis.

5.7.1. $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon events

At the trigger level, especially for the level-1, $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated using only the calorimeter components, as described in section 5.2, thus the muon momentum is not propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation at the trigger level. This fact makes the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers very helpful for this analysis, in which the muons are not always caught by the muon triggers.

In the muon channel, the primary triggers are SMT: EF_mu24vhi_medium1 and EF_mu36_medium1. However, the plateaux of these trigger efficiencies are around 80% as can be seen in fig 5.8. This is because of the limited coverage of level-1 muon chambers in the barrel region of the detector (see fig 5.9), and also the muon trigger's efficiency loss for the high p_{T} muons[c]. The large quantity of cables and cooling pipes of the inner detector and barrel calorimeter systems has to go through the muon spectrometer: this was accomplished by routing most of them radially outwards at $z = 0$ and at fixed azimuthal locations, where the radiation damage is small in comparison with the forward region, as illustrated in fig 5.9. In order to recover the events missed by SMT, $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are introduced. They are a combination of EF_xe80_tclcw, EF_xe80T_tclcw_loose and EF_xe80_tclcw_loose. Details of the combination of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are described in section 5.2. In case of $p_{\mathrm{T}}^{W} > 120$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are used only when the SMT mentioned above are not fired.

---

[c] When EF applies matching between the tracks reconstructed by the Inner Detector and the Muon Spectrometer, the energy deposited by the muons in the calorimeter is always estimated as minimum ionizing particle. This constant estimation differs from the actual energy deposited by high p_{T} muons due to the multiple scattering, lowering the muon trigger efficiency for the high p_{T} muons. In the offline reconstruction algorithms, a p_{T} dependent estimation is used.

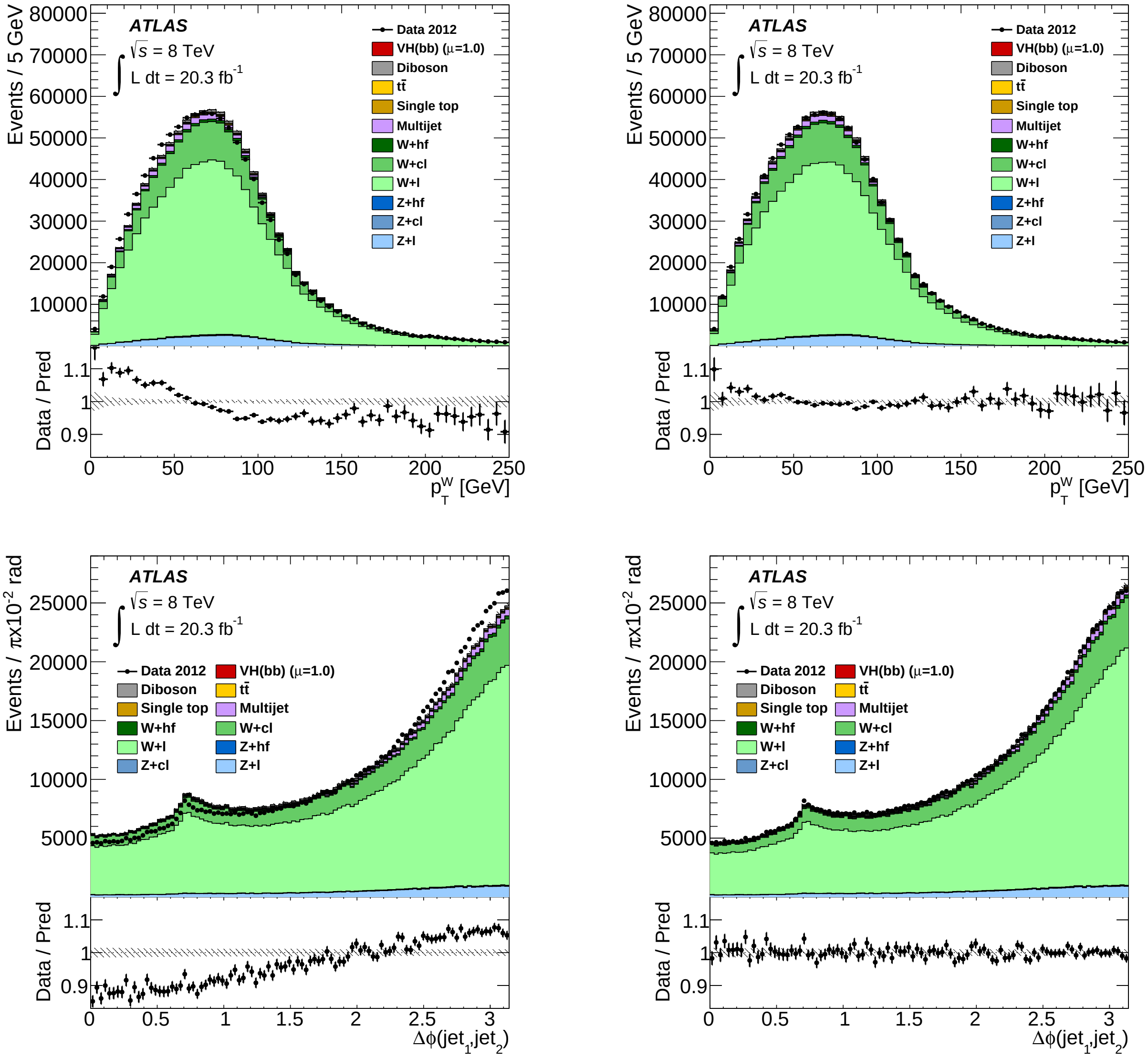


Fig. 5.7: The p_{T}^{W} and $\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2)$ distributions on the top and bottom row, respectively. Data (points with error bars) and simulation (histograms) distributions are shown for the 2-jet 0-tag control region . before (left) and after (right) $\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2)$ reweighting, with MVA selection (see section 5.4). The multijet and simulated-background normalizations are provided by the multijet fits. The size of the statistical uncertainty is indicated with the shaded band. The data-to-prediction ratios are shown in the lower panels [90].

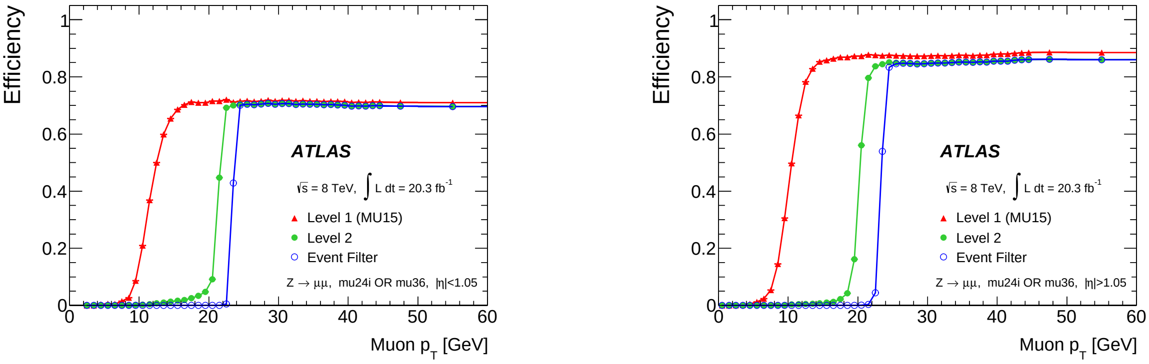


Fig. 5.8: Efficiency of passing either the mu24i or mu36 trigger chain as functions of the probe muon p_{T}, separately for the three trigger levels. The efficiency was measured from data, and is shown separately for the barrel region (left) and the endcap region (right). The error bars show the statistical uncertainties only [109].

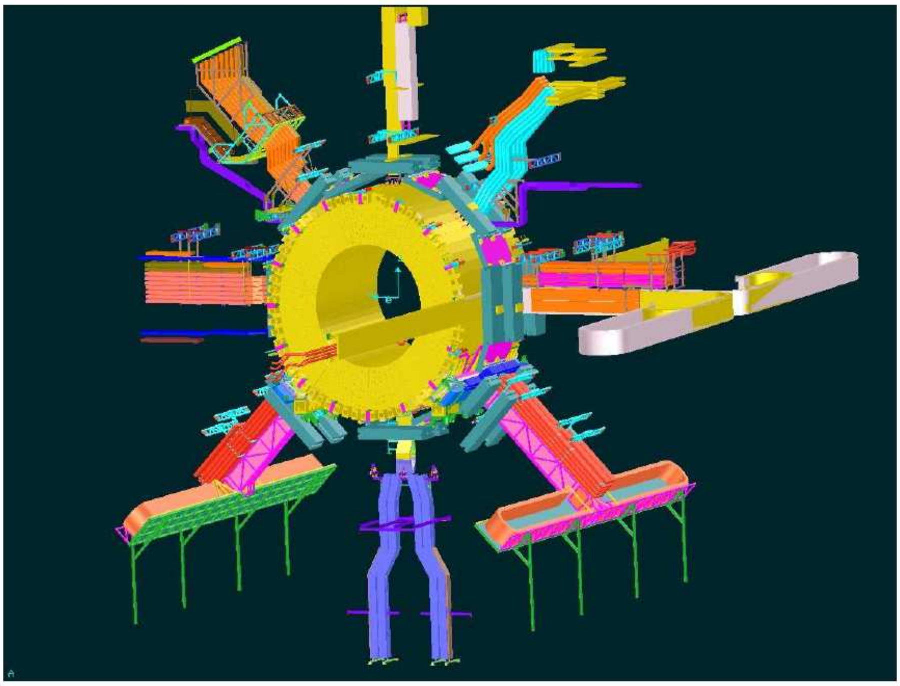


Fig. 5.9: Detailed three-dimensional layout and routing of cables and services for the ATLAS barrel system. The inner barrel muon chambers in the central region are shown. One clearly sees the holes in the acceptance caused by the considerable volume of services exiting the detector at $\eta \sim 0$ [3].

5.7.2. $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency and scale factor

The trigger scale factor is parametrized by offline-reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$, using the $ZH \to \nu\nu b\bar{b}$ analysis selection. However to emulate this scale factor, $E_{\mathrm{T}}^{\mathrm{miss}}$ should be calculated without muon momentum in the $WH \to \mu\nu b\bar{b}$ channel, because muons are not taken into account in the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$ at trigger level, and the efficiency measurements are done without using the muon term using $W \to \mu\nu$+jets and $Z \to \mu\mu$+jets events. Hence, the data/MC agreement of $E_{\mathrm{T}}^{\mathrm{miss}}$ without muon contribution ($E_{\mathrm{T}}^{\mathrm{miss}}(\not\mu)$) should be checked.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency is measured with the following steps:

1. Measure the trigger efficiency turn-on curves in the level-1, level-2, and EF separately in data and simulation.

2. Obtain total efficiency turn-on curves by multiplying these three separate turn-on curves both in data and simulation.

3. The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency scale factor is given by data/simulation ratio of the total efficiency turn-on curves.

Figures 5.10 and 5.11 show the distribution of $E_{\mathrm{T}}^{\mathrm{miss}}(\not\mu)$, after MVA selection, for the 2- and 3-jet bins, respectively. Figure 5.12 shows the trigger efficiency scale factor as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$, obtained from the data-to-simulation ratio of the trigger efficiency turn-on curves.

Additional selections for $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel

In the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel, further requirements are applied in addition to the nominal selection for adjustment of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger application which is optimized for the $ZH \to \nu\nu b\bar{b}$ analysis. These cuts are introduced to remove the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger biases found for these variables. The events must have $\Sigma p_{\mathrm{T}}^{\mathrm{jet}} > 120\ (150)$ GeV, where $\Sigma p_{\mathrm{T}}^{\mathrm{jet}}$ is the scalar p_{T} sum of two (three) central jets in 2(3)-jet events. In the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel, the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}(\not\mu)$ does not always match with the reconstructed p_{T}^{W} even though it should match in ideal conditions. This inconsistency occurs because of the difference between $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}(\not\mu)$ mainly in low p_{T}^{W} or $E_{\mathrm{T}}^{\mathrm{miss}}(\not\mu)$ since the modeling of calorimeter energy deposits by muons in this region is difficult. To avoid using such events reconstructed with a large difference between $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}(\not\mu)$, the events in this channel are required to have $E_{\mathrm{T}}^{\mathrm{miss}}(\not\mu) > 120$ GeV.

5.7.3. Background estimation

Background estimation in this category is done in the same way as the SLT-only analysis, using the MC simulations with various corrections applied.

For the MJ background, it is also estimated from data with MJ fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$as described in section 5.6.1, but using the same template obtained for the SMT analysis since the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel has no p_{T}^{W} interval which is the main source of the MJ background. However, the size of the MJ background is negligible in this analysis category, due to the high p_{T}^{W} requirement and also the additional requirements made for this category.

5.7.4. Sensitivity gain

Figures 5.13–5.15 (5.16–5.18) show the m_{jj} distribution for each p_{T}^{W} bin and b-tagging category in the 2 (3) jet bin. Based on this distribution in each category and the SLT-only analysis regions,

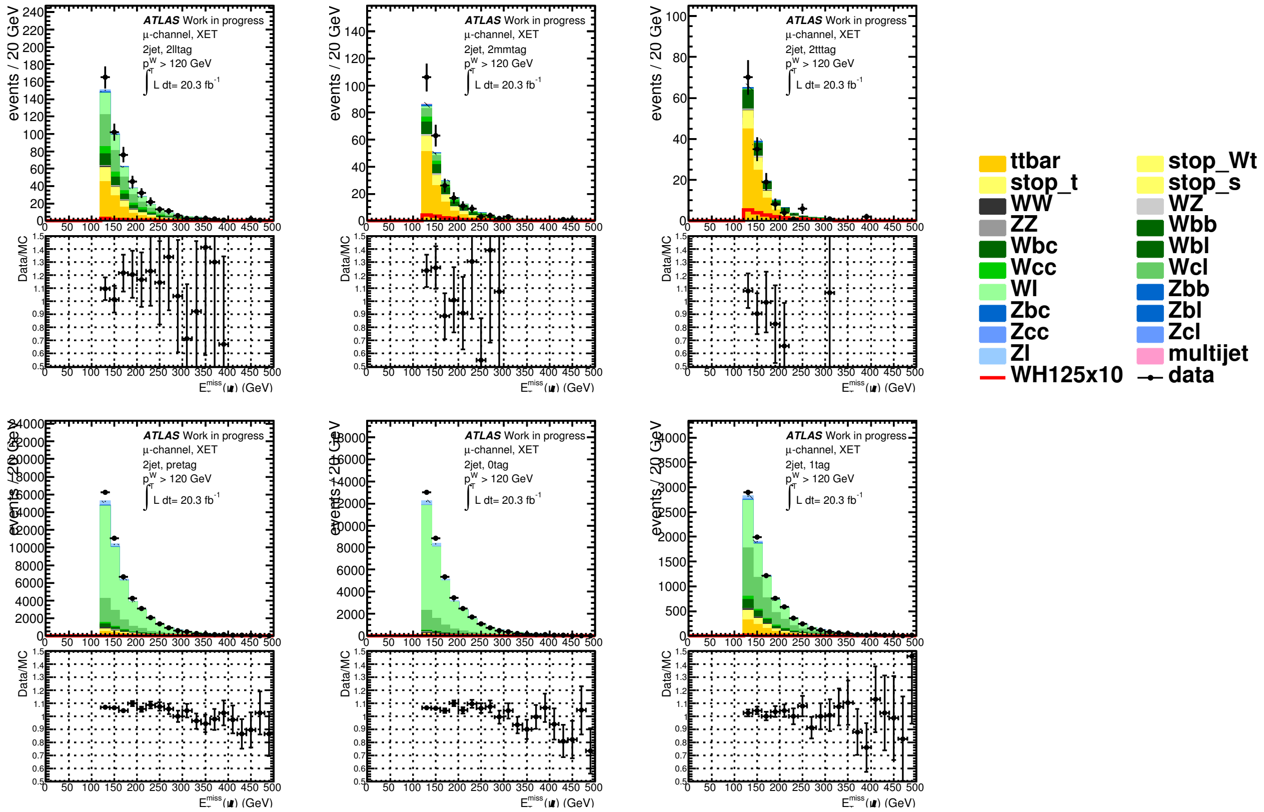


Fig. 5.10: The pre-fit $E_{\mathrm{T}}^{\mathrm{miss}}(\not\mu)$ distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered events in 2 jet bin, $p_{\mathrm{T}}^{W} > 120$ GeV after MVA selection. Each b-tagging category is shown in 2 loose tag (top left), 2 medium tag (top middle), 2 tight tag (top right), 1 tag (bottom right), 0 tag (bottom middle) and pre-tag (bottom left).

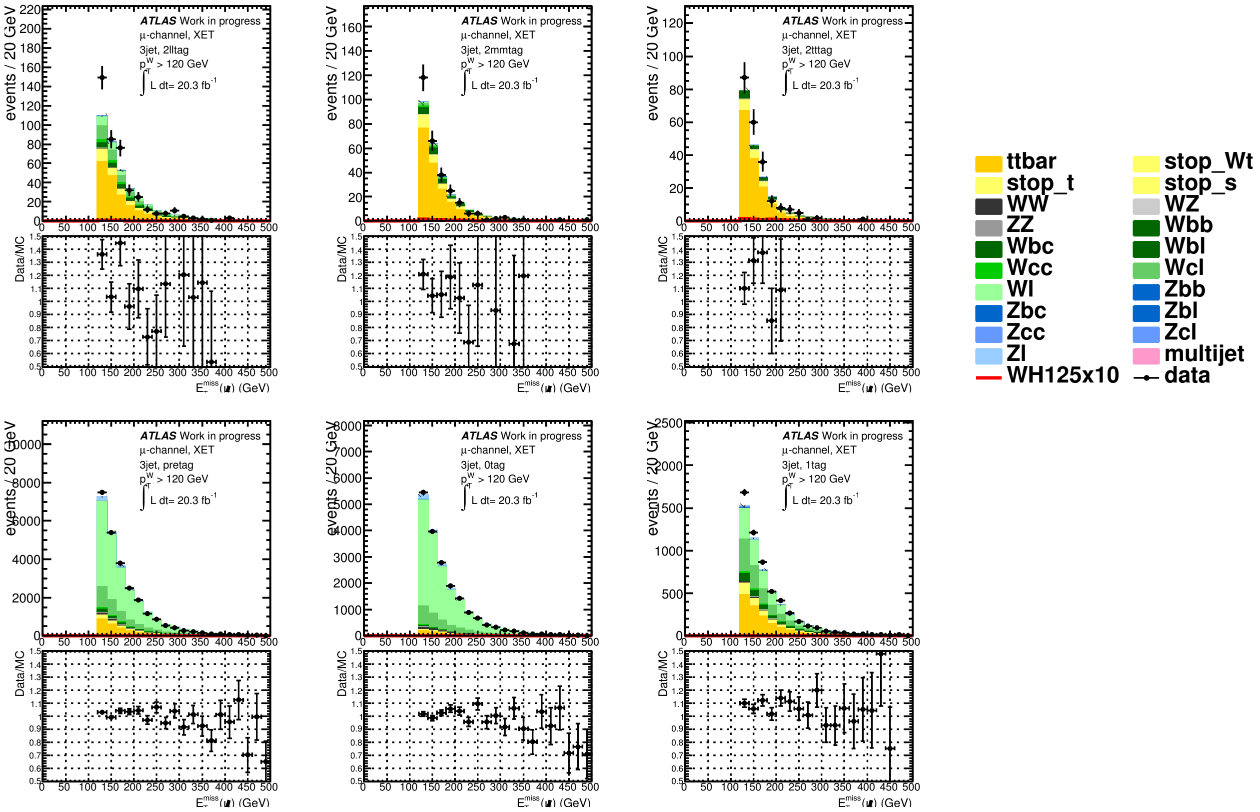


Fig. 5.11: The pre-fit distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ without muon contribution ($E_{\mathrm{T}}^{\mathrm{miss}}(\not\mu)$) of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered events in the 3-jet bin, $p_{\mathrm{T}}^{W} > 120$ GeV after MVA selection. Each b-tagging category is shown in 2 loose tag (top left), 2 medium tag (top middle), 2 tight tag (top right), 1 tag (bottom right), 0 tag (bottom middle) and pre-tag (bottom left).

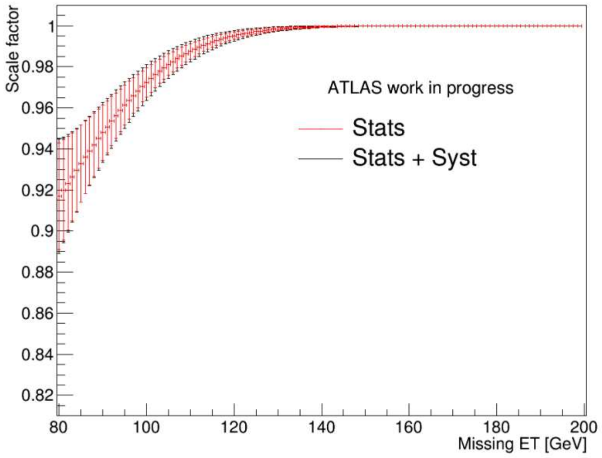


Fig. 5.12: $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency scale factor for $W \rightarrow \mu\nu$+jet samples with statistical and statistical+systematic uncertainties [74].

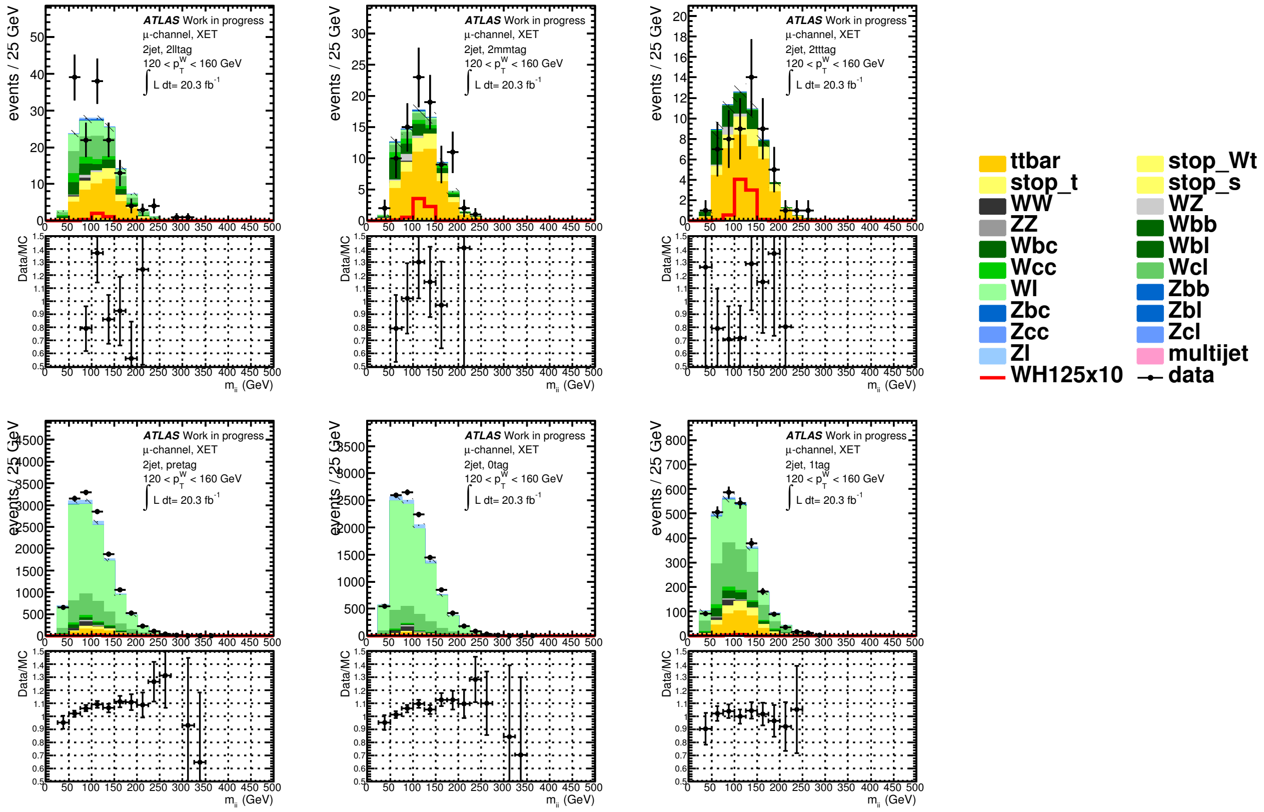


Fig. 5.13: Pre-fit m_{jj} distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered events in 2-jet bin, $120 < p_{\mathrm{T}}^{W} < 160$ GeV. Each b-tagging category is shown in 2 loose tag (top left), 2 medium tag (top middle), 2 tight tag (top right), 1 tag (bottom right), 0 tag (bottom middle) and pre-tag (bottom left).

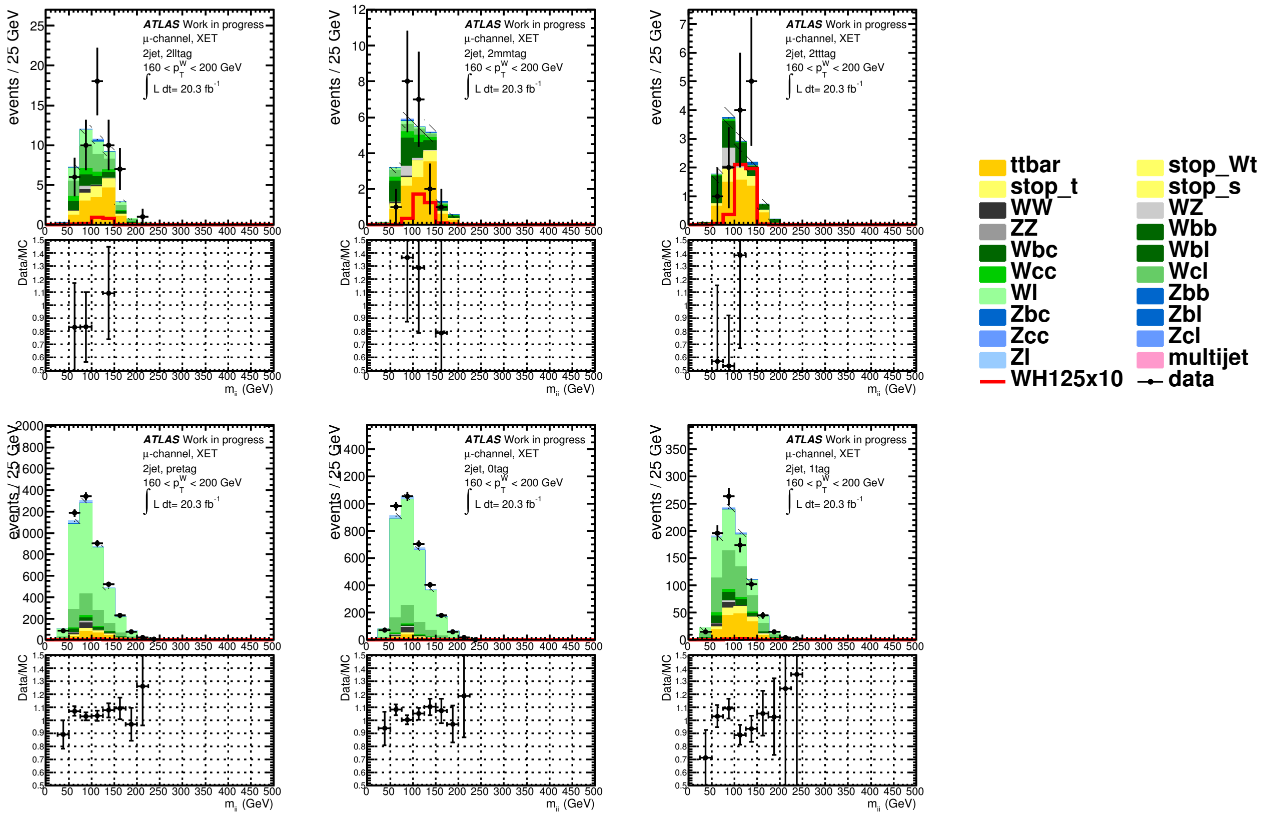


Fig. 5.14: Pre-fit m_{jj} distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered events in 2-jet bin, $160 < p_{\mathrm{T}}^{W} < 200$ GeV. Each b-tagging category is shown in 2 loose tag (top left), 2 medium tag (top middle), 2 tight tag (top right), 1 tag (bottom right), 0 tag (bottom middle) and pre-tag (bottom left).

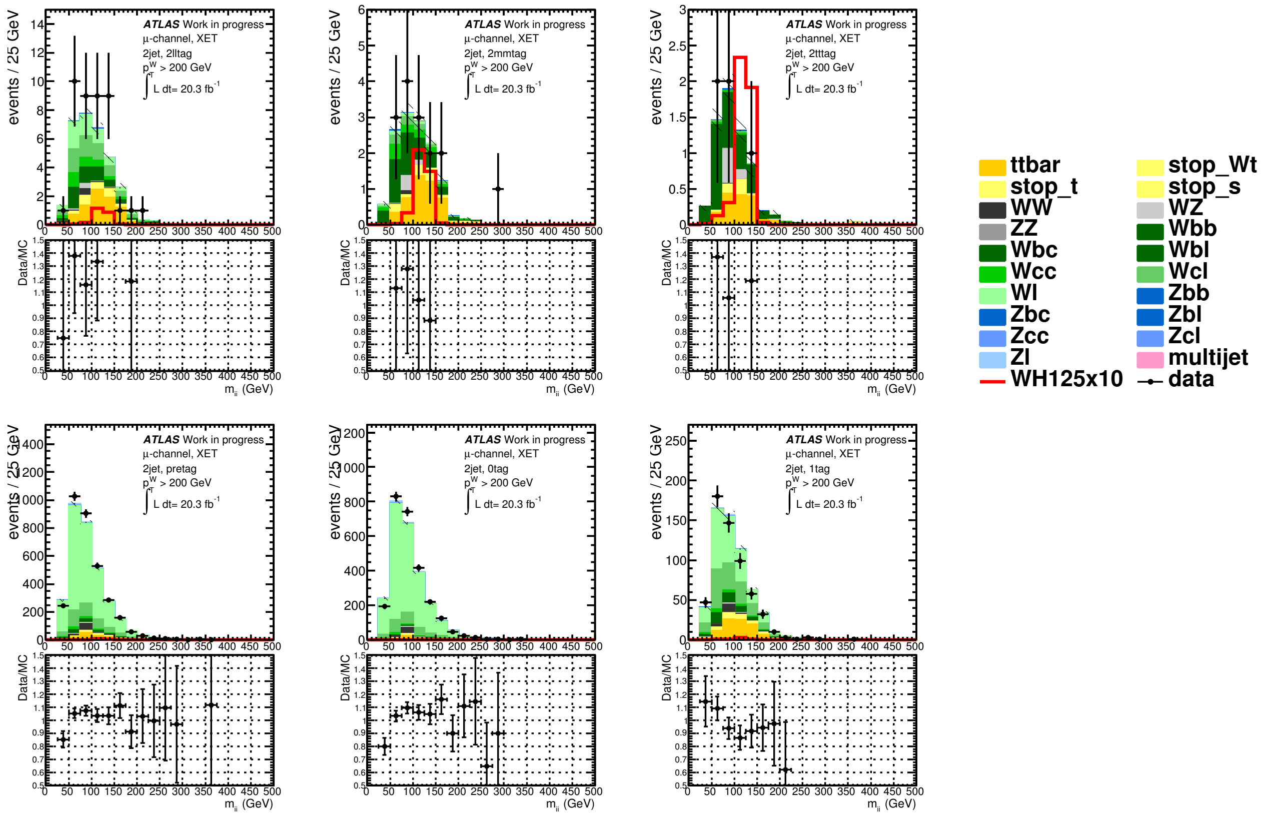


Fig. 5.15: Pre-fit m_{jj} distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered events in 2-jet bin, $p_{\mathrm{T}}^{W} > 200$ GeV. Each b-tagging category is shown in 2 loose tag (top left), 2 medium tag (top middle), 2 tight tag (top right), 1 tag (bottom right), 0 tag (bottom middle) and pre-tag (bottom left).

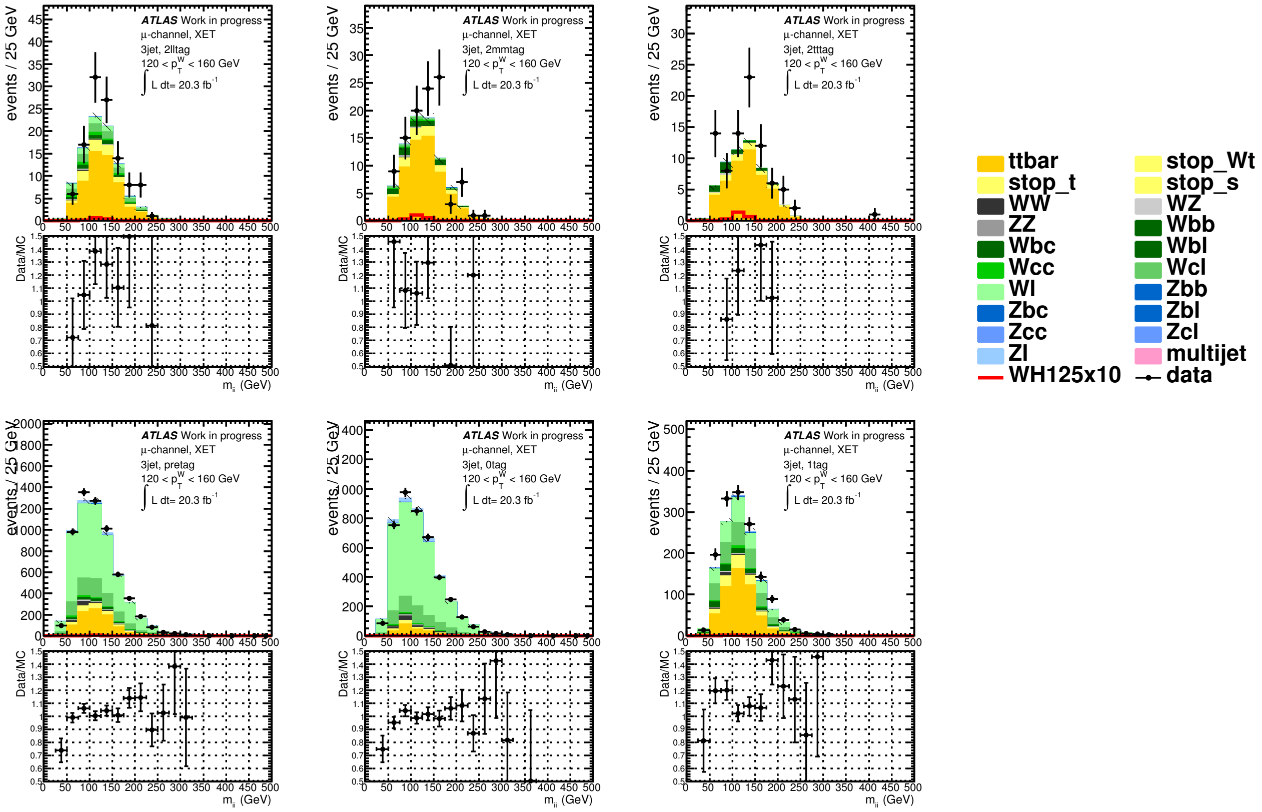


Fig. 5.16: Pre-fit m_{jj} distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered events in 3 jet bin, $120 < p_{\mathrm{T}}^{W} < 160$ GeV. Each b-tagging category is shown in 2 loose tag (top left), 2 medium tag (top middle), 2 tight tag (top right), 1 tag (bottom right), 0 tag (bottom middle) and pre-tag (bottom left).

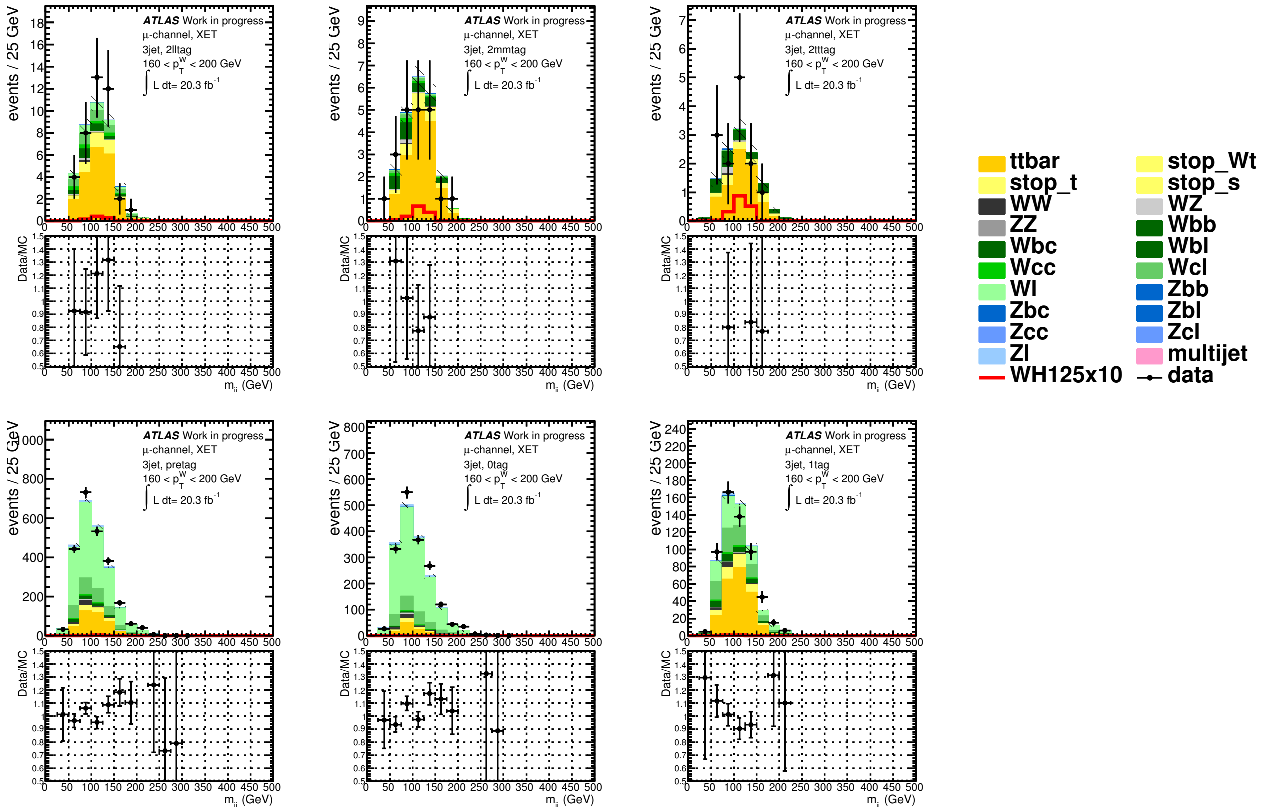


Fig. 5.17: Pre-fit m_{jj} distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered events in 3 jet bin, $160 < p_{\mathrm{T}}^{W} < 200$ GeV. Each b-tagging category is shown in 2 loose tag (top left), 2 medium tag (top middle), 2 tight tag (top right), 1 tag (bottom right), 0 tag (bottom middle) and pre-tag (bottom left).

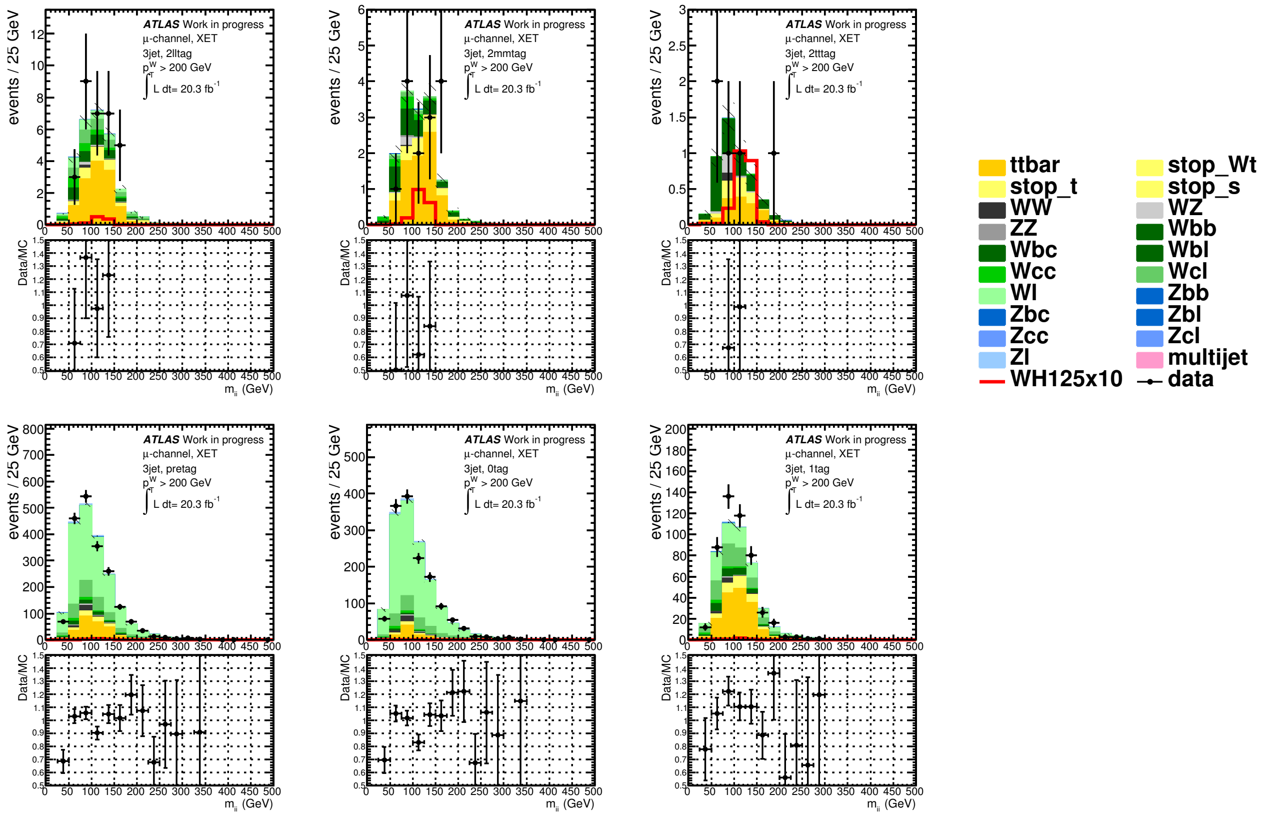


Fig. 5.18: Pre-fit m_{jj} distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered events in 3 jet bin, $p_{\mathrm{T}}^{W} > 200$ GeV. Each b-tagging category is shown in 2 loose tag (top left), 2 medium tag (top middle), 2 tight tag (top right), 1 tag (bottom right), 0 tag (bottom middle) and pre-tag (bottom left).

sensitivity [117] is calculated by:

$$\sqrt{2((s+b)\log(1+\frac{s}{b})-s)}, \quad (5.2)$$

where s (b) is the number of signal (background) events within the range $100 < m_{jj} < 150$ GeV with the cut-based analysis selection. The combined significance of the three 2-tag categories (loose, medium and tight) is shown in table 5.7 for 2- and 3-jet bins with the final combined sensitivity of 2- and 3-jet bins. Comparing the SMT-only analysis with the SMT+XET analysis, the overall gain was found to be 10%, while it can go up to 20% in particular analysis regions.

Table 5.7: Significance in 2 and 3 jet bins, and in the combination.

p_{T}^{W} [GeV]	$0-90$	$90-120$	$120-160$	$160-200$	$200-$	total
2 jet bin						
Single muon trigger	0.55	0.34	0.35	0.38	0.61	1.03
$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger	–	–	0.19	0.21	0.33	0.46
Total	0.55	0.34	0.40	0.43	0.70	1.13
Gain	–	–	14%	14%	14%	9.6%
3 jet bin						
Single muon trigger	0.135	0.095	0.11	0.14	0.24	0.34
$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger	–	–	0.05	0.07	0.16	0.18
Total	0.135	0.095	0.12	0.16	0.29	0.39
Gain	–	–	9.1%	11%	20%	14%
Combination of 2 and 3 jet bins						
Single muon trigger	0.565	0.35	0.37	0.41	0.66	1.1
$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger	–	–	0.20	0.22	0.37	0.5
Total	0.565	0.35	0.42	0.46	0.75	1.2
Gain	–	–	11%	13.5%	14%	10%

5.8. Systematic uncertainties

The various sources of systematic uncertainties, which will give different results for the signal strength measurement, are discussed in this section. The experimental and modeling uncertainties are considered as well as the uncertainties arising from the MJ background.

5.8.1. Experimental uncertainties

The systematic uncertainties arising from the measurement of the physics objects in the experiments are discussed in this section.

- For the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, an efficiency correction is derived in $W \to \mu\nu$+jets and $Z \to \mu\mu$+jets events. The correction amounts to below 1% for events with $E_{\mathrm{T}}^{\mathrm{miss}} > 120$ GeV, the threshold required in the analysis and becomes negligible in the range $E_{\mathrm{T}}^{\mathrm{miss}} > 160$ GeV. The associated uncertainties arise from the statistical uncertainties of this method and differences observed in the two event classes, which is less than 1%.

- Sets of uncertainties are considered for lepton trigger, reconstruction, identification and isolation efficiencies. Their impact is typically very small, less than 1%. Some of them are pruned due to the negligible impact on the final results.

- Jet energy scale (JES) uncertainties considered are originated in several sources: in-situ calibration analyses, pileup dependent corrections and the flavor composition of the jets in different event classes. The total relative systematic uncertainties on the JES range from about 3% to 1% for central jets with a p_{T} of 20 and 1000 GeV, respectively.

- An additional specific uncertainty of about 1–2% is taken into account for the b-jet energy scale.

- Corrections and uncertainties are also considered for the jet energy resolution (JER) [118]. Total relative JER uncertainty is in the 10%–20% range depending on the η range for $p_{\mathrm{T}}^{\mathrm{jet}} = 20$ GeV. The uncertainty goes down as $p_{\mathrm{T}}^{\mathrm{jet}}$ increases and becomes less than 5% for $p_{\mathrm{T}}^{\mathrm{jet}} > 200$ GeV.

- An uncertainty arising from JVF, which is used to reject jets from the pileup interaction, is considered.

- The $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement is impacted by the jet related systematic uncertainties, as well as the lepton related systematic uncertainties. The uncertainties on the energy calibration (8%) and resolution (2.5%) of energy clusters not associated to any reconstructed object [69] are also taken into account for the $E_{\mathrm{T}}^{\mathrm{miss}}$ systematic uncertainty.

- The MV1c algorithm is used to tag b-jet against the c- and light jet background. The b-tagging related uncertainties are mixtures of statistical error from data, experimental error (JER), and modeling errors ($t\bar{t}$ events used to measure the b-tagging uncertainties). These uncertainties depend on the jet p_{T}, η and the b-tagging operating points used in the analysis. The uncertainties are decomposed to uncorrelated components and the 10 most impacting uncertainties are considered for this analysis for b-jets, 15 for c-jets and 10 for light jets.

- The uncertainty on the integrated luminosity uncertainty is 2.8%. It is derived, following the same methodology as that described in ref [119], from a preliminary calibration of the luminosity scale derived from beam-separation scans performed in November 2012. It is applied to the signal and to those backgrounds estimated from simulation. A 4% uncertainty on the average number of interactions per bunch crossing is included.

5.8.2. Uncertainties on the multijet background

Normalization uncertainties on the MJ background estimation arise from two sources: statistical uncertainties of the multijet-fits and uncertainties on the subtraction of the EW background in the control region. Normalization uncertainties are assessed in the LL, MM, and TT categories to cover differences between MJ fits performed inclusively in the 2-tag regions and in the individual categories. The overall normalization uncertainties are 11%, 14%, and 22% in the LL, MM, and TT categories of the electron channel. The corresponding uncertainties are about three times larger in the muon channel due to the smaller size of the background.

Shape uncertainties are estimated by comparing the tighter and looser control region compared to the nominal control region definition. In the electron channel, an alternative template is constructed with a track-based isolation in the 0.12 to 0.5 interval, and another alternative template with a calorimeter-based isolation in the 0 to 0.04 interval. In the muon channel, the results obtained with the nominal MJ template are compared with those obtained with the tighter or looser isolation requirements, defined by track-based isolation intervals of 0.07–0.95 and 0.95–0.5, respectively.

These variations are also summarized in fig 5.1. Furthermore, the $\Delta R(\mathrm{jet}_1, \mathrm{jet}_2)$ and p_{T}^{W} reweighting mentioned in section 5.6.1 for the electron channel are varied by 50% for the systematic variations.

5.8.3. Uncertainties on the modeling of simulated background

The uncertainties related to the simulated samples are discussed: these are modeling uncertainties on $t\bar{t}$, (W/Z)+jets, single-top, diboson, and signal samples.

$t\bar{t}$ and single-top quark production

As explained in section 5.6.2, the p_{T} of the $t\bar{t}$ system is re-weighted at generator level to bring it in agreement with the measurement [115]. A half of this correction is taken as systematic uncertainty across channels. The nominal $t\bar{t}$ simulated samples are made using POWHEG+PYTHIA. The various generators, parton showering, hadronization scheme and PDFs are compared to these nominal events. The systematic variations are obtained using the ALPGEN generator [48], since the most significant deviation is generally found with it.

The theoretical uncertainties on the cross-sections of the three process contributing to single-top production are 4%, 4%, and 7% for the s-channel, t-channel, and Wt production, respectively [120].

W+jet production

As explained in section 5.6.2, a $\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2)$ correction is applied to the Wl and Wc samples. A half of the correction is assigned as systematic uncertainties for each of the 2- and 3-jet bins. Although no correction is applied to the Wcc and Wb components, a systematic uncertainty is assigned, equal to the full correction applied to the Wl and Wcl samples, uncorrelated between 2- and 3-jet events.

Based on the agreement observed between data and prediction in the 0-tag sample, a 10% uncertainty is considered for the 3-to-2-jet ratio for the Wl and Wcl backgrounds. The normalizations of the Wcl and the Wbb backgrounds are left freely floating in the global fit.

Further systematic uncertainties are assessed for the Wbb background, for which dedicated control regions are not available in the data. Generator level comparisons are made using various generators, PDFs, hadronization, etc. As a result, a 10% uncertainty is assigned to the 3-to-2-jet ratio, taken as correlated between all W+heavy flavor processes. Shape uncertainties are also assessed for the m_{bb} and p_{T}^{W} distributions.

For the flavor composition uncertainty, comparison of the nominal events to the inclusively generated SHERPA and MC@NLO samples is performed after full reconstruction and event selection, resulting in uncertainties assigned in the W+hf samples: 35% for bl/bb, and 12% for each of bc/bb and cc/bb.

Z+jet production

Although the contribution from Z+jet background is very small in WH signal regions, a dedicated background study has been made in the $ZH \rightarrow \ell\ell b\bar{b}$ analysis. Similar to the W+jet sample as described in section 5.6.2, but not identical correction has been applied to correct the mis-modeling found in Z+jet samples. The $\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2)$ reweighting is applied to the Zl component, and the p_{T}^{Z} reweighting is applied to the Zb and Zc components. The half (full) correction has been considered as a systematic uncertainty for Zl ($Zb + Zc$) component, separately being done for 2- and 3-jet bins. These uncertainties are treated as uncorrelated, however correlated in the $Zb + Zc$ component.

The normalization of the Zl background is determined from data in the 0-tag region of the $ZH \rightarrow \ell\ell b\bar{b}$ analysis, resulting in 5% of uncertainty while the normalization of Zcl and Zbb backgrounds are left free in the global fit.

The simulated shape of the m_{bb} distribution is compared to that in data using the mass side-band

($m_{bb} < 100$ and m_{bb} >150 GeV) in the 2-tag regions of the $ZH \to \ell\ell b\bar{b}$ channel. When increasing m_{bb} by 3% at 50 GeV, it decreases by 5% at 200 GeV. This uncertainty is treated as uncorrelated for the Zl and $Zb + Zc$ components.

Diboson production

The uncertainties on the cross-sections for diboson production (WW, WZ, and ZZ) are assessed at parton level using MCFM [121] at NLO in QCD. The sources of uncertainty considered are the renormalization and factorization scales and the choice of PDFs. Other than normalization uncertainties, the shape-only systematic uncertainty is assessed by comparing the shapes obtained with the nominal generator (POWHEG+HERWIG) and with HERWIG.

WH signal process

The signal samples are corrected according to p_{T}^{W} as described in section 5.1. The signal cross-section uncertainty is assessed from the choice of normalization and factorization scales and PDFs. The scale uncertainties of 1% and 3% are assessed for $q\bar{q} \to WH$ and $gg \to ZH$ production, respectively. Acceptance uncertainties due to the choice of scales are determined from signal samples generated with POWHEG interfaced with PYTHIA8. The uncertainty on the branching ratio of $H \to b\bar{b}$ is about 3%.

5.9. Statistical analysis

The relative signal strength to the SM prediction, $\mu = \sigma/\sigma_{\mathrm{SM}}$, is extracted from data, using a statistical fitting procedure based on the ROOSTATS framework [122, 123]. A binned maximum likelihood function is constructed like the following equation, as the product of Poisson-probability terms over the bins of the input distributions including the numbers of data events and expected signal and background yields, taking into account the effects of the floating background normalizations and systematic uncertainties:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{j} \prod_{i=\mathrm{bin}} \mathrm{Poisson}\left(N_i(j) | \mu \left(s_i^{WH}(j) + s_i^{ZH}(j)\right) + b_i(j)\right) \prod_{\boldsymbol{\theta}} \mathrm{func}(t|\boldsymbol{\theta}, 1), \tag{5.3}$$

where μ is the signal strength, $\boldsymbol{\theta}$ is the nuisance parameters (NPs), func is given by Gaussian or log-normal probability density functions (PDFs), and t is the measured values in the calibrations. $N_i(j)$ is the number of observed events in the ith bin in the jth signal region, and $s_i^{WH}(j)$, $s_i^{ZH}(j)$ and $b_i(j)$ are the expected number of WH, ZH signals and background events, respectively. Here, $s_i(j)$ and $b_i(j)$ are given by a function of NP. The impact of systematic uncertainties on the signal and background expectations is described by NPs, $\boldsymbol{\theta}$, which are constrained by Gaussian or log-normal PDFs, the latter being used for normalization uncertainties to prevent normalization factors from becoming negative in the fit. The expected number of signal and background in each bin are given by functions of $\boldsymbol{\theta}$. The NP parameterization is chosen such that the predicted signal and background yields in each bin are log-normally distributed for a normally distributed θ.

The test statistic, q_μ, is defined by the profile likelihood ratio,

$$q_\mu = -2\ln(\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}}_\mu)/\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})), \tag{5.4}$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the parameters that maximize the likelihood (with the constraint $0 \leq \hat{\mu} \leq \mu$), and $\hat{\hat{\boldsymbol{\theta}}}_\mu$ are NP values that maximize the likelihood for a given μ. This test statistic is used for exclusion intervals derived with the CL_s method [117, 124]. To measure the compatibility of the background-only hypothesis with the observed data, the test statistic used is $q_0 = -2\ln(\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}_0)/\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}}))$.

When the maximized likelihood for background-only (signal+background) hypothesis is given by $\hat{\boldsymbol{\theta}}_0^{\mathrm{obs}}$ ($\hat{\boldsymbol{\theta}}_\mu^{\mathrm{obs}}$), PDFs expressing the observed data are given by the functions of $f(q_0|0,\hat{\boldsymbol{\theta}}_0^{\mathrm{obs}})$ ($f(q_\mu|\mu,\hat{\boldsymbol{\theta}}_\mu^{\mathrm{obs}})$). Then compatibility between observed data and a given hypothesis (μ) is measured by a p-value, which is defined as below:

$$p_\mu = P(q_\mu \geq q_\mu^{\mathrm{obs}} \mid \mathrm{signal+background}) = \int_{q_\mu^{\mathrm{obs}}}^{\infty} f(q_\mu|\mu,\hat{\boldsymbol{\theta}}_\mu^{\mathrm{obs}})dq_\mu \tag{5.5}$$

$$1 - p_b = P(q_\mu \geq q_\mu^{\mathrm{obs}} \mid \mathrm{background-only}) = \int_{q_\mu^{\mathrm{obs}}}^{\infty} f(q_0|0,\hat{\boldsymbol{\theta}}_0^{\mathrm{obs}})dq_\mu \tag{5.6}$$

Using these two, CL_s is calculated:

$$CL_s(\mu) = \frac{p_\mu}{1-p_b}. \tag{5.7}$$

The CL_s upper limit on μ is obtained by solving $CL_s(\mu)$ =5%, labeled as μ_{up}. The value of μ is considered to be excluded at 95% confidence level if $\mu > \mu_{\mathrm{up}}$.

Fit categories

The cut-based analysis uses the dijet invariant mass values for the inputs of the “global-fit”, while the MVA uses the BDT discriminants. In the cut-based analysis, the five p_{T}^W intervals, two jet-bins (2- and 3-jet), and three b-tagging categories (LL, MM, and TT) were used. In the MVA, the number of p_{T}^W intervals is two instead of five, however, the p_{T}^W information is used for the classification. Additional input distributions are used, which are the $MV1c$ distributions from the 1-tag control regions.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel have been merged at the input preparation level before the global fit. The SLT-only results are compared to the one with $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel.

Systematic uncertainties

The number of regions, NPs and the number of systematic-variation template pairs ($\pm 1\sigma$) amount close to twenty thousand, which renders the fits highly time consuming. To reduce the fitting time, systematic uncertainties with a negligible impact on the final results are pruned away. These pruned sources are muon momentum scale and resolution, electron energy resolution, and the quark gluon composition component of the JES.

The ranking of systematic uncertainties obtained with the MVA applied to 8 TeV data is shown in fig 5.19 with the NPs ordered by decreasing post-fit impact on $\hat{\mu}$. The most impacting sources of systematic uncertainties are the normalization and the m_{bb} shape uncertainty of the W+hf background and the signal acceptance due to the parton shower modeling. With and without the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel, no significant difference on the most impacting systematic uncertainties were observed.

5.10. Results

Figures 5.20 and 5.21 show the post-fit $MV1c$ distribution in the 1-tag category and BDT distributions in the three b-tagging categories (LL, MM, and TT) for $p_{\mathrm{T}}^W < 120$ GeV in the 2- and 3-jet bins, respectively. Figure 5.22–5.24 shows the MV1c distributions in 1-tag category and the BDT distributions in the three 2-tag categories in the 2- and 3-jet bins, comparing before and after the addition of the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel.

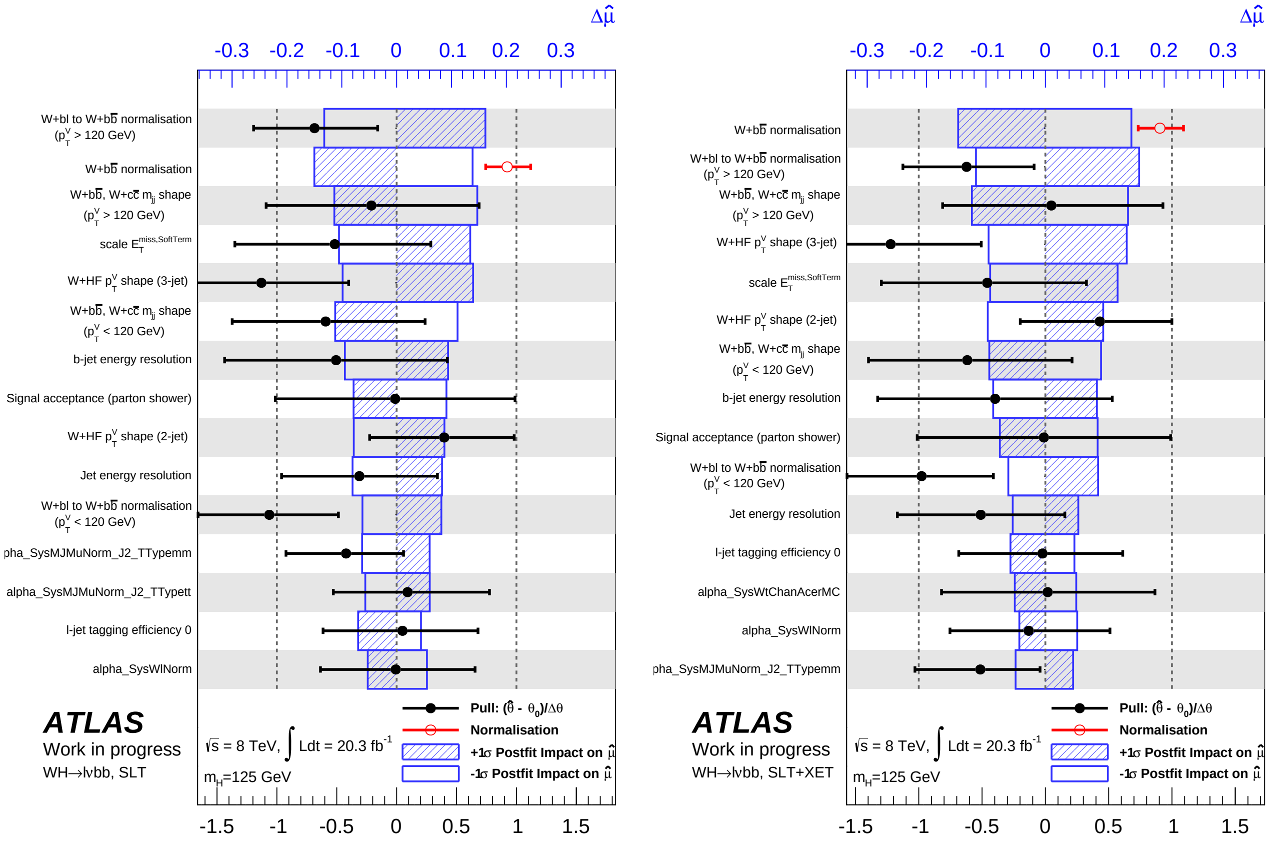


Fig. 5.19: The impact of systematic uncertainties on the fitted signal-strength parameter $\hat{\mu}$ for the MVA applied to the 8 TeV data for SLT-only results (left) and SLT+XET (right). The systematic uncertainties are listed in decreasing order of their impact on $\hat{\mu}$ on the y-axis. The boxes show the variations of $\hat{\mu}$ (the top x-axis), with the corresponding NP θ fixed to $\pm 1\sigma$ of its post-fit value $\hat{\theta}$ while the other NPs are kept at their post-fit values. The black points (the bottom x-axis) show the deviation of the fitted NPs $\hat{\theta}$ with respect to their nominal value θ_0 ($(\hat{\theta} - \theta_0)/\Delta\theta$ where $\Delta\theta$ is their nominal uncertainty). The associated error bars are the relative post-fit uncertainties of the NPs, relative to their nominal uncertainties. The red points and error bars are fitted values of the normalization parameters.

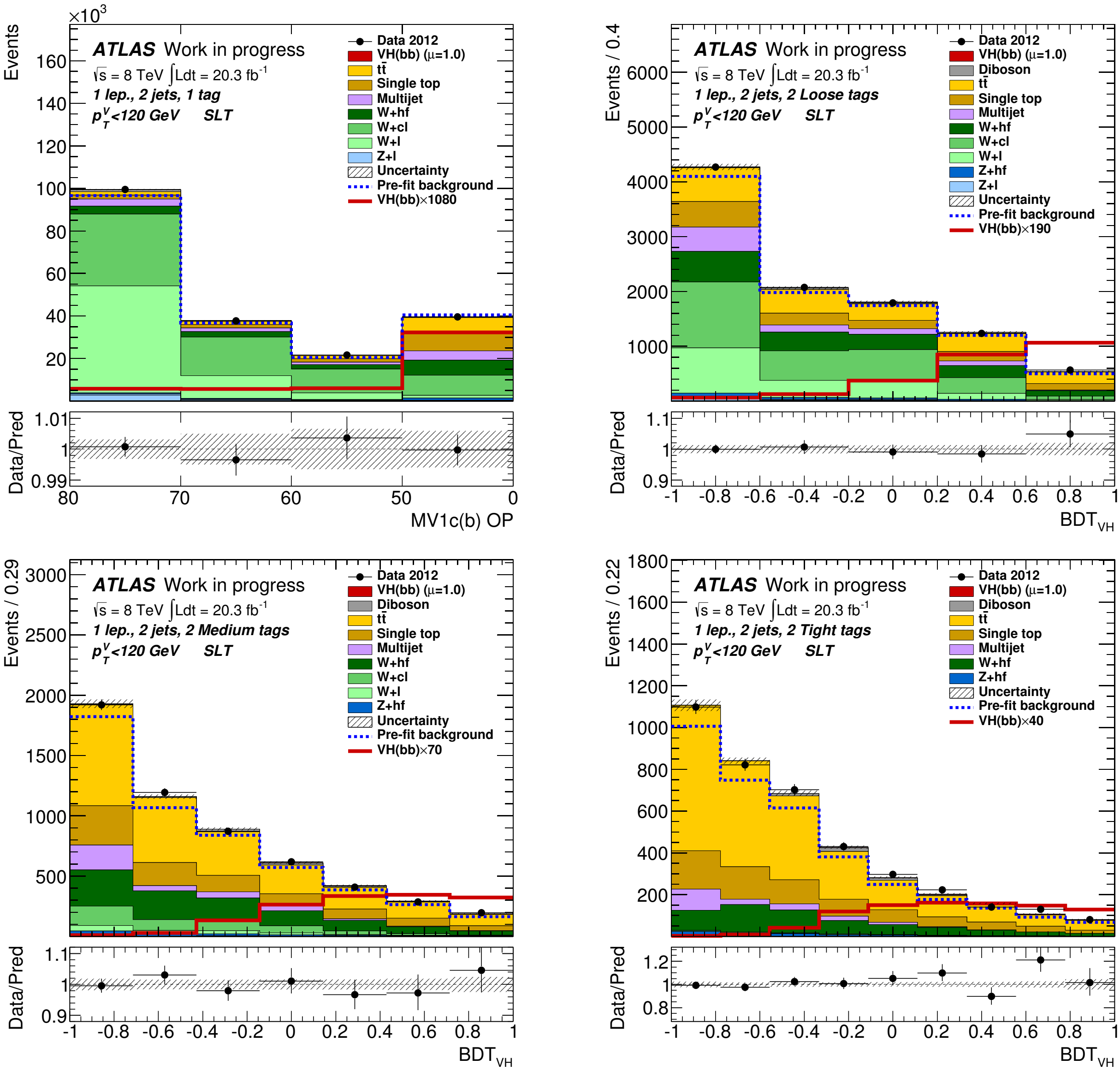


Fig. 5.20: The post-fit $MV1c$ distribution in the 1-tag b-tagging category (top-left) and BDT distributions in the LL (top-right), MM (bottom-left), and TT (bottom-right) b-tagging categories, in the 2-jet low p_{T}^{W} regions. The data/prediction ratios are shown in the lower panels.

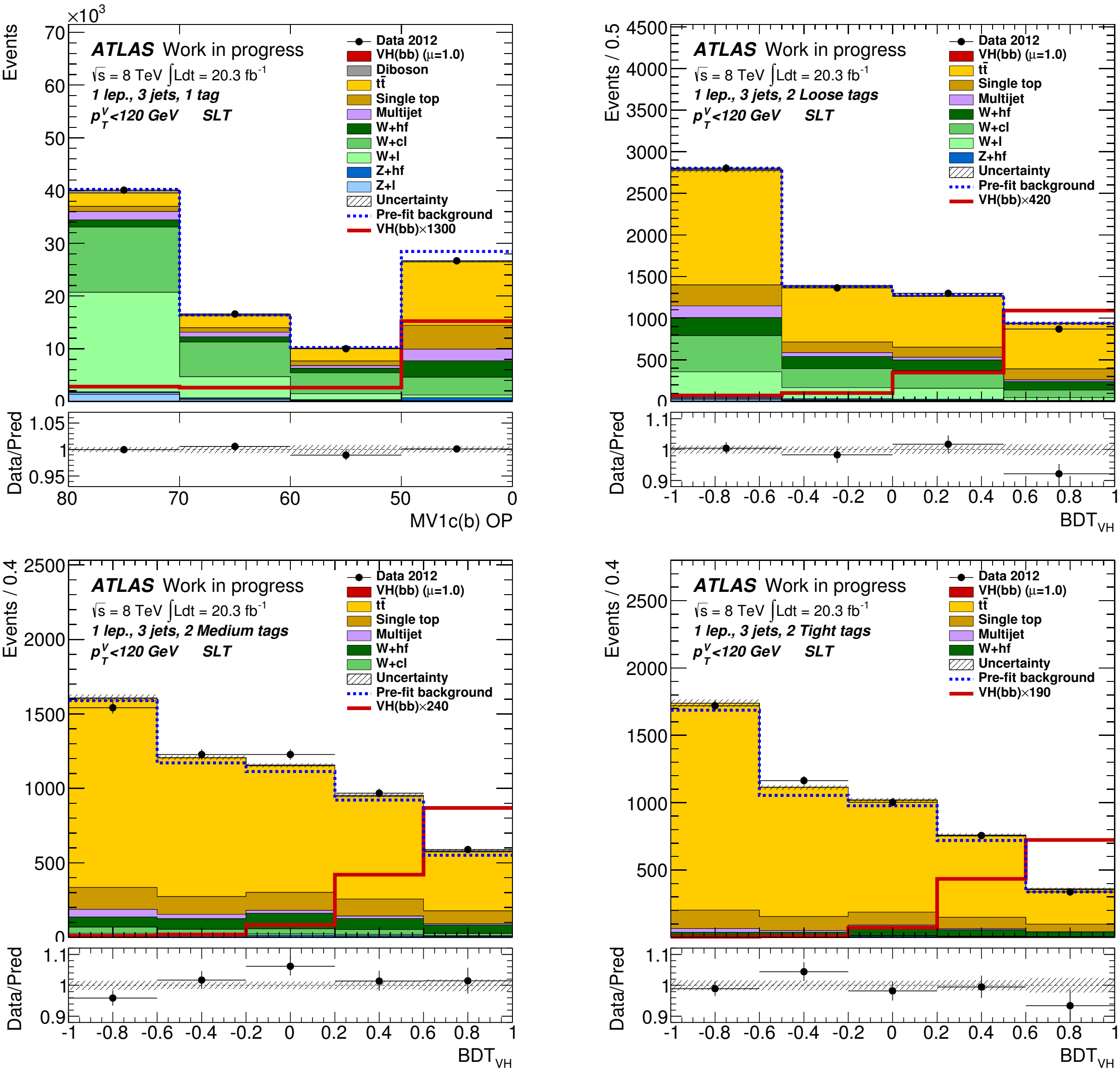


Fig. 5.21: The post-fit $MV1c$ distribution in the 1-tag b-tagging category (top-left) and BDT distributions in the LL (top-right), MM (bottom-left), and TT (bottom-right) b-tagging categories, in the 3-jet low p_{T}^{W} regions. The data/prediction ratios are shown in the lower panels.

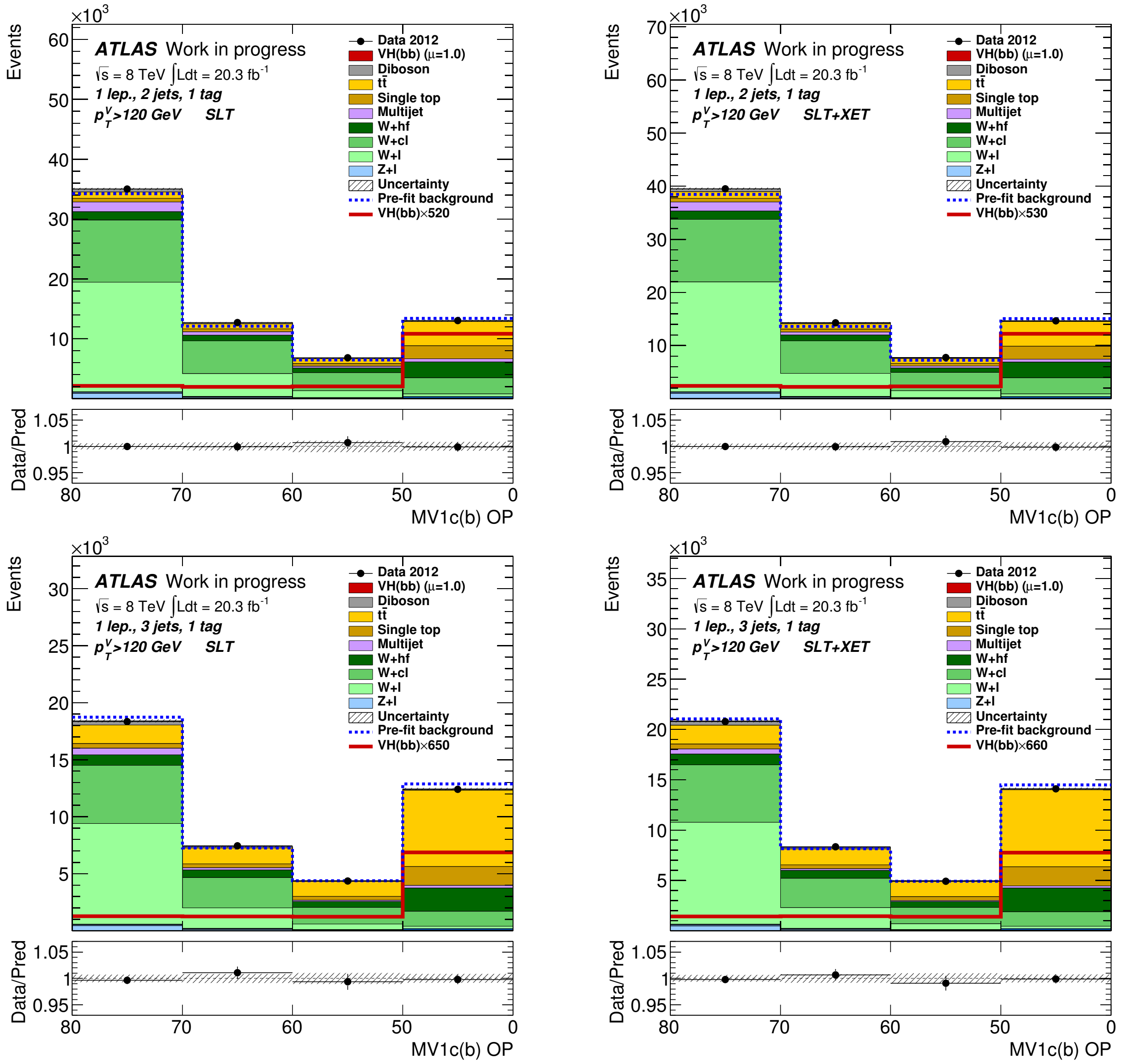


Fig. 5.22: The post-fit $MV1c$ distributions in the high p_{T}^{W} regions 1-tag b-tagging category. The top (bottom) row shows the distributions in the 2-jet (3-jet) bin, and the SLT-only and SLT+XET events are shown on the left and right columns, respectively. The data/prediction ratios are shown in the lower panels.

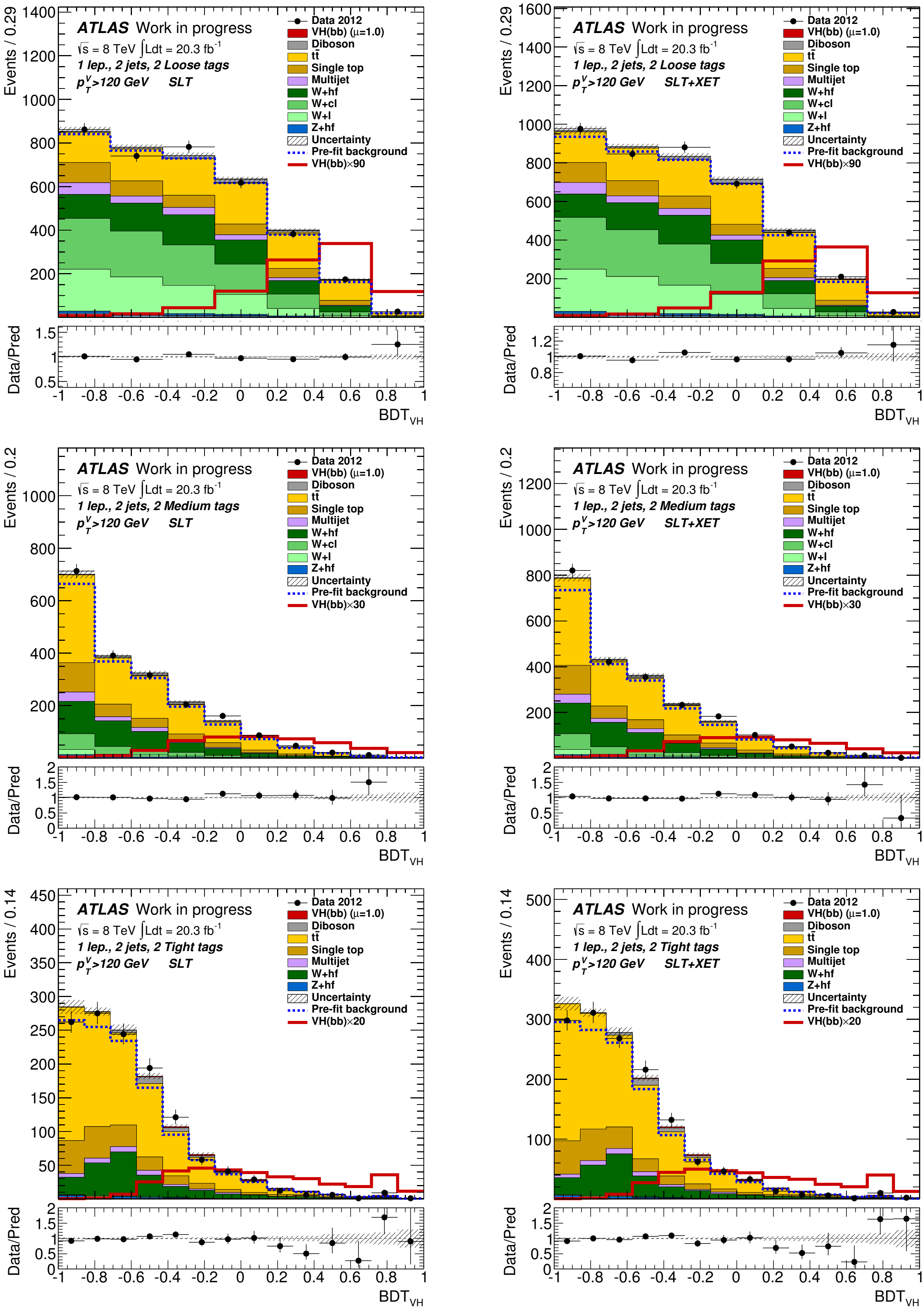


Fig. 5.23: The post-fit BDT distributions in the high p_{T}^{W} regions of 2-jet bin in the LL, MM, and TT b-tagging categories are shown in the top, middle, and bottom rows, respectively. The SLT-only (SLT+XET) events are shown on the left (right) column. The data/prediction ratios are shown in the lower panels.

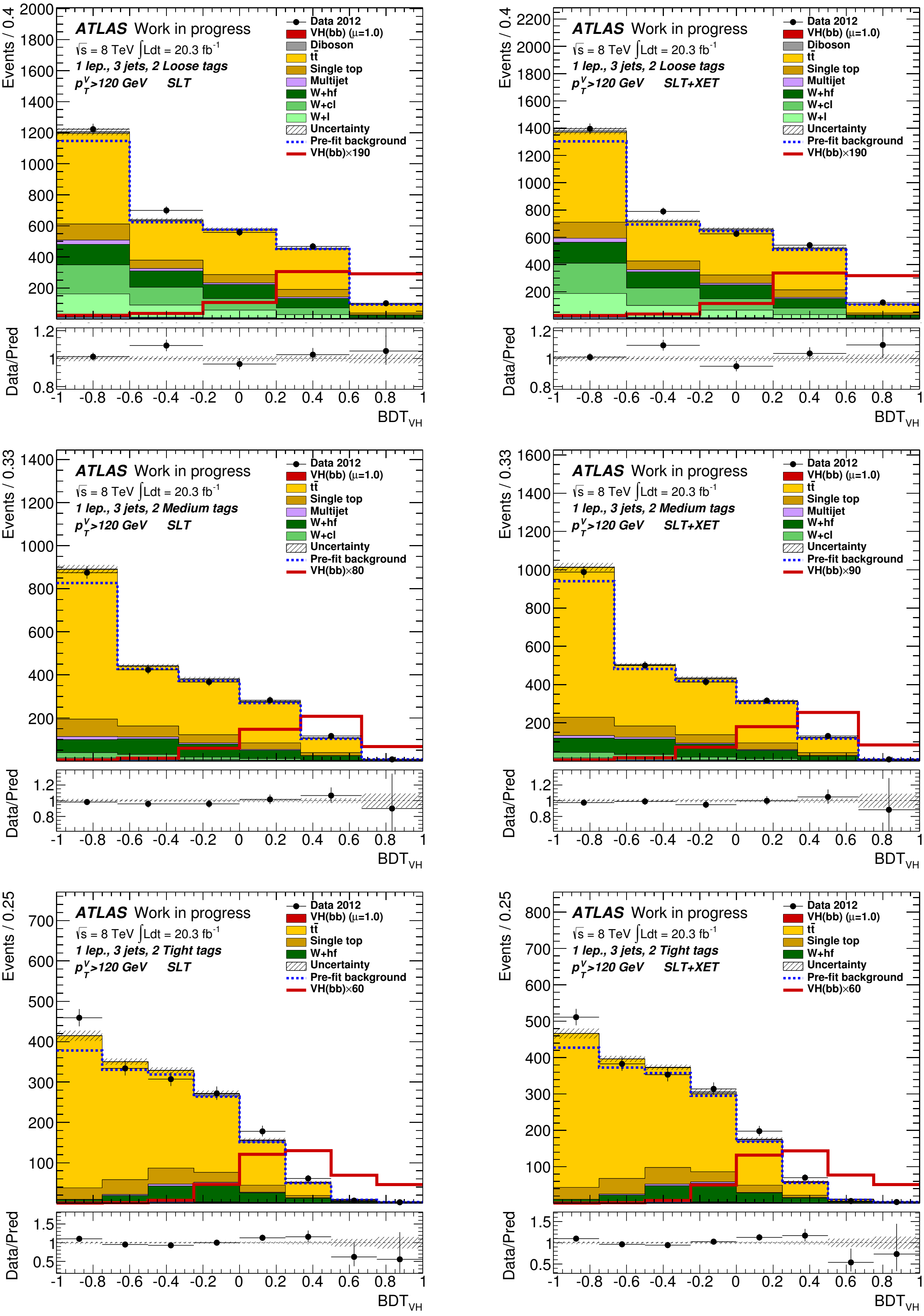


Fig. 5.24: The post-fit BDT distributions in the high p_{T}^{W} regions of 3-jet bin in the LL, MM, and TT b-tagging categories are shown in the top, middle, and bottom rows, respectively. The SLT-only (SLT+XET) events are shown on the left (right) column. The data/prediction ratios are shown in the lower panels.

Figure 5.25 shows the expected and observed 95% CL cross-section upper limits normalized to the SM Higgs boson production cross-section as a function of m_H, showing the $WH \to \ell\nu b\bar{b}$ combination results and the $VH \to Vb\bar{b}$ combination results. The expected and observed 95% CL upper limit

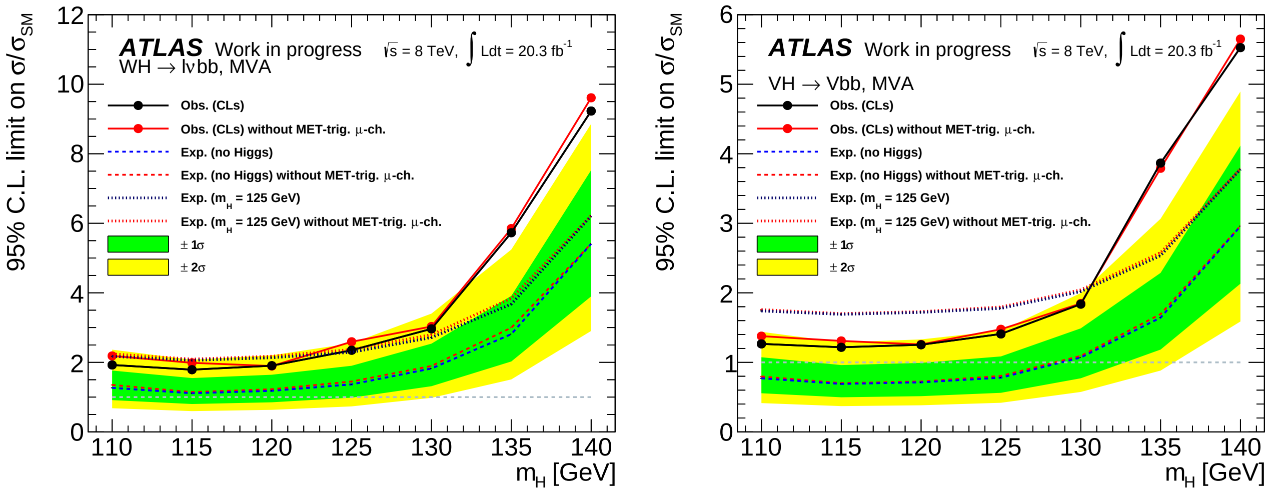


Fig. 5.25: Observed (solid line) and expected (dashed line) 95% CL cross-section upper limits, normalized to the SM Higgs boson production cross-section, as a function of the Higgs boson mass (m_H). The WH (VH) combination result is shown in the left (right) plot, using 8 TeV data. The expected upper limit is given for the background-only hypothesis (dashed) and with the injection of a SM Higgs boson signal at a mass of $125\ \text{GeV}$ (dotted). For these expected limits, the blue (red) curves are with (without) the $E_{\text{T}}^{\text{miss}}$ triggered muon channel. The green (yellow) error band is corresponding to the $1(2)\sigma$ ranges of the expectation in the absence of a signal, with the $E_{\text{T}}^{\text{miss}}$ triggered muon channel. The black (red) points show the observed limit with (without) the $E_{\text{T}}^{\text{miss}}$ triggered muon channel.

for the SM Higgs boson production cross-section times branching ratio at $m_H = 125$ GeV is listed in table 5.8, comparing the SLT-only results to the SLT+XET results. The gain from the $E_{\text{T}}^{\text{miss}}$ triggered muon channel can clearly be seen for the expected sensitivity. The improvements were 3.1% for the cut-based analysis and 5.8% for the MVA. The uncertainties are also reduced due to the additional statistics from the $E_{\text{T}}^{\text{miss}}$ triggered muon channel.

The expected and observed significance over the background-only hypothesis for the Higgs boson at $m_H = 125\ \text{GeV}$ is listed in table 5.9, comparing the SLT-only results to the SLT+XET results. The improvement due to the introduction of the XET signal region is similar to the one observed in cross-section limits.

5.11. Conclusions

A search for the $b\bar{b}$ decay of the Higgs boson in the $WH \to \ell\nu b\bar{b}$ channel has been performed and presented. The dataset used for this analysis corresponds to an integrated luminosity of 20.3 fb^{-1} from pp collisions at 8 TeV, recorded by the ATLAS experiment during LHC Run 1.

The analysis has been carried out in event categories based on the number of jets (two or three), b-tagged jets (one or two), and p_{T} of the W boson candidate. Two analysis methods, cut-based and

MVA ones, were performed to cross-check each other.

An excess over the background-only hypothesis has been observed, corresponding to the Higgs boson with $m_H = 125$ GeV. For a Higgs boson with $m_H = 125$ GeV, the observed (expected) significance corresponds to 1.8σ (1.5σ) and 2.7σ (1.3σ) for MVA and cut-based methods, respectively. The relative signal yields to the SM expectations are found to be $\mu = 1.06^{+0.73}_{-0.67}$ and $\mu = 2.2^{+0.97}_{-0.87}$ for MVA and cut-based methods, respectively.

An additional signal region, introduced by the author, was implemented in the published results, presented with other $ZH \to Zb\bar{b}$ searches [90]. The author contributed in all aspects of the $WH \to \ell\nu b\bar{b}$ channel, including an important work: the construction of the estimation method for the multijet background, which was not the primary one but had sizable contribution in the analysis regions. The additional signal region, the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel, improved the expected sensitivity for both cut-based and MVA methods. The gain by this additional signal region for the MVA (cut-based) method is 5.8% (3.1%) for the expected sensitivity for the $WH \to \ell\nu b\bar{b}$ search at $m_H = 125$ GeV. The combined results for the $VH \to Vb\bar{b}$ search are also improved by the additional $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered signal region and the sensitivity gain is 1.25% (0.94%) for the MVA (cut-based) method.

Table 5.8: The expected and observed 95% CL upper limits normalized to the SM Higgs boson production for $pp \to WH \to \ell\nu b\bar{b}$ and $pp \to VH \to Vb\bar{b}$ searches at $m_H = 125$ GeV using 8 TeV data. The total (stat.+syst.) uncertainties are shown for expected limits.

Combination	Analysis	exp. / obs.	SLT+XET	SLT-only	Gain
$WH \to \ell\nu b\bar{b}$	Cut-based	exp.	$1.63^{+0.64}_{-0.46}$	$1.68^{+0.66}_{-0.47}$	3.1%
		obs.	3.88	4.28	–
	MVA	exp.	$1.37^{+0.54}_{-0.38}$	$1.45^{+0.57}_{-0.40}$	5.8%
		obs.	2.35	2.59	–
$VH \to Vb\bar{b}$	Cut-based	exp.	$1.06^{+0.42}_{-0.30}$	$1.07^{+0.42}_{-0.30}$	0.94%
		obs.	2.14	2.29	–
	MVA	exp.	$0.78^{+0.31}_{-0.22}$	$0.80^{+0.32}_{-0.22}$	1.25%
		obs.	1.41	1.48	–

Table 5.9: The expected and observed significance over the background-only hypothesis for $pp \to VH \to Vb\bar{b}$ at $m_H = 125$ GeV using 8 TeV data. Numbers in the table is given in units of standard deviation, σ.

Combination	Analysis	exp. / obs.	SLT+XET	SLT-only	Gain
$WH \to \ell\nu b\bar{b}$	Cut-based	exp.	1.31	1.28	2.3%
		obs.	2.69	2.99	–
	MVA	exp.	1.52	1.44	5.6%
		obs.	1.81	1.69	–
$VH \to Vb\bar{b}$	Cut-based	exp.	1.94	1.93	0.5%
		obs.	2.02	2.23	–
	MVA	exp.	2.52	2.47	2.0%
		obs.	1.84	1.71	–

6. Conclusion

This thesis presented two major studies: development of high performance b-tagging algorithms and the search for the Higgs boson in the $WH \to \ell\nu b\bar{b}$ channel.

The identification of jets stemming from the b-quark's fragmentation, the so-called b-tagging, is a crucial ingredient for many physics analyses performed in the ATLAS experiment, especially for $H \to b\bar{b}$ searches. This thesis presented the development of high performance b-tagging algorithms using boosted decision trees (BDT). The input variables, reweighting scheme for kinematical difference ($p_{\mathrm{T}}^{\mathrm{jet}}$and η^{jet}) in the signal and background events, and the BDT structure were optimized, and the new algorithm developed in this thesis improves the b-tagging performance significantly. At the working point of 70% b-jet efficiency, the new algorithm, MV2c20 with the best c-jet rejection, improves the c-jet rejection by 11.5% while keeping the same light jet rejection in Run 1 condition. The MV2c20 is currently being used as the baseline b-tagging algorithm for Run 2 analyses.

After the discovery, followed by the confirmation for Higgs boson, all results available currently show the compatibility of the Higgs boson with the SM prediction at $m_H \sim 125$ GeV. The $H \to b\bar{b}$ decay mode has yet to be observed due to the overwhelming multijet background production in pp collisions. Associated production of vector bosons (W/Z) decaying leptonically ($Z \to \nu\nu, \ell\ell$ and $W \to \ell\nu$) was exploited by the $H \to b\bar{b}$ search for background reduction and triggering purposes. In this thesis, a search for the Higgs boson in the $WH \to \ell\nu b\bar{b}$ decay channel has been presented using pp collision data at $\sqrt{s} = 8$ TeV collected from the ATLAS detector in 2012, corresponding to an integrated luminosity of 20.3 fb^{-1}. This analysis is particularly important to fully confirm the nature of the Higgs boson. In spite of the largest branching ratio of $H \to b\bar{b}$ (~58%) at $m_H \sim 125$ GeV, this analysis is quite challenging due to the lower production cross-section of the associated (WH/ZH) production compared to the other production modes. The additional analysis region of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel, introduced by the author, improved the $WH \to \ell\nu b\bar{b}$ search sensitivity by 5% at $m_H = 125$ GeV and was included in the published results presented together with the ZH search results. Although the improvement was not very large, this gain was very important since the data taking had been stopped during the long shutdown while no clear $H \to b\bar{b}$ signal has been observed. An excess over the background-only hypothesis has been found with a significance of 1.8 σ while 1.5 σ was expected using multivariate analysis. The observed (expected) upper limit on the cross-section times branching ratio for $WH \to \ell\nu b\bar{b}$ at 95% confidence level was found to be 2.35 ($1.37^{+0.54}_{-0.38}$ with total uncertainty) times the SM prediction at $m_H = 125$ GeV. The results were found to be consistent with the SM prediction.

Bibliography

[1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. **B716** (2012) 1–29, `arXiv:1207.7214 [hep-ex]`.

[2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61, `arXiv:1207.7235 [hep-ex]`.

[3] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[4] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[5] Particle Data Group Collaboration, K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

[6] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214–215.

[7] G. Altarelli, *Partons in Quantum Chromodynamics*, Phys. Rept. **81** (1982) 1.

[8] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588.

[9] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168–171.

[10] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266.

[11] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, `arXiv:hep-ex/9807003 [hep-ex]`.

[12] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531–533.

[13] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657.

[14] P. Langacker, *Structure of the standard model*, Adv. Ser. Direct. High Energy Phys. **14** (1995) 15–36, `arXiv:hep-ph/0304186 [hep-ph]`.

[15] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323.

[16] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132–133.

[17] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[18] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585–587.

[19] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, `arXiv:1307.1347 [hep-ph]`.

[20] Tev4LHC Higgs working group Collaboration, *Standard Model Higgs cross sections at hadron colliders*,. `http://maltoni.home.cern.ch/maltoni/TeV4LHC/SM.html`.

[21] OPAL, DELPHI, LEP Working Group for Higgs boson searches, ALEPH, L3 Collaboration, R. Barate et al., *Search for the standard model Higgs boson at LEP*, Phys. Lett. **B565** (2003) 61–75, `arXiv:hep-ex/0306033 [hep-ex]`.

[22] CDF, D0 Collaboration, T. Aaltonen et al., *Higgs Boson Studies at the Tevatron*, Phys. Rev. **D88** no. 5, (2013) 052014, `arXiv:1303.6346 [hep-ex]`.

[23] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. **B726** (2013) 120–144, `arXiv:1307.1432 [hep-ex]`.

[24] CDF, D0 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*, Phys. Rev. Lett. **109** (2012) 071804, `arXiv:1207.6436 [hep-ex]`.

[25] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[26] T. Behnke, J. E. Brau, B. Foster, J. Fuster, M. Harrison, J. M. Paterson, M. Peskin, M. Stanitzki, N. Walker, and H. Yamamoto, *The International Linear Collider Technical Design Report - Volume 1: Executive Summary*, `arXiv:1306.6327 [physics.acc-ph]`.

[27] M. Aicheler, M. Aicheler, P. Burrows, M. Draper, T. Garvey, P. Lebrun, K. Peach, N. Phinney, H. Schmickler, D. Schulte, et al., *A Multi-TeV Linear Collider Based on CLIC Technology*, CERN-2012-007. `https://cds.cern.ch/record/1500095`.

[28] J. Stirling, *Parton Luminosity and Cross-section plots*,. `http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html`.

[29] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[30] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[31] ATLAS Collaboration, *Luminosity recorded by the ATLAS Experiment*,. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`.

[32] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, ATLAS-CONF-2014-032. `http://cds.cern.ch/record/1706245`.

[33] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007.

[34] C. Lapoire, *Operational Experience with the ATLAS Pixel Detector at the LHC*, ATL-INDET-PROC-2012-006. `https://cds.cern.ch/record/1425050`.

[35] ATLAS Collaboration, *dE/dx measurement in the ATLAS Pixel Detector and its use for particle identification*, ATLAS-CONF-2011-016. `https://cds.cern.ch/record/1336519`.

[36] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013. `https://cds.cern.ch/record/1291633`.

[37] ATLAS Collaboration, *ATLAS event at 13 TeV – First stable beam, 3 June 2015 – run: 266904, evt: 25884805*, ATLAS-PHO-Event-2015-016. `https://cds.cern.ch/record/2022202`.

[38] ATLAS Japan SCT working group Collaboration, T. Kondo, *Selected Photos of ATLAS Silicon Microstrip Modules*,. `http://atlas.kek.jp/sub/photos/SCTJ/SCTJModule/SCTJModule.html`.

[39] ATLAS Japan SCT working group Collaboration, T. Kondo, *Selected Photos of ATLAS SCT: Design*,. `http://atlas.kek.jp/sub/photos/SCTJ/SCTJDesign/SCTJDesign.html`.

[40] ATLAS Collaboration, *Commissioning of the ATLAS Muon Spectrometer with Cosmic Rays*, Eur.Phys.J. **C70** (2010) 875–916, `arXiv:1006.4384 [physics.ins-det]`.

[41] ATLAS Collaboration, *ATLAS: technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*, CERN-LHCC-94-43. `https://cds.cern.ch/record/290968`.

[42] ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report*, CERN-LHCC-98-014. `https://cds.cern.ch/record/381429`.

[43] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur.Phys.J. **C70** (2010) 823–874, `arXiv:1005.4568 [physics.ins-det]`.

[44] T. Sjostrand, S. Mrenna, and P. Z. Skands, PYTHIA *6.4 Physics and Manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175`.

[45] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput.Phys.Commun. **178** (2008) 852–867, `arXiv:0710.3820 [hep-ph]`.

[46] G. Corcella et al., *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including super-symmetric processes*, JHEP **01** (2001) 010.

[47] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. (2008) 639.

[48] M. L. Mangano et al., ALPGEN*, a generator for hard multi-parton processes in hadronic collisions*, JHEP **07** (2003) 001.

[49] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, Comput.Phys.Commun. **184** (2013) 919–985, `arXiv:hep-ph/0405247 [hep-ph]`.

[50] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, `arXiv:hep-ph/0204244 [hep-ph]`.

[51] Gleisberg, T. and Hoeche, Stefan. and Krauss, F. and Schonherr, M. and Schumann, S. and Siegert, F. and Winter, J., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, `arXiv:0811.4622 [hep-ph]`.

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

[53] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, `arXiv:0709.2092 [hep-ph]`.

[54] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, `arXiv:1002.2581 [hep-ph]`.

[55] Z. Was, *TAUOLA the library for tau lepton decay, and KKMC / KORALB / KORALZ /... status report*, Nucl.Phys.Proc.Suppl. **98** (2001) 96–102, `arXiv:hep-ph/0011305 [hep-ph]`.

[56] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152–155.

[57] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A506** (2003) 250–303.

[58] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, ATLAS-CONF-2012-042. `https://cds.cern.ch/record/1435196`.

[59] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment*, ATLAS-CONF-2010-069. `https://cds.cern.ch/record/1281344`.

[60] ATLAS Collaboration, *ATLAS Stand-Alone Event Displays*,. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayStandAlone`.

[61] ATLAS Collaboration, W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, et al., *Calorimeter clustering algorithms: Description and performance*, ATL-LARG-PUB-2008-002. `https://cds.cern.ch/record/1099735`.

[62] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. **C74** no. 10, (2014) 3071, `arXiv:1407.5063 [hep-ex]`.

[63] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, Eur. Phys. J. **C74** no. 11, (2014) 3130, `arXiv:1407.3935 [hep-ex]`.

[64] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV*, Eur. Phys. J. **C75** no. 7, (2015) 303, `arXiv:1412.7086 [hep-ex]`.

[65] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, `arXiv:0802.1189 [hep-ph]`.

[66] *Pile-up subtraction and suppression for jets in ATLAS*, ATLAS-CONF-2013-083. `https://cds.cern.ch/record/1570994`.

[67] *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018. `https://cds.cern.ch/record/1700870`.

[68] ATLAS Collaboration, *Measurement of the missing transverse momentum based on tracks in proton-proton collisions at $\sqrt{s} = 900$ GeV centre-of-mass energy with the ATLAS detector*, ATLAS-CONF-2010-020. `https://cds.cern.ch/record/1277652`.

[69] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV with ATLAS*, Eur. Phys. J **C72** (2012) 1844, `arXiv:1108.5602 [hep-ex]`.

[70] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*, ATLAS-CONF-2013-082. `https://cds.cern.ch/record/1570993`.

[71] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, ATLAS-PERF-2012-04-002,CERN-PH-EP-2015-216. `https://cds.cern.ch/record/2042204`. Currently ATLAS internal, soon to be public.

[72] Linear Collider ACFA Working Group Collaboration, K. Abe et al., *Particle physics experiments at JLC*, `arXiv:hep-ph/0109166 [hep-ph]`.

[73] ATLAS Collaboration, *Performance of Impact Parameter-Based b-tagging Algorithms with the ATLAS Detector using Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2010-091. `https://cds.cern.ch/record/1299106`.

[74] L. Alio, *Search for the Higgs boson produced in association with a Z boson and decaying to a pair of bottom quarks with the ATLAS experiment at LHC*. PhD thesis, Marseille, CPPM, 2014. `https://cds.cern.ch/record/2012603`. presented 12 Nov 2014.

[75] P. Jaiswal, K. Kopp, and T. Okui, *Higgs Production Amidst the LHC Detector*, Phys.Rev. **D87** no. 11, (2013) 115017, `arXiv:1303.1181 [hep-ph]`.

[76] *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022. `https://cds.cern.ch/record/2037697`.

[77] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics*, `arXiv:0901.0512`.

[78] Hoecker, A. et al., *TMVA - Toolkit for Multivariate Data Analysis*, `arXiv:physics/0703039`.

[79] Breiman, Leo and Friedman, Jerome and Olshen, R. A. and Stone, Charles J., *Classification and regression trees*, , Stanford, 1984. Wadsworth.

[80] Y. Coadou, *Boosted Decision Trees and Applications*, EPJ Web of Conferences **55** (2013).

[81] Y. Freund and R. E. Schapire, *Experiments with a New Boosting Algorithm*, , San Fransisco, 1996, in *Proceedings of the Thirteenth International Conference on Machine Learning*, L. Saitta, ed., 148. Morgan Kaufmann.

[82] C. Peterson, T. Rognvaldsson, and L. Lonnblad, *JETNET 3.0: A Versatile artificial neural network package*, Comput.Phys.Commun. **81** (1994) 185–220.

[83] M. Ughetto, *Recherche de production électrofaible supersymétrique dans des cascades de désintégration contenant un boson de Higgs avec le détecteur ATLAS au LHC*. PhD thesis, Marseille, CPPM, 2014. presented 7 Nov 2014.

[84] ATLAS Collaboration, *Measuring the b-tag efficiency in a top-pair sample with 4.7* fb^{-1} *of data from the ATLAS detector*, ATLAS-CONF-2012-097. `https://cds.cern.ch/record/1460443`.

[85] ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, ATLAS-CONF-2014-004. `https://cds.cern.ch/record/1664335`.

[86] ATLAS Collaboration, *Calibration of the b-tagging efficiency for c jets with the ATLAS detector using events with a W boson produced in association with a single c quark*, ATLAS-CONF-2013-109. `http://cds.cern.ch/record/1640162`.

[87] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046Jul. `http://cds.cern.ch/record/1741020`.

[88] ATLAS Collaboration, *Continuous b-tagging for the ATLAS experiment*, ATL-COM-PHYS-2014-035. `https://cds.cern.ch/record/1644430`. Currently ATLAS internal, soon to be public.

[89] *Commissioning of the ATLAS b-tagging algorithms using $t\bar{t}$ events in early Run-2 data*, ATL-PHYS-PUB-2015-039. `https://cds.cern.ch/record/2047871`.

[90] 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*, JHEP **1501** (2015) 069, `arXiv:1409.6212 [hep-ex]`.

[91] K.Mochizuki, *Search for the Higgs boson in fermionic channels using ATLAS detector*, J. Phys. Conf. Ser. **623** no. 1, (2015) 012020.

[92] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, `arXiv:hep-ph/0201195 [hep-ph]`.

[93] ATLAS Collaboration, *ATLAS tunes of PYTHIA6 and PYTHIA8 for MC11*, ATL-PHYS-PUB-2011-009. `http://cdsweb.cern.ch/record/1363300`.

[94] ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data*, ATL-PHYS-PUB-2011-008. `https://cdsweb.cern.ch/record/1345343`.

[95] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, Eur.Phys.J. **C45** (2006) 97–107, `arXiv:hep-ph/0506026 [hep-ph]`.

[96] G. Ferrera, M. Grazzini, and F. Tramontano, *Associated WH production at hadron colliders: a fully exclusive QCD calculation at NNLO*, Phys.Rev.Lett. **107** (2011) 152003, `arXiv:1107.1164 [hep-ph]`.

[97] M. Ciccolini, S. Dittmaier, and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys.Rev. **D68** (2003) 073003, `arXiv:hep-ph/0306234 [hep-ph]`.

[98] A. Denner, S. Dittmaier, S. Kallweit, and A. Muck, *EW corrections to Higgs strahlung at the Tevatron and the LHC with HAWK*, PoS **EPS-HEP2011** (2011) 235, `arXiv:1112.5258 [hep-ph]`.

[99] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56–74.

[100] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024.

[101] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$*, Phys.Rev. **D74** (2006) 114017, `arXiv:hep-ph/0609070 [hep-ph]`.

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

[103] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys.Rev. **D83** (2011) 091503, `arXiv:1103.2792 [hep-ph]`.

[104] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys.Rev. **D82** (2010) 054018, `arXiv:1005.4451 [hep-ph]`.

[105] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys.Rev. **D81** (2010) 054028, `arXiv:1001.5034 [hep-ph]`.

[106] P. Nason and G. Zanderighi, *W^+W^-, WZ and ZZ production in the POWHEG-BOX-V2*, Eur.Phys.J. **C74** no. 1, (2014) 2702, `arXiv:1311.1365 [hep-ph]`.

[107] A. Sherstnev and R. Thorne, *Parton Distributions for LO Generators*, Eur. Phys. J. **C55** (2008) 553–575, `arXiv:0711.2473 [hep-ph]`.

[108] ATLAS Collaboration, *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\sqrt{s} = 7$ TeV in 2011*, ATLAS-CONF-2012-048. `https://cds.cern.ch/record/1450089`.

[109] ATLAS Collaboration, *Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV*, Eur.Phys.J. **C75** no. 3, (2015) 120, `arXiv:1408.3179 [hep-ex]`.

[110] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur.Phys.J. **C74** no. 7, (2014) 2941, `arXiv:1404.2240 [hep-ex]`.

[111] ATLAS Collaboration, *Muon reconstruction efficiency and momentum resolution of the ATLAS experiment in proton-proton collisions at $\sqrt{s}$ = 7 TeV in 2010*, Eur.Phys.J. **C74** no. 9, (2014) 3034, `arXiv:1404.4562 [hep-ex]`.

[112] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys.Lett. **B659** (2008) 119–126, `arXiv:0707.1378 [hep-ph]`.

[113] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Eur.Phys.J. **C73** no. 3, (2013) 2304, `arXiv:1112.6426 [hep-ex]`.

[114] J. Boudreau, *Instrumental Backgrounds to $t\bar{t}$ and Single Top Production at Hadron Colliders*, ATL-PHYS-PROC-2013-077. `https://cds.cern.ch/record/1528551`.

[115] ATLAS Collaboration, *Measurements of normalized differential cross sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys.Rev. **D90** no. 7, (2014) 072004, `arXiv:1407.0371 [hep-ex]`.

[116] ATLAS Collaboration, *Search for the bb decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector*, ATLAS-CONF-2013-079. `https://cds.cern.ch/record/1563235`.

[117] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, `arXiv:1007.1727v2 [physics.data-an]`.

[118] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur.Phys.J. **C73** no. 3, (2013) 2306, `arXiv:1210.6210 [hep-ex]`.

[119] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur.Phys.J. **C73** no. 8, (2013) 2518, `arXiv:1302.4393 [hep-ex]`.

[120] N. Kidonakis, *Differential and total cross sections for top pair and single top production*, `arXiv:1205.3453 [hep-ph]`.

[121] J. M. Campbell and R. K. Ellis, *MCFM for the Tevatron and the LHC*, Nucl. Phys. Proc. Suppl. **205-206** (2010) 10–15, `arXiv:1007.3492 [hep-ph]`.

[122] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, *The RooStats Project*, PoS**ACAT2010** 057, `arXiv:1009.1003 [physics.data-an]`.

[123] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, in *2003 Computing in High Energy and Nuclear Physics, CHEP03*. 2003. `arXiv:physics/0306116`.

[124] A. L. Read, *Presentation of search results: The CL(s) technique*, J.Phys. **G28** (2002) 2693–2704.

A.Auxiliary materials for MV2 development

A.1. MV2 input variable distributions in Run-1 condition

In the MV2 algorithm, the 24 input variables are fed to BDT and all of them were listed in table 4.3. Here, the input variable distributions are shown, comparing b-, c-, and light jet flavors.

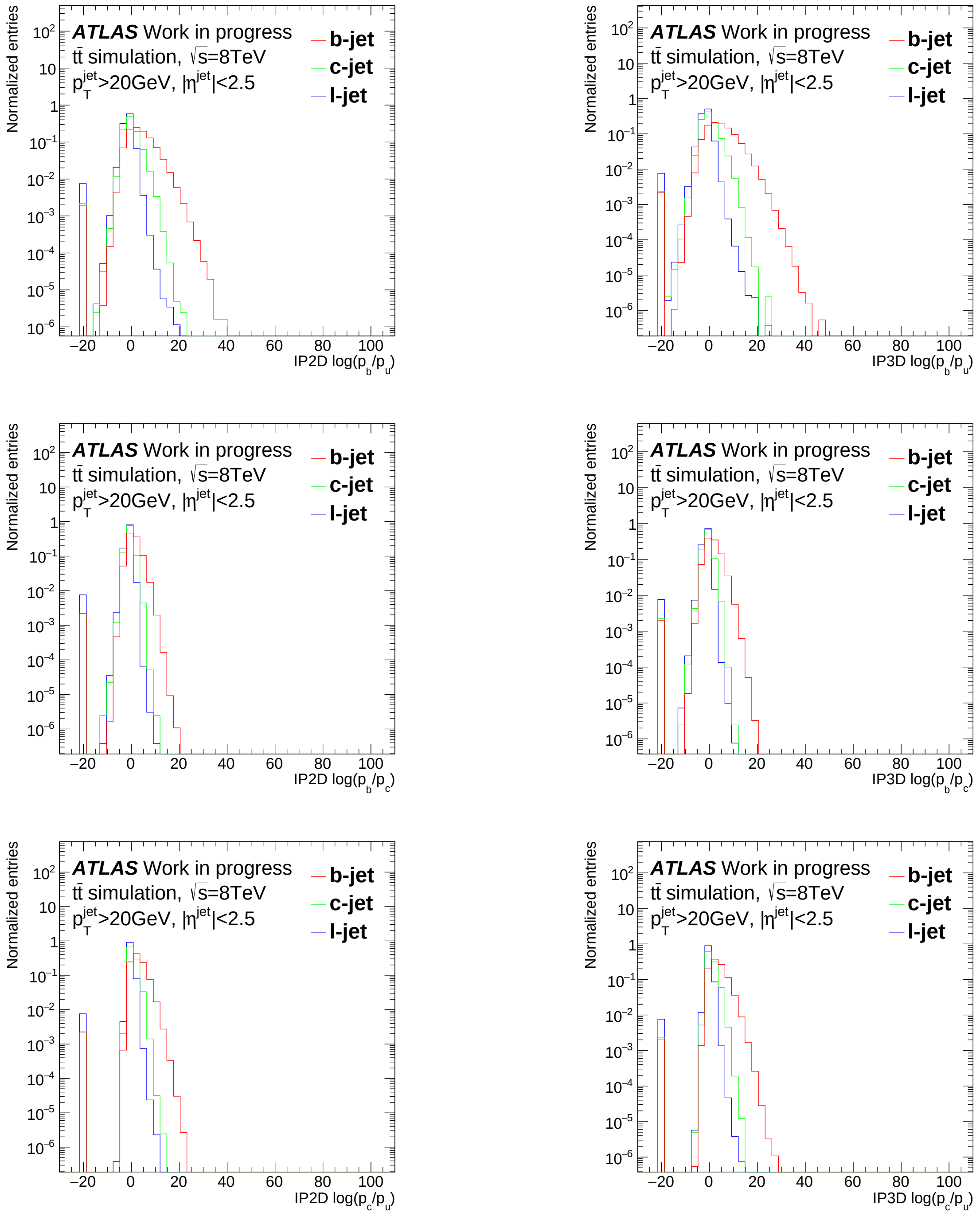


Fig. A.1: The MV2 input variable distributions from IP taggers. The left (right) distributions show the IP2D (IP3D) tagger's log-likelihood ratios, comparing b-, c-, and light jet flavors shown in red, green, and blue, respectively. The top, middle, bottom plots show the log-likelihood ratios of $log(p_b/p_u)$, $log(p_b/p_c)$, and $log(p_c/p_u)$, respectively.

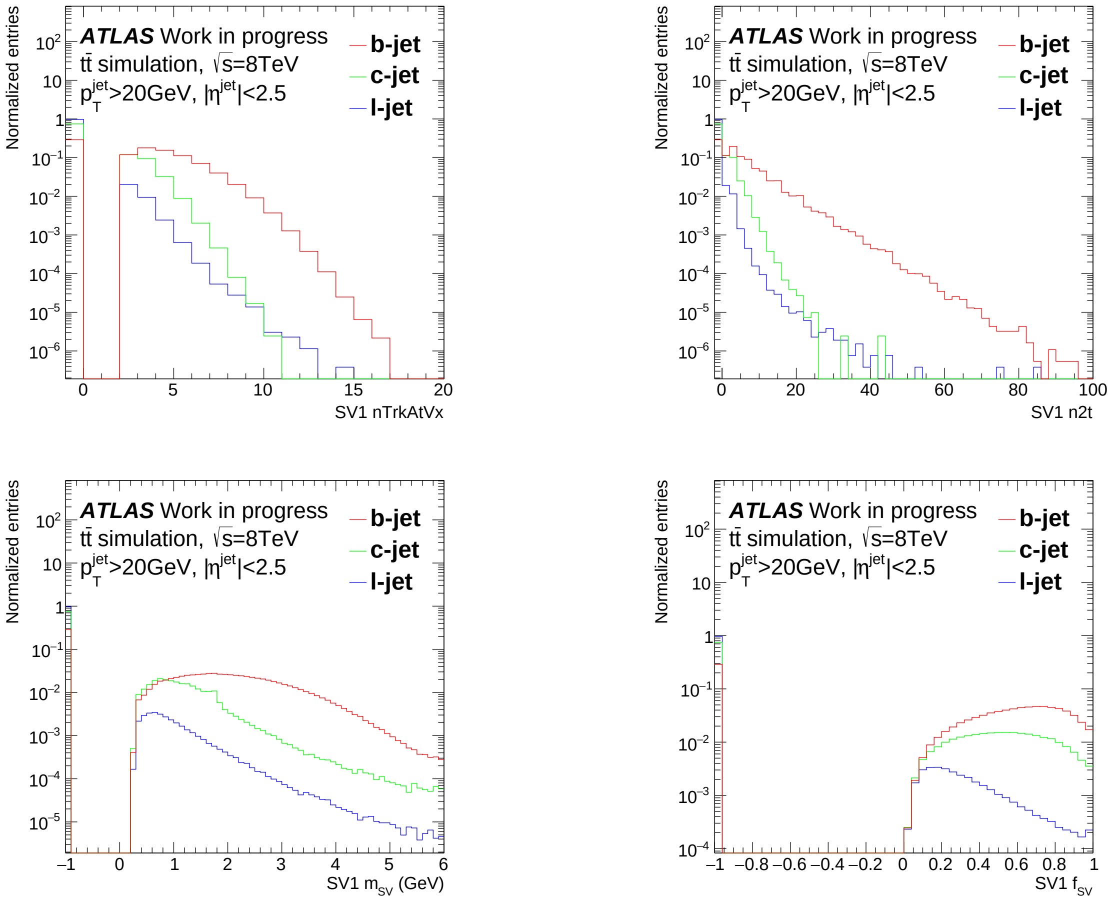


Fig. A.2: The MV2 input variable distributions from SV tagger, comparing b-, c-, and light jet flavors shown in red, green, and blue, respectively. The top left and right plots show the distributions of number of tracks associated to the reconstructed secondary vertex, and the number of two-track-pairs to form a vertex, respectively. The bottom left and right plots show the distributions of m_{SV} and f_{SV} for the reconstructed secondary vertex, respectively.

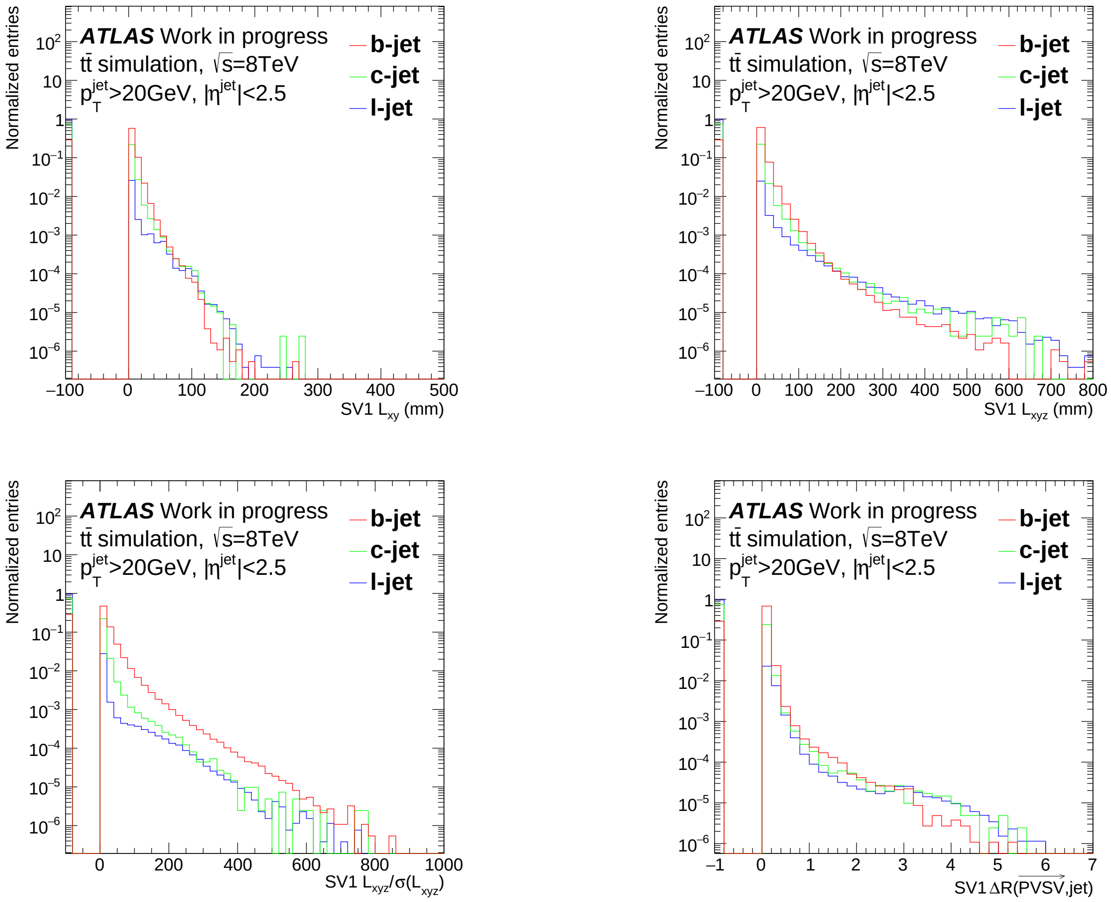


Fig. A.3: The MV2 input variable distributions from SV tagger, comparing b-, c-, and light jet flavors shown in red, green, and blue, respectively. The top left and right plots show the distributions of transverse and 3D decay length of the reconstructed secondary vertex, respectively. The bottom left plot shows the distribution of 3D decay length significance, and the bottom right plot shows the angular distance between the jet-axis and a vector from primary vertex to the reconstructed secondary vertex.

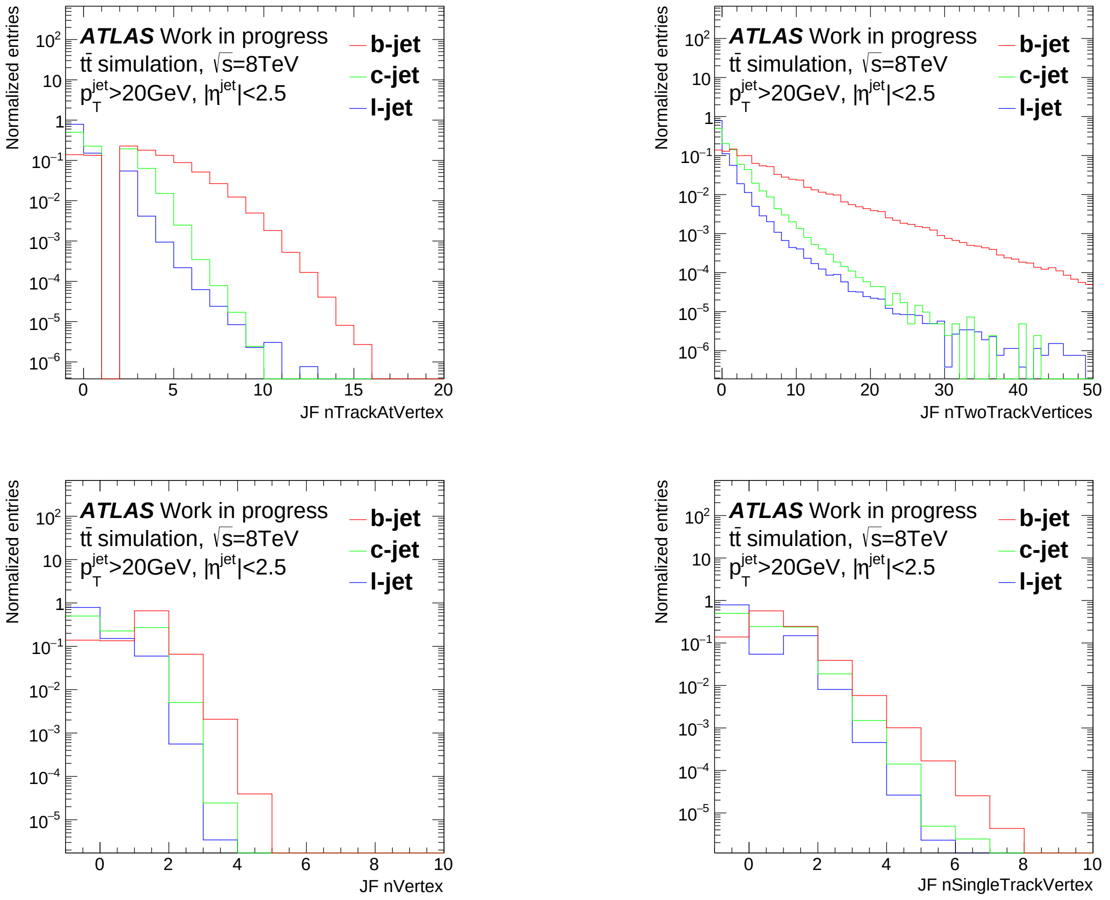


Fig. A.4: The MV2 input variable distributions from JetFitter, comparing b-, c-, and light jet flavors shown in red, green, and blue, respectively. The top left and right plots show the distributions of the number of tracks associated to the reconstructed secondary vertex and the number of 2-track-pair forming a vertex candidate, respectively. The bottom left plot shows the distribution of the number of reconstructed vertices with more than one tracks, and the bottom right plot shows the distribution of the number of single-track-vertices.

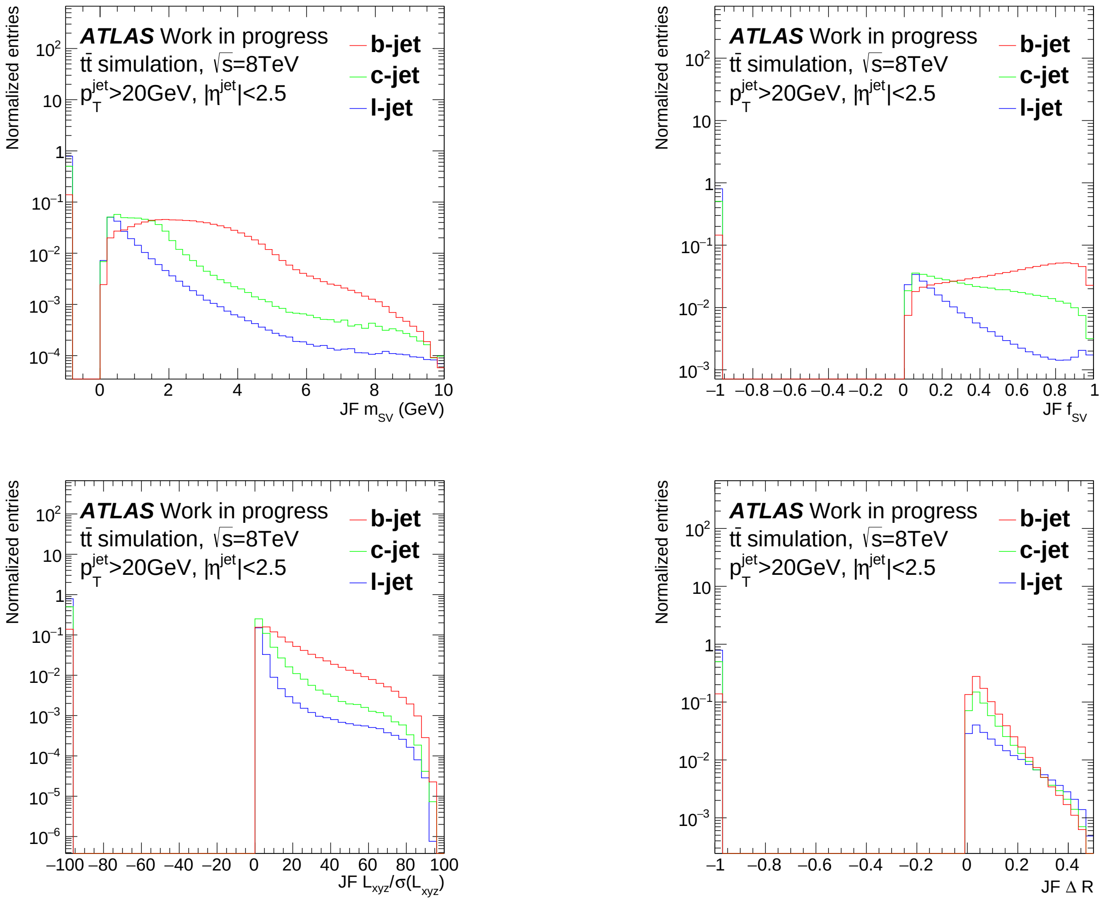


Fig. A.5: The MV2 input variable distributions from JetFitter, comparing b-, c-, and light jet flavors shown in red, green, and blue, respectively. The top left and right plots show the distributions of m_{SV} and f_{SV}, respectively. The bottom left plot shows the distribution of the decay length significance and the bottom right plot shows the distribution of the angular distance between the jet-axis and the sum of the vectors from the primary vertex to the reconstructed secondary vertices.

A.2. Multi-classification tagger: MV2m

The TMVA package also provides multi-classification MVA methods based on NN and BDT. Due to the huge training time required for the NN method, the BDT method was chosen for the baseline of a multi-classification tagger.

Using the same set of training samples and conditions as the MV2c20 tagger's setting, the **MV2m** tagger was developed with the multi-classification BDT applied for b-, c-, and light jet classifications. It provides the three probability-like output, p_b, p_c, and p_u, for b-, c-, and light jet flavors, and the sum is always kept at one: $p_b + p_c + p_u = 1$. The three probability distributions are shown in fig A.6.

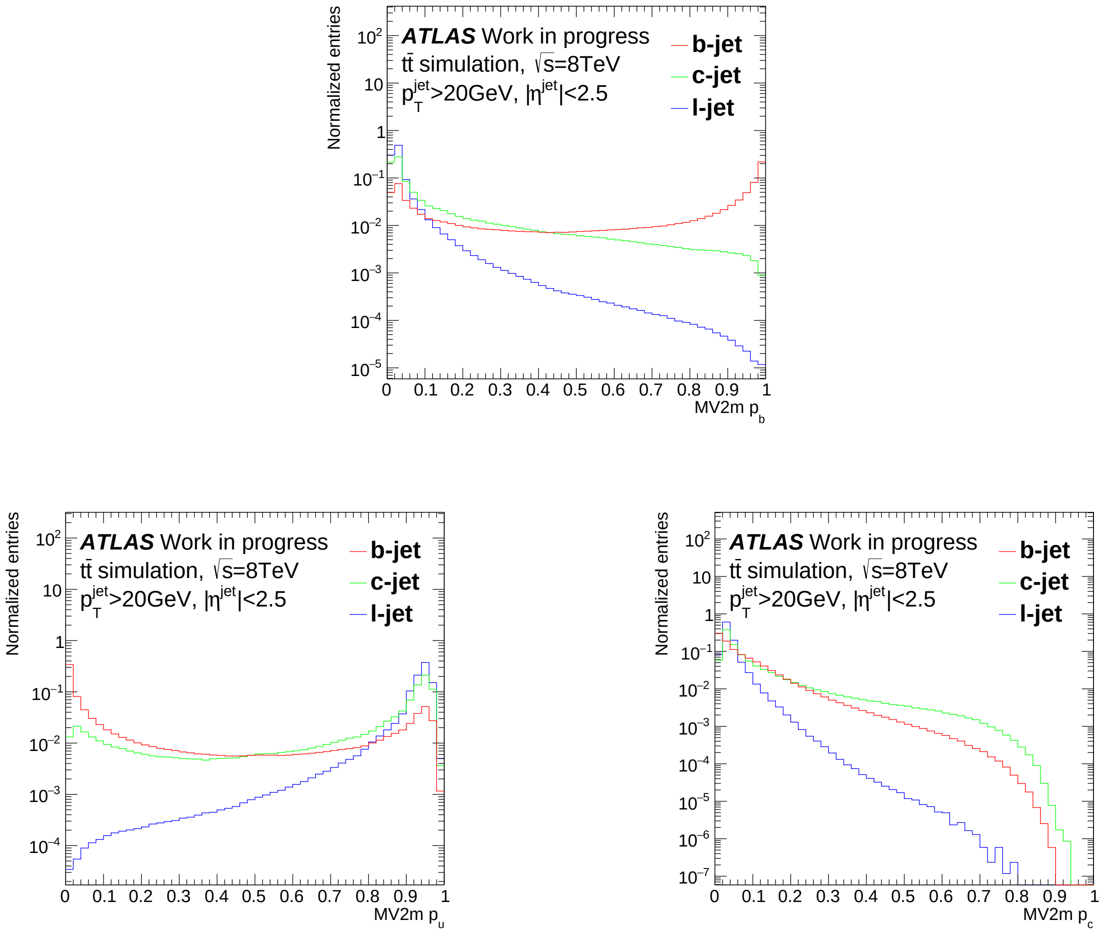


Fig. A.6: The three output put probability distributions of MV2m tagger, comparing b-, c-, and light jet flavors in red, green, and blue, respectively. The top, bottom left, and bottom right plots show the p_b, p_u, and p_c distributions, respectively.

Although more study is required for the optimization of MV2m usage in actual analyses, for example like 2D-cut in the $\log(p_c/p_u)$-$\log(p_c/p_b)$ plane (see fig A.7), the simple combinations of three output values by $\log(p_c/(p_u+p_b))$, and $\log(p_c \cdot p_c/(p_u \cdot p_b))$ show suitable c-tagging performance, which are shown in table A.1. Figure A.8 shows the output distribution of the MV2m tagger.

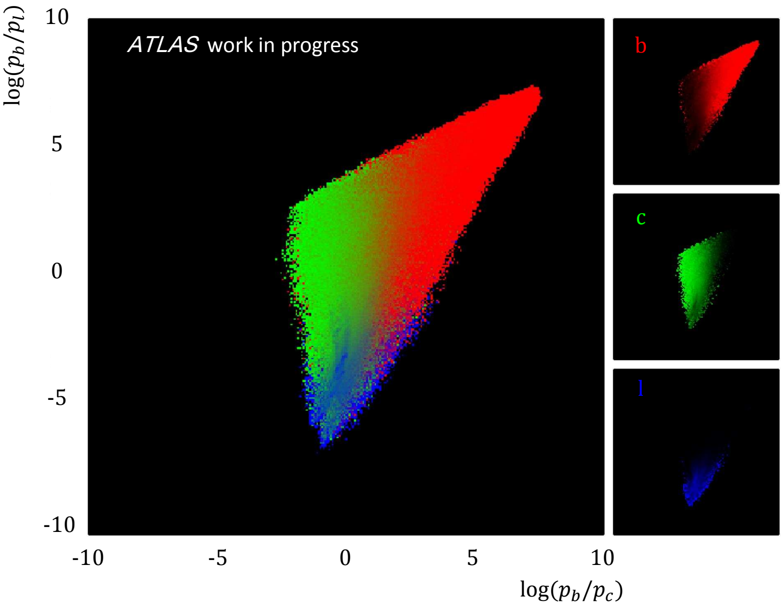


Fig. A.7: A 2D Map of b-, c-, and light flavors in the $\log(p_b/p_c)$ vs $\log(p_b/p_l)$ plane using MV2m tagger.

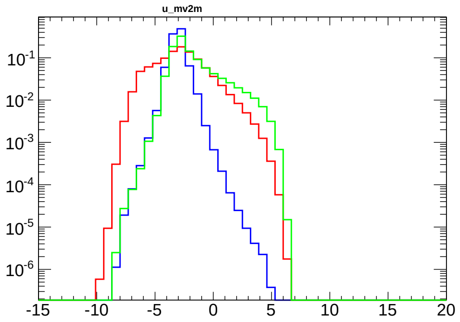


Fig. A.8: An example of combination of three output probabilities into 1D variable by $\log(p_c \cdot p_c/(p_u \cdot p_b))$. The b-, c-, light jet tracks are shown with red, green, and blue, respectively.

Table A.1: The rejection factors for b-jet and light jet background for the c-jet efficiency 20 and 30% working points. The two combination ways of $\log(p_c/(p_u+p_b))$ and $\log(p_c \cdot p_c/(p_u \cdot p_b))$ are shown. Results using MV2c20 tagger are shown only for the reference.

Combination	c-jet efficiency	b-jet rejection	light jet rejection
MV2m $\log(p_c/(p_u+p_b))$	20 %	7.54	396
	30 %	4.25	61.7
MV2m $\log(p_c \cdot p_c/(p_u \cdot p_b))$	20 %	10.3	267
	30 %	5.42	66.7
MV2c20	30 %	1.2	74
	20 %	1.3	204

B.Auxiliary materials for $WH \to \ell\nu b\bar{b}$ analysis

B.1. The expected yield of signal and fitted background

The following shows the number of expected signal and background events after the global fit. The MVA yield table is first shown followed by the yield number in cut-based analysis in 1-tag, LL, MM, and TT b-tagging categories, shown in this order.

B.2. BDT analysis input variables distributions

This section shows the important input distributions for BDT. In figs B.1–B.6, the m_{jj}, ΔR_{jj}, p_{T}^{W}, m_{T}^{W}, $E_{\mathrm{T}}^{\mathrm{miss}}$, and $\Delta\phi(W, bb)$ distributions in low p_{T}^{W} region are shown, respectively.

In figs B.7–B.12, the m_{jj}, ΔR_{jj}, p_{T}^{W}, m_{T}^{W}, $E_{\mathrm{T}}^{\mathrm{miss}}$, and $\Delta\phi(W, bb)$ distributions in high p_{T}^{W} region are shown, respectively.

Sample	1-tag		2L-tag		2M-tag		2T-tag	
	$p_{\mathrm{T}}^{V} < 120\ GeV$	$p_{\mathrm{T}}^{V} > 120\ GeV$	$p_{\mathrm{T}}^{V} < 120\ GeV$	$p_{\mathrm{T}}^{V} > 120\ GeV$	$p_{\mathrm{T}}^{V} < 120\ GeV$	$p_{\mathrm{T}}^{V} > 120\ GeV$	$p_{\mathrm{T}}^{V} < 120\ GeV$	$p_{\mathrm{T}}^{V} > 120\ GeV$
	2-jet							
VH	45.3 ± 26.6	36.2 ± 21.3	12.8 ± 7.5	11.1 ± 6.5	20.2 ± 11.8	17.2 ± 10.0	22.5 ± 13.1	19.7 ± 11.4
VV	1932.7 ± 52.4	1073.8 ± 117.5	102.9 ± 2.8	61.4 ± 1.7	64.1 ± 2.7	36.4 ± 6.0	50.6 ± 4.5	29.7 ± 5.8
$t\bar{t}$	11645.7 ± 887.5	7812.4 ± 641.1	1913.2 ± 138.6	1039.9 ± 77.9	2421.6 ± 169.0	1078.7 ± 75.4	2287.6 ± 157.1	882.9 ± 59.7
s-top	13553.3 ± 2482.5	4164.5 ± 380.8	1105.3 ± 178.7	380.0 ± 33.9	936.9 ± 134.8	289.3 ± 26.6	712.7 ± 99.9	205.4 ± 22.3
W+l	65796.8 ± 4616.8	26844.0 ± 1779.2	1620.6 ± 171.9	706.6 ± 74.8	122.2 ± 21.7	48.6 ± 8.0	3.1 ± 0.1	0.9 ± 0.0
W+cl	72171.0 ± 5453.5	24399.2 ± 2025.9	2647.9 ± 186.0	970.3 ± 74.8	460.4 ± 34.9	156.9 ± 13.1	21.2 ± 3.8	6.4 ± 1.2
W+hf	15662.1 ± 3015.8	6397.8 ± 1203.9	1492.3 ± 278.3	638.2 ± 126.3	1012.2 ± 193.7	413.7 ± 84.0	535.8 ± 109.7	240.7 ± 51.5
Z+l	3864.7 ± 355.0	1378.5 ± 137.7	101.6 ± 12.3	37.1 ± 1.0	7.3 ± 0.2	2.8 ± 0.1	0.1 ± 0.0	0.0 ± 0.0
Z+cl	1530.7 ± 41.5	423.6 ± 11.5	60.0 ± 1.6	19.9 ± 0.5	11.5 ± 0.3	3.5 ± 0.1	0.6 ± 0.0	0.2 ± 0.0
Z+hf	1854.9 ± 436.3	481.2 ± 13.0	147.6 ± 4.0	40.9 ± 1.1	108.9 ± 29.9	27.3 ± 5.1	75.0 ± 21.6	19.1 ± 5.5
MJ$_e$	–	3028.3 ± 1100.0	–	160.4 ± 54.1	–	86.7 ± 36.2	–	38.5 ± 19.6
MJ$_\mu$	10485.5 ± 1260.7	164.1	739.7 ± 140.0	16.0	336.2 ± 86.9	5.1 ± 0.0	184.2 ± 73.4	1.3 ± 0.9
Total	198542.9 ± 486.0	76203.8 ± 306.7	9943.8 ± 91.2	4081.7 ± 51.2	5501.7 ± 66.0	2166.1 ± 33.3	3893.4 ± 55.8	1444.8 ± 26.9
Data	198544.0 ± 445.6	76219.0 ± 276.1	9941.0 ± 99.7	4072.0 ± 63.8	5499.0 ± 74.2	2199.0 ± 46.9	3923.0 ± 62.6	1405.0 ± 37.5
$\frac{\mathrm{Sim-Data}}{\sigma}$	-0.0	-0.04	0.02	0.12	0.03	-0.57	-0.35	0.85
	3-jet							
VH	17.6 ± 10.4	18.2 ± 10.7	3.8 ± 2.2	4.4 ± 2.6	5.8 ± 3.4	6.9 ± 4.1	6.4 ± 3.8	7.8 ± 4.6
VV	1087.9 ± 29.5	756.7 ± 84.3	49.5 ± 1.3	39.3 ± 1.1	22.5 ± 0.6	17.7 ± 1.8	14.0 ± 0.4	13.8 ± 1.9
$t\bar{t}$	19141.9 ± 733.5	12594.2 ± 736.5	3196.4 ± 94.7	1639.4 ± 76.0	4134.4 ± 87.7	1691.5 ± 63.1	4189.5 ± 84.4	1400.1 ± 47.2
s-top	7228.3 ± 463.5	3176.8 ± 322.7	633.5 ± 39.3	312.1 ± 32.0	589.8 ± 39.2	248.3 ± 26.8	479.3 ± 33.1	177.9 ± 19.9
W+l	24843.2 ± 1828.1	12999.8 ± 926.1	590.6 ± 65.5	337.8 ± 37.6	39.4 ± 1.1	22.2 ± 4.0	1.0 ± 0.0	0.4 ± 0.0
W+cl	25843.0 ± 2152.1	11690.9 ± 1041.3	957.1 ± 72.6	477.9 ± 39.3	158.9 ± 12.8	72.4 ± 6.4	7.2 ± 0.7	3.2 ± 0.3
W+hf	6427.9 ± 1247.5	4806.4 ± 897.4	595.6 ± 112.8	469.5 ± 88.0	374.4 ± 76.2	294.5 ± 56.4	201.7 ± 42.1	166.2 ± 34.5
Z+l	1811.4 ± 204.4	662.3 ± 80.7	43.0 ± 1.2	18.0 ± 0.5	2.8 ± 0.1	1.4 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Z+cl	785.7 ± 21.3	243.0 ± 6.6	30.8 ± 0.8	11.2 ± 0.3	5.1 ± 0.1	1.8 ± 0.0	0.2 ± 0.0	0.1 ± 0.0
Z+hf	940.4 ± 25.5	274.8 ± 7.5	70.7 ± 1.9	23.7 ± 0.6	49.2 ± 1.3	17.1 ± 4.3	37.1 ± 1.0	12.3 ± 4.1
MJ$_e$	–	880.5 ± 398.7	–	60.1 ± 23.8	–	30.9 ± 14.5	–	12.9 ± 5.7
MJ$_\mu$	5209.1 ± 528.4	91.1	237.7 ± 39.9	4.9	121.8 ± 17.4	3.4 ± 0.6	59.8 ± 8.8	0.8 ± 0.0
Total	93336.5 ± 307.0	48194.9 ± 213.1	6408.5 ± 60.9	3398.4 ± 43.7	5504.0 ± 51.5	2408.0 ± 33.8	4996.3 ± 56.1	1795.5 ± 30.7
Data	93359.0 ± 305.5	48145.0 ± 219.4	6336.0 ± 79.6	3472.0 ± 58.9	5551.0 ± 74.5	2356.0 ± 48.5	4977.0 ± 70.5	1838.0 ± 42.9
$\frac{\mathrm{Sim-Data}}{\sigma}$	-0.05	0.16	0.72	-1.01	-0.52	0.87	0.21	-0.81

Table B.1: Table of post unconditional fit yields for 1-lepton 2 and 3-jet events in the MVA selection. The uncertainties are the full postfit errors including all nusicance parameters with priors, floating normalizations, and the correlations deduced from the data. The difference in the data and the postfit simulation is compared to σ which is obtained from the sum in simulation as the sum in quadrature of the postfit and Poissonian error.

Sample	1-tag	
	$p_{\mathrm{T}}^{V} < 120\ GeV$	$p_{\mathrm{T}}^{V} > 120\ GeV$
	2-jet	
VH	75.0 ± 30.3	47.2 ± 19.2
VV	1670.5 ± 45.3	768.4 ± 90.2
$t\bar{t}$	8802.5 ± 584.7	5023.6 ± 341.9
s-top	9475.2 ± 1950.3	1792.6 ± 147.4
W+l	47088.5 ± 3590.8	11437.8 ± 872.9
W+cl	55227.5 ± 4279.1	11192.6 ± 965.9
W+hf	13904.3 ± 2561.6	3314.3 ± 552.2
Z+l	2944.5 ± 293.3	474.9 ± 57.5
Z+cl	1191.5 ± 32.3	166.0 ± 4.5
Z+hf	1465.1 ± 330.2	198.6 ± 5.4
MJ_e	–	1126.1 ± 490.5
MJ_μ	6555.6 ± 764.0	55.5
Total	148400.1 ± 415.7	35597.5 ± 203.4
Data	148390.0 ± 385.2	35618.0 ± 188.7
$\frac{\mathrm{Sim-Data}}{\sigma}$	0.02	-0.07
	3-jet	
VH	28.7 ± 11.8	22.0 ± 9.0
VV	892.9 ± 24.2	381.5 ± 48.7
$t\bar{t}$	15790.4 ± 568.8	8043.2 ± 393.7
s-top	5649.4 ± 398.1	1614.6 ± 205.6
W+l	17812.5 ± 1391.1	5062.6 ± 379.3
W+cl	18588.5 ± 1634.7	4511.3 ± 411.1
W+hf	5355.4 ± 967.1	1824.5 ± 365.0
Z+l	1360.4 ± 161.9	203.7 ± 5.5
Z+cl	587.4 ± 15.9	86.6 ± 2.3
Z+hf	689.4 ± 18.7	96.7 ± 2.6
MJ_e	–	329.3 ± 124.2
MJ_μ	3205.9 ± 341.4	25.3
Total	69960.8 ± 263.1	22201.4 ± 142.7
Data	69956.0 ± 264.5	22164.0 ± 148.9
$\frac{\mathrm{Sim-Data}}{\sigma}$	0.01	0.18

Table B.2: Table of post unconditional fit yields for 1-lepton 2 and 3-jet 1-tag events in the dijet mass selection. The uncertainties are the full postfit errors including all nuisance parameters with priors, floating normalizations, and the correlations deduced from the data. The difference in the data and the postfit simulation is compared to σ which is obtained from the sum in quadrature (in simulation) of the postfit and Poissonian error.

Sample	2L-tag				
	$p_{\mathrm{T}}^{V} < 90\ GeV$	$90 < p_{\mathrm{T}}^{V} < 120\ GeV$	$120 < p_{\mathrm{T}}^{V} < 160\ GeV$	$160 < p_{\mathrm{T}}^{V} < 200\ GeV$	$p_{\mathrm{T}}^{V} > 200\ GeV$
	2-jet				
VH	16.6 ± 6.7	5.8 ± 2.3	8.5 ± 3.4	3.9 ± 1.6	4.8 ± 1.9
VV	68.8 ± 1.9	22.1 ± 0.6	25.0 ± 0.7	10.0 ± 1.9	11.5 ± 3.0
$t\bar{t}$	995.6 ± 66.7	383.8 ± 25.4	416.9 ± 26.8	109.5 ± 7.3	60.1 ± 5.9
s-top	648.8 ± 117.9	154.9 ± 26.0	113.5 ± 10.1	27.5 ± 2.1	17.2 ± 1.5
W+l	877.1 ± 103.4	245.4 ± 28.9	189.4 ± 22.1	54.4 ± 6.6	46.3 ± 5.9
W+cl	1548.8 ± 113.7	404.8 ± 31.1	296.6 ± 23.4	76.0 ± 6.1	48.2 ± 4.3
W+hf	1078.2 ± 194.1	288.7 ± 46.3	231.5 ± 33.8	67.2 ± 9.9	55.4 ± 9.7
Z+l	54.7 ± 7.1	19.8 ± 2.6	10.3 ± 0.3	1.7 ± 0.0	0.7 ± 0.0
Z+cl	34.9 ± 0.9	11.9 ± 0.3	5.8 ± 0.2	1.0 ± 0.0	0.4 ± 0.0
Z+hf	81.8 ± 2.2	28.7 ± 0.8	12.7 ± 0.3	2.4 ± 0.1	1.0 ± 0.1
MJ_e	–	–	42.8 ± 13.3	9.2 ± 4.4	1.9 ± 1.0
MJ_μ	459.0 ± 93.7	54.5 ± 11.3	5.1 ± 0.0	1.0	–
Total	5864.4 ± 68.7	1620.3 ± 25.9	1358.2 ± 20.6	363.7 ± 6.9	247.4 ± 8.4
Data	5891.0 ± 76.8	1570.0 ± 39.6	1355.0 ± 36.8	375.0 ± 19.4	239.0 ± 15.5
$\frac{\mathrm{Sim-Data}}{\sigma}$	-0.26	1.05	0.08	-0.56	0.47
	3-jet				
VH	4.8 ± 2.0	1.8 ± 0.8	2.9 ± 1.2	1.7 ± 0.7	2.3 ± 0.9
VV	31.3 ± 0.8	9.6 ± 0.3	11.5 ± 0.3	5.2 ± 0.1	5.4 ± 1.0
$t\bar{t}$	1977.1 ± 61.8	553.5 ± 17.8	567.6 ± 22.9	173.6 ± 9.1	111.1 ± 10.0
s-top	405.6 ± 26.9	97.8 ± 6.8	87.1 ± 10.5	30.1 ± 4.2	24.7 ± 3.6
W+l	329.6 ± 39.8	81.0 ± 9.7	74.1 ± 8.7	24.3 ± 2.9	23.9 ± 3.0
W+cl	543.8 ± 44.6	125.8 ± 10.3	109.9 ± 9.0	33.8 ± 2.8	27.1 ± 2.6
W+hf	406.8 ± 75.7	104.7 ± 18.3	105.7 ± 18.2	39.8 ± 7.5	44.3 ± 9.6
Z+l	24.5 ± 0.7	6.8 ± 0.2	3.7 ± 0.1	0.8 ± 0.0	0.5 ± 0.0
Z+cl	17.9 ± 0.5	4.6 ± 0.1	2.5 ± 0.1	0.5 ± 0.0	0.3 ± 0.0
Z+hf	41.8 ± 1.1	9.7 ± 0.3	6.8 ± 0.2	1.2 ± 0.0	0.7 ± 0.0
MJ_e	–	–	12.3 ± 1.0	2.5 ± 0.0	0.8 ± 0.0
MJ_μ	138.6 ± 18.9	10.8 ± 1.5	1.2	–	–
Total	3921.7 ± 43.7	1006.1 ± 16.0	985.3 ± 16.1	313.5 ± 6.6	241.1 ± 8.6
Data	3882.0 ± 62.3	1004.0 ± 31.7	1051.0 ± 32.4	313.0 ± 17.7	238.0 ± 15.4
$\frac{\mathrm{Sim-Data}}{\sigma}$	0.52	0.06	-1.86	0.03	0.17

Table B.3: Table of post unconditional fit yields for 1-lepton 2 and 3-jet 2L-tag events in the dijet mass selection. The uncertainties are the full postfit errors including all nusicance parameters with priors, floating normalizations, and the correlations deduced from the data. The difference in the data and the postfit simulation is compared to σ which is obtained from the sum in simulation as the sum in quadrature of the postfit and Poissonian error.

Sample	2M-tag				
	$p_{\mathrm{T}}^{V} < 90\ GeV$	$90 < p_{\mathrm{T}}^{V} < 120\ GeV$	$120 < p_{\mathrm{T}}^{V} < 160\ GeV$	$160 < p_{\mathrm{T}}^{V} < 200\ GeV$	$p_{\mathrm{T}}^{V} > 200\ GeV$
			2-jet		
VH	27.1 ± 10.8	9.0 ± 3.6	13.6 ± 5.4	6.6 ± 2.6	7.9 ± 3.1
VV	46.4 ± 2.9	15.2 ± 1.3	18.6 ± 2.5	7.3 ± 1.4	6.9 ± 1.8
$t\bar{t}$	1235.7 ± 77.1	469.5 ± 28.6	456.3 ± 26.6	85.4 ± 5.5	34.7 ± 3.3
s-top	604.0 ± 97.7	139.0 ± 21.4	97.5 ± 9.3	20.6 ± 1.8	11.4 ± 1.1
W+l	66.7 ± 13.0	18.3 ± 3.6	13.2 ± 2.5	3.6 ± 0.7	3.0 ± 0.6
W+cl	261.1 ± 21.8	68.9 ± 5.9	49.8 ± 4.6	11.5 ± 1.0	7.2 ± 0.8
W+hf	730.6 ± 126.7	206.9 ± 30.7	163.1 ± 22.3	46.0 ± 6.4	41.1 ± 6.6
Z+l	4.3 ± 0.1	1.5 ± 0.0	0.7 ± 0.0	0.1 ± 0.0	0.0 ± 0.0
Z+cl	7.3 ± 0.2	2.4 ± 0.1	1.2 ± 0.4	0.1 ± 0.0	0.0 ± 0.0
Z+hf	66.9 ± 12.7	23.4 ± 4.9	10.2 ± 2.5	1.6 ± 0.4	0.7 ± 0.1
MJ_e	–	–	25.2 ± 9.2	4.7 ± 2.7	1.2 ± 0.7
MJ_μ	202.8 ± 64.0	15.9 ± 5.2	1.1 ± 0.4	–	–
Total	3253.0 ± 49.1	969.9 ± 16.6	850.7 ± 16.5	187.6 ± 4.6	114.2 ± 5.4
Data	3230.0 ± 56.8	984.0 ± 31.4	868.0 ± 29.5	199.0 ± 14.1	116.0 ± 10.8
$\frac{\mathrm{Sim}-\mathrm{Data}}{\sigma}$	0.31	-0.4	-0.52	-0.79	-0.15
			3-jet		
VH	7.3 ± 3.0	3.0 ± 1.2	4.8 ± 1.9	3.0 ± 1.2	3.6 ± 1.4
VV	14.5 ± 0.4	4.9 ± 0.1	6.3 ± 0.2	2.8 ± 0.4	3.3 ± 0.5
$t\bar{t}$	2595.2 ± 58.4	674.2 ± 17.2	589.1 ± 19.9	127.6 ± 5.9	63.9 ± 6.0
s-top	397.8 ± 28.3	88.9 ± 6.5	71.5 ± 8.1	21.0 ± 3.1	17.2 ± 2.5
W+l	22.6 ± 0.6	5.8 ± 0.2	5.1 ± 0.1	1.5 ± 0.0	1.4 ± 0.3
W+cl	91.4 ± 7.8	22.2 ± 2.1	17.9 ± 1.6	5.3 ± 0.6	4.1 ± 0.5
W+hf	260.8 ± 49.9	68.7 ± 11.3	80.3 ± 12.9	26.9 ± 4.6	31.4 ± 6.3
Z+l	1.6 ± 0.0	0.5 ± 0.0	0.3 ± 0.0	0.1 ± 0.0	0.0 ± 0.0
Z+cl	3.0 ± 0.1	0.8 ± 0.0	0.4 ± 0.0	0.1 ± 0.0	0.0 ± 0.0
Z+hf	29.9 ± 0.8	8.0 ± 0.2	5.0 ± 0.1	1.2 ± 0.0	0.5 ± 0.1
MJ_e	–	–	6.0 ± 0.4	2.2 ± 0.3	0.4 ± 0.1
MJ_μ	70.7 ± 9.6	12.1 ± 9.4	–	0.6	–
Total	3494.9 ± 37.8	889.2 ± 16.2	786.7 ± 14.6	192.1 ± 4.7	125.8 ± 5.5
Data	3519.0 ± 59.3	936.0 ± 30.6	780.0 ± 27.9	181.0 ± 13.5	123.0 ± 11.1
$\frac{\mathrm{Sim}-\mathrm{Data}}{\sigma}$	-0.34	-1.38	0.21	0.76	0.22

Table B.4: Table of post unconditional fit yields for 1-lepton 2 and 3-jet 2M-tag events in the dijet mass selection. The uncertainties are the full postfit errors including all nusicance parameters with priors, floating normalizations, and the correlations deduced from the data. The difference in the data and the postfit simulation is compared to σ which is obtained from the sum in simulation as the sum in quadrature of the postfit and Poissonian error.

Sample			2T-tag		
	$p_{\mathrm{T}}^{V} < 90\ GeV$	$90 < p_{\mathrm{T}}^{V} < 120\ GeV$	$120 < p_{\mathrm{T}}^{V} < 160\ GeV$	$160 < p_{\mathrm{T}}^{V} < 200\ GeV$	$p_{\mathrm{T}}^{V} > 200\ GeV$
			2-jet		
VH	30.1 ± 11.8	10.0 ± 3.9	15.6 ± 6.1	8.3 ± 3.2	8.6 ± 3.4
VV	36.8 ± 3.1	12.4 ± 1.3	15.4 ± 2.4	6.8 ± 1.5	6.1 ± 1.9
$t\bar{t}$	1174.1 ± 73.3	431.3 ± 25.9	377.4 ± 21.0	44.5 ± 3.0	10.3 ± 1.2
s-top	490.8 ± 77.3	107.5 ± 16.6	72.8 ± 8.3	13.2 ± 1.3	4.6 ± 0.5
W+l	2.0 ± 0.1	0.4 ± 0.0	0.3 ± 0.0	0.1 ± 0.0	0.1 ± 0.0
W+cl	13.5 ± 2.7	3.4 ± 0.5	2.3 ± 0.3	0.5 ± 0.2	0.3 ± 0.1
W+hf	403.0 ± 73.4	115.1 ± 16.6	105.6 ± 14.9	29.7 ± 4.2	27.3 ± 4.5
Z+l	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Z+cl	0.4 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	–
Z+hf	44.0 ± 12.7	15.4 ± 4.4	6.8 ± 2.0	1.7 ± 0.5	0.5 ± 0.1
MJ_e	–	–	11.7 ± 5.0	2.0 ± 1.8	1.1 ± 0.5
MJ_μ	135.7 ± 65.4	12.7 ± 5.9	–	–	–
Total	2330.5 ± 48.6	708.3 ± 15.0	608.0 ± 14.5	106.7 ± 3.8	59.0 ± 4.5
Data	2364.0 ± 48.6	700.0 ± 26.5	591.0 ± 24.3	112.0 ± 10.6	59.0 ± 7.7
$\frac{\text{Sim}-\text{Data}}{\sigma}$	-0.49	0.27	0.59	-0.48	0.0
			3-jet		
VH	8.4 ± 3.4	3.3 ± 1.3	5.6 ± 2.2	3.4 ± 1.3	4.1 ± 1.6
VV	9.2 ± 0.2	3.4 ± 0.1	5.8 ± 0.2	2.8 ± 0.5	2.6 ± 0.4
$t\bar{t}$	2637.1 ± 56.9	642.9 ± 16.4	488.1 ± 16.5	61.5 ± 3.5	15.6 ± 1.8
s-top	336.3 ± 24.8	71.5 ± 5.5	50.3 ± 6.2	10.4 ± 1.6	7.3 ± 1.1
W+l	0.7 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
W+cl	4.4 ± 0.5	1.1 ± 0.1	1.0 ± 0.1	0.2 ± 0.0	0.1 ± 0.0
W+hf	142.1 ± 26.4	41.4 ± 6.7	48.3 ± 7.7	19.6 ± 3.4	21.7 ± 4.4
Z+l	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Z+cl	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Z+hf	22.2 ± 0.6	6.1 ± 0.2	3.7 ± 0.1	0.9 ± 0.3	0.4 ± 0.2
MJ_e	–	–	2.4 ± 0.1	0.7 ± 0.1	0.3 ± 0.0
MJ_μ	35.6 ± 4.8	1.9 ± 0.2	–	0.3 ± 0.0	–
Total	3196.1 ± 42.4	771.8 ± 15.6	605.2 ± 14.4	99.9 ± 3.8	52.3 ± 4.0
Data	3161.0 ± 56.2	779.0 ± 27.9	590.0 ± 24.3	101.0 ± 10.0	53.0 ± 7.3
$\frac{\text{Sim}-\text{Data}}{\sigma}$	0.5	-0.23	0.53	-0.1	-0.09

Table B.5: Table of post unconditional fit yields for 1-lepton 2 and 3-jet 2T-tag events in the dijet mass selection. The uncertainties are the full postfit errors including all nusicance parameters with priors, floating normalizations, and the correlations deduced from the data. The difference in the data and the postfit simulation is compared to σ which is obtained from the sum in simulation as the sum in quadrature of the postfit and Poissonian error.

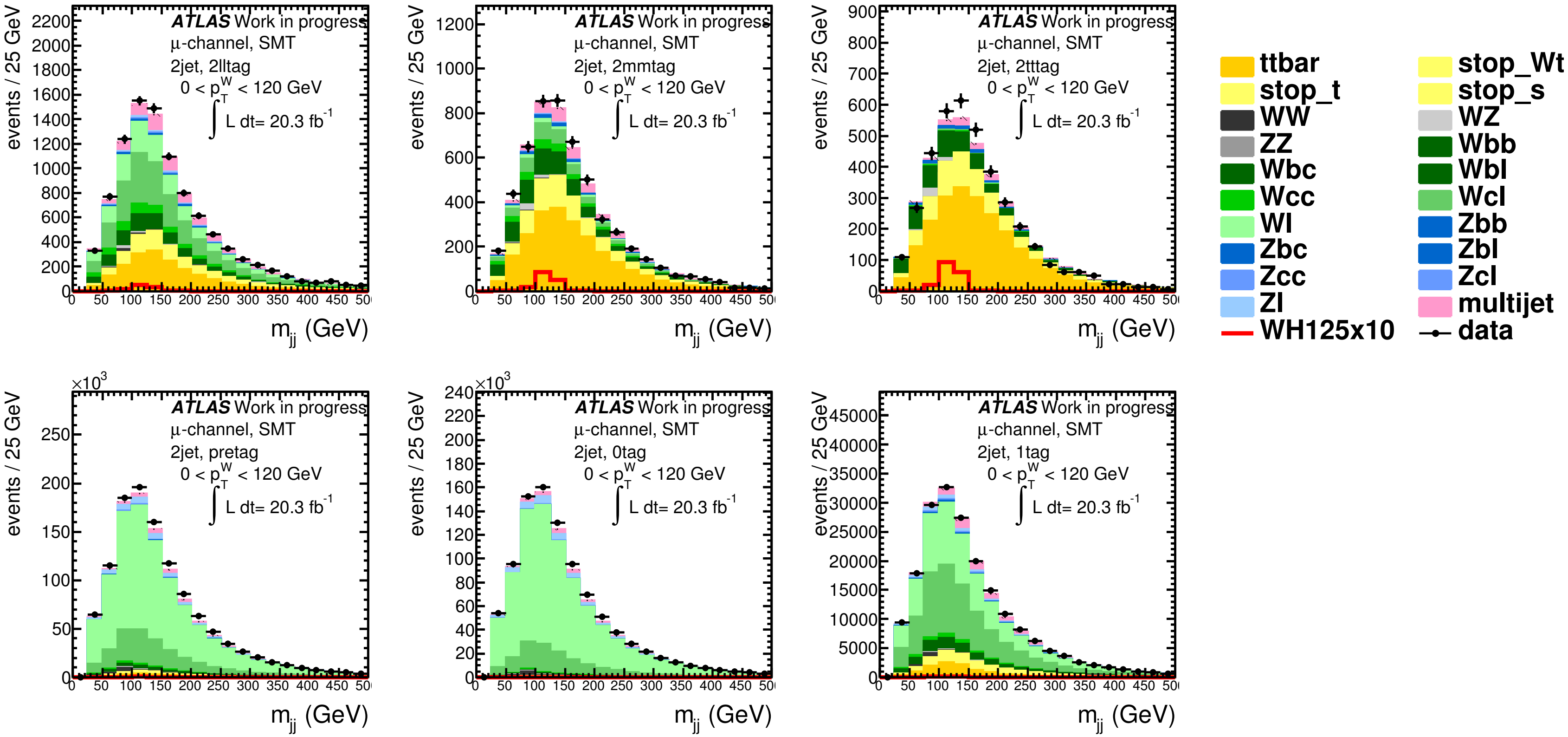


Fig. B.1: The pre-fit m_jj distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

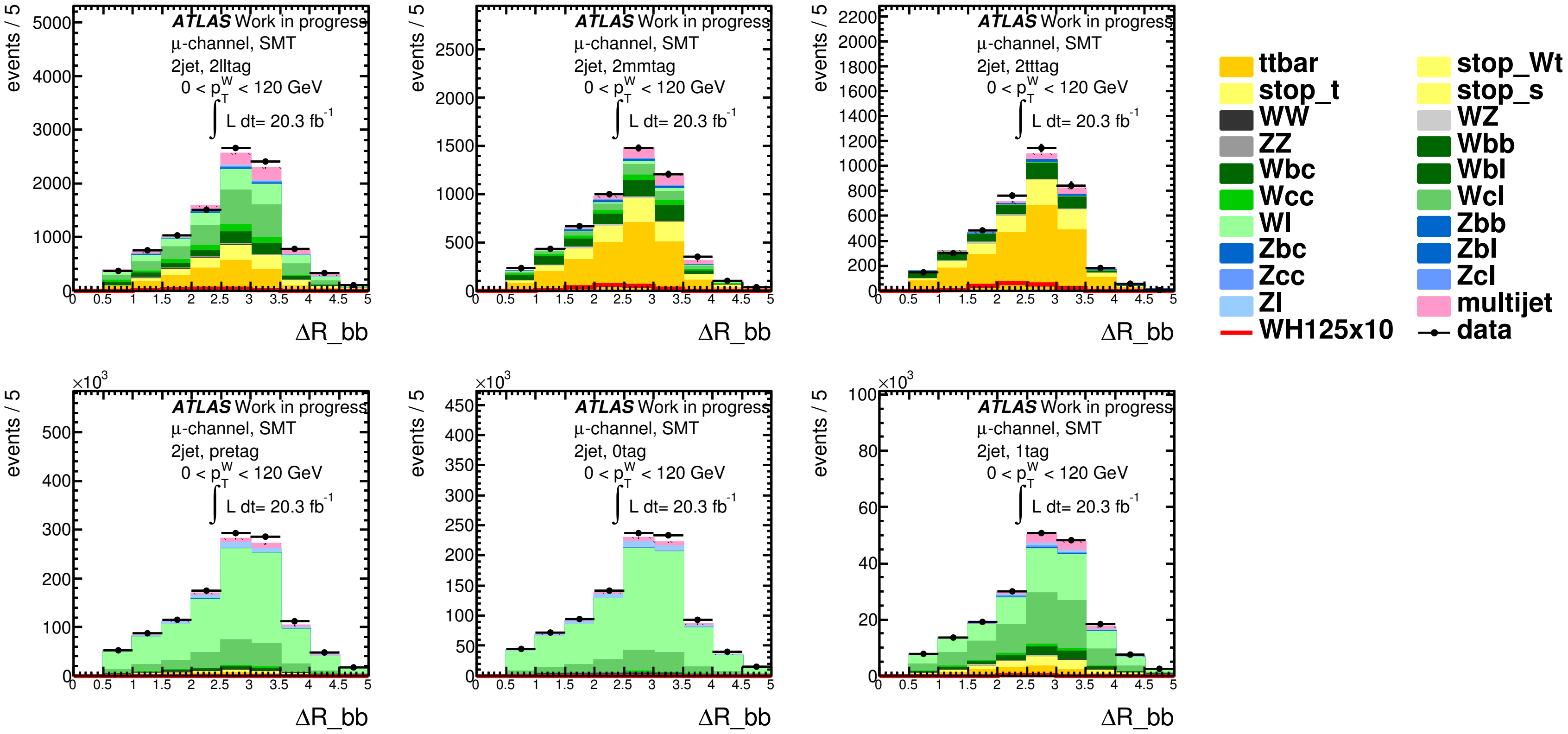


Fig. B.2: The pre-fit ΔR_{jj} distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

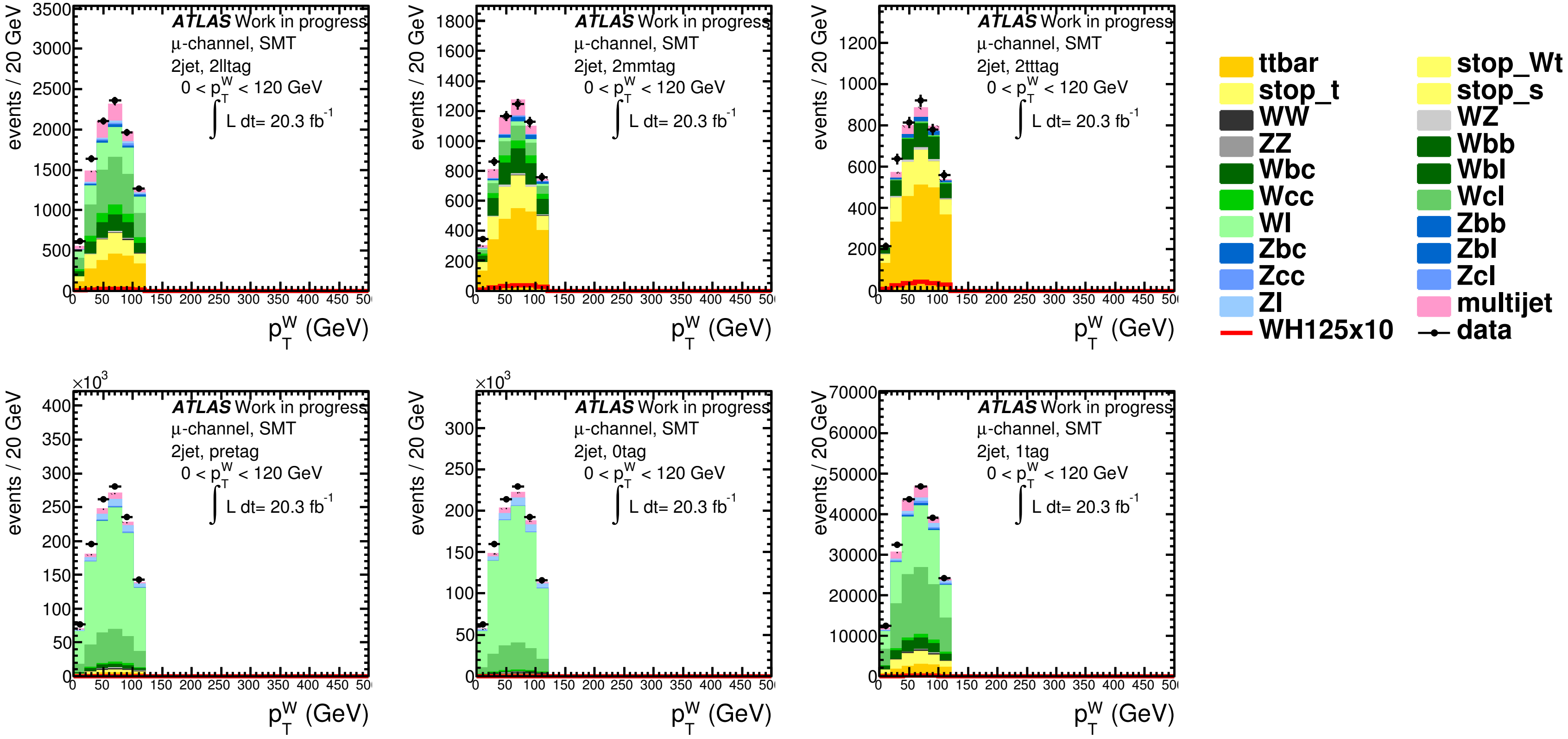


Fig. B.3: The pre-fit p_{T}^{W} distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

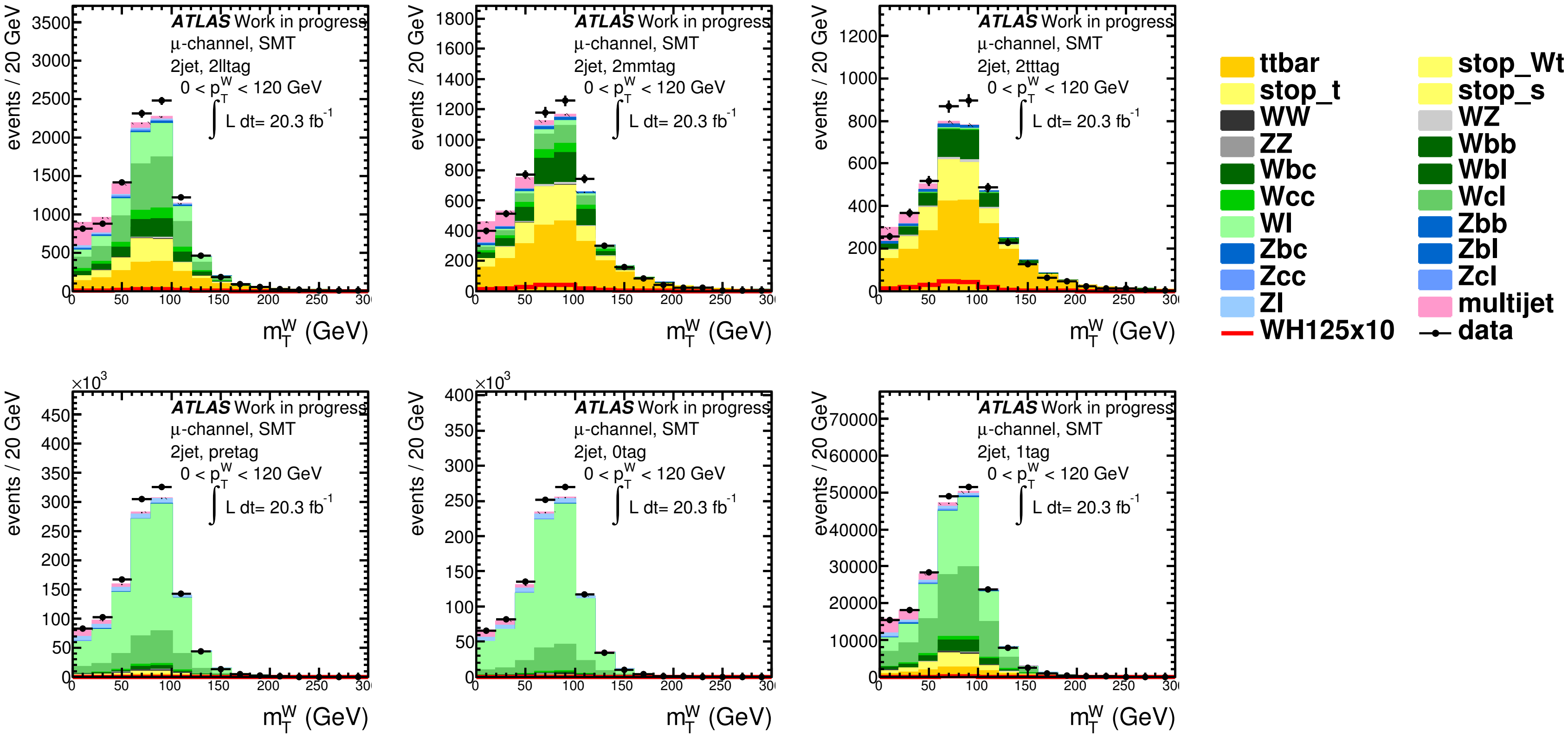


Fig. B.4: The pre-fit m_{T}^{W} distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

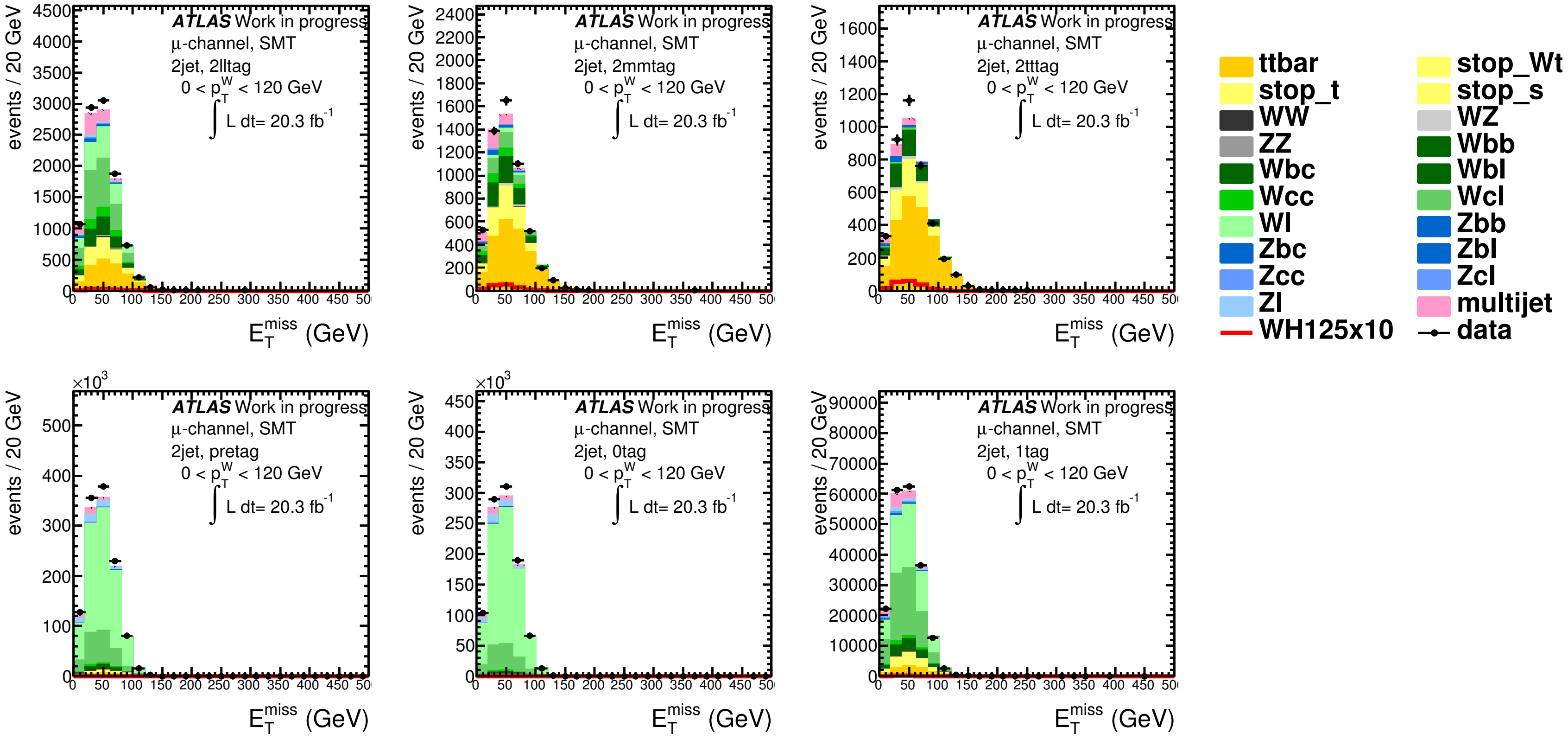


Fig. B.5: The pre-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

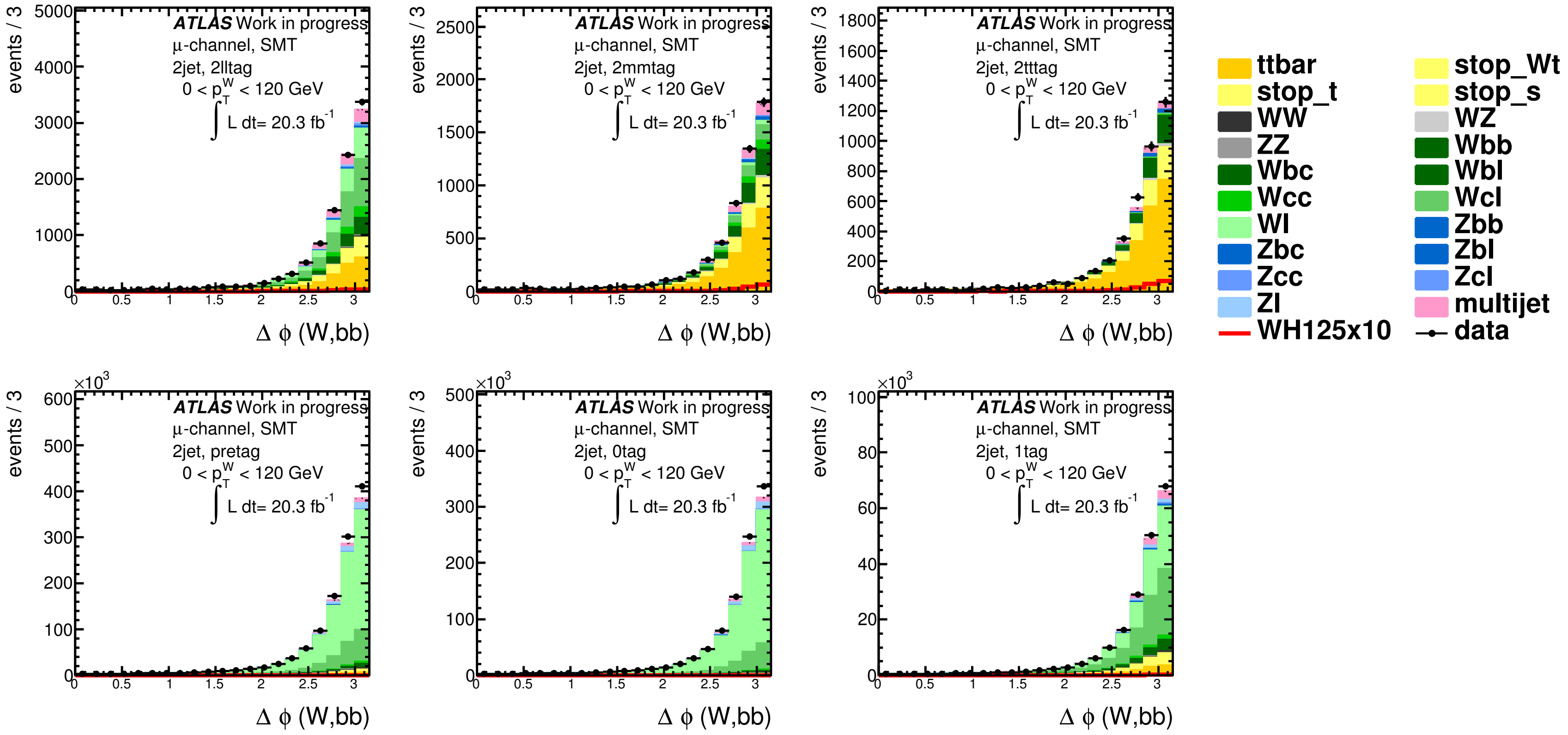


Fig. B.6: The pre-fit $\Delta\phi(W, bb)$ distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

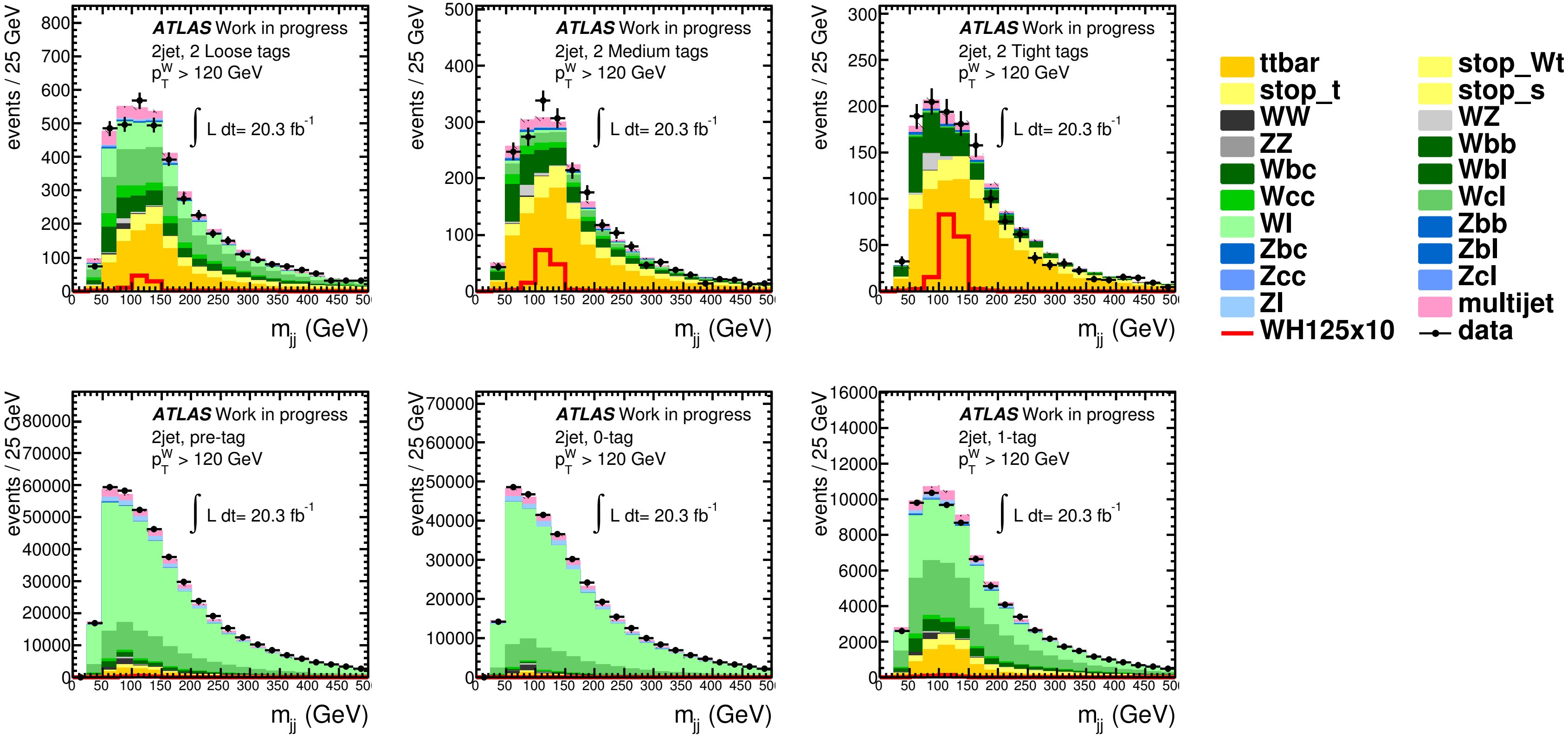


Fig. B.7: The pre-fit m_jj distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

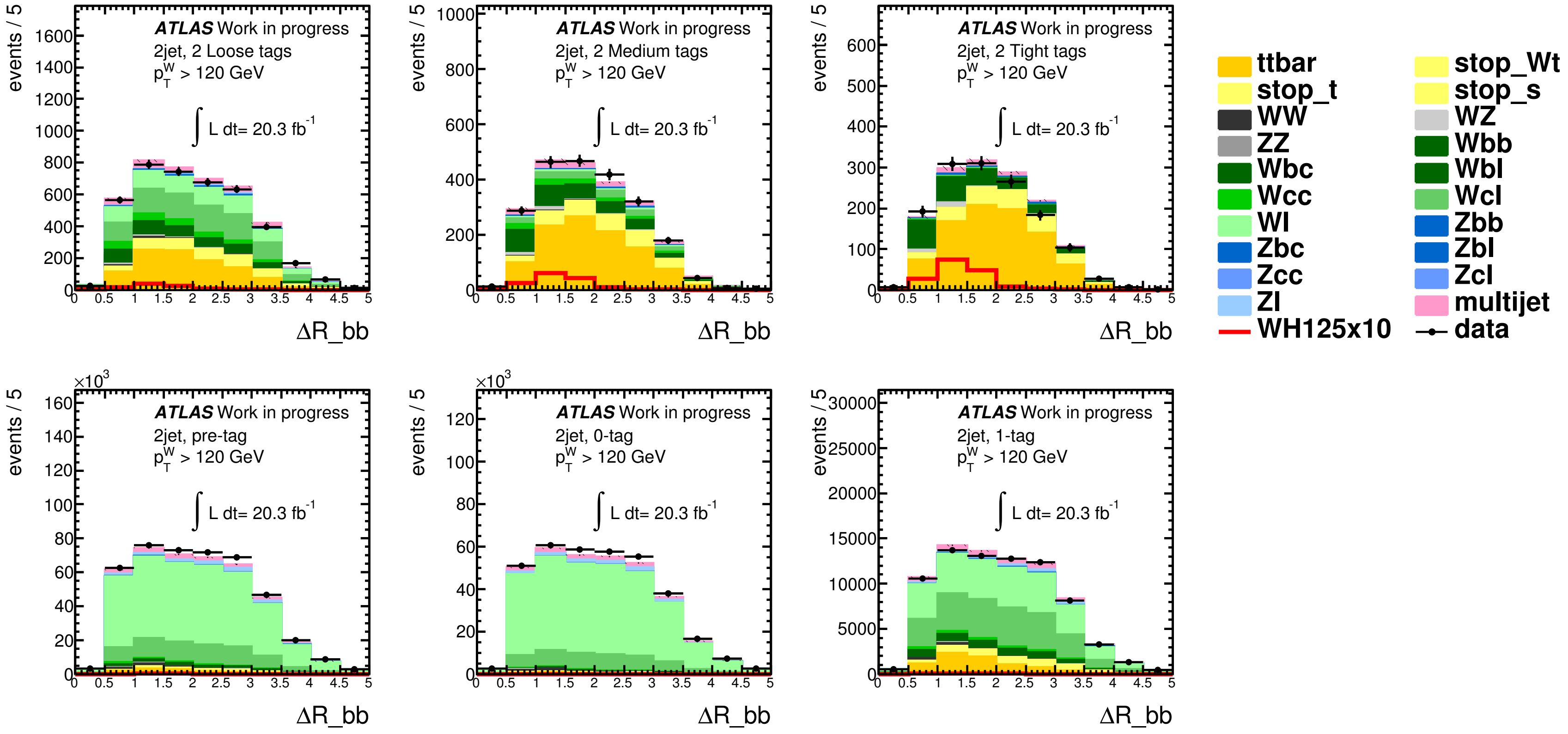


Fig. B.8: The pre-fit ΔR_{jj} distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

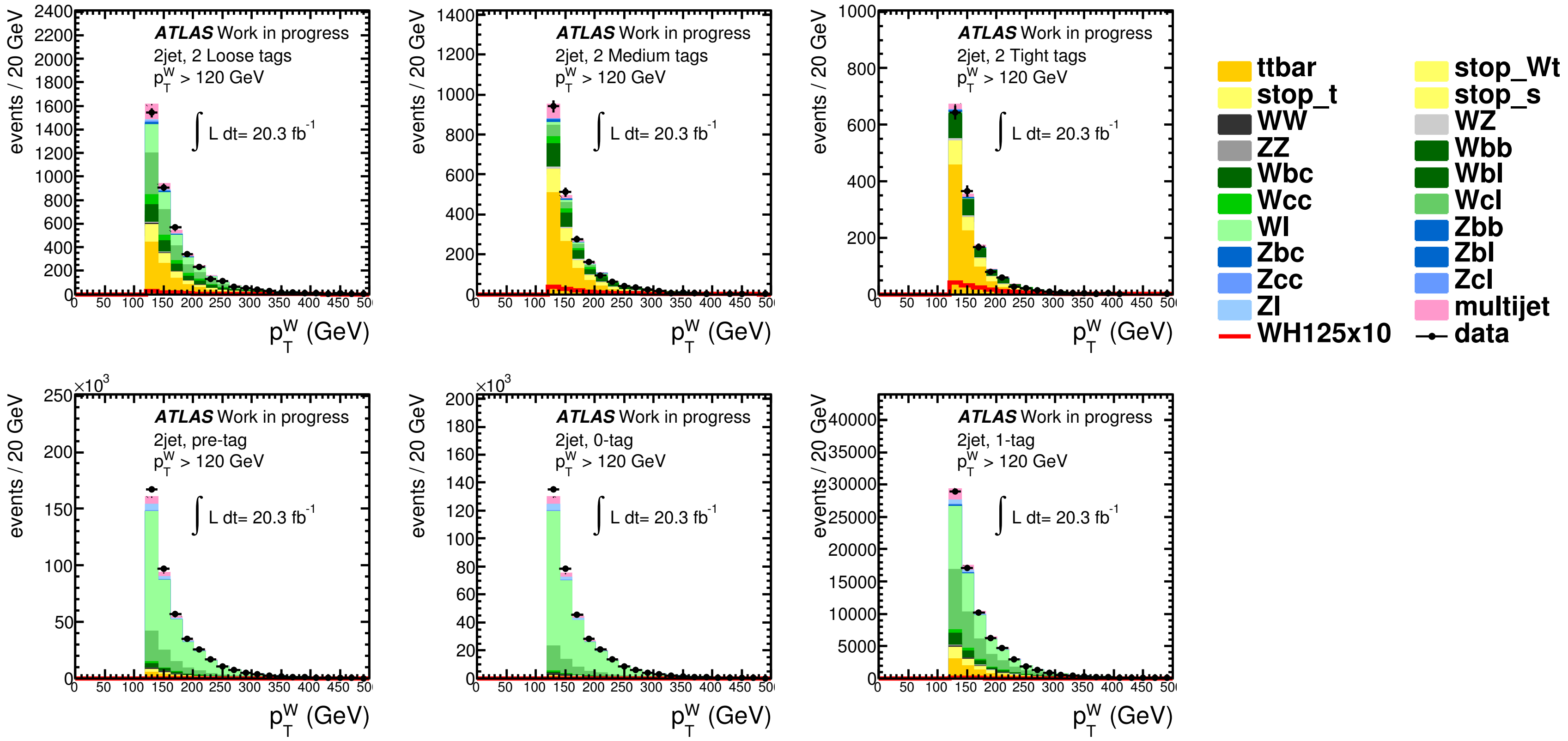


Fig. B.9: The pre-fit p_{T}^{W} distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

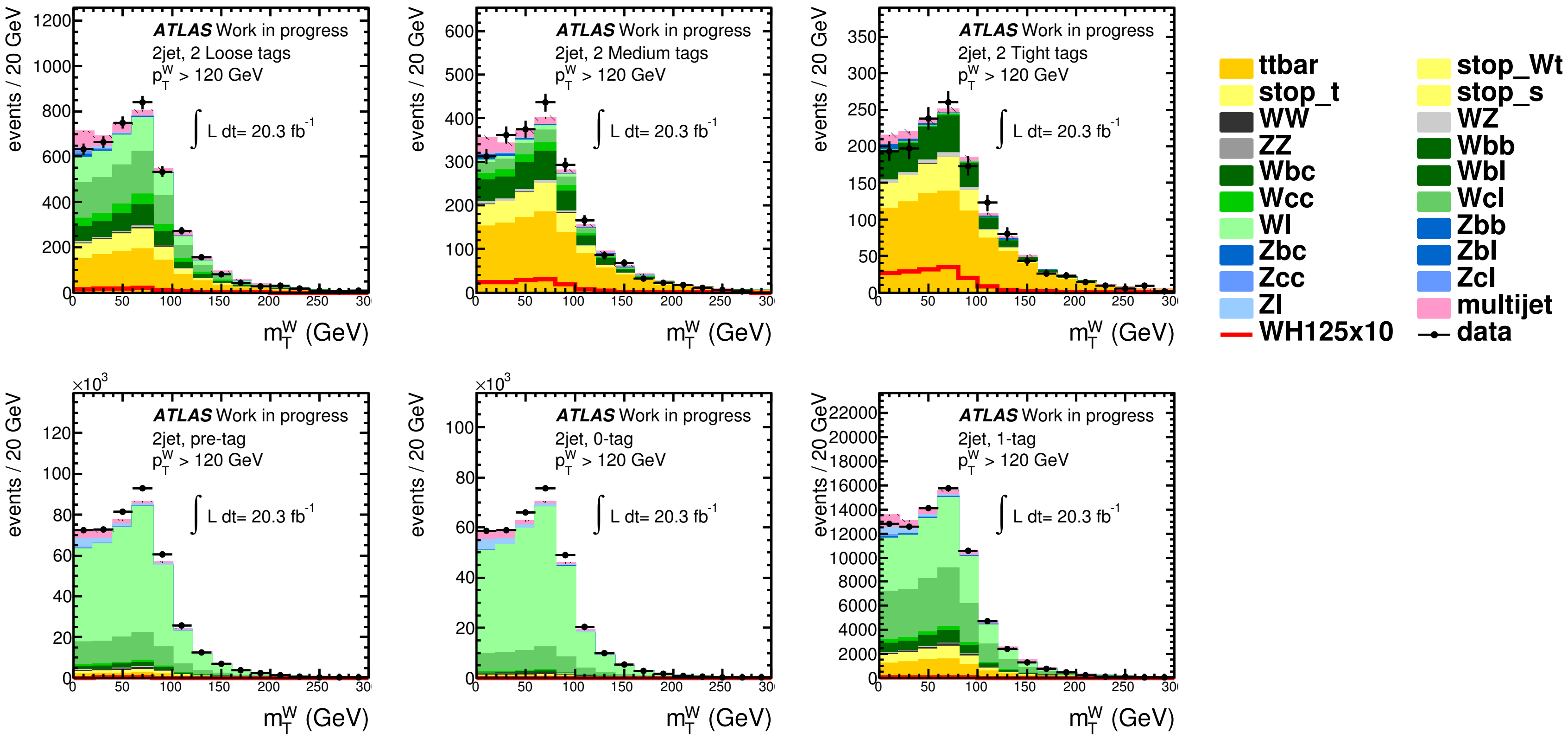


Fig. B.10: The pre-fit m_{T}^{W} distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

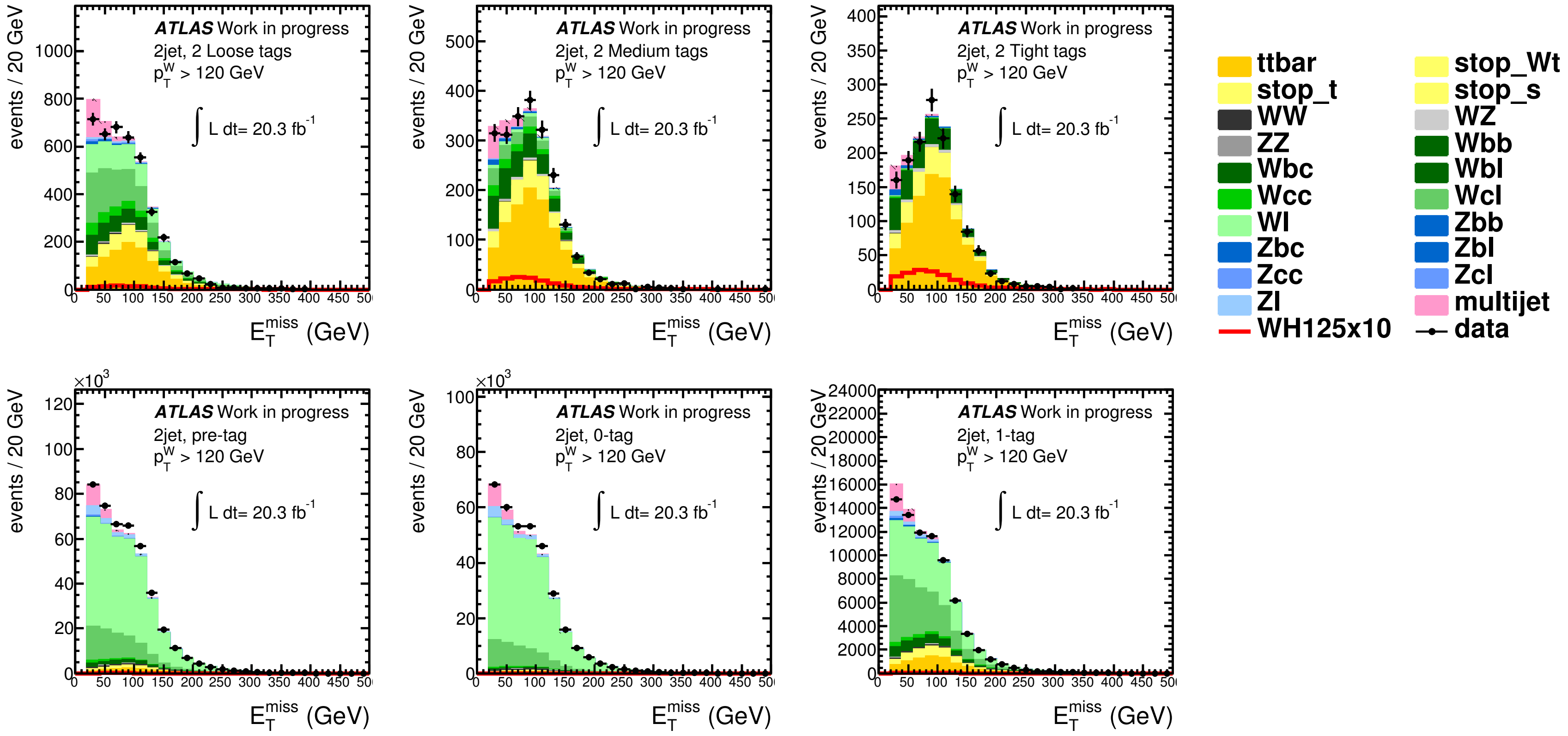


Fig. B.11: The pre-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

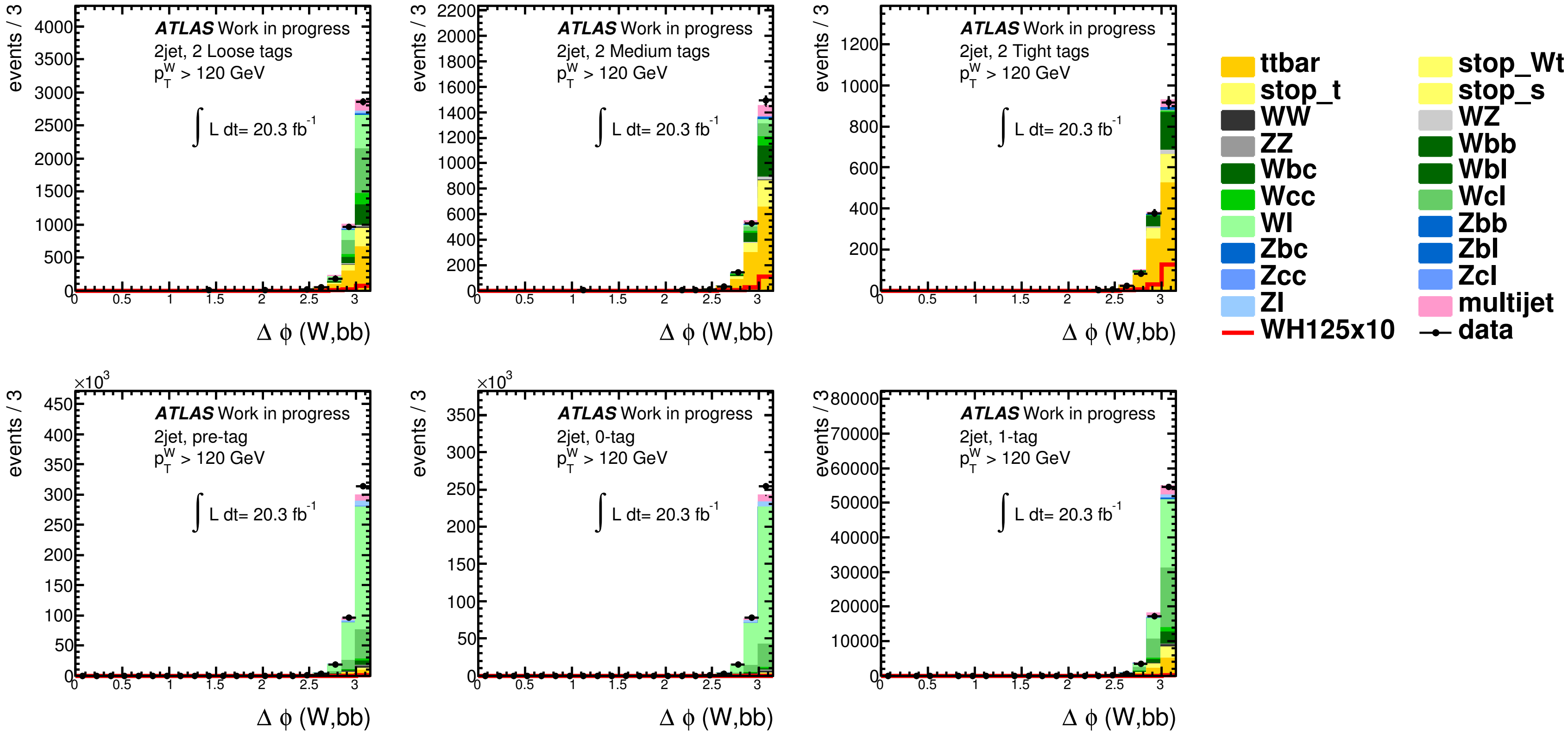


Fig. B.12: The pre-fit $\Delta\phi(W, bb)$ distribution after MVA selection in the pre-tag (bottom-left), 0-tag (bottom-center), 1-tag (bottom-right), 2 loose tags (top-left), 2 medium tags (top-center), and 2 tight tags (top-right) categories.

C. Sensitivity gain study using loosely selected leptons

This chapter summarizes a study to gain sensitivity in the $WH \to \ell\nu b\bar{b}$ analysis using additional lepton categories. The study here focuses on the muon channel, in which the multi-jet background is smaller than the electron channel. Except the loosened lepton selection criteria, object and event selection criteria used in this study are the ones described in chapter 5.

C.1. Loosely selected leptons

In order to gain in analysis sensitivity, additional lepton categories are considered by loosening the isolation and p_{T} requirements:

A) $p_{\mathrm{T}} > 25$ GeV and track isolation greater than 0.04 but smaller than 0.1

B) $15 < p_{\mathrm{T}} \leq 25$ GeV and track isolation smaller than 0.04

C) $15 < p_{\mathrm{T}} \leq 25$ GeV and track isolation greater than 0.04 but smaller than 0.1

and these three categories are illustrated in fig C.1 together with the tight (signal) region, on the (trk-iso, p_{T}^{ℓ}) plane.

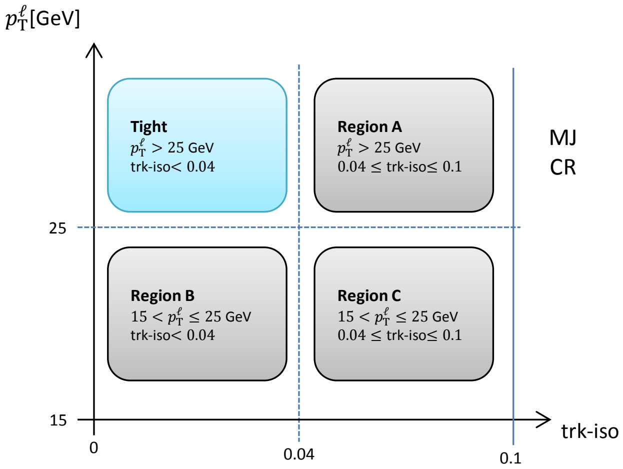


Fig. C.1: Definition of additional lepton categories concerned, according to the lepton p_{T} and track isolation. The multi-jet background control region is also indicated (MJ CR).

C.2. Sensitivity gain

The signal acceptance gain using the three regions defined in the previous section are studied with the cut-based selection of the $WH \to \ell\nu b\bar{b}$ analysis. Table C.1 summarizes the signal acceptance gain in each region. The numbers were estimated in the dijet mass region of $100 < m_{jj} < 150$ GeV, adding the three 2-tag categories (LL, MM, and TT). Region A has the largest gain in signal acceptance, since the leptons from W boson decay have normally high enough p_{T} above the 25 GeV requirement.

Table C.1: Expected signal yield in each region, in 2- and 3-jet bins separately. For the additional lepton categories, relative gain in the signal acceptance are shown in parentheses.

	2-jet	3-jet
Tight lepton	33.4	11.6
Region A	3.69 (11%)	1.22 (11%)
Region B	1.53 (4.5%)	0.43 (4.1%)
Region C	0.46 (1.4%)	0.25 (2.3%)

The significance in the three regions are calculated using eq 5.2, and the gain from the three regions are summarized in table C.2. The significance is estimated using the expected yield of signal and background events in the dijet mass region of $100 < m_{jj} < 150$ GeV. The expected yield of the signal and background events were counted separately in the five p_{T}^{W} intervals and the three 2-tag categories (LL, MM, and TT) to calculated the significance.

Since the study was done at the early stage of the analysis, the multi-jet background yield was not taken into account. However, the analysis significance is mainly lead by the high p_{T}^{W} region, where the multi-jet background contributes minimally. The significance in the first p_{T}^{W} bin ($p_{\mathrm{T}}^{W} < 90$ GeV) in the additional lepton region was not counted since this bin is dominated by the multi-jet background. As can be seen in also the signal acceptance gain, the largest gain in significance was found in region A, with 2.3% and 5.0% gain in the 2- and 3-jet bins, respectively.

Table C.2: Significance in 2 and 3 jet bins. For the additional lepton categories, relative gain in significance are shown. The number in the parenthesis refers to the combined significance with the tight lepton category.

	2-jet		3-jet	
	significance	gain	significance	gain
Tight lepton	0.886		0.291	
Region A	0.193 (0.906)	2.3%	0.093 (0.305)	5.0%
Region B	0.115 (0.893)	0.8%	0.033 (0.294)	0.6%
Region C	0.065 (0.888)	0.2%	0.026 (0.293)	0.3%

C.3. Summary

Although it is found there is a sensitivity gain using loosely selected leptons, the sensitivity improvement was small compared to the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel. Also, providing a proper estimate of the multi-jet background would need more study and it was highly time consuming. Due to the limited time, the decision taken was to discard this analysis category for the Run 1 analysis and to keep the efforts focused on the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered muon channel. The gain seen in region A shows a few % level of improvement and is correlated to the optimization of the lepton isolation requirement. Therefore, it can be one of the check points to re-optimize the analysis in Run 2.

Abstract

Keywords : LHC, ATLAS, Higgs boson, $H \to b\bar{b}$, jet flavor tagging, boosted decision trees

In July 2012, the ATLAS and CMS collaborations announced the discovery of a new particle at a mass of about $125\ \text{GeV}$, compatible with the Higgs boson predicted by the Standard Model (SM) of particle physics. Although all measurements as of Summer 2015 show a full consistency with the SM predictions, the fermionic decay channels ($H \to b\bar{b}$ and $\tau^+\tau^-$) have not been seen yet as clearly as the bosonic decay channels. This thesis presents a search for the Higgs boson in the $WH \to \ell\nu b\bar{b}$ channel, using proton-proton collision data at $\sqrt{s} = 8\ \text{TeV}$ taken with the ATLAS detector in the year 2012, corresponding to an integrated luminosity of 20.3 fb^{-1}. This document details in particular one of the contributions made by the author in this search: the additional analysis region with muon events triggered by missing transverse momentum. An excess over the background-only hypothesis has been found with a significance of 1.8 σ while 1.5 σ was expected. The observed (expected) upper limit on the cross-section times branching ratio for $WH \to \ell\nu b\bar{b}$ at 95% confidence level is found to be 2.35 (1.37) times the SM prediction at $m_H = 125\ \text{GeV}$. This search highly relies on the identification of jets originating in b-quark fragmentation, so-called b-tagging. In order to improve the $H \to b\bar{b}$ search and other physics analyses using b-tagging in ATLAS, the development of high performance b-tagging algorithms has been performed and is presented also in this thesis. A new b-tagging algorithm, called MV2, is introduced: it improves the b-tagging performance significantly and is now the baseline b-tagging algorithm in ATLAS for the Run-2 analyses.

En juillet 2012, les collaborations ATLAS et CMS ont annoncé la découverte d'une nouvelle particule à une masse d'environ $125\ \text{GeV}$, compatible avec le boson de Higgs prédit par le Modèle Standard (MS) de la physique des particules. Bien qu'à ce jour toutes les mesures montrent une pleine cohérence avec les prédictions du MS, les canaux de désintégration fermioniques ($H \to b\bar{b}$ and $\tau\tau$) n'ont pas encore été vus aussi clairement que les canaux de désintégration en bosons. Cette thèse présente une recherche du boson de Higgs dans le canal $WH \to \ell\nu b\bar{b}$, en utilisant les données de collisions proton-proton à $\sqrt{s} = 8\ \text{TeV}$ prises avec le détecteur ATLAS en 2012, correspondant à une luminosité intégrée de $20,3$ fb^{-1}. Ce document détaille en particulier l'une des contributions faites par l'auteur à cette recherche : la région d'analyse supplémentaire avec des événements de muons déclenchés par l'impulsion transverse manquante. Un excès par rapport à l'hypothèse de fond seulement a été trouvé avec une signification statistique de $1,8\sigma(1,5\sigma)$ observée (attendue). La limite supérieure observée (attendue) sur le produit section efficace par rapport d'embranchement pour $WH \to \ell\nu b\bar{b}$ à un niveau de confiance de 95% est de 2,35 (1,37) fois la prédiction du MS pour $m_H = 125\ \text{GeV}$. Cette recherche repose fortement sur l'identification des jets issus de la fragmentation du quark b. Afin d'améliorer la recherche $H \to b\bar{b}$ et d'autres analyses de physique utilisant l'identification des jets b dans ATLAS, le développement d'algorithmes de haute performance a été conduit et est présenté également dans cette thèse. Un nouvel algorithme, appelé MV2, est introduit : il améliore la performance d'identification des jets b de manière significative et est maintenant l'algorithme de référence pour les analyses du Run 2 avec ATLAS.